Application of Deep Learning as a Noninvasive Tool to Determine Pathological Diagnosis of Enlarged Cervical Lymph Nodes with PET/CT

Yuhan Yang
West China School of Medicine: Sichuan University West China Hospital
Bo Zheng
West China School of Medicine: Sichuan University West China Hospital
Yixi Wang
West China School of Medicine: Sichuan University West China Hospital
Xuelei Ma (drmaxuelei@gmail.com)
State Key Laboratory of Biotherapy and Cancer Center,West China Hospital,Sichuan University,Chengdu,PR China https://orcid.org/0000-0002-6853-1348

Original research

Keywords: Deep learning, Convolutional neural network, Malignant lymphoma, Lymph node metastases, Computed tomography, Positron emission tomography/computed tomography.

Posted Date: September 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-75481/v1

Abstract

Objective: To construct a deep-learning convolution neural network (DL-CNN) system for pathological diagnosis of cervical lymph nodes by using computed tomography (CT), fluorine-18 fluorodeoxyglucose (^{18}F-FDG) positron emission tomography (PET), and fused PET/CT images.

Materials and methods: A total of 1020 cross-sectional images for each imaging modality was obtained from 211 patients (153 patients with lymphomas and 116 patients with metastases) with enlarged cervical lymph nodes from January 2014 to June 2018. All eligible images were distributed randomly into the training, validation, and testing cohorts with ratios of 70%, 15%, and 15%. We applied eight DL-CNN algorithms with pretrained bases from ImageNet dataset on CT, PET, and fused PET/CT imaging datasets to differentiate lymphomatous nodes from metastatic nodes, respectively. Attention heatmaps of PET and fused PET/CT images generated by class activation mapping (CAM) were used in visualization of class specific regions recognized by the prediction model with best performance.

Results: The accuracy of eight deep learning algorithms with pretrained base ranged from 0.650 to 0.981 on PET testing cohort, and from 0.738 to 0.981 on fused PET/CT testing cohort. The VGG16 model on PET images and DenseNet121 model on fused PET/CT images had the best diagnostic performance among all eight algorithms with sensitivity and specificity of 1.000 and 0.963. Class-specific discriminative subregions were highlighted by attention maps for clinical review.

Conclusion: A DL-CNN system was developed for classifying metastatic and lymphomatous involvement with favorable diagnostic performance on PET and PET/CT images in patients with enlarged cervical lymph nodes. The further clinical practice of this system may improve quality of the following therapeutic interventions and optimize patients' outcomes.

Introduction

The enlargement of cervical lymph nodes has existed in many diseases such benign lymphadenitis, nodal metastases from head and neck cancer or other solitary cancer, and malignant lymphoma. Clinicians have faced problems in distinguishing histological compositions of enlarged nodes by physical examinations. With assistance of routine imaging examination such as computed tomography (CT), magnetic resonance imaging (MRI), and ultrasonography (US), the difficulties have been eliminated in diagnosis of benign and malignant nodes. However, it is still challenging to differentiate whether malignant lymphomas invade cervical lymph nodes involvement or solid tumors metastasize to cervical lymph nodes [1] due to similar radiological presences in size, contents, margin, and distribution. Further, some invasive assessment tools have been applied to increase diagnostic accuracy like fine needle aspiration cytology (FNAC) with sonography. Nevertheless, the success of FNAC depends on the physicians' experience and location of suspicious lymph nodes. Otherwise, lymph node swelling interferes the proceedings of puncture with low-quality and inadequate tissue sampling. We expected some easy-to-reach and noninvasive examination tools could occur to determine histological diagnosis of lymphadenopathy or metastases with satisfactory performance.

Positron emission tomography (PET) /CT, as an advanced imaging modality was applied in diagnosis and staging of various malignancies using fluorine-18 fluorodeoxyglucose (^{18}F-FDG)[2–4], where PET provided functional activity of lymph node status and CT provided high-resolution anatomical localization[5]. Facing with distinctive standard uptake values (SUVs) of abnormal nodes, manual interpretation of PET/CT images was considered to be accurate[6] which could identify nodal status with respect to involvement of lymphoma or metastases of solid malignancies that some highly malignant solid tumors indeed reacted strongly active lymph nodes. However, pitfalls exist that the diagnostic accuracy takes radiologists' experience into account when judging the status of moderately active lymph nodes. Thus, development of an automatic prediction model for lymph node identification of PET/CT images could assist histological diagnosis of enlarged cervical lymph nodes which would recognize the pathological components of targeted nodes whether lymphomas nor metastases of solid tumors from multi-objective imaging patterns.

Computer-aided diagnosis (CAD) consists of two major strategies: handcrafted feature-based and feature learning-based strategies[7]. Handcrafted feature-based strategies mainly focused on extracting a large number of quantitative features from original images manually and constructing prediction models to solve various problems[8–10] with shallow machine learning methods. Nowadays, there has been a shift to feature learning-base strategies which could recognize images directly and represent accurate and consistent interpretations of imaging features with deep learning techniques, especially multilayered convolutional neural networks (CNN). With this method, there have been various clinical tasks achieved including tumor staging, treatment outcome prediction, and survival analysis[11–13]. The performance and reliability of deep learning-CNN (DL-CNN) could be improved as the training datasets of algorithms expanded[14, 15]. To our knowledge, there have been no studies to predict histological components of enlarged cervical lymph nodes on neither PET images nor combinations of PET and CT images.

In this study, we aimed to determine the utility of deep learning-based model in predicting histological diagnosis of enlarged cervical lymph nodes by analyzing features from CT, PET, and PET/CT images.

Materials And Methods

Patients

This retrospective study was approved by the Ethics Committee of the Institutional Review Board. We retrieved 3060 cross-sectional images including three modalities: plain CT (1020 slices), PET (1020 slices), and PET/CT (slices) images respectively. Each modality contained 510 slices attributing to lymphomas and 510 slices attributing to solid tumors from 211 patients (153 patients with lymphomas and 116 patients with solid tumors) (Table.1) undergoing ^{18}F-FDG-PET/CT scanning before clinical interventions. Each image only contained one enlarged cervical lymph node. All eligible plain CT, PET, and PET/CT images were randomly distributed into training (350 lymphomas and 350 solid tumors) (70%), validation (80 lymphomas and 80 solid tumors) (15%), and test (80

lymphomas and 80 solid tumors) (15%) cohorts. All eligible patients received PET/CT examinations before individual therapy between January 2014 and June 2018. For patients with malignant lymphomas, the diagnoses were completed by FNAC or biopsy on peripheral lymph nodes. The patients were diagnosed as lymphomas with positive histological and immunohistochemical evidences on examined nodes despite original locations. In some cases, there was no positive observation on samples of cervical nodes. As for patients with solid tumors, all eligible subjects were confirmed histologically by primary lesions. Only few patients received cervical lymph node biopsy of FNAC. We gave preference to solid tumor metastases when facing enlarged cervical lymph node with evidence of primary tumor.

We performed standard whole-body ^{18}F-FDG PET/CT using a Gemini GXL PET/CT scanner (Philips, Amsterdam, Netherlands). The blood glucose level was evaluated immediately before the administration of ^{18}F-FDG that PET/CT scanning would be completed till the blood glucose level less than 150 mg/dL. Fasting for at least 6 hours was required before the examination. Approximately 5 MBq of ^{18}F-FDG per kilogram (up to 550 MBq) of body weight was administered intravenously, and patients received low-dose CT scanning (40 mA, 120 kVp) after resting in a quiet, dark environment for approximately 60 minutes. After initial low-dose CT, emission images were obtained from the top of the skull to the middle of the thigh, with acquisition times of 2 minutes per bed position in the 3D mode. PET images were reconstructed iteratively with CT based attenuation correction. In this study, we concentrated on patients' cervical region for the proof of concept (Figure.1).

Image preprocessing

All CT, PET, and fused PET/CT images were reviewed by one experienced radiologist and one nuclear medicine physician who depicted the axial regions of enlarged cervical lymph node lesions with rectangular masks which could enter the ultimate DL-CNN, respectively. In order to avoid disturbing noises from surrounding organ tissue and blood vessels, the area around the bladder cancer lesions were regulated up to 15% of all masked regions. Moreover, to control the balance of images recognition, all CT imaged were transformed into final masking images under stable CT window [200~300 Hounsfiled Unit (HU)] and window level (20~30 HU) which were augmented into 150 × 150 pixels to standardize the distance scale and avoid distortion simultaneously. Through injected activity and body weight to SUV, all PET images were intensity normalized for injecting activity and body weight to SUVs. The real-time data augmentation was applied to the deep learning models which were exposed to multi-aspects of existing data by random transformation from limited samples in the training phase under dynamic control to improve model generalization and decrease the degree of overfitting.

Model development and evaluation

The development of DL-CNN contained three sections: CT, PET, and fused PET/CT images. The models were developed on different types of images by extracting salient features in axial scans, respectively. To optimize DL-CNN diagnostic performance, eight algorithms were constructed to classify histological components of cervical lymph nodes on fundaments of pretrained models on the ImageNet database including VGG16, Xception, VGG19, InceptionV3, InceptionResNetV2, DenseNet121, DenseNet169, and DenseNet201[16-22]. These networks ended with a layer_dense of size 1 with a sigmoid activation because of this binary classification problem. The RMSprop optimizer was applied with binary crossentropy as the loss in the compilation step. In addition, generated activation maps by class activation mapping (CAM) for the prediction model with best performance among eight algorithms on the test dataset were applied to evaluate the region of interest of PET and PET/CT images for further clinical review. The process of CAM was completed by linking 2D grid of an input image with a specific output class to compute every location in any input image, and indicating how important each location is with respect to the class under consideration. In this study, we performed Grad-CAM to interpret DL-CNN procedures consisting steps of taking the output feature map of a convolution layer, given an input image, and weighing every channel in that feature map by the gradient of the class with respect to the channel[23]. This workflow scheme was presented concisely in Fig. 1.

A total of twenty-four DL-CNN (eight algorithms for three kinds of images) were trained in the same training sample through deep learning package Keras (http://keras.io/). During the fitting process, the batch size was set to 30 and the epochs was 15~20 times until when no further obvious fluctuation of loss function was observed after tuning the models in the validation cohort (Figure.2). The performances of the eight algorithms were measured by the accuracy and loss function of the training and validation sets, and the accuracy, loss function, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the test set. The receiver operating characteristic curve (ROC) plots were completed by pROC package35 (https://cran.r-project.org/web/packages/pROC). The whole processes of DL-CNN including development, validation, testing and visualization were performed by R statistical software 3.6.1 (https://www.r-project.org/).

Results

Diagnostic performance

The results of diagnostic performance on eight deep learning algorithms were present in the Table. 2. The accuracy of the tested algorithms ranged from 0.550 to 0.650 for CT images, from 0.650 to 0.981 for PET images, and from 0.738 to 0.981 for fused PET/CT images. The AUROC (area under the receiver operating characteristic curve) fluctuated from 0.512 to 0.684 for CT images, from 0.628 to 0.995 for PET images, and from 0.795 to 0.998 for fused PET/CT images. Of eight CT models, the DenseNet121 algorithm achieved better performance with AUROC of 0.684 compared with other algorithms on the testing dataset. For the PET models, VGG16 and DenseNet169 algorithms were the best-performing models for the testing dataset with AUROC of 0.995 and 0.990, respectively. As for prediction models on fused PET/CT images, DenseNet121 and VGG16 algorithms were considered as models with satisfactory predictive power with AUROC of 0.998 and 0.991, respectively. The ROC (receiver operating characteristic) curves of eight pretrained algorithms and their associated AUROC were shown in Figure.2. Eight models based on fused PET/CT images represented stable fluctuation between different algorithms. The DenseNet121 and VGG16 algorithms achieved top three prediction performances no matter kinds of images considering AUROC on the testing dataset. The sensitivity, specificity, PPV, and NPV of DenseNet121 algorithm on the testing cohort were 0.750, 0.550, 0.625, and 0.688 on CT images, 0.975, 0.975, 0.975, and 0.975 on

PET images, and 1.000, 0.963, 0.964, and 1.000 on fused PET/CT images, respectively. The sensitivity, specificity, PPV, and NPV of VGG16 algorithm on the testing cohort were 0.913, 0.375, 0.593, and 0.811 on CT images, 1.000, 0.963, 0.964, and 1.000 on PET images, and 1.000, 0.900, 0.909, and 1.000 on fused PET/CT images, respectively. Meanwhile, the diagnostic performance of the testing groups was not found significant correlations with the parameters and depth with p value > 0.05 for all eight algorithms.

Class activation mapping development

Due to satisfactory performance of VGG16 algorithm, we took this algorithm as the pretrained base for CAM. The attention heatmap of the lymphoma involved and metastatic lymph nodes was generated by CAM and then superimposed on the original PET and PET/CT image so that the location of the actual lymph node and the region highlighted by the model could be compared. The attention heatmaps based on CT images were omitted because of unfavorable prediction power. As shown in Fig. 3, the attention heatmap (region in red) emphasized important subregions for further clinical review. This represented that the salient subregions of enlarged lymph nodes recognized by CNN had been translated as abnormal characteristics of metastases and applied into classification of cervical nodes for proof of histological diagnosis.

Discussion

This study focused on determining histological diagnosis of enlarged cervical lymph nodes through ^{18}F-FDG PET/CT scanning by deep learning-based CAD systems. With the highest accuracy of 0.888 in the validation cohort, the VGG16 model could classify cervical lymph nodes into lymphadenopathy and metastases on its accuracy of 0.950 in the testing cohort on fused PET/CT images. Considering PET images, the accuracy of VGG16 model could reach 0.981 based on the testing dataset. According such high accuracy of DL-CNN models, we may consider their practicability in identifying the histological compositions when detecting enlarged cervical lymph nodes.

Ultrasonography has been considered as a conventional and useful imaging tool in diagnosis of enlarged cervical lymph nodes or masses including benign lymph nodes, lymphadenopathy, metastasis of head and neck cancer[24–26]. Under review of sonograph images, lymphomatous nodes always present as round, hypoechoic nodes without echogenic hilus and tends to show intranodal reticulation with sharp margin[27, 28]. As for metastatic lymph nodes, reactive presence of hilar and peripheral vascularity with nodal necrosis has been considered as valuable signs for diagnosis with high specificity[29]. However, due to relatively high rates of hilar vascularity (28.5%)[30] and comparable reactive signs (10%)[31] in lymphomatous nodes, an increasing risk of misdiagnosis has existed in differentiation of metastatic lymph nodes. Therefore, additional imaging examination has been recommended to provide more information to determine some subtypes of malignant lymphomas difficult to concern using sonography such as CT, MRI and ^{18}F-FDG PET[32]. In previous studies, supplementary CT scans assisted identifying lymphomatous lymph nodes with an accuracy of 87.5%[1]. Moreover, a meta-analysis demonstrated high specificity (86%) and sensitivity (85%) of ^{18}F-FDG PET in determination for nodal metastases[33]. In the present study, the prediction model with best performance on CT images reached an accuracy of 65.0% that the same prediction model on fused PET/CT images improved accuracy to 98.1% with assistance of ^{18}F-FDG PET scanning.

Deep learning serves as an important branch of artificial intelligence which mainly discovers intricate structures or relationships in existing datasets rather than programming[11]. With backpropagation algorithms for transformation of original data points, DL-CNN models have been expected to solve ultimate decision-making tasks such as risk stratification and diagnosis[34, 35]. In this study, we applied a relatively simple architectures with 8 pretrained bases to recognize and extract imaging features for final classification between lymphomatous and metastatic lymph nodes in the neck. The architecture we constructed implemented both conventional VGG net[21] and advanced DenseNet[13] to optimize prediction performance and develop CAD system on selected images. The results showed little improvement of DenseNet with AUROC of 0.998 compared with that of VGG net with AUROC of 0.991 using fused PET/CT images. Comparing with conventional diagnostic method through sampling, the diagnostic performance of CAD system might be more reliable and cost effective[36] that uncertainty of sampling procedure depended on the surgeons' experience no matter biopsy or FNAC. Therefore, we realized that the CAD system with proposed DL-CNN may act as a useful and noninvasive tool in pathological diagnosis of cervical lymph nodes.

To our knowledge, this is the first study to incorporate convolutional neural network model to identify metastatic and lymphomatous nodes of enlarged cervical lymph nodes on PET/CT images. The DL-CNN CAD systems may have some strengths over manual reviewing in recognition for compositions of cervical lymph nodes. First, direct recognition of selected images may lead to consistent and accurate results when eliminating the obstacle of inter-observer variability. Second, the attention heatmap drawn by CAM increased the interpretability which pointed out important area analyzed by tedious algorithms. This process indirectly revealed internal relationship between input data and output labels. Third, the rigorous setting of training, validation, and testing cohorts adjusted the scalability of the CAD system by incorporating misclassification individuals. There existed some limitations in our study. At first, this study did not contain auto-localization process to amplify 512 × 512 pixels images reconstructed by scanners into images of region of interests. The additional steps of lymph node region depiction and augmentation were warranted before data input into the DL-CNN models causing unavoidable observing errors. Secondly, this was a retrospective study based on a single center dataset. The next step is to develop prospective diagnostic models on multi-center datasets with external validation. Thirdly, the further transformation from our diagnostic model to formal clinical practice was uncertain. The robust optimization and integration of present models need thorough evaluation and technique improvement in the future.

In conclusion, we developed a deep learning-convolutional neural network system for classifying lymphomatous and metastatic involvement with favorable diagnostic performance on ^{18}F-FDG PET and fused PET/CT images in patients with enlarged cervical lymph nodes. The clinical practice of this system may improve quality of the following therapeutic interventions and optimize patients' outcomes.

Conclusion

A DL-CNN system was developed for classifying metastatic and lymphomatous involvement with favorable diagnostic performance on PET and PET/CT images in patients with enlarged cervical lymph nodes. The further clinical practice of this system may improve quality of the following therapeutic interventions and optimize patients' outcomes.

Abbreviations

DL-CNN
deep-learning convolution neural network; CT:computed tomography; ^{18}F-FDG:fluorine-18 fluorodeoxyglucose; PET:positron emission tomography; CAM:class activation mapping; MRI:magnetic resonance imaging; US:ultrasonography; FNAC:fine needle aspiration cytology; CAD:computer-aided diagnosis; PPV:positive predictive value; NPV:negative predictive value; ROC:receiver operating characteristic curve; AUROC:area under the receiver operating characteristic curve.

Declarations

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study was established, according to the ethical guidelines of the Helsinki Declaration and was approved by the Ethics Committee of the Institutional Review Board, West China hospital, Sichuan University. Written informed consent was not applicable as a retrospective cohort study in our country.

CONSENT FOR PUBLICATION

Written informed consent for publication was obtained from all participants.

AVAILABILITY OF DATA AND MATERIAL

Some data generated or used during the study are available from the corresponding author by request.

COMPETING INTERESTS

Not applicable

FUNDING

Not applicable

AUTHORS' CONTRIBUTIONS

XM, YY, BZ and YW made substantial contributions to conception and design, and revised the manuscript critically for important intellectual content. XM revised the manuscript and gave final approval of the version to be published. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

The authors declared that they had full access to all of the data in this study and the authors take complete responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Predictors of the necessity. for lymph node biopsy of cervical lymphadenopathy. Journal of Cranio-Maxillofacial Surgery. 2015;43(10):2200–4.
2. Adams S, Baum RP, Stuckensen T, Bitter K, Hör G. Prospective comparison of 18F-FDG PET with conventional imaging modalities (CT, MRI, US) in lymph node staging of head and neck cancer. Eur J Nucl Med. 1998;25:1255–60.
3. Braams J, Pruim J, Freling N, Nikkels P, Roodenburg JLN, Boering G, et al. Detection of Lymph Node Metastases of Squamous-Cell Cancer of the Head and Neck with FDG-PET and MRI. Journal of nuclear medicine: official publication. Society of Nuclear Medicine. 1995;36:211–6.
4. Laubenbacher C, Saumweber D, Wagner-Manslau C, Kau R, Herz M, Avril N, et al. Comparison of fluorine-18-fluorodeoxyglucose PET, MRI and endoscopy for staging head and neck squamous-cell carcinomas. Journal of nuclear medicine: official publication Society of Nuclear Medicine. 1995;36:1747–57.
5. Yoon DY, Hwang H, Chang S, Rho Y-S, Ahn H, Kim J, et al. CT, MR, US, 18F-FDG PET/CT, and their combined use for the assessment of cervical lymph node metastases in squamous cell carcinoma of the head and neck. Eur Radiol. 2008;19:634–42.
6. Schöder H, Yeung H, Gonen M, Kraus D, Larson S. Head and Neck Cancer: Clinical Usefulness and Accuracy of PET/CT Image Fusion1. Radiology. 2004;231:65–72.
7. Zhou Z, Chen L, Sher D, Zhang Q, Shah J, Pham N-L, et al. Predicting Lymph Node Metastasis in Head and Neck Cancer by Combining Many-objective Radiomics and 3-dimensioal Convolutional Neural Network through Evidential Reasoning*2018. 1–4 p.
8. Ganeshan B, Abaleke S, Young R, Chatwin C, Miles K. Texture analysis of non-small cell lung cancer on unenhanced computed tomography: Initial evidence for a relationship with tumour glucose metabolism and stage. Cancer imaging: the official publication of the International Cancer Imaging Society. 2010;10:137–43.

9. Mak, Raymond H, Lee, Stephanie W, et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. Radiotherapy Oncology Journal of the European Society for Therapeutic Radiology Oncology. 2016.
10. Huang Y-Q, Liang C-H, He L, Tian J, Liang C-S, Chen X, et al. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. Journal of Clinical Oncology. 2016;34.
11. LeCun Y, Bengio Y, Hinton G. Deep Learning Nature. 2015;521:436–44.
12. Yamashita R, Nishio M, Do R, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights into Imaging. 2018;9.
13. Essa I, Kang SB, Pollefeys M. Guest Editors' Introduction to the Special Section on Award-Winning Papers from the IEEE Conference on Computer Vision and Pattern Recognition 2009 (CVPR 2009). IEEE Trans Pattern Anal Mach Intell. 2011;33:2339-40.
14. Suzuki K. Overview of deep learning in medical imaging. Radiological Physics and Technology. 2017;10.
15. Sollini M, Antunovic L, Chiti A, Kirienko M. Towards clinical application of image mining: a systematic review on artificial intelligence and radiomics. European Journal of Nuclear Medicine and Molecular Imaging. 2019;46.
16. Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions. 2016.
17. Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. AAAI Conference on Artificial Intelligence. 2016.
18. Sun M, Farhadi A, Seitz S. Ranking Domain-Specific Highlights by Analyzing Edited Videos2014. 787–802 p.
19. Implementing Efficient ConvNet Descriptor Pyramids
DenseNet
Iandola F, Moskewicz M, Karayev S, Girshick R, Darrell T, Keutzer K. DenseNet: Implementing Efficient ConvNet Descriptor Pyramids. 2014.
20. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition2016. 770–8 p.
21. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 14091556. 2014.
22. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision. 2014;115.
23. Rs R, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. International Journal of Computer Vision. 2019;128.
24. Matsumoto F, Itoh S, Ohba S-i, Yokoi H, Furukawa M, Ikeda K. Biopsy of cervical lymph node. Auris nasus larynx. 2008;36:71–4.
25. Niedzielska G, Kotowski M, Niedzielski A, Dybiec E, Wieczorek P. Cervical lymphadenopathy in children–incidence and diagnostic management. Int J Pediatr Otorhinolaryngol. 2007;71(1):51–6.
26. Shin JH, Baek JH, Chung J, Ha EJ, Kim JH, Lee YH, et al. Ultrasonography Diagnosis and Imaging-Based Management of Thyroid Nodules: Revised Korean Society of Thyroid Radiology Consensus Statement and Recommendations. Korean J Radiol. 2016;17(3):370–95.
27. Ahuja AT, Ying M. Sonographic evaluation of cervical lymph nodes. AJR American journal of roentgenology. 2005;184(5):1691–9.
28. Gritzmann N, Hollerweger A, Macheiner P, Rettenbacher T. Sonography of soft tissue masses of the neck. Journal of clinical ultrasound: JCU. 2002;30(6):356–73.
29. Kaji AV, Mohuchy T, Swartz JD. Imaging of cervical lymphadenopathy. Seminars in ultrasound. CT MR. 1997;18(3):220–49.
30. Ying M, Ahuja A, Brook F. Accuracy of sonographic vascular features in differentiating different causes of cervical lymphadenopathy. Ultrasound Med Biol. 2004;30(4):441–7.
31. Orita Y, Nose S, Sato Y, Miki K, Domae S, Hirai M, et al. Cervical lymph node extirpation for the diagnosis of malignant lymphoma. Surgery today. 2012;43.
32. King AD, Ahuja AT, Yeung DK, Fong DK, Lee YY, Lei KI, et al. Malignant cervical lymphadenopathy: diagnostic accuracy of diffusion-weighted MR imaging. Radiology. 2007;245(3):806–13.
33. Kyzas PA, Evangelou E, Denaxa-Kyza D, Ioannidis JP. 18F-fluorodeoxyglucose positron emission tomography to evaluate cervical node metastases in patients with head and neck squamous cell carcinoma: a meta-analysis. J Natl Cancer Inst. 2008;100(10):712–20.
34. Ha EJ, Baek JH, Na DG. Risk Stratification of Thyroid Nodules on Ultrasonography: Current Status and Perspectives. Thyroid: official journal of the American Thyroid Association. 2017;27(12):1463–8.
35. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115–8.
36. Hehn ST, Grogan TM, Miller TP. Utility of fine-needle aspiration as a diagnostic technique in lymphoma. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2004;22(15):3046–52.

Tables

Table.1: A detailed list of patients with enlarged cervical lymph nodes attributing to malignant lymphomas and solid tumors in this study.

Malignant lymphoma (N=153)		Solid tumor (N=116)	
Subtype	n	Subtype	n
Burkitt lymphoma	2 patients	Nasopharyngeal carcinoma	20 patients
Hodgkin's lymphoma	40 patients	Lung cancer	10 patients
Follicular lymphoma	21 patients	Hepatocellular carcinoma	18 patients
Mantle cell lymphoma	10 patients	Laryngocarcinoma	4 patients
B immunoblastic lymphoma	2 patients	Esophageal carcinoma	13 patients
T immunoblastic lymphoma	5 patients	Gastric cancer	11 patients
Diffuse large B-cell lymphoma	73 patients	Breast cancer	2 patients
		Thyroid carcinoma	38 patients

Table.2: Diagnostic performance of eight deep learning algorithms for differentiation between lymphomatous and metastatic lymph nodes on CT, PET, and PET/CT images in training, testing, and validation cohorts.

Model	Modality	Training dataset		Validation dataset		Testing dataset					
		Accuracy	Loss	Accuracy	Loss	Accuracy	Loss	Sensitivity	Specificity	PPV	NPV
Xception	CT	0.631	0.659	0.581	0.720	0.550	0.692	0.725	0.375	0.537	0.577
	PET	0.799	0.447	0.631	0.710	0.788	0.587	0.788	0.788	0.788	0.788
	PET/CT	0.531	1.664	0.500	1.998	0.838	1.624	0.750	0.925	0.909	0.787
VGG16	CT	0.668	0.660	0.538	0.677	0.644	0.680	0.913	0.375	0.593	0.811
	PET	0.945	0.195	0.944	0.259	0.981	0.204	1.000	0.963	0.964	1.000
	PET/CT	0.934	0.275	0.888	0.356	0.950	0.268	1.000	0.900	0.909	1.000
VGG19	CT	0.562	0.668	0.538	0.679	0.588	0.693	0.613	0.563	0.583	0.592
	PET	0.945	0.202	0.875	0.329	0.956	0.251	0.950	0.963	0.962	0.951
	PET/CT	0.916	0.257	0.875	0.318	0.944	0.255	0.963	0.925	0.928	0.961
InceptionV3	CT	0.504	0.730	0.519	0.722	0.594	0.760	0.363	0.825	0.674	0.564
	PET	0.408	0.865	0.469	0.824	0.713	0.855	0.913	0.513	0.652	0.854
	PET/CT	0.665	0.645	0.588	0.759	0.806	0.623	0.888	0.725	0.763	0.881
InceptionResnetV2	CT	0.512	0.836	0.506	0.918	0.638	0.923	0.450	0.825	0.720	0.600
	PET	0.512	2.712	0.506	2.857	0.650	2.913	0.363	0.938	0.853	0.595
	PET/CT	0.515	1.621	0.500	1.885	0.738	1.845	0.538	0.938	0.896	0.670
DenseNet121	CT	0.520	0.716	0.488	0.768	0.650	0.722	0.750	0.550	0.625	0.688
	PET	0.509	0.999	0.506	1.107	0.975	1.118	0.975	0.975	0.975	0.975
	PET/CT	0.636	0.675	0.581	0.851	0.981	0.797	1.000	0.963	0.964	1.000
DenseNet169	CT	0.536	0.700	0.469	0.752	0.556	0.727	0.863	0.250	0.535	0.645
	PET	0.822	0.376	0.688	0.503	0.975	0.435	0.963	0.988	0.987	0.963
	PET/CT	0.724	0.554	0.656	0.680	0.938	0.720	0.975	0.900	0.907	0.973
DenseNet201	CT	0.508	0.737	0.525	0.758	0.563	0.742	0.875	0.250	0.538	0.667
	PET	0.901	0.248	0.881	0.273	0.956	0.193	0.963	0.950	0.951	0.962
	PET/CT	0.872	0.327	0.850	0.374	0.944	0.339	0.913	0.900	0.973	0.918

*Abbreviations: CT: computed tomography; PET: positron emission tomography; PPV: positive predictive value; NPV negative predictive value.

Figures

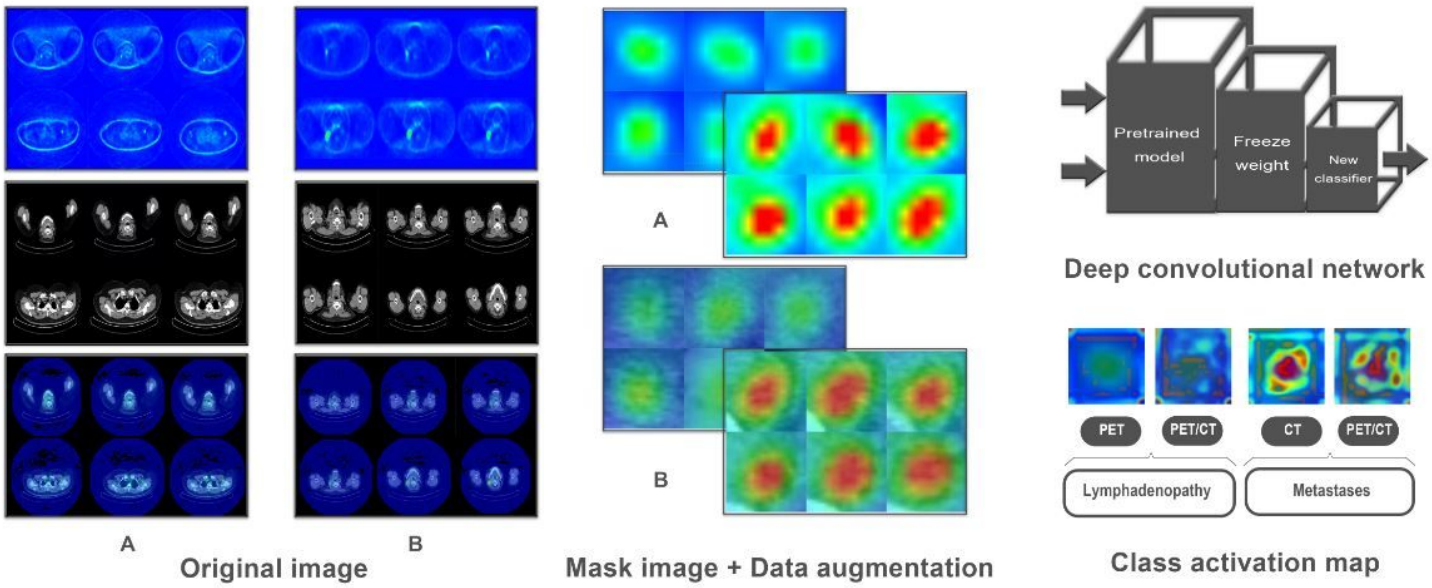

Figure 1

The flowchart of deep learning-convolutional neural networks on enlarged cervical lymph nodes on CT, PET, and fused PET/CT images. Original images, as inputs of neural networks, contained lymphomatous (A) and metastatic (B) nodes of PET, CT and fused PET/CT images (from top to bottom). Data augmentation was completed in both PET (A) and fused PET/CT (B) images. This scheme represented the workflow of data preprocessing, model development and validation by pretrained algorithms, and generation of visible attention heatmaps.

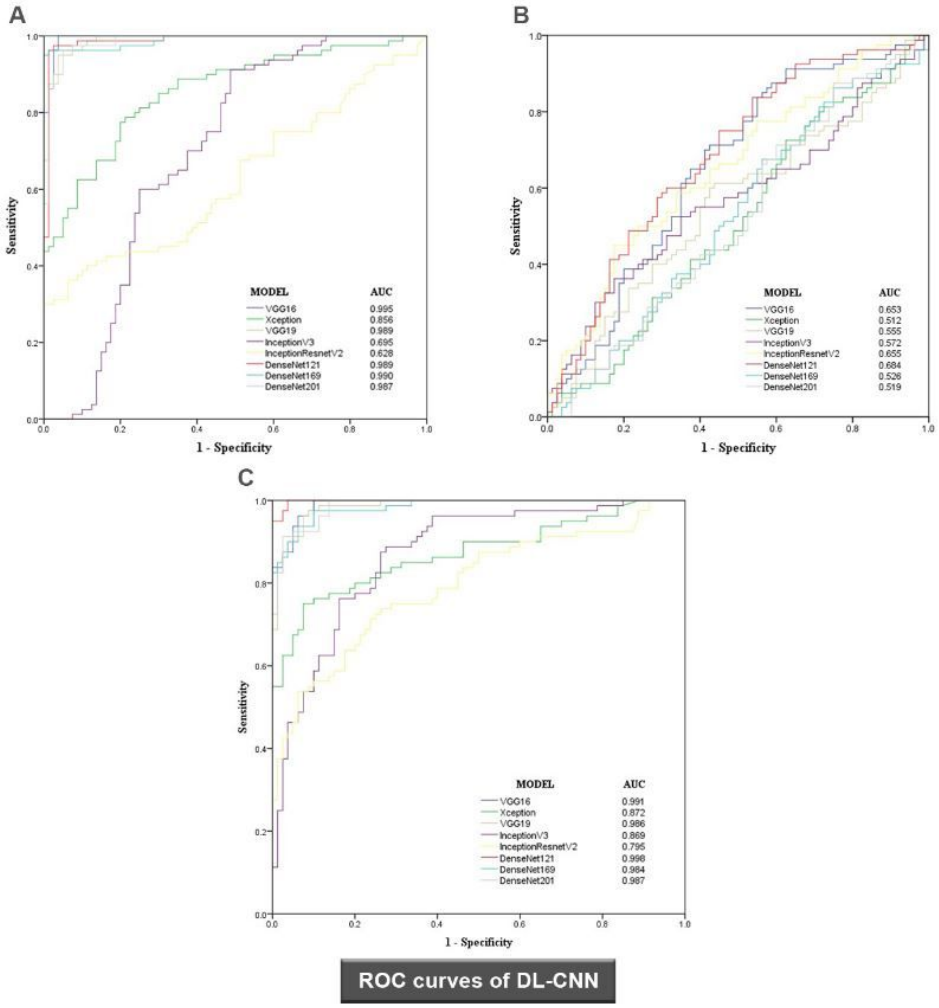

Figure 2

Receiver operating characteristic curves of eight deep learning algorithms with pretrained bases on PET (A), CT (B), and fused PET/CT (C) images of enlarged cervical lymph nodes in the testing dataset.

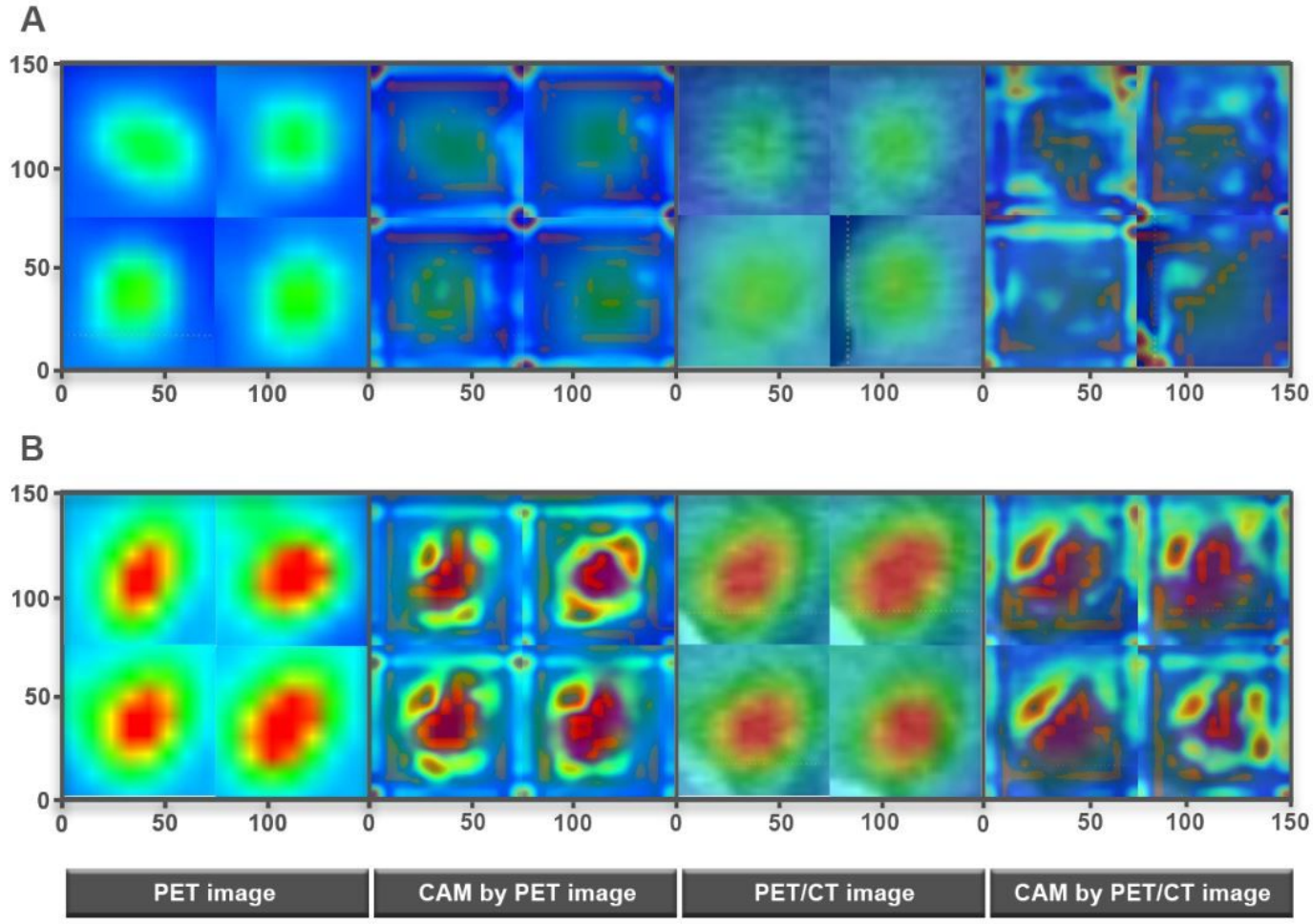

Figure 3

Attention heatmap depicted by class activation mapping. The red color highlighted the region of interest to classify lymphomatous and metastatic lymph nodes. The red color focuses on a diffusing area weakly on lymphomatous lymph nodes (A) and on nodal boundaries intensively on metastatic lymph nodes (B).