Deep learning: a turning point in acute neurology

Published Online
May 14, 2020
https://doi.org/10.1016/
S2589-7500(20)30106-0

See Articles page e314

Deep learning is a form of artificial intelligence, mimicking the structure and organisation of neurons and human intelligence in the brain. In the past decade, deep learning has been applied enthusiastically in the field of medicine, outperforming other established methods. In the different branches of neuroscience, for instance, deep learning algorithms have proven their worth in many ways including aiding the anatomical segmentation of specific brain structures,[1] the delineation of brain lesions including tumours,[2] and the image-based prediction of different neurological diseases.[3] Thanks to the optimisation of algorithms, improved computational hardware, and access to a large amount of imaging data, deep learning is no longer just an academic exercise, but a valuable tool in clinical practice.

In *The Lancet Digital Health*, Miguel Monteiro, Virginia Newcombe, and colleagues[4] provide a clinically significant, open-access deep learning algorithm, based on convolutional neural networks (CNN), with the ability to detect, segment, quantify, and differentiate brain lesion types on head CT images in patients with traumatic brain injury (TBI). TBI is a condition of vast proportions; worldwide it is the leading cause of mortality in young adults and a major cause of death and disabilities across all ages.[5] The major clinical challenges in patients with TBI include how to appropriately evaluate the patients, predict patient outcomes, and identify a suitable treatment strategy and rehabilitation programme. Almost universally, patients with TBI undergo CT scans as a conventional procedure to aid diagnosis. CT also plays a fundamental role for triage in the acute phase to identify patients who require urgent neurosurgical intervention.[6] Considering such background, this work[4] explored two important questions. First, how does the deep learning algorithm fit in the routine clinical setting to replace and improve time-consuming neuroradiological procedures? And second, how can this artificial intelligence method help in the medical and surgical treatment decision-making of individual patients with TBI?

Monteiro, Newcombe, and colleagues[4] used a dataset of 98 manually segmented scans to train and validate an initial CNN. This CNN was then used to segment a new dataset of 839 scans, which the authors then corrected manually, obtaining accurate ground truth data. A subset of 184 scans from the second dataset was used as a training set of the final CNN for multiclass voxel-wise segmentation of lesion types, and the remaining 655 scans were used as a testing set, evaluating the performance of this CNN. Finally, they used an independent set of 500 scans to externally validate their CNN. They found that compared with manual reference, CNN-derived lesion volumes showed a mean difference of 0·86 mL (95% CI –5·23 to 6·94) for intraparenchymal haemorrhage, 1·83 mL (–12·01 to 15·66) for extra-axial haemorrhage, 2·09 mL (–9·38 to 13·56) for perilesional oedema, and 0·07 mL (–1·00 to 1·13) for intraventricular haemorrhage.

Within the context of the overall flow of neuroradiology work, the CNN developed by Monteiro, Newcombe, and colleagues[4] addressed onerous tasks such as detecting and segmenting multiclass haemorrhagic lesions, both of which have strong implications in clinical practice, in particular in centres where radiological expertise is less easily available. Indeed, automation of lesion detection and segmentation would minimise inter-observer variability and discordance in reporting lesions, increase reliability and generalisability of findings, and reduce the amount of time-consuming assessments neuroradiologists would need to do. Previous studies[7,8] attempted to automate acute intracranial haemorrhage segmentation on relatively small datasets, without ensuring robustness and generalisability of the models. Major strengths of the study by Monteiro, Newcombe, and colleagues[4] are the use of large numbers of CT scans from multicentre, heterogeneous datasets to train, validate, and test the CNN, and the further external validation of the model on a large, independent, unseen dataset. This validation was essential to show consistency and efficacy of the algorithm performance in real-word application. Additionally, the authors used a so-called ground truth reference of manually annotated and manually corrected automatic segmentation of CT scans to more accurately assess the deep learning predictions. The availability of high-quality noiseless ground truth data enabled the algorithm to improve its learning capability.

What would be the impact of the CNN on the future medical and surgical treatment decision-making of patients with TBI? Previous approaches[9,10] to automate assessment of CT images for TBI were mostly limited to the undifferentiated binary detection of lesions with no volumetric analysis. Identifying and quantifying different types of haemorrhagic lesions and perilesional oedema in patients with TBI is crucial to unravel the heterogeneous spectrum of pathology involved and improve prediction of the clinical progression of a patient's TBI. In the quantitative multiclass segmentation assessment, the CNN developed by Monteiro, Newcombe, and colleagues[4] detected four lesion types—intraparenchymal haemorrhage, extra-axial haemorrhage, intraventricular haemorrhage, and perilesional oedema—and measured their volume with high accuracy. These features are particularly valuable in the TBI context for image-based diagnosis, for assessing injury type, and for quantifying its burden and progression. The capability of the aforementioned algorithm to complete a fine-grained segmentation of lesions and consequently to estimate their volumes allows accurate monitoring of lesion progression, which is a major predictor of clinical evolution and a key target for therapies in acute phases.

The study by Monteiro, Newcombe, and colleagues[4] stands out also for having dealt with the challenging delineation of hypointense oedema. Although the performance of the CNN for this delineation was suboptimal, automated quantification of perilesional oedema can provide useful additional information for prognosis and for delineation of surrogate outcome variables to facilitate the design of clinical trials aimed at reducing cerebral oedema and contusion growth.

One possible limitation of this algorithm is the inability to localise lesions. This would give more strength mainly in the assessment of small haemorrhagic lesions, whose effect depends mostly on number and regional distribution, rather than on the volume of each lesion. Moreover, implementing the algorithm to also distinguish extra-axial haemorrhagic lesions would increase the value of prognostic models.

The findings by Monteiro, Newcombe, and colleagues[4] highlight the potential gain in applying a CNN to obtain quantitative and objective neuroradiological metrics in patients with TBI. Future multimodal algorithms integrating neuroradiological, clinical, blood, and cerebrospinal fluid data are now needed to make artificial intelligence a precious tool capable of helping, supporting, and interacting with the medical staff decisions in the management of patients with TBI in ordinary settings and broader populations.

MF has received personal fees from Bayer, Biogen Idec, Merck-Serono, Novartis, Roche, Sanofi Genzyme, Takeda, and Teva; and MF's institution has received grants from Biogen Idec, Merck-Serono, Novartis, Roche, and Teva. FA has received personal fees from Biogen Idec, Novartis, and Philips. CC declares no competing interests.

*Massimo Filippi, Camilla Cividini, Federica Agosta
filippi.massimo@hsr.it

Neurology Unit (MF), Neurophysiology Unit (MF), and Neuroimaging Research Unit (MF, CC, FA), Institute of Experimental Neurology, Division of Neuroscience, IRCCS San Raffaele Scientific Institute, Milan, Italy; and Vita-Salute San Raffaele University, Milan, Italy (MF, CC, FA)

1 Ataloglou D, Dimou A, Zarpalas D, Daras P. Fast and precise hippocampus segmentation through deep convolutional neural network ensembles and transfer learning. *Neuroinformatics* 2019; **17:** 563–82.
2 Wu S, Li H, Quang D, Guan Y. Three-plane-assembled deep learning segmentation of gliomas. *Radiol Artif Intell* 2020; **2:** e190011.
3 Basaia S, Agosta F, Wagner L, et al. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *Neuroimage Clin* 2019; **21:** 101645.
4 Monteiro M, Newcombe VFJ, Mathieu F, et al. Multiclass semantic segmentation and quantification of traumatic brain injury lesions on head CT using deep learning: an algorithm development and multicentre validation study. *Lancet Digital Health* 2020; published online May 14. https://doi.org/10.1016/S2589-7500(20)30085-6.
5 Maas AIR, Menon DK, Adelson PD, et al. Traumatic brain injury: integrated approaches to improve prevention, clinical care, and research. *Lancet Neurol* 2017; **16:** 987–1048.
6 Mutch CA, Talbott JF, Gean A. Imaging evaluation of acute traumatic brain injury. *Neurosurg Clin N Am* 2016; **27:** 409–39.
7 Bardera A, Boada I, Feixas M, et al. Semi-automated method for brain hematoma and edema quantification using computed tomography. *Comput Med Imaging Graph* 2009; **33:** 304–11.
8 Farzaneh N, Reza Soroushmehr SM, Williamson CA, et al. Automated subdural hematoma segmentation for traumatic brain injured (TBI) patients. *Conf Proc IEEE Eng Med Biol Soc* 2017; **2017:** 3069–72.
9 Jain S, Vyvere TV, Terzopoulos V, et al. Automatic quantification of computed tomography features in acute traumatic brain injury. *J Neurotrauma* 2019; **36:** 1794–803.
10 Chilamkurthy S, Ghosh R, Tanamala S, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *Lancet* 2018; **392:** 2388–96.