Automated Histopathologic Classification of Brain Tumors Using Artificial Intelligence

Quin Yuhui Xie, Kevin Faust, Dominick Han, Zadeh Gelareh, Kenneth Aldape, Zoya Volynskaya, Ugljesa Djuric, Phedias Diamandis

Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada, M5S 1A8
Department of Computer Science, University of Toronto, 40 St. George Street, Toronto, Ontario M5S 2E4
Princess Margaret Cancer Centre, MacFeeters-Hamilton Brain Tumour Centre, College Street 101, Toronto, ON, Canada, M5G 1L7
Laboratory Medicine Program, University Health Network, 200 Elizabeth Street, Toronto, ON, Toronto, Ontario, Canada, M5G 2C4

Abstract

Purpose: Brain tumors represent a diverse group of neoplasms with highly variable therapies and clinical outcomes. Early personalized management and initiation of precision-based molecular studies still heavily relies on morphologic interpretation of hematoxylin and eosin (H&E)-stained slides. Unfortunately, due to its qualitative nature, histopathological classification is prone to well-recognized inter-observer variability. To overcome this limitation, we developed an objective morphology-based brain tumor classifier using a deep convolutional neural network (CNN).

Methods: Our CNN is trained on a dataset of over one million pathologist- and molecularly-annotated image patches from H&E slides spanning over 20 common brain tumor classes. Importantly, our tool is fully automated, compatible with standard pathology workflows and provides prompt whole-slide annotation and lesion classification in under 5 minutes.

Results: The performance of our CNN-based tumor classifier is highly concordant with board-certified pathologists and confirmatory immunohistochemical stains. Testing reveals an area under the receiver operator characteristic (AUC) of > 0.95 for multiple classification tasks, including lesion localizing and differentiating among different brain tumor types. In certain scenarios, it also offers objective predictions of molecular alterations (IDH mutations and 1p19q co-deletions). Lastly, we use cloud-computing to provide our classifier as a web-based tool capable of rendering timely second opinions and quality assurance to remote cancer centers requiring additional subspecialized neuropathological expertise.

Conclusions: This study demonstrates the efficacy of utilizing artificial intelligence to create an autonomous histologic brain tumor classifier. Acutely, our compact tool aims to provide prompt, intra-operative information to help tailor surgical resections and personalized therapies. In the sub-acute setting, our CNN can provide objective triaging of molecular tests to help reduce diagnostic work-up times, costs and subjective interpretative errors. Our classifier thus has immediate translational potential as a rapid, precise and cost-effective tool to help guide personalized care in neuro-oncology.

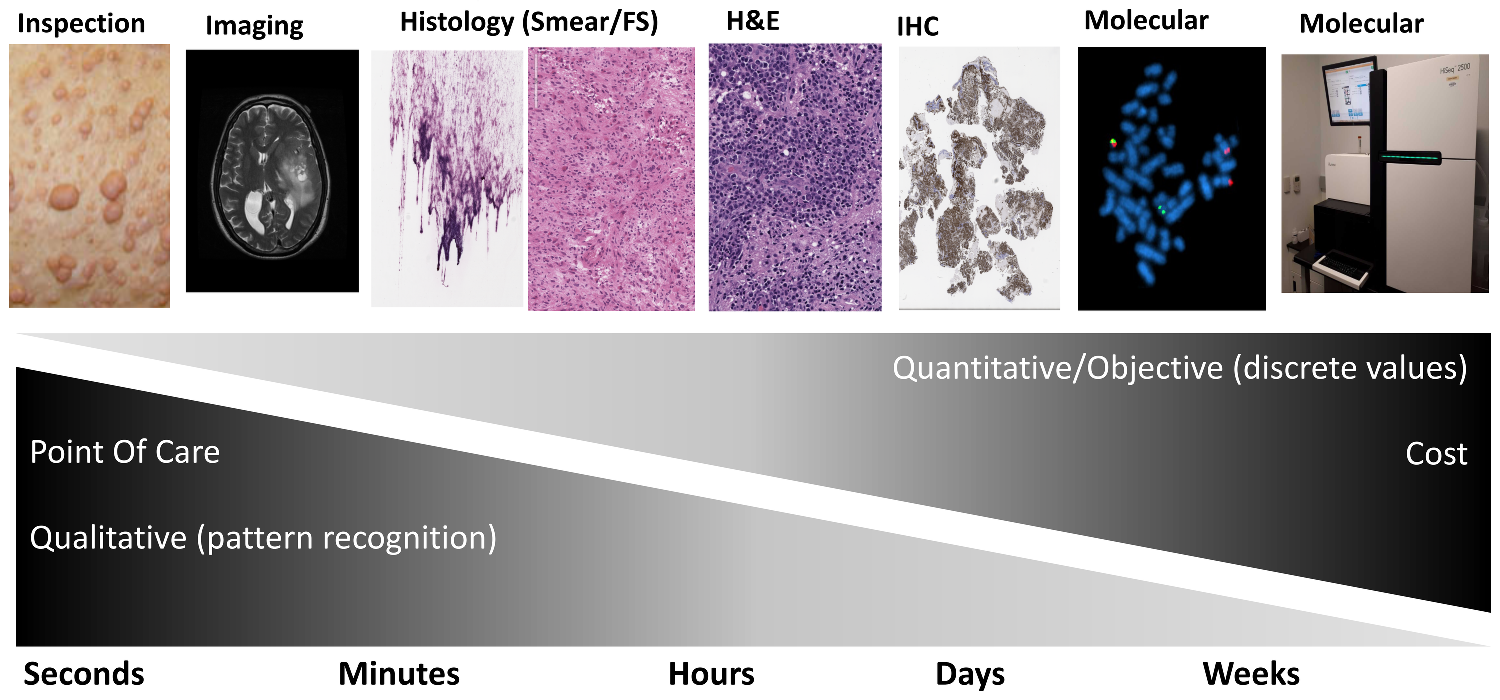
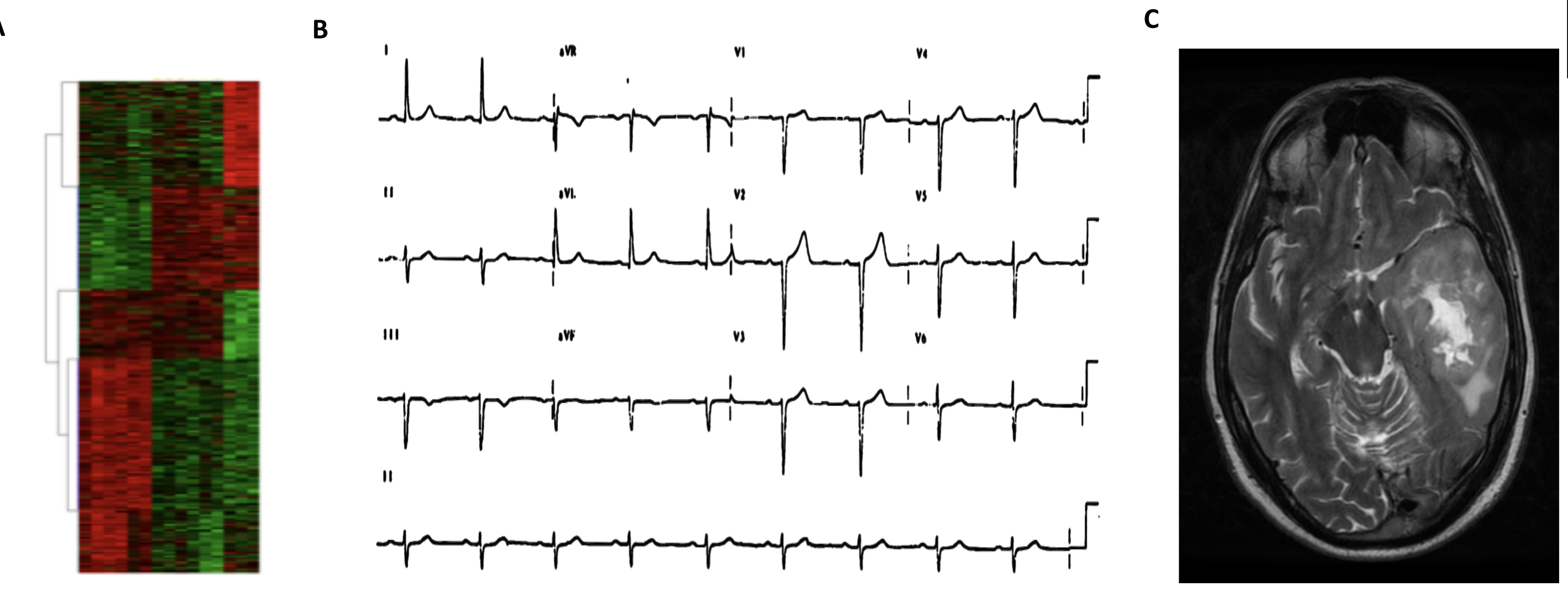


Figure 1 | Pattern recognition tasks in diagnostic medicine. Unlike the non-acute setting where medical information is usually generated as discrete, objective and quantitative values, most clinical data accessible early and close to the point of care, is typical visual. This information often requires highly trained personnel for interpretation and even then, classifications are usually broad, qualitative and prone to inter-observer reliability.



D

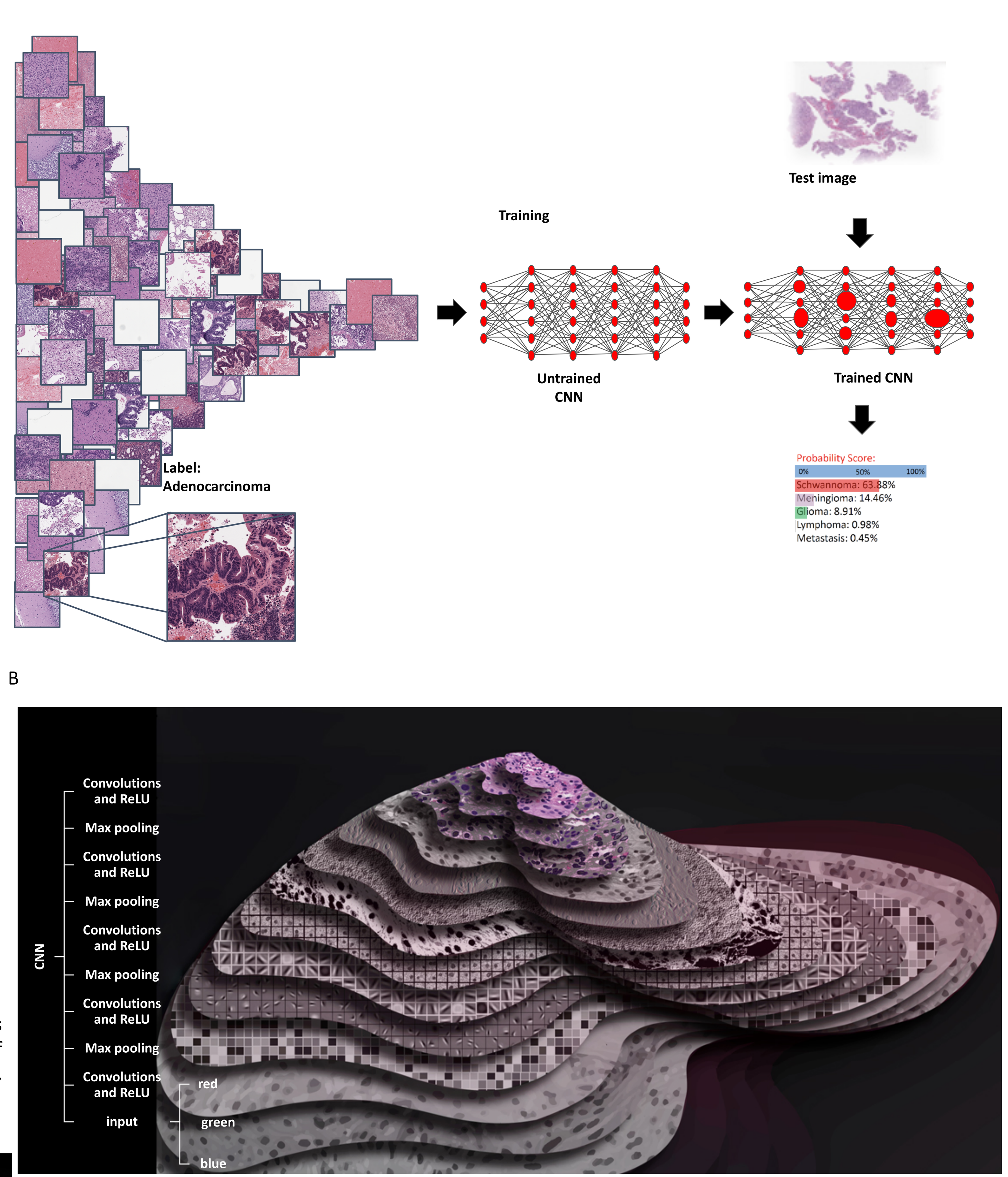
n	Combination C(100,n)	Permutations P(100,n)
1	100	100
4	$3.9x10^{7}$	$9.4x10^{7}$
16	$1.4x10^{18}$	$2.8x10^{31}$
32	$1.4x10^{26}$	$3.8x10^{61}$
64	$2.0x10^{27}$	$2.5x10^{116}$
75	$2.4x10^{23}$	$6.0x10^{132}$

Figure 2 | Computational analysis in medical data. Different data structures provide different orders of magnitude of computational complexity. **A.** Traditional machine learning approaches are able to efficiently handle combination of gene and protein expression value patterns where each data variable can be considered independent from one another. Permutational data however, where data points are temporally (Panel **B**, electrocardiogram (ECG)) or spatial (Panel **C**, MRI images) organized, pose progressively more challenging computational task. Panel **D** shows a table highlighting possible combination and permutations of a hypothetical task of arranging "n" data points chosen from a set of 100 to illustrate this. When the positional order of data matters (eg. images), the number of possible arrangements quickly grows orders of magnitude higher than other data types.

Objectives and Hypothesis:

Development of a CNN model for identification of most commonly seen brain tumors based on histology.
Validate performance of CNN classification to ensure that overfitting of model due to limited or non-representative training images did not occur.

We hypothesize that the application of trained CNN in the sphere of neuropathology will significantly improve the accuracy and efficiency of morphologic analysis, and essentially, living conditions of the patients.



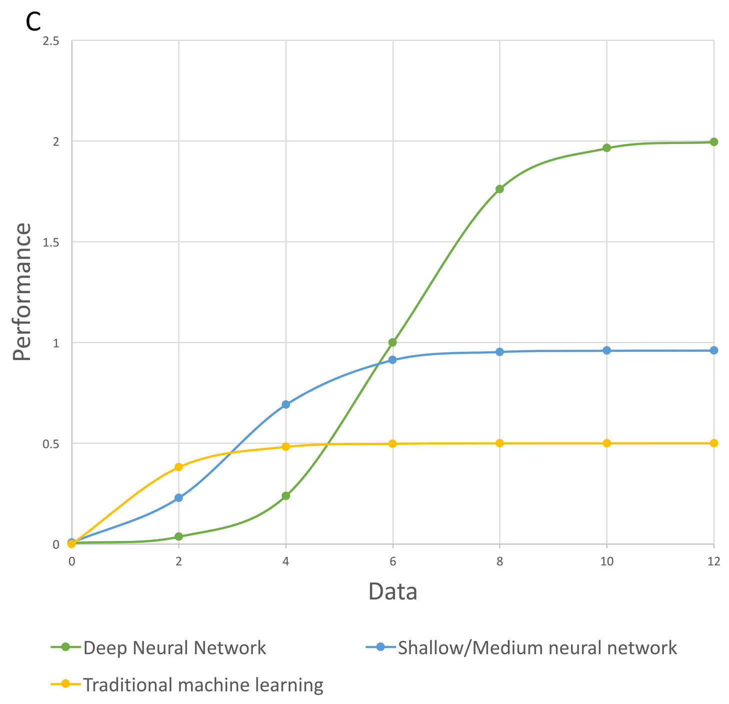


Figure 3 | Precision in medical image analysis using deep neural networks . Advancements in computer vision now aim to now bring objectivity to image analysis. Computational approaches such as convolutional neural networks are now able to accept massive cohorts of annotated images and discover informative patterns in a data-driven process independant of hand crafted features. Panel **B**. During deep learning, training images are first individually analyzed pixel by pixels to first find simple elementary patterns such as lines and circles. These features are then feed to deep layers where these objects are used to create more complex features. Feature selection occurs through a process called backpropagation where features of diagnostic importance are prioritized by changing "weightings" within the neural networks. By searching for these features in future images, computer can provide objective diagnostic interpretations of spatially dependent data.

C. Unlike traditional machine learning approaches that generally use thousands of human "hand-crafted features" for pattern recognition, neural networks are able to develop "data-driven" features. This is a highly data intensive approach and usually requires large amounts of data before performance benefits can be achieved. As the layers of a neural network are increased, more complex features can be generated and extracted which leads to massive improvements in performance. This emerging pattern recognition tool now dominates computer vision.

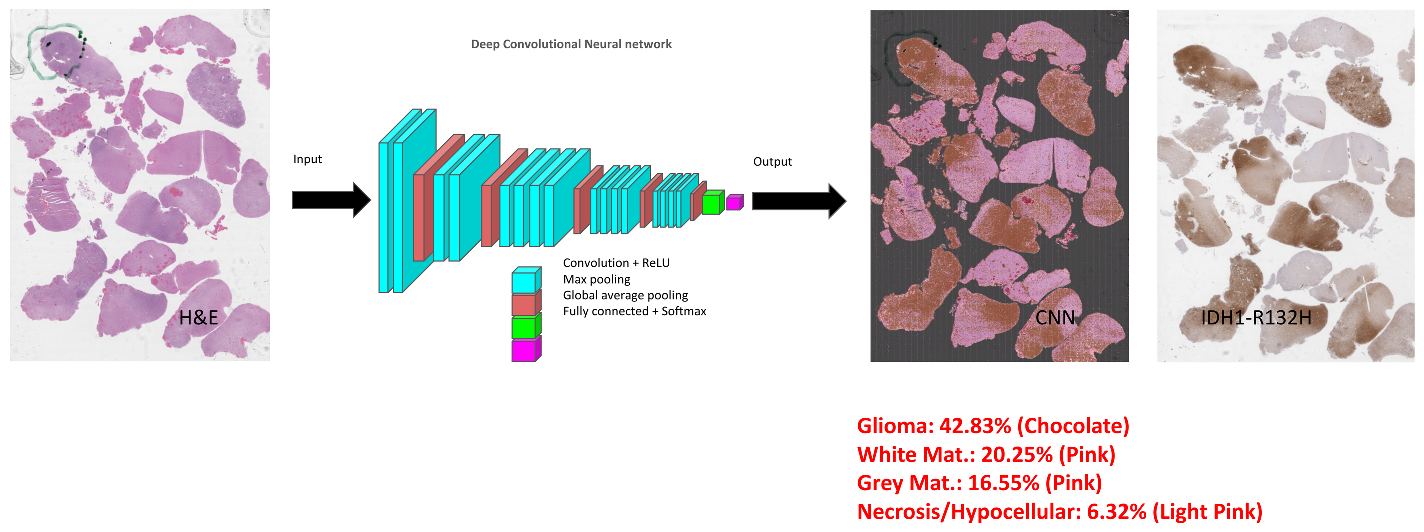


Figure 4 | Visualization of histopathological image classification using CNNs. A. Class activation map (CAM) showing the localization of lesional tissue on a WSI over 1 gigabite in size as detected by a trained CNN. For comparison an immunostain showing strong concordance with the ground truth. Classes with the highest probability score can be used to provide annotations and a diagnosis for the entire WSI

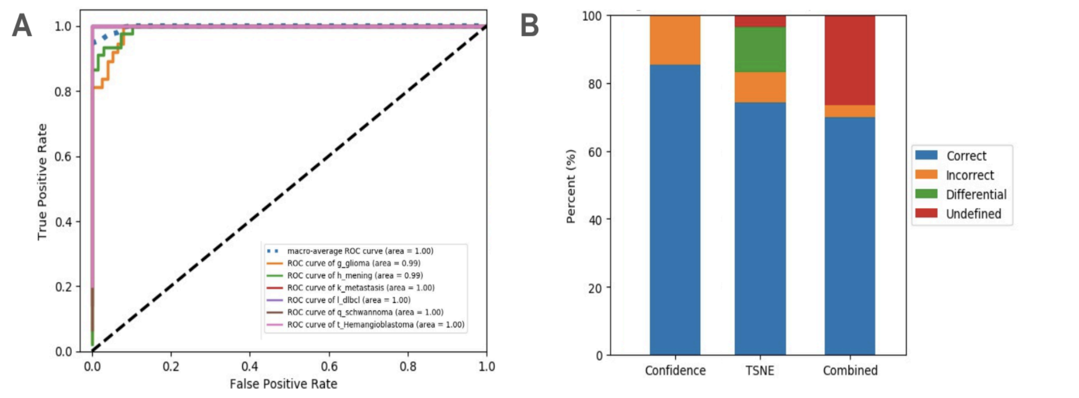
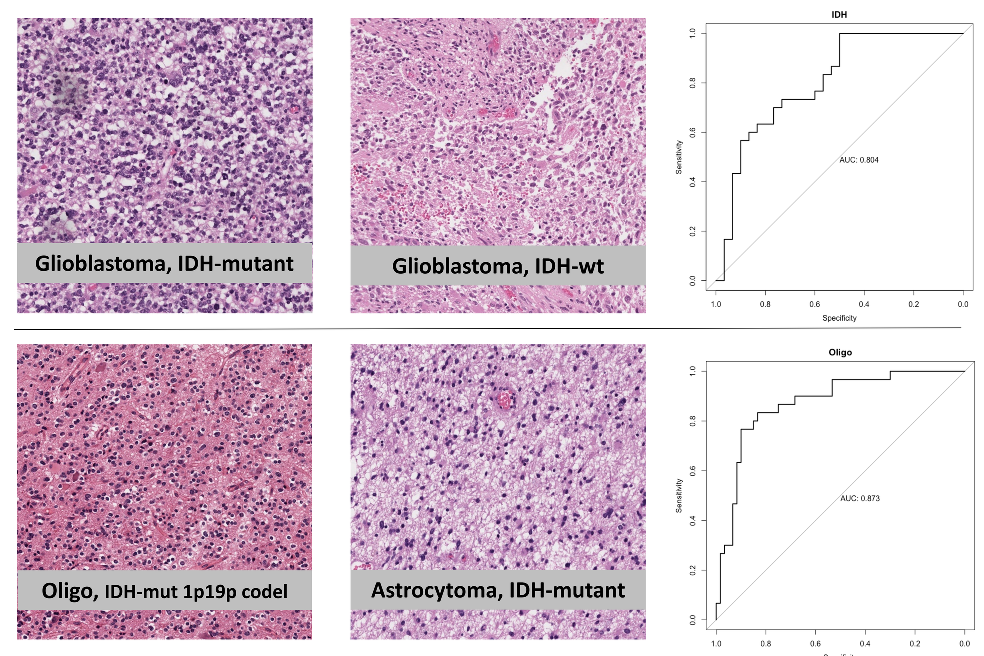


Figure 6: Biostatistics and Performance Testing. Performance of our automated lesion segmentation and classification workflow for 136 prospective and randomly selected WSIs of cerebral lesions.
A. Multi-class ROC curves were empirically generated by deriving the sensitivity (fraction of detected true positives) and specificity (fraction of detected true negatives) of the test WSIs at different probability score distribution thresholds. The displayed area under the curve is a measure of performance with a minimum value of 0.50 (random predictions) and 1.0 (all correct predictions). **B.** Performance displayed as accuracy of the top classification output of different approaches. For t-SNE and combined approaches, classification was determined by using tile distributions.



Other Output Parameters:
- ❑ Histologic Diagnosis
- ❑ Molecular Features
- ❑ Clinical Outcome
- ❑ Therapy Response

❑ **(tile-level ROC)**

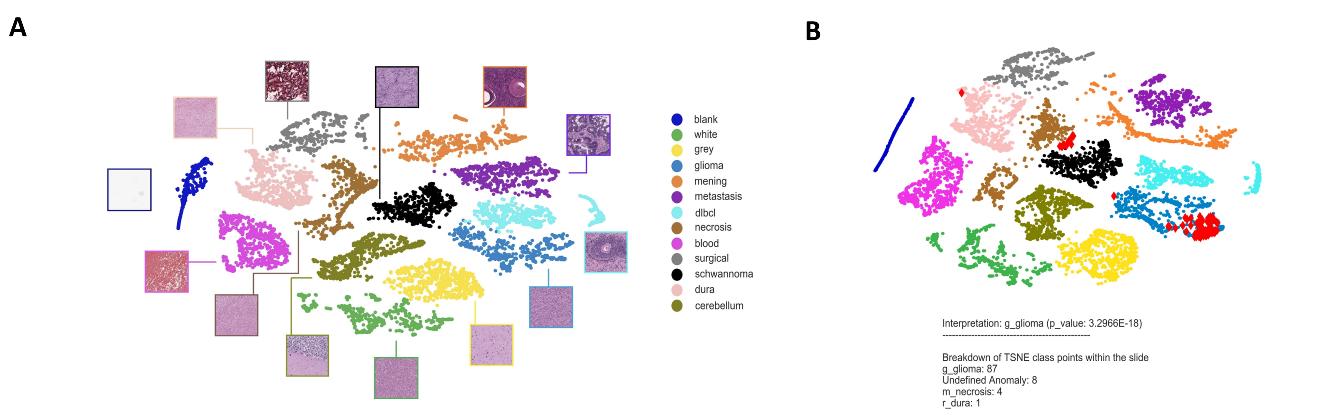


Figure 5 | Visualizing CNN image classification. Neural networks are often described as black boxes as the features used for classification are often not accessible to humans. **A, B.** CNN learning (**C**) and classification of test slides (**D**) can be visualized using dimensionality reduction techniques. Collectively, these tools help provide transparency to CNN-based classification disease and will be an important component to the implementation of these tool in a clinical setting.

Conclusions

This study demonstrates the efficacy of utilizing artificial intelligence to create an autonomous histologic brain tumor classifier. Specifically, our compact tool aims to provide prompt, intra-operative information to help tailor surgical resections and personalized therapies. In the sub-acute setting, our CNN can provide objective triaging of molecular tests to help reduce diagnostic work-up times, costs and subjective interpretative errors. Our classifier thus has immediate translational potential as a rapid, precise and cost-effective tool to help guide personalized care in neuro-oncology.

Selected References
1. E. Abrahams, M. Silver, The case for personalized medicine., *J. Diabetes Sci. Technol.* **3**, 680–4 (2009).
2. A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, *Nature* **542**, 115–118 (2017).
3. N. C. Dolan, Skin cancer control in the primary care setting: are we making any progress?, *J. Gen. Intern. Med.* **16**, 342–3 (2001).
4. M. Wilhelm, J. Schlegl, H. Hahne, A. Moghaddas Gholami, M. Lieberenz, M. M. Savitski, E. Ziegler, L. Butzmann, S. Gessulat, H. Marx, T. Mathieson, S. Lemeer, K. Schnatbaum, U. Reimer, H. Wenschuh, M. Mollenhauer, J. Slotta-Huspenina, J.-H. Boese, M. Bantscheff, A. Gerstmair, F. Faerber, B. Kuster, Mass-spectrometry-based draft of the human proteome., *Nature* **509**, 582–7 (2014).
5. M. Wilhelm, J. Schlegl, H. Hahne, A. Moghaddas Gholami, M. Lieberenz, M. M. Savitski, E. Ziegler, L. Butzmann, S. Gessulat, H. Marx, T. Mathieson, S. Lemeer, K. Schnatbaum, U. Reimer, H. Wenschuh, M. Mollenhauer, J. Slotta-Huspenina, J.-H. Boese, M. Bantscheff, A. Gerstmair, F. Faerber, B. Kuster, Mass-spectrometry-based draft of the human proteome., *Nature* **509**, 582–7 (2014).
6. M.-S. Kim, S. M. Pinto, D. Getnet, R. S. Nirujogi, S. S. Manda, R. Chaerkady, A. K. Madugundu, D. S. Kelkar, R. Isserlin, S. Jain, J. K. Thomas, B. Muthusamy, P. Leal-Rojas, P. Kumar, N. A. Sahasrabuddhe, L. Balakrishnan, J. Advani, B. George, S. Renuse, L. D. N. Selvan, A. H. Patil, V. Nanjappa, A. Radhakrishnan, S. Prasad, T. Subbannayya, R. Raju, M. Kumar, S. K. Sreenivasamurthy, A. Marimuthu, G. J. Sathe, S. Chavan, K. K. Datta, Y. Subbannayya, A. Sahu, S. D. Yelamanchi, S. Jayaram, P. Rajagopalan, J. Sharma, K. R. Murthy, N. Syed, R. Goel, A. A. Khan, S. Ahmad, G. Dey, K. Mudgal, A. Chatterjee, T.-C. Huang, J. Zhong, X. Wu, P. G. Shaw, D. Freed, M. S. Zahari, K. K. Mukherjee, S. Shankar, A. Mahadevan, H. Lam, C. J. Mitchell, S. K. Shankar, P. Satishchandra, J. T. Schroeder, R. Sirdeshmukh, A. Maitra, S. D. Leach, C. G. Drake, M. K. Halushka, T. S. K. Prasad, R. H. Hruban, C. L. Kerr, G. D. Bader, C. A. Iacobuzio-Donahue, H. Gowda, A. Pandey, A draft map of the human proteome., *Nature* **509**, 575–81 (2014).