Editorial: The use of deep learning in mapping and diagnosis of cancers

Abhishek Mahajan[1]* and Nivedita Chakrabarty[2]

[1]Department of Radiology, The Clatterbridge Cancer Liverpool, Liverpool, United Kingdom,
[2]Department of Radiodiagnosis, Tata Memorial Hospital, Homi Bhabha National Institute (HBNI),
Mumbai, India

Editorial on the Research Topic
The use of deep learning in mapping and diagnosis of cancers

Deep Learning (DL) is a subset and an augmented version of Machine Learning (ML), which in turn is a subgroup of Artificial Intelligence (AI), that uses layers of neural networks, similar to human brain, for performing complex tasks quickly and accurately. AI can recognize patterns in a large volume of data and extract characteristics imperceptible to the human eye (1). Convolutional Neural Network (CNN) is the most commonly used network of DL, which contains multiple layers, with weighted connections between neurons that are trained iteratively to improve performance. DL can be supervised or unsupervised, but most of the practical uses of DL in cancer has been with supervised learning where labelled images are used for data training (2). Despite the growing number of uses of DL in cancer mapping and diagnosis, there are uncharted territories in DL which remain to be explored to utilize it to its full capacity. Also, in spite of the revolution in cancer research that DL has ushered in, there are a lot of challenges to overcome, before DL can be widely used and accepted in every corner of the world.

Role of DL in oncology

There has been an unprecedented surge in DL based research in oncology due to the availability of big data, powerful hardware and robust algorithms. Screening and diagnosis of cancer, prediction of treatment response, and survival outcome and recurrence prediction, are the various roles of ML and DL in cancer management. AI algorithms integrated with clinical decision support (CDS) tools can automatically mine electronic health record (EHR) and identify cohort that would benefit maximum from

cancer screening programmes (3). For successful implementation of AI in cancer diagnosis, it is imperative for the radiologists and pathologists to collaborate with the key stakeholders, industrial partners and scientists (4). With ever increasing cancer burden worldwide, and availability of molecular targeted therapies, DL has served as an elixir, by its ability to screen, detect and diagnose tumours rapidly, and predict biomarkers non-invasively on imaging (5). Studies have shown that DL can be used to stage and grade tumours quickly and provide non-invasive histopathological diagnosis in cases where obtaining an invasive sample is risky. Patients, clinicians, radiologists and the pathologists, all have the potential to be benefitted by this DL technology as the utility of DL is no longer limited to tumour diagnosis, but to the cancer care as a whole. Prediction of overall survival, progression free survival, and disease free survival, assessment of response to treatment and outcome prediction are few of the many ways DL can benefit patients afflicted with cancer, the mere thought of which was previously unfathomable (5). Treatment planning and patient management can be hastened through the wider applications of DL based image interpretation, for example, non-responders to treatment detected on DL based baseline image interpretation, can be spared of further invasive treatment, and a change in management strategy may be considered for them.

Major applied uses of DL technology

Image classification and regression

DL can be used for classifying a lesion into benign or malignant, for treatment response evaluation and survival prediction. If DL models can be trained using a large dataset from a source domain, then it can be used in a target domain with a small sample size (2).

Object detection

DL can be used in tumour localization.

Semantic segmentation

DL can mark specific areas of concern on an image and assist the radiologists in decision making (2).

Image registration

Images acquired at different times can be accurately linked using DL, thus, enabling the radiologists to compare the images (2).

Federated learning

Robust deployable model can be built notwithstanding geographic boundaries, if multiple organizations/institutions/hospitals jointly train a model on a large data after de-identification of patient information (6).

Systematic review and meta-analysis data

A systematic review and meta-analysis from 1st January 2012 to 6th June 2019, comparing the diagnostic accuracy of health-care professionals with deep learning algorithms using imaging, found 10 studies on breast cancer, 9 studies on skin cancer, 7 studies on lung cancer, 5 studies on gastroenterological or hepatological cancers, 4 studies on thyroid cancer, 2 studies on oral cancer, and 1 study on nasopharyngeal cancer (7). Another systematic review on AI techniques in cancer diagnosis and prediction from articles published from 2009 to April 2021, revealed 10 articles pertaining to brain tumours, 13 articles related to breast cancer, 8 articles each related to cervical, liver, lung, and skin cancers, 6 articles related to colorectal cancer, 5 articles each related to renal and thyroid cancers, 2 articles each related to oral and prostate cancers, 7 articles related to stomach cancer, and 1 article each related to neuroendocrine tumours and lymph node metastasis (8). Few studies involving AI in cancer diagnosis and management include:

a. Histology prediction and screening of breast cancer on mammography (9, 10).
b. Brain tumour segmentation (11–14).
c. Lung nodule segmentation on computed tomography (CT) (15–17).
d. Liver tumour segmentation on CT (17, 18).
e. Prostate gland tumour detection on magnetic resonance imaging (MRI) (19, 20).
f. Brain tumour survival prediction (21–23).
g. F. Glioblastoma recurrence prediction (24).

Challenges and limitations of DL

a. Requirement of a large data: DL models need a large data (in thousands) to be trained and availability of such a huge data may not be possible in every institution.
b. Precise data annotation: Tumour region needs to be annotated or labelled accurately without contamination from surrounding non-tumour regions. This may not always be possible as many a times, tumours are

c. There is need for equal representation of data on training and test sets failing which data gets skewed and bias is introduced (2).

d. Heterogeneity of data: Difference in training set of images and deployable image sets may affect the performance of a model, for example if the CT scanner used while acquiring images for training is different from the one on which the model is validated, then performance may be reduced.

e. Patient privacy concerns: Despite the available methods for deidentification of patient information, the problems of patient privacy still loom large (2).

f. Problem of hidden layers: DL uses multiple layers of neural network to analyse data, which remain hidden, and the exact reasoning of outcome is not decipherable, which makes it difficult to be relied upon and convincingly used.

g. Infrastructure: Use of DL requires a robust infrastructure which may not be available everywhere.

h. Lack of trained personnel and expertise and lack of awareness about collaboration for implementation of AI projects (25).

Imaging biobanks

Repositories of human tissue sample stored in an organized manner for research purpose is known as "biobank", and collection of medical image data for long term storage and retrieval for research is known as "imaging biobank" (26, 27) Digital Imaging and Communications in Medicine (DICOM) is the universal format for Picture Archiving and Communication System (PACS) storage and data sharing across all institutions (26). The data needs to be de-identified and informed consent of the patient obtained prior to data archiving (28). Few examples of such open-source platforms include The Cancer Genome Atlas (TCGA) program, The Cancer Imaging Archive (TCIA), and European Genome–phenome Archive (EGA) (29, 30). In India, collaboration between the Department of Biotechnology (Government of India) under the guidance of the National Institution for Transforming India (NITI) Aayog, and Tata Memorial Centre has led to the creation of The Tata Memorial Center Imaging Biobank (31). World's biggest multi-modality imaging study was commenced by the UK Biobank in 2014 to have a repository of neuro, cardiac, and abdominal MRI imaging, dual energy x-ray absorptiometry (DEXA) and carotid ultrasonography (32). Similarly, CAN-I-AID (Cancer Imaging Artificial Intelligence Database) biobank project has been initiated by Dr. Abhishek Mahajan at the Clatterbridge Cancer Centre, Liverpool, United Kingdom (UK). Such imaging biobanks for public use should be encouraged as it fulfils the requirement of large image data to promote DL based research across the globe.

Articles in research topic

In this Research Topic, we present 20 topics, 19 of which are original articles and one is a systematic review. *There is one article on cervical cancer screening:* Sun et al. used Stacking-Integrated Machine Learning Algorithm based on demographic, behavioural, and clinical factors to accurately identify women at high risk of developing cervical cancer and suggested the use of this model to personalise cervical cancer screening programme. *Three articles on lung cancer:* Shen et al. showed that DL based CT images have the potential to accurately predict malignancy and invasiveness of pulmonary subsolid nodules on CT Images and thus aid in management decisions. Sun et al. conducted a study to establish the role of Convolutional Neural Network-Based Diagnostic Model to differentiate between benign and malignant lesions manifesting as a solid, indeterminate solitary pulmonary nodule (SPN) or mass (SPM) on computed tomography (CT). Xia et al. compared and fused DL and Radiomics features of ground-glass nodules to predict the invasiveness risk of stage-I lung adenocarcinomas in CT scan and concluded that fusion of DL and radiomics features can refine the classification performance for differentiating non-invasive adenocarcinoma (non-IA) from IA and the prediction of invasiveness risk of GGNs is similar to or better than radiologists using AI scheme. *One article on thyroid cancer:* Wu et al. combined ACR TI-RADS with DL by training three commonly used deep learning algorithms to differentiate between benign and malignant in TR4 and TR5 thyroid nodules with available pathology and concluded that irrespective of the type of TI-RADS used for the classification competition, DL algorithms outperformed radiologists. *One article on bladder cancer:* Zhang et al. proposed a DL model based on CT images to predict muscle-invasive status of bladder carcinoma pre-operatively and concluded that DL model exhibited relatively good prediction ability with capability to enhance individual treatment of bladder carcinoma. *One article on periampullary region:* Tang et al. used DL to identify periampullary regions on MRI images and achieved optimal accuracies in the segmentation of the peri-ampullary regions on both T1 and T2 MRI images concordant with manual human assessment. *One article on rectal cancer:* Zhang et al. segmented rectal cancer *via* 3D V-Net on T2WI and DWI and then compared the radiomics performance in predicting KRAS/NRAS/BRAF status between DL-based auto segmentation and manual-based segmentation. They concluded that 3D V-Net architecture could conduct reliable rectal cancer segmentation on T2WI and DWI images. *One article on jaw lesions:* Chai et al. showed that AI-based

cone-beam CT can distinguish between Ameloblastoma and Odontogenic Keratocyst with better accuracy than the surgeons. *Two articles on spine:* Ouyang et al. evaluated the efficiency of DL-based automated detection of primary spine tumours on MRI using the turing test. Hallinan et al. developed a DL model for classifying metastatic epidural spinal cord compression on MRI and which had comparable agreement to a subspecialist radiologist and clinical specialists. *One article on kidney tumour:* Sun et al. conducted a study on kidney tumour segmentation based on FR2PAttU-Net model. *One article on brain tumour:* Kandalgaonkar et al. conducted a study predicting IDH subtype of Grade 4 Astrocytoma and Glioblastoma from tumour radiomic patterns extracted from Multiparametric MRI using a machine learning approach and inferred that it may be used in either escalating or de-escalating adjuvant therapy for gliomas or for using targeted agents in future. *One article on survival rate prediction in cancer patients:* Sinzinger et al. developed Spherical Convolutional Neural Networks for survival rate prediction in cancer patients and concluded that it is beneficial in cases where expert annotations are not available or difficult to obtain. *One systematic review and meta-analysis:* Guha et al. performed a systematic review and meta-analysis differentiating primary central nervous system lymphoma (PCNSL) from glioblastoma (GBM) using deep learning and radiomics based ML approach. *There are five non-imaging related articles:* Zhu et al. developed transparent machine learning pipeline to efficiently predict Microsatellite instability (MSI), thus, helping pathologists to guide management decisions. Wang et al. conducted a study to reveal the heterogeneity in the tumor microenvironment of pancreatic cancer and analyze the differences in prognosis and immunotherapy responses of distinct immune subtypes. Menon et al. explored the histological similarities across cancers from a deep learning perspective. Huang et al. studied the effects of biofilm nano-composite drugs OMVs-MSN-5-FU on cervical lymph node metastases from oral squamous cell carcinoma (OSCC) on the animal model. Zormpas-Petridis et al. prepared a DL pipeline for mapping tumour heterogeneity on low-resolution whole-slide digital histopathology images. Figure 1 shows the list of authors based on type of articles submitted towards Research Topic.

Conclusions

DL has ushered in revolution in the field of oncology research, from cancer screening and diagnosis, to response assessment and survival prediction, thus positively influencing patient management. With the increasing cancer burden and limited number of specialized healthcare providers, there is a growing inclination to use DL at various levels of cancer diagnosis to cater to the needs of patients and the healthcare providers alike. Despite the umpteen benefits, there are a few challenges that DL needs to conquer, before it can be

FIGURE 1
List of authors based on type of articles submitted towards research topic.
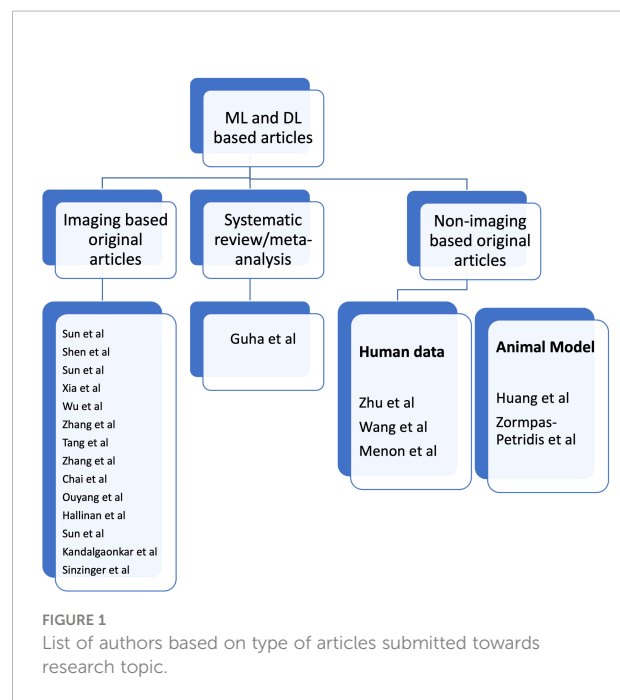
ubiquitously used. Through this Research Topic, we wish to acquaint the readers with the latest ongoing DL based research in cancer diagnosis, which can pave the way for further innovations and research in this field, as full potential of DL is still underutilized.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Choy G, Khalilzadeh O, Michalski M, Do S, Samir AE, Pianykh OS, et al. Current applications and future impact of machine learning in radiology. *Radiology* (2018) 288(2):318–28. doi: 10.1148/radiol.2018171820

2. Cherian Kurian N, Sethi A, Reddy Konduru A, Mahajan A, Rane SUA. 2021 update on cancer image analytics with deep learning. *WIREs Data Min Knowl Discov* (2021) 11:e1410. doi: 10.1002/widm.1410

3. Bizzo BC, Almeida RR, Michalski MH, Alkasab TK. Artificial intelligence and clinical decision support for radiologists and referring providers. *J Am Coll Radiol* (2019) 16(9 Pt B):1351–6. doi: 10.1016/j.jacr.2019.06.010

4. Tang A, Tam R, Cadrin-Chênevert A, Guest W, Chong J, Barfett J, et al. Canadian association of radiologists (CAR) artificial intelligence working group. canadian association of radiologists white paper on artificial intelligence in radiology. *Can Assoc Radiol J* (2018) 69(2):120–35. doi: 10.1016/j.carj.2018.02.002

5. Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med* (2021) 13(1):152. doi: 10.1186/s13073-021-00968-x

6. Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. *NPJ Digit Med* (2020) 3(1):119. doi: 10.1038/s41746-020-00323-1

7. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *Lancet Digit Health* (2019) 1(6):e271–97. doi: 10.1016/S2589-7500(19)30123-2

8. Kumar Y, Gupta S, Singla R, Hu Y-C. A systematic review of artificial intelligence techniques in cancer prediction and diagnosis. *Arch Comput Methods Eng* (2022) 29(4):2043–70. doi: 10.1007/s11831-021-09648-w

9. Sapate S, Talbar S, Mahajan A, Sable N, Desai S, Thakur M. Breast cancer diagnosis using abnormalities on ipsilateral views of digital mammograms. *Biocybernetics Biomed Eng* (2020) 40(1):290–305. doi: 10.1016/j.bbe.2019.04.008

10. Sapate SG, Mahajan A, Talbar SN, Sable N, Desai S, Thakur M. Radiomics based detection and characterization of suspicious lesions on full field digital mammograms. *Comput Methods Programs Biomed* (2018) 163:1–20. doi: 10.1016/j.cmpb.2018.05.017

11. Baid U, Ghodasara S, Mohan S, Bilello M, Calabrese E, Colak E, et al. The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. *arXiv preprint* (2021). arXiv:2107.02314.

12. Baid U, Talbar S, Rane S, Gupta S, Thakur MH, Moiyadi A, et al. A novel approach for fully automatic intra-tumor segmentation with 3D U-net architecture for gliomas. *Front Comput Neurosci* (2020) 14:10. doi: 10.3389/fncom.2020.00010

13. Mehta R, Filos A, Baid U, Sako C, McKinley R, Rebsamen M, et al. QU-BraTS: MICCAI BraTS 2020 challenge on quantifying uncertainty in brain tumor segmentation–analysis of ranking metrics and benchmarking results. *arXiv e-prints* (2021).

14. Pati S, Baid U, Zenk M, Edwards B, Sheller M, Reina GA, et al. The federated tumor segmentation (fets) challenge. *arXiv preprint* (2021).

15. Singadkar G, Mahajan A, Thakur M, Talbar S. Deep deconvolutional residual network based automatic lung nodule segmentation. *J Digit Imaging* (2020) 33(3):678–684. doi: 10.1007/s10278-019-00301-4

16. Singadkar G, Mahajan A, Thakur M, Talbar S. Automatic lung segmentation for the inclusion of juxtapleural nodules and pulmonary vessels using curvature based border correction. *J King Saud University-Computer Inf Sci* (2021) 33 (8):975–87. doi: 10.1016/j.jksuci.2018.07.005

17. Kumar YR, Muthukrishnan NM, Mahajan A, Priyanka P, Padmavathi G, Nethra M, et al. Statistical parameter-based automatic liver tumor segmentation from abdominal CT scans: A potential radiomic signature. *Proc Comput Science* (2016) 93:446–52. doi: 10.1016/j.procs.2016.07.232

18. Rela M, Krishnaveni BV, Kumar P, Lakshminarayana G. Computerized segmentation of liver tumor using integrated fuzzy level set method. *AIP Conf Proc* (2021) 2358(1):60001. doi: 10.1063/5.0057980

19. Hambarde P, Talbar SN, Sable N, Mahajan A, Chavan SS, Thakur M. Radiomics for peripheral zone and intra-prostatic urethra segmentation in MR imaging. *Biomed Signal Process Control* (2019) 51:19–29. doi: 10.1016/j.bspc.2019.01.024

20. Bothra M, Mahajan A. Mining artificial intelligence in oncology: Tata memorial hospital journey. *Cancer Res Stat Treat* (2020) 3:622–4. doi: 10.4103/CRST.CRST_59_20

21. Davatzikos C, Barnholtz-Sloan JS, Bakas S, Colen R, Mahajan A, Quintero CB, et al. AI-Based prognostic imaging biomarkers for precision neuro-oncology: theReSPOND consortium. *Neuro-oncology* (2020) 22(6):886–8. doi: 10.1093/neuonc/noaa045

22. Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, et al. MahIdentifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. *arXiv preprint* (2018).

23. Baid U, Rane SU, Talbar S, Gupta S, Thakur MH, Moiyadi A, et al. Overall survival prediction in glioblastoma with radiomic features using machine learning. *Front Comput Neurosci* (2020) 14:61. doi: 10.3389/fncom.2020.00061

24. Akbari H, Mohan S, Garcia JA, Kazerooni AF, Sako C, Bakas S, et al. Prediction of glioblastoma cellular infiltration and recurrence using machine learning and multi-parametric mri analysis: Results from the multi-institutional respond consortium. *Neuro-Oncology* (2021) 23(Supplement_6):vi132–3. doi: 10.1093/neuonc/noab196.522

25. Mahajan A, Vaidya T, Gupta A, Rane S, Gupta S. Artificial intelligence in healthcare in developing nations: The beginning of a transformative journey. *Cancer Research Statistics Treat* (2019) 2(2):182. doi: 10.4103/CRST.CRST_50_19

26. Mantarro A, Scalise P, Neri E. Imaging biobanks, big data, and population-based imaging biomarkers. In: *Imaging biomarkers: Development and clinical integration* (2017). Switzerland:Springer International Publishing. p. 153–7.

27. Woodbridge M, Fagiolo G, O'Regan DP, et al. MRIdb: Medical image management for biobank research. *J Digit Imaging* (2013) 26(5):886–90. doi: 10.1007/s10278-013-9604-9

28. Available at: https://car.ca/news/new-car-white-paper-on-ai-provides-guidance-on-de-identification-of-medical-imaging-data/ (Accessed on 17/10/2022).

29. Geis JR, Brady A, Wu CC, Spencer J, Ranschaert E, Jaremko JL, et al. Ethics of artificial intelligence in radiology: Summary of the joint European and north American multisociety statement. *Insights Into Imaging* (2019) 293(2):436–440. doi: 10.1148/radiol.2019191586

30. Available at: https://www.cancerimagingarchive.net/ (Accessed on 17/10/2022).

31. Bothra M, Mahajan A. Mining artificial intelligence in oncology: Tata memorial hospital journey. *Cancer Res Stat Treat* (2020) 3:622–4. doi: 10.4103/CRST.CRST_59_20

32. Littlejohns TJ, Holliday J, Gibson LM, Garratt S, Oesingmann N, Alfaro-Almagro F, et al. The UK biobank imaging enhancement of 100,000 participants: Rationale, data collection, management and future directions. *Nat Commun* (2020) 11(1):2624. doi: 10.1038/s41467-020-15948-9