



Experimental investigation of neural network estimator and transfer learning techniques for K-edge spectral CT imaging

Kevin C. Zimmerman, Gayatri Sharma, Abdul Kareem Parchur, Amit Joshi, and Taly Gilat Schmidt^{a)}

Department of Biomedical Engineering, Marquette University Medical College of Wisconsin, Milwaukee WI53233, USA

(Received 27 March 2019; revised 16 October 2019; accepted for publication 14 November 2019; published 6 January 2020)

Purpose: Spectral computed tomography (CT) material decomposition algorithms require accurate physics-based models or empirically derived models. This study investigates a machine learning algorithm and transfer learning techniques for Spectral CT imaging of K-edge contrast agents using simulated and experimental measurements.

Methods: A feed forward multilayer perceptron was implemented and trained on data acquired using a step wedge phantom containing acrylic, aluminum, and gadolinium materials. The neural network estimator was evaluated by scanning a rod phantom with varying dilutions of gadolinium oxide nanoparticles and by scanning a rat leg specimen with injected nanoparticles on a bench-top photon-counting computed tomography system. The algorithm decomposed each spectral projection measurement into path lengths of acrylic and aluminum and mass lengths of gadolinium. Each basis material sinogram was reconstructed into basis material images using filtered backprojection. Machine learning techniques of data standardization, transfer learning from aggregated pixel data, and transfer learning from simulations were investigated to improve image quality. The algorithm was compared to a previously published empirical method based on a linear approximation and calibration error look-up tables.

Results: The combined transfer learning techniques did not improve quantification in the rod phantom and provided only a small qualitative improvement in ring artifacts. Transfer learning from aggregated pixel data and from simulations improved the qualitative image quality of the rat specimen, for which the calibration data were limited. Transfer learning from aggregated pixel data and simulations estimated 3.26, 6.26, and 12.45 mg/mL Gd concentrations compared to true 2.72, 5.44, and 10.88 mg/mL concentrations in the rod phantom. Additionally, the neural networks were able to separate the soft tissue, bone, and gadolinium nanoparticles of the *ex vivo* rat leg specimen into the different basis images.

Conclusions: The results demonstrate that empirical K-edge imaging from calibration measurements using machine learning and transfer learning is possible without explicit models of material attenuations, incident spectra, or the detector response. © 2019 American Association of Physicists in Medicine [https://doi.org/10.1002/mp.13946]

Key words: contrast agent, machine learning, material decomposition, photon-counting, spectral CT

1. INTRODUCTION

Artificial intelligence and machine learning methods have been growing at a rapid pace in many industries, especially in the medical imaging field. Several medical imaging fields of study have had machine learning methods applied to them including computer-aided diagnosis, reconstruction, image denoising, preprocessing, among others.^{1–4} The broad deployment of machine learning techniques in medical imaging has given researchers a new perspective on solving the complicated problems in the field.

Energy-resolved photon-counting detectors enable spectral computed tomography (CT) and continue to show promising developments.⁵ Several research prototypes have been developed, and studies comparing photon-counting CT scanners to those with conventional energy integrating detectors have been performed.^{6,7} As the sensor and electronic technologies continue to advance, the clinical advantages of photon-

counting CT, such as lower dose and increased contrast and material separability, may surpass that of scanners with conventional detectors.

An important application of energy-resolved photon-counting x ray or CT acquisition is material decomposition, that is, the ability to estimate the amount of certain materials that the x rays traveled through before being detected.⁸ This is possible due to the multiple spectral measurements acquired simultaneously by the detector. In contrast, conventional detectors measure the integrated energies of the incident spectra creating the possibility for different materials and thicknesses to have the same measured value. Spectral CT with conventional detectors can be performed with multiple spectral measurements by repeated acquisitions, modifications to the x-ray source and generator, modifications to the detector, or some combination thereof. Photon-counting detectors can benefit spectral CT applications by simultaneously acquiring more than two spectral measurements with

potentially higher signal-to-noise ratio compared to conventional detectors due to improved detection efficiency.⁹

K-edge imaging, or the detection and quantification of contrast agents with heavy elements, is made possible with photon-counting detectors due to their ability to acquire more than two spectral measurements and the ability to tune the energy bin content. Nanoparticles with gold, iodine, or gadolinium have been investigated for tumor detection,^{10,11} cardiovascular imaging,^{12,13} and theranostic applications.¹⁴

Material decomposition from photon-counting CT data can be performed from either the acquired spectral projection measurements or from the reconstructed spectral images.^{15,16} Decomposition from the projection measurements has the benefit of compensating for beam hardening effects. If the measurement model and statistics are known, projection-domain decomposition can be performed using an optimization-based algorithm, such as maximum likelihood.¹⁷ However, it is difficult to accurately model the detector response and pulse pileup effects, especially with pixel variations associated with the semiconductor crystal heterogeneity and electronics.

Empirical material decomposition methods utilize actual system measurements that tune model parameters for improved estimation.^{18–21} These methods require a model of the relationship between the basis material path lengths and the spectral measurements and also require calibration measurements to estimate the model parameters. It can be difficult for a parametric model to accurately represent the complex photon-counting measurement process, therefore additional correction steps have been proposed.¹⁹ Practically, the number of measurements for empirical calibrations should be minimal. Feed forward neural networks are a highly parameterized model that can approximate any continuous function.²² Our previous work developed a neural network estimator for two-material decomposition from photon-counting CT data.¹⁸ The method provided accurate decomposition values, but resulted in prominent ring artifacts due to each detector channel having an independently trained neural network with varying bias for different basis material path lengths.

Three-material decomposition that includes a K-edge basis material poses additional challenges compared to two-material decomposition. Adding a third basis function increases the instability of the estimation problem, increasing the susceptibility to noise and inaccuracies in the decomposed values. Also, the signal from the K-edge signature of the contrast agent is small, due to low concentrations of the contrast material. In addition, three-material decomposition theoretically increases the number of required calibration measurements, which could pose practical challenges for empirical material decomposition. Our previous study of neural network material decomposition did not investigate this more challenging case of three-material decomposition.

In this study, we used machine learning techniques such as data standardization and transfer learning to develop a neural network that can perform three-material, K-edge imaging from photon-counting data. The network was evaluated experimentally on a bench-top photon-counting CT scanner

using a rod phantom containing varying dilutions of gadolinium oxide nanoparticles and a rat leg specimen with a nanoparticle gel injection.

2. MATERIALS AND METHODS

2.A. Theory

A photon-counting measurement consists of a set of photon counts n_k in each energy bin, k , along a ray path, l , that terminates on a detector element. When neglecting the effect of pileup, the photon-counting measurement model can be represented as,

$$n_{k,l} = n_{k,0} \int S(E) \left[\int_{U_k}^{U_{k+1}} R(E, U) dU \right] \exp \left[- \int_l \mu(E, \vec{r}) dl \right] dE \quad (1)$$

where $n_{k,0}$ is the number of photons in the k^{th} energy bin for an air scan, $\mu(E, \vec{r})$ are the spatially and energy varying linear attenuation coefficients of the object being scanned, $R(E, U)$ is the probability that a photon of energy E is measured at energy U , $[U_k, U_{k+1})$ are the thresholds that define the k^{th} energy bin window, $S(E)$ is the energy distribution of the source spectrum, and $n_{k,l}$ is the number of photons detected in the energy bin window along the ray l through the object.

The detector deviates from its ideal behavior at both the sensor and electronic level. The detector response, $R(E, U)$, incorporates different semiconductor and electronic effects such as finite energy resolution, charge trapping, fluorescence, K-edge absorption, etc.

A line integral through an object, $\int_l \mu(E, \vec{r}) dl$, can be approximated as the unique path lengths of some basis materials.⁸ Suppose the line integral through an object contains a contrast agent, such as gadolinium, which has a unique K-absorption edge, as shown in Fig. 1. When Compton scatter, photoelectric absorption, and the K-absorption edge contribute to the attenuation, the approximation contains three basis functions,

$$\int_l \mu(E, \vec{r}) dl \approx f_1(E)x_1 + f_2(E)x_2 + f_3(E)x_3. \quad (2)$$

Projection-domain decomposition is the process of estimating the basis function coefficients, $\vec{x} = [x_1, x_2, x_3]$, from each spectral measurement, $n_{k,l}$. When all spectral CT measurements are decomposed, the result is three basis sinograms. In this work, we assume f_1 and f_2 are the linear attenuation coefficients of two basis materials without K-edges in the scanned energy range, for example acrylic and aluminum. We assume f_3 is the mass attenuation coefficient of the contrast agent element with the K-edge signature, such as gadolinium. By assuming linear attenuation coefficients for the first two basis materials and mass attenuation coefficient for the K-edge material, the estimated basis material components are $\vec{x} = [x_1, x_2, \rho_3 x_3]$, where x_1 and x_2 represent the pathlength through the two basis materials, and $\rho_3 x_3$ is

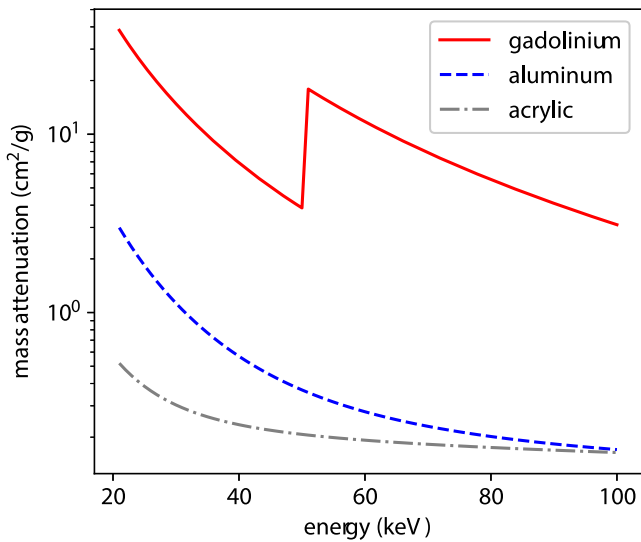


FIG. 1. Mass attenuation coefficients of the basis materials used in this study: gadolinium, aluminum, and acrylic. The discontinuity above 50 keV in the gadolinium attenuation is the K-absorption edge and is a unique characteristic of gadolinium. [Color figure can be viewed at wileyonlinelibrary.com]

the mass length of the K-edge contrast agent. Because we expect the density of the K-edge contrast agent to vary, we must use the mass length to estimate the K-edge material. The mass length sinogram is reconstructed into a basis image in units of density (concentration). The other basis sinograms are reconstructed into unitless basis images representing the contribution of each material in each voxel. The process of estimating an image representing the density of a K-edge material (e.g., gadolinium) in each image pixel (or voxel) is referred to as K-edge imaging.

Material decomposition can be considered a transformation from a measurement space, $\vec{n} \in \mathcal{N}$, to a basis material space, $\vec{x} \in \mathcal{R}$, and is an inverse problem. In this work, we use a neural network for K-edge material decomposition, specifically, a feed forward fully connected shallow multilayer perceptron. The neural network is trained to learn the inverse transformation between the spectral measurements, \vec{n} , and basis material components, \vec{x} .

2.B. Counts data preprocessing

There is an exponential relationship between the basis material components to be estimated, \vec{x} , and the photon counts measurements, n_k . Rather than have the neural network learn this exponential relationship, log-normalization with an air measurement is performed on each energy bin measurement, similar to conventional detector processing. Therefore,

$$p_k = -\ln \frac{n_k}{n_{k,0}} = \ln(n_{k,0}) - \ln(n_k) \quad (3)$$

where p_k are the components of the vector \vec{p} , which is the log-transformed measurement, and $n_{k,0}$ is the number of photons counted by bin k for a measurement through air.

2.C. Training data standardization

The neural network undergoes a supervised training process whereby known spectral measurements, \vec{p}^i , resulting from known basis parameters, \vec{x}^i , are used as the network inputs and targets, respectively. Training is an iterative process that starts by evaluating each spectral input using the initial network weights and compares the network prediction to the specified target. The error is calculated between the network prediction and the specified target value and is used to update the network weights, thereby creating a network that will produce a smaller error in a future prediction. This iterative process continues through the input/target pairs of the training data until convergence criteria are satisfied. The neural network training process should consider errors in each component of the output as being equally undesirable. Data standardization by mean removal and variance scaling achieves this goal.

The outputs of the neural network estimator are basis material parameters, \vec{x} , which consist of the path lengths of aluminum and acrylic and the mass length of gadolinium. Depending on the data used for training, the range of the magnitude of each component is likely to be different. For example, this study measures training data from acrylic up to 10.16 cm, aluminum up to 2.54 cm, and gadolinium mass length up to 79 mg/cm². When the magnitudes of the components are different, the larger basis parameters will dominate the cost function during training, and the network will preferentially favor accuracy of that basis parameter. Therefore, the mean is subtracted and the variance of the training data is scaled to unity for training. The network estimates are then scaled back to their original range for predictions after training. This is effective in giving equal treatment to all basis materials during training.

2.D. Network architecture and parameter selection

Neural networks have several hyperparameters which must be chosen as part of the design. These hyperparameters can include the optimization algorithm, the number of hidden processing elements, the number of hidden layers, etc. A fully connected network with a single hidden layer was chosen because it has previously shown promising results in material decomposition^{18,23} and the universal approximation theorem²² suggests this neural network model may be sufficient for the material decomposition problem. The network architecture is illustrated in Fig. 2. The inputs to the network are the log-normalized energy bin measurements, p_k , and the outputs are the estimated scaled basis material components, \vec{x}_{scaled} . The neural network outputs, \vec{x}_{scaled} , are scaled back to their original ranges to produce the final estimate of the basis material components, \vec{x} .

The Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm²⁴ was chosen for training because it is more suited for smaller datasets and the training dataset used in this study was relatively small, see Section 2.F.2. A leave-one-out cross-validation method²⁴ determined that 100 hidden

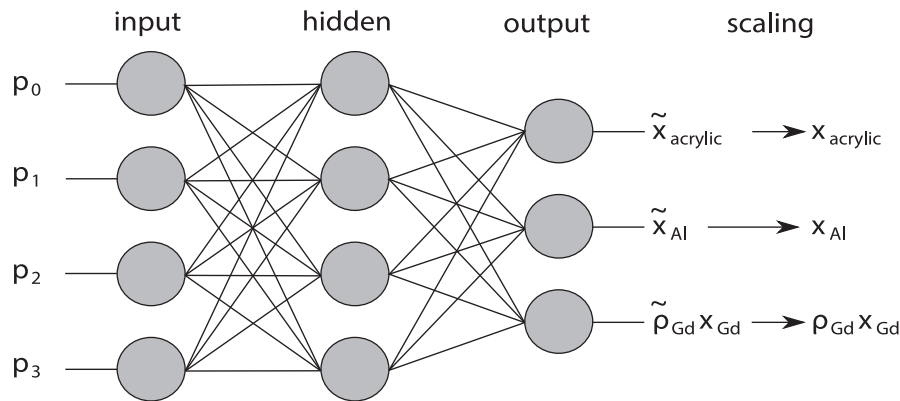


FIG. 2. Feed forward neural network architecture with four energy bin inputs from a single spectral measurement, three scaled basis material outputs, and one hidden layer. The network outputs are scaled back based on scaling parameters determined from the training dataset. The network used in this study had 100 processing elements in the hidden layer (the illustration only shows four).

processing elements and L2 regularization penalty of 0.001 provided good accuracy without overfitting effects in the validation data.

The cost function that is minimized during training is non-convex and multimodal. The training algorithm does not guarantee global convergence and is sensitive to the initial network weights. Using the same training data, a neural network may perform differently with different random initial weights.

2.E. Transfer learning

Neural network material decomposition previously demonstrated sensitivity to ring artifacts due to variations across photon-counting detector pixels.¹⁸ Another practical challenge of applying neural network material decomposition to K-edge imaging is the need for additional calibration measurements containing the K-edge basis material. There is a trade-off between the number of calibration measurements and the acquisition time and complexity of the calibration phantom. This study investigates whether two machine learning transfer learning techniques can overcome these challenges of neural network K-edge material decomposition.

Transfer learning is a machine learning technique in which a neural network is pre-trained on a different set of data prior to training on the intended dataset. Therefore, the initial network parameters are the result of a previous training rather than random values. Transfer learning can reduce the training time, can accommodate fewer and noisier training data, and can result in more consistent network performances. Two types of transfer learning are proposed in this study: transfer learning from aggregated pixel data and transfer learning from a simulated detector.

2.E.1. Transfer learning from aggregated to individual pixels (M1)

The detector response varies between detector pixels due to effects such as sensor heterogeneity, pixel variations in electronic dead time, and nonideal energy bin threshold

calibration. Therefore, each pixel will have a different spectral measurement for the same basis material parameters. For this reason, a separate neural network should be trained for each detector pixel. However, variations in the bias of the networks across pixels can lead to ring artifacts.

The magnitude of the ring artifacts depend on the difference between estimates from one pixels network and its neighbor's. For example, if all pixels returned the same biased estimate, the image would be free of ring artifacts but quantitatively inaccurate. Pixels that have the same magnitude of bias but different polarity of bias could introduce ring artifacts even if the bias was small. Here, we propose a transfer learning method to promote similar biases across pixels by starting each individual pixels training with a network that has already been trained to the average pixel response.

More specifically, we investigated a transfer learning method to overcome the challenges of a small calibration dataset and pixel to pixel variations. The calibration measurements from all detector pixels were aggregated to train a neural network from random initial weights. This network learns the average response of the detector pixels. The learned network parameters are used as the initial weights when training specific networks using the calibration data for the individual pixels. This transfer learning promotes neural network models that are similar in characteristics (i.e., biases) but are fine-tuned for the given pixel's unique detector response inferred from the calibration data. We refer to this approach of transfer learning from the aggregated pixel data to individual pixel data as method M1.

The method of transfer learning from the aggregated pixel data to individual pixel data is compared to a baseline approach (referred to as Baseline) which consists of training a neural network for each pixel from random initial weights using only the calibration measurements measured by the pixel.

2.E.2. Transfer learning from simulations (M2)

The additional dimension representing the K-edge material requires more calibration measurements, increasing the

acquisition time and complexity of the calibration phantom. Depending on the calibration phantom used, the number of unique basis material combinations sampled may be limited. Transfer learning from simulated data is proposed to enable good network performance from a relatively small experimental training data size. A simple detector response, illustrated in Fig. 3 was simulated and used in Eq. (1) to generate many simulated measurements from known basis material parameters. The detector response model is similar to those in other studies.^{4,17} The detector response consists of the fluorescence escape and capture peaks of cadmium and tellurium, an energy resolution of 3 keV full-width at half-maximum (FWHM) for all incident energies, and a constant background to model charge sharing and cross-talk effects. Using simulations, a large training dataset can be generated to more extensively sample the three-material decomposition space compared to the experimental training dataset. However, the simulated detector response is only an approximate model for the detector, and does not include pixel-to-pixel variations or pulse-pileup effects. The simulated model does not accurately represent the experimental detector but has similar characteristics. Therefore, we propose an initial training step to learn the approximate model from a large set of simulated calibration data, followed by additional training from the limited experimental calibration measurements.

More specifically, a network is first trained using the simulated data only. Then transfer learning is used to update the network through additional training using the experimental calibration measurements from the aggregation of all pixel data, which transfers the network from the simulated detector response to the average experimental detector response. During this additional training, the network learns the differences between the simulated and an actual detector. Finally, the network is customized for each detector pixel by additional training using only that pixel's calibration measurements (method M1), to accommodate the additional variations across pixels. We refer to this approach of combining transfer learning from simulations and from the aggregated pixel data to the individual pixel data as method M2.

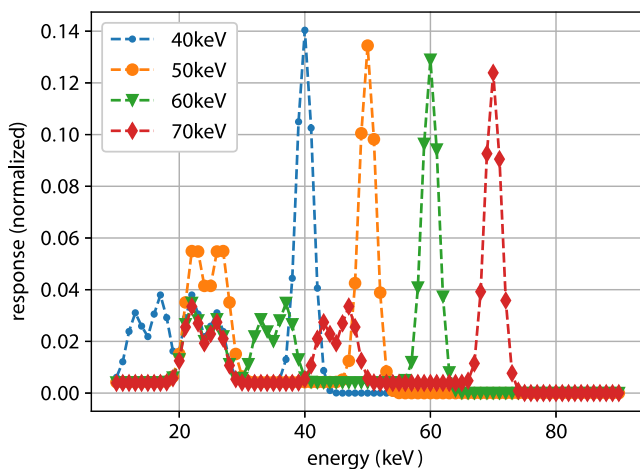


FIG. 3. Simulated detector response function, $R(E,U)$ for a few incident energies. [Color figure can be viewed at wileyonlinelibrary.com]

2.F. Experimental study

2.F.1. Photon-counting CT system

The bench-top energy-resolved CT system consists of a microfocus x-ray tube (L9181-02, Hamamatsu), and a CdTe photon-counting detector (DxRay Inc., Northridge CA) with 1 mm detector elements in a 4×64 array located approximately 73 cm from the source. The detector has four programmable energy bin thresholds, which were placed at 25, 51, 57, 65 keV. The x-ray source was operated at a tube voltage of 90 kVp. Every CT projection or x-ray calibration transmission measurement was acquired at a tube setting of 0.6 mAs with a count rate of approximately 140 kcps/mm² through air. CT acquisitions were performed by acquiring 120 view angles over 360°. Data from the fourth row of the detector were used in this study, though all rows yielded similar results.

2.F.2. Step wedge phantom

The network must be trained from data acquired using known basis material components that cover the extent of components that will be encountered within the scanned objects. For this study, the non-K-edge basis materials (μ_1 and μ_2) were chosen to be aluminum and acrylic because they are readily available, easy to machine, and their effective atomic numbers and electron densities are significantly different.

Training data were obtained by acquiring energy bin transmission measurements through a step wedge phantom that consisted of 25 combinations of aluminum (0–2.54 cm in steps of 0.635 cm) and acrylic (0–10.16 cm in steps of 2.54 cm).

Each of the 25 steps was imaged with and without a 0.1-mm thick gadolinium foil in the x-ray beam path, for a total of 50 combinations of basis material parameters. An illustration of the step wedge phantom with the gadolinium foil is shown in Fig. 4. The training data did not have repeated measurements in order to reduce calibration time.

2.F.3. Synthesis of $Gd_2O_3:Nd$ Nanoparticles

A modified protocol based on prior literature was employed to synthesize Gd_2O_3 nanoparticles doped with 1% Nd nanoparticles for NIR emission^{25,26}: 1 mL gadolinium (III) nitrate hexahydrate (1 M) containing 1% of neodymium (III) nitrate hexahydrate and 10 g polyvinylpyrrolidone (MW = 2000) was dissolved in 2 L water, heated at 80°C for 30 min, and then urea (200 mm) was added. The reaction was continued for 1 h via further heating and then the reaction mixture was allowed to cool to room temperature. Particles were collected by centrifugation and then allowed to dry at room temperature. Resulting particles were then annealed at 600°C for 1 h.²⁵ Finally, Gd_2O_3 nanoparticles (300 mg) were ground using an agate mortar pestle and dispersed by sonication in 100 mL of NaOH (5 mm) solution for 10 min followed by neutralization with HCl. These particles further

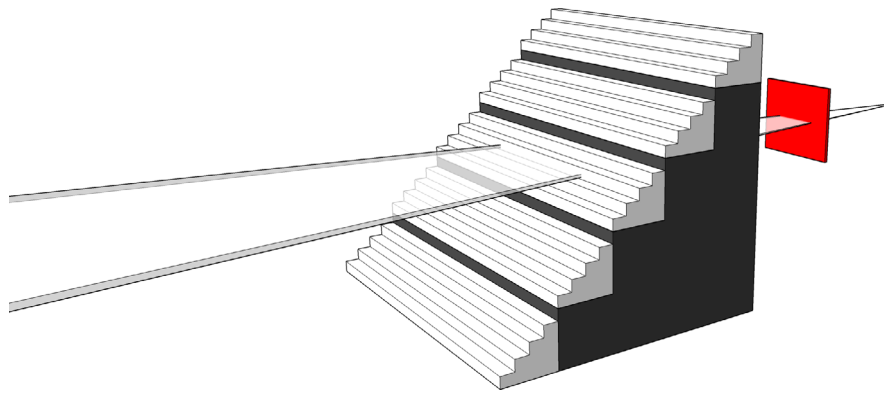


FIG. 4. An illustration of the step wedge phantom consisting of aluminum (white) and acrylic (black) path length variations at each step and the gadolinium foil (red) in the beam path. The step wedge phantom is translated so the beam path goes through each step incrementally. [Color figure can be viewed at wileyonlinelibrary.com]

diluted tenfold and sonicated for 1 h, followed by centrifugation at 600 g for 5 min to eliminate the large aggregates. The suspension was then centrifuged at 3,000 g for 15 min.²⁶ The size and charge of Gd_2O_3 nanoparticles were characterized using Malvern Zetasizer Nano ZS (Malvern Instruments Ltd.). The hydrodynamic size and zeta potential of the particles was found to be 144 nm and -42 mV, respectively. NPs demonstrated strong NIR emission at 1064 nm under 808 nm excitation.

2.F.4. Rod phantom

A rod phantom was used to evaluate the neural network estimator performance for different densities of Gd_2O_3 nanoparticles. The phantom consisted of a 6.35 cm diameter acrylic body and four, 1.9 cm diameter rods of different materials. One rod consisted of Teflon, with the other three rods consisting of 3.125, 6.25, and 12.5 mg/mL dilutions of Gd_2O_3 in glass vials. Figure 5 illustrates the rod phantom components. The rod phantom was placed 47 cm from the source. During CT acquisition, the detector was translated to two positions for each view to image the full phantom extent.

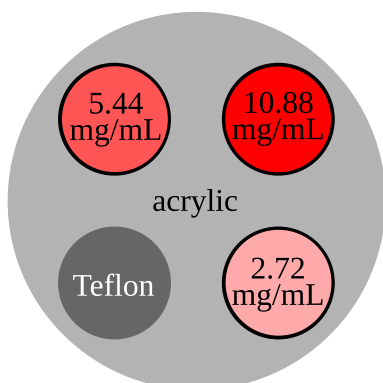


FIG. 5. An illustration of the rod phantom imaged in this study. The rod phantom consists of a 6.35-cm diameter acrylic body with a 1.9-cm Teflon rod insert and 3 glass vials containing low, medium, and high densities of gadolinium oxide nanoparticles of 2.72, 5.44, and 10.88 mg/mL Gd, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

2.F.5. Ex vivo tumor model

An *ex vivo* bone-adjacent tumor mimicking phantom was generated from an excised hind limb from a Wistar rat. The tumor-like scaffold on the rat leg adjacent to femur was prepared by composite hydrogels (alginate and gelatin). The alginate solution (8%, w/v) and gelatin solution (6%, w/v) were dissolved in phosphate-buffered saline. The final homogeneous composite hydrogel was prepared by mixing the hydrogels in final concentration of alginate solution (7%, w/v) and gelatin solution (3%, w/v) with 0.1 mL of $\text{Gd}_2\text{O}_3\text{:Nd}$ nanoparticles. This mixture was then loaded into a sterilized syringe with 18-gauge needle and injected in the rat leg. After injection, the scaffold was allowed to settle for 10 min and then used for further imaging experiments.

2.F.6. Data processing and analysis

For all CT acquisitions, the count sinograms were log-normalized (Section 2.B) and decomposed ray-by-ray using the decomposition methods specified in (Table I). After estimation by the neural networks, the basis material sinograms were reconstructed into basis material images using a filtered backprojection algorithm.

To evaluate the performance of the proposed method, regions of interest (ROIs) were placed within the three nanoparticle rods in the gadolinium image and the Teflon rod. The sample means of the nanoparticle rods were compared to the nominal gadolinium density which is approximately 87% of the gadolinium oxide density. The ROI placed within the Teflon rod in each of the three basis images was compared to the ground truth acrylic, aluminum, and gadolinium contributions obtained from a least squares fitting of their linear attenuation coefficients provided by the NIST database.²⁷ The ROI sample means ($n = 498$) between different neural network methods were compared using two-sample *t*-tests with a level of significance of 5%. The null hypothesis tested was the equivalence of population means.

The simulated training dataset consisted of 8000 combinations of aluminum (0–2.54 cm in steps of 0.134 cm) and

TABLE I. A summary of the compared methods and their abbreviated names.

Abbreviation	Summary
Baseline	Train individual neural networks for each pixel using calibration data from each pixel
M1	Train a network using the calibration data aggregated from all pixels and use transfer learning from the single network to train individual networks for each pixel using calibration data from each pixel.
M2	First train a network from simulated data only. Then use transfer learning to further train a network from the aggregated pixel data and finally to the individual pixel using method M1.
SimOnly	Train a single neural network using calibration data from a simulated detector.
Atable	A linear model approximating the basis material parameters and the spectral measurements. The basis material parameter estimates are corrected by interpolating error look-up tables created from the calibration measurements. ¹⁹

acrylic (0–10.16 cm in steps of 0.535 cm) and gadolinium mass lengths (0 g/cm² to 79 mg/cm² in steps of 4.16 mg/cm²).

The approaches in this study were then compared to another empirical decomposition algorithm, the A-table method, proposed by Alvarez.¹⁹ The method assumes a linear approximation between basis material parameters and spectral measurements. The errors in the linear approximation using the calibration data are stored in multidimensional look-up tables and interpolation of the errors is subtracted from the linearly approximated basis material parameters to compute the final estimates.

3. RESULTS

Figure 6 displays the reconstructed aluminum, acrylic, and gadolinium rod phantom basis material images from the decomposition methods described in Table I. The Baseline, M1, and M2 methods produced similar qualitative images of the rod phantom. Method M1 did not reduce the ring artifact as expected. The SimOnly method (fourth row) incorrectly distributed the glass vials and Teflon across the basis material images, demonstrating the approximate nature of the simulation model.

The gadolinium rod estimates in the Baseline and M2 methods were similar, with no significant difference in their means (low density: $P = 0.589$; medium density: $P = 0.342$; high density: $P = 0.123$). However, there were differences in the accuracy of the Teflon rod decomposition between the Baseline and M2 methods. The Baseline method was more accurate in estimating the contribution of acrylic ($P < 0.05$) but the M2 method was more accurate in estimating the aluminum ($P < 0.05$) and gadolinium ($P < 0.05$) contributions.

Similarly, the ROI sample means were compared between the transfer learning from simulations (M2) and the Atable method. Although the Atable method has more image artifacts, the ROI sample means were significantly lower and

closer to the ground truth than the transfer learning method for all concentration nanoparticle rods ($P < 0.05$). The sample means of the Teflon rod for the M2 method were significantly closer to the ground truth in the acrylic and aluminum images compared to the Atable method ($P < 0.05$). However, the sample mean of the Teflon rod for the Atable method in the gadolinium image was significantly closer to the ground truth compared to method M2 ($P < 0.05$). The ROI sample means and standard deviations of the compared methods can be seen in Fig. 7.

The acrylic, aluminum, and gadolinium basis images of the rat leg specimen are displayed in Fig. 8. The enhanced region in the gadolinium image of the neural network method represents the injected Gd₂O₃ gel. Transfer learning from combined pixel data (M1) reduced the ring artifacts in the rat leg images compared to the Baseline method. Transfer learning from simulations (M2) further reduced the ring artifacts at the center of the image and in the gadolinium basis image compared to method M1. The Atable method did not depict bone contribution in the aluminum image or nanoparticle contribution in the gadolinium image, demonstrating material decomposition error.

A photon-counting image, which is the image reconstructed from the sum of all detected counts, is compared to virtual monoenergetic images (VMI) synthesized at 50 keV for the neural network and Atable methods and is illustrated in Fig. 9. The transfer learning methods appear to remove the ring artifacts in the rat monoenergetic image. The SimOnly method is able to depict the rat anatomy in the VMI despite the material decomposition errors in Fig. 8. Due to the material decomposition errors shown in Fig. 8, the Atable VMI has a dark area of hypo-attenuation in the region where the nanoparticle gel is located. The photon-counting image is not a monoenergetic image and is therefore expected to have a different contrast level compared to the VMIs. It is included as a reference for the expected cross sectional anatomy of the rat leg specimen.

4. DISCUSSION

The ability to detect Gd₂O₃:Nd nanoparticles via spectral CT will have a direct impact on interventional radiology (IR)-based procedures for tumor treatment with theranostic nanoparticles which include multiple contrasts such as NIR luminescence and T1-weighted MR imaging provided by rare-earth-doped Gd₂O₃.²⁸ NIR and MR imaging can be utilized for therapy planning and optimization studies on small animals, while the CT contrast can be included in standard IR workflows for image guided thermal ablation procedures by clearly visualizing Gd-contrast accumulating lesions during IR procedures. Further extension of the proposed methods in identifying other elements such as Au, will enable clinical translation of emerging therapies such as image guided plasmonic photothermal ablation.^{28–30}

This study proposed two transfer learning techniques to address the challenges of pixel variations in photon-counting

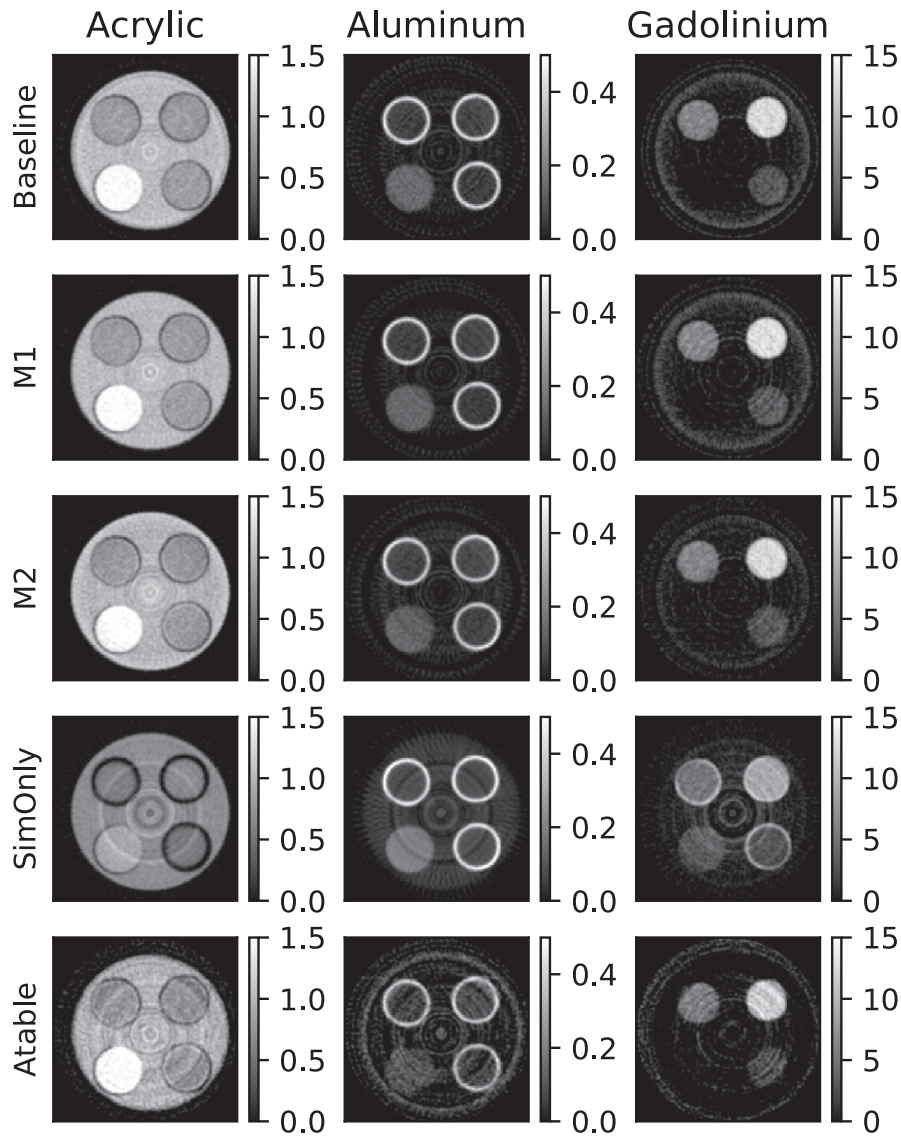


FIG. 6. Reconstructed basis material images of the rod phantom using the various decomposition methods described in Table I.

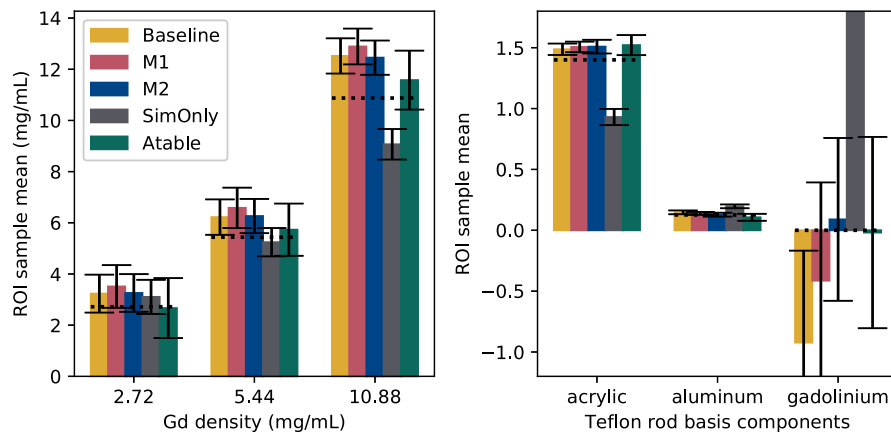


FIG. 7. A comparison of the ROI sample means from the various decomposition methods described in Table I. The dotted line represents the ground truth values. The error bars represent one sample standard deviation. [Color figure can be viewed at wileyonlinelibrary.com]

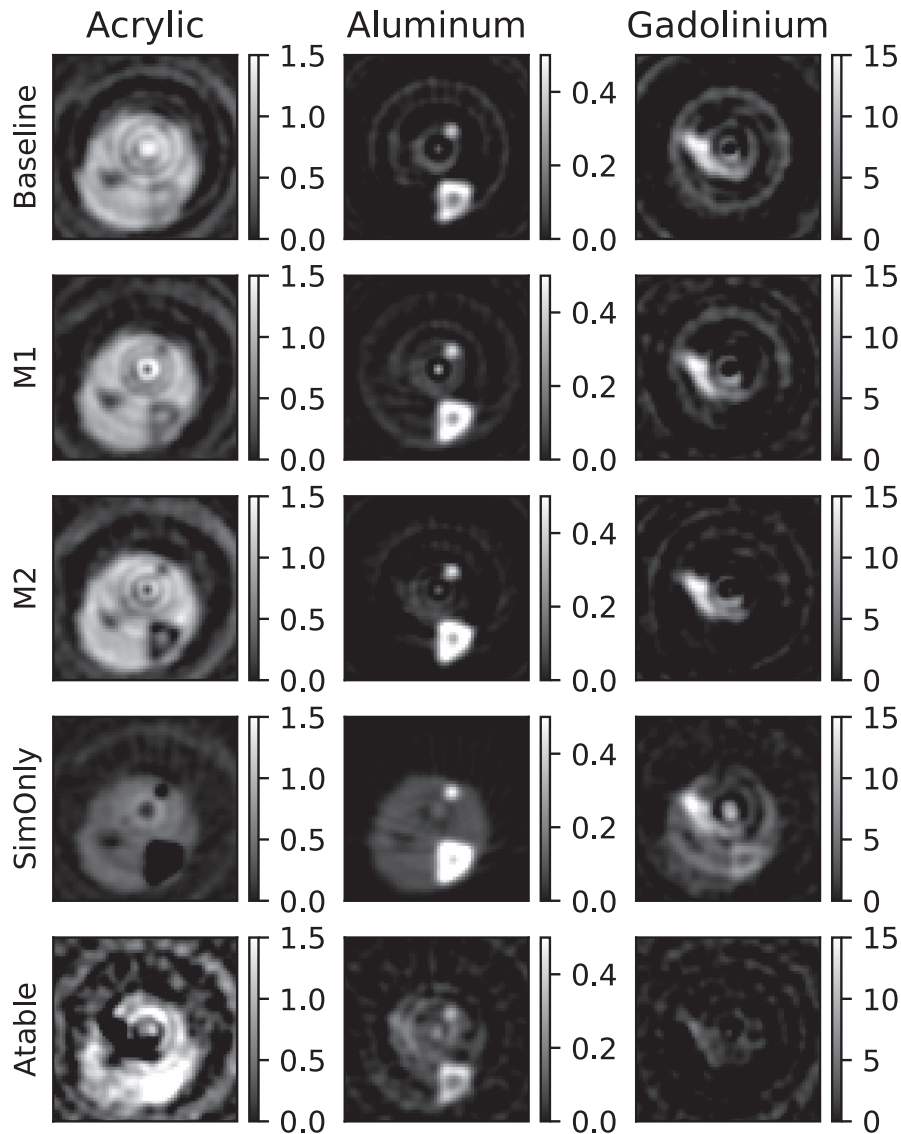


FIG. 8. Reconstructed basis material images of the rat leg specimen using the various decomposition methods described in Table I. The nanoparticle injection is visible in the gadolinium basis image as the area of hyperattenuation.

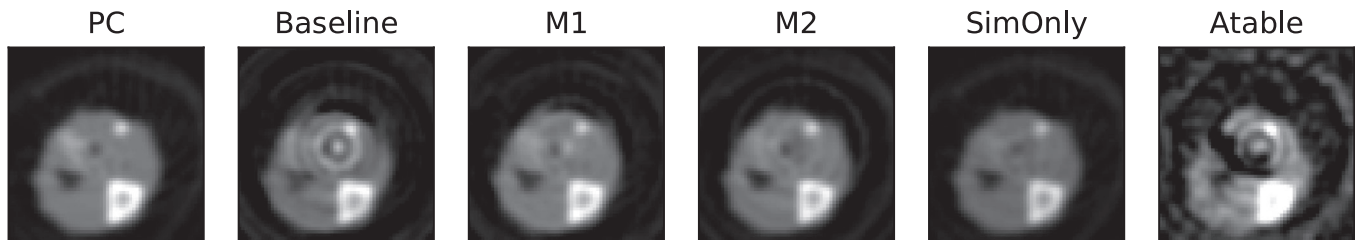


FIG. 9. A photon-counting image of the rat leg specimen compared with synthesized virtual monochromatic images of the rat leg specimen at an energy of 50 keV using the various decomposition methods described in Table I. (WL/WW 500/3000 HU)

detectors and the challenges of acquiring large training datasets in practice. Transfer learning from the aggregated pixel data effectively reduced ring artifacts in the rat leg specimen images but not rod phantom images (Figs. 6 and 8). The rat leg specimen, with an approximate diameter of 1 cm, contains a small range of basis material path lengths compared to

the larger extent of the calibration step wedge phantom. Transfer learning from simulations provided more training data at the smaller path lengths relevant to the rat leg specimen, which likely caused the improved gadolinium identification in Fig. 8. The rat leg specimen results suggest that transfer learning from simulations and from aggregated pixel

data can be a beneficial supplement to an experimental training dataset for objects for which relevant calibration data is limited.

The Atable had statistically lower and closer sample ROI means to the ground truth concentration in the nanoparticle rods than the neural network method M2. The Teflon rod was less accurately decomposed into the acrylic, aluminum, and gadolinium images compared to the neural network methods. The Atable method had noticeable artifacts in the rod phantom and was unable to separate the bone and gadolinium in the rat phantom. The Atable method is most accurate when the basis material parameters are close to the ones in the calibration data. The rat specimen results suggest that the Atable method has limitations in generalizing away from the calibration points, while the neural networks have the benefit of improved generalization when experimental calibration data are limited.

Quantitative photon-counting K-edge imaging was previously demonstrated using a projection-domain maximum likelihood algorithm and a detector response model developed from synchrotron measurements.¹⁷ This previous study imaged higher concentrations of gadolinium (11–40 mg/mL) compared to the current study and demonstrated errors of around 9 mg/mL. In contrast, the neural network methods developed in this study used a set of 50 calibration transmission measurements and was able to estimate gadolinium concentrations as low as 3 mg/mL with errors <1 mg/mL. One benefit of the neural network approach, compared to maximum likelihood or other optimization-based material decomposition algorithms, is that the decomposition requires a small number of arithmetic operations.

One limitation of this study is that calibration data was only acquired at one noise realization, which occurred because of practical limitations with this specific experimental bench-top system. The quantitative accuracy and ring artifacts may be further improved with more calibration datasets. The improvement provided by transfer learning may be reduced as more calibration data is acquired. There may be cases in practice where calibration time is limited. The results of this study demonstrate accurate K-edge material decomposition despite the limited calibration dataset.

The gadolinium oxide nanoparticles in the rod phantom were diluted from a 50 mg/mL solute with distilled water. The nominal densities were shown in Fig. 5 but there is some uncertainty in the dilution process. In addition, the nanoparticles may have settled in the glass vials during the imaging process, causing inhomogeneities within the vials. For these reasons, there is some uncertainty in the ground truth concentrations of the rod phantom nanoparticles, which may contribute to the bias in the ROI estimates.

Scattered radiation can affect material decomposition but was neglected in this study. The imaging system did not utilize a post-object antiscatter grid. A pre-object lead slit collimator was used to collimate the beam in the cone direction only, effectively illuminating only the four rows of the

detector. The fan angle was not collimated. However, scatter in the fan direction is believed to be small because of the small extent of the imaged objects.

This study did not use a bowtie filter in the imaging system. However, the methods presented in this paper should be effective in the presence of a bowtie filter because each pixel is trained individually using calibration measurements.

The measured spectrum for a given detector element is a result of the spectra incident on that detector element and its neighbors.³¹ However, the neural network developed in this study estimates the basis material components for each ray independently. Furthermore, the neural network was trained from step wedge data where all detector elements imaged the same object simultaneously resulting in a similar number of counts detected in all pixels. When acquiring data of an arbitrary object, neighboring pixels at an object boundary will be exposed to different numbers of photons, changing the amount of charge sharing between the two pixels. In the future, a more robust neural network estimator could be developed to compensate for charge sharing effects by considering the measurements of the detector element of interest as well as its neighbors.

The neural network was both trained on step wedge data and evaluated with phantom and rat specimen data acquired using the same scanning parameters (tube voltage, tube current, integration time, etc.). The neural network model implicitly learns the relationship between the measurements and the basis material parameters. Therefore, the network is specific to the scanning parameters used in training and requires retraining for a different set of scan parameters. An interesting area of future work is to investigate whether a single network model could be generalized to estimate basis material parameters from many different scanning parameters provided sufficient data for training. The ability to have a general neural network decomposition model for many typical tube output and filter combinations would be beneficial in a clinical product. Theoretically, an accurate physics-based model could overcome this limitation, if one could be developed. This limitation is the focus of future work.

5. CONCLUSIONS

This study investigated neural network methods using transfer learning for K-edge imaging of gadolinium oxide nanoparticles. The networks were trained using data acquired from a step wedge phantom containing path lengths of aluminum, acrylic, and gadolinium basis materials. The transfer learning techniques did not result in significant quantitative improvement for the rod phantom. Transfer learning from aggregated pixel data and from simulations improved the qualitative image quality of the rat specimen, for which representative calibration data was limited. The results of rod phantom and rat specimen experiments demonstrate that quantitative K-edge imaging using machine learning is possible with photon-counting x-ray detectors.

ACKNOWLEDGMENTS

This work was supported by NIH Grant number R21EB015094. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health. Amit Joshi and Abdul Parchur acknowledge the support by MCW Cancer Center: State of Wisconsin Tax-check off project. Finally, we thank the reviewers for their feedback which improved the methods of this paper.

CONFLICT OF INTEREST

The authors have no conflict to disclose.

^{a)} Author to whom correspondence should be addressed. Electronic mail: tal.gilat-schmidt@marquette.edu.

REFERENCES

- Hoo-Chang S, Roth HR, Gao M, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging*. 2016;35:1285.
- Jin KH, McCann MT, Froustey E, Unser M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans Image Process*. 2017;26:4509-4522.
- Kang E, Min J, Ye JC. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. *Med Phys*. 2017;44:e360-e375.
- Touch M, Clark DP, Barber W, Badea CT. A neural network-based method for spectral distortion correction in photon counting x-ray CT. *Phys Med Biol*. 2016;61:6132.
- Willeink MJ, Persson M, Pourmorteza A, Pelc NJ, Fleischmann D. Photon counting CT: technical principles and clinical prospects. *Radiology*. 2018; 289:172656.
- Gutjahr R, Halaweish AF, Yu Z, et al. Human imaging with photon-counting-based CT at clinical dose levels: contrast-to-noise ratio and cadaver studies. *Invest Radiol*. 2016;51:421.
- Yu Z, Leng S, Jorgensen SM, et al. Initial results from a prototype whole-body photon-counting computed tomography system. In: *Medical Imaging 2015: Physics of Medical Imaging*. Vol. 9412. . International Society for Optics and Photonics; 2015:94120W.
- Alvarez RE, Macovski A. Energy-selective reconstructions in x-ray computerised tomography. *Phys Med Biol*. 1976;21:733-744.
- Taguchi K, Iwanczyk JS. Vision 20/20: single photon counting x-ray detectors in medical imaging. *Med Phys*. 2013;40:100901.
- Yeh BM, FitzGerald PF, Edic PM, et al. Opportunities for new CT contrast agents to maximize the diagnostic potential of emerging spectral CT technologies. *Adv Drug Deliv Rev*. 2017;113:201-222.
- Müllner M, Schlattl H, Hoeschen C, Dietrich O. Feasibility of spectral CT imaging for the detection of liver lesions with gold-based contrast agents-A simulation study. *Physica Med*. 2015;31:875-881.
- Cormode DP, Roessler E, Thran A, et al. Atherosclerotic plaque composition: analysis with multicolor CT and targeted gold nanoparticles. *Radiology*. 2010;256:774-782.
- Jafer FA, Weissleder R. Seeing within: molecular imaging of the cardiovascular system. *Circ Res*. 2004;94:433-445.
- Bardhan R, Lal S, Joshi A, Halas NJ. Theranostic nanoshells: from probe design to imaging and treatment of cancer. *Acc Chem Res*. 2011;44:936-946.
- Roessler E, Proksa R. K-edge imaging in x-ray computed tomography using multi-bin photon counting detectors. *Phys Med Biol*. 2007;52:4679-4696.
- Taguchi K, Zhang M, Frey EC, Xu J, Segars WP, Tsui BM. Image-domain material decomposition using photon-counting CT. In: *Medical Imaging 2007: Physics of Medical Imaging*. Vol. 6510. International Society for Optics and Photonics; 2007:651008.
- Schlomka J, Roessler E, Dorscheid R, et al. Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography. *Phys Med Biol*. 2008;53:4031-4047.
- Zimmerman KC, Schmidt TG. Experimental comparison of empirical material decomposition methods for spectral CT. *Phys Med Biol*. 2015;60:3175-3191.
- Alvarez RE. Efficient, non-iterative estimator for imaging contrast agents with spectral x-ray detectors. *IEEE Trans Med Imaging*. 2016;35:1138-1146.
- Ehn S, Sellerer T, Mechlem K, et al. Basis material decomposition in spectral CT using a semi-empirical, polychromatic adaption of the Beer-Lambert model. *Phys Med Biol*. 2016;62:N1-N17.
- Wu D, Zhang L, Zhu X, Xu X, Wang S. A weighted polynomial based material decomposition method for spectral x-ray CT imaging. *Phys Med Biol*. 2016;61:3749-3783.
- Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Netw*. 1989;2:359-366.
- Zimmerman KC, Schmidt TG. Comparison of quantitative k-edge empirical estimators using an energy-resolved photon-counting detector. In: *SPIE Medical Imaging, International Society for Optics and Photonics*; 2016:97831S.
- Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn machine learning in Python. *J Mach Learn Res*. 2011;12:2825-2830.
- Chen H, Qi B, Moore T, et al. Multifunctional Yolk-in-shell nanoparticles for pH-triggered drug release and imaging. *Small*. 2014;10:3364-3370.
- de Chermont QIM, Chanéac C, Seguin J, et al. Nanoprobes with near-infrared persistent luminescence for in vivo imaging. *Proceedings of the National Academy of Sciences*. 2007;104:9266-9271.
- Xcom N. Photon cross sections database. Retrieved April 30. 2013.
- Parchur AK, Sharma G, Jagtap JM, et al. Vascular interventional radiology-guided photothermal therapy of colorectal cancer liver metastasis with theranostic gold nanorods. *ACS Nano*. 2018;12:6597-6611.
- Ayala-Orozco A, et al. Au nanomatryoshkas as efficient near-infrared photothermal transducers for cancer treatment: benchmarking against nanoshells. *ACS Nano*. 2014;8:6372-6381.
- Bardhan R, Chen W, Bartels M, et al. Tracking of multimodal therapeutic nanocomplexes targeting breast cancer in vivo. *Nano Lett*. 2010;10:4920-4928.
- Veale M, Bell S, Duarte D, et al. Measurements of charge sharing in small pixel CdTe detectors. *Nucl Instrum Methods Phys Res Sect A*. 2014;767:218-226.