



Artificial intelligence in molecular imaging

Edward H. Herskovits[^]

Department of Diagnostic Radiology and Nuclear Medicine, The University of Maryland, Baltimore, School of Medicine, Baltimore, MD, USA

Correspondence to: Edward H. Herskovits, MD, PhD. Department of Diagnostic Radiology and Nuclear Medicine, The University of Maryland, Baltimore, School of Medicine, 22 S Greene Street, Baltimore, MD 21201, USA. Email: eherskovits@som.umaryland.edu.

Abstract: AI has, to varying degrees, affected all aspects of molecular imaging, from image acquisition to diagnosis. During the last decade, the advent of deep learning in particular has transformed medical image analysis. Although the majority of recent advances have resulted from neural-network models applied to image segmentation, a broad range of techniques has shown promise for image reconstruction, image synthesis, differential-diagnosis generation, and treatment guidance. Applications of AI for drug design indicate the way forward for using AI to facilitate molecular-probe design, which is still in its early stages. Deep-learning models have demonstrated increased efficiency and image quality for PET reconstruction from sinogram data. Generative adversarial networks (GANs), which are paired neural networks that are jointly trained to generate and classify images, have found applications in modality transformation, artifact reduction, and synthetic-PET-image generation. Some AI applications, based either partly or completely on neural-network approaches, have demonstrated superior differential-diagnosis generation relative to radiologists. However, AI models have a history of brittleness, and physicians and patients may not trust AI applications that cannot explain their reasoning. To date, the majority of molecular-imaging applications of AI have been confined to research projects, and are only beginning to find their ways into routine clinical workflows via commercialization and, in some cases, integration into scanner hardware. Evaluation of actual clinical products will yield more realistic assessments of AI's utility in molecular imaging.

Keywords: Artificial intelligence (AI); machine learning; deep learning; nuclear medicine

Submitted Aug 30, 2020. Accepted for publication Nov 27, 2020.

doi: 10.21037/atm-20-6191

View this article at: <http://dx.doi.org/10.21037/atm-20-6191>

Introduction

Molecular imaging—the noninvasive interrogation of molecules involved in a biological process—involves a complex series of steps, the end result of which is qualitative or quantitative characterization of the target molecules. Molecular imaging, in its physiological specificity, complements structural imaging modalities such as computed tomography and magnetic resonance imaging, in which abnormal structures are delineated with high spatial, temporal and contrast resolution. Whereas structural imaging is good at answering questions about a particular region of the body, such as “Is there something abnormal

there?”, which relate to sensitivity, molecular imaging answers questions such as “What types of molecules exist there?”, addressing specificity. Molecular imaging shares with structural imaging its reliance on complex acquisition hardware, such as positron emission tomography (PET) and magnetic-resonance scanners, as well as the generation of complex spatiotemporal data sets.

Although many aspects of molecular imaging are independent of any particular modality, for the purposes of this review I will adopt the perspective of PET-based molecular imaging: molecular-probe design; time-of-flight (TOF) estimation and image reconstruction; image

[^] ORCID: 0000-0001-9445-1774.

quantification; image registration; image segmentation; and synthesis of image and non-image features to generate a differential diagnosis, to quantify disease burden, or to provide prognostic information relative to candidate treatments. Each of these processes involves the analysis of data representing multivariate nonlinear associations among variables; it is no surprise, then, that each has been affected, to varying degrees, by the advent of AI methods well suited to delineating such associations.

Artificial intelligence (AI)

Although there are as many different definitions of AI as there are AI researchers and users, for the purposes of this review we can define AI as the computational manifestation of complex behavior, such as inference or perception (1). Computer scientists further distinguish between strong and weak AI, the former behaving exactly like a human (e.g., manifesting consciousness), and the latter performing well on a given task without regard to implementation details. All methods described in this review are examples of weak AI.

Machine learning

I will further focus on machine learning, the subdomain of AI in which an algorithm takes as input a set of training data, and constructs a model that embodies the associations among variables in those data relevant to a particular task or outcome, such as classification. Machine-learning researchers further subdivide this discipline into unsupervised and supervised machine learning, and reinforcement learning. Unsupervised machine-learning algorithms take as input only the training data, and generate classes; that is, they partition the samples into mutually exclusive groups that may, or may not, have meaning to people. For 2-dimensional data, our visual system is an excellent example of real-time, massively parallel unsupervised machine learning: we immediately perceive clusters of pixels, even if they are unusual in shape and have no meaning to us. In contrast, during supervised machine learning each training-data point is labeled with the class it represents (e.g., tumor versus normal); this label represents the supervision, as it were—an analogy to the knowledge that a professor might impart to a trainee regarding what a particular collection of bright voxels represents in a PET examination. The majority of machine-learning algorithms applied to biomedical image data are supervised. Finally, reinforcement learning presents the machine-learning

algorithm with a series of decisions accompanied by rewards or penalties; the reinforcement-learning algorithm reinforces associations that lead to long-term rewards and weakens associations that lead to penalties (2). These algorithms are most often applied in settings in which there is frequent feedback, such as video games [e.g., (3)].

The models generated by supervised learning allow end users to analyze data not used during the training process, just as a nuclear-medicine trainee who has been provided with many examples of lymphoma, and normal cases, on PET examination during training can then independently distinguish between such cases in medical practice. The superficial resemblance between medical training and machine learning belies the vast gulf in sophistication between the consciousness, and medical and practical knowledge possessed by the trainee, and the relatively impoverished models generated by even the most advanced AI algorithms.

The vast majority of software running on the world's devices, including servers and portable devices, has been written by people. In many domains, people understand the specifications of the software to be developed, and developers have access to domain expertise to ensure that the software is performing as expected. For example, in developing an application to present test questions to students, the developers have access to a database of questions, and to educational experts who tell them how to decide which question to present next, and how to record the students' answers. In domains in which we do not know the optimal way to solve a problem, such as segmenting a tumor on PET examination, but have many examples of accurately segmented tumors based on expert consensus, supervised machine learning offers us the potential to construct models that may approach, or even exceed, the accuracy manifest in the training data. In these cases, the only code written by people is the implementation of the training and inference components of a particular machine-learning algorithm; the algorithm, in effect, uses the data to complete the model.

Neural networks

There is a vast range of machine-learning approaches to building classification models; historically, some of the most common approaches have been statistical (e.g., regression) models, support vector machines, and random forest models. Although these approaches worked well across a variety of domains, image-classification performance has lagged successes achieved in non-spatial domains. Neural-

network research originated decades ago (4,5); however, it was only in the last decade that improvements in hardware allowed the implementation of so-called deep neural networks. Neural networks are crude analogs of neurons and how they intercommunicate. Each node, or neuron, in a neural-network model has parent nodes that influence it, an activation function, a firing threshold, and an output value. Combining parent-activation levels via the activation function yields the neuron's output value, which in turn is propagated to its children. The original neural network design, called a perceptron, had only two layers: input and output; such networks are capable of discriminating between linearly separable classes. Deep networks, due to their many intermediate, or hidden, layers, can model nonlinear multivariate associations and thereby perform complex image-classification or segmentation tasks.

One problem with standard neural networks, such as multilayer perceptrons (MLPs), is that they become unmanageable as the number of inputs increases. For example, an FDG examination of the head with $160 \times 160 \times 96$ voxels would yield approximately 2.5 million input voxels, which could not be accommodated under an MLP architecture. Even if we could scale MLPs to millions of neurons, any change to an input image, such as translation or scaling, would appear as completely different images to the learning algorithm. In contrast, a convolutional neural network (CNN) (6,7) applies a neural-network layer to a subset of an image, systematically traversing across the entire image volume during learning and classification; the output of this layer is then convolved with the next-deeper CNN layer, until output is produced at the last layer. The word convolution refers to the traversal process, which renders CNNs much more parsimonious with respect to parameters and computational requirements, and more robust to perturbations of the input image, relative to MLPs. Stacking convolutional layers in this fashion typically leads to layers with progressively complex features, such as edges, shapes, and objects.

In addition, CNNs alternate convolution layers with pooling layers, which are specified, rather than learned. A pooling layer downsamples its input, typically by a factor of 2, in effect summarizing features and improving spatial invariance (e.g., to translation). One or more fully connected layers complete the network; a fully connected layer is one in which each node is connected to each of the nodes in the previous layer, but not to any of the nodes in the same layer. These fully connected layers summarize features, and map features onto output values; for example,

the softmax function maps features onto output classes such that outputs sum to 1.0, i.e., they act as pseudo-probabilities. In this way, CNN learning algorithms capture relevant features automatically, from data, and tend to be robust to minor perturbations of the input image. Furthermore, since convolution operates on only a small subset of the image at a time for training or classification, the number of parameters to be learned, and hence the computational requirements, of CNNs are much lower than those of MLPs. The vast majority of neural-network architectures applied to image data are deep CNN variants. It is important to note that there is a vast range of neural-network architectures beyond CNNs, including recurrent neural networks and deep belief networks (8).

Generative adversarial networks (GANs)

Much of nuclear medicine centers on providing patients and their physicians with diagnostic or prognostic information from PET or SPECT image data, tasks to which CNNs are well suited. However, there are cases in which researchers wish to generate novel examples of a class, such as a PET examination positive for lymphoma, rather than classify a PET examination to determine a diagnosis or segment a lesion or normal structure of interest. Goodfellow's invention of GANs in 2014 has revolutionized this theretofore relatively underdeveloped research area (9). A GAN consists of two networks—a generator and a discriminator—that learn together in the game-theoretic context of a zero-sum game (10). The generator generates an example of the input data, with the goal of minimizing the difference between the true input feature distribution and the distribution of features in samples that it generates (i.e., its goal is to generate realistic counterfeits). The discriminator, in turn, maximizes accuracy (i.e., its goal is to classify real examples and counterfeits perfectly). During this simultaneous supervised-learning process, the generator examines the characteristics that the discriminator uses, and thereby learns to produce increasingly realistic samples, even as the discriminator becomes better able to detect even subtle differences between real and counterfeit samples. GANs have been used to synthesize strikingly realistic pictures of faces and inanimate objects (11,12). With respect to medical image data, GANs have shown promise across a wide range of applications (13), including simulated modality transformations (14-18), artifact reduction (19,20), and synthetic-image generation for supervised machine learning, thereby obviating patient-privacy protection of

training data (21-23).

Probe design

This aspect of molecular imaging is the least explored, and one of the more controversial. As described in (24), a molecular imaging probe typically consists of various combinations of a signal agent, which can be detected by the scanner; a targeting moiety, which specifically interacts or binds with the molecule of interest; and a linker, which binds the other two components. The *in vivo* interactions of the probe with the molecule of interest, as well as with the remainder of the organism, underlie the vast complexity of probe design. As with drug design, high-throughput screening, including combinatorial chemistry and the creation of phage-display peptide libraries, allow chemists to test large numbers of potential targeting moieties in parallel (24).

Medicinal chemists have also employed simulation software to predict interactions between a candidate probe and a target of interest. This approach is very computationally intensive, but allows chemists to evaluate a large number of candidates without having to synthesize them (25-28). To increase the efficiency of docking software, scientists have applied swarm intelligence (27,29), a branch of AI involving the interaction of large numbers of relatively simple agents, much as ants or bees cooperate in nature, to solve complex tasks. Although these techniques have not been directly applied to design molecular imaging probes, the drug-design literature indicates potential novel methods for applying AI and simulation software to vastly increase the number of candidate probes, and to increase the probability that an individual probe will have desired characteristics, such as bioavailability, stability, and safety (30-32). However, the application of AI techniques to drug design is in its early stages, and there remains skepticism regarding its promise (33,34).

TOF estimation and image reconstruction

TOF estimation is central to localizing PET annihilation events along the line of response in TOF PET. Historically, TOF has been estimated using signal-processing methods such as constant fraction discrimination, leading-edge discrimination, or linear fitting. The ready availability of TOF ground-truth data enabled Berg *et al.* to train a CNN that computes timing information directly from coincidence waveforms (35); they observed approximately 20%

improvement in timing resolution compared to leading-edge discrimination and constant-fraction discrimination.

PET image data are acquired as sinograms, which contain count data for a discrete time interval. As these data are not visually interpretable, they must be reconstructed into images. However, as with CT image reconstruction, noise in the acquired data prevents reconstruction of a unique optimal image; this is an example of the inverse problem, i.e., determining the cause (e.g., 3D distribution of a probe) of a set of observations (e.g., sinograms) (36). Due to noise in the sinogram data, solutions to the inverse problem are unstable, i.e., they may change in unpredictable ways with even small perturbations of the sinogram data. Analytic methods, such as backprojection, backprojection with subsequent filtering, and rebinning, are the oldest and most computationally tractable reconstruction approach (36). However, these approaches are based on simplifying assumptions that limit reconstruction quality. Iterative algorithms allow more complex models of PET data, at the cost of considerably greater computational complexity. These approaches typically include an image model, a data model (statistical distribution of observations given their true values), a system model that mathematically relates the image and data models, an objective function that formalizes the properties of a “good” reconstruction, and an optimization algorithm that optimizes the objective function, thereby generating an image that corresponds to the sinogram data (37).

Due to the extensive computational requirements of iterative methods, and their requirement for an objective function, AI researchers have sought to apply deep learning methods as an alternative to reconstruct PET images directly from sinogram data (38). Häggström *et al.* (39) demonstrated that a deep-learning network, trained on a large corpus of phantom-derived simulated data, outperformed ordered subset expectation maximization and filtered backprojection, with respect to efficiency and image quality. Other image-mapping methods, such as those based on GANs (40), have also shown promise in generating images directly from sinograms. Cui *et al.* (41) trained a sparse autoencoder to reconstruct dynamic PET images, improving upon maximum likelihood expectation maximization at the cost of greater computational requirements.

Researchers have also sought to generate full-dose PET images from low-dose images, in effect removing noise inherent to low-dose technique. The most common approach is based on a GAN-like architecture, in which

an estimator generates a full-dose equivalent image from low-dose input, and a discriminator, which is trained simultaneously to distinguish between true full-dose and estimated PET images (42-46). Preliminary results indicate the potential for considerable dose reduction, with reported good results for low-dose images ranging from 1/4 dose to 1/200 dose. However, this application of GANs is in very early stages, and will require much more extensive validation before approval by regulatory agencies and adoption by scanner manufacturers.

Image segmentation

Anyone who has manually traced the outline of a tumor, lesion, or anatomic structure understands how laborious and error-prone such tasks are; although intra- and inter-rater reliability can be high for such tasks under optimal conditions (47), manual methods do not scale to thousands of image volumes. Furthermore, if people without extensive segmentation experience participate, low intra- and inter-observer reliability may significantly affect results (48,49). To automate all or portions of the PET image-segmentation process, researchers have implemented non-AI approaches based on intensity thresholding, active contours, or spatial models, among others (50,51). In addition, both unsupervised machine learning methods, such as k-nearest neighbor (52,53), and supervised machine learning methods, such as support vector machines (54) and neural networks (53,55-61) have been applied successfully to segment PET or PET/CT tumor volumes automatically. In addition, researchers have combined CNNs and radiomics to detect lymph-node metastases (62). In general, deep-learning approaches have outperformed other AI and non-AI approaches for segmentation of CT and MR images in open competitions (7,63).

Differential diagnosis and prognosis

Historically, AI approaches to differential diagnosis based partially or completely on image-derived features required a two-stage approach consisting of feature extraction, and differential diagnosis based on those features. However, with the advent of CNN-based approaches, researchers have demonstrated the feasibility of generating a differential diagnosis directly from images (64,65), and in some cases outperforming radiologists. More commonly, however, even approaches based on neural networks employ a multi-stage approach, in which statistics or radiomics-derived

metrics are calculated from volumes segmented by a neural network.

CNNs have also been applied to predict outcomes; Ding *et al.* (66) demonstrated that a CNN applied to ^{18}F -FDG PET images could predict the onset of Alzheimer's disease years in advance, exceeding the accuracy of experts. Similarly, Shen *et al.* trained a deep belief network with ^{18}F -FDG PET images, and achieved approximately 86% accuracy in predicting conversion of mild cognitive impairment to Alzheimer's disease (67).

Discussion

Images, with their complex multiscale features, have presented immense challenges to biomedical engineers and AI researchers. Prior to the advent of deep neural networks powered by inexpensive graphics processing units (68), most medical-image-analysis algorithms performed well only on tightly constrained domains; for example, the atlas-registration approach to brain-image segmentation worked well for subjects with normal or atrophied brains, but did not generalize to those with mass effect. Deep learning models have dominated image-segmentation competitions because they can model complex nonlinear multiscale interactions among image features, driven purely by training data, rather than expert-derived equations or constraints (69).

Some of the aspects of deep learning that contribute to its strong performance also carry significant disadvantages. Despite their parsimony relative to MLPs, CNNs may still require very large training data sets, depending on the complexity of the features required for classification. In general, as with any machine-learning approach, the number of training samples required increases with the number of parameters, the number of output classes, the complexity of the associations, and their subtlety. Transfer learning, in which features and parameters learned in one domain (e.g., classifying photographs of everyday objects) are used to initialize a CNN to be trained in a (perhaps distantly) related domain, has proved useful in reducing data requirements in medical and nonmedical applications of AI (70-72). In practice, it can be difficult to know how many training samples to collect. Undersampling, i.e., training on fewer samples than parameters, risks overfitting, in which very specific features are learned that do not generalize well to novel images. This risk is best addressed by enlarging the training sample, and can be detected by evaluating a CNN model with external data, i.e., data that were

collected independently of the training data, preferably at a different site.

Although CNNs are more robust than MLPs, deep-learning models may still be brittle; that is, they may fail unexpectedly with only minor perturbations, such as rotation, of the input. The AI literature is replete with examples of minimal noise (even one voxel!), or minimal rotation, leading to misclassification with high confidence, such as a turtle being labeled as a rifle, or a cat being labeled as guacamole, with upwards of 95% confidence (73-75). Although improving robustness to so-called adversarial attacks is a major focus of AI research, this problem has not been solved, leaving CNN-based models vulnerable to manipulation or inadvertent corruption due to noise (76). Related to this is the lack of introspection and memory regarding training; current machine-learning algorithms, including those for CNNs, generate a classification model that is suited for a narrow task, but does not include information about the training process. Whereas if we showed a PET examination of a patient with a rare disorder such as Erdheim-Chester disease to a nuclear-medicine physician, she would either recall having seen a similar case, or remark that she had never seen anything with this distribution of ^{18}F -FDG avidity. In contrast, a CNN trained to recognize lymphoma on ^{18}F -FDG PET examination will readily classify a CT scan, a PET phantom, or a volume of noise, as it lacks the ability to state that the input does not conform to its training history. The introduction of introspection, or a similar quality, to deep learning will constitute an important step in increasing trust in classification results.

Another limitation of deep-learning methods is their opaqueness. Unlike most older image-analysis methods, which have been based largely on equations that developers implemented, neural networks classify based on connection parameters derived from data, rendering their mechanisms opaque to computer scientists and domain experts. Although visualization tools exist to highlight features relevant to a particular CNN classification task (77), in general it is much more difficult to debug a CNN classifier than one based on a Bayesian network or random forest, for example.

In summary, AI, and in particular deep learning, has profoundly altered the medical-image-analysis landscape. Although most of the advances presented herein have yet to be adopted in the clinic, the rapid pace of advances in this field augur well for eventually increased robustness, transparency and trustworthiness, which will ultimately lead to improved patient care. In the near term, it is clear that

superior segmentation results, which can be readily verified visually, will free trainees and practicing nuclear-medicine physicians from drawing volumes of interest, and will thereby foster the implementation of quantitative radiology. In addition, although AI applications to probe design are in their early stages, results in drug discovery indicate the promise of this approach.

Acknowledgments

Funding: NIH grants (R21 AG058118 and R21 NS108811).

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editor (Dr. Steven P. Rowe) for the series “Artificial Intelligence in Molecular Imaging” published in *Annals of Translational Medicine*. The article has undergone external peer review.

Conflicts of Interest: The author has completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/atm-20-6191>). The series “Artificial Intelligence in Molecular Imaging” was commissioned by the editorial office without any funding or sponsorship. EHH reports that he has Patent US 20160048956: System and method for medical image analysis and probabilistic diagnosis, with royalties paid to University of Pennsylvania. He is the founder of RadOptimal Inc., which makes AI-based software for analyzing radiology reports during dictation. He is a cofounder and board member of Galileo CDS Inc, which makes AI-based software for generating a radiology differential diagnosis.

Ethical Statement: The author is accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Russell SJ, Norvig P. Artificial intelligence: a modern approach. Fourth edition. Hoboken: Pearson; 2021. (Pearson series in artificial intelligence).
- Szepesvári C. Algorithms for Reinforcement Learning. *Synth Lect Artif Intell Mach Learn* 2010;4:1-103.
- Mnih V, Kavukcuoglu K, Silver D, et al. Playing Atari With Deep Reinforcement Learning. NIPS Deep Learning Workshop 2013.
- Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 1958;65:386.
- McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 1943;5:115-33.
- O'Shea K, Nash R. An introduction to convolutional neural networks. ArXiv Prepr ArXiv151108458 2015 [cited 2020 Aug 29]; Available online: <https://arxiv.org/pdf/1511.08458.pdf>
- Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017 Dec;42:60-88.
- Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. *Neural Comput* 2006;18:1527-54.
- Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets. In: Ghahramani Z, Welling M, Cortes C, et al. editors. *Advances in Neural Information Processing Systems 27*. Curran Associates, Inc., 2014:2672-80.
- Wang K, Gou C, Duan Y, et al. Generative adversarial networks: introduction and outlook. *IEEECAAJ Autom Sin* 2017;4:588-98.
- Karras T, Aila T, Laine S, et al. Progressive growing of GANs for improved quality, stability, and variation. ArXiv Prepr ArXiv171010196 2017 [cited 2020 Aug 29]; Available online: <https://arxiv.org/abs/1710.10196>
- This Person Does Not Exist [Internet]. [cited 2020 Aug 28]. Available online: <https://www.thispersondoesnotexist.com/>
- Sorin V, Barash Y, Konen E, et al. Creating Artificial Images for Radiology Applications Using Generative Adversarial Networks (GANs) - A Systematic Review. *Acad Radiol* 2020;27:1175-85.
- Wolterink JM, Dinkla AM, Savenije MHF, et al. Deep MR to CT Synthesis Using Unpaired Data. In: Tsafaris SA, Gooya A, Frangi AF, et al. editors. *Simulation and Synthesis in Medical Imaging*. Cham: Springer International Publishing, 2017:14-23.
- Hiasa Y, Otake Y, Takao M, et al. Cross-Modality Image Synthesis from Unpaired Data Using CycleGAN. In: Gooya A, Goksel O, Oguz I, et al. editors. *Simulation and Synthesis in Medical Imaging*. Cham: Springer International Publishing, 2018:31-41.
- Zhang Z, Yang L, Zheng Y. Translating and Segmenting Multimodal Medical Volumes With Cycle- and Shape-Consistency Generative Adversarial Network. In *Computer Vision and Pattern Recognition* 2018. p. 9242-51. Available online: https://openaccess.thecvf.com/content_cvpr_2018/html/Zhang_Translating_and_Segmenting_CVPR_2018_paper.html
- Ben-Cohen A, Klang E, Raskin SP, et al. Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection. *Eng Appl Artif Intell* 2019;78:186-94.
- Jin CB, Kim H, Liu M, et al. Deep CT to MR Synthesis Using Paired and Unpaired Data. *Sensors* 2019;19:2361.
- Shitrit O, Raviv TR. Accelerated magnetic resonance imaging by adversarial neural network. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 2017:30-8.
- Johnson PM, Drangova M. Conditional generative adversarial network for 3D rigid-body motion correction in MRI. *Magn Reson Med* 2019;82:901-10.
- Chuquicusma MJM, Hussein S, Burt J, et al. How to fool radiologists with generative adversarial networks? A visual Turing test for lung cancer diagnosis. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* 2018:240-4.
- Bermudez C, Plassard AJ, Davis LT, et al. Learning implicit brain MRI manifolds with deep learning. In: *Medical Imaging 2018: Image Processing*. International Society for Optics and Photonics; 2018:105741L.
- Cronin NJ, Finni T, Seynnes O. Using deep learning to generate synthetic B-mode musculoskeletal ultrasound images. *Comput Methods Programs Biomed* 2020;196:105583.
- Chen K, Chen X. Design and Development of Molecular Imaging Probes. *Curr Top Med Chem* 2010;10:1227-36.
- Yi X, Zhang Y, Wang P, et al. Ligands Binding and Molecular Simulation: the Potential Investigation of a Biosensor Based on an Insect Odorant Binding Protein. *Int J Biol Sci* 2015;11:75-87.
- Banerjee SR, Foss CA, Castanares M, et al. Synthesis and Evaluation of Technetium-99m- and Rhenium-Labeled Inhibitors of the Prostate-Specific Membrane Antigen

- (PSMA). *J Med Chem* 2008;51:4504-17.
27. Pagadala NS, Syed K, Tuszynski J. Software for molecular docking: a review. *Biophys Rev* 2017;9:91-102.
 28. Cheng CS, Jia KF, Chen T, et al. Experimentally Validated Novel Inhibitors of Helicobacter pylori Phosphopantetheine Adenyltransferase Discovered by Virtual High-Throughput Screening. *PLoS One* 2013;8:e74271.
 29. Beni G, Wang J. Swarm Intelligence in Cellular Robotic Systems. In: Dario P, Sandini G, Aebischer P. editors. *Robots and Biological Systems: Towards a New Bionics?* Berlin, Heidelberg: Springer, 1993:703-12.
 30. Zhao Z, Qin J, Gou Z, et al. Multi-task learning models for predicting active compounds. *J Biomed Inform* 2020;108:103484.
 31. Jing Y, Bian Y, Hu Z, et al. Deep Learning for Drug Design: an Artificial Intelligence Paradigm for Drug Discovery in the Big Data Era. *AAPS J* 2018;20:58.
 32. Hessler G, Baringhaus KH. Artificial Intelligence in Drug Design. *Molecules* 2018;23:2520.
 33. Schneider P, Walters WP, Plowright AT, et al. Rethinking drug design in the artificial intelligence era. *Nat Rev Drug Discov* 2020;19:353-64.
 34. Jordan AM. Artificial Intelligence in Drug Design—The Storm Before the Calm? *ACS Med Chem Lett* 2018;9:1150-2.
 35. Berg E, Cherry SR. Using convolutional neural networks to estimate time-of-flight from PET detector waveforms. *Phys Med Biol* 2018;63:02LT01.
 36. Tong S, Alessio AM, Kinahan PE. Image reconstruction for PET/CT scanners: past achievements and future challenges. *Imaging Med* 2010;2:529-45.
 37. Reader AJ, Zaidi H. Advances in PET Image Reconstruction. *PET Clin* 2007;2:173-90.
 38. Reader AJ, Corda G, Mehranian A, et al. Deep Learning for PET Image Reconstruction. *IEEE Trans Radiat Plasma Med Sci* 2020;1-1.
 39. Häggström I, Schmidlein CR, Campanella G, et al. DeepPET: A deep encoder-decoder network for directly solving the PET image reconstruction inverse problem. *Med Image Anal* 2019;54:253-62.
 40. Liu Z, Chen H, Liu H. Deep Learning Based Framework for Direct Reconstruction of PET Images. In: Shen D, Liu T, Peters TM, et al. editors. *Medical Image Computing and Computer Assisted Intervention - MICCAI 2019*. Cham: Springer International Publishing, 2019:48-56.
 41. Cui J, Liu X, Wang Y, et al. Deep reconstruction model for dynamic PET images. *PLoS One* 2017;12:e0184667.
 42. Kaplan S, Zhu YM. Full-Dose PET Image Estimation from Low-Dose PET Image Using Deep Learning: a Pilot Study. *J Digit Imaging* 2019;32:773-8.
 43. Wang Y, Yu B, Wang L, et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. *NeuroImage* 2018;174:550-62.
 44. Zhou L, Schaefferkoetter JD, Tham IWK, et al. Supervised learning with cycleGAN for low-dose FDG PET image denoising. *Med Image Anal* 2020;65:101770.
 45. Ouyang J, Chen KT, Gong E, et al. Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss. *Med Phys* 2019;46:3555-64.
 46. 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;267:406-16.
 47. Chow TW, Takeshita S, Honjo K, et al. Comparison of manual and semi-automated delineation of regions of interest for radioligand PET imaging analysis. *BMC Nucl Med* 2007;7:2.
 48. Vorwerk H, Beckmann G, Bremer M, et al. The delineation of target volumes for radiotherapy of lung cancer patients. *Radiother Oncol* 2009;91:455-60.
 49. Pfähler E, Burggraaff C, Kramer G, et al. PET segmentation of bulky tumors: Strategies and workflows to improve inter-observer variability. *PLoS One* 2020;15:e0230901.
 50. Foster B, Bagci U, Mansoor A, et al. A review on segmentation of positron emission tomography images. *Comput Biol Med* 2014;50:76-96.
 51. Comelli A, Bignardi S, Stefano A, et al. Development of a new fully three-dimensional methodology for tumours delineation in functional images. *Comput Biol Med* 2020;120:103701.
 52. Hagos YB, Minh VH, Khawaldeh S, et al. Fast PET Scan Tumor Segmentation Using Superpixels, Principal Component Analysis and K-Means Clustering. *Methods Protoc* 2018;1:7.
 53. Xu L, Tetteh G, Lipkova J, et al. Automated Whole-Body Bone Lesion Detection for Multiple Myeloma on 68Ga-Pentixafor PET/CT Imaging Using Deep Learning Methods. *Contrast Media Mol Imaging* 2018;2018:2391925.
 54. Wu B, Khong PL, Chan T. Automatic detection and classification of nasopharyngeal carcinoma on PET/CT with support vector machine. *Int J Comput Assist Radiol Surg* 2012;7:635-46.

55. Sharif MS, Abbod M, Amira A, et al. Artificial Neural Network-Based System for PET Volume Segmentation. *Int J Biomed Imaging* 2010;Article ID 105610, 2010.
56. Zhao L, Lu Z, Jiang J, et al. Automatic Nasopharyngeal Carcinoma Segmentation Using Fully Convolutional Networks with Auxiliary Paths on Dual-Modality PET-CT Images. *J Digit Imaging* 2019;32:462-70.
57. Zhong Z, Kim Y, Zhou L, et al. 3D fully convolutional networks for co-segmentation of tumors on PET-CT images. *Proc IEEE Int Symp Biomed Imaging* 2018; 2018:228-31.
58. Leung K, Ashrafinia S, Sadaghiani MS, et al. A fully automated deep-learning based method for lesion segmentation in 18F-DCFPyL PSMA PET images of patients with prostate cancer. *J Nucl Med* 2019;60:399.
59. Blanc-Durand P, Gucht AVD, Schaefer N, et al. Automatic lesion detection and segmentation of 18F-FET PET in gliomas: A full 3D U-Net convolutional neural network study. *PLoS One* 2018;13:e0195798.
60. Guo Z, Li X, Huang H, et al. Deep Learning-Based Image Segmentation on Multimodal Medical Imaging. *IEEE Trans Radiat Plasma Med Sci* 2019;3:162-9.
61. Papandrianos N, Papageorgiou E, Anagnostis A, et al. Bone metastasis classification using whole body images from prostate cancer patients based on convolutional neural networks application. *PLoS One* 2020;15:e0237213.
62. Chen L, Zhou Z, Sher D, et al. Combining many-objective radiomics and 3D convolutional neural network through evidential reasoning to predict lymph node metastasis in head and neck cancer. *Phys Med Biol* 2019;64:075011.
63. Xue Y, Xu T, Zhang H, et al. SegAN: Adversarial Network with Multi-scale L1 Loss for Medical Image Segmentation. *Neuroinformatics* 2018;16:383-92.
64. Wu P, Roy AG, Yakushev I, et al. Deep Learning on 18F-FDG PET Imaging for Differential Diagnosis of Parkinsonian Syndromes. *J Nucl Med* 2018;59:624.
65. Liu M, Cheng D, Yan W, et al. Classification of Alzheimer's Disease by Combination of Convolutional and Recurrent Neural Networks Using FDG-PET Images. *Front Neuroinform* 2018;12:35.
66. 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* 2019;290:456-64.
67. Shen T, Jiang J, Lu J, et al. Predicting Alzheimer Disease From Mild Cognitive Impairment With a Deep Belief Network Based on 18F-FDG-PET Images. *Mol Imaging* 2019;18:1536012119877285.
68. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*. Red Hook, NY, USA: Curran Associates Inc., 2012:1097-105.
69. O'Mahony N, Campbell S, Carvalho A, et al. Deep Learning vs. Traditional Computer Vision. In: Arai K, Kapoor S. editors. *Advances in Computer Vision*. Cham: Springer International Publishing, 2020:128-44.
70. Shao L, Zhu F, Li X. Transfer Learning for Visual Categorization: A Survey. *IEEE Trans Neural Netw Learn Syst* 2015;26:1019-34.
71. Antony J, McGuinness K, Connor NEO, et al. Quantifying Radiographic Knee Osteoarthritis Severity using Deep Convolutional Neural Networks. *ArXiv160902469 Cs* 2016 Sep 8 [cited 2020 Aug 29]; Available online: <http://arxiv.org/abs/1609.02469>
72. Kim DH, MacKinnon T. Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. *Clin Radiol* 2018;73:439-45.
73. Su J, Vargas DV, Kouichi S. One pixel attack for fooling deep neural networks. *IEEE Trans Evol Comput* 2019;23:828-41.
74. Akhtar N, Mian A. Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey. *ArXiv180100553 Cs* 2018 Feb 26 [cited 2020 Aug 29]; Available online: <http://arxiv.org/abs/1801.00553>
75. Ilyas A, Engstrom L, Athalye A, . Black-box Adversarial Attacks with Limited Queries and Information. *ArXiv180408598 Cs Stat* 2018 Jul 11 [cited 2020 Aug 29]; Available online: <http://arxiv.org/abs/1804.08598>
76. Antun V, Renna F, Poon C, et al. On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proc Natl Acad Sci U S A* 2020;117:30088-95.
77. Qin Z, Yu F, Liu C, et al. How convolutional neural network see the world - A survey of convolutional neural network visualization methods. *ArXiv180411191 Cs* 2018 May 31 [cited 2020 Aug 29]; Available online: <http://arxiv.org/abs/1804.11191>

Cite this article as: Herskovits EH. Artificial intelligence in molecular imaging. *Ann Transl Med* 2021;9(9):824. doi: 10.21037/atm-20-6191