

X-ray Imaging Meets Deep Learning

Ge Wang*

Department of Biomedical Engineering, School of Engineering, Biomedical Imaging Center,
Center for Biotechnology & Interdisciplinary Studies, Rensselaer Polytechnic Institute,
110 8th St, Troy, New York 12180, USA

ABSTRACT

Deep learning, the mainstream of artificial intelligence (AI), has made progresses in computer vision, exacting information of multi-scale features from images. Since 2016, deep learning methods are being actively developed for tomography, reconstructing images of internal structures from their integrative features such as line integrals. There are both excitements and challenges in the Wild West of AI at large, and AI-based imaging in particular, involving accuracy, robustness, generalizability, interpretability, among others. Based on the author's plenary speech at SPIE Optics + Photonics, August 2, 2021, here we provide a background where x-ray imaging meets deep learning, describe representative results on low-dose CT, sparse-data CT, and deep radiomics, and discuss opportunities to combine data-driven and model-based methods for x-ray CT, other imaging modalities, and their combinations so that imaging service can be significantly improved for precision medicine.

Keywords: Deep learning, machine learning, artificial intelligence (AI), biomedical imaging, x-ray imaging, computed tomography (CT), multi-modality imaging, precision medicine

1. INTRODUCTION

The field of x-ray imaging has been steadily developed since Roentgen's discovery in 1895 (Nobel Prize in 1901). Remarkably, Hounsfield and Cormack pioneered the field of x-ray tomography (Nobel Prize in 1979), which is invaluable for medicine. In 2009, the Science Museum in London invited its curators to select ten important items in its collection and then invited the public to vote for the most impactful discovery among them. With 9,581 out of 50,000 votes, x-ray imaging was recognized as the number one¹. Today, x-ray radiography and tomography are routinely performed for diverse applications covering six orders of magnitude in terms of image resolution and object size, and in different imaging modes (attenuation, photon-counting, phase-contrast, and dark-field imaging, as well as circular, helical, and robotic scanning), with biomedical imaging as the primary example. Over the past years, deep learning has been successful in a number of fields, including biomedical imaging especially x-ray tomographic imaging²⁻⁵.

Kuhn proposed in his book "*Structure of Scientific Revolutions*" that scientific progresses are made through paradigm shifts. Basically, the four paradigms are considered as (1) Empirical (to describe phenomena), (2) Theoretical (to extract laws and formulate models), (3) Computational (to simulate phenomena and processes), and (4) Data-driven (to extract knowledge from big data by deep networks with high-performance computing techniques). Currently, we are working in the 4th paradigm, also referred to as data science, artificial intelligence, machine learning, and deep learning.

Let us first review the idea of deep learning (assuming you just heard it), aided by Figure 1. *Deep learning* means to use *deep artificial neural networks* for machine learning. While there are many artificial neural networks, their building blocks are simply *artificial neurons*. An artificial neuron is a model of a biological neuron. In an artificial neuron, a number of input signals are respectively weighted and added, which is a linear operation in the form of the inner product. The inner product is then processed via thresholding or a nonlinear activation to decide if the neuron should be activated or not, which means if it should take a high or low value. There are various activation functions, such as *Sigmoid* $S(x)=1/(1+e^{-x})$ with the range $(0, 1)$ and *ReLU* $R(x)=\max(0,x)$. The *feed-forward network* is common, in which there are the input layer, the output layer, and hidden layers between the input and output layers. An input vector can be processed layer-wise (one layer's output serves as the input to the next layer), and finally the output layer presents a result. In addition to the feed-forward network, many other types of architectures are available. Roughly speaking, a deep network contains many hidden layers, which facilitates multi-scale analysis and knowledge representation.

*ge-wang@iee.org; <https://biotech.rpi.edu/centers/bic/people/faculty/ge-wang>

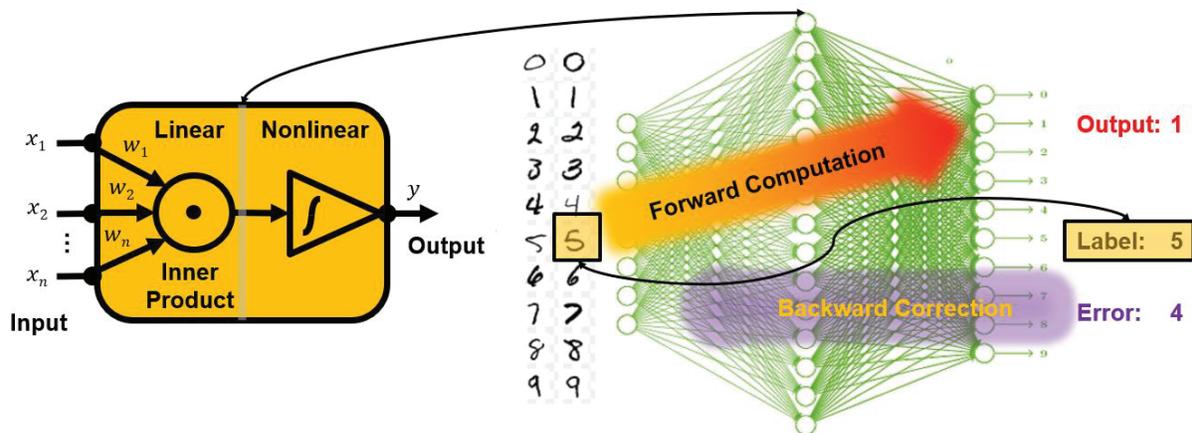


Figure 1. Essential idea of deep learning. A neural network consists of inter-connected neurons, each performing a linear computation (inner product) and then a nonlinear operation (thresholding, such as ReLU). The weighting parameters for the links between the neurons are randomly initialized and iteratively adjusted to minimize a loss function (say, the absolute error shown as “4” in that iteration) using training data to output the right answer eventually (i.e., “5” in this case) .

The specific feed-forward network in Figure 1 contains only one hidden layer, but it can do an excellent job already for digit recognition (given an image of a handwritten digit, the network will recognize it). The input is a vectorized digit image. The output layer is the so-called *one-hot vector*, whose one and only one element will be activated to take a higher value than the other elements, and the activated element/neuron indicates what the input digit is (each digit is assigned to one and only one neuron in the one-hot vector). Initially, the parameters of the network were randomly initialized. For a given image, the output of the hidden layer can be computed. Then, the output vector of the network can be obtained. Generally, it is unlikely that this output will match our desired result. Then, the error can be reduced by adjusting the current weighting parameters, which can be done via gradient descent search in a *backpropagation* manner. If there are many hidden layers, the forward computation and backpropagation processes can be similarly performed.

Currently, there are a number of popular deep learning platforms such as TensorFlow, Torch, Caffe, Keras, and MatLab. On these platforms, a network architecture can be easily defined, trained, tested, and deployed. In other words, it is not difficult to use deep learning techniques in many tasks. However, a governing theory of deep learning is still under development, and many fundamental questions are yet to be answered. In a big picture, deep learning involves novel algorithmic design, complicated non-convex optimization, and unique challenges such as robustness, generalizability, uncertainty, interpretability, regulation, privacy, fairness, and so on.

In this article, we would like to focus on the intersection of x-ray imaging and deep learning where there are great synergy and numerous opportunities. On one hand, deep learning has a remarkable potential to empower x-ray imaging performance. On the other hand, x-ray imaging suggests interesting possibilities for deep learning research and applications. In the next section, we describe deep tomographic imaging aided by bibliographic analysis, and representative promising results of deep imaging, with an emphasis on x-ray imaging especially computed tomography (CT). In the third section, we discuss some current challenges of deep learning as related to tomography, mainly instabilities of deep reconstructions and lack of generalizability of deep radiomics when a trained network is applied to out of distribution datasets. Finally, we present some research directions, which include combination of data-driven and rule-based machine learning as well as integration of cost-effective tomographic imagers in the deep learning framework.

2. PROMISES OF DEEP IMAGING

X-ray radiography and CT are major medical imaging modalities and responsible for a majority of imaging procedures. Medical imaging, artificial intelligence and deep learning are important topics with high-level public interest. By a Google Trends search, the Google-defined public attention to medical imaging in percentile has been fairly constant over the past decade. In comparison, artificial intelligence and deep learning attracted rapidly increasing attention since a few of years ago. In particular, the attention to deep learning exceeded that to medical imaging in 2016, as illustrated in Figure 2.

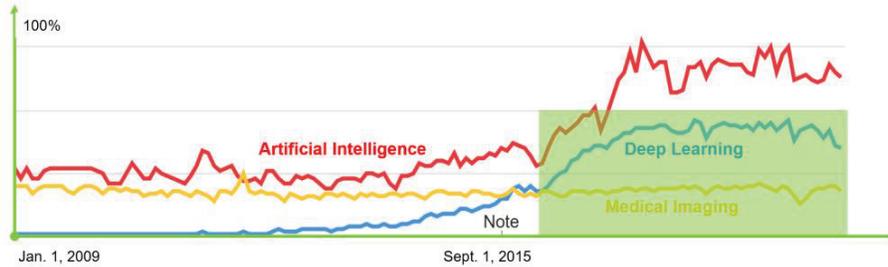


Figure 2. While the attention to medical imaging remains constant, the attention to artificial intelligence and its mainstream approach “deep learning” has enjoyed a recent surge. The attention to deep learning exceeded medical imaging in 2016.

In 2016, we wrote a perspective article on deep learning-based tomographic imaging⁶. This is a new frontier of AI and machine learning. Deep learning is widely used in computer vision and image analysis, which deal with existing images and produce features. By deep learning for tomographic imaging, we produce images of multi-dimensional structures from externally measured data in the form of various transforms such as integrals and harmonic coefficients. In brief, while computer vision and image analysis go *from images to features*, deep tomographic imaging transforms *features to images*. Based on this perspective, we filed the first patent on deep tomographic imaging in 2016, which was granted in 2021⁷. Also, in 2019 we published the first book dedicated to this emerging field⁸.

Traditionally, there were only two kinds of image reconstruction algorithms: analytic and iterative. While the former is based on a closed-form solution, the latter performs model-based optimization iteratively. With the emerging deep learning-based algorithms, we train a deep network iteratively but use the trained network in the feed forward fashion like a close-form formula. Currently, deep imaging has become a main stream of tomographic imaging research. On July 18, 2021, a Scopus search retrieved 944 publications using the *TITLE-ABS-KEY rule A = "deep learning" AND medical AND image AND reconstruct**) and 373 publications with the rule *B = A AND (x-ray OR CT OR "computed tomography")*). That is, roughly speaking, among all deep tomographic imaging papers, nearly 40% is related to x-ray imaging or CT.

A quick analysis on the bibliographic data retrieved using the above-mentioned Scopus rule *B* is summarized in Figure 3. It can be observed that deep x-ray imaging and CT has covered a wide array of deep networks and clinical tasks, and deep reconstruction research has placed an emphasis on low-dose CT and sparse data reconstruction. Also, many deep learning techniques have been developed for segmentation, detection, classification, and artifact reduction.

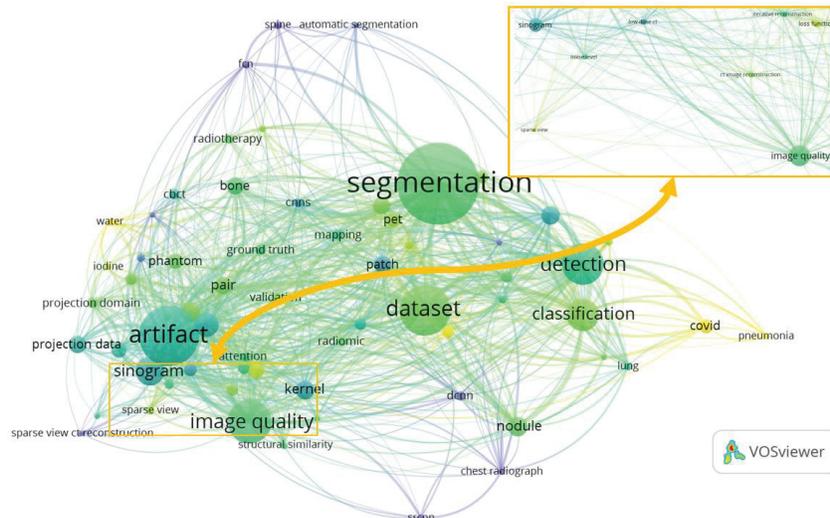


Figure 3. Visualization of the Scopus search results using the TITLE-ABS-KEY rule = ("deep learning" AND medical AND image AND reconstruct*) AND (x-ray OR CT OR "computed tomography").

Based on the above bibliographic screening, let us describe the following examples from our recent studies on deep learning-based low-dose CT post-processing, sparse data reconstruction, and deep radiomics. Needless to say, excellent results on these and other topics were also produced by our peers worldwide; for more details, see the review articles²⁻⁵.

The first example is low-dose CT denoising. X-ray CT is the method of choice for lung cancer screening, and often times used in many other tasks before any major procedure is taken. However, x-ray radiation dose potentially induces cancer and genetic damage. Hence, it is highly desirable to minimize radiation dose, but a reduced dose CT scan is associated with a degraded signal-to-noise ratio, and a compromised image quality. In collaboration with Massachusetts General Hospital, we designed a modularized adaptive processing deep network with radiologists included in the training loop to optimize the denoising strength⁹, as shown in Figure 4. In a double blind reader study, our deep approach was shown to be superior or comparable to the commercial iterative low-dose CT algorithms on major vendors' CT scanners⁹.

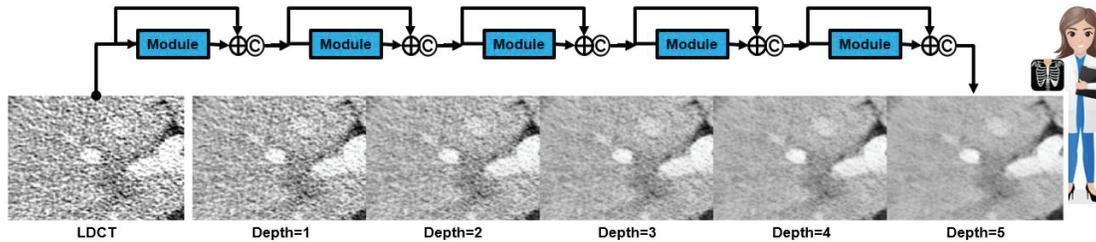


Figure 4. Our modularized deep denoising network trained in the radiologists-in-loop fashion⁹. The network produces denoised images at different depths for radiologists to decide the optimal number of steps specific to a diagnostic task.

The second example is sparse-data tomography. Specifically, in C-arm CT devices, futuristic stationary or robotic CT systems, there are only a limited number of projections. Without deep reconstruction, contemporary algorithms via sparsity promotion (such as total variation (TV) minimization) could do a good job from at least over one hundred views. With deep reconstruction, comprehensive prior knowledge on the image content in an application domain can be utilized for unprecedented image quality from a much-reduced amount of data. We recently designed a deep reconstruction network referred to as SUGAR for “split unrolled grid-like alternative reconstruction”¹⁰. In the SUGAR network, we perform deep learning from multiple inter-connected angles. First, we learn to reconstruct a low resolution version of an underlying image from a sparse sinogram. Then, we use an up-sampled low-resolution image and the sinogram to estimate a high-resolution image. Furthermore, we iteratively refine the current image in a format inspired by the split Bregman scheme. In the extremely challenging case of only 36 projections, the SUGAR network produced encouraging initial result as compared to the counterparts from TV minimization (as well as existing deep reconstruction networks; for details, please see¹⁰), as shown in Figure 5.

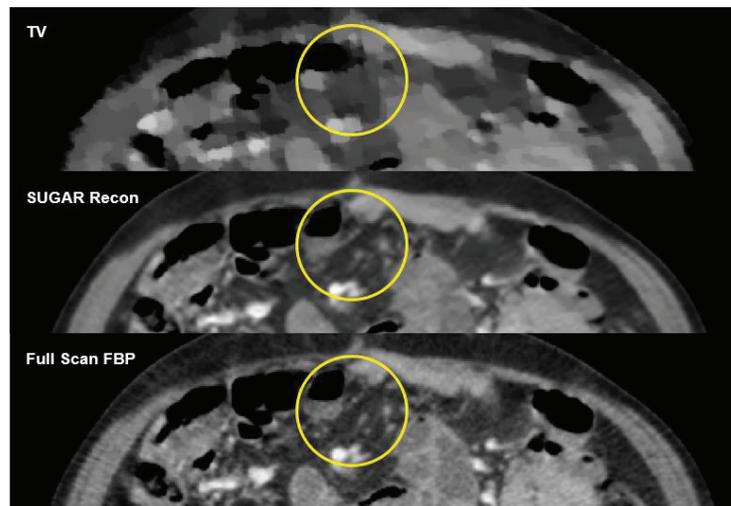


Figure 5. Our recently designed SUGAR network showing promising initial tomographic results from rather sparse data¹⁰.

While low-dose CT is mostly used for lung cancer screening, in collaboration with Massachusetts General Hospital we analyzed low-dose CT images to assess cardiovascular disease risks¹¹. Our deep learning model processed numerous low-dose CT images and obtained the ROC curve comparable to that based on radiologists' reports of ECG-gated cardiac CT scans. That is to say, low-dose CT images can be deeply analyzed for dual-screening of both lung cancer and cardiovascular disease risks.

Although the above three examples are all on x-ray CT, deep tomographic imaging results are very promising for other modalities including MRI, nuclear tomography, ultrasound, optical imaging, and more. For details, please check the literature, for example, the review article⁵.

Beyond medical imaging, there are other kinds of deep imaging opportunities. As a rather interesting example, with an intense synchrotron radiation beam, a biological target, such as a protein molecule, could be immediately smashed to emit scatter waves. How to reconstruct the 3D structure of an object from such a single shot is fundamentally important for biological research and drug design. Efforts were made to perform 3D reconstruction from 2D information on a spherical wave front¹²⁻¹⁴. We did some work to perform this task with polychromatic radiation¹⁴, instead of monochromatic radiation. Very recently, the AlphaFold networks take a protein sequence as the input, and reconstruct the 3D structure of the protein¹⁵. Actually, the AlphaFold network represents the-state-of-the-art prior biological information on the underlying protein to be reconstructed. Along this direction, we can imagine a combination of deep reconstruction and AlphaFold-based prior knowledge. The hope is shown in Figure 6 where an excellent reconstruction will be enabled at the intersection of the solution space determined by scatter data, the space of sparse images, and the space described by the AlphaFold prior.

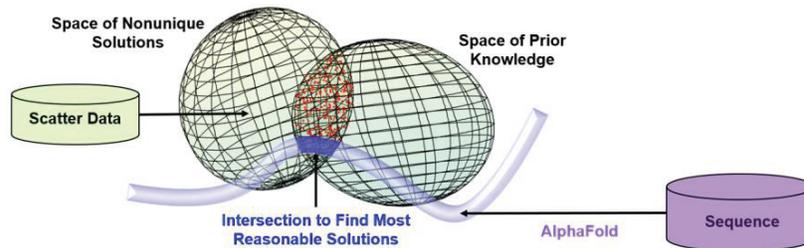


Figure 6. AlphaFold-based x-ray tomography of non-crystalline targets such as protein molecules via deep learning. Since the intense radiation required for ultrahigh resolution destroys the target immediately, only a single-shot scatter wave front data are available for tomographic reconstruction, representing a most challenging imaging task from extremely sparse data.

3. CHALLENGES IN IMAGING APPLICATIONS

In the preceding section, we have presented deep learning-based tomographic imaging very positively. However, that is only one side of the coin. Now, let us emphasize challenges in real-world imaging applications.

First, the instability (also referred to as vulnerability or brittleness) of deep networks is a primary concern¹⁶. Figure 7 is an example. In this example, a purposely computed subtle noise image in the middle was added to this left digit image “4” and formed the right digit image “4”. For human eyes, there is little difference between the left and right images, but our trained classification network was easily tricked to misclassify the right image to a digit “2” with high confidence while the network classified the left image correctly as “4”.

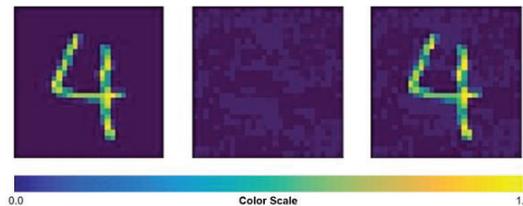


Figure 7. Adversarial attacks putting a trained deep network at a high risk. While the original digit image (Left) was correctly classified by a deep classifier, a subtle adversarial noise (Middle) was computed and added to the original image to form a visually indifferent image (Right) which was misclassified by the same classifier.

There is a landmark paper published in PNAS on instabilities of deep learning in image reconstruction¹⁷. In this study, three types of instabilities of deep reconstruction networks were reported for both deep CT and MRI reconstructions. The three Instabilities include (1) strong artefacts from perturbation; (2) small features undetectable; and (3) increased data compromising reconstruction quality. These instabilities are believed to come from lacking kernel awareness and be “nontrivial to overcome”, but compressed sensing algorithms were reported to be stable due to their kernel awareness¹⁷. The first two types of instabilities are the types I and II errors: false alarm and missing target errors. The third type of instabilities is somehow like a “fed-up” problem (over-weight is not good); that is, more data confused a network that had been trained to handle a certain amount of data, leading to a network performance poorer than that without additional data, which is kind of counter-intuitive.

How to address the above-identified instabilities? A general answer was mentioned in our perspective article⁶. The key idea is that complementary algorithms, such as compressed a model-based iterative or compressed sensing inspired algorithm and a deep network, can be combined for synergy of image reconstruction. For example, a maximum likelihood image reconstruction can be further improved by a deep network. Also, we have a specific answer to address the problems posed in the PNAS paper¹⁷. This is called the Analytic, Compressive, Iterative Deep (or ACID) network that integrates analytic reconstruction, compressed sensing, iterative refinement and deep learning together for accurate and stable image reconstruction¹⁸.

Let us briefly explain this ACID framework with Figure 8. First, we have a reconstruction network Φ that transforms the original measurement (a sinogram or a k-space dataset) into an initial image. This initial image reflects strong image prior knowledge extracted from big data (CT or MRI data). This initial image is subject to errors and instabilities. Then, the current image is improved by a compressed sensing inspired sparsity promotion module Θ to output a sparsified image. Furthermore, based on this sparse image we can estimate measurement data using the system model A , and compare the estimated and measured data to compute the residual data that reflect observable errors, indicating if the current image needs correction. Any significant residual is taken by the same deep reconstruction network Φ to produce an incremental image on the top of the current image to form an updated image. This updated image is again processed by the compressed sensing module Θ . So on and so forth to improve the current image gradually.

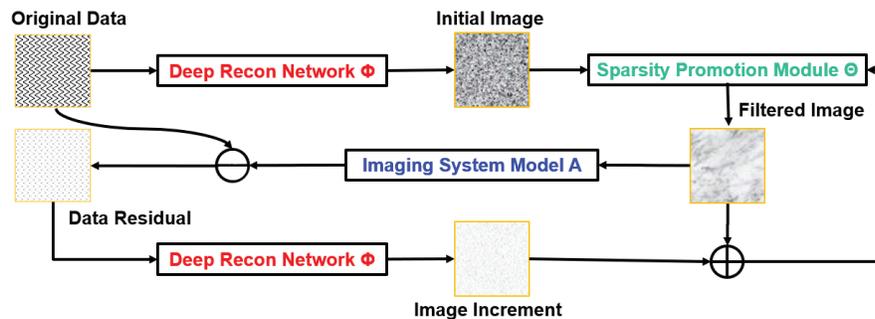


Figure 8. General idea behind the Analytic, Compressive, Iterative Deep (ACID) network as an exemplary embodiment to integrate strengths of analytic reconstruction, sparsity promotion, iterative refinement, and deep learning.

Furthermore, we analyzed the convergence of our ACID network¹⁸. Our analysis assumes the bounded error norm (BEN) property. Basically, the BEN property means that the error of a deep reconstruction is less than the ground truth in the sense of the L_2 norm. The BEN property is a special case of the Lipschitz continuity, and serves as a simple model to demonstrate the feasibility of hybrid image reconstruction. Assuming the validity of BEN, the convergence of the ACID network becomes heuristically clear. The initial deep reconstruction network output is not perfect but the error component is less than the ground truth. The error component has observable and unobservable parts. Because of the observable part, the predicted and real measurement datasets are not the same. Their difference, or the data residual, can be used to reduce the observable error component. Given the BEN property, after an ACID iteration the observable error component will be reduced. This residual-based iterative refinement can be repeated. In the limiting case, the observable error in the final reconstruction will be eliminated at an exponential rate. The only error remains in the limiting case will be unobservable. It is underlined that if the deep network is properly designed, well trained on big and diverse data, and the input is not out of the assumed distribution, the unobservable error of the output image should be typically small, as reported in many deep reconstruction studies. In other words, an ACID-type deep reconstruction approach is to seek a

best solution at the intersection of the space of data-driven solutions, the space of sparse solutions, and the space of solutions subject to data constraints. This idea is in the same as what has been mentioned in the preceding section to reconstruct the 3D structure of a protein molecule aided by AlphaFold.

In addition to the instability issue, the generalizability is also important in applications of deep networks, which is actually related to the instability of deep networks. In a recent Nature Medical paper on how medical AI devices are evaluated by FDA, some interesting findings were reported¹⁹. In a remarkable case study on the potential problem with out of distribution datasets, a well-established AI model for detecting collapsed lungs was respectively trained on each of the three public big datasets. When the deep network was trained on one of the three datasets and then tested on the other two, the deep model suffered from ~10% accuracy drops¹⁹. Currently, we are actively working on this problem.

4. DIRECTIONS FOR FURTHER RESEARCH

Deep learning-based tomographic imaging enjoy a strong research momentum. There are a good number of promising directions of research, including robustness, generalizability, uncertainty, interpretability, regulation, privacy, fairness, and so on, as mentioned in the Introduction section.

4.1 Extension of Artificial Neurons to Diverse Forms

Among many research opportunities, here we underline the interpretability as an important direction for AI in general and deep learning-based imaging in particular. A deep network is typically considered as a black box, and extensive research is in progress to make the black box more transparent, making the box “grey” or “white”²⁰. Most of these interpretation methods are heuristic^{21–23}. For example, we can ask how the network prediction will change without a given training point? This exercise will help identify which samples are informative for the prediction of a network. As the second example, we can ask what is the important data routing path in a network? The class activation map is yet another example, which shows features relevant to the classification results. We can also model the network training process as a dynamic system in the form of different equations or under extreme conditions such as assuming very wide and/or very deep network layers²².

Instead of commenting on representative methods for network interpretation, here we mention some of our pilot results. Currently, almost all deep networks use the type of artificial neurons we described earlier that rely on a linear operation, *i.e.*, the inner product. Interestingly, in biology there are many types of neurons with diverse forms and functions. Hence, we are motivated to innovate deep networks at the cellular level by introducing new types of artificial neurons. Specifically, we proposed to replace the inner product in the conventional artificial neuron with a quadratic operation by multiplying two different inner products together for the same input vector, and then add another quadratic function of the input vector^{24,25}. This modification will increase the number of parameters of the conventional neuron from n to $3n$. The rationale for our design can be appreciated in the case of a two-dimensional logic function. It is well known that a conventional neuron cannot implement the *XOR* gate. In contrast, our quadratic neuron can implement a hyperbolic function and directly implement the *XOR* gate. As a matter of fact, our quadratic neuron can be adapted to implement any two-dimensional logic function. Along this direction, we believe that other kinds of artificial neurons can be designed to improve deep learning performance. Hinton’s Capsule idea is in a similar spirit²⁶, since a capsule can be viewed as a complicated neural computing unit.

According to the De Morgan’s law, the negation of a disjunction is the conjunction of the negations. Inspired by this duality, we systematically analyzed the family of *ReLU* networks, and proved that the width and depth of *ReLU* networks are quasi-equivalent for either classification or regression tasks. That is, for a given network complexity, we can have a wide network, a deep network, and many network variants between, all of which can perform essentially the same classification or regression task. If you are interested, please read our paper²⁷ for mathematical details, including several theorems and proofs.

4.2 Integration of Connectionism and Symbolism

Since our quadratic neurons are continuous instead of being binary, we tend to consider them as generalized fuzzy logic gates. Hence, a deep neural network is nothing but a deep fuzzy logic system²⁴. In a case study, we reported a digit recognition network as shown in Figure 9, where the size of a quadratic neuron indicates its relative importance, while the color of a quadratic neuron shows the type of its fuzzy logic operation, for example, hyperbolic or parabolic, as determined via spectral analysis²⁴. A fuzzy logic system is essentially a rule-based system. That is to say, we could obtain, in a good sense, a rule-based system in a data-driven fashion. To be a desirable rule-based system, in many cases

we should compress or transform a deep fuzzy logic network of numerous parameters into a relatively compact knowledge graph to enjoy sparsity or regularity. This idea should be a way to link the connectionism and the symbolism, and if so, this will offer a great interpretability of a deep neural network, as also shown in Figure 9. We believe that the knowledge graph field is critically important in the years to come^{28,29}. If we could develop good methods to combine data-driven and rule-based learning, and perform multi-task learning, semi-supervised and unsupervised learning, we should be one major step closer to artificial general intelligence or strong AI.

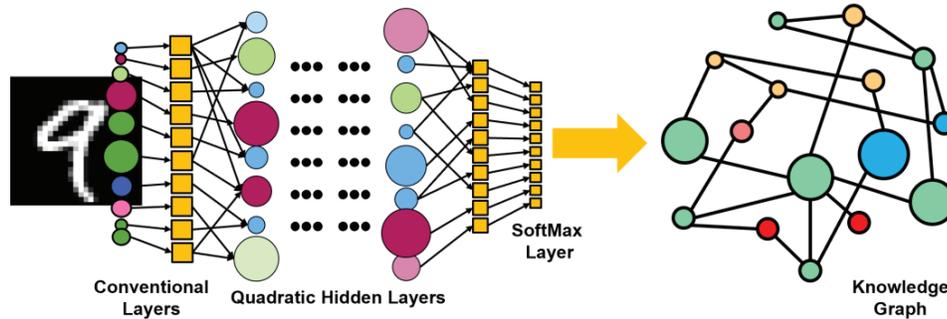


Figure 9. Combination of data-driven and rule-based methods enabled by the deep fuzzy logic interpretation and knowledge distillation using knowledge graph techniques.

4.3 Fusion of Relevant Tomographic Modalities Cost-effectively

Toward strong AI, an AI system should behave more like a human. Last month, a deep omni-perception system was published³⁰, which can process images, text and audio files simultaneously. This is achieved with token-based embeddings, multi-task pretext learning, and multi-granularity integration for an image-text-audio multi-modal representation, and produced very encouraging results³⁰. This excellent omni-perception work reminds us of our earlier omni-tomography proposal³¹, in which we hope to achieve a grand fusion of relevant tomographic imaging modalities including CT, MRI and PET/SPECT. By such a fusion, multi-modalities are integrated into a single gantry, and collect diverse datasets at the same time. As a result, spatiotemporal correlations are ideally captured, and complementary information can be synergized for cardiac and cancer imaging³². However, it appeared years ago that such omni-tomographic instrumentation would be rather challenging and expensive, given the bulkiness of individual imagers and physical limitations for hardware integration.

Thanks to the rapid development of medical imaging and deep learning technologies, the situation is now quite different. The companies NanoX and others reported cost-effective x-ray imagers for tomosynthesis. Hyperfine is the brand name of low-field portable MRI scanners. Also, photon-counting detectors can record both x-ray and gamma ray signals to potentially unify CT and SPECT. Hence, it becomes feasible to integrate these cost-effective imagers together into a cost-effective omni-tomographic system in the deep learning framework, as sketched in Figure 10. This could be a next highlight of medical imaging, featured by portability, point of care, and power unleashed with auto-driving cars to deliver imaging services wherever and whenever needed. This will be instrumental for underdeveloped countries and remote areas.

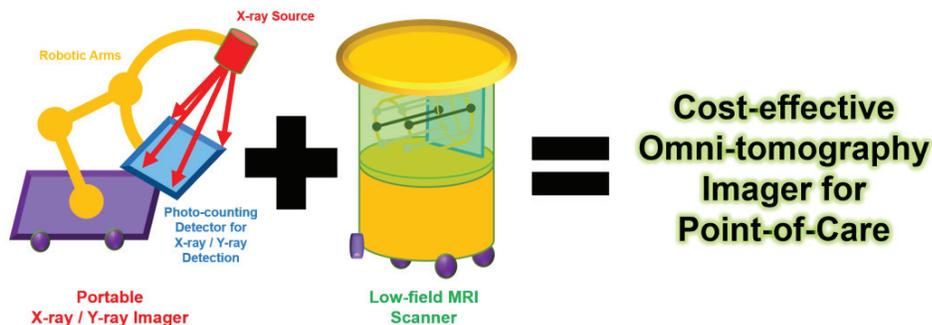


Figure 10. Cost-effective construction of portable and mobile hybrid imagers becomes feasible for point-of-care.

4.4 Concluding Remarks

In Figure 11, the current landscape of medical imaging and radiation therapy is on the left. The future for omni-tomography and image-guided therapy is presented on the right. We believe that tighter fusion of imaging modalities in a more cost-effective way is happening. We envision a paradigm shift in medical imaging, from hospital/clinic/center-oriented to decentralized, intelligent and integrated services. A futuristic possibility is what we call “Auto-driving Vehicle-based Affordable Tomography-Analytics Robots (AVATAR)”³³, which are most desirable on natural disaster spots, after terrorists’ attacks, and near battle fields. Also, AVATAR are advantageous in routine healthcare imaging such as for cancer screening because of a potentially much-reduced cost, full automation, and greatly-improved convenience.

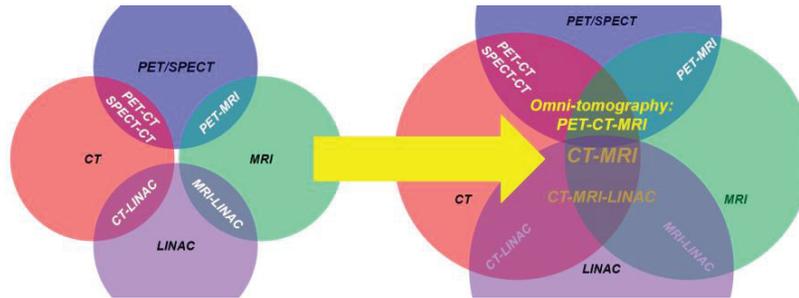


Figure 11. Tighter fusion of imaging modalities and more cost-effective instrumentation in the deep learning framework.

Rail/bus stations made milestones in the transportation history but quickly overwhelmed by privately-owned cars. Today, the Uber-type service becomes popular, and auto-driving cars are on the horizon. Similarly, supermarkets/malls are where we shop but the trend is moving towards internet-shopping and door-to-door delivery, accelerated by the pandemic. Not long ago, we used to go to cinemas and theaters for entertainment. Then, we built home theaters. Now, we use smart phones to watch TVs most of time. Yet, many believe that the blockchain technology will revolutionize the society; say, eliminating banks. After “Internet of information”, “Internet of things” is under active development, and “Internet of service” seems the next wave. The individualized and optimized use of information, products, services demands decentralization, interconnection, artificial intelligence especially deep learning, promoting democracy and improving quality of life. We are blessed by so many research opportunities towards the future of imaging services.

In the scientific fiction “Siren of Titan”, Kurt Vonnegut wrote, “Once upon a time ... there were creatures who weren’t anything like machines. ... And these poor creatures were obsessed by the idea that everything that existed had to have a purpose, ... And every time they found out what seemed to be a purpose of themselves, ..., the creatures would make a machine to serve it.” We certainly found that longevity and well-being seemed to be one of our purposes, and we are endeavoring to invent a brighter future. As shown in Figure 12, the efforts to synergize x-ray imaging and deep learning are made on a global scale, suggesting information exchange and international collaboration.

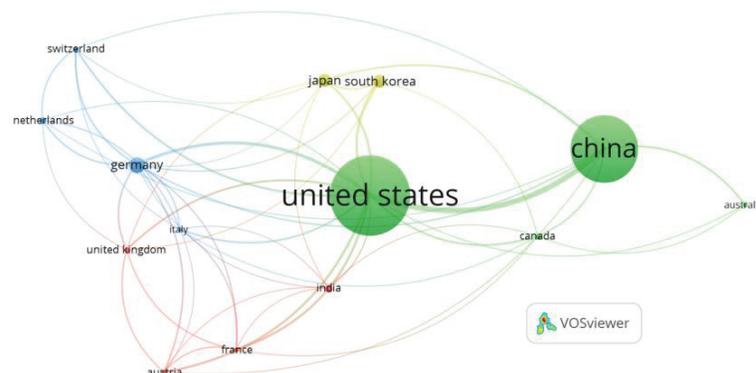


Figure 12. Global landscape at the intersection of x-ray imaging and deep learning based on data using Scopus TITLE-ABS-KEY rule = (“deep learning” AND medical AND image AND reconstruct*) AND (x-ray OR CT OR “computed tomography”).

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health under the grants R01CA237267, R01HL151561, R01EB026646, and R01CA233888. The author thanks his lab members for research contributions cited below, and Dr. Michael Vannier, Dr. Mannudeep Kalra, Dr. Lizhi Sun, Dr. Hengyong Yu, Dr. Xuanqin Mou, Mr. Christopher Wiedeman, and many others for discussions.

REFERENCES

- [1] Kermeliotis, T., “X-ray voted top modern discovery,” Cable News Network, <https://www.cnn.com/2009/WORLD/europe/11/04/xray.machine.science.museum/index.html> (2009).
- [2] Wang, G., Ye, J. C. and De Man, B., “Deep learning for tomographic image reconstruction,” *Nat. Mach. Intell.* 2(12), 737–748 (2020).
- [3] Lell, M. M. and Kachelrieß, M., “Recent and upcoming technological developments in computed tomography: High speed, low dose, deep learning, multienergy,” *Invest. Radiol.* 55(1), 8–19 (2020).
- [4] Maier, A., Syben, C., Lasser, T., Riess, C., “A gentle introduction to deep learning in medical image processing | Eine sanfte Einführung in Tiefes Lernen in der Medizinischen Bildverarbeitung,” *Z. Med. Phys.* 29(2), 86–101 (2019).
- [5] Sahiner, B., Pezeshk, A., Hadjiiski, L. M., Wang, X., Drukker, K., Cha, K. H., Summers, R. M., Giger, M. L., “Deep learning in medical imaging and radiation therapy,” *Med. Phys.* 46(1), e1–e36 (2019).
- [6] Wang, G., “A perspective on deep imaging,” *IEEE Access* 4, 8914–8924 (2016).
- [7] Wang, G., Cong, W. X., Yang, Q. S., “Tomographic image reconstruction via machine learning,” US 10,970,887 B2 (Provisional application No. 62 / 354,319, June 24, 2016), issued April 6, 2021, <https://patents.google.com/patent/US10970887B2/en?q=10970887> (2021).
- [8] Wang, G., Zhang, Y., Ye, X. J., Mou, X. Q., [Machine Learning for Tomographic Imaging], IOP Publishing Ltd (2019).
- [9] Shan, H., Padole, A., Homayounieh, F., Kruger, U., Khera, R. D., Nitiwarangkul, C., Kalra, M. K., Wang, G., “Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction,” *Nat. Mach. Intell.* 1(6), 269–276 (2019).
- [10] Wu, W. W., Niu, C., Ebrahimian, S., Yu, H. Y., Kalra, M., Wang, G., “AI-enabled ultra-low-dose CT reconstruction”, <https://arxiv.org/abs/2106.09834> (2021).
- [11] Chao, H., Shan, H., Homayounieh, F., Singh, R., Khera, R. D., Guo, H., Su, T., Wang, G., Kalra, M. K., Yan, P. K., “Deep learning predicts cardiovascular disease risks from lung cancer screening low dose computed tomography,” *Nat. Commun.* 12(1) (2021).
- [12] Raines, K. S., Salha, S., Sandberg, R. L., Jiang, H., Rodríguez, J. A., Fahimian, B. P., Kapteyn, H. C., Du, J., Miao, J., “Three-dimensional structure determination from a single view,” *Nature* 463(7278), 214–217 (2010).
- [12] Wei, H., “Fundamental limits of ‘ankylography’ due to dimensional deficiency,” *Nature* 480(7375) (2011).
- [14] Wang, G., Yu, H., Cong, W. X., Katsevich, A., “Non-uniqueness and instability of ‘ankylography,’” *Nature* 480(7375) (2011).
- [15] Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Židek, A., Nelson, A. W. R., Bridgland, A., Kavukcuoglu, K., Hassabis, D., “Improved protein structure prediction using potentials from deep learning,” *Nature* 577(7792), 706–710 (2020).
- [16] Akhtar, N., Mian, A., “Threat of adversarial attacks on deep learning in computer vision: A survey,” *IEEE Access* 6, 14410–14430 (2018).
- [17] Antun, V., Renna, F., Poon, C., Adcock, B., Hansen, A. C., “On instabilities of deep learning in image reconstruction and the potential costs of AI,” *Proc. Natl. Acad. Sci. U. S. A.* 117(48), 30088–30095 (2020).
- [18] Wu, W. W., Hu, D. L., Cong, W. X., Shan, H. M., Wang, S. Y., Niu, C., Yan, P. K., Yu, H. Y., Vardhanabhuti, V., Wang, G., “Stabilizing deep tomographic reconstruction networks”, <https://arxiv.org/abs/2008.01846> (v1, v2, and v3, 2020, v4, 2021).
- [19] Wu, E., Wu, K., Daneshjou, R., Ouyang, D., Ho, D. E., Zou, J., “How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals,” *Nat. Med.* 27(4), 582–584 (2021).

- [20] Chan, K. H. R., Yu, Y. D., You, C., Qi, H. Z., Wright, J., Ma, Y., “ReduNet: A white-box deep network from the principle of maximizing rate reduction” (2021).
- [21] Fan, F. L., Xiong, J. J., Li, M. Z., Wang, G., “On interpretability of artificial neural networks: A survey,” *IEEE Trans. Radiat. Plasma Med. Sci.*, 10.1109/TRPMS.2021.3066428 (2021).
- [22] Shamshirband, S., Fathi, M., Dehzangi, A., Chronopoulos, A. T., Alinejad-Rokny, H., “A review on deep learning approaches in healthcare systems: Taxonomies, challenges, and open issues,” *J. Biomed. Inform.* 113 (2021).
- [23] Montavon, G., Samek, W., Müller, K.-R., “Methods for interpreting and understanding deep neural networks,” *Digit. Signal Process. A Rev. J.* 73, 1–15 (2018).
- [24] Fan, F. L., Wang, G., “Fuzzy logic interpretation of quadratic networks,” *Neurocomputing* 374, 10–21 (2020).
- [25] Fan, F. F., Xiong, J. J., Wang, G., “Universal approximation with quadratic deep networks,” *Neural Networks* 124, 383–392 (2020).
- [26] Sabour, S., Frosst, N., Hinton, G. E., “Dynamic routing between capsules,” *Adv. Neural Inf. Process. Syst.* 2017-Decem, 3857–3867 (2017).
- [27] Fan, F. L., Lai, R. J., Wang, G., “Quasi-equivalence of width and depth of neural networks”, <https://arxiv.org/abs/2002.02515> (2020).
- [28] Nickel, M., Murphy, K., Tresp, V., Gabrilovich, E., “A review of relational machine learning for knowledge graphs,” *Proc. IEEE* 104(1), 11–33 (2016).
- [29] Wang, Q., Mao, Z., Wang, B., Guo, L., “Knowledge graph embedding: A survey of approaches and applications,” *IEEE Trans. Knowl. Data Eng.* 29(12), 2724–2743 (2017).
- [30] Liu, J. Zhu, X. X., Liu, F., Guo, L. T., Zhao, Z. J., Sun, M. Z., Wang, W. N., Lu, H. Q., Zhou, S. Y., Zhang, J. J., Wang, J. Q., “OPT: Omni-perception pre-trainer for cross-modal understanding and generation”, <https://arxiv.org/abs/2107.00249> (2021).
- [31] Wang, G., Zhang, J., Gao, H., Weir, V., Yu, H. Y., Cong, W. X., Xu, X., Shen, H., Bennett, J., Furth, M., Wang, Y., Vannier, M., “Towards omni-tomography – Grand fusion of multiple modalities for simultaneous interior tomography,” *PLoS One* 7(6) (2012).
- [32] Wang, G., Kalra, M., Murugan, V., Xi, Y., Gjestebj, L., Getzin, M., Yang, Q. S., Cong, W. X., Vannier, M., “Vision 20/20: Simultaneous CT-MRI – Next chapter of multimodality imaging,” *Med. Phys.* 42(10), 5879–5889 (2015).
- [33] Dineley, J., “Tackling the silent crisis in cancer care”, <https://www.lindau-nobel.org/blog-tackling-the-silent-crisis-in-cancer-care-with-innovation/> (2018).