ARTICLE   **OPEN**

Check for updates

Intraoperative margin assessment for basal cell carcinoma with deep learning and histologic tumor mapping to surgical site

Joshua J Levy [1,2,3,4,5,6 ✉], Matthew J Davis[3], Rachael S Chacko[7], Michael J Davis[3], Lucy J Fu[7], Tarushii Goel[8,9], Akash Pamal [8,10], Irfan Nafi [8,11], Abhinav Angirekula[8,12], Anish Suvarna[8], Ram Vempati[8], Brock C Christensen [3,13,14], Matthew S Hayden[3], Louis J Vaickus[4] and Matthew R LeBoeuf[3]

Successful treatment of solid cancers relies on complete surgical excision of the tumor either for definitive treatment or before adjuvant therapy. Intraoperative and postoperative radial sectioning, the most common form of margin assessment, can lead to incomplete excision and increase the risk of recurrence and repeat procedures. Mohs Micrographic Surgery is associated with complete removal of basal cell and squamous cell carcinoma through real-time margin assessment of 100% of the peripheral and deep margins. Real-time assessment in many tumor types is constrained by tissue size, complexity, and specimen processing / assessment time during general anesthesia. We developed an artificial intelligence platform to reduce the tissue preprocessing and histological assessment time through automated grossing recommendations, mapping and orientation of tumor to the surgical specimen. Using basal cell carcinoma as a model system, results demonstrate that this approach can address surgical laboratory efficiency bottlenecks for rapid and complete intraoperative margin assessment.

npj Precision Oncology (2024)8:2 ; https://doi.org/10.1038/s41698-023-00477-7

INTRODUCTION

Complete surgical resection is first line treatment for many solid tumors, which typically requires excision of the clinically evident tumor and the rim of surrounding normal tissue followed by closure with subsequent post-operative histologic analysis of tissue margins (POMA). In the pathology laboratory, the specimen is commonly grossed in a breadloafed or radial fashion, embedded, sectioned, stained, and examined by the pathologist. While POMA of breadloafed sections is the current standard, removal and histologic analysis of the margin in this manner has three major pitfalls: (1) post-operative identification of positive margins (tumor identified at the tissue edge), necessitating a repeat procedure, (2) false negative or "missed" margins, where the tumor is present at a portion of the margin not evaluated due to sampling error, and (3) excessive tissue is removed to limit the possibility of pitfalls 1 and 2 as above, which can result in the removal of critical structures. Standard excisions and POMA for the treatment of skin cancer reveals a combined positive margin or tumor recurrence rate of at least 20%[1–4]. Either of these outcomes (i.e., post-operative positive margins or false negative margins) requires additional surgery, radiation, chemotherapy, or some combination thereof, resulting in patient morbidity, mortality, and a significant cost to our healthcare system[5–10]. These pitfalls have been addressed in some settings through the use of intraoperative frozen sections or analyzing a larger percentage of tissue margins, which, in comparison to standard excisions and POMA, can reduce positive margin or recurrence rates to less than 1-2% in certain surgical subspecialties[3,11–13].

Successful intraoperative treatment of solid tumors requires the combined efforts of multiple highly trained individuals. Tumor removal is performed by the surgeon; cryofreezing and sectioning of the tissue and then staining by the histotechnologist; and histologic analysis by the pathologist. In the current surgical workflow, the surgeon, histotechnologist, and pathologist are often separated by time and space. For example, communication of histological findings between pathologist and surgeon may occur over the phone. This separation presents an obstacle to evaluating intraoperative frozen sections (Fig. 1). Prior studies have shown that breadloaf grossing of tissue sections results in analysis of approximately 1–2% of the margin[6]. Increasing the percentage of tissue margins analyzed requires either: (1) more tissue blocks and sections, or (2) an alternative grossing method. These approaches require more time and/or expertise on the part of both the histotechnician and pathologist.

Mohs Micrographic Surgery (MMS) is used for the treatment of skin cancers of the head, neck, and special sites[14,15]. Tumor removal is performed under local anesthesia with real-time margin assessment using frozen tissue sections that are cut by a histotechnician in an on-site laboratory. Tissue is grossed in a manner that allows the peripheral and deep margin to sit in the same plane, allowing analysis of 100% of tissue margins. The Mohs Micrographic Surgeon performs tumor removal, histologic analysis, and the creation of a surgical tumor map to inform additional tumor removal if necessary. As compared to standard excisions and POMA (≥20% recurrence, as aforementioned), MMS results in a significantly lower tumor recurrence rate (less than 1–2%, as

---

[1]Department of Pathology and Laboratory Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA. [2]Department of Computational Biomedicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA. [3]Department of Dermatology, Geisel School of Medicine at Dartmouth, Hanover, NH 03756, USA. [4]Emerging Diagnostic and Investigative Technologies, Clinical Genomics and Advanced Technologies, Department of Pathology and Laboratory Medicine, Dartmouth Hitchcock Medical Center, Lebanon, NH 03756, USA. [5]Department of Epidemiology, Geisel School of Medicine at Dartmouth, Hanover, NH 03756, USA. [6]Program in Quantitative Biomedical Sciences, Geisel School of Medicine at Dartmouth, Hanover, NH 03756, USA. [7]Geisel School of Medicine at Dartmouth, Hanover, NH 03755, USA. [8]Thomas Jefferson High School for Science and Technology, Alexandria, VA 22312, USA. [9]Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [10]University of Virginia, Charlottesville, VA 22903, USA. [11]Stanford University, Palo Alto, CA 94305, USA. [12]University of Illinois Urbana-Champaign, Champaign, IL 61820, USA. [13]Department of Molecular and Systems Biology, Geisel School of Medicine at Dartmouth, Hanover, NH 03756, USA. [14]Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth, Hanover, NH 03756, USA. ✉email: Joshua.levy@cshs.org
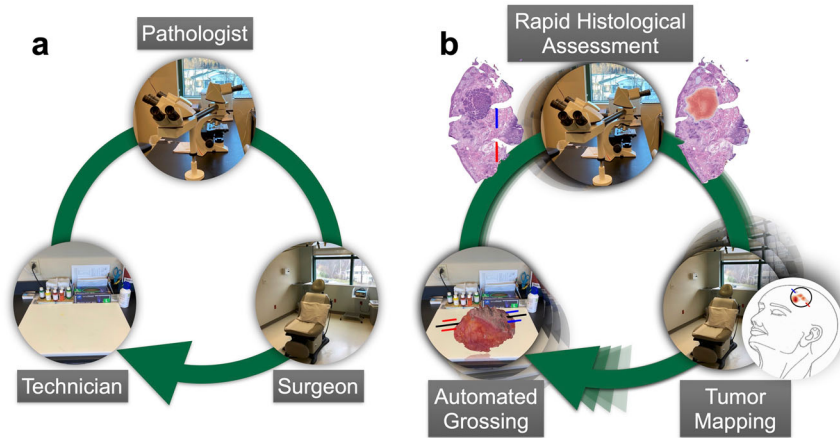
npj

THE HORMEL INSTITUTE
UNIVERSITY OF MINNESOTA

Fig. 1 Visual description of intraoperative surgical excision setting demonstrates potential use cases for integrating artificial intelligence. a Surgeon removes tumor in the operating room, tissue is prepared in the gross room, margins are assessed by pathologist in slide room for margin assessment and findings are mapped back to orientation of surgical site. **b** 3D modeling for automated tissue grossing, computer vision and graph neural networks for margin assessment, and morphing techniques orient histological findings to a surgical tumor map to inform surgeon where to cut additional tumor.

aforementioned) while minimizing the size of the surgical defect and sparing normal surrounding tissue. The advantages conferred by MMS are largely possible due to the size and location of the tumors being excised. These characteristics facilitate the use of local anesthesia and allow the entire margin to be efficiently processed and analyzed. There are numerous obstacles to the application of real-time 100% margin analysis in other surgical practices, including: (1) time under general anesthesia, (2) tissue specimens of prohibitive size and complexity, (3) availability of expert pathologists, and (4) clear mapping of histological findings back to the resection site. Additionally, suboptimal preparation of frozen sections can impact the location of positive margins and is often cited as precluding real-time margin assessment in higher risk settings. Highly trained histotechnicians are required to create quality tissue sections but are in short supply[16]. Thus, investing in methods that can improve the speed of specimen preparation, ensure high-quality tissue sections, and promote rapid and accurate histological assessment of tissue margins are of paramount importance.

Emerging artificial intelligence (AI) technologies have demonstrated the capacity to model complex medical processes and may soon fundamentally transform healthcare delivery through incorporation of non-autonomous diagnostic decision making. These technological advancements have been propelled through the advent of artificial neural networks (ANN) including deep learning methodologies[17]. ANN are inspired by central nervous system processes and represent data input to the algorithm through a collection of nodes, where, given the appropriate activation energy, the signal from these nodes may be passed or shared to a hidden set of nodes organized into multiple processing layers which represent an object through multiple layers of abstraction. For instance, ANN have been widely applied to tasks in digital pathology[18–26], from simulating application of chemical staining reagents[27–34], to predicting prognostic molecular information from digitized representations of tissue slides (Whole Slide Images; WSI), and predicting the origin of tumors with unknown primary site[35]. Recent ANN methods have been proposed for margin assessment across multiple surgical sub-specialties though have only focused on identifying tumor[36–39] while ignoring other issues that are critical to MMS, including: (1) assessing 100% tissue margins intraoperatively, (2) tissue preparation, (3) tissue section quality, (4) mapping findings to surgical tumor site, and (5) operational efficiency.

To address these crucial concerns, a better comprehension of the margin in three dimensions with respect to the surgical site's

position and orientation needs to be developed. Three-dimensional (3D) reconstructive histopathological platforms have been developed to understand the spatial arrangement of structural and functional elements in processed tissue[40–45]. Such tools have already been developed for visualization of physiologic and pathologic liver and prostate tissue. Applications in dermatopathology include iso-surface plot construction of histological BCC sections to understand invasion and growth patterns as well as epidermal inclusion cyst visualization to assess anatomical relationship to surrounding structures. While existing 3D reconstruction approaches (e.g., CODA and VALIS) enable comprehensive tissue characterization at single-cell resolution, the compute time to co-register serial sections far exceeds what is allowable to complete a rapid, intraoperative margin assessment. Furthermore, mapping these results back to correct anatomic position and orientation at the surgical site is non-trivial, as it requires leveraging inking patterns identifiable on the gross tissue/histological sections to establish common coordinate system that is also drawn by the surgeon at the surgical site. The mapping of histological findings at every serial section to a common coordinate system, established by identifying and utilizing inking patterns, can facilitate rapid co-registration and orientation of serial sections. This can further enable the mapping of these histological findings to the surgical site (Supplementary Fig. 1).

The techniques developed for 3D histology can only evaluate tissue after it has been sectioned, not the gross specimen that is measured, inked, and sectioned to prepare the tissue for histological examination. Although recent studies have examined 3D reconstruction of the gross specimen, none have focused on its applications in the intraoperative surgical context[46–51]. Real-time guidance on optimal inking and sectioning patterns, as well as reporting of tissue dimensions, could speed up tissue preparation before histological analysis. However, existing methods provide inadequate reconstructions of tissue at a slow pace, making them unsuitable for practical use. To the best of our knowledge, there are no available platforms that offer 3D reconstruction of the gross BCC specimen, which can aid in tissue grossing and inking to expedite histopathological analysis (Supplementary Fig. 1).

We have designed and developed a non-autonomous digital assessment tool, *ArcticAI*, that can expedite tissue preparation, histological inspection, and tumor mapping to improve solid tumor removal (Fig. 2) using MMS for removal of basal cell carcinoma as a model system. This computational workflow places the surgeon, histotechnician, and pathologist in the same virtual space to: (1) reduce the amount of time a histotechnician takes to
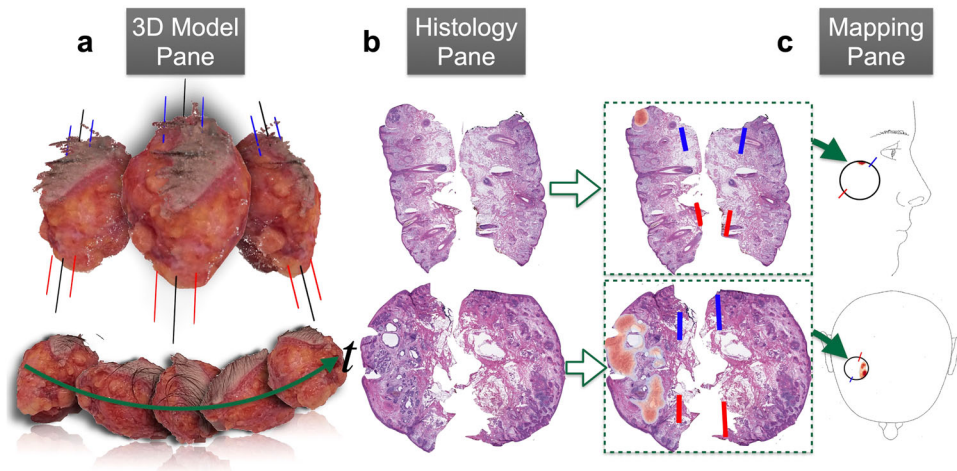
Fig. 2   The margin assessment tool features three panes for modeling the gross specimen, histological assessment of tumor of tumor and mapping histological findings back to the surgical site. Workflow overview: **a** gross tissue measurements and inking recommendations are made using the *3D Model Pane*, which reconstructs 3D models of tissue from video of tissue rotating around a turntable set up, **b** rapid margin assessment is accomplished through the *Histology Pane*, which localizes holes/tears (completeness), tumor, and calculates spatial statistics on ink for orientation (blue is 12 o'clock, red is 6 o'clock), and **c** *Mapping Pane* maps margin assessment results to surgical specimen through morphing to user defined surgical tumor map and by leveraging the orientation calculated in the *Histology Pane* to reorient the results to a format understandable by the surgeon.

process tissue and generate pathology reports through 3D modeling techniques and smart grossing recommendations (e.g., reporting of tissue size and where to ink), (2) improve the efficiency of pathologic analysis through a collection of sophisticated graph neural networks to map tumor and artifacts on whole slide images (WSI) acquired from serial tissue sections, and (3) automatically generate a descriptive and visual pathology report easily interpreted by the surgeon either in real-time or postoperatively. A web application demonstrating 3D specimen grossing recommendations, histological examination, and mapping of histological results back to the surgical site for a few select cases can be accessed at the following URL: https://arcticai.demo.levylab.host.dartmouth.edu.

RESULTS
Data collection and study population
After Institutional Review Board approval (see *Methods* section "Study Design"), we assessed specimens from 194 patients undergoing tumor excision in the Mohs Micrographic Surgery (MMS) setting for the treatment basal cell carcinoma (BCC). Tissue from 16 patients (17 specimens) were used for tissue grossing algorithms, while tissue from the remaining 178 patients were used for histological assessment and tumor mapping algorithms. All specimens first underwent accessioning and gross measurement. For the tissue grossing algorithm, the gross specimen was placed on a turntable and imaged using low resolution video capture. The remaining cases underwent grossing, inking, processing, cryoembedding (frozen section), sectioning, and staining with hematoxylin and eosin (H&E). From these 178 cases, 351 slides corresponding to 1065 serial sections and 1537 tissue pieces were scanned at 20X resolution using the Leica Aperio AT2 scanner and stored as Whole Slide Images (WSI) in either SVS or TIFF file format. A total of 3,754,730 image patches (256×256 pixels) were extracted from the WSI for further analysis. Using the ASAP annotation software (https://computationalpathologygroup.github.io/ASAP/, v1.9), all cases were annotated for tumor (BCC), benign structures, inflammatory aggregates, holes/tears, ink color and location, and major compartments (epithelium, dermis, fat, etc.). The data was then divided into training/validation sets (65% of cases) responsible for

algorithm training and finetuning and a held-out test set (35% of cases) (Supplementary Table 1). Follicles and individual nuclei were annotated in a subset of the training/validation slides as annotation of these smaller structures on all training/validation slides was intractable.

Results overview
In the following subsections, the impact of a digital assessment on the surgical workflow will be demonstrated through description of the accuracy and execution time of: (1) tumor removal and specimen preparation, (2) histological assessment, and (3) tumor mapping.

Tissue preparation results
Following tumor removal, a specimen is sent to the pathology laboratory for accessioning, grossing, inking, sectioning, and staining. Tissue grossing and inking decisions are made by the histotechnician depending on the size and shape of the tissue specimen. These decisions are not standardized and require a high level of training and rigorous documentation. To determine if tissue characteristics could be autogenerated, seventeen surgical specimens were collected, and the superior pole was delineated by either tissue ink or placement of a suture. Three-dimensional reconstructions of excised tissue combined multiple views of the tissue through captured smartphone videos for one revolution around a turn table setup (Fig. 3a, b; Supplementary Video 1).

Tissue size measurements. The 3D tissue model was used to measure the tissue proportions (length, width and height), compared to manual measurements[52] (Fig. 3c). Representing the tissue specimen as a 3D point cloud, automated tissue measurements of the reconstructed tissue varied by 0.29 cm on average as compared to manual measurements for each tissue dimension (Table 1, Supplementary Table 2; Supplementary Video 1). Our results improved further when using a neural network modeling-based approach (neural radiance fields, NeRF)[53]—automated measurements of reconstructed tissue varied by 0.22 cm from manual measurements and took less than 30 s per specimen to render as compared to rendering the 3D point cloud, which took over one and a half minutes per specimen (Table 1, Supplementary Tables 2, 3; Supplementary Video 2).
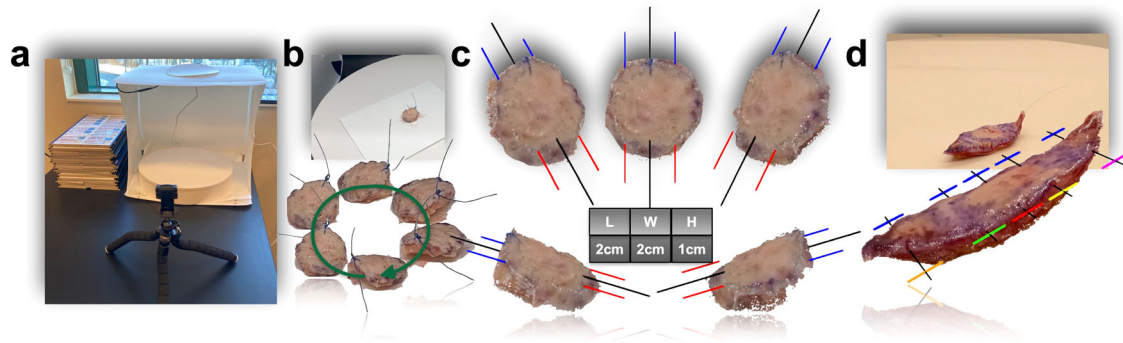
Fig. 3 Tissue preparation workflow enables precise 3D modeling of a gross specimen for accurate tissue measurements and tailored grossing recommendations. Tissue Grossing Measurement and Recommendations via the *3D Model Pane*: **a** image of turntable setup, where phone camera is placed on mount to record gross specimen revolve around table; **b** still frame from phone video of rotating tissue specimen; also depicted on the bottom are automated segmentations of the tissue and suture using the *3D Model* preprocessing subroutine, images were selected from a representative set of still frames to demonstrate multiple object viewpoints; **c** these viewpoints are integrated together for 3D reconstruction of the gross specimen, pictured here are screenshots while using the interactive *3D Model Pane* to rotate the specimen to various orientations; inking recommendations are deposited via the addition of red/blue and black lines, which denote inking of 12 o'clock (blue) near the suture (which has been removed) and 6 o'clock (red) after bisecting the tissue (black); length, width and height measurements are automatically reported on the scale of centimeters while operating the display; **d** tissue grossing and ink recommendations for radial sectioning of a wide local excision specimen.

Inking recommendations. Taking into account: (1) autogenerated tissue size, (2) preferred grossing approach (MMS versus radial) as dictated by surgeon/pathologist, and (3) size of a glass slide, automated grossing and inking measurements/recommendations were generated to maximize the amount of margin assessed per tissue block/slide. This resulted in lines being placed through the 3D model recapitulating expert domain knowledge, using suture/ink locations for guidance. For MMS specimens, the algorithm placed 3D black lines, which identified the location of grossing cuts, and blue and red lines, denoting 12 and 6 o'clock ink placement, respectively (Fig. 3c). Figure 3d demonstrates tissue grossing and ink recommendations for radial sectioning of a wide local excision specimen. Grossing cuts are identified by black lines through the body and two tips of the specimen based on tissue size. Inking recommendations include unique color combinations for each tissue piece allowing multiple pieces to be put into a single cassette, resulting in fewer tissue blocks for the histotechnician to section.

Histological assessment results

Tissue completeness assessment. Effective histologic analysis of tumor margins relies on high-quality tissue sections that are devoid of holes or tears. In the absence of a complete tissue section, it is not possible to definitively declare a margin free of tumor. To address this, a 'Tissue Completeness' algorithm was developed using a combination of convolutional and graph convolutional neural networks (CNN-GNN)[54]. The algorithm was trained and validated on 381 annotated tissue sections using PyTorch to segment holes/tears (tissue artifacts) in the tissue[54,55]. The tissue completeness algorithm was trained to delineate between the following macro-architectural features: (1) holes/tears, (2) epidermis, (3) dermis, and (4) subcutaneous fat. The GNN algorithm successfully identified tissue defects in our test set with an AUC of 0.84 (Table 1). Interestingly, the CNN model initially trained to identify holes and tears demonstrated better performance than the GNN in identifying incomplete tissue sections (test-set AUC = 0.92 vs. 0.84). An example output of the 'Completeness' algorithm is shown in Fig. 4a, b, where sporadically placed holes/tears are highlighted by the algorithm, while regions of fat or significant gaps introduced by hair follicles or less structured dermis are ignored (Supplementary Fig. 2).

Tumor localization. To evaluate for the presence or absence of tumor, a CNN-GNN was trained and validated on 1065 tissue sections containing a variety of BCC histologic subtypes reflective of clinical practice, including the following subtypes: nodular, superficial, infiltrative, micronodular, and sclerodermiform (Supplementary Table 1). Annotation subgroups included: tumor, benign skin structures (e.g., hair follicles), and cell populations that may be confused with or a harbinger of tumor (e.g., inflammation). With the aim of enhancing tumor classification specificity, a comparative analysis involving two CNN-GNN models was conducted within an internal validation set for model selection—one focused on distinguishing tumor from benign tissue (two-class), and the other incorporating inflammation as a distinct class (three-class). The exclusion of inflammation from the model elevated the risk of misclassifying inflamed regions as tumor, thereby diminishing specificity. Inclusion of inflammation modeling, however, not only bolstered accuracy (AUC = 0.98 versus 0.96 for three and two classes respectively) but also significantly curtailed tumor misclassification within inflamed regions (4% versus 38% misclassification of inflamed regions as tumor). This approach reinforces the critical role of modeling inflammation in refining tumor classification outcomes (Supplementary Table 4). An image detection model removed hair follicles from conflated tumor-predicted regions (Table 1, Supplementary Fig. 3). On a test set on 121 held out slides, the CNN-GNN obtained an AUC of 0.97 for the task of tumor localization across sections (Table 1; Supplementary Video 2). These results varied by histological subtype (Supplementary Table 5)– for instance, achieving an AUC of 0.96 for superficial BCC and 0.98 for squamatized and nodular BCC tumors. Across all subtypes, results reveal that the GNN exhibited superior performance (AUC = 0.97) compared to the CNN (AUC = 0.90) in detecting tumors at a 50-micron resolution. Example displays of the tumor detection output across serial sections of two test set cases is shown in Fig. 4a, b. We have included example displays of hair follicle and inflammation-predicted regions on held-out slides in the supplementary information file (Supplementary Figs. 3, 4), which further demonstrate how exclusion of these regions can inform tumor localization.

Nuclei detection and classification. To rule out rare tumor cells in regions predicted to be inflammatory aggregates, a cell detection neural network and graph neural network were trained and validated on 32,763 cell annotations to provide high-resolution tumor maps designating precisely which cells correspond to the

Table 1. Model performance and concordance with pathologist/surgeon annotations for the gross measurements, tissue orientation/mapping, completeness, and margin assessment tasks.

Task	Evaluation metric	Subtask	Estimate	2.5% CI	97.5% CI
Gross Measurements	3D Point Cloud: MAD[a] (cm)	L	0.36	0.21	0.53
		W	0.23	0.12	0.52
		H	0.29	0.19	0.50
		Overall	0.29	0.2	0.47
	NeRF: MAD (cm)	L	0.21	0.11	0.53
		W	0.19	0.17	0.22
		H	0.27	0.17	0.36
		Overall	0.22	0.18	0.33
	3D Point Cloud: MPC[b] (%)	L	12.4	7.9	17.7
		W	13.0	8.3	32.0
		H	31.7	20.0	58.0
		Overall	19.0	12.6	33.2
	NeRF: MPC (%)	L	11.3	4.5	14.3
		W	11.3	8.3	16.9
		H	34	6.5	58.3
		Overall	18.9	8.2	28.5
	3D Point Cloud: Correlation	L	0.848	0.544	0.973
		W	0.833	0.467	0.971
		H	0.812	0.456	0.955
		Overall	0.831	0.602	0.937
	NeRF: Correlation	L	0.946	0.814	0.988
		W	0.875	0.576	0.976
		H	0.893	0.646	0.984
		Overall	0.905	0.782	0.96
Tissue orientation	MAD (°)		4.883	4.04	5.451
	Proportion Correct Orientation	≤45°	94.7%	92.3%	96.6%
		≤15°	86.0%	82.6%	88.9%
		≤5°	51.3%	46.7%	55.7%
GNN: Tissue completeness	AUC		0.839	0.825	0.855
	Macro-AUC		0.851	0.839	0.863
GNN: Margin assessment	AUC	Original Slides	0.967	0.960	0.979
		Follicle Removal	0.965	0.959	0.975
	Macro-AUC	Original Slides	0.962	0.954	0.970
		Follicle Removal	0.957	0.949	0.964
CNN: Tissue completeness	AUC		0.923	0.917	0.930
	Macro-AUC		0.916	0.907	0.925
CNN: Margin assessment	AUC		0.900	0.887	0.912
	Macro-AUC		0.877	0.859	0.893
Tumor Mapping	Accuracy		0.992	0.915	0.999

Macro-AUC represents reporting of AUC statistic on slide level and averaging across slides, giving each slide equal weight, while normal AUC statistic is calculated for subimages across all slides. 95% confidence intervals were acquired using 1000-sample non-parametric bootstrap, where bootstrapping was done on the WSI level to account for variation in concordance across the cases.
[a]Median Absolute Deviation.
[b]Median Proportion Change.

BCC annotation subgroup[56,57]. The cell detection neural network was used to locate nuclei in the slide, recording their positional coordinates. Based on their positional coordinates, a k-nearest neighbors graph was constructed. Individual cells were represented as nodes in the graph– numerical information extracted using a CNN represents their nodal attributes and cells are connected (i.e., edges) by their k-nearest neighbors. Cell graph neural networks pass messages between adjacent cells to provide contextual information, capturing the relationships between different cell populations in the tissue, including tumor cells and surrounding inflammation, which can be used to improve the accuracy of our predictions. Annotation subgroups to train the cell detection algorithm included: "BCC", "hair follicle", "inflammatory", "fibroblast", and "epidermal keratinocyte". Results demonstrate the ability to accurately localize cells (Dice=0.85) in an internal test set, while predicting with high accuracy the corresponding cell type (F1-Score=0.86) (Supplementary Table 6; Supplementary Fig. 5).
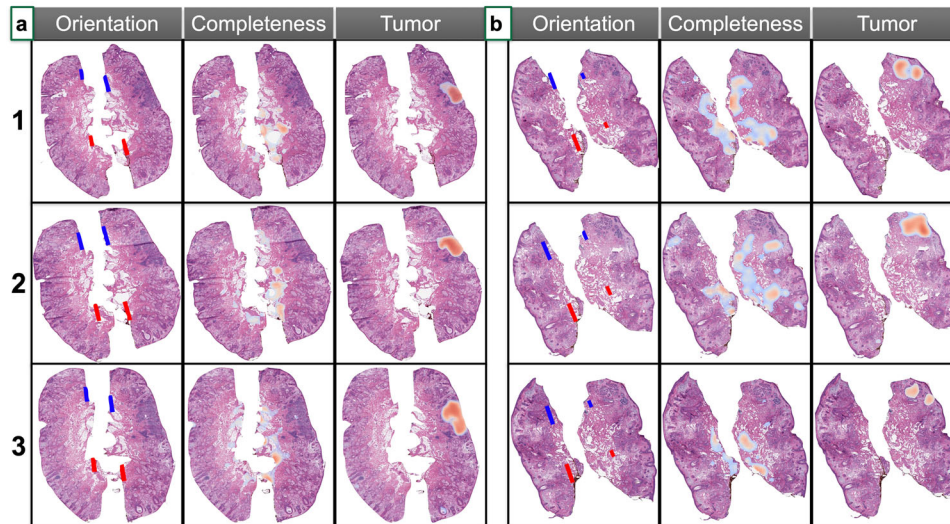
Fig. 4 Histological margin assessment workflow captures tissue orientation, completeness, and tumor localization in serial sections. Margin assessment via the *Histology Pane* for two test cases: Patients 1 and 2 represented by panels (**a**) and **b**, respectively—three serial sections per patient are numbered accordingly (1,2, and 3). Results are plotted on top of each WSI for assessments of orientation, tissue completeness, and tumor localization. Screenshots of WSI from the *Histology Pane*. For tissue orientation, blue and red lines are drawn over center of mass positions to define 12 o'clock and 6 o'clock respectively for each tissue section. High resolution completeness and tumor results represented by thresholded heatmaps (where patches removed from display if failing to surpass probability threshold), where red indicates whether part of tissue is incomplete or positive margins, and blue indicates lower yet non-negligible probability of incompleteness/tumor.

Mapping histological findings to surgical tumor site

Tissue orientation. Histologic ink location is critical to tissue orientation and subsequent tumor mapping. In this study, MMS specimens were inked blue (12 o'clock), and red (6 o'clock) which was reflected in the surgeon's hand-drawn diagrams. These terms (e.g., 12 o'clock, 6 o'clock) are commonly used in histopathology to indicate the position of the tissue section on the slide in relation its anatomic positioning and orientation. For instance, the surgeon will indicate the anatomic position of the examined tissue. After histological sections are examined, this information needs to be mapped back to the original location of the tumor and rotated correctly (using inks to denote the tissue's orientation on the tumor map). This step is important to enable the surgeon to utilize the histological information during excision without having to rely on constant communication with the pathologist (typically over phone). Ink detection was automated through segmentation of tissue edges through filters, color thresholding, and connected component analyses to remove spurious applications (i.e., where ink is erroneously applied / seeps)[58]. Subsequently, a line was plotted between detected blue and red ink on tissue sections and stored for use later to calculate the relative orientation to lines drawn either with surgeon annotated inks on histological slides as comparison or to inks drawn in the surgical tumor map for mapping histological results back to the specimen (Figs. 4, 5). On a subset of held out test slides, the relative angular difference between predicted and surgeon-annotated blue-red ink lines were measured and compared. Findings indicate that 95% of tissue sections were oriented with less than 45° difference between annotated/predicted lines with an average relative angular difference of 4° (Table 1, Supplementary Fig. 6a, c, Supplementary Table 7). More than 85% of sections were oriented correctly with less than a 15° angular difference, deemed an acceptable amount of variation for accurate tissue orientation. Complete performance characteristics can be found in Supplementary Table 7. Sections without correct orientation demonstrated relative lack or spurious applications of ink, highlighting the importance of proper tissue inking (Supplementary Fig. 6b)[59].

Tumor mapping. Accurate tumor mapping is critical to inform additional tumor removal if needed. Tumor mapping relies on anatomic identification of surgical site, accurate tissue size measurement, tissue orientation, and tumor identification. To build a tumor map, a template is selected by the surgeon based on the anatomical surgical site with blue and red lines used to indicate 12 o'clock and 6 o'clock, respectively (Supplementary Fig. 7). The tissue sections are then fit to the tumor map through an algorithm that morphs and rotates the WSI histologic tissue sections into the shape and orientation (i.e., aligned blue-red ink line) at the anatomic site as drawn on the templated map (Fig. 5a–c illustrates mapping at arbitrary orientations, Supplementary Videos 2, 3 demonstrates concordance between hand-drawn and digital tumor maps)[60–62]. To determine the accuracy of automated tumor mapping, 28 test set cases (selected at random from cases with known positive margins) were used to compare platform generated tumor maps to the surgeon hand drawn maps. This showed 99.2% (95% CI: 91.5–99.9%) correspondence between the surgeon and algorithm generated maps, respectively (Supplementary Figs. 8–12, Table 1, Supplementary Video 3, Supplementary Data 1).

Margin assessment speed

For broad applicability of this approach for tumors in patients where anesthesia is required, the platform must perform with efficiency and speed[63]. This was accomplished by parallelizing the histological processing workflow across all tissue sections and WSI for each case (Supplementary Fig. 13a; Supplementary Table 8). Overall, margin assessment using this platform across the entire test set (*n* = 41 cases, 121 slides) had an average execution time of 72 s per slide and 78 (95% CI: [66–88]) seconds per case (i.e., many slides/sections per case), consisting of 48 s for preprocessing and 24 s for parallel performance of image stitching, CNN-GNN analysis, and tissue orientation (Supplementary Fig. 13b). Execution of the platform in series would take five to seven times longer than parallelized, 494 (95% CI: [367–553]) seconds per case. For computational systems with lower computational power (i.e., no
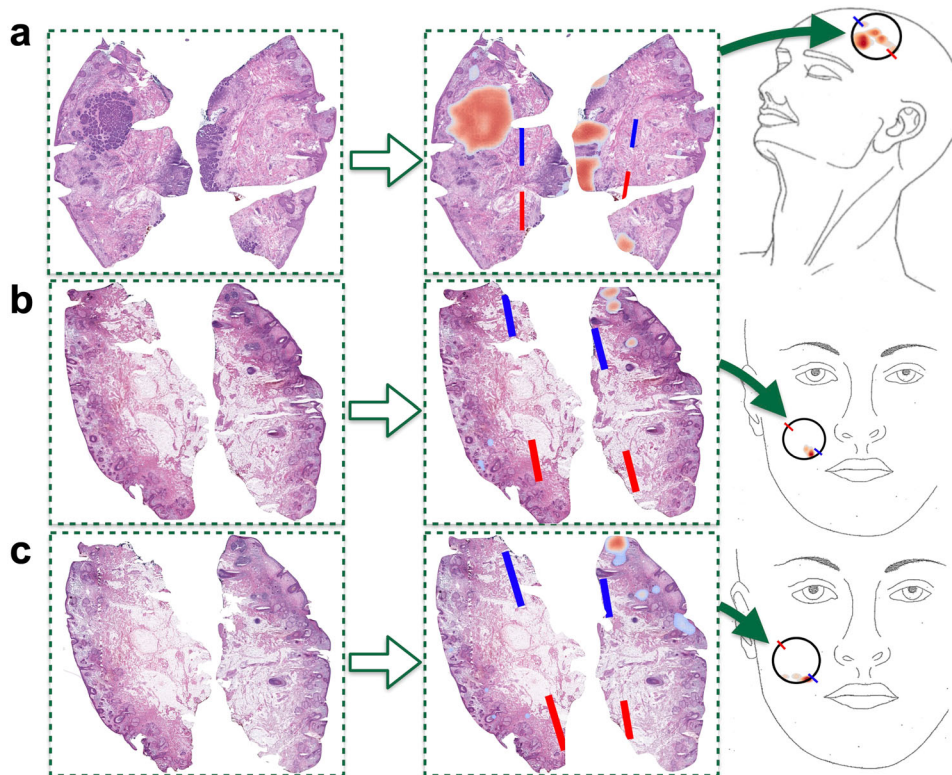
Fig. 5 Surgical tumor mapping workflow transforms histological findings into patient-tailored, surgeon-drawn surgical tumor maps in correct anatomical location/orientation for real-time surgical recommendations. Mapping margin assessment results to surgical tumor map via *Mapping Pane*: representation of workflow using three separate sections, the first (**a**) from one case, and the second two (**b**, **c**) are serial sections in another case; first, margins are assessed via *Histology Pane* and tissue orientation and tumor localization results are plotted over the WSI; then, results are mapped to surgical tumor diagram selected by the user (**a** features top of scalp, while case **b**, **c** are of front of face), where circle is drawn by user to represent surgical site anatomic location and an arbitrary orientation is defined via user drawing of blue/red lines. Note how tumor results are morphed and rotated to match circle interior and orientation in surgical tumor maps, where density map in surgical tumor map represents tumor at user defined threshold. For **b**, **c**, note how tumor is automatically rotated close to 180 degrees to preserve orientation of margin on surgical map.

GPU), we have developed and timed execution of a CPU-based workflow, which had an average execution time of 96 (95% CI: [79–133]) seconds per case when computing on tissue sections in parallel (92 s per WSI; 49 s for parallel performance of image stitching, CNN-GNN analysis, and tissue orientation). Execution in series through a CPU based workflow would take significantly longer than parallel execution, requiring an average of 1392 (95% CI: [907–1817]) seconds per case (Supplementary Table 9).

DISCUSSION
Tumor excision with real-time intraoperative 100% margin assessment results in low recurrence rate and efficient delivery of surgical care in MMS. Real-time margin analysis and/or expanding the area of assessed tissue margins has the potential to eliminate adverse outcomes associated with post-operative positive margins (which necessitate additional treatment; i.e., repeat procedure) and false-negative margins (which can lead to recurrence) across surgical oncology procedures. However, broad applicability of real-time total margin analysis is relatively limited outside the MMS setting. There are several logistical constraints that contribute to this including separation of multiple experts in time and space, inefficient manual laboratory processes, and labor-intensive pathologic analysis of histologic specimens. Factors such as the size and location of the tumor, as well as access to real-time pathologic care, can impact the feasibility of real-time margin analysis. In this study, a rapid tissue margin assessment tool was designed and tested to address the rate

limiting steps in the current surgical tumor removal workflow. The software developed in the present study aims to facilitate the use of MMS (100% and real-time) margin analysis in a wider range of surgical contexts, although we acknowledge that MMS is not a replacement for all resection protocols. Additionally, it can be used to support the current standard of care, rather than replacing it entirely, by providing more accurate and efficient margin assessment.

The developed digital assessment tool performs automated tissue measurements aimed at improving laboratory workflow through efficient grossing and inking recommendations. These recommendations aim to maximize the amount of tissue per block while decreasing the number of tissue blocks to be cut by the histotechnician. However, it is important to note that the platform does not aim to reduce the amount of time for the histotechnician to gross and ink, embed, section, and stain the tissue, as these rate-limiting steps require additional innovation to further improve surgical care delivery. Additionally, unique predetermined ink combinations allow the tissue sections to be reconstructed and mapped to the 3D tissue model. These features have the potential to standardize grossing and inking in the surgical pathology laboratory thereby decreasing the time of processing and required level of expertise. Future works will attempt to encapsulate this 3D gross specimen modeling functionality into an "augmented reality" cell phone application that provides both tissue dimensions and where to section/ink for MMS/radial grossing. However, the development and validation of this application are outside the scope of this study.

Histotechnicians are highly trained and currently in high demand in our healthcare system. Increasing the efficiency of the histotechnician as well as decreasing the training required to expertly process and section a tissue excision specimen are two solutions to address the demand for more histotechnicians[64]. In addition to the importance of the 'Completeness' algorithm for tumor margin analysis, this digital assessment aid can also be used as a training tool to assess the competency of histotechnicians either in training or as part of an annual review of performance. Taken together, the platform provides significant support in training, standardization, and workflow efficiency for histotechnicians.

The current study resulted in an AUC of 0.97 for tumor detection at 256-pixel resolution (~64 microns at 40x resolution), indicating high accuracy in identifying/localizing remaining tumor, which is anticipated to increase with subsequent and expanded training sets. These results are comparable to findings from previous studies, which have reported AUCs between 0.9 and 0.99, validated at different microscopic magnifications/resolutions (e.g., patient level presence of BCC), which can make it challenging to place these study findings in the context of the prior art[36,37,65,66]. Positive margins were correctly identified/located in the correct anatomical orientation/position in 99.2% of our test set cases. Our platform does differ in the following ways from previous studies: (1) 3D gross specimen reconstruction/recommendations, (2) tumor completeness assessment, (3) addition of a high-resolution cell-level analysis, (4) identification of follicles and inflammatory regions, and (5) mapping histological findings back to the surgical site by using inking patterns for orientation. Furthermore, previously reported compute times for margin assessment (e.g., 13 min for 3 sections)[65] were significantly slower than that reported in the present study (78 s for more than 20 tissue sections). Identification of specific histologic BCC subtypes or normal structures that present particular challenges for the CNN-GNN will help to identify surgical cases that will be most impactful for further improvement of the algorithm. In general, we found that our algorithm performed well across different histological subtypes. Interestingly, this study elucidates the importance of the relationship between tumor and additional cell populations including surrounding inflammation as an indicator of the presence of tumor. Further delineation of the tumor niche involves the classification of associated/confounding cell types and stromal changes which if unaccounted for could reduce the specificity of the algorithm. This step is essential in order to enhance the specificity of the algorithm to delineate tumor from surrounding benign tissue or structures such as hair follicles which can have similar structures and nuclear morphologies to BCC. Accurately identifying stromal changes and associated cell populations at tumor margins is crucial to avoid overcalling tumor or positive margins– overcalling may result in a less specific model. The creation of CNN-GNN for the margin analysis of other solid tumors will require the identification of both varied histologic tumor types and tissue specific cell populations or tissue structures that aid in tumor identification.

MMS is possible because the surgeon removes the tumor and examines the slides in a laboratory in close proximity to the operating room. In other surgical settings, performing frozen section margin analysis requires an onsite pathologist to assess the slides and relay the results back to the surgeon in the operating room. In these settings the laboratory and operating room are separated by time and space. This prevents the use of frozen sections in many healthcare settings, particularly smaller rural hospitals, where caseload or demand may not support the presence of an on-demand highly trained expert pathologist. By creating an algorithm that enables rapid and accurate identification of tumor combined with a virtual platform allowing for remote whole slide imaging and result viewing, this software obviates the pathologist being in physical proximity to the operating suite. This allows a pathologist with expertise in one particular tumor type or organ system to maintain a high case load while providing highly specialized pathologic care to healthcare settings that might not otherwise have such access. High quality, complete tissue sections are critical to accurate pathologic analysis and prior research has demonstrated that tissue holes/tears, common to frozen sections, have been shown to be the largest contributor to cases of local tumor recurrence[67,68]. If the margins come back clear with incomplete tissue/tears at the margin of the resected specimen, it raises the possibility of false negatives and additional assessment is required– irrespective of whether an algorithm was used to assess tissue completeness, similar to existing procedures for margin assessment. Integration of the 'Completeness' algorithm, which identifies holes and tears, will determine low quality or incomplete sections prior to pathologic analysis and allow the histotechnician to create additional sections as needed, prior to final pathologist review. This will minimize recuts and allow rapid sign out and reporting. Integration of the vast amount of data present in a pathology report through automation will decrease the amount of work on the back end and also provide both written and visual outputs that can be used either in real-time or post-operatively. Excessive charting and documentation result in pathologist burnout, limiting the amount of manual documentation will both increase productivity and decrease administrative burden[69–71].

For skin cancer of the head and neck, which is more challenging to assess than the model system featured in this work (BCC) due to increased tissue size and complexity, decreased positive margin and recurrence rates have been shown with real-time complete margin analysis[72–79]. As many healthcare systems are functioning with decreased staffing and disruption of supply chains, delivery of efficient surgical care is critical for patient access and maximizing hospital resources. Limiting positive margins, tumor recurrence, or the need for adjuvant treatments will decrease the burden on the surgical and medical system[80–82]. Access to remote pathologists using informatics augmented platforms will allow both hospitals that may otherwise not be able to offer a surgical service to do so and increase the productivity of remote pathologists. Providing histotechnicians with a platform that automates their tedious tasks and makes grossing/inking recommendations that are reflected in the pathology report will allow them to focus their time and energy on embedding and sectioning the tissue, thereby increasing the number of specimens processed. Taken together, the use of technology in the delivery of surgical care will not just provide better outcomes for the patients, but also improve the efficiency of surgical care delivery in an unprecedented time of resource shortages (e.g., access to care).

Limitations of this study include that it was performed at a single site. Whole slide images in the training, validation, and test set were generated in a single laboratory with a standardized sectioning and staining protocol. Therefore, the next steps will include creation of an external test set from outside MMS units. This will be accomplished through a multicenter study, currently outside of the scope of this work, which is proof-of-concept informatics-augmented surgical workflow. The innovations featured and validated in this study establishes feasibility of a larger multi-center study in the future. Validation of this workflow in this context and in the real-world setting will also require adapting virtual inking recommendations to the specific requirements of each hospital, including their existing standardized inking protocols.

As this proof-of-concept workflow was developed and validated on BCC subtypes only, it will be necessary to incorporate the detection and localization of collision lesions or other skin cancer subtypes (e.g., Squamous Cell Carcinoma; SCC; Melanoma in Situ, etc.) into this workflow to account instances where these tumors may occur in the same field. Since incidental diagnoses are not infrequent, future work is planned to identify additional tumor

types to assess for this possibility. For these tumors, our approach does not yet account for other histological structures such as actinic keratosis (AK) and perineural invasion patterns which are more common to these other skin tumor subtypes (e.g., aggressive SCC; tumor differentiation) and are relevant considerations when improving our approach. Furthermore, while we accounted for the presence of resident hair follicles, which are often challenging to distinguish from BCC on the nose, through a "post-hoc" adjustment using a follicle detection algorithm, the follicle detection algorithm will be improved in future iterations of this work through additional data collection in various tissue sites (e.g., nose). We acknowledge that there are additional algorithmic methods (e.g., image transformers, self-supervised pretraining) which could be leveraged and compared to improve the precision of the histological assessment, which will be the subject of future work, though such comparisons were outside of the study scope in favor of illustrating how these components (gross specimen 3D modeling, histological assessment, and tumor mapping) can be integrated together to augment current surgical practices.

Finalized tumor detection and completeness algorithms will likely require an input of whole slide images from multiple laboratories. Alternatively, sites aiming to use the platform could adopt a standardized workflow including reagents and process similar to those used to generate the tissue sections incorporated into the algorithms in this study. Obstacles to the usage of the platform include the availability of whole slide scanners, as these are currently costly to obtain, and large file uploads which requires robust computing infrastructure or a workstation capable of handling high throughput assessment. With time, the cost of scanners will decrease, while computing power will increase with advances and cost reductions in graphics processing units (GPUs). Timely tissue processing and analysis is critical to seamless integration of the platform into the surgical workflow. The timing featured in this study considers parallel execution of workflow elements in optimal computing infrastructure. However, many high-performance computing environments are bottlenecked by the time it takes to submit and start simultaneous compute jobs as well as communication bottlenecks which may be workflow specific. In future studies, all aspects of the surgical workflow and platform including: (1) tissue transport and processing, (2) slide scanning, (3) image upload and processing, and (4) pathologist review, will be timed via a simulated clinical trial to provide practical time estimates. We have performed an initial evaluation on the downstream effects of the algorithm on staff efficiency, which has demonstrated that automated histologic analysis results in significant reductions in staff waiting time and more efficient delivery of surgical care. While these findings are outside of the scope of the current study, which has focused on the technical innovations introduced through this algorithmic workflow, initial findings have demonstrated time savings of more than an hour per case. These findings will be discussed in detail in a follow-up work focused on operational efficiency improvements conferred through usage of this approach.

As multiple serial sections are assessed using this method, corresponding WSIs could be co-registered to facilitate a 3-dimensional histopathological assessment to visualize the spatial arrangement of structural and functional elements within a single processed tissue. While this is a consideration for follow up work, 3D reconstruction applied in this context will drastically slow down our workflow, as sections still need to be mapped to the anatomical orientation and position after a nuanced, complex, and time-intensive co-registration process. We instead opted to register tissue findings at each serial sections to a common coordinate system defined by the inking patterns (e.g., 12 o'clock and 6 o'clock) which could be more easily mapped to the patient as defined by the surgical tissue map, which can summarize information across multiple sections to report the location of true positive margins. We would expect 3D co-registration of histological sections, once mapped to the patient, to perform similarly as mapping individual sections to a common coordinate system, at slower execution times.

Complete surgical removal of solid tumors remains most patients' best chance at achieving a cure. In this study, MMS removal of BCC is used as a model system to highlight the integration of informatics technologies, including the incorporation of artificial intelligence where appropriate, into the surgical workflow to address critical bottlenecks that might otherwise prevent real-time and/or complete tumor margin analysis. This model has the ability to improve surgical care delivery through technology driven standardization and automation as one approach to solve the significant labor and resource shortages and mismatches in the current system. This can be accomplished by: (1) improving the efficiency of the individuals and processes within the system and (2) increasing the number of individuals capable of performing a critical task. Nonetheless, adopting a digital aid requires stakeholder buy-in and a readiness for changing established practices, which carries significant barriers for entry. Dissemination and implementation of such technologies requires educational alignment and qualitative assessment of stakeholder interests and values. In order for such technology to be adopted, it will need to demonstrate significant improvement in efficiency over traditional methods while meeting the needs of surgeons, pathologists, and histotechnicians.

METHODS

Study design

The overall goal of our study was to demonstrate where artificial intelligence technologies could provide efficiency gains for intraoperative margin assessment by: (1) generating a 3-dimensional model of tissue for grossing recommendations, (2) rapid localization of remaining tumor, and (3) mapping histological findings to a display output familiar to the surgeon. We first compared tissue gross measurements (length, width, height) to computer generated measurements. Then, we calculated AUCs to communicate the accuracy of a CNN-GNN for tumor localization. Finally, the surgeon reported the number of cases which exhibited concordance between hand-drawn and computer-generated tumor maps at the correct anatomical position and orientation. Throughout the study, we consulted with end-users (pathologists, surgeons) for the design of the digital displays, and we utilized an independent test set of randomly selected cases. Patients were randomized into training, validation and held-out test set cohorts (i.e., serial sections and WSI subimages were included in the same cohort to avoid target leakage and inflating test set statistics). Sample size was based on data availability to calculate concordance and accuracy. Confidence intervals for the study findings were calculated using nonparametric bootstrapping and posterior distributions. The authors complied with all relevant ethical regulations including the Declaration of Helsinki. Human Research Protection Program (institutional review board, IRB) of Dartmouth Hitchcock Medical Center gave ethical approval for this work. All necessary patient/ participant consent was obtained, including written consent, and the appropriate institutional forms have been archived, and that any patient/participant/sample identifiers included were not known to anyone (e.g., hospital staff, patients or participants themselves) outside the research group so cannot be used to identify individuals.

Technology overview

ArcticAI is an AI-based software platform for the rapid assessment of tumor margins. The functions of *ArcticAI* are encapsulated in several modules including:

Tissue grossing via the 3D Model Pane (Fig. 2a)*:* When tissue arrives at the pathology laboratory, it undergoes accessioning, description, measurement, grossing, inking, processing, embedding, sectioning, and staining prior to pathologic analysis. To expedite this process, we have prototyped a mobile application that takes multiple images / video of the tissue and synthesizes them to form a 3D model of the tissue. This allows the system to: (1) determine tissue size (e.g., length, width, height) and orientation automatically and add this data to the pathology report, (2) create optimal grossing guides for the histotechnician, and (3) create optimum inking diagrams for the specimen (e.g., blue ink indicates piece is at 12 o'clock; ink used to establish a "coordinate system" for tissue "map").

Histological Assessment via the Histology Pane (Fig. 2b)*:* Following tissue processing, slides are scanned to generate high resolution WSIs which are uploaded into the ArcticAI platform where they are assessed for (1) tissue orientation by detecting inking patterns, (2) tissue quality assessment (e.g. holes and tears from processing and sectioning), and (3) presence or absence of tumor, where (4) tumor confounders (e.g., identification of hair follicles) and (5) nuclei are classified to provide further clarification of histological findings (e.g., residual tumor within large pockets of inflammation).

Mapping of Results to Surgical Specimen via the Mapping Pane (Fig. 2c)*:* Outputs of data from the aforementioned algorithms, notably tissue inking/orientation and the presence or absence of tumor, are used to automatically transpose tumor predictions onto hand-drawn surgical maps. Automated mapping has the advantage of providing the precise location of remaining tumor to inform the surgeon if and where additional tumor needs to be removed. A pathology text report with information on tissue preprocessing is automatically generated and piped to the patient's electronic health record. This information is communicated back to the surgeon, and tumor mapping results (graphics which resemble surgeon drawn tumor maps) are exported to the EHR system to update the automatically generated pathology report.

Workflow automation: Intraoperative resection with 100% margin analysis typically involves the inspection of 6–10 serial tissue sections and can take upwards of 30 min per patient under general/local anesthesia. *ArcticAI* was optimized to reduce histological inspection and tumor mapping time using a sophisticated workflow engine that can be executed in both high performance computing environments and local workstations using Toil and Singularity. The pipeline is additionally capable of processing multiple tissue sections across multiple whole slide images in parallel.

Web Application: Histotechnicians, pathologists, and surgeons can interact with the results in real-time as an interactive/exportable pathology report through a dynamic web application which contains the following panes: (1) Case upload and execution (*Selection Pane*), (2) 3D specimen modeling and pathology report generation (*3D Model Pane*), (3) histological findings and quality report (*Histology Pane*), (4) tumor mapping and orientation to surgical specimen (*Mapping Pane*).

ArcticAI Software Framework

The aforementioned functionality of *ArcticAI* is accomplished through a self-contained software framework, comprised of:

1. A *pip-installable* Python package (*arctic_ai*) which contains an Application Programming Interface (API) and command line interface (*CLI*) that are organized into a collection of modules:

    a. 3D tissue modeling via photogrammetry (*arctic_ai.model_3d*)

    b. Tissue Preprocessing (*arctic_ai.preprocessing*)
    c. Histological Findings
        i. Tissue Quality Prediction (*arctic_ai.cnn_embeddings*, *generate_graph*, *gnn_prediction*, set to *macro_map* mode)
        ii. ii. Tumor Margin Assessment (*arctic_ai.cnn_embeddings*, *generate_graph*, *gnn_prediction*, set to *tumor_map* mode)
        iii. Ink detection and spatial statistics for tissue orientation (*arctic_ai.ink_detection*)
    d. Tumor Confounder Identification
        i. Follicle detection (*arctic_ai.follicle_detection*)
        ii. Cell classification (*arctic_ai.nuclei_detection*)
    e. Tumor and quality mapping onto surgical specimen (*arctic_ai.tumor_map*)
    f. Image stitching (*arctic_ai.image_stitch*)

2. A collection of docker and singularity containers that host various subcomponents of the software to enable interoperability and ease installation/dependency conflicts through self-contained linux subkernels.
3. Toil job scheduling tool and workflow engine for massive parallelization across local and cloud computing clusters (*arctic_ai.workflow*).
4. A dockerized dynamic web framework that can be hosted online and interacts with the aforementioned software elements and results output through *Plotly Dash*, which contain the following panes:

    a. Patient selection and workflow job submission (*Selection Pane*)
    b. 3D tissue model, size and orientation measurements, smart grossing recommendations and report generation (*3D Model Pane*)
    c. Histological findings–tumor and tissue quality (e.g., holes and tears), optional nuclei/follicle detection results, and detected inks placed atop slide images (*Histology Pane*)
    d. Mapping of tumor and/or hole/tear results back to original specimen via computer generated surgical maps (*Mapping Pane*)

In the following sections, we will elaborate on the functionality of each of the ArcticAI modules, with reference to supplementary methods if necessary.

Patient selection pane

A log-in pane allows for the selection of a patient/case. A database containing the patient and file paths to existing results data are searched and if results do not exist for the patient, the user is prompted to upload data for the *3D Model* and *Histology* panes, whichever may exist. Upon uploading, jobs are deployed to a high-performance computing cluster or within a GPU-capable device for parallel execution, which dynamically updates the database as results become available. If results exist, the *3D Model*, *Histology*, and *Mapping* panes become available for navigation. Here, the user is also instructed to supply the number of sections and tissue pieces per WSI based on their placement prior to image scanning.

3D tissue modeling and grossing recommendations in *3D Model Pane*

Three-dimensional tissue modeling prior to histological assessment provides smart grossing recommendations while automating the report of tissue size and orientation[83]. We utilized photogrammetry techniques which triangulate image features across multiple viewpoints/images to 3D coordinates in order to

generate a 3D model of the tissue. We developed a low-cost photogrammetry studio using a phone camera placed at a fixed distance away from a turntable. Immediately after resection, the tissue is placed on a turntable, from which a video of the tissue is recorded on a smart phone as it revolves around the table for one revolution. The video is then uploaded to the *ArcticAI* web app interface. Three-dimensional modeling is accomplished using the following two methodologies:

Point cloud based workflow.

1. Tissue Localization: First, the area of the turntable is approximated using RANSAC-based ellipse finding algorithm, which defines a static search area for tissue across the video frames. Then, image segmentation is performed on each video frame, which separates tissue from background using intensity thresholding, a connected component analysis for image labeling and an object size filter (Supplementary Fig. 14a) with background removal using the *grabcut* algorithm[84,85]. Under diverse imaging conditions, intensity thresholding can return many objects; however, only the gross specimen should follow an elliptical pattern as it completes a revolution. As such, RANSAC ellipse fitting and various fit statistics are again used to remove non-specimen objects through consideration of gross specimen's temporal trajectory. Alternatively, segmentation neural networks, which return pixelwise coordinates of the tissue location, can also accomplish this task given training data. Here, only one-tenth the number of segmented still frames are selected for inclusion in the reconstruction algorithms to reduce the compute time. This limits reconstruction quality, but the number of frames used for reconstruction can be varied based on speed/accuracy preferences.
2. Feature Matching: Is accomplished with image matching (e.g., SIFT, SURF, ORB, deep feature matching[86,87]), which can find correspondent features across different viewpoints. We utilized Colmap's SIFT implementation, which was accelerated using graphics processing units[88–90].
3. 3D Reconstruction: Three-dimensional scene reconstruction using colmap's structure from motion (SFM) framework after image pairing (i.e., match features between images from similar perspective), registration, and triangulation of pixel coordinates in a 3D cartesian coordinate system, which yields a sparse point cloud[91], after which a dense point cloud is generated using a Multi-View Stereo (MVS) framework via depth estimation.
4. Distance Calibration: Distance calibration (i.e., conversion of pixel distance to physical distance) by measuring the diameter of the turntable and fitting an ellipse (RANSAC) to the edges of the turntable, where edges were detected using a *scharr* filter (Supplementary Fig. 14b)[52,92,93].
5. Measuring orientation: Since the 3D model is oriented randomly upon creation, the 3D model is reoriented such that the flat surface at the tissue bottom is fixed in the downward ("negative-z" direction) position and the tissue is translated to the (0,0,0) cartesian coordinate system. First, a k-nearest neighbor's outlier detection subroutine is used to refine the point cloud. Calculation of the bottom tissue surface is accomplished through RANSAC plane fitting, where the normal plane vector is used to calculate a rotation matrix[94]. Finally, the tissue's 12 o'clock is calculated through segmentation of the tissue suture as a point of reference, which is used to rotate the tissue such that 12 o'clock aligns with the "positive-y" direction (Supplementary Fig. 14c). In the absence of the suture or potential slight misalignment, the web application features a slider to allow minor rotational adjustments.

6. Measuring Tissue Size: Measurements of tissue size (e.g., length, width, height) are captured by calculating the maximal x-y-z extents of the tissue respectively after tissue orientation (Supplementary Fig. 14d).
7. Further Model Refinement: The output 3D model retains the original color and texture of the excised tissue. The model is further refined using a Radius Neighbor's regression algorithm, which interpolates color and texture from adjacent points while estimating the z-coordinates from a closely spaced x-y grid. Alternatively, Poisson mesh reconstruction after estimation of triangle normals and/or Delauney triangulation and alpha hull construction present alternative refinement approaches (Supplementary Fig. 14e)[95].

Automated neural network 3D modeling method with neural radiance fields.

1. Automated Tissue Localization: This step was accomplished using segmentation neural networks, which was used to isolate tissue in individual images after training a general-purpose neural network using hand annotations. Only 20–30 frames per video were captured to simulate a faster one-second revolution around the turntable.
2. Estimation of Camera *Intrinsics*: Neural Radiance Fields (NeRF) is a machine learning method that helps us understand the 3D structure of tissues by learning to generate images of tissue taken from previously unseen angles and positions. These angles and positions used to train each tissue-specific NeRF model were estimated for each video through a GPU implementation of Colmap using the previous described methodology[90].
3. 3D Modeling with NeRF: One of the key innovations of NeRF is its use of neural graphics primitives, which greatly reduces its computational complexity, allowing for rendering of 3D scenes in real-time, even on low-cost devices like cell phones[53]. Our workflow uses a NeRF model trained with hash encodings, which speeds up the learning process, generating detailed 3D models from just a few 2D photos. These high-quality tissue images are devoid any gaps or holes, something that other techniques struggle to achieve. By providing specific camera angles and orientations relative to the tissue, the model can quickly measure the tissue's length, width, and height through assessment of the segmented tissue at supplied side, front and top-down views.

It should be noted that the *3D Modeling* step does not model or image deep margins since the bottom of the tissue sits on the turntable, though this modeling step is entirely separate from the histological findings (which do model deep margins) and mapping those results to the surgical tumor map but may be integrated with the other two modules in future iterations.

Grossing Recommendations and Size Report in 3D Model Pane. The 3D tissue model is displayed using an interactive web application using the *dash_vtk* package along with exportable technical readouts on the tissue size measurements (*3D Model Pane*)[96,97]. ArcticAI features two grossing recommendation tools, one for Mohs and another for traditional excisions with breadloafing. For the Mohs configuration, a 3D line is drawn from 12 o'clock to 6 o'clock in the web application. The 12 o'clock portion of the line is colored blue while the 6 o'clock portion is colored red. If the tissue is to be bisected, two pairs of blue-red lines are drawn parallel to a black line, which is drawn in the middle of the orientation lines. For breadloafing, the surgical excision is arranged such that the Burow's triangles/cones (i.e., superior/inferior or lateral/medial triangular excisions adjacent to resection used as a skin graft to repair surgical defects) point in

the forward/backward ("positive/negative-y") direction. Colored lines are placed across the specimen in the side-to-side direction at regular 0.5 to 1-centimeter increments (or set by the user; based on distance to the center) to represent placement of the breadloaf section cuts. Lines to the left of the specimen are colored blue to maintain orientation, while lines to the right are colored red, yellow, green, purple and orange to denote unique sections. The tissue can be inked in accordance with grossing recommendations.

Whole slide image preprocessing

After uploading tissue image sections in Whole Slide Image (WSI) format (TIFF/SVS file format, unsigned 8-bit color), slide images are prepared for both tumor and hole/tear prediction subroutines. First, tissue mask is created using a collection of image filters via the *PathFlowAI* package[98]. The tissue mask is generated using the following subroutine:

1. Using an intensity threshold filter, where objects of too high intensity are removed/set to white and filtering out large gray objects which may be artifactual (e.g., image scanner text, background black pen, etc.).
2. Morphology (binary closing) and blurring operations to smooth out the mask.
3. Removal of small objects and small holes.

Patch Extraction and Assignment of Tissue Piece and Section Identifier. WSI are typically partitioned into patches/subimages because they are too large to predict on using modern high performance computing resources with limited GPU memory. Therefore, subimages (256-pixel by 256-pixel) were extracted from the source image. Patches were extracted given that they had a significant overlap with the tissue mask as defined by a set threshold of tissue present. Patches were appended with patch metadata (e.g., x-y coordinate in WSI). The patch metadata also contains which serial section the patch belongs to (multiple serial sections per WSI). Oftentimes, each of the tissue sections were bisected or cut into four quadrants and inked separately. We refer to the resulting fragments as tissue pieces (one or more pieces per section). Each piece was placed separately in the WSI and physically close to other pieces in the same section, though there were instances where pieces either overlapped or highly separated which made it difficult to properly tag a patch with the relevant section. Tagging patches with the section they belong to is essential for tumor mapping, such that each section can be isolated from the others and then mapped by itself to the surgical tumor map after predicting the histological findings. As many WSI may be extracted per excision stage/depth, and multiple stages may be extracted during the excision procedure, the naming convention for each section denotes the depth in the specimen. Inadequate separation of sections and/or tissue pieces by failing to tag patches with the correct section identifier may degrade the performance of the tissue completeness, orientation and mapping algorithms because patches will be extracted from the space between the two conjoined section, which may contain excess whitespace and distort the ink and shape statistics. However, estimating which sections certain tissue pieces belong to is non-trivial since neighboring pieces may be conjoined, which resembles a section with fewer pieces that may fail to map well. To this end, the *preprocess* module features a robust automated section/piece splitting algorithm which can assign patches to the appropriate tissue piece/section using the following subroutine (Supplementary Fig. 15a):

1. Tissue patches are connected based on spatial proximity, building a radius nearest neighbors graph. Sections comprised of patches are established using a connected component analysis that finds and labels contiguous sets of

patches. Sections are assigned based on a large neighborhood of patches, large enough (within a 4096-pixel radius) to connect patches between neighboring tissue pieces within the section but small enough to not incorporate adjacent sections. At this point, it is difficult to delineate pieces from the section especially if they are conjoined, because what constitutes a contiguous element is defined at a large distance, and it is assumed that tissue pieces are the tissue sections, which now must be further subdivided. Then, pieces are initially broken based on connectivity at a smaller distance (within a 512-pixel radius), small enough to break apart pieces when they are separable in a section but not small enough to separate conjoined pieces. This will generate in most cases multiple pieces per section.

2. For each tissue section, if the number of tissue pieces for the section matches expectations (input parameter), then piece/section assignment for that section is complete. If the number of pieces per section does not match expectations, then the algorithm would divide largest candidate piece into the expected section pieces using Spectral Clustering, a technique that divides conjoined sections into separate areas by regions of weak connectivity between conjoined pieces. Repeat this step until optimal results are achieved.

The initial set of tissue subimages that remain after the above procedure serve as input to the tumor prediction algorithm i.e., *tumor_map* configuration, which predicts presence of tumor on a patch-by-patch basis. This configuration defines areas on a section where tissue is present from which to predict location of tumor but purposefully omits candidate holes and tears to avoid predicting those regions as benign. Thus, this set of patches is insufficient to predict where tissue is absent or incomplete (i.e., tissue quality, location of holes and tears that determine whether the section should be assessed) since patches are, by definition, absent.

Patches correspondent to candidate holes and tears are extracted for each section piece using an *alpha shape* object finding algorithm which outlines the section piece in a way that is both tightly fit to the piece while also bridging tears that are connected to the exterior of the object and thus are not normally estimated using traditional hole finding algorithms (Supplementary Fig. 15b)[95]. These patches are added to the *tumor_map* patches to form the *macro_map* configuration for tissue quality / tissue completeness assessment. If tissue is incomplete, it is inadvisable to assess margins. All tissue piece subimage patches, tissue masks and their corresponding metadata are written to NPY format (numpy array[99]) and serialized into pickle format respectively for storage.

Feature extraction using convolutional neural networks

We trained *ResNet-50* convolutional neural network (CNN) models to extract predictors from the tissue subimages to be used in our prediction workflow[100]. *ResNet-50* was selected after comparing the performance on the validation set between a myriad of CNN models available in the *PathFlowAI* package and after conducting a random/coarse hyperparameter search. First, convolutional neural networks were trained and internally validated on a subset of image patches ($n = 122$ WSI; 1,988,841 patches) for the following prediction tasks, with a batch size of 32 patches, learning rate of 1e–4, modulated with a cosine annealing learning rate scheduler for 100 training epochs:

1. *Tumor CNN:* Tumor localization, where regions of tumor were delineated from benign structures and inflammation. If a patch contained both malignant and inflammatory cells, the patch was labeled as having contained tumor.
2. *Completeness CNN:* Delineation of macro-architectural subcompartments, including: (1) hole/tear, (2) fat, which if not

explicitly annotated, could closely resemble hole/tears, (3) epidermis and (4) dermis. Here, patches of dermis containing wispy white patterns were removed from the training/validation set to avoid conflating regions of dermis with hole/tear.

After training, the two ResNet50 CNN models were used to extract *embeddings* of image features from the penultimate layer of the model as images passed through the neural networks. CNNs were organized into multiple processing layers, each of which represent objects/images at increasing levels of abstraction (i.e., each input register corresponds to a more complex image feature at a deeper layer). Whereas the final CNN layer is used to output the probability of the presence of a specific tissue architecture, the penultimate layers output a rich feature set (*embeddings*) which can be used as a generic representation of the image features and if plotted could demonstrate how specific images cluster together with dimensionality more expressive than the output layer alone. The trained CNN models, *Tumor CNN* and *Completeness CNN*, are configured to output 2048-dimensional *embeddings* for all *tumor_map* and *macro_map* patches respectively for a given tissue section. CNN models were configured using the *PyTorch* package (v1.8.0) using Python v3.7[101].

Graph neural networks for final histological assessment

WSI contain significant white space and the placement of tissue on a slide is relatively arbitrary. The dimensions of WSI are typically very large which necessitates dividing the tissue into smaller subimages (tiles). However, prediction using a CNN assumes that neighboring image patches are unassociated, which undervalues their spatial context within the surrounding tissue architecture. Graphs represent image patches and their spatial dependencies as nodes and edges, capturing both feature and spatial information. As such, graph-based neural networks (GNN) have emerged as premier methods for histological assessment. Using GNNs, predictions are invariant to the positioning and orientation of the tissue and are enhanced by the incorporation of spatial information encoded in the edges. The motivation for adopting this methodology was based on previous reports of improved performance by utilizing contextual information from surrounding tissue[54]. The GNN architecture (e.g., number of layers, type of layers– graph attention, etc.) was determined after a randomized hyperparameter search, comparing performance on the internal validation set. We fit two GNN models corresponding to the following prediction tasks, with a batch size of 16 WSI graphs, learning rate of 1e-2, modulated with a cosine annealing learning rate scheduler for 1500 training epochs:

1. Tumor GNN: Analogous to Tumor CNN.
2. Completeness GNN: Analogous to Completeness CNN. However, all patches are included in this analysis, including wispy dermis which is now contextualized by the surrounding dermis and not subject to conflation with holes/tears.

Graphs were defined using a radius neighbors algorithm, which connected patches (nodes) to their immediate neighbors (edges) using their positional x-y coordinates[54]. Attributes of the graph nodes were set to the *embeddings* extracted by the relevant CNN. Node attributes (CNN features) were shared/passed to adjacent patches using three graph attention layers of dimensionality 32, 32, and 64. The graph convolution layers were interspersed with *DropEdge* and *Dropout* layers, which randomly pruned patch-wise connections and graph-learned node features during training (to enhance model robustness to noise). After running the graph convolutional layers, these features were piped to a prediction output layer that would return predicted probabilities (and their logits) of each class for the respective tasks. GNN models were configured using the *PyTorch-Geometric* package (v1.7.1) using Python v3.7[55].

Ink detection and calculation of spatial statistics for tissue orientation

The orientation of the WSI tissue section with respect to the original specimen / surgical tumor map was inferred using spatial statistics / tissue orientation algorithms. These algorithms were developed to automatically identify ink colors and orient a WSI tissue section based on a collection of applied inks: blue and red, though subroutines exist in *ArcticAI* to additionally calculate: yellow, green, orange, purple, and black. First, tissue edges were segmented using a Sobel filter with morphological dilation and opening operations. Then, sensitivity analysis over thresholds in Hue, Saturation, Value (HSV) color space yielded optimal color thresholds to detect inks, which were paired with a connected component analysis to identify contiguous regions and remove spurious applications of ink within the tissue edges (i.e., where ink is erroneously applied / seeps). Alternatively, semantic segmentation algorithms based on annotations, and conditional random field[102] can further improve ink detection. After detecting ink, orientation of the WSI section piece is inferred through calculation of the center of mass of the x-y coordinates of each detected ink color (using either the mean, median, or trimmed mean of each pixel coordinate). In our practice setting blue defines 12 o'clock, red defines 6 o'clock, in accordance with the 3D model of the Mohs specimen and with the surgical tumor map. The line between the blue and red ink defines tissue orientation relative to blue and red locations defined in the surgeon's hand-drawn tumor map, where the relative angular difference between the blue-red lines from the histology and blue-red lines from the tumor map dictate the relative rotation required for the histology results to match the same angle of the tumor map.

Image stitching

Input WSI are prepared for viewing using a subroutine which converts the images of individual sections, extracted using the preprocessing workflow, to a "Deep Zoom Image" (DZI) format, a pyramidal file format which interfaces with *openseadragon*, a WSI viewer. The aforementioned rapid histological assessment steps (Preprocessing, *CNN-GNN*, *Ink Detection*) return positional predictions for their respective coordinates. These positional predictions are piped and prepared for display through a dynamic json export of the histological results and imported into an *openseadragon* SVG overlay component for viewing across the slide[103].

Removal of potential tumor confounders and identifying residual tumor in regions of predominant inflammation

Two R101-FPN neural network models using the computer vision framework *detectron2* (for panoptic segmentation) were trained for the task of localizing follicles across a slide and residual tumor within pockets of inflammation as an added layer of auditing. Panoptic segmentation models can detect objects in an image and their image class through proposals of bounding boxes using neural network detected features while simultaneously segmenting the object using a segmentation architecture which operates dynamically on the proposed regions[56,57].

First, 672 follicles were annotated across 16 WSI (60 tissue sections) from the training/validation sets, where 595 non-overlapping 1024 pixel by 1024-pixel subimages were extracted and assigned to training and validation sets whether they belonged to WSI of the training or validation set for the CNN-GNN algorithms (i.e., different set of patients within these cohorts). A panoptic segmentation network was fit to the data at a starting learning rate of 1e−3 and was trained for 1000 epochs. The trained neural network was applied to patches suspected to contain tumor in the test set WSI (i.e., different set of patients from the training and validation sets) to eliminate patches with significant confounding. This was done using an adjustment scheme in which

tumor scores were reduced more for patches with greater proximity to the follicle based on the overlap between the predicted follicle and three concentric circles (128, 256, 512-pixel radii respectively) around each patch. The percentage overlap with the follicles and each circle multiplied by a circle specific penalty (higher for the 128-pixel circle and lower for the 512-pixel radius circle), normalized to a scale between zero and one was used to determine the tumor prediction probability to dock from the original score.

Inflammatory patches were assessed at the level of individual nuclei via the creation of a workflow to detect, classify, and segment nuclei. Using the nuclei annotations that were manually annotated by four pathologists using the Automated Slide Analysis Platform (ASAP), we extracted 795 patches of size 128 by 128, correspondent to approximately 32,763 nuclei, across three whole slide images (WSI) from the training/validation set. The model was trained to detect and delineate the following cell types: (1) fibroblasts, (2) hair follicles, (3) inflammation, (4) malignant basal cells, and (5) benign epidermal keratinocytes. For our train-validation split, 80% of these patches were randomly chosen for our training/validation dataset and the algorithm was evaluated using the remaining 20% of the image patches prior to prediction on the held-out test WSI. The held-out test set was comprised of an entirely different/unique set of patients than the training and validation sets. We reported the predictive statistics for detection accuracy on an internal validation set using the Dice score (related to Intersection over Union). Predicted cell types were refined from the detected nuclei using a CNN (ResNet-50 architecture) and GNN model using the same training/validation sets. F1-score statistics were recorded as a final measure of fit across the test set cells, separately comparing the detection, CNN and GNN models for their prediction accuracy, bootstrapping on the patch and slide level since the nuclei are nested within patches.

Optimal hyperparameters were identified from evaluated results on this held out validation set prior to application across test slides. We timed both algorithms through evaluation across test set WSI and similarly recorded the uncertainty through non-parametric bootstrapping while accounting for clustering on the WSI level.

For these assessments, algorithms were trained, validated and evaluated on training, validation and test sets respectively which were comprised of a different set of patients to avoid target leakage from having serial sections placed in different cohorts. Data within each patient was not included in both the training and test sets. These cohorts were partitioned using the *Group-ShuffleSplit* iterator available in the scikit-learn package.

Compilation of histological assessment results into *Histology Pane*

The results from the histological assessment models (*Tumor CNN-GNN, Completeness CNN-GNN*) are passed to an OpenSeadragon plugin that features the DZI image correspondent to the selected case, resection site/stage and section depth. This plugin operates within a *plotly dash* environment that interfaces the WSI viewer with the results data[104]. The user has the option of selecting whether to display a heatmap containing the tumor or completeness (holes/tears) prediction results over the slide using the SVG plugin as aforementioned. A slider controls the minimal prediction probability for inclusion in the heatmap to filter out irrelevant regions. The patches' color intensity (blue to red and opacity) is determined by their prediction scores. The predictions from image patches can be optionally refined using an interpolation method which uses a custom prediction propagation GNN to yield refined predicted probabilities that exist between the original patches (e.g., if we had four patches in locations (1024,256), (1024,512), (1024,768), (1024, 1024), we could infer information at tile position

(1024, 640) by leveraging information from all four tiles, though primarily from adjacent tiles)[105]. The SVG display can be toggled on and off. The tissue orientation may also be toggled on and off, where red and blue lines may be automatically placed on the slide based on detected inking patterns and associated spatial statistics.

Mapping tumor and completeness results to surgical specimen using the *Mapping Pane*

Results (tumor/completeness) from user selected tissue sections for each case can be mapped to the surgical specimen that is featured on a hand-drawn surgical tumor map at arbitrary locations using an interactive image display. First, the user selects from a set of prepopulated surgical map templates representing various anatomical positions (e.g., back of hand, neck). After selecting the position template, the user draws a black ellipse on the image template representing the removal site. The user also defines tissue orientation by drawing blue and red inks at the circle's edge to define 12 o'clock and 6 o'clock respectively, which is correspondent to inking patterns recommended/selected from the *3D Model* and *Histology* panes. Tissue sections comprised of 1-2 tissue pieces are represented by a 2D point cloud or a collection of points, where each point is tagged with positional x-y coordinates within the WSI, the tumor/hole/tear predicted probabilities and ink locations. These points are "morphed" or registered to locations in the interior of the circle using an optimizer for optimal transport, which minimizes the cost or effort required to match the points to the interior of the circle while maintaining the relative positioning of the coordinates comprising the histological section. The distributional difference between the histological section and surgical mapping ellipse is estimated using the sliced Wasserstein ("Earth Movers") distance and minimized using gradient descent via *pytorch* and *python optimal transport (POT)* libraries[60–62,101]. In sum, this methodology morphs the arbitrary shapes of the histological specimen, which are dependent on serial sectioning of the gross specimen, to the elliptical shape drawn by the Mohs surgeon. The relative positioning of ink is preserved during this transformation and the angular difference between the ink after tissue morphing and that defined via the *Mapping pane* are used as a final rotational adjustment to match the surgical tumor map. Finally, the histological section results are placed in the circle on the *Mapping pane* in the correct orientation, where a kernel density contour map defined over the tumor/completeness results is placed to highlight tumor/holes/tears. Like the *Histology pane*, the density map can be thresholded with arbitrary probability cutoffs by the user to yield specific tumor locations and the finalized map can be exported to the pathology report.

Differentiating 3D modeling approaches

In the supplementary information file, we have included several supplementary figures (Supplementary Figs. 1, 9–11; Supplementary Videos 1–4) illustrating methods to study tissue in 3D, which have been summarized below:

1. 3D reconstruction of the gross specimen: Use of photogrammetry techniques to image tissue at numerous angles prior to histological examination and using triangulation of imaging features or use of neural radiance fields to measure tissue dimensions and suggest inking/sectioning recommendations (Supplementary Videos 1, 2)[53].

2. 3D histopathology: Co-registering WSI corresponding to serial sections through image registration methods and then identifying histological structures persistent through these serial layers– does not consider position and orientation of tissue at surgical site and due to high dimensionality of images convergence time can be slow across many serial sections[40,41]. This technique was not employed in this

computational workflow due to limitations precluding real-time analysis though 3D meshes of the tumor margin were reconstructed from alignment of serial sections to demonstrate application of these methods (Supplementary Video 4).

3. Tumor mapping: Mapping each serial section to a common coordinate system through identification of inks as means to perform rapid co-registration and orientation to the original surgical site (Supplementary Video 3, Supplementary Data 1).

Workflow specification

All aforementioned *ArcticAI* jobs execute using a Toil job executor[63], which can run jobs in parallel either locally (on a GPU-capable machine) or in an HPC environment using a Slurm or alternative job submission system. Here, we will enumerate which components execute in parallel based on their respective workflows, where we have denoted which set of subcomponents execute in series or parallel.

1. 3D model pane (series)

a. Tissue preprocessing (series)
b. 3D Reconstruction (series)
c. Final tissue filtering (series)

2. Histology pane (series)

a. Tissue preprocessing and section assignment (series)
b. Parallel components, where final subworkflow time is assessed by the subcomponent and tissue section which took the longest time to execute (below subpoints are **parallel**)

    i. CNN-GNN subcomponents for tumor/completeness prediction, comprised of CNN embedding creation, graph generation and GNN prediction (parallel)
    ii. Ink detection and orientation (parallel)
    iii. Image stitching (parallel)

c. Based on CNN-GNN results (parallel)

    i. Follicle detection (parallel)
    ii. Nuclei detection (parallel)

Finally, we have included a *Histology* workflow diagram which illustrates how results from various workflow components feed into subsequent steps (Supplementary Fig. 13). It should be noted that after tissue preprocessing, tissue pieces/sections all execute in parallel, from which the aforementioned subworkflows also execute in parallel.

Experimental Objectives

Tissue Grossing Measurement Concordance: Length, width and height measurements of the 3D reconstruction of resected tissue was compared to hand measurements of the original specimen using median absolute deviation and spearman correlation statistics. These statistics were also recalculated under the assumption that the calculated tissue dimensions were off on all 3 dimensions by a proportional constant (i.e., improper calibration of the video with distance measurements).

Concordance of Algorithm to Hand-Drawn Maps from Surgeon: To assess the accuracy of the Arctic grossing, completeness, and tumor detection algorithms and mapping of histological findings to the surgical tumor map, we included the following comparisons, assessing accurate: (1) calculation of tissue size, (2) analysis of tissue quality or completeness of tissue as judged by the localization of holes and tissue tears, common to frozen specimens, (3) localization of tumor in WSI, (4) orientation of tissue section, (5) mapping of tumor to surgical tumor map, and (6) a prediction of whether and where additional tumor removal is required.

In comparison to previous studies which examine the diagnostic significance of positive margins on post-operative BCC sections, we assessed whether pathologists would manually map tumors similarly in digital versus analog mediums by comparing hand-drawn tumor maps to digital ones. After establishing concordance between histological findings via pathologist annotations and BCC predictions, we established the concordance between the automated tumor map with the intraoperative hand generated map that was produced by the Mohs Micrographic Surgeon at the time of surgery. If there were discrepancies between surgeon-generated and ArcticAI-generated tumor maps, both the original glass slides and WSI were manually reviewed.

A receiver operating characteristic curve (ROC; sensitivity analysis) was performed to establish predictive probability cutoffs which result in high sensitivity to minimize the potential for false negatives. Results are reported with a 1000-sample non-parametric bootstrap 95% confidence interval with bootstrapping performed on the WSI level to capture clustering on the slide level (i.e., variation in performance statistics between and across slides).

Separately, concordance between hand-drawn and digital tumor maps was calculated based on the proportion of cases the surgeon subjectively rated as equivalent (orientation of map and position of tumor) to the original map. Uncertainty in this proportion was assessed through calculation of the 95% credible interval (CI; like the confidence interval) of a Beta posterior distribution ($Beta(a = 0.5 + 28, b = 0.5 + 0)$), updated through a beta-binomial conjugate prior, with a Jeffrey's prior ($Beta(0.5, 0.5)$) and Binomial likelihood ($Binomial(n = 28, p = 1.0)$; 28 cases with positive margin, 28 successful trials; three cases had clear margins).

Execution Time: To demonstrate the timely execution of the ArcticAI system, the following steps in the process were precisely timed: (1) image preprocessing; (2) tissue quality and tumor CNN-GNN prediction; and (3) tumor mapping and pathology report output. We report median times across slides to account for outliers, with 1000-sample nonparametric bootstrap 95% confidence intervals. Details on the calculation of timing given optimal parallelization can be found in the methods section, section "Workflow Specification". Test cases were evaluated using four compute nodes in the Dartmouth Discovery computing cluster which shared between them 13 Nvidia v100 GPUs (32 Gb memory each), 272 CPUs, and 1.9 TB RAM.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

The datasets presented in this article are not readily available because of participant privacy concerns. A subset data produced in the present study may be made available upon reasonable request to the authors. A subset of data is also readily accessible through our web application: https://arcticai.demo.levylab.host.dartmouth.edu/. Requests to access the datasets should be directed to joshua.j.levy@dartmouth.edu or joshua.levy@cshs.org.

CODE AVAILABILITY

ArcticAI is a publicly available package that can be installed using pip (pypi: *arctic-ai*) and is hosted on GitHub at the following URL: https://github.com/jlevy44/ArcticAI/. To ensure reproducible computing, ArcticAI is also available as a Docker/Singularity container, with installation instructions on the GitHub website. For readers who are interested in developing their own intraoperative margin assessment workflows, documentation for operating the application programming interface (API) and command line interface (CLI) can be found on our GitHub wiki (https://github.com/jlevy44/ArcticAI/wiki) and on our accompanying ReadtheDocs website (https://jlevy44.github.io/ArcticAI/). Please note that as the neural network models were developed at a single site– separate models will need to be trained at each institution and models trained for this manuscript can be made available upon reasonable request. A tutorial for training graph neural networks can be found at the following GitHub URL: https://github.com/jlevy44/WSI-GTFE. To learn more about the ArcticAI approach and how to operate it, we have provided a publicly available web application at the following URL: https://arcticai.demo.levylab.host.dartmouth.edu/. This application demonstrates 3D specimen grossing recommendations, histological examination, and mapping of histological results back to the surgical site for select cases. We have also included a video demonstrating operation of this web application at the following URL (Supplementary Figs. 16–18): https://www.youtube.com/watch?v=I1Gfe6xl3Yg. *ArcticAI* relies on the following Python packages: alphashape (v1.3.1), click (v8.1.3), dask (v2023.2.1), fire (v0.5.0), kornia (v0.6.10), pathflowai/pathpretrain (latest version), numpy (v1.24.2), pandas (v1.5.3), tifffile (v2021.11.2), opencv (v4.5.5.62), scikit-learn (v1.2.1), scikit-image (v0.18.3), pytorch (v1.13.1), torchvision (v0.14.1), torch-geometric (all related packages with versions compatible with pytorch), detectron2 (latest version), seaborn (v0.12.0), shapely (v2.0.1), matplotlib (v3.7.0), plotly (v5.17.0), dash (v2.14.0), toil (v5.12.0), instant-ngp (latest version), and POT (python optimal transport). Jupyter notebooks were used for custom prototyping, training and evaluation of the model results while *ArcticAI* was used for large scale inference across the test set slides. Default training/inference input arguments can be found in the GitHub repository. Credible intervals for concordance assessments between hand-drawn and digital tumor maps developed using R v4.1.

REFERENCES

1. Thomas Robbins, K. et al. Surgical margins in head and neck cancer: Intra- and postoperative considerations. *Auris Nasus Larynx* **46**, 10–17 (2019).
2. Chen, T. Y., Emrich, L. J. & Driscoll, D. L. The clinical significance of pathological findings in surgically resected margins of the primary tumor in head and neck carcinoma. *Int. J. Radiat. Oncol.*Biol.*Phys.* **13**, 833–837 (1987).
3. Bialy, T. L. et al. Mohs micrographic surgery vs traditional surgical excision: A cost comparison analysis. *Arch. Dermatol.* **140**, 736–742 (2004).
4. van Loo, E. et al. Surgical excision versus Mohs' micrographic surgery for basal cell carcinoma of the face: A randomised clinical trial with 10year follow-up. *Eur. J. Cancer* **50**, 3011–3020 (2014).
5. Walker, P. & Hill, D. Surgical treatment of basal cell carcinomas using standard postoperative histological assessment. *Australas. J. Dermatol.* **47**, 1–12 (2006).
6. van Delft, L. C. J., Nelemans, P. J., van Loo, E., Abdul Hamid, M. & Kelleners-Smeets, N. W. J. The illusion of conventional histological resection margin control. *Br. J. Dermatol.* **180**, 1240–1241 (2019).
7. Berlin, J., Katz, K. H., Helm, K. F. & Maloney, M. E. The significance of tumor persistence after incomplete excision of basal cell carcinoma. *J. Am. Acad. Dermatol.* **46**, 549–553 (2002).
8. Lane, J. E. & Kent, D. E. Surgical margins in the treatment of nonmelanoma skin cancer and Mohs micrographic. *Surg. Curr. Surg.* **62**, 518–526 (2005).
9. Nolan, G. S. et al. Incomplete surgical excision of keratinocyte skin cancers: a systematic review and meta-analysis*. *Br. J. Dermatol.* **184**, 1033–1044 (2021).
10. Wolf, D. J. & Zitelli, J. A. Surgical margins for basal cell carcinoma. *Arch. Dermatol.* **123**, 340–344 (1987).
11. Cook, J. & Zitelli, J. A. Mohs micrographic surgery: A cost analysis. *J. Am. Acad. Dermatol.* **39**, 698–703 (1998).
12. Essers, B. A. B. et al. Cost-effectiveness of Mohs micrographic surgery vs surgical excision for basal cell carcinoma of the face. *Arch. Dermatol.* **142**, 187–194 (2006).
13. Smeets, N. W. et al. Surgical excision vs Mohs' micrographic surgery for basal-cell carcinoma of the face: randomised controlled trial. *Lancet* **364**, 1766–1772 (2004).
14. Wong, E., Axibal, E. & Brown, M. Mohs micrographic surgery. *Facial Plast. Surg. Clin.* **27**, 15–34 (2019).
15. Shriner, D. L., McCoy, D. K., Goldberg, D. J. & Wagner, R. F. Mohs micrographic surgery. *J. Am. Acad. Dermatol.* **39**, 79–97 (1998).
16. LeBoeuf, M., Bieber, K., Cooper, S. & Isenor, A. Quantitative analysis of frozen section histology in Mohs micrographic surgery. *Dermatol. Surg.* **46**, 876–879 (2020).
17. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
18. Acs, B., Rantalainen, M. & Hartman, J. Artificial intelligence as the next step towards precision pathology. *J. Intern. Med.* **288**, 62–81 (2020).
19. Levy, J. J. & Vaickus, L. J. Artificial intelligence in anatomic pathology. *Adv. Mol. Pathol.* **4**, 145–171 (2021).
20. Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V. & Madabhushi, A. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nat. Rev. Clin. Oncol.* **16**, 703–715 (2019).
21. Niazi, M. K. K., Parwani, A. V. & Gurcan, M. N. Digital pathology and artificial intelligence. *Lancet Oncol.* **20**, e253–e261 (2019).
22. Sarwar, S. et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. *npj Dig. Med.* **2**, 1–7 (2019).
23. Cui, M. & Zhang, D. Y. Artificial intelligence and computational pathology. *Lab. Investig.* **101**, 412–422 (2021).
24. Mahmood, F. et al. Deep Adversarial Training for Multi-Organ Nuclei Segmentation in Histopathology Images. *IEEE Trans. Med. Imaging* https://doi.org/10.1109/TMI.2019.2927182 (2020).
25. Chen, R. J. et al. Pathomic fusion: An integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Trans. Med. Imaging* 1–1 https://doi.org/10.1109/TMI.2020.3021387 (2020).
26. Lu, M. Y. et al. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat. Biomed. Eng.* 1–16 https://doi.org/10.1038/s41551-020-00682-w (2021).
27. Zhang, Y. et al. Digital synthesis of histological stains using micro-structured and multiplexed virtual staining of label-free tissue. *Light. Sci. Appl.* **9**, 78 (2020).
28. Rivenson, Y., de Haan, K., Wallace, W. D. & Ozcan, A. Emerging advances to transform histopathology using virtual staining. *BME Front.* **2020**, 9647163 (2020).
29. Rivenson, Y. et al. PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning. *Light Sci. Appl* **8**, 1–11 (2019).
30. de Haan, K. et al. Deep learning-based transformation of H&E stained tissues into special stains. *Nat. Commun.* **12**, 4884 (2021).
31. Bayramoglu, N., Kaakinen, M., Eklund, L. & Heikkilä, J. Towards Virtual H E Staining of Hyperspectral Lung Histology Images Using Conditional Generative Adversarial Networks. in *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)* 64–71. https://doi.org/10.1109/ICCVW.2017.15 (2017).
32. Rivenson, Y. et al. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nat. Biomed. Eng.* **3**, 466–477 (2019).
33. Jackson, C. R., Sriharan, A. & Vaickus, L. J. A machine learning algorithm for simulating immunohistochemistry: development of SOX10 virtual IHC and evaluation on primarily melanocytic neoplasms. *Mod. Pathol.* 1–11 https://doi.org/10.1038/s41379-020-0526-z (2020).
34. Levy, J. J. et al. A large-scale internal validation study of unsupervised virtual trichrome staining technologies on nonalcoholic steatohepatitis liver biopsies. *Mod. Pathol.* **34**, 808–822 (2021).
35. Lu, M. Y. et al. AI-based pathology predicts origins for cancers of unknown primary. *Nature* **594**, 106–110 (2021).
36. Sohn, G. K., Sohn, J. H., Yeh, J., Chen, Y. & Jiang, S. I. B. A deep learning algorithm to detect the presence of basal cell carcinoma on Mohs micrographic surgery frozen sections. *J. Am. Acad. Dermatol.* **84**, 1437–1438 (2021).
37. Campanella, G. et al. A deep learning algorithm with high sensitivity for the detection of basal cell carcinoma in Mohs micrographic surgery frozen sections. *J. Am. Acad. Dermatol.* **85**, 1285–1286 (2021).
38. Thomas, S. M., Lefevre, J. G., Baxter, G. & Hamilton, N. A. Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer. *Med. Image Anal.* **68**, 101915 (2021).
39. Hollon, T. C. et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nat. Med.* **26**, 52–58 (2020).
40. Kiemen, A. L. et al. CODA: quantitative 3D reconstruction of large tissues at cellular resolution. *Nat. Methods* **19**, 1490–1499 (2022).
41. Gatenbee, C. D. et al. Virtual alignment of pathology image series for multi-gigapixel whole slide images. *Nat. Commun.* **14**, 4502 (2023).
42. Roberts, N. et al. Toward routine use of 3D histopathology as a research tool. *Am. J. Pathol.* **180**, 1835–1842 (2012).
43. Sood, R. R. et al. 3D Registration of pre-surgical prostate MRI and histopathology images via super-resolution volume reconstruction. *Med. image Anal.* **69**, 101957 (2021).

44. Scheibe, P. et al. Image-processing chain for a three-dimensional reconstruction of basal cell carcinomas. *Exp. Dermatol* **19**, 689–691 (2010).

45. Choi, J. E., Kwon, I. H., Seo, S. H., Kye, Y. C. & Ahn, H. H. Pathogenesis of plantar epidermal cyst: Three-dimensional reconstruction analysis. *Ann. Dermatol.* **28**, 133–135 (2016).

46. Perez, A. N. et al. Ex vivo 3D scanning and specimen mapping in anatomic pathology. *J. Pathol. Inf.* **14**, 100186 (2023).

47. Bois, M. C. et al. Three-dimensional surface imaging and printing in anatomic pathology. *J. Pathol. Inf.* **12**, 22 (2021).

48. Mahmoud, A. & Bennett, M. Introducing 3-dimensional printing of a human anatomic pathology specimen: potential benefits for undergraduate and post-graduate education and anatomic pathology practice. *Arch. Pathol. Lab. Med.* **139**, 1048–1051 (2015).

49. Saturno, M. P. et al. Utilizing 3D head and neck specimen scanning for intraoperative margin discussions: Proof of concept of our novel approach. *Head. Neck* **45**, 10–21 (2023).

50. Kulak, O. et al. Smart glasses as a surgical pathology grossing tool. *Arch. Pathol. Lab. Med.* **145**, 457–460 (2020).

51. Xie, W. et al. Prostate cancer risk stratification via nondestructive 3D pathology with deep learning–assisted gland analysis. *Cancer Res.* **82**, 334–345 (2022).

52. Schnabel, R., Wahl, R. & Klein, R. Efficient RANSAC for point-cloud shape detection. *Comput. Graph. forum* **26**, 214–226 (2007).

53. Müller, T., Evans, A., Schied, C. & Keller, A. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph. (ToG)* **41**, 1–15 (2022).

54. Levy, J., Haudenschild, C., Barwick, C., Christensen, B. & Vaickus, L. Topological feature extraction and visualization of whole slide images using graph neural networks. *Pac. Symp. Biocomput.* 285–296 https://doi.org/10.1101/2020.08.01.231639 (2021).

55. Fey, M. & Lenssen, J. E. Fast graph representation learning with PyTorch geometric. *arXiv:1903.02428 [cs, stat]* (2019).

56. Cheng, B. et al. Panoptic-DeepLab: A simple, strong, and fast baseline for bottom-up panoptic segmentation. *arXiv:1911.10194 [cs]* (2020).

57. Kirillov, A., He, K., Girshick, R., Rother, C. & Dollar, P. Panoptic segmentation. in 9404–9413 (2019).

58. Maleki, D. et al. (Springer International Publishing, 2020) 359–368. https://doi.org/10.1007/978-3-030-64556-4_28.

59. Barlow, J. O. Using en face ink to guide the conservative sectioning of a Mohs margin. *Dermatol Surg.* **37**, 1730–1739 (2011).

60. Flamary, R. et al. Pot: Python optimal transport. *J. Mach. Learn. Res.* **22**, 1–8 (2021).

61. Bonneel, N., Rabin, J., Peyré, G. & Pfister, H. Sliced and radon wasserstein barycenters of measures. *J. Math. Imaging Vis.* **51**, 22–45 (2015).

62. Liutkus, A., Simsekli, U., Majewski, S., Durmus, A. & Stöter, F-R. Sliced-Wasserstein flows: Nonparametric generative modeling via optimal transport and diffusions. in *International Conference on Machine Learning* 4104–4113 (PMLR, 2019).

63. Vivian, J. et al. Toil enables reproducible, open source, big biomedical data analyses. *Nat. Biotechnol.* **35**, 314–316 (2017).

64. Garcia, E., Kundu, I., Kelly, M. & Soles, R. The American Society for Clinical Pathology 2020 Vacancy Survey of Medical Laboratories in the United States. *Am. J. Clin. Pathol.* aqab197 https://doi.org/10.1093/ajcp/aqab197 (2021).

65. Bonnefille, H. et al. A deep-learning algorithm to localize basal cell carcinoma foci on Mohs surgery frozen sections. *J. Eur. Acad. Dermatol. Venereol.* **37**, e631–e632 (2023).

66. van Zon, M. C. M., van der Waa, J. D., Veta, M. & Krekels, G. A. M. Whole-slide margin control through deep learning in Mohs micrographic surgery for basal cell carcinoma. *Exp. Dermatol.* **30**, 733–738 (2021).

67. Cataldo, P. A., Stoddard, P. B. & Reed, W. P. Use of frozen section analysis in the treatment of basal cell carcinoma. *Am. J. Surg.* **159**, 561–563 (1990).

68. Dellon, A. L., DeSilva, S., Connolly, M. & Ross, A. Prediction of recurrence in incompletely excised basal cell carcinoma. *Plast. Reconstr. Surg.* **75**, 860–871 (1985).

69. Lam, C. et al. Burnout and resiliency in Mohs surgeons: A survey study. *Int. J. Women's Dermatol.* **7**, 319–322 (2021).

70. Kohli, N. & Golda, N. Practice patterns and job satisfaction of Mohs surgeons. *Dermatol. Surg.* **44**, 42–47 (2018).

71. Lam, C. et al. Impact of medical scribes on physician and patient satisfaction in dermatology. *Arch. Dermatol. Res.* **314**, 71–76 (2022).

72. Bricca, G. M., Brodland, D. G., Ren, D. & Zitelli, J. A. Cutaneous head and neck melanoma treated with Mohs micrographic surgery. *J. Am. Acad. Dermatol.* **52**, 92–100 (2005).

73. Ciążyńska, M. et al. The incidence and clinical analysis of non-melanoma skin cancer. *Sci. Rep.* **11**, 4337 (2021).

74. Hanson, J., Demer, A., Liszewski, W., Foman, N. & Maher, I. Improved overall survival of melanoma of the head and neck treated with Mohs micrographic surgery versus wide local excision. *J. Am. Acad. Dermatol.* **82**, 149–155 (2020).

75. Egeler, S. A. et al. Regional incidence of and reconstructive management patterns in melanoma and nonmelanoma skin cancer of the head and neck: A 3-year analysis in the inpatient setting. *J. Plast., Reconstructive Aesthetic Surg.* **73**, 507–515 (2020).

76. Ibrahim, A. M. et al. Common patterns of reconstruction for Mohs defects in the head and neck. *J. Craniofacial Surg.* **25**, 87–92 (2014).

77. Patel, S. A. et al. Complication rates in delayed reconstruction of the head and neck after Mohs micrographic surgery. *JAMA Facial Plast. Surg.* **18**, 340–346 (2016).

78. Swanson, N. A., Grekin, R. C. & Baker, S. R. Mohs surgery: techniques, indications, and applications in head and neck surgery. *Head. Neck Surg.* **6**, 683–692 (1983).

79. Van Lee, C. B. et al. Recurrence rates of cutaneous squamous cell carcinoma of the head and neck after Mohs micrographic surgery vs. standard excision: a retrospective cohort study. *Br. J. Dermatol.* **181**, 338–343 (2019).

80. Ridha, H. et al. Intraoperative use of Mohs' surgery for the resection of major cutaneous head and neck cancer under general anaesthetic: Initial experiences, efficiency and outcomes. *J. Plast., Reconstructive Aesthetic Surg.* **68**, 1706–1712 (2015).

81. TERI McGILLIS, S., WHEELAND, R. G. & Sebben, J. E. Current issues in the performance of Mohs micrographic surgery. *J. dermatol. Surg. Oncol.* **17**, 681–684 (1991).

82. Weber, P. J., Moody, B. R., Dryden, R. M. & Foster, J. A. Mohs surgery and processing: novel optimizations and enhancements. *Dermatol. Surg.* **26**, 909–914 (2000).

83. Aslam, A. & Aasi, S. Z. Frozen-section tissue processing in Mohs surgery. *Dermatol. Surg.* **45**, S57 (2019).

84. Tang, M., Gorelick, L., Veksler, O. & Boykov, Y. GrabCut in One Cut. In 1769–1776 (2013).

85. Ünver, H. M. & Ayan, E. Skin lesion segmentation in dermoscopic images with combination of YOLO and GrabCut algorithm. *Diagnostics* **9**, 72 (2019).

86. Bay, H., Tuytelaars, T. & Van Gool, L. SURF: Speeded Up Robust Features. in *Computer Vision – ECCV 2006* (eds. Leonardis, A., Bischof, H. & Pinz, A.) 404–417 (Springer, 2006). https://doi.org/10.1007/11744023_32.

87. Sarlin, P-E, DeTone, D., Malisiewicz, T. & Rabinovich, A. SuperGlue: Learning feature matching with graph neural networks. *arXiv:1911.11763 [cs]* (2019).

88. Acharya, K. A., Venkatesh Babu, R. & Vadhiyar, S. S. A real-time implementation of SIFT using GPU. *J. Real.-Time Image Process.* **14**, 267–277 (2018).

89. Heymann, S., Müller, K., Smolic, A., Froehlich, B. & Wiegand, T. SIFT implementation and optimization for general-purpose GPU. (2007).

90. Schonberger, J. L. & Frahm, J-M Structure-from-motion revisited. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 4104–4113 (2016).

91. Sun, J., Xie, Y., Chen, L., Zhou, X. & Bao, H. NeuralRecon: Real-time coherent 3D reconstruction from monocular video. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 15598–15607 (2021).

92. Van der Walt, S. et al. scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).

93. Scharr, H. Optimal filters for extended optical flow. in *International Workshop on Complex Motion* 14–29 (Springer, 2004).

94. Rusu, R. B. & Cousins, S. 3d is here: Point cloud library (pcl). In *2011 IEEE international conference on robotics and automation* 1–4 (IEEE, 2011).

95. Akkiraju, N. et al. Alpha shapes: definition and software. in *Proceedings of the 1st international computational geometry software workshop* vol. 63 (1995).

96. Modern Analytic Apps for the Enterprise. *Plotly* https://plot.ly.

97. Hanwell, M. D., Martin, K. M., Chaudhary, A. & Avila, L. S. The visualization toolkit (VTK): Rewriting the rendering code for modern graphics cards. *SoftwareX* **1**, 9–12 (2015).

98. Levy, J., Salas, L. A., Christensen, B. C., Sriharan, A. & Vaickus, L. J. PathFlowAI: A high-throughput workflow for preprocessing, deep learning and interpretation in digital pathology. *Pac. Symp. Biocomput.* **25**, 403–414 (2020).

99. Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).

100. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. *arXiv:1512.03385 [cs]* (2015).

101. Paszke, A. et al. PyTorch: An imperative style, high-performance deep learning library. *arXiv:1912.01703 [cs, stat]* (2019).

102. Li, Y. et al. A comprehensive review for MRF and CRF approaches in pathology image analysis. *Arch. Comput. Methods Eng.* https://doi.org/10.1007/s11831-021-09591-w (2021).

103. OpenSeadragon. http://openseadragon.github.io/.

104. Nokeri, T. C. Python Web Frameworks and Apps. In *Web App Development and Real-Time Web Analytics with Python* 79–85 (Springer, 2022).

105. Klicpera, J., Bojchevski, A. & Günnemann, S. Predict then propagate: Graph neural networks meet personalized pagerank. *arXiv preprint arXiv:1810.05997* (2018).

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

J.J.L., M.R.L. and L.J.V. conceived of study. M.J.D., R.S.C., M.J.D., L.J.F., M.R.L., and L.J.V. curated and annotated data. T.G., A.P., I.N., A.A., A.S., R.V., and J.J.L. performed analyses. J.J.L. and M.R.L. drafted initial manuscript, with initial feedback from B.C.C. and M.S.H. All authors contributed to the writing and editing of this work. All authors approved the completed version of this work.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41698-023-00477-7.

Correspondence and requests for materials should be addressed to Joshua J. Levy.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.