Invited Perspective

Deep learning convolutional neural network (DLCNN): unleashing the potential of $^{18}$F-FDG PET/CT in lymphoma

Ke Li$^{1,2,3}$, Ran Zhang$^{1}$, Weibo Cai$^{1,2,3}$

$^1$Department of Medical Physics, University of Wisconsin-Madison, 1111 Highland Avenue, Madison, WI, USA; $^2$Department of Radiology, University of Wisconsin-Madison, 600 Highland Avenue, Madison, WI, USA; $^3$University of Wisconsin Carbone Cancer Center, 600 Highland Avenue, Madison, WI, USA

Received June 3, 2021; Accepted June 7, 2021; Epub August 15, 2021; Published August 30, 2021

Abstract: This perspective briefly reviewed the applications of $^{18}$F-FDG PET/CT in the clinical management of lymphoma and the need for lesion segmentation in those applications. It discussed the limitations of existing segmentation technologies and the great potential of using deep learning convolutional neural network (DLCNN) to accomplish automatic lymphoma segmentation and characterizations. Finally, the authors shared perspectives on the technical challenges that need be addressed to fully unleash the potential of DLCNN and $^{18}$F-FDG PET/CT in the diagnosis, prognosis, and treatment of lymphoma.

Keywords: PET/CT, $^{18}$F-FDG, lymphoma, deep learning, artificial intelligence (AI), convolutional neural network (CNN)

Introduction

$^{18}$F-FDG PET/CT is a hybrid imaging technique that brings the benefits of both functional and anatomical imaging to the diagnosis, prognosis, and treatment response assessment of lymphoma [1-3] and other diseases [4-10]. Prior studies have demonstrated clear advantages of FDG PET/CT over contrast-enhanced CT to detect disease in small (sub-centimeter) or normal-sized involved nodes, in unaltered bones, and at extranodal sites [11, 12]. FDG PET/CT may also detect more occult lymphomatous sites than bone marrow biopsy and CT [13]. By providing a more complete depiction of disease sites, FDG PET/CT leads to more accurate staging and restaging [14-16]. Literature has also shown the potential of using FDG PET/CT to obviate the invasive bone marrow biopsy in patients with newly diagnosed diffuse large B-cell lymphoma [17]. For the interim treatment response assessment of lymphoma patients, FDG PET/CT indicates whether there is still some metabolically active disease to help physicians determine whether to escalate the chemo intensity or introduce radiotherapy [18, 19]. In the post-therapy setting, FDG PET/CT is strongly encouraged in patients with Hodgkin lymphoma and diffused large B-cell non-Hodgkin lymphoma for which a complete treatment response is required to achieve a curative outcome [20-22]. FDG PET/CT has also been used the prediction of the progression-free survival and the overall survival [23], identification of the most suitable biopsy sites [24, 25], radiotherapy treatment planning [26] and other applications [4, 5].

At the core of PET/CT image postprocessing for lymphoma evaluations is the segmentation of malignant foci throughout the body. Once segmented, the tumors can be characterized using quantitative or semi-quantitative metrics such as the total metabolic tumor volume (TMTV), the maximum standardized uptake value ($SUV_{\text{max}}$), and the total lesion glycolysis (TLG). The classical segmentation approach replies on an experienced reader to manually detect and contour each focal tumor in PET/CT images. However, manual segmentation is time consuming and prone to inter-reader variabilities. In the past decade, a series of meth-
Deep learning-based PET/CT image segmentations

Methods have been developed for automated segmentation of lymphomas, although so far no consensus has been reached regarding the best method. The simplest segmentation approach is to use a fixed SUV value (e.g., 40% of SUV$_{max}$) as the threshold to differentiate abnormal and normal tissues [27, 28]. Such threshold-based methods can be quickly calculated and easily implemented. However, they lack the needed flexibility for discriminating lesions from tissues with normal physiological FDG uptake or normal FDG excretion. Alternative to fixed-value thresholding is adaptive thresholding that utilizes signal processing techniques such as region growing and Bayesian statistics [29-31]. Despite the improved flexibility, there are still some disadvantages of these adaptive thresholding methods such as high computational burden and poor sensitivity to small or heterogeneous tumors [32]. Consequently, there remains an unmet clinical need to develop a fast and robust method to automatically segment lymphoma lesions in FDG PET/CT images.

Motivated by the recent success of convolutional neural networks (CNNs) in other image segmentation applications, researchers began their endeavors to employ CNNs for lymphoma segmentation in FDG PET/CT images. In the work by Zhou et al. published in this issue of the AJNMMI [33], the popular U-Net CNN architecture [34] was adapted to segment mantle cell lymphoma (MCL) in 18F-FDG PET/CT images. To jointly utilize the complementary anatomical and functional information from PET/CT images, the authors constructed two symmetric convolutional encoder channels to extract feature representations from PET and CT images separately. At the end of the network, the PET and CT feature maps were combined along the decoder pathway to generate the final classification mask. To train and test the network, the authors retrospectively collected 142 sets (110 internal, 32 external) of baseline 18F-FDG PET/CT images of MCL patients. The network training solely used the internal dataset that was partitioned into 64% training, 16% validation, and 20% testing. In addition, the authors used a technique of image random affine transformations to further augment the training data [35]. The performance of the network was evaluated for each test patient in terms of true positive (TP) and false positive (FP) MCL lesions, where the true lesion contours were manually established by 3 physicians. The results showed that for the internal image data, the network generated a median sensitivity of 88% (IQR: 25%) and 15 (IQR: 10) FP lesions per patient; for scans from external institutions, the network generated a median sensitivity of 84% (IQR: 24%) and 14 (IQR: 10) FP lesions per patient. The authors also evaluated how well the network can differentiate organs with physiologic FDG activities (i.e., brain, heart, liver, kidneys, and bladder) from true MCL lesions: the specificity of the network for those organs with physiologic FDG activities were found to be above 90% except for the kidneys (77%). Sub-group analysis of the results further showed that the sensitivity of the network is higher for lesions with larger sizes or a higher SUV$_{max}$. No statistically significant dependence on the lesion location was observed.

Besides MCL segmentation, CNNs have found potential applications in the segmentation and characterization of other types of lymphomas such as diffuse large B-cell lymphoma [36-40], follicular lymphoma [39], and Hodgkin lymphoma [40]. CNNs were also employed to automate the discrimination of lymphoma lesions from organs with normal physiological FDG uptakes [41]. Improved accuracy over other state-of-the-art methods were reported in multiple studies [36, 41] and the segmentation time has been reduced to only a few seconds by CNNs [33]. Despite the promising results reported in literature, additional research is needed before CNNs can be reliably deployed in clinical practice. For example:

- The sensitivities of the networks to lesions with lower SUVs are relatively low (e.g., at 62% in [33]). Similarly, the sensitivities to smaller lesions are lower than to larger lesions (e.g., 71% for sub-cm MCL lesions vs. 84% for MCL lesions greater than 1 cm [33]). Consequently, existing networks may overlook indolent/early-stage lesions. Some studies discarded lesions with volumes below 2 ml despite these lesions’ clinical values.

- While some works used relatively large and multi-institution/multi-vendor patient cohorts to train and test their CNNs [37-39], other studies used small datasets (e.g., <100 scans) that were collected from a single institution and even with a single PET/CT scanner model and scan protocol. This prevents any conclu-
Deep learning-based PET/CT image segmentations

In addition to the use of larger and more external datasets, the rigor of the network evaluation method can be further strengthened in the following aspects: Most studies published so far lack comparisons (in terms of both accuracy and computation speed) with the classical thresholding method and state-of-the-art segmentation methods to justify the advantages of CNNs. The way the ground truth for the lesion contour was established varies across studies: some use qualified nuclear medicine physicians while others used existing semi-automated segmentation methods. To enable a direct and fair comparison of different studies, consensus needs to be reached on the ground truth establishment method.

Efforts also need to be spent on how to optimally utilize the complementary information from PET and CT images. The work in [33] used two network branches to extract features from PET and CT separately before combining their features at the final output stage. It is yet unclear how much benefit it brings compared to a simple PET-only approach or using concatenated PET and CT images as network inputs. How to use multi-modality data is an interesting and important topic for not only lymphoma segmentation but also the general deep learning field.

Looking beyond the automatic extraction of explicit image features such as lesion size and location, we see opportunities for using CNNs to extract more advanced information about lymphoma from PET/CT image data. As shown in [44], CNNs hold promise of directly predicting progression-free survival for individual lymphoma patients. With continued development and optimization of the network algorithm as well as more extensive network trainings and evaluations, we believe the full potential of CNN-based automatic PET/CT image post-processing can be unleashed so that it can be reliably used in the clinical care of lymphoma.

Acknowledgements

The authors are grateful for financial support from the University of Wisconsin-Madison.

Disclosure of conflict of interest

None.

Address correspondence to: Ke Li, Department of Medical Physics, University of Wisconsin-Madison, 1111 Highland Avenue, Madison, WI, USA. Tel: 01-608-262-4566; E-mail: ke.li@wisc.edu

References


Deep learning-based PET/CT image segmentations


