Am J Nucl Med Mol Imaging 2021;11(4):327-331
www.ajnmmi.us /ISSN:2160-8407/ajnmmi0135990

*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 to 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 ($SUV_{max}$), 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-

ods have been developed for automated segmentation of lymphomas, although so far no consensus has been reached yet 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 $^{18}$F-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 $^{18}$F-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: 12) 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-

sive evaluation about the network's robustness and generalizability. As pointed out in a recent article [42], researchers are recommended to fulfill the checklist for artificial intelligence in medical imaging (CLAIM) [43] in order to minimize the risk of bias and avoid systematic methodological flaws. Independent external validations and robustness or sensitivity analysis, proper considerations of biological variables such as sex and ethnicity are among the CLAIM criteria and should be implemented in future research on this topic.

• 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

[1] Kanoun S, Rossi C and Casasnovas O. [18F] FDG-PET/CT in Hodgkin lymphoma: current usefulness and perspectives. Cancers 2018; 10: 145.

[2] Trotman J and Barrington SF. The role of PET in first-line treatment of Hodgkin lymphoma. Lancet Haematol 2021; 8: e67-e79.

[3] Seam P, Juweid ME and Cheson BD. The role of FDG-PET scans in patients with lymphoma. Blood 2007; 110: 3507-16.

[4] Borra A, Morbelli S, Zwarthoed C, Bianchi A, Bergesio F, Chauvie S, Zaucha JM, Taszner M, Malkowski B, Biggi A, Thyss A, Darcourt J and Gallamini A. Dual-point FDG-PET/CT for treatment response assessment in Hodgkin lymphoma, when an FDG-avid lesion persists after treatment. Am J Nucl Med Mol Imaging 2019; 9: 176-84.

[5] Paul DM, Ghiuzeli CM, Rini J, Palestro CJ, Fung EK, Ghali M, Ben-Levi E, Prideaux A, Vallabhajosula S and Popa EC. A pilot study treatment of malignant tumors using low-dose (18)F-fluorodeoxyglucose ((18)F-FDG). Am J Nucl Med Mol Imaging 2020; 10: 334-41.

[6] Gelezhe PB, Blokhin IA, Marapov DI and Morozov SP. Quantitative parameters of MRI and (18)F-FDG PET/CT in the prediction of breast cancer prognosis and molecular type: an original study. Am J Nucl Med Mol Imaging 2020; 10: 279-92.

[7] Zirakchian Zadeh M, Raynor WY, Østergaard B, Hess S, Yellanki DP, Ayubcha C, Mehdizadeh Seraj S, Acosta-Montenegro O, Borja AJ, Gerke O, Werner TJ, Zhuang H, Revheim ME, Abildgaard N, Høilund-Carlsen PF and Alavi A. Correlation of whole-bone marrow dual-timepoint (18)F-FDG, as measured by a CT-based method of PET/CT quantification, with respon-

se to treatment in newly diagnosed multiple myeloma patients. Am J Nucl Med Mol Imaging 2020; 10: 257-64.
[8] Al-Zaghal A, Aras M, Borja AJ, Moghbel M, Demir Y, Houshmand S, Ciftci E, Werner TJ, Høilund-Carlsen PF, Torigian DA, Han Y and Alavi A. Detection of pulmonary artery atherosclerosis by FDG-PET/CT: a new observation. Am J Nucl Med Mol Imaging 2020; 10: 127-34.
[9] Baffour FI, Wenger DE and Broski SM. (18)F-FDG PET/CT imaging features of lipomatous tumors. Am J Nucl Med Mol Imaging 2020; 10: 74-82.
[10] Borja AJ, Hancin EC, Dreyfuss AD, Zhang V, Mathew T, Rojulpote C, Werner TJ, Patil S, Gonuguntla K, Lin A, Feigenberg SJ, Swisher-McClure S, Alavi A and Revheim ME. (18)F-FDG-PET/CT in the quantification of photon radiation therapy-induced vasculitis. Am J Nucl Med Mol Imaging 2020; 10: 66-73.
[11] Paes FM, Kalkanis DG, Sideras PA and Serafini AN. FDG PET/CT of extranodal involvement in non-Hodgkin lymphoma and hodgkin disease. Radiographics 2010; 30: 269-91.
[12] Kostakoglu L and Goldsmith SJ. Fluorine-18 fluorodeoxyglucose positron emission tomography in the staging and follow-up of lymphoma: is it time to shift gears? Eur J Nucl Med 2000; 27: 1564-78.
[13] Moog F, Bangerter M, Kotzerke J, Guhlmann A, Frickhofen N and Reske SN. 18-F-fluorodeoxyglucose-positron emission tomography as a new approach to detect lymphomatous bone marrow. J Clin Oncol 1998; 16: 603-9.
[14] Rini JN, Leonidas JC, Tomas MB and Palestro CJ. 18F-FDG PET versus CT for evaluating the spleen during initial staging of lymphoma. J Nucl Med 2003; 44: 1072-4.
[15] Freudenberg LS, Antoch G, Schutt P, Beyer T, Jentzen W, Muller SP, Görges R, Nowrousian MR, Bockisch A and Debatin JF. FDG-PET/CT in re-staging of patients with lymphoma. Eur J Nucl Med Mol Imaging 2004; 31: 325-9.
[16] Isasi CR, Lu P and Blaufox MD. A metaanalysis of 18F-2-deoxy-2-fluoro-D-glucose positron emission tomography in the staging and restaging of patients with lymphoma. Cancer 2005; 104: 1066-74.
[17] Berthet L, Cochet A, Kanoun S, Berriolo-Riedinger A, Humbert O, Toubeau M, Dygai-Cochet I, Legouge C, Casasnovas O and Brunotte F. In newly diagnosed diffuse large B-cell lymphoma, determination of bone marrow involvement with 18F-FDG PET/CT provides better diagnostic performance and prognostic stratification than does biopsy. J Nucl Med 2013; 54: 1244.
[18] Radford J, Illidge T, Counsell N, Hancock B, Pettengell R, Johnson P, Wimperis J, Culligan D, Popova B, Smith P, McMillan A, Brownell A, Kruger A, Lister A, Hoskin P, O'Doherty M and Barrington S. Results of a trial of PET-directed therapy for early-stage Hodgkin's lymphoma. N Engl J Med 2015; 372: 1598-607.
[19] Johnson P, Federico M, Kirkwood A, Fosså A, Berkahn L, Carella A, d'Amore F, Enblad G, Franceschetto A, Fulham M, Luminari S, O'Doherty M, Patrick P, Roberts T, Sidra G, Stevens L, Smith P, Trotman J, Viney Z, Radford J and Barrington S. Adapted treatment guided by interim PET-CT scan in advanced Hodgkin's lymphoma. N Engl J Med 2016; 374: 2419-29.
[20] Isohashi K, Tatsumi M, Kato H, Fukushima K, Maeda T, Watabe T, Shimosegawa E, Kanakura Y and Hatazawa J. Prognostic value of FDG-PET, based on the revised response criteria, in patients with malignant lymphoma: a comparison with CT/MRI evaluations, based on the International Working Group/Cotswolds Meeting Criteria. Asia Ocean J Nucl Med Biol 2015; 3: 91-8.
[21] Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, Coiffier B, Fisher RI, Hagenbeek A, Zucca E, Rosen ST, Stroobants S, Lister TA, Hoppe RT, Dreyling M, Tobinai K, Vose JM, Connors JM, Federico M and Diehl V; International Harmonization Project on Lymphoma. Revised response criteria for malignant lymphoma. J Clin Oncol 2007; 25: 579-86.
[22] Juweid ME, Stroobants S, Hoekstra OS, Mottaghy FM, Dietlein M, Guermazi A, Wiseman GA, Kostakoglu L, Scheidhauer K, Buck A, Naumann R, Spaepen K, Hicks RJ, Weber WA, Reske SN, Schwaiger M, Schwartz LH, Zijlstra JM, Siegel BA and Cheson BD; Imaging Subcommittee of International Harmonization Project in Lymphoma. Use of positron emission tomography for response assessment of lymphoma: consensus of the Imaging Subcommittee of International Harmonization Project in Lymphoma. J Clin Oncol 2007; 25: 571-8.
[23] Aldin A, Umlauff L, Estcourt LJ, Collins G, Moons KG, Engert A, Kobe C, von Tresckow B, Haque M, Foroutan F, Kreuzberger N, Trivella M and Skoetz N. Interim PET-results for prognosis in adults with Hodgkin lymphoma: a systematic review and meta-analysis of prognostic factor studies. Cochrane Database Syst Rev 2019; 16: CD012643.
[24] Wu MH, Xiao LF, Liu HW, Yang ZQ, Liang XX, Chen Y, Lei J and Deng ZM. PET/CT-guided versus CT-guided percutaneous core biopsies in the diagnosis of bone tumors and tumor-like lesions: which is the better choice? Cancer Imaging 2019; 19: 69.
[25] Fei B and Schuster DM. PET molecular imaging-directed biopsy: a review. AJR Am J Roentgenol 2017; 209: 255-69.

[26] Mikhaeel NG, Milgrom SA, Terezakis S, Berthelsen AK, Hodgson D, Eich HT, Dieckmann K, Qi SN, Yahalom J and Specht L. The optimal use of imaging in radiation therapy for lymphoma: guidelines from the International Lymphoma Radiation Oncology Group (ILROG). Int J Radiat Oncol Biol Phys 2019; 104: 501-12.

[27] Hellwig D, Graeter TP, Ukena D, Groeschel A, Sybrecht GW, Schaefers HJ and Kirsch CM. 18F-FDG PET for mediastinal staging of lung cancer: which SUV threshold makes sense? J Nucl Med 2007; 48: 1761.

[28] Nestle U, Schaefer-Schuler A, Kremp S, Groeschel A, Hellwig D, Rübe C and Kirsch CM. Target volume definition for 18F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer. Eur J Nucl Med Mol Imaging 2007; 34: 453-62.

[29] Desbordes P, Petitjean C and Ruan S. 3D automated lymphoma segmentation in PET images based on cellular automata. 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA); 2014. pp. 1-6.

[30] Grenier T, Revol-Muller C, Costes N, Janier M and Gimenez G. 3D Robust Adaptive Region Growing for segmenting [18F] fluoride ion PET images. 2006 IEEE Nuclear Science Symposium Conference Record; 2006. pp. 2644-8.

[31] Hatt M, Rest CCl, Turzo A, Roux C and Visvikis D. A fuzzy locally adaptive bayesian segmentation approach for volume determination in PET. IEEE Trans Med Imaging 2009; 28: 881-93.

[32] Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rübe C and Kirsch CM. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. J Nucl Med 2005; 46: 1342.

[33] Zhou Z, Lu PJY, Macapinlac H, Wang ML, Son JB, Pagel MD, et al. Computer-aided detection of mantle cell lymphoma on 18F-FDG PET/CT using deep learning convolutional neural networks. Am J Nucl Med Mol Imaging 2021.

[34] Ronneberger O, Fischer P and Brox T. U-net: convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI): Springer; 2015. pp. 234-41.

[35] Mikołajczyk A and Grochowski M. Data augmentation for improving deep learning in image classification problem. 2018 International Interdisciplinary PhD Workshop (IIPhDW); 2018. pp. 117-22.

[36] Yuan C, Zhang M, Huang X, Xie W, Lin X, Zhao W, Li B and Qian D. Diffuse large B-cell lymphoma segmentation in PET-CT images via hybrid learning for feature fusion. Med Phys 2021; [Epub ahead of print].

[37] Capobianco N, Meignan M, Cottereau A-S, Vercellino L, Sibille L, Spottiswoode B, Zuehlsdorff S, Casasnovas O, Thieblemont C and Buvat I. Deep-learning 18F-FDG uptake classification enables total metabolic tumor volume estimation in diffuse large B-cell lymphoma. J Nucl Med 2021; 62: 30.

[38] Blanc-Durand P, Jégou S, Kanoun S, Berriolo-Riedinger A, Bodet-Milin C, Kraeber-Bodéré F, Carlier T, Le Gouill S, Casasnovas RO, Meignan M and Itti E. Fully automatic segmentation of diffuse large B cell lymphoma lesions on 3D FDG-PET/CT for total metabolic tumour volume prediction using a convolutional neural network. Eur J Nucl Med Mol Imaging 2021; 48: 1362-70.

[39] Jemaa S, Fredrickson J, Carano RAD, Nielsen T, de Crespigny A and Bengtsson T. Tumor segmentation and feature extraction from whole-body FDG-PET/CT using cascaded 2D and 3D convolutional neural networks. J Digit Imaging 2020; 33: 888-94.

[40] Weisman AJ, Kieler MW, Perlman SB, Hutchings M, Jeraj R, Kostakoglu L and Bradshaw TJ. Convolutional neural networks for automated PET/CT detection of diseased lymph node burden in patients with lymphoma. Radiol Artif Intell 2020; 2: e200016.

[41] Bi L, Kim J, Kumar A, Wen L, Feng D and Fulham M. Automatic detection and classification of regions of FDG uptake in whole-body PET-CT lymphoma studies. Comput Med Imaging Graph 2017; 60: 3-10.

[42] Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nature Machine Intelligence 2021; 3: 199-217.

[43] Mongan J, Moy L and Kahn CE Jr. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers. Radiol Artif Intell 2020; 2: e200029.

[44] Guo R, Hu X, Song H, Xu P, Xu H, Rominger A, Lin X, Menze B, Li B and Shi K. Weakly supervised deep learning for determining the prognostic value of 18F-FDG PET/CT in extranodal natural killer/T cell lymphoma, nasal type. Eur J Nucl Med Mol Imaging 2021: 1-11.