



# Deep-Learning-Based Molecular Imaging Biomarkers: Toward Data-Driven Theranostics

Hongyoon Choi

*Department of Nuclear Medicine, Seoul National University Hospital, Seoul, Korea*

**Received** 18 April 2019

**Revised** 11 May 2019

**Accepted** 11 May 2019

**Corresponding author**

Hongyoon Choi  
(chy1000@gmail.com)  
Tel: 82-2-2072-3347  
Fax: 82-2-745-7690

Deep learning has been applied to various medical data. In particular, current deep learning models exhibit remarkable performance at specific tasks, sometimes offering higher accuracy than that of experts for discriminating specific diseases from medical images. The current status of deep learning applications to molecular imaging can be divided into a few subtypes in terms of their purposes: differential diagnostic classification, enhancement of image acquisition, and image-based quantification. As functional and pathophysiologic information is key to molecular imaging, this review will emphasize the need for accurate biomarker acquisition by deep learning in molecular imaging. Furthermore, this review addresses practical issues that include clinical validation, data distribution, labeling issues, and harmonization to achieve clinically feasible deep learning models. Eventually, deep learning will enhance the role of theranostics, which aims at precision targeting of pathophysiology by maximizing molecular imaging functional information.

**Keywords:** Deep learning, Molecular imaging, Theranostics, Medical imaging, Imaging biomarker

## Introduction

Deep learning rapidly begins to be applied in the medical field. Recently, several deep learning-related medical devices and softwares have been developed and started to be applied in the clinical fields.<sup>1)</sup> The major contribution of deep learning to medical data was to objectively evaluate high-dimensional medical data and remarkably reduce laborious works such as segmentation and object detection from high-resolution images. The major medical application is medical imaging fields as a boom of deep learning was started from the computer vision field initiated by ImageNet Challenge.<sup>2,3)</sup> The methods and neural network architectures developed for ImageNet Challenge have been applied to medical images including radiologic and pathologic exams as well as natural photographic images. These

approaches based on computer vision fields have showed remarkable performance in differential diagnosis. For natural photographic images such as skin images and funduscopy deep learning techniques were relatively easily adopted as convolutional neural network (CNN) models developed for ImageNet Challenge were directly transferred to such images.<sup>4,5)</sup> Moreover, CNN which show good performance on image classification and processing have been applied to radiologic exams such as chest X-ray and mammography.<sup>6-8)</sup> Subsequently, CNN models have been used for image-based diagnosis as well as image processing.<sup>9)</sup> The application of deep learning included 3-dimensional images such as CT, PET and MRI data as well as 2-dimensional radiologic exams. The purpose of clinical use was also expanded to include various applications such as image-based differential diagnosis, segmentation, and image en-

hancement. Because of the substantial different features of molecular imaging including PET and SPECT from natural images, there have been various concerns with regard to application of deep learning. Nonetheless, various deep learning techniques have suggested feasible applications to enhance molecular imaging and solved problems such as image resolution and sensitivity.<sup>10)</sup> In this review, current deep learning models for nuclear medicine and molecular imaging are summarized according to the clinical purposes. In order to develop robust deep learning models and guide their appropriate direction for clinical use, practical issues of current deep learning are introduced in this review.

### Current Deep Learning Models for Molecular Imaging

Current deep learning models particularly for molecular imaging have focused on various different applications: Image-based diagnosis, enhancing image reconstruction and image quality, and deep learning application for image-based quantification (Table 1).

Intuitively, one of the most important applications of deep learning in medical fields was differential diagnosis. For molecular imaging studies, as deep learning models generally require a large dataset for the training, several models have used PET or SPECT images which routinely

acquired in the clinical setting. One of the major applications was differentiating disorders from normal status. Recently, using FDG PET images, a few deep CNN models for the differential diagnosis were suggested. For example, using FDG PET images, a deep learning model was developed to differentiate metastatic mediastinal lymph nodes from benign lymph nodes in lung cancer.<sup>11)</sup> Using a deep CNN, diagnostic accuracy for differentiating metastatic lymph nodes was 86%, which was higher than conventional machine learning algorithms.<sup>11)</sup> Another CNN model to differentiate T-stages from lung cancer showed comparable results to identify pathologic T-staging.<sup>12)</sup> Area of receiver-operating-characteristic curve (ROC) was 0.68 for differentiating advanced T-stage tumors in an independent test set. Deep CNN models have been developed for differential diagnosis of brain disorders using brain SPECT or PET images. As a binary classification problem, dopamine transporter imaging has been interpreted by experts' reading, thus, it was a good candidate for the deep CNN application. A 3-dimensional CNN model showed high accuracy for differentiating <sup>123</sup>I-FP-CIT SPECT images of Parkinson's disease from those of controls.<sup>19)</sup> As accurate image-based diagnosis and the prediction of future cognitive decline in Alzheimer's disease (AD) and mild cognitive impairment (MCI) patients have been clinically important issues, several deep learning models using MRI and PET have been suggested. One of the first research of deep learning application to medical images was representation learning for PET and MRI images for diagnosing AD.<sup>17,18)</sup> Though these pioneer studies did not use CNN, regarded as a *de facto* standard model in recent application, these models extract discriminative features automatically and showed higher performance for classifying brain images of AD compared with conventional algorithms. Recently developed models use deep CNN models for differentiating AD from controls, and showed high accuracy for the differentiation.<sup>13,45)</sup>

Another important application is enhancement of image reconstruction and image quality. For example, CNN models were incorporated into iterative reconstruction framework and showed better performance than conventional denoising algorithms.<sup>27)</sup> As a generalized approach, deep learning was used to solve the inverse function of signals encoded by sensors including MRI and PET with regard to

**Table 1.** Types of current deep learning applications for nuclear medicine and molecular imaging

Types of applications	Examples	References
Image-based diagnosis	Cancer staging (T- and N-staging)	11,12
	Diagnosis of Alzheimer's disease using PET and/or MRI	13-18
	Diagnosis of Parkinson's disease using dopamine transporter imaging	19-21
	Prediction of coronary heart disease	22-24
Enhancement of image reconstruction and image quality	Image reconstruction	25-29
	Attenuation correction	30-34
Image-based quantification	Recovery of low-dose PET images	35-37
	Segmentation	38-42
	Image generation for quantification	43,44

the image reconstruction, which resulted in fully-automated and flexible reconstruction framework.<sup>28)</sup> Furthermore, attenuation correction, a crucial step of PET image reconstruction, was aided by deep learning-based attenuation maps. While CT incorporated in fusion PET/CT scanners can provide attenuation information, recent PET/MR requires synthetic CT attenuation maps. Because of the difficulty in the estimation of attenuation map without CT, there have been various issues regarding PET quantification.<sup>46,47)</sup> Recently suggested deep learning-based CT image synthesis using MR or PET images is promising to solve the quantification issues caused by attenuation correction.<sup>30-34)</sup> Additionally, deep learning has been used to enhance image quality for low dose PET images.<sup>35-37)</sup> By combining the algorithms for image reconstruction with low-dose radiotracers and PET- or MR-based attenuation correction can dramatically reduce radiation exposure in the future. Such an ultra-low dose PET may be used for new clinical purposes including disease screening which has been difficult to obtain benefits due to radiation hazards.

As molecular imaging provides quantitative value related to pathophysiology, studies have focused on the application of deep learning to obtain accurate quantification. The most common application of deep learning to medical images is segmentation.<sup>9)</sup> The segmentation methods are usually based on anatomical images such as CT and MRI. As recent clinical molecular imaging modalities provides fusion images such as PET/CT, PET/MR, and SPECT/CT, deep learning-based segmentation methods can be used to calculate quantitative values such as the accumulation of radiotracer in a specific tissue delineated by anatomical imaging.<sup>39,48)</sup> The quantification can be improved by generative models such as generative adversarial networks (GAN). For example, pseudo-MR images were generated by AV-45 PET using GAN for the quantification of cortical radiotracer uptake without structural MR acquisition.<sup>43)</sup>

## Clinically Feasible Deep Learning-Based Biomarkers and Practical Issues

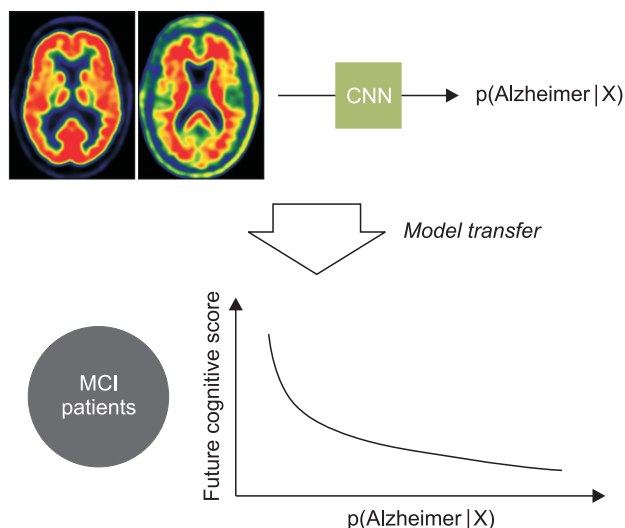
### Necessity of deep learning-based biomarker

Even though various deep learning techniques have ap-

plied to molecular imaging for differential diagnosis, image enhancement, and accurate quantification, there are many issues that need to be solved in order to be clinically used. One of the gaps between deep learning approaches for natural image recognition and medical images, particularly molecular imaging, is placed on the purpose of imaging. While the image recognition task has simple labels, clinicians often require various types of information from medical images. They include prediction of prognostic outcome and treatment response as well as differential diagnosis.<sup>10)</sup> In a narrower range, differential diagnosis is similar with labels of natural images; however, many diagnostic classifications are not simple classification. Because many disorders have a spectrum ranged from healthy to fully-blown disease status, ground-truth labels widely used in deep learning training are ambiguous in medical images. Furthermore, a gold standard of diagnostic classification is variable according to disease types as well as clinical situations.<sup>49)</sup> Thus, if we think more deeply, the eventual purpose of deep learning application to the medical field is not just for simple diagnosis, but for looking to play a critical role in clinical decision.<sup>50)</sup> As molecular imaging intrinsically provides molecular and pathophysiologic properties with noninvasive manner, deep learning algorithms should more emphasize on the acquisition of objective quantitative value which can predict future outcome and treatment response. Instead of the achievement of the state-of-the-art in classification accuracy, we should find appropriate clinical application of the output of deep learning. For example, a deep learning model was developed for discriminating Alzheimer's disease and normal aged subjects, however, the importance of the application of this model was to transfer to the MCI subjects who would rapidly progress to full-blown dementia.<sup>13)</sup> The output of the CNN model represents a probability of Alzheimer's disease in a cohort consisting of Alzheimer's disease and normal subjects. As the output of the CNN was estimated by patterns of FDG and amyloid deposit in the brain, these patterns could be associated with a predictive biomarker for the outcome of MCI subjects (Fig. 1).

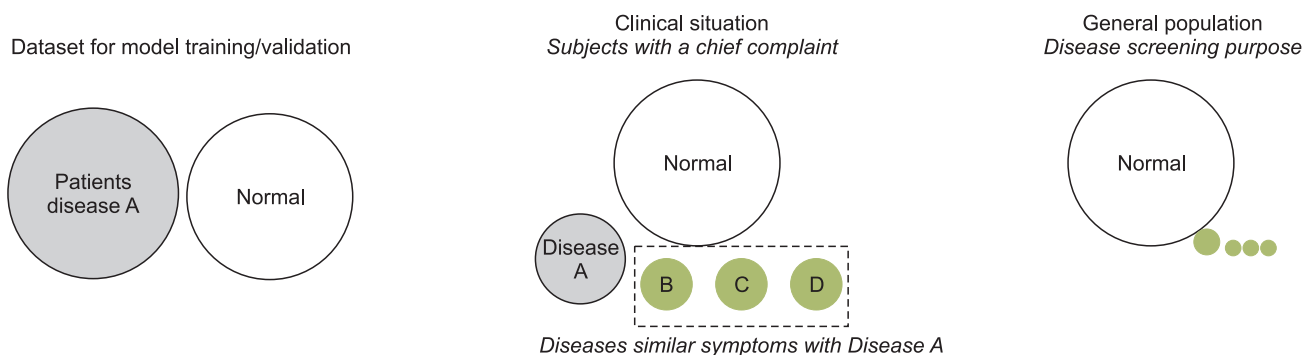
## Data distribution and validation

Even though many deep learning models show remarkable performance on the classification problem, such as discriminating funduscopy images or brain PET images, most models are not validated in the real-world clinical settings. It is related to the evaluation of the performance when a suggested deep learning model tries to be used in the clinical setting. To achieve this validation issue, deep



**Fig. 1.** The output of deep learning model as a predictive biomarker. A deep convolutional neural network (CNN) model was developed to differentiate brain PET of Alzheimer's disease from healthy subjects. This model was applied to another cohort, mild cognitive impairment patients to predict future cognitive outcome. The output of the model represents a probability of Alzheimer's disease, which can be used as a predictive biomarker for predicting cognitive outcome in preclinical disorders.

learning models should be tested in an independent test set from the training and internal validation data. The most commonly used method is the application to datasets obtained from different centers.<sup>51)</sup> Even though deep learning models are validated in an external dataset and show good performance on diagnostic classification or prediction for clinical outcome, they can hardly guarantee the same performance in the heterogeneous clinical environment. That is because the cohort used for the development of deep learning models are different from clinical trials, in which subjects are recruited with specific criteria defined for a clinical setting.<sup>52)</sup> The problem is placed on the fact that patients in the clinical setting are highly heterogeneous and clinical decision should be made under various situations. For example, deep learning models were mostly developed by a training cohort which consists of patients with a particular disorder and healthy controls. Training and even more validation cohorts usually include similar number of patients and controls. However, in the clinical situation, differential diagnosis or clinical decision is made under the patients' symptoms and signs instead of the simple classification. There are different disorders similar to a given disease status which aims at a deep learning model, even more, a few types of rare disorders. The ratio of disease status and healthy status can be considerably different from the cohort for the training. The problem with data distribution is a bigger factor when we use the deep learning model for disease screening purposes in general population (Fig. 2). This is the reason why deep learning models should be subjected to clinical trials in spite of the high accuracy, and



**Fig. 2.** A gap between training and real-world data. Most of deep learning models are developed by patients' data with specific disorders and controls. The problem of deep learning application to the clinic is the difference between real-world data and the training cohort. Real-world data in the clinic included heterogeneous patients different from training cohorts. Furthermore, the distribution of disease and normal is considerably different. This data distribution issue become a bigger factor when deep learning aims at general population.

it is necessary to make appropriate use criteria and use it clinically under limited clinical situations.

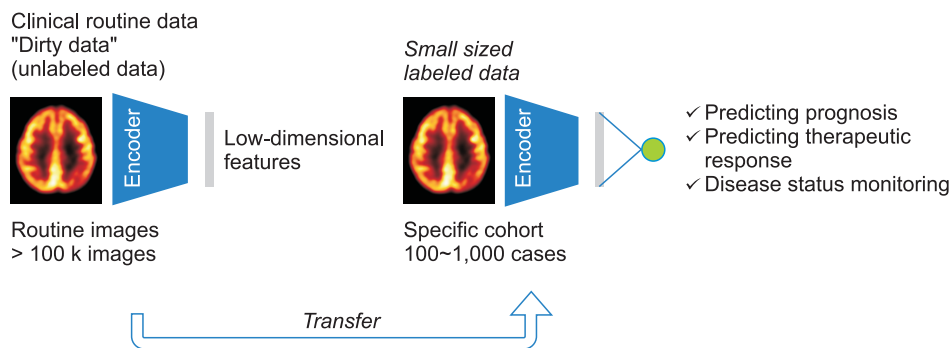
### Uncertainty and unseen data

The issues regarding data distribution and 'unseen data' in training cohorts can be extended to uncertainty. Under the current approaches of supervised learning from big data and their labels, deep learning-based diagnosis and clinical outcome prediction requires diagnostic uncertainty due to unseen and rare cases. Furthermore, clinical decision is not made by differential diagnosis of high probability, but the exclusion of critical diagnosis related to life-threatening. Lowering the uncertainty of a fatal disease is one of the most important factors in diagnostic testing and one of the most important elements of clinical decision to be achieved through biomarkers.<sup>53)</sup> Thus, deep learning models should provide uncertainty in its decision to determine whether subjects need additional diagnostic tests. Bayesian approximation with DL for uncertainty measurement is a good example for supervised learning models.<sup>54)</sup> Another way to bypass the issue regarding uncertainty and unseen data, particularly rare disorders, is to employ unsupervised learning for the anomaly detection. As deep learning is representation learning, latent features in imaging data could show distribution according to training datasets. After the definition of distribution of latent features in the training data, unseen data can be identified by the definition in the latent space.<sup>55,56)</sup> As conditional generative models such as conditional generative adversarial networks (GAN) or variational autoencoders (VAE) synthesize virtual data of specific conditions, it can be used to define a population distribution of specific conditions. For example, by training a generative model for normal aging changes in brain metabolism, a pseudo-population distribution of brain metabolism at each age can be generated.<sup>57)</sup> This generated population distribution will be used to find abnormal patterns taking age information into consideration from a given brain image. This type of anomaly detection can bypass the issue related to deep learning models for heterogeneous disorders.

### Labeling of data: leveraging unlabeled data

Unsupervised learning is an important approach to solve practical issues in labels of imaging data. The labeling of image data, particularly for medical imaging is expansive as well as time-consuming. It requires experts to interpret the images or to decide clinical diagnosis. To obtain 'gold standard' diagnosis, many cases require clinical follow-up interpretations, which need a complex professional review process for medical records. Obviously, ethical issues with regard to the acquisition of large data and their label are inevitable. It is a big obstacle to deep learning application that the data with such labels are limited and labeling as a large scale is much more difficult. In addition, many nuclear medicine and molecular imaging data are more difficult to obtain with large scale with labels as various imaging techniques are used according to the clinical purposes.

One of the ways to overcome this labeling issue will be found in the property of medical imaging data. It is relatively easy to collect heterogeneous image data obtained for clinical routine. By using these clinical routine data and unsupervised learning methods, representative features can be obtained. These representative features will be visualized by dimension reduction methods to intuitively identifying patterns of large imaging data. Furthermore, these features obtained by unsupervised learning can be transferred to relatively small datasets which contain both labels and images. This transfer learning can produce a robust deep learning model even if the well-labeled data is relatively small (Fig. 3).<sup>58,59)</sup> The flexible application of unsupervised learning and transfer learning can be extended to semi-supervised learning. As aforementioned, a database clinically routinely obtained can be relatively easily obtained and a few data in the large unlabeled data can be labeled with the clinical outcome or diagnosis. In spite of a small labeled samples, various deep learning approaches employ unlabeled data to find discriminative representations for small labeled samples.<sup>60,61)</sup> For example, a study was aimed at prediction of FDG uptake estimated by PET using gene expression data for lung cancer, while a small number of subjects include both PET and gene expression data. By employing a larger gene expression dataset without PET data, a prediction model of FDG uptake can



**Fig. 3.** Leveraging unlabeled data as a clinical routine for facilitating deep learning development. As labeling for medical data is too expensive and time-consuming, it is a bottleneck for developing deep learning models. Since it is relatively easy to collect heterogeneous image data obtained for clinical routine, unsupervised learning can leverage these unlabeled 'dirty' data. Unsupervised learning-based feature extraction can be transferred to relatively small cohorts which contain both labels and images to predict clinical outcome as well as differential diagnosis according to the clinical purposes.

be developed.<sup>62)</sup> As many clinical data are placed on the situation of 'large unlabeled data and small labeled data', the deep learning model which can enhance performance through unsupervised learning and unlabeled data will be widely used in future molecular imaging and medical data research.

Another feasible way to overcome the labeling issue is to employ multiple unstructured data corresponding to imaging data. For example, clinical imaging data include text reports which included human interpretation results with natural languages. Even though these reports are mostly unstructured, they have a lot of information of image labels, including differential diagnosis, abnormal findings and disease locations. Data mining of the semantic interactions of medical images and texts will be a feasible approach to develop a deep learning model based on real-world clinical data.<sup>63)</sup> As self-supervised learning of imaging representations using a deep learning model for semantic context can be already used in natural image data, medical imaging data will be trained by representations of text reports.<sup>64)</sup> The learning of representations of the imaging data and finding their clinical significance can be a data-driven approach to develop biomarker without a priori knowledge. The self-supervised learning will be one of the future directions of a data-driven approach and will be achieved by using a text report or intrinsic information, such as age and gender matched with image data.

### Data harmonization

One of the overlooked practical issues is data harmonization. Molecular imaging routinely used in the clinical setting has various types. Numerous tracers can be used to obtain imaging data according to their clinical purposes. Furthermore, image acquisition protocols are varied according to the centers, which may reduce the accuracy of deep learning models when they aim at generalized application for multiple centers. Different imaging textures related to different detector types and image reconstruction algorithms can affect the performance of deep learning. Furthermore, the distribution of tracer has temporal dynamics, image acquisition at different time points may influence on the acquisition of deep learning-based biomarkers. Recently, deep learning has been used to analyze kinetics of dynamic imaging data,<sup>65)</sup> however, most imaging data routinely obtained in the clinic are static images, which require harmonization for multiple centers. The different tracers which aim at same molecular targets also cause a harmonization problem. For example, to obtain the information of brain amyloid deposits, several radiotracers are available, e.g. <sup>11</sup>C-PIB, <sup>18</sup>F-Florbetapir, <sup>18</sup>F-Florbetaben, and <sup>18</sup>F-Flutemetamol. These PET imaging show similar results though different quantification results.<sup>66,67)</sup> While classical amyloid quantification can be overcome by linear correction, deep learning models using heterogeneous image data with these different tracers are challenging.

## Future Direction to Data-Driven Theranostics

In this review, current deep learning models developed for molecular imaging have been briefly introduced in terms of their purposes. As molecular imaging has information of molecular changes regarding pathophysiology, accurate and objective quantification is a critical step to use in the clinic. This quantitative information is linked to clinical decision and prediction of outcome as well as differential diagnosis. Thus, instead of simple diagnostic classification, we should focus on the discovery of biomarkers by extracting functional information of molecular imaging using deep learning. This information can contribute to theranostic approaches, which aim at the combination of diagnostics and therapeutics using same molecular targets. Deep learning models will summarize the status of patients with quantitative value. The models should be clinically validated under the clinical situation with unbiased data instead of limited datasets. Clinically validated molecular imaging-based biomarker can be used to monitor the disease status in terms of functional information. By predicting the outcome of the patient at the individual level using imaging data, therapeutic plans including dose and schedule as well as treatment methods can be personalized. To facilitate the clinically feasible deep learning models, it is promising to leverage unlabeled data and unsupervised learning. This approach will be used to considerably untangle the issues induced by supervised learning approaches which have been employed by most of deep learning models for imaging data. These issues included the heterogeneous data distribution, unseen data and uncertainty of decisions. Furthermore, unsupervised learning followed by transfer learning can develop various types of deep learning models with relatively small samples. Because of the distinctiveness of the medical field and the various purposes of molecular imaging, the development of a deep learning model that meets the particular clinical goals will be necessary, and the result will be an objective biomarker that plays an important role in objective clinical decision.

## Acknowledgements

None.

## Conflicts of Interest

The author has nothing to disclose.

## Availability of Data and Materials

All relevant data are within the paper and its Supporting Information files.

## Ethics Approval and Consent to Participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

## Informed Consent

For this study formal consent is not required.

## References

1. Ravı D, Wong C, Deligianni F, et al. Deep learning for health informatics. *IEEE journal of biomedical and health informatics*. 2017;21(1):4-21.
2. Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*. 2015;115(3):211-252.
3. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Paper presented at: *Advances in neural information processing systems*2012.
4. Weber GM, Mandl KD, Kohane IS. Finding the missing link for big biomedical data. *Jama*. 2014;311(24):2479-2480.
5. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118.
6. Rajpurkar P, Irvin J, Ball RL, et al. Deep learning for chest

- radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS medicine*. 2018;15(11):e1002686.
7. Dhungel N, Carneiro G, Bradley AP. Automated mass detection in mammograms using cascaded deep learning and random forests. Paper presented at: Digital Image Computing: Techniques and Applications (DICTA), 2015 International Conference on 2015.
  8. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*. 2017; 284(2):574-582.
  9. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Medical image analysis*. 2017;42:60-88.
  10. Choi H. Deep learning in nuclear medicine and molecular imaging: current perspectives and future directions. *Nuclear medicine and molecular imaging*. 2018:1-10.
  11. Wang H, Zhou Z, Li Y, et al. Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18 F-FDG PET/CT images. *EJNMMI research*. 2017;7(1):11.
  12. Kirienko M, Sollini M, Silvestri G, et al. Convolutional Neural Networks Promising in Lung Cancer T-Parameter Assessment on Baseline FDG-PET/CT. *Contrast Media & Molecular Imaging*. 2018;2018.
  13. Choi H, Jin KH, Initiative AsDN. Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging. *Behavioural brain research*. 2018;344:103-109.
  14. Ding Y, Sohn JH, Kawczynski MG, et al. A deep learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the brain. *Radiology*. 2018;290(2):456-464.
  15. Liu M, Cheng D, Yan W. Classification of Alzheimer's Disease by Combination of Convolutional and Recurrent Neural Networks Using FDG-PET images. *Frontiers in neuroinformatics*. 2018;12:35.
  16. Liu S, Liu S, Cai W, et al. Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE Transactions on Biomedical Engineering*. 2015;62(4): 1132-1140.
  17. Suk H-I, Lee S-W, Shen D, Initiative AsDN. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage*. 2014;101: 569-582.
  18. Li F, Tran L, Thung K-H, Ji S, Shen D, Li J. A robust deep model for improved classification of AD/MCI patients. *IEEE journal of biomedical and health informatics*. 2015; 19(5):1610-1616.
  19. Choi H, Ha S, Im HJ, Paek SH, Lee DS. Refining diagnosis of Parkinson's disease with deep learning-based interpretation of dopamine transporter imaging. *NeuroImage: Clinical*. 2017.
  20. Martinez-Murcia FJ, Gorritz JM, Ramirez J, Ortiz A. Convolutional Neural Networks for Neuroimaging in Parkinson's Disease: Is Preprocessing Needed? *International journal of neural systems*. 2018:1850035-1850035.
  21. Kim DH, Wit H, Thurston M. Artificial intelligence in the diagnosis of Parkinson's disease from ioflupane-123 single-photon emission computed tomography dopamine transporter scans using transfer learning. *Nuclear medicine communications*. 2018;39(10):887-893.
  22. Betancur JA, Hu L-H, Commandeur F, et al. Deep Learning Analysis of Upright-Supine High-Efficiency SPECT Myocardial Perfusion Imaging for Prediction of Obstructive Coronary Artery Disease: A Multicenter Study. *Journal of Nuclear Medicine*. 2018;jnumed. 118.213538.
  23. Xu C, Xu L, Gao Z, et al. Direct detection of pixel-level myocardial infarction areas via a deep-learning algorithm. Paper presented at: International Conference on Medical Image Computing and Computer-Assisted Intervention 2017.
  24. Betancur J, Commandeur F, Motlagh M, et al. Deep learning for prediction of obstructive disease from fast myocardial perfusion SPECT: a multicenter study. *JACC: Cardiovascular Imaging*. 2018;11(11):1654-1663.
  25. Kim K, Wu D, Gong K, et al. Penalized PET reconstruction using deep learning prior and local linear fitting. *IEEE transactions on medical imaging*. 2018;37(6):1478-1487.
  26. Gong K, Catana C, Qi J, Li Q. PET Image Reconstruction Using Deep Image Prior. *IEEE transactions on medical imaging*. 2018.
  27. Gong K, Guan J, Kim K, et al. Iterative PET image reconstruction using convolutional neural network representation. *IEEE transactions on medical imaging*. 2019;38(3): 675-685.
  28. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image re-



- construction by domain-transform manifold learning. *Nature*. 2018;555(7697):487.
29. Pfaehler E, De Jong JR, Dierckx RA, van Velden FH, Boellaard R. SMART (SiMulAtion and ReconsTruction) PET: an efficient PET simulation-reconstruction tool. *EJNMMI physics*. 2018;5(1):16.
30. Hwang D, Kang SK, Kim KY, et al. Generation of PET attenuation map for whole-body time-of-flight 18F-FDG PET/MRI using a deep neural network trained with simultaneously reconstructed activity and attenuation maps. *Journal of Nuclear Medicine*. 2019;jnumed. 118.219493.
31. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. *Medical physics*. 2017;44(4):1408-1419.
32. Liu F, Jang H, Kijowski R, Bradshaw T, McMillan AB. Deep learning MR imaging-based attenuation correction for PET/MR imaging. *Radiology*. 2017;286(2):676-684.
33. Leynes AP, Yang J, Wiesinger F, et al. Direct pseudoCT generation for pelvis PET/MRI attenuation correction using deep convolutional neural networks with multi-parametric MRI: zero echo-time and dixon deep pseudoCT (ZeDD-CT). *Journal of Nuclear Medicine*. 2017;jnumed. 117.198051.
34. Hwang D, Kim KY, Kang SK, et al. Improving the accuracy of simultaneously reconstructed activity and attenuation maps using deep learning. *Journal of Nuclear Medicine*. 2018;59(10):1624-1629.
35. Xiang L, Qiao Y, Nie D, et al. Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI. *Neurocomputing*. 2017; 267:406-416.
36. Chen KT, Gong E, de Carvalho Macruz FB, et al. Ultra-Low-Dose 18F-Florbetaben Amyloid PET Imaging Using Deep Learning with Multi-Contrast MRI Inputs. *Radiology*. 2018:180940.
37. Wang Y, Yu B, Wang L, et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. *Neuroimage*. 2018;174:550-562.
38. Wang T, Lei Y, Tang H, et al. A learning-based automatic segmentation and quantification method on left ventricle in gated myocardial perfusion SPECT imaging: A feasibility study. *Journal of Nuclear Cardiology*. 2019:1-12.
39. Belal SL, Sadik M, Kaboteh R, et al. Deep Learning for Segmentation of 49 Selected Bones in CT Scans: First Step in Automated PET/CT-based 3D Quantification of Skeletal Metastases. *European Journal of Radiology*. 2019.
40. Chen L, Shen C, Zhou Z, et al. Automatic PET cervical tumor segmentation by combining deep learning and anatomic prior. *Physics in medicine and biology*. 2019.
41. Zhong Z, Kim Y, Plichta K, et al. Simultaneous cosegmentation of tumors in PET-CT images using deep fully convolutional networks. *Medical physics*. 2019;46(2):619-633.
42. Huang B, Chen Z, Wu P-M, et al. Fully Automated Delineation of Gross Tumor Volume for Head and Neck Cancer on PET-CT Using Deep Learning: A Dual-Center Study. *Contrast media & molecular imaging*. 2018;2018.
43. Choi H, Lee DS. Generation of structural MR images from amyloid PET: Application to MR-less quantification. *Journal of Nuclear Medicine*. 2018;59(7):1111-1117.
44. Kang SK, Seo S, Shin SA, et al. Adaptive template generation for amyloid PET using a deep learning approach. *Human brain mapping*. 2018;39(9):3769-3778.
45. Ding Y, Sohn JH, Kawczynski MG, et al. A deep learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the brain. *Radiology*. 2018:180958.
46. Samarin A, Burger C, Wollenweber SD, et al. PET/MR imaging of bone lesions—implications for PET quantification from imperfect attenuation correction. *European journal of nuclear medicine and molecular imaging*. 2012;39(7): 1154-1160.
47. Choi H, Cheon GJ, Kim H-J, et al. Segmentation-based MR attenuation correction including bones also affects quantitation in brain studies: an initial result of 18F-FP-CIT PET/MR for patients with parkinsonism. *Journal of Nuclear Medicine*. 2014;55(10):1617-1622.
48. Park J, Bae S, Seo S, et al. Measurement of Glomerular Filtration Rate using Quantitative SPECT/CT and Deep-learning-based Kidney Segmentation. *Scientific reports*. 2019;9(1):4223.
49. Beam AL, Kohane IS. Translating artificial intelligence into clinical care. *Jama*. 2016;316(22):2368-2369.
50. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nature medicine*. 2019;25(1):30.
51. Saria S, Butte A, Sheikh A. Better medicine through machine learning: What's real, and what's artificial? : Public

- Library of Science; 2018.
52. Park SH, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology*. 2018;286(3):800-809.
  53. Redelmeier DA, Shafir E. Medical decision making in situations that offer multiple alternatives. *Jama*. 1995;273(4):302-305.
  54. Gal Y, Ghahramani Z. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. Paper presented at: international conference on machine learning2016.
  55. Wei Q, Ren Y, Hou R, Shi B, Lo JY, Carin L. Anomaly detection for medical images based on a one-class classification. Paper presented at: Medical Imaging 2018: Computer-Aided Diagnosis2018.
  56. Schlegl T, Seeböck P, Waldstein SM, Schmidt-Erfurth U, Langs G. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. Paper presented at: International Conference on Information Processing in Medical Imaging2017.
  57. Choi H, Kang H, Lee DS, Initiative AsDN. Predicting aging of brain metabolic topography using variational autoencoder. *Frontiers in aging neuroscience*. 2018;10.
  58. Le QV, Ranzato MA, Monga R, et al. Building high-level features using large scale unsupervised learning. *arXiv preprint arXiv:11126209*. 2011.
  59. Bengio Y. Deep learning of representations for unsupervised and transfer learning. Paper presented at: Proceedings of ICML Workshop on Unsupervised and Transfer Learning2012.
  60. Rasmus A, Berglund M, Honkala M, Valpola H, Raiko T. Semi-supervised learning with ladder networks. Paper presented at: Advances in neural information processing systems2015.
  61. Odena A. Semi-supervised learning with generative adversarial networks. *arXiv preprint arXiv:160601583*. 2016.
  62. Choi H, Na KJ. Integrative analysis of imaging and transcriptomic data of the immune landscape associated with tumor metabolism in lung adenocarcinoma: Clinical and prognostic implications. *Theranostics*. 2018;8(7):1956.
  63. Shin H-C, Lu L, Kim L, Seff A, Yao J, Summers RM. Interleaved text/image deep mining on a very large-scale radiology database. Paper presented at: Proceedings of the IEEE conference on computer vision and pattern recognition2015.
  64. Gomez L, Patel Y, Rusiñol M, Karatzas D, Jawahar C. Self-supervised learning of visual features through embedding images into text topic spaces. Paper presented at: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition2017.
  65. Pan L, Cheng C, Haberkorn U, Dimitrakopoulou-Strauss A. Machine learning-based kinetic modeling: a robust and reproducible solution for quantitative analysis of dynamic PET data. *Physics in Medicine & Biology*. 2017;62(9):3566.
  66. Landau SM, Breault C, Joshi AD, et al. Amyloid- $\beta$  imaging with Pittsburgh compound B and florbetapir: comparing radiotracers and quantification methods. *Journal of Nuclear Medicine*. 2013;54(1):70-77.
  67. Klunk WE, Koeppe RA, Price JC, et al. The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. *Alzheimer's & dementia*. 2015;11(1):1-15. e14.