Automated histologic diagnosis of CNS tumors with machine learning

Siri Sahib S Khalsa*,1 [ID], Todd C Hollon1, Arjun R Adapa2, Esteban Urias2, Sudharsan Srinivasan2, Neil Jairath2, Julianne Szczepanski3, Peter Ouillette3, Sandra Camelo-Piragua3 & Daniel A Orringer4

1Department of Neurosurgery, University of Michigan, Ann Arbor, MI 48109, USA
2Medical School, University of Michigan, Ann Arbor, MI 48109, USA
3Department of Pathology, University of Michigan, Ann Arbor, MI 48109, USA
4Department of Neurosurgery, New York University, New York, NY 10012, USA
*Author for correspondence: skhalsa@umich.edu

The discovery of a new mass involving the brain or spine typically prompts referral to a neurosurgeon to consider biopsy or surgical resection. Intraoperative decision-making depends significantly on the histologic diagnosis, which is often established when a small specimen is sent for immediate interpretation by a neuropathologist. Access to neuropathologists may be limited in resource-poor settings, which has prompted several groups to develop machine learning algorithms for automated interpretation. Most attempts have focused on fixed histopathology specimens, which do not apply in the intraoperative setting. The greatest potential for clinical impact probably lies in the automated diagnosis of intraoperative specimens. Successful future studies may use machine learning to automatically classify whole-slide intraoperative specimens among a wide array of potential diagnoses.

CNS masses

A new mass discovered in the CNS is a common reason for referral to a neurosurgeon. CNS masses are typically discovered on MRI or computed tomography (CT) scans after a patient presents with new neurologic symptoms. Presenting symptoms depend on the location of the tumor and can include headaches, seizures, difficulty expressing or comprehending language, weakness affecting extremities, sensory changes, bowel or bladder dysfunction, gait and balance changes, vision changes, hearing loss and endocrine dysfunction.

A mass in the CNS has a broad differential diagnosis, including tumor, infection, inflammatory or demyelinating process, infarct, hemorrhage, vascular malformation and radiation treatment effect. The most likely diagnoses can be narrowed based on patient demographics, medical history, imaging characteristics and adjunctive laboratory studies. However, accurate histopathologic interpretation of tissue obtained at the time of surgery is frequently required to make a diagnosis and guide intraoperative decision making.

Over half of CNS tumors in adults are metastases from systemic cancer originating elsewhere in the body [1]. An estimated 9.6% of adults with lung cancer, melanoma, breast cancer, renal cell carcinoma and colorectal cancer have brain metastases [2]. The remainder of CNS tumors in adults originate within the CNS and are called primary CNS tumors. The most common adult primary CNS tumors are meningiomas (36.8%), pituitary tumors (16.2%) and glioblastomas (14.9%) [3,4].

Among children and adolescents, the CNS is the most common cancer site overall. The prevalence of each type of tumor varies from adults. Whereas metastases are the most common type of CNS tumor in adults, primary CNS tumors are the most common in children. In patients aged 0–14 years with CNS tumors, pilocytic astrocytoma (17.8%) is the most common, followed by high-grade glioma (16.9%) and embryonal tumors (13.1%) [3].

Future Medicine

Tumors within the CNS are most likely to occur in the brain and meninges, whereas the spinal cord, spinal meninges and spinal nerve roots account for only 3.1% of adult and 4.9% of pediatric CNS tumors [3]. Spinal lesions can be further divided by location into intramedullary, intradural extramedullary and extradural. Intramedullary lesions arise within the spinal cord, and therefore the CNS proper. The most common intradural tumors are ependymomas and astrocytomas. Intradural extramedullary lesions occur outside the spinal cord but are in close proximity within the spinal thecal sac; these are typically meningiomas or nerve sheath tumors. In adults, meningiomas (38.8%), nerve sheath tumors (29.5%) and ependymal tumors (17.6%) are the most common intradural/intramedullary tumors [3]. In children, ependymal tumors (20.6%), astrocytomas (19.7%), nerve sheath tumors (17.8%), other neuroepithelial tumors (16.8%) and meningiomas (15.7%) are the most common intradural/intramedullary tumors [3]. Extradural masses are discussed separately below.

Spinal vertebral & extradural masses

The vertebral column and spinal epidural space, although outside the CNS, are important locations to consider given their proximity to the spinal cord. The vertebral and spinal epidural space are frequent sites of cancer metastases, which may present as neurosurgical emergencies in the setting of spinal cord compression and/or spinal instability.

Although the most common presenting symptom is back pain, the presence of cancer in the epidural space can compress the spinal cord, causing numbness, gait imbalance, weakness or paralysis in the extremities, and bowel and bladder incontinence. Metastatic epidural spinal cord compression (MESCC), which is present in approximately 5% of cancer patients, is therefore a source of significant morbidity in the cancer population [5].

Among adults, the most common primary sources of MESCC are prostate, breast and lung cancer (15–20%), followed by lymphoma, renal cell carcinoma and multiple myeloma (5–10%). The remaining cases are comprised of other primaries, including colorectal cancer, sarcomas, melanoma and seminoma [5,6]. Tumors within the thecal sac or spinal cord are rare relative to MESCC.

The importance of CNS tumor histology

The specific cell type, or histology, comprising a CNS tumor is a critical piece of information in the determination of an appropriate treatment plan and the estimation of a patient's prognosis.

A common category of brain mass appearance on MRI is the presence of rim-enhancement. The radiologic differential diagnosis of a rim-enhancing mass includes primary malignant glioma, metastasis, lymphoma, demyelination and infection; each of which has different surgical indications, treatment and prognosis. If the mass is a large accessible solitary brain metastasis, then the recommended treatment is typically complete surgical resection, followed by radiation therapy [7]. Conversely, in patients with primary CNS lymphoma, surgical resection is usually contraindicated in favor of chemotherapy with or without radiation [8]. A high-grade glioma, if accessible, will benefit from significant surgical debulking followed by radiation and chemotherapy [9]. In some cases, the imaging appearance of a tumor may be mimicked by demyelinating disease, which is treated with immunotherapy, without any role for surgical resection [10]. Infectious lesions may benefit from microbiology cultures and surgical debridement and will also require targeted antimicrobial therapy depending on the identity of the infectious organism.

Among gliomas, for example, the histologic grade highly informs a patient's predicted overall survival. The highest grade glioma, known as glioblastoma, confers a poor prognosis with median survival of 14.6 months with the maximal treatment of surgical debulking, followed by radiation and temozolamide chemotherapy [9]. Low-grade gliomas, however, are associated with an overall survival of over 11 years after surgical treatment [11]. Obtaining viable representative tumor is also important for proper molecular classification, even if only from a needle biopsy.

In the pediatric population, the distinction between ependymoma and medulloblastoma can guide the aggressiveness with which the neurosurgeon resects the tumor. In both cases, the goal of surgery is maximal safe resection. However, gross total resection, as opposed to near-total resection, is thought to result in a greater improvement in overall survival for ependymomas as compared with medulloblastomas [12,13]. In all patient populations, surgical resection has little role in the treatment of germinomas, which are typically highly responsive to radiotherapy and chemotherapy [14]. Therefore, the diagnosis of germinoma would typically guide the neurosurgeon to cease efforts to surgically resect the tumor.

The importance of tumor histology in metastatic epidural spinal cord compression

Metastatic epidural spinal cord compression typically requires emergency neurosurgical consultation to prevent irreversible spinal cord injury. The appropriate treatment plan considers multiple patient factors, including the presence of neurologic deficits, the radiation sensitivity of the tumor, any disruption in mechanical stability of the spine and the overall state of the patient's systemic disease [6,15].

The expected radiation response of a tumor is determined by its histology and is a critical component in determining whether surgical resection is required. Lymphoma, myeloma, seminoma and small cell carcinoma are radiosensitive. Many types of sarcomas, melanoma, non-small-cell lung cancer, renal cell carcinoma and gastrointestinal cancers are radioresistant. Prostate cancer and breast cancer are considered intermediate in terms of their radio-responsiveness.

In radiosensitive cases, surgical tumor resection can be avoided, and conventional external beam radiation therapy is typically sufficient to achieve local control. In radioresistant cases a neurosurgeon can offer separation surgery, wherein the tumor is partially resected away from the spinal cord in order to allow high-dose focused radiation to improve local control (stereotactic radiosurgery or stereotactic body radiotherapy). Attempts at gross-total tumor resection are often deferred in the setting of metastatic disease to avoid significant surgical morbidity [6,15].

The method of surgical pathology

Before surgery, a neurosurgeon can formulate a prediction of the identity of a CNS mass using a patient's medical history, the imaging characteristics of the mass and adjunctive tests. However, a mass that is thought preoperatively to represent a brain metastasis, for example, may be discovered by the pathologist to be a glioblastoma, lymphoma, infection or inflammatory/demyelinating process [7]. The gold standard for determination of the final diagnosis is examination of a slide of tissue by a neuropathologist. To prepare a permanent specimen, a sample of tissue is fixed in formalin, embedded in wax and stained with hemotoxylin and eosin (H&E). The sample may also be analyzed further with immunohistochemical stains and molecular testing for accurate diagnosis and proper classification. Overall, this process of obtaining a final pathologic diagnosis lasts multiple days and is important for specific treatment and prognosis.

However, as described above, the goals of surgery and subsequent treatment can be significantly influenced by the specific histology of the mass. For this reason, neurosurgeons will typically send a sample of the mass to the pathologist for an expedited intraoperative consultation (IOC) to decide the next surgical step. The standard options for preparing tissue for IOC are cytologic preparations (smear and touch preps) and frozen sections. A cytologic preparation allows the pathologist to study the cellular features of the lesion, but the overall histologic architecture is lost. Frozen sections involve freezing the specimen to allow slicing and staining with rapid H&E, which preserves the underlying tissue architecture but may introduce several freezing artifacts. Many neuropathologists advocate for using both frozen and smear preparations together in a complementary fashion to improve interpretation [16]. Intraoperative tissue preparation is more difficult to interpret than permanent sections and is interpretable within only 20 minutes, as opposed to multiple days for a permanent specimen [17]. Depending on the diagnosis, the results of this IOC may steer the neurosurgeon to either abort the operation or pursue aggressive surgical resection.

The need for automated histology

The ability to formulate an expedited preliminary diagnosis during surgery is a critical service during any brain surgery, as described above. In the USA, the number of board-certified neuropathologists is only 56% the number of hospitals performing brain tumor surgery [18]. Hospitals without board certified neuropathologists often rely on general pathologists who may not be fully experienced with the subtleties of CNS lesions. This has led to recent attempts to automate the process of interpreting slides with machine learning, which are reviewed below. In addition, automation has the potential to increase the speed at which a preliminary diagnosis is made. A shortened time to intraoperative diagnosis is particularly beneficial if it is able to reduce the duration of surgery while remaining as accurate as the current gold standard, as reduced time in surgery is associated with less perioperative risk and healthcare costs [19–21].

To assist in the discussion, a glossary of machine learning terms described in this article are defined in Table 1.

An introduction to machine learning

Machine learning is a general term that refers to computer algorithms that adapt themselves to improve accuracy during pattern recognition tasks. A common application of machine learning is the task of image classification,

Table 1. Glossary of machine learning terms addressed in this article, as applied to image classification.		
Term	**Description**	**Ref.**
Machine learning	Computer algorithms that automatically adapt themselves by experience in order to improve accuracy	
Conventional machine learning	A machine learning approach wherein the features used to characterize an image are predesigned by a human engineer	[22]
Convolutional neural network	A deep learning neural network architecture that utilizes the discrete convolution as a mathematical filtering operation	[22]
Deep learning	A machine learning approach that automatically discovers the necessary features needed to characterize an image	[22]
Image classification	An application of machine learning to automatically determine the correct category, label, or class, of a digital image (e.g., dog, table, bicycle)	
Image segmentation	The process of partitioning an image into its various labeled components (e.g. separating cell nuclei from other components in a digitized histology slide).	
Neural network	A digital multilayered interconnected network of data calculation points (nodes or neurons) which together performs a machine learning task	
Pixel	A number that represents the intensity of a color at a specific location in the image	
Random forest	A conventional machine learning approach that functions by generating a large number of randomly generated decision trees, which collectively vote on the most likely class	[45]
Support vector machine	A simple machine learning technique to separate two classes of data, by calculating the best hyperplane which separates data of one class from the other class	[47]
Transfer learning	A method of training a neural network wherein the network is pretrained for a certain task, and then partially retrained on a new set of data for a new purpose, which may allow for training with a smaller dataset as compared with training a network from scratch	[22]

such as determining whether an image contains a dog, table or bicycle. A digital image is simply a matrix of pixels, each of which is a number that represents the intensity of a color at a specific location in the image. To confront the task of automatically classifying an image, the computer algorithm must first extract a collection of features that describe the image more abstractly than pixel values. These features might represent edges arranged in certain orientations or texture patterns, for example. The machine learning algorithm is trained on several example images to determine the patterns of features needed to classify new images.

The choice of features to characterize an image may be predesigned by a human engineer, in which case the process falls under conventional machine learning. On the other hand, the algorithm can be designed to discover all the necessary image features itself, in a process called deep learning [22]. Interest in deep learning has risen sharply in recent years after the AlexNet deep convolutional neural network (CNN) achieved remarkable accuracy in classifying real-world images into 1000 different classes [23]. The performance of deep learning networks for image classification has since improved rapidly [24–27].

Machine learning applied to histopathology

Machine learning has shown promise for the automated interpretation of both imaging and histopathology of CNS tumors. With regard to brain MRI, machine learning has been applied for automated tumor identification, quantification of tumor burden and assessment of treatment response [28–30]. MRI-based automated brain tumor classification has also been an area of active research [31,32], including efforts to use machine learning to classify gliomas by molecular subtypes using MRI prior to tissue acquisition and histopathologic analysis, such as the status of *IDH1* mutation [33,34], *MGMT* promoter methylation [33] and 1p19q codeletion [34]. With regard to histopathology, most of the research effort has focused on automated classification and segmentation of formalin-fixed paraffin-embedded permanent specimens in organ systems outside the CNS [35–37]. The success of machine learning techniques in the automated classification of specimens outside the CNS has generated significant interest in applying them to CNS lesions. In this section, we review the current literature on automated interpretation of CNS histopathology, with a focus on intraoperative specimens.

A framework for the general method of using deep machine learning for classifying whole slide images is illustrated in Figure 1. A training set consisting of labeled histopathologic images is used to train a CNN. A whole slide with unknown diagnosis is then split into smaller fields of view, which are fed into the trained CNN for automated classification. The classified fields of view are then summated into a whole-slide diagnosis.
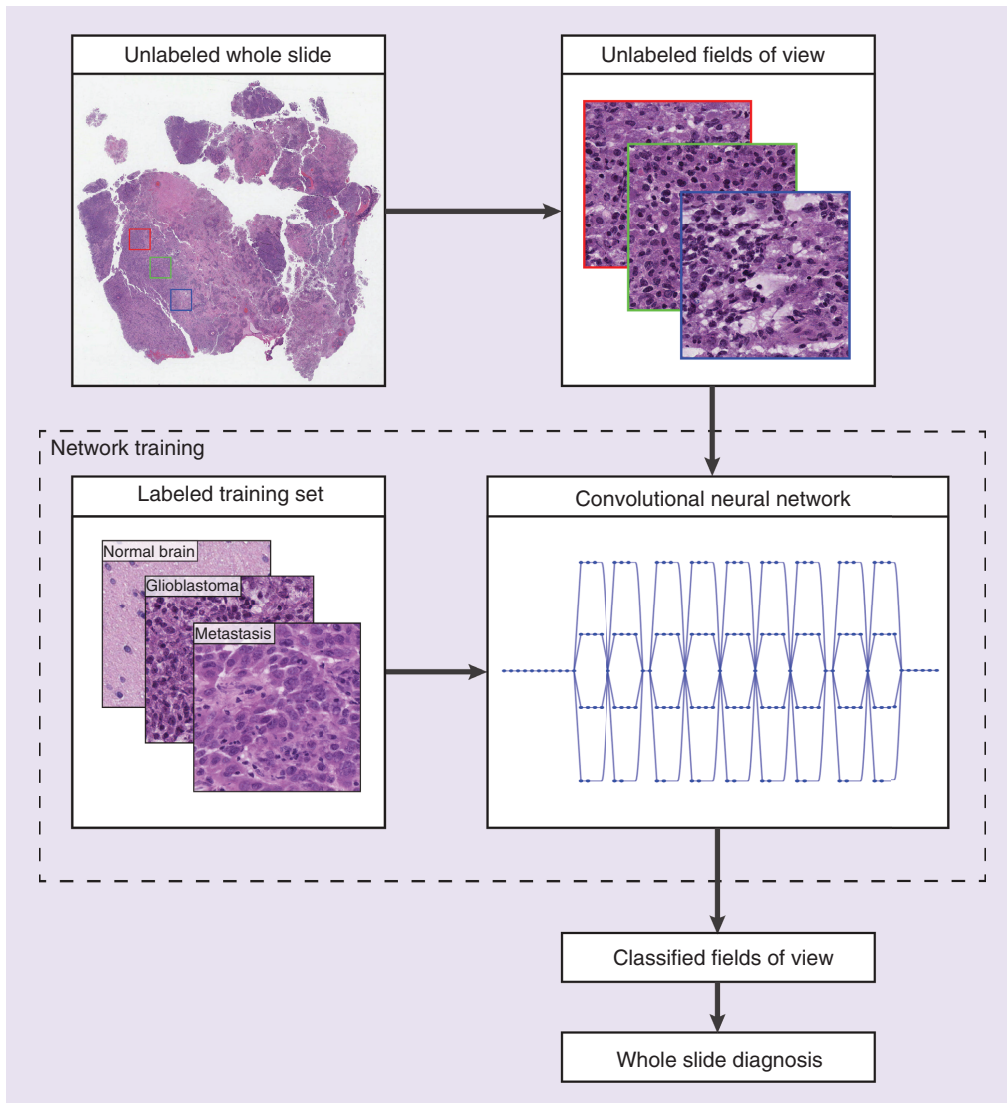
Figure 1.   General framework for automated whole-slide classification using a deep convolutional neural network. A training set consisting of labeled histopathologic images is used to train a convolutional neural network. A whole slide with unknown diagnosis is then split into smaller fields of view, which are fed into the trained convolutional neural network for automated classification. The classified fields of view are then summated into a whole-slide diagnosis.

Machine learning for fixed specimens

Most of the published efforts in automated histopathologic classification among all organ systems have focused on fixed, or permanent, specimens. These specimens have the advantage of fewer artifacts than frozen or smear preparations, and a large amount of labeled data are available to research groups. Glioma grading has been the focus for studies that have applied machine learning to permanent whole-slide specimens in the CNS.

Barker *et al.* [38] achieved 93.1% accuracy in classifying whole-slide permanent H&E specimens as either glioblastoma or low-grade glioma (binary classification). They divided whole slides into 1024 × 1024 pixel tiles at 20× magnification, which were analyzed with a conventional machine learning model of feature extraction followed by an Elastic Net [39] linear regression model.

Xu *et al.* [40] used a deep learning approach for the similar task of classifying whole-slide permanent H&E specimens as either glioblastoma or low-grade glioma (binary classification) and achieved 97.8% accuracy. They also automatically segmented necrotic regions in the glioblastoma specimens with 84% accuracy. Their approach used transfer learning with the AlexNet CNN trained on the ImageNet dataset [23]. In a transfer learning approach,

Table 2. Summary of publications with machine learning approaches to the automated diagnosis of intraoperative histopathologic specimens of CNS tumors.

Study	Image Type	Classes	Method	Ref.
Abas *et al.* (2016)	Smear preparation	Neoplastic astrocytes and non-neoplastic astrocytes	Conventional machine learning	[42]
Orringer *et al.* (2017)	Stimulated Raman histology	Low-grade glioma, high-grade glioma, non-neoplastic cortex, metastatic adenocarcinoma, metastatic melanoma and meningioma	Conventional machine learning	[18]
Hollon *et al.* (2018)	Stimulated Raman histology	Normal brain, ganglioglioma, pilocytic astrocytoma, primitive neuroectodermal tumor, dysembryoplastic neuroepithelial tumor, ependymoma and medulloblastoma	Conventional machine learning	[44]
Hollon *et al.* (2020)	Stimulated Raman histology	Malignant glioma, diffuse low-grade glioma, pilocytic astrocytoma, ependymoma, lymphoma, metastases, medulloblastoma, meningioma, pituitary adenoma, gliosis, white matter, gray matter and nondiagnostic tissue	Convolutional neural network	[46]

the network used is one that is already trained to extract a set of features, usually those necessary to classify real-world scenes and natural objects [41]. Histopathology slides are then used to train the last few layers of this pretrained neural network, such that fewer training images are needed as compared with training the network from scratch.

Machine learning for intraoperative histopathology

There have been fewer attempts to apply machine learning to intraoperative specimens, especially in the CNS. Intraoperative specimens pose unique challenges, including increased artifacts introduced by tissue preparation, variability in preparation techniques among institutions and a relative lack of digitized data available for research. Table 2 summarizes studies that have applied machine learning to classify intraoperative histopathologic specimens of CNS tissue and tumors.

Abas *et al.* [42] reported one of the earliest attempts at machine learning for intraoperative CNS pathology. They used a conventional machine learning approach to distinguish neoplastic astrocytes from non-neoplastic astrocytes on intraoperative smear preparations from patients with suspected recurrent glioma. Slides were divided such that one cell was analyzed per image. Geometric and textural features were extracted from the image and fed into five different classifiers for comparison. The authors' best performance was achieved with a binary classifier and support vector machine with F1 scores 0.92–0.94 [42].

Orringer *et al.* [18] introduced a novel rapid approach to intraoperative pathology for neurosurgical patients. A label-free microscopy method called stimulated Raman scattering was coupled with an image processing technique, stimulated Raman histology (SRH), to generate virtual H&E slides. These digital SRH images provided the same diagnostic histopathologic features as permanent H&E specimens, but without any tissue processing. A conventional machine learning technique was also applied for automated classification of the SRH specimens, which included diagnoses of low-grade glioma, high-grade glioma, non-neoplastic cortex, metastatic adenocarcinoma, metastatic melanoma and meningioma. Each specimen was divided into 400×400 micron fields of view. A set of image features were extracted with the software WND-CHARM [43] and classified with a machine learning technique called a multilayer perceptron. Specimens were classified as normal, low-grade glioma, high-grade glioma or nonglial tumor with 90% accuracy using this automated approach [18].

The same group applied SRH and conventional machine learning to automated intraoperative diagnosis of pediatric brain tumors [44]. Specimens consisted of normal brain, ganglioglioma, pilocytic astrocytoma, primitive neuroectodermal tumor, dysembryoplastic neuroepithelial tumor, ependymoma and medulloblastoma. Nuclear density, tumor-associated macrophage infiltration and nuclear morphology parameters were extracted from the fresh SRH specimens and used to train a classification technique called random forest. They were able to distinguish lesional versus nonlesional and high-grade versus low-grade specimens with 100% accuracy [44]. The random forest technique functions by generating a large number of randomly generated decision trees, which collectively vote on the most likely class [45]. As with other conventional machine learning techniques, the features used to generate the decision trees are deliberately selected and extracted from the images prior to training the classifier.

Most recently, Hollon *et al.* [46] trained a deep CNN to distinguish among 13 different classes of brain tumors and non-neoplastic brain tissue using intraoperative SRH specimens obtained during brain tumor surgery. The algorithm was able to distinguish malignant glioma, diffuse low-grade gliomas, pilocytic astrocytoma, ependymoma, lymphoma, metastases, medulloblastoma, meningioma, pituitary adenoma, gliosis, white matter, gray matter and nondiagnostic tissue. This CNN-based automated interpretation of intraoperative SRH images was tested in a

multicenter prospective fashion and found to arrive at a diagnosis in less than 150 seconds and with a diagnostic accuracy of 94.6% without human input, which was noninferior to pathologist-based interpretation of conventional intraoperative frozen and smear preparations with an overall accuracy of 93.9% [46].

Hollon *et al.* [46] provided additional insight into the potential role for deep neural networks to aid neuropathologists in improving intraoperative diagnostic accuracy. Of 278 specimens, 17 were misdiagnosed by expert board-certified neuropathologists (93.9% overall accuracy), but the CNN correctly classified all of these 17 difficult specimens. Similarly, 14 of 278 specimens were misdiagnosed by the CNN (94.6% overall accuracy), but the neuropathologists correctly interpreted these 14 cases. Therefore, there was no overlap in errors between the CNN and the neuropathologists in this study. This highlights the possibility that deep neural network-based automated interpretation may be used as an aid for neuropathologists in improving diagnostic accuracy for difficult cases, but also indicates the importance of validation by a trained neuropathologist to avoid misclassification by the neural network.

Conclusion

The accuracy with which machine learning algorithms can automatically classify real-world images has grown rapidly. Several groups have applied machine learning techniques to cancer histopathology, and some have focused on brain tumors. Multiple attempts have been successful with binary tasks, such as high-grade versus low-grade glioma or necrosis versus non-necrosis. Progress in automatically distinguishing among a wider selection of core diagnoses has been made using stimulated Raman histology specimens. To make the greatest impact on surgical decision-making, an automated system would need to read any whole-slide intraoperative specimen and automatically provide a diagnosis within minutes among a wide set of core diagnoses in brain and spine tumor specimens. A system that could automatically interpret intraoperative frozen and/or smear preparations has the potential to immediately impact tumor surgery in the CNS and spinal column, especially in settings with limited access to intraoperative neuropathology consultation.

Future perspective

Histologic diagnosis is key for decision-making during surgery for tumors affecting the brain and spine. The ideal automated system would be able to distinguish between all the core diagnoses that would influence intraoperative decision-making, when provided a digitized whole slide.

The closest achievement of this goal for brain tumors utilized a deep learning approach to interpreting stimulated Raman histology [46]. SRH has the potential to revolutionize intraoperative histopathology, although the specialized microscope is currently available at a limited number of research institutions. Most centers performing surgery on the CNS currently have the infrastructure to produce frozen and cytologic preparations; however, to the authors' knowledge, machine learning has not yet been applied for clinically significant automated interpretation of frozen sections for the CNS. One group did address smear preparations, but the task was limited to the binary task of distinguishing neoplastic from non-neoplastic cells in glioma specimens [42]. The ability to perform automated diagnosis on frozen and smear preparations would thus be a significant contribution to centers without ready access to neuropathologists in the intraoperative setting.

Overall, the number of different tumor types classified by machine learning-based algorithms and the classification accuracy are both increasing as automated histopathologic classification techniques are refined by multiple research groups. This raises the question of how automated classifiers might be used in the clinical setting. As described in this article, rapid automated diagnosis is most likely to influence clinical decision making in the intraoperative setting for brain and spine tumors, where histology can greatly influence immediate surgical goals.

We suspect that the automated classifiers will not be used as a replacement for human pathologists, but rather as a tool to improve the speed and accuracy with which a preliminary diagnosis is made. Hollon *et al.* [46] showed in a large prospective brain tumor series that the diagnostic errors made by board-certified neuropathologists were mutually exclusive with those made by their convolutional neural network. This suggests that a pathologist assisted by artificial intelligence has the potential to achieve greater accuracy than either interpreter alone; a hypothesis that would benefit from prospective validation.

Furthermore, a current limitation of machine learning algorithms in general is their reliance on a large number of training examples, which limits their ability to identify rare tumor types that are underrepresented in the training set. It will be important for automated classifiers in the clinical setting to produce a confidence metric along with the whole-slide diagnosis, such that difficult or rare cases are flagged for further expert review.

Finally, real-world neuropathologists do not interpret the tumor specimen in isolation in order to render a diagnosis. Additional factors, including patient demographics, clinical history and imaging, are all important in informing the pretest probability for various diagnoses. The field may benefit from more holistic classifiers that consider these other important medical and radiological parameters when calculating the most likely histopathologic diagnosis.

Practice points

- The appropriate treatment strategy for a new brain or spine mass is dependent on the histologic diagnosis.
- A preliminary histopathologic diagnosis is often established intraoperatively, via a pathologist's interpretation of a frozen section or cytologic preparation obtained during a biopsy surgery or at the beginning of a surgery intended to resect a mass.
- The limited availability of neuropathologists, especially in resource-limited settings, has prompted some groups to investigate artificial intelligence techniques for automated interpretation of histopathologic specimens. Deep learning is an artificial intelligence technique that has shown promise for automated diagnosis of digitized pathology specimens.
- Most groups have investigated automated diagnosis of formalin-fixed, paraffin-embedded specimens, which do not apply to intraoperative decision-making.
- The literature is sparse on attempts to automatically classify intraoperative specimens. These specimens pose unique challenges to automated interpretation, including an increased number of artifacts introduced by tissue preparation, variability in preparation techniques among institutions and a relative lack of digitized data available for research.
- Future studies have the potential to significantly impact brain and spine tumor surgery by developing algorithms to automatically provide a diagnosis for digitized intraoperative specimens, such as frozen and cytologic preparations.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Nayak L, Lee EQ, Wen PY. Epidemiology of brain metastases. *Curr. Oncol. Rep.* 14(1), 48–54 (2012).

2. Barnholtz-Sloan JS, Sloan AE, Davis FG, Vigneau FD, Lai P, Sawaya RE. Incidence proportions of brain metastases in patients diagnosed (1973 to 2001) in the Metropolitan Detroit Cancer Surveillance System. *J. Clin. Oncol.* 22(14), 2865–2872 (2004).

3. Ostrom QT, Cioffi G, Gittleman H *et al.* CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2012–2016. *Neuro Oncol.* 21(Suppl.5), v1–v100 (2019).

4. Ostrom QT, Gittleman H, Liao P *et al.* CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2010–2014. *Neuro Oncol.* 19(Suppl.5), v1–v88 (2017).

5. Cole JS, Patchell RA. Metastatic epidural spinal cord compression. *Lancet Neurol.* 7(5), 459–466 (2008).

6. Laufer I, Rubin DG, Lis E *et al.* The NOMS framework: approach to the treatment of spinal metastatic tumors. *Oncologist* 18(6), 744–751 (2013).

7. Patchell RA, Tibbs PA, Walsh JW *et al.* A randomized trial of surgery in the treatment of single metastases to the brain. *N. Engl. J. Med.* 322(8), 494–500 (1990).

8. Morris PG, Correa DD, Yahalom J *et al.* Rituximab, methotrexate, procarbazine, and vincristine followed by consolidation reduced-dose whole-brain radiotherapy and cytarabine in newly diagnosed primary CNS lymphoma: final results and long-term outcome. *J. Clin. Oncol.* 31(31), 3971–3979 (2013).

9. Stupp R, Mason WP, Van Den Bent MJ *et al.* Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N. Engl. J. Med.* 352(10), 987–996 (2005).

10. Hardy TA, Reddel SW, Barnett MH, Palace J, Lucchinetti CF, Weinshenker BG. Atypical inflammatory demyelinating syndromes of the CNS. *Lancet Neurol.* 15(9), 967–981 (2016).

11. Houillier C, Wang X, Kaloshi G *et al.* IDH1 or IDH2 mutations predict longer survival and response to temozolomide in low-grade gliomas. *Neurology* 75(17), 1560–1566 (2010).

12. Merchant TE, Li C, Xiong X, Kun LE, Boop FA, Sanford RA. Conformal radiotherapy after surgery for paediatric ependymoma: a prospective study. *Lancet Oncol.* 10(3), 258–266 (2009).

13. Thompson EM, Hielscher T, Bouffet E *et al.* Prognostic value of medulloblastoma extent of resection after accounting for molecular subgroup: a retrospective integrated clinical and molecular analysis. *Lancet Oncol.* 17(4), 484–495 (2016).

14. Calaminus G, Kortmann R, Worch J *et al.* SIOP CNS GCT 96: final report of outcome of a prospective, multinational nonrandomized trial for children and adults with intracranial germinoma, comparing craniospinal irradiation alone with chemotherapy followed by focal primary site irradiation for patients with localized disease. *Neuro Oncol.* 15(6), 788–796 (2013).

15. Spratt DE, Beeler WH, De Moraes FY *et al.* An integrated multidisciplinary algorithm for the management of spinal metastases: an International Spine Oncology Consortium report. *Lancet Oncol.* 18(12), e720–e730 (2017).

16. Plesec TP, Prayson RA. Frozen section discrepancy in the evaluation of central nervous system tumors. *Arch. Pathol. Lab. Med.* 131(10), 1532–1540 (2007).

17. Novis DA, Zarbo RJ. Interinstitutional comparison of frozen section turnaround time. A College of American Pathologists Q-Probes study of 32868 frozen sections in 700 hospitals. *Arch. Pathol. Lab. Med.* 121(6), 559–567 (1997).

18. Orringer DA, Pandian B, Niknafs YS *et al.* Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. *Nat. Biomed. Eng.* 1, 0027 (2017).

19 . Bekelis K, Coy S, Simmons N. Operative duration and risk of surgical site infection in neurosurgery. *World Neurosurg.* 94, 551–555.e6 (2016).

20. Cheng H, Clymer JW, Po-Han Chen B *et al.* Prolonged operative duration is associated with complications: a systematic review and meta-analysis. *J. Surg. Res.* 229, 134–144 (2018).

21. Golebiowski A, Drewes C, Gulati S, Jakola AS, Solheim O. Is duration of surgery a risk factor for extracranial complications and surgical site infections after intracranial tumor operations? *Acta Neurochir. (Wien)* 157(2), 235–240 (2015).

22. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature* 521(7553), 436–444 (2015).

•   **A review article that provides an excellent overview of deep machine learning.**

23. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60(6), 84–90 (2017).

24. Erhan D, Szegedy C, Toshev A, Anguelov D. Scalable object detection using deep neural networks. *Presented at: IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* OH, USA (2014).

25. Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, Inception-ResNet and the impact of residual connections on learning. *arXiv* arXiv:1602.07261 (2016).

•   **Technical description of state-of-the-art deep convolutional neural networks developed by Google, which exceed human performance in classifying real-world images.**

26. Szegedy C, Liu W, Jia YQ *et al.* Going deeper with convolutions. *Presented at: IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* MA, USA (2015).

27. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception architecture for computer vision. *arXiv* arXiv:1512.00567 (2015).

28. Chang K, Beers AL, Bai HX *et al.* Automatic assessment of glioma burden: a deep learning algorithm for fully automated volumetric and bidimensional measurement. *Neuro Oncol.* 21(11), 1412–1422 (2019).

29. Grovik E, Yi D, Iv M, Tong E, Rubin D, Zaharchuk G. Deep learning enables automatic detection and segmentation of brain metastases on multisequence MRI. *J. Magn. Reason. Imaging* 51(1), 175–182 (2020).

30. Kickingereder P, Isensee F, Tursunova I *et al.* Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. *Lancet Oncol.* 20(5), 728–740 (2019).

31. Grist JT, Withey S, Macpherson L *et al.* Distinguishing between paediatric brain tumour types using multi-parametric magnetic resonance imaging and machine learning: A multi-site study. *Neuroimage Clin.* 25, 102172 (2020).

32. Macyszyn L, Akbari H, Pisapia JM *et al.* Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques. *Neuro Oncol.* 18(3), 417–425 (2016).

33. Chang P, Grinband J, Weinberg BD *et al.* Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. *AJNR Am. J. Neuroradiol.* 39(7), 1201–1207 (2018).

34. Zhou H, Chang K, Bai HX *et al.* Machine learning reveals multimodal MRI patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low- and high-grade gliomas. *J. Neurooncol.* 142(2), 299–307 (2019).

35. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *J. Pathol. Inform.* 7, 29 (2016).

36. Komura D, Ishikawa S. Machine learning methods for histopathological image analysis. *Comput. Struct. Biotechnol. J.* 16, 34–42 (2018).

37. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: challenges and opportunities. *Med. Image Anal.* 33, 170–175 (2016).

38. Barker J, Hoogi A, Depeursinge A, Rubin DL. Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles. *Med. Image Anal.* 30, 60–71 (2016).

39. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J. R. Statist. Soc. B* 67(2), 301–320 (2005).

40. Xu Y, Jia Z, Wang LB *et al.* Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics* 18(1), 281 (2017).

41. Torrey L, Shavlik J. Transfer learning. In: *Handbook of Research on Machine Learning Applications.* Soria E, Martin J, Magdalena R, Martinez M, Serrano A (Eds). IGI Global, PA, USA (2009).

42. Abas FS, Gokozan HN, Goksel B, Otero JJ, Gurcan MN. Intraoperative neuropathology of glioma recurrence: Cell detection and classification. *Proc. SPIE* 9791, 979109 (2016).

•   **A conventional machine learning approach to distinguish neoplastic from non-neoplastic astrocytes on smear preparations. The application of machine learning to automated interpretation of intraoperative smear preparations of brain tumors is a relatively unexplored area.**

43. Orlov N, Shamir L, Macura T, Johnston J, Eckley DM, Goldberg IG. WND-CHARM: Multi-purpose image classification using compound image transforms. *Pattern Recognit. Lett.* 29(11), 1684–1693 (2008).

44. Hollon TC, Lewis S, Pandian B *et al.* Rapid intraoperative diagnosis of pediatric brain tumors using stimulated Raman histology. *Cancer Res.* 78(1), 278–289 (2018).

45. Breiman L. Random Forests. *Mach. Learn.* 45(1), 5–32 (2001).

46. Hollon TC, Pandian B, Adapa AR *et al.* Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nat. Med.* 26(1), 52–58 (2020).

••   **A prospectively validated deep machine learning approach to automatically distinguish among 13 different classes of brain tumors and non-neoplastic brain tissue, using intraoperative stimulated Raman histology specimens.**

47. Christianini N, Shawe-Taylor J. Support vector machines. In: *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods.* Cambridge University Press, Cambridge, UK, 93–121 (2000).