

# Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

## **Robust multimaterial decomposition of spectral CT using convolutional neural networks**

Zhengyang Chen  
Liang Li

# Robust multimaterial decomposition of spectral CT using convolutional neural networks

Zhengyang Chen<sup>a,b</sup> and Liang Li<sup>a,b,\*</sup>

<sup>a</sup>Tsinghua University, Department of Engineering Physics, Beijing, China

<sup>b</sup>Ministry of Education, Key Laboratory of Particle and Radiation Imaging (Tsinghua University), Beijing, China

**Abstract.** Spectral computed tomography (CT) can reconstruct scanned objects at different energy-bins and thus solve the multimaterial decomposition (MMD) problem. Because the linear attenuation coefficients of different basis materials may be extremely close, the decomposition problem is often ill-conditioned. Meanwhile, traditional material decompositions with image-domain algorithms are usually voxelwise based. Therefore, these algorithms rely heavily on image quality. Ring artifacts often exist in the reconstructed images of spectral CT due to the inconsistency feature of energy-resolved detectors and beam-hardening effect. Considering the enlargement of the receptive field and taking advantage of the modeling ability of convolutional neural networks in deep learning, we proposed a convolutional material decomposition algorithm to solve the MMD problem through a basis of patches instead of pixels of the spectral CT images. Simulations and physical experiments were performed to validate the proposed algorithm, and its quality was compared with a traditional MMD algorithm in the image domain. Results show that the proposed method achieves good accuracy, reduces mean squared errors by one to two orders, and exhibits robustness in the MMD of spectral CT images even in the case that obvious ring artifacts is presented. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.OE.58.1.013104](https://doi.org/10.1117/1.OE.58.1.013104)]

Keywords: spectral CT; multimaterial decomposition; convolutional neural network; image-domain; deep learning.

Paper 180823 received Jun. 8, 2018; accepted for publication Dec. 11, 2018; published online Jan. 12, 2019.

## 1 Introduction

The ability of spectral computed tomography (CT) to distinguish photons in terms of their different energies allows this technique to decompose scanned objects into their basis materials.<sup>1-5</sup> Compared with linear attenuation coefficients from traditional single energy CT, the distribution of basis materials can effectively reflect inner information. For example, in clinical practice, liver fat quantification requires the calculation of the components of fat, tissue, blood, and contrast agent.<sup>6</sup>

With the upgrading of product applications and the popularity of spectral CT, many material decomposition algorithms have been proposed. These algorithms can be divided into three types, namely, projection-domain processing methods, image-domain processing methods, and direct iterative reconstruction methods. Projection-domain methods use the projections from detectors as the input data and bring the decomposition model into nonlinear polychromatic projection process.<sup>7-9</sup> Therefore, they are sensitive to the mismatch of the projection data in different energy-bins. Image-domain methods assume that the effective attenuation coefficients of each pixel of the spectral CT images can be decomposed into the linear combination of several basis functions.<sup>10-16</sup> They are widely using in current preclinical and clinical applications. However, the performance of image-domain method is usually affected by the ring artifacts and beam-hardening artifacts within the CT images especially for the spectral CT using the photon-counting detectors. The direct iterative reconstruction method incorporates the models of the material decomposition and the physics of

spectral CT transmission.<sup>17-20</sup> For example, in Ref. 20, the iterative reconstruction method can even perform MMD of traditional integral detectors using angle-dependent filters. Theoretically, it has the potentials of decreasing noise and artifacts by taking the advantages of accurate modeling. However, its computational cost of the forward- and backward-projections between the material images and the projections in different energy-bins is expensive. In this paper, we focus on the MMD problem in image-domain.

Traditional material decomposition algorithms in spectral CT use the linear combination of a set of basis functions to represent the linear attenuation coefficient of location  $\vec{x}$  and then apply these functions in material decomposition.<sup>21</sup> Different basis functions lead to different material decomposition methods. For example, Ref. 22 uses the linear attenuation curves of preselected materials as the basis functions, leading an analytical solution of MMD problem. However, regardless of the choice of basis functions, these traditional material decomposition algorithms are voxelwise based, using no more than dozens of pixels around the target pixel. Therefore, their performances usually depend on the reconstructed image quality of spectral CT. They are often affected by the beam-hardening artifacts and ring artifacts in the spectral CT using the photon-counting detectors. An idea to avoid these problems is using the data-driven methods, for example, machine learning, instead of the analytical solution such as Ref. 22. In this way, we can expand the receptive field of data-driven algorithms to provide more information for their learning.

Wang<sup>23</sup> published an outlook article on deep learning and presented a roadmap of CT imaging methods in the deep learning framework. Inspired by his work, we imported the deep learning technique into spectral CT and proposed

\*Address all correspondence to Liang Li, E-mail: [lliang@tsinghua.edu.cn](mailto:lliang@tsinghua.edu.cn)

a multimaterial decomposition (MMD) method in image-domain.<sup>24</sup> Deep learning has made considerable progress in many fields. In image analyses, deep neural networks can extract the features of images without manual marking and are thus effective in image classification.<sup>25–27</sup> Recently, deep learning has shown its effectiveness in medical imaging,<sup>28</sup> such as auxiliary diagnosis,<sup>29–31</sup> lesion localization,<sup>32,33</sup> and image noise reduction.<sup>34,35</sup> By using deep learning, we can break through the restriction of sampling theorem and realize reconstruction using few views.<sup>36</sup> Using deep learning to acquire priori knowledge of images can also enhance the results of traditional MMD algorithm.<sup>37</sup> As a universal approximator,<sup>38</sup> neural networks can be trained to solve the MMD problem by given adequate and suitable training data. In this paper, we propose a convolutional material decomposition (CMD) algorithm on the basis of patches instead of voxels of the spectral CT images by building two convolutional neural networks (CNN) based on existing networks. Section 2 presents the CMD algorithm. Then, the results of simulation and physical experiment are presented in Sec. 3 to validate it. Sections 4 and 5 are the discussion and conclusion.

## 2 Method

### 2.1 MMD Model

As an image-domain method, the reconstructed image of spectral CT is used as the input data. The reconstructed image is the x-ray linear attenuation coefficient  $\mu(\vec{x}, E)$  of an object scanned at different spectra. Here,  $\vec{x}$  is the spatial position, and  $E$  is the spectrum or energy-bin.<sup>35</sup>

Knowing  $\mu(\vec{x}, E)$  by the spectral CT reconstruction algorithm, we choose the attenuation coefficient  $\mu_l(E)$  of the basis material  $l$  at different spectra  $E$  as the basis function. Thus, we have

$$\mu(\vec{x}, E) = \sum_{l=1}^{N_l} \eta(\vec{x}, l) \mu_l(E), \quad E = 1, \dots, N_E, \quad (1)$$

where  $\eta(\vec{x}, l)$  denotes the volume fraction of basis material  $l$  at position  $\vec{x}$ ,  $N_l$  is the number of basis materials, and  $N_E$  is the number of spectra. As proven in Ref. 4, we apply the constraints on volume and mass conservation, all of which lead to the following constraints:

$$\begin{cases} \sum_{l=1}^{N_l} \eta(\vec{x}, l) = 1, & \forall \vec{x} \\ 0 \leq \eta(\vec{x}, l) \leq 1, & \forall \vec{x}, l \end{cases} \quad (2)$$

The constraints on each  $\vec{x}$  have the same form as the result of the image classification, that is, the possibility of each image type. Therefore, a CNN designed for image classification can be transformed to fit the material decomposition problem.

### 2.2 Deep Learning Based on CNN

Equations (1) and (2) describe the relationship between the reconstructed image and the decomposition result. Traditional material decomposition algorithms in image-domain are based on these voxelwise equations. They often suffer from the spatial-specific errors or artifacts such as beam-hardening artifacts, metal artifacts, ring artifacts,

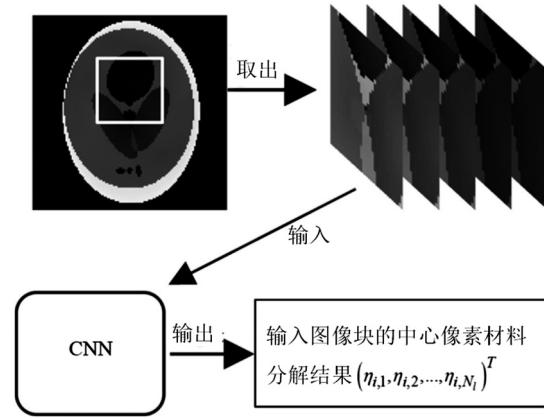


Fig. 1 Work flowchart of CMD.

and so on. Although some decomposition algorithms add regularization terms based on *a priori* knowledge, e.g., total variation, into the cost function to improve image quality, the results are still unsatisfactory. These spatial-specific features are difficult to describe using explicit mathematical formulas; nevertheless, they can be learned using deep neural networks. As a universal approximator, neural networks have been proved effective if enough neurons are given.<sup>38</sup> This paper will show that neural networks can determine the patterns of material decomposition if the neural networks are trained with sufficient samples.

Enlarging the receptive field of each pixel can improve material decomposition performance as well. As shown in Fig. 1, we train the CNN to fit each patch and the exact decomposition result of this patch's central voxel. Then, by moving this receptive field voxel-by-voxel among the reconstructed CT image, we obtain the decomposition result. We refer to this method as the CMD algorithm.

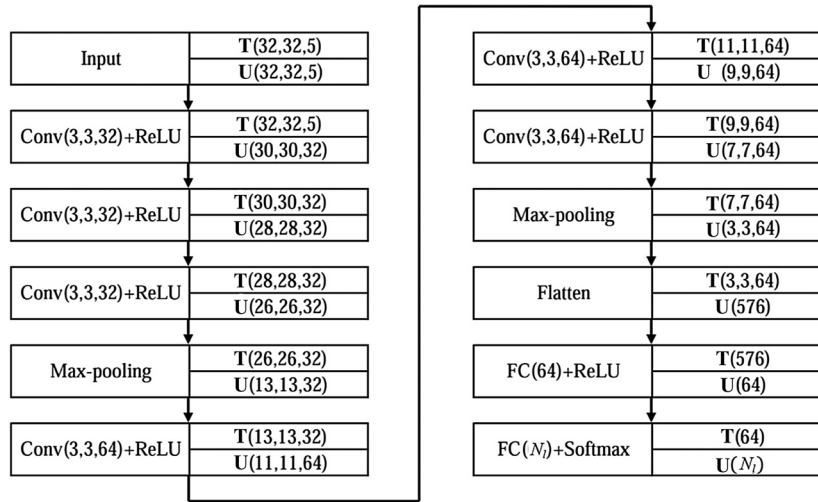
CNN has become one of the most important methods of image analysis.<sup>26,38</sup> A CNN consists of several simple operations or so-called layers. By sending an image into the CNN and going through several layers, such as convolutional, activation, pooling, and fully connected layers, we can transform the image into the result we want, including classification or material decomposition.

The image can be considered as a three-rank tensor  $\mathbf{T}(x, y, k)$ , whose first two parameters  $x$  and  $y$  describe the position of a voxel and  $k$  is the channel.<sup>39</sup> For a chromatic image,  $k$  ranges from 1 to 3, and for the reconstructed image of spectral CT,  $k$  ranges from 1 to  $N_E$ . All the layers mentioned above are the derivable operations on this tensor, transforming tensor into another tensor. Because these parameters are derivable, most of parameters in the layers can be trained automatically by stochastic gradient descent (SGD) or other training algorithms.<sup>38</sup>

Details of these layers can be found in the [Appendix](#).

### 2.3 CNN Used in CMD

We build two CNNs for the decomposition task. Both CNNs are simplified versions of the CNNs that work effectively in image classification. Let  $x = 32$  and  $y = 32$ .  $k = 5$  is the size of the input patch  $\mathbf{T}(x, y, k)$ , which is the receptive field with size  $x$ ,  $y$  and channel number  $k = N_E$ .



**Fig. 2** Structure of the first CNN based on VGG used in this work. The left parts of the rectangles are the layers' names, and the right parts are the sizes of the input tensors  $\mathbf{T}$  and output tensors  $\mathbf{U}$ . The numbers after the convolutional layers are  $(c_x, c_y, l)$ , which refer to the size of the convolutional kernels. Parameter  $k$  of convolutional layers is determined by the input tensor.

### 2.3.1 Visual geometry group network

The first CNN is based on a visual geometry group (VGG) network.<sup>25</sup> The VGG network is stacked by convolution layers with small kernel. Its structure is shown in Fig. 2. All the activation layers are ReLUs except for the last one, which is Softmax.

### 2.3.2 Deep residual network

To compare the performances of the different networks, we build the second CNN based on a deep residual network (DRN),<sup>27</sup> which is more complicated than a VGG network. The network comprises three types of blocks that are shown in Fig. 3. Block A does not change the size of the input tensor. Block B halves the size of the input image by down-sampling the first two parameters. Block C halves the size while doubling its channels. The full structure of the DRN is shown in Fig. 4.

### 2.4 Loss Function and SGD with Momentum

As shown in Fig. 1,  $(\mathbf{T}, \mathbf{U})$  is the training sample, where  $\mathbf{T}$  is a three-rank tensor and  $\mathbf{U}$  is a one-rank tensor.  $\mathbf{U}$  represents the material decomposition result of the central voxel of  $\mathbf{T}$ . The CNN is initialized randomly and trained using the minibatch SGD method. To establish a regression problem, we use mean square error (MSE)

$$L = \frac{1}{N_S} \sum_{\mathbf{T} \in \mathbf{S}} \frac{1}{N_I} [\mathbf{U} - \mathbf{F}(\mathbf{T})]^2, \quad (3)$$

as the loss value, where  $\mathbf{F}(\mathbf{T})$  is the output of the CNN.  $\mathbf{S}$  is a set of  $\mathbf{T}$ , who is the minibatch in the training step, the validation set in validation step, and test set in test step.  $N_S$  is the number of tensors in  $\mathbf{S}$ . As defined above, every parameter in the CNN layers is derivable or assigned with a gradient. According to the chain rule of derivation,<sup>38</sup> the partial derivatives of  $L$  to the parameters in the CNN

are calculable. For a parameter  $p$  of the CNN, the updating function is

$$p = p - lr \times \frac{\partial L}{\partial p}, \quad (4)$$

where  $lr$  is the learning rate or the step length of the gradient descent algorithm. To improve the stability of the convergence, we add the momentum term  $\rho$  to Eq. (4)

$$v = \rho v - lr \times \frac{\partial L}{\partial p}, \quad p = p + v. \quad (5)$$

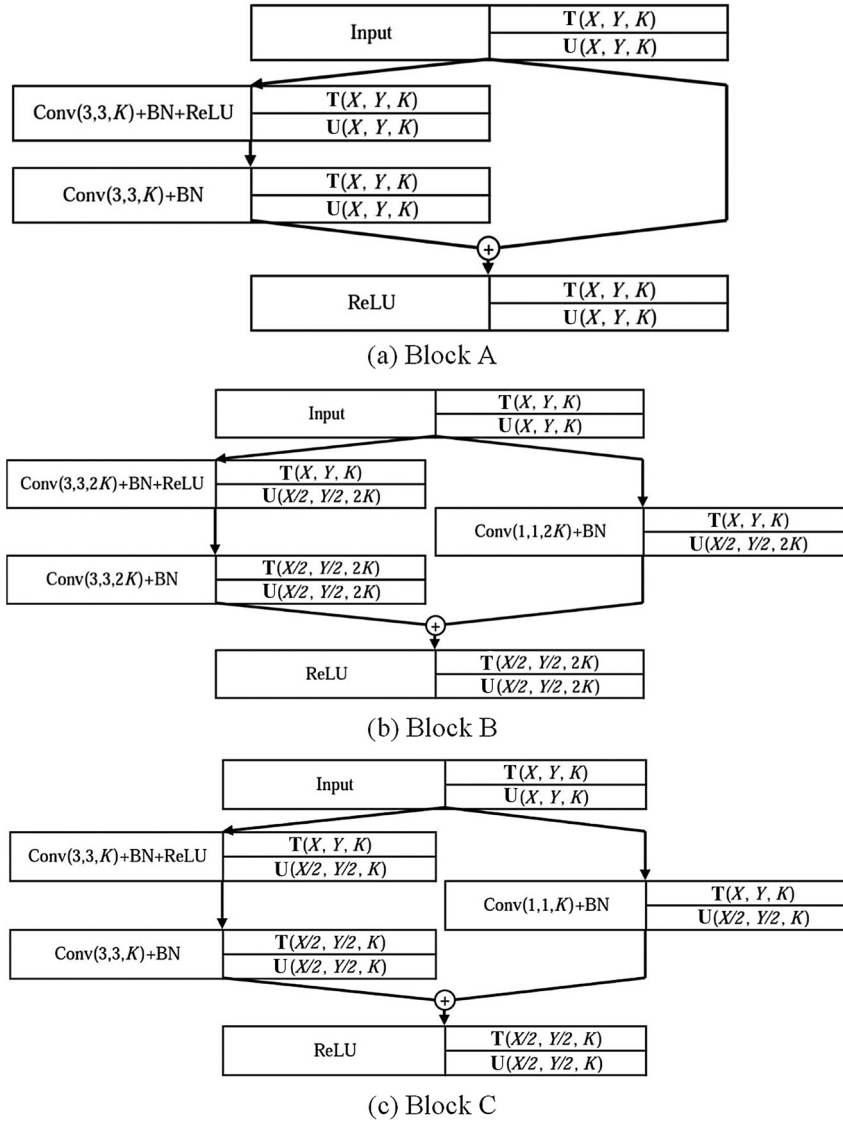
However, no specific method can be used for training CNNs. In this study, a learning rate decrease policy is adopted to train the CNN and achieve satisfactory convergence.

Initially, we set  $lr = LR$  and  $patience = P$  and then pick  $VAL\_NUM$  patches randomly from the validation phantoms as the validation set. We then pick  $BATCH\_NUM$  samples from the training set randomly and use them to train the CNN through SGD. For every  $ITER\_NUM$  iteration (the training patches are reselected for every iteration), we send the validation set to calculate the MSE of the CNN output. If the MSE is greater than the  $lastMSE$  for more than  $patience$  times sequentially, we divide  $lr$  into two, add  $P\_STEP$  to  $patience$ , repick the validation set, and continue the iteration until the  $MAX\_ITERATION\_TIMES$  or while  $lr$  is less than  $\epsilon$ . The pseudocode is shown in Fig. 5.

### 2.5 Convergence Property of Neural Network

The SGD with momentum method is widely used in deep learning optimization and turns out to be effective among various applications.

In the case of a local minimum  $p^*$  nearing  $p$ . At time step  $t$ , Eq. (5) can be rewritten as



**Fig. 3** Three substructures in DRN. The furcation indicates two copies of the related tensor. The plus sign in the circle represents the addition of a corresponding element in the tensor with the same shape.

$$v_{t+1} = \rho v_t - lr \times \frac{\partial L_t}{\partial p_t}, \quad (1 - \sqrt{\rho})^2 \leq lr \cdot \frac{\partial L_t}{\partial p_t} \leq (1 + \sqrt{\rho})^2, \quad (9)$$

$$p_{t+1} = p_t + v_{t+1} = p_t + \rho v_t - lr \times \frac{\partial L_t}{\partial p_t} = p_t + \rho(p_t - p_{t-1}) - lr \times \frac{\partial L_t}{\partial p_t}. \quad (6)$$

we may get

$$\lambda(\mathbf{A}_t) = \sqrt{\rho}, \quad (10)$$

Define momentum operator  $\mathbf{A}_t$

$$\mathbf{A}_t = \begin{pmatrix} lr \cdot \frac{\partial L_t}{\partial p_t} & -\rho \\ 1 & 0 \end{pmatrix}. \quad (7)$$

Equation (11) is equivalent to

$$\begin{pmatrix} p_{t+1} - p^* \\ p_t - p^* \end{pmatrix} = \mathbf{A}_t \times \begin{pmatrix} p_t - p^* \\ p_{t-1} - p^* \end{pmatrix}. \quad (8)$$

Reference 40 proved that when  $p_t$  is close to the local minimum  $p^*$ , if

where  $\lambda$  is the spectral radius of  $\mathbf{A}_t$ . When  $\mathbf{A}_t$  is a constant matrix,  $\lambda = \sqrt{\rho} < 1$  which ensures Eq. (8) converges. Actually,  $\mathbf{A}_t$  may be changed in the above iterations. In this case, Ref. 40 verified that this momentum method could robustly yield linear convergence with rate  $\sqrt{\rho}$  empirically around a local minimum  $p^*$ .

The difference between the method in Ref. 40 and ours is that we evaluate the gradient  $\partial L / \partial p$  by sampling a minibatch from the training set instead of calculating over the whole training set, which is a widely used trick to accelerate the convergence without losing too much accuracy. In this way, the SGD with momentum method is converged in the expectation sense.

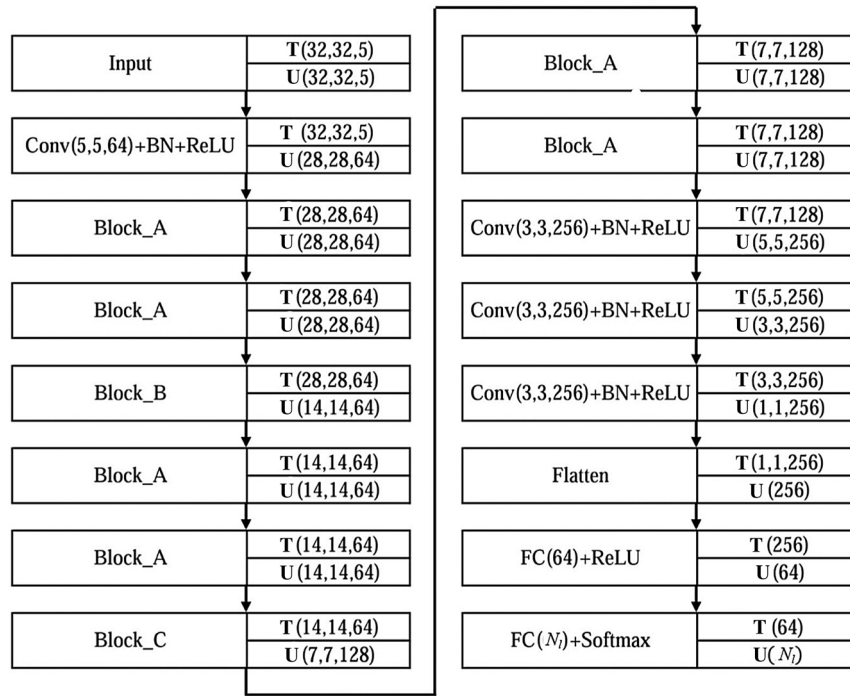


Fig. 4 Structure of the second CNN used in this paper.

```

lr = LR, patience = P, counter = 0, lastMSE = M
generate VAL_NUM validation patches from validation phantoms randomly
do
  for iter in 1...ITER_NUM
    generate BATCH_NUM training patches from training phantoms randomly
    use mini-batch SGD to train CNN with these training patches
  end
  calculate MSE on validation patches
  if MSE <= lastMSE
    counter = 0
    lastMSE = MSE
  else
    counter = counter + 1
    if counter > patience
      lr = lr / 2, patience = patience + P_STEP, counter = 0, lastMSE = M
      generate VAL_NUM validation patches from validation phantoms randomly
    else
      lastMSE = MSE
    end
  end
end
until lr <  $\epsilon$  or reaches MAX_ITERATION_TIMES

```

Fig. 5 Pseudocode of learning rate decrease policy.

However, because the fitting ability of deep neural network is nearly infinite, the search for the global optimal solution in most cases will lead to overfitting.<sup>41</sup> On the other hand, critical point is quite rare in a deep neural network. The SGD method tends to stop in some “flat minima.”<sup>42</sup>

The loss of flat minima that SGD could reach is related with the structure of neural network.<sup>43</sup> As stated in the paper, the CNNs we used are based on VGG and DRN, which are widely verified in image classification. As a task immigration, these CNNs perform well according to our experiments.

### 3 Numerical Simulation and Physical Experiment Results

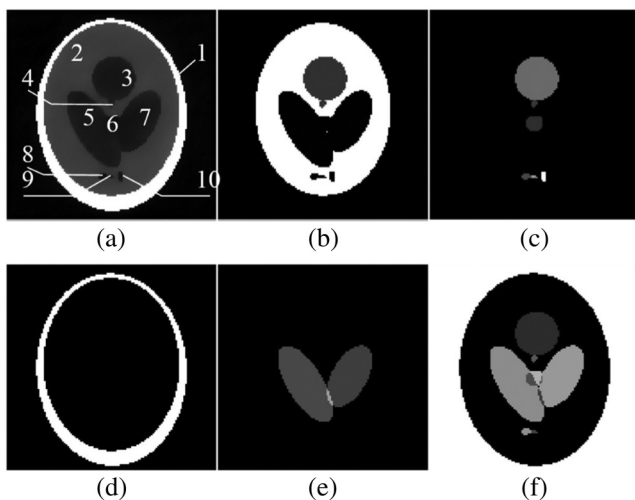
To train the CNN to solve the MMD problem, we need organize three datasets, namely, the training dataset for SGD to optimize the parameters in the CNNs, the validation dataset for adjusting the hyperparameters (such as the number of mini-batches,  $l_r$ , and so on), and the test dataset for evaluating the performance of the VGG-CMD and DRN-CMD. The three datasets should satisfy the same distribution.<sup>38</sup> The ground truth values for the three datasets are also necessary as the supervised signal for training or the evaluation criterion. The phantom generation, projection, and reconstruction are all coded in MATLAB R2014a. The construction, training, and testing of CNN are coded by deep learning library Keras 2 in Python.<sup>44</sup>

#### 3.1 Numerical Simulation Results

In numerical simulations, we used 115 digital phantoms based on the 2-D Shepp–Logan phantom to train and evaluate the proposed CMD algorithm. All of these phantoms comprise ten ellipses and five materials such as soft tissue, lung, bone, blood, and air. Their linear attenuation coefficients are obtained from Ref. 5, as shown in Table 1. One of the phantoms is shown in Fig. 6. The concentrations of these ellipses are shown in Table 2. “×” in Table 2 means the concentration is 0. “~” means the concentration is 1 minus the other materials’ concentrations for meeting the constraints of Eq. (2). “±0.1” means that the concentration of that ellipse satisfies 0.1 normal distribution among

**Table 1** ID of materials used this paper and in Ref. 5.

This paper [the order in Figs. 6(b)–6(f)]	1	2	3	4	5
Material	Soft tissue	Lung	Bone	Blood	Air
Reference 5 (the indexes in Fig. 4)	3	2	7	4	1



**Fig. 6** An example of the digital phantoms. (a) One of the phantoms who comprises 10 ellipses. (b)–(f) The concentrations of five materials: soft tissue, lung, bone, blood, and air, respectively.

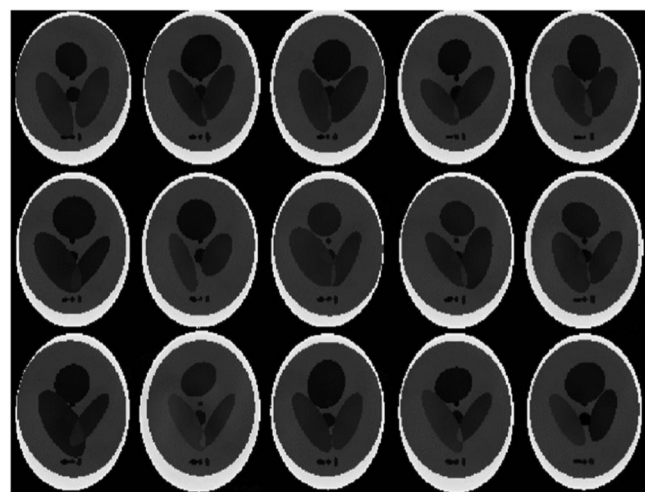
different phantoms. Then, we randomly changed the position, size, and direction of the ellipses, and the concentrations of the materials to generate other similar 114 phantoms, as shown in Fig. 7.

Note that the linear attenuation coefficients of materials 1, 4 and 2, 5 in this paper are both relatively close, which worsens the ill-condition of MMD problem.

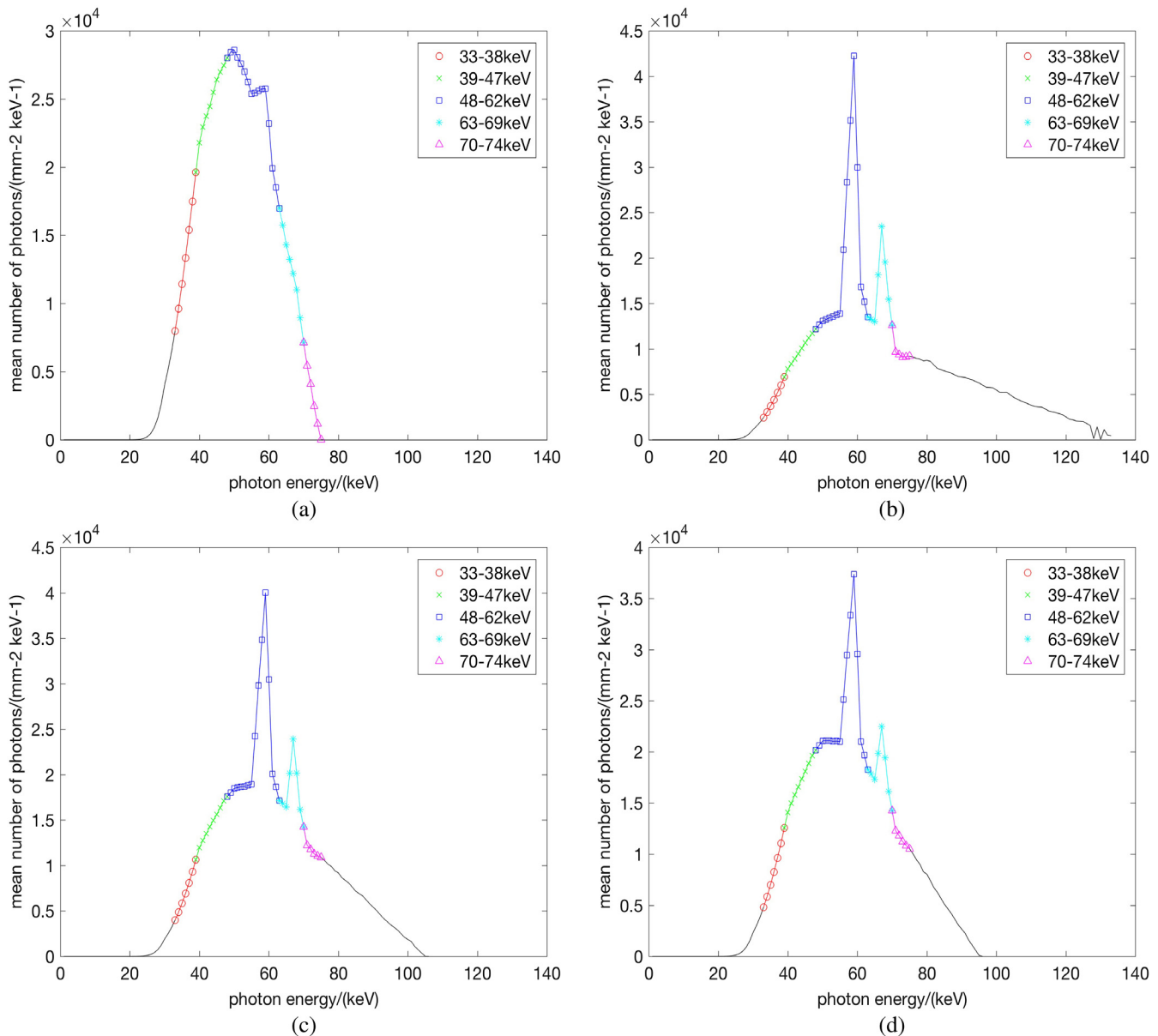
The x-ray beam spectra used in numerical simulations were generated with a Siemens simulator<sup>45</sup> at 75, 135, 105, and 95 kilovolts peak (kVp), respectively. A 12-mm thickness aluminum filter was added. To simulate five-energy-bin photon-counting detectors, each spectrum is divided into five energy-bins as shown in Fig. 8. The x-ray fan-beam covered a 20-mm-diameter field of view with 256 detector elements. The number of views over 360 deg is 360. The reconstructed images were discretized into a 128 × 128 grid. We use

**Table 2** Mean concentrations of materials in each ellipse of the generated phantoms.

ID of ellipses	ID of materials				
	1	2	3	4	5
1	×	×	1	×	×
2	1	×	×	×	×
3	0.1 ± 0.1	0.5 ± 0.1	×	×	~
4	0.1 ± 0.1	0.25 ± 0.1	×	×	~
5	0.2 ± 0.1	×	×	0.35 ± 0.1	~
6	0.2 ± 0.1	0.25 ± 0.1	×	×	~
7	0.2 ± 0.1	×	×	0.25 ± 0.1	~
8	×	0.3 ± 0.1	×	×	~
9	×	0.6 ± 0.1	×	×	~
10	×	1	×	×	×



**Fig. 7** Some phantoms used in the simulation. The size, position, and direction of every ellipse and the concentrations of the five materials are different (but close).



**Fig. 8** In numerical simulations, four spectra were generated using a Siemens simulator. Each spectrum was divided into five different energy-bins. Twenty energy-bins of the four spectra used in this paper. The maximum energies of (a–d) are 75, 135, 105, and 95 kVp, respectively. The scales of the energy-bins are marked in each figure.

the method in Ref. 46 to obtain the projection, which can simulate the statistical fluctuation and scattering in the imaging process, and can simulate the unsatisfactory situation such as pile up and charge sharing of the detectors based on real detectors. The spectral CT images in different energy-bins were reconstructed with the ASD-POCS algorithm with 25 iterations.<sup>47</sup>

Two different simulation studies were carried out. The first one only used the reconstructed spectral CT images of all of 115 phantoms with the 75 kVp spectrum of Fig. 8(a). We divided them randomly into three parts with the number of 85, 15, and 15, and used them to generate image patches, which were used as training sets, verification sets, and test sets, respectively. The patches we used has  $32 \times 32 \times 5$  voxels. As stated in Sec. 2.4, we set  $LR =$

$0.1$ ,  $P = 3.5$ ,  $P\_STEP = 0.5$ ,  $VAL\_NUM = 16384$ ,  $BATCH\_NUM = 128$ ,  $ITER\_NUM = 64$ ,  $\varepsilon = 1e-4$ , and  $\rho = 0.9$ .

The MSE of decomposition results of VGG-CMD and DRN-CMD algorithms are shown in Fig. 9 for the 15 test phantoms. As a contrast, we used an extended direct inversion algorithm in image-domain algorithm to calculate the MMD voxel-by-voxel (EDI-MMD). It can be seen as a spectral CT version of the original one proposed by Mendonca et al.<sup>4,22</sup> The procedure of EDI-MMD is shown in Fig. 10. The basis material triplets set is in Table 3.

On our GPU (GTX 1070), training VGG-CMD and DRN-CMD took about 3 and 12 h, respectively. After separation, the average MSE of all the 15 testing phantoms are  $6.5908e-4 \text{ mm}^{-2}$  for VGG-CMD,  $4.7651e-4 \text{ mm}^{-2}$  for DRN-CMD, and  $1.0030e-2 \text{ mm}^{-2}$  for EDI-MMD,



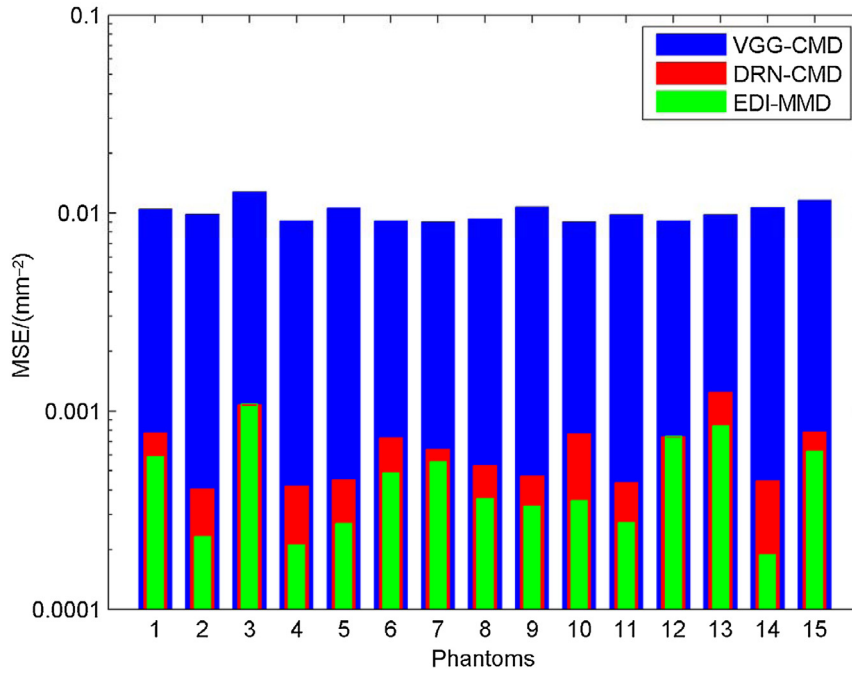


Fig. 9 MSEs for 15 testing phantoms based on three methods.

form the triplets set  $\Omega$  according to the phantoms:

```

for  $i = 1 \dots N_{pixel}$ 
    team = [];
    for  $t = 1 \dots N_{triplet}$ 
        assume  $pixel_i$  consists of the basis materials in material  $\Omega_t$  only, and solve the least
        square solution of (1) with sum-to-one constraint;
        if the solution satisfies the block constraint [0, 1]
            append the solution into team;
        end
    end
    if team is empty
        apply the least MSE solution as the result of  $pixel_i$ ;
    else
        apply the least MSE solution in team as the result of  $pixel_i$ ;
    end
end
filter the decomposition result with a  $3 \times 3$  median filter;
    
```

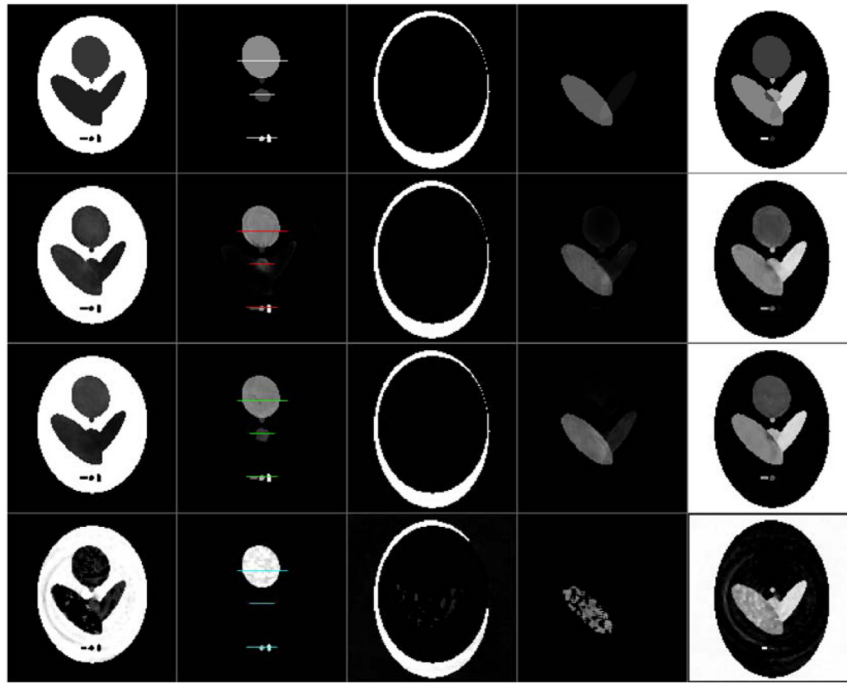
Fig. 10 Pseudocode of EDI-MMD.

Table 3 Triplets set for phantom in simulation.

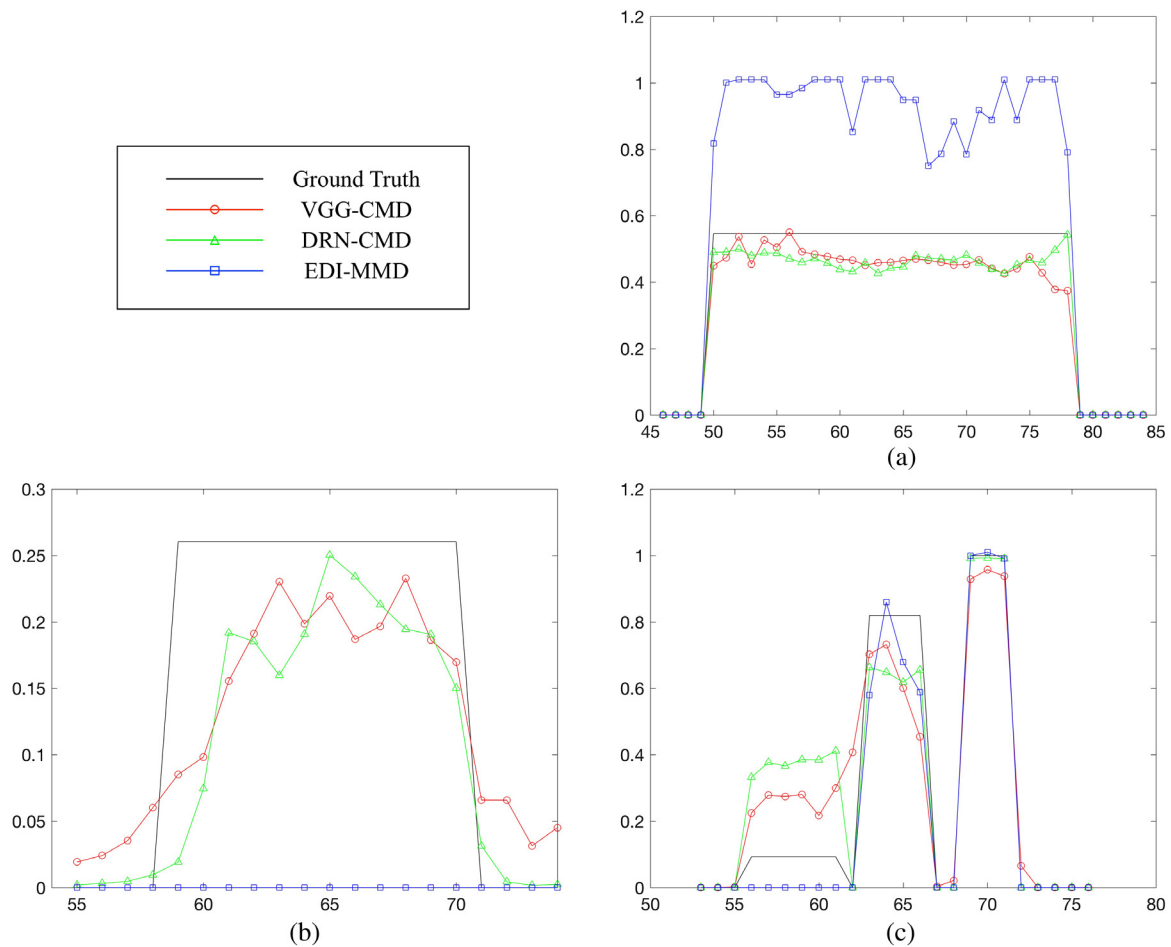
Triplet ID	Soft tissue	Lung	Bone	Blood	Air
1			✓		✓
2		✓			✓
3	✓				✓
4	✓			✓	✓
5	✓	✓			✓
6	✓	✓		✓	✓

respectively. The MSE of EDI-MMD algorithm of different testing phantoms are very close because the illcondition of MMD problems of the test phantoms are similar with the same five basis materials as shown in Figs. 6(b)–6(f).

There are some differences among the MSE of VGG-CMD and DRN-CMD algorithms for different phantoms due to their different structures. However, compared to the results of EDI-MMD, the MSE are reduced more than two orders in general. The decomposition results of one phantom are shown in Fig. 11. The rows from top to bottom are the decomposition ground truth and the results of VGG-CMD, DRN-CMD, and EDI-MMD, respectively. The columns from left to right are the concentrations of five materials: soft tissue, lung, bone, blood, and air, respectively.



**Fig. 11** Ground truth and decomposition results of three algorithms. The rows from top to bottom are ground truth, results of VGG-CMD, DRN-CMD, and EDI-MMD algorithms. The columns from left to right are the concentrations of soft tissue, lung, bone, blood, and air.



**Fig. 12** Profiles of the decomposition results of lung. (a-c) Mean the different positions of the 44th, 69th, and 102nd rows.

The EDI-MMD algorithm decomposes the soft tissue, lung, and blood wrong due to the close linear attenuation coefficients.

We chose the 44th, 69th, and 102nd rows of the decomposition result of lung to draw profiles, which divide the ellipses halves. The positions of these profiles are showed in the second column of Fig. 11, and the section lines are drawn in Fig. 12. The MSE of the three methods on the profiles are presented in Table 4.

Equation (1) indicates that the decomposition result should only depend on single voxel if we ignore the unideal factors in spectral CT image. Although it uses patches as the input data, the output of CMD algorithm should rely mainly on the central voxel of every patch. We chose a test phantom. A  $3 \times 3$  square in the center of each patch of its reconstructed images is set to zero. The decomposition results are shown in Fig. 13. Although the zero region only occupies  $3^2/32^2 = 0.879\%$  of the whole patch, the degradation of the output of the CMD is drastic. This result indicates that the deep

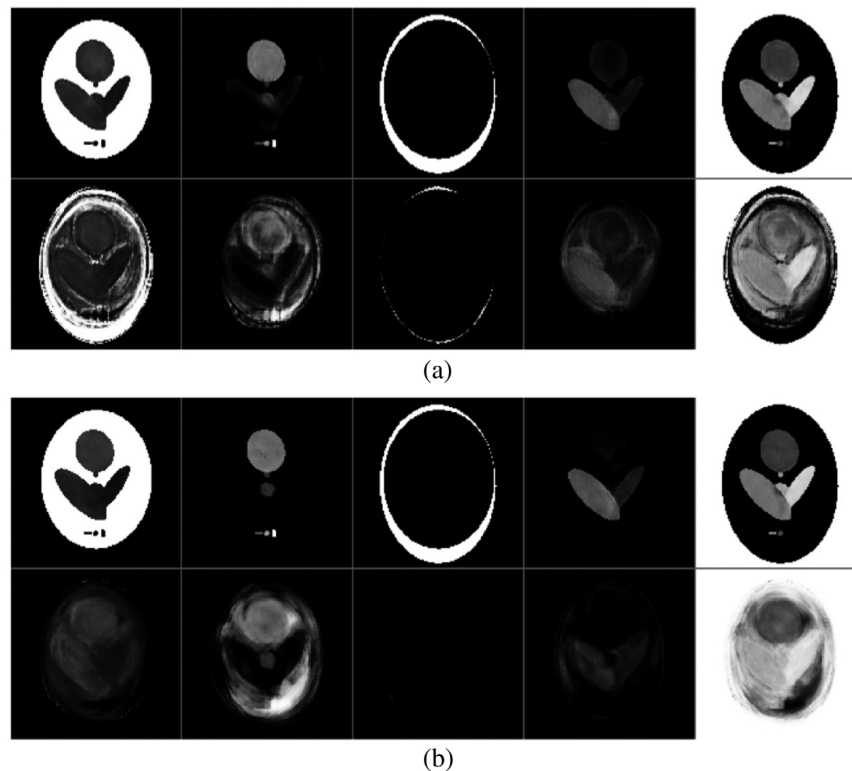
**Table 4** MSE of the decomposition section lines of material 2.

MSE ( $\text{mm}^{-2}$ )	Profiles (a)	Profiles (b)	Profiles (c)
VGG-CMD	0.0061	0.0060	0.0235
DRN-CMD	0.0051	0.0072	0.0253
EDI-MMD	0.1215	0.0407	0.0076

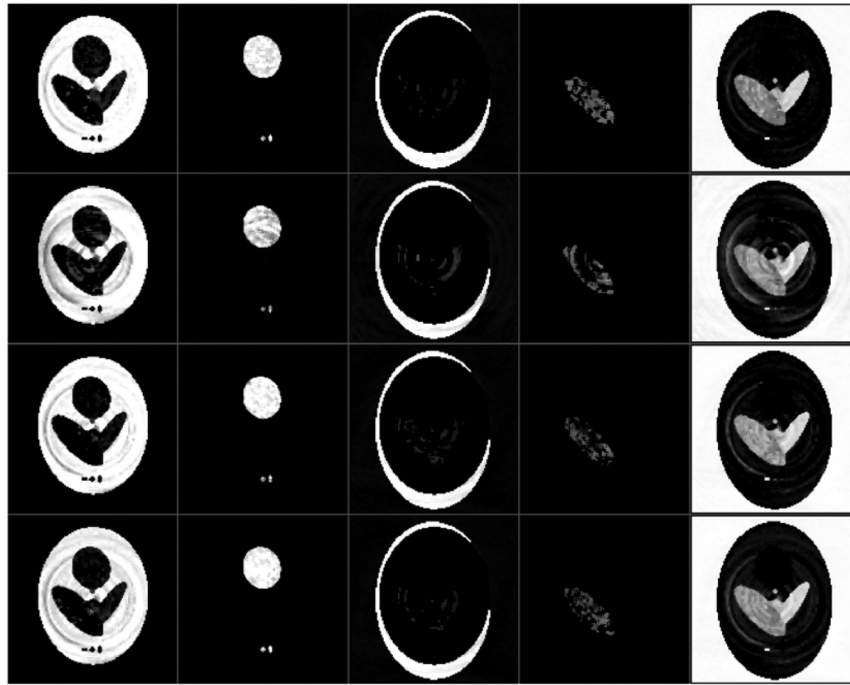
learning procedure of CMD algorithm can insure that the central voxel of each patch plays a more important role.

In the second simulation, we studied the CMD performance on the spectral CT images from an unknown or never trained spectrum. Traditional material decomposition algorithms need to measure  $\mu_l(E)$  of the basis materials and spectral CT images of the scanned object at the same x-ray spectrum.  $\mu_l(E)$  relates to not only the physical properties of material itself but also the spectrum of incident x-ray beam and the detection efficiency of the detector. Therefore, their decomposition results are sensitive to the variation of the effective spectrum of spectral CT systems. Figure 14 shows the decomposition results of EDI-MMD algorithm. The first, second, and third rows are the results when the x-ray spectra and  $\mu_l(E)$  of the basis materials within the corresponding energy-bins are all exactly provided. If the x-ray spectrum used in CT scanning is changed from the one measuring  $\mu_l(E)$  of the basis materials, the decomposition results of EDI-MMD will deteriorate. As showed in the last row of Fig. 14,  $\mu_l(E)$  of the basis materials used in EDI-MMD algorithm were generated by performing interpolations among the  $\mu_l(E)$  of the basis materials of the spectra of Figs. 8(a)–8(c). The average MSE of the results of spectra Figs. 8(a)–8(c) is  $1.1751\text{e-}2 \text{ mm}^{-2}$ , while the MSE of the results of spectrum Fig. 8(d) is  $1.2002\text{e-}2 \text{ mm}^{-2}$ .

Trained by patches from multiple different spectra rather than one specific spectra, CMD shows its potential to solve the MMD problem of unknown x-ray spectra. Because training for every CT system and its configuration on kVp/filtering is sometimes uneconomical. In this simulation, we randomly picked 85 phantoms and reconstructed their spectral CT



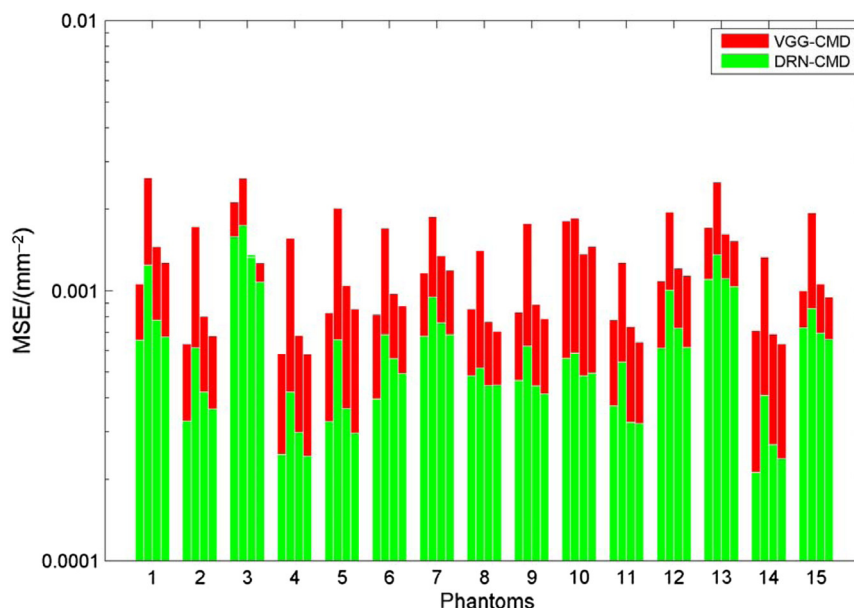
**Fig. 13** Decomposition results of (a) VGG-CMD and (b) DRN-CMD. The top row of each figure is the results with normal patches. The bottom row of each figure is the results with the patches setting the central  $3 \times 3$  square to zero. The columns from left to right are the concentrations of soft tissue, lung, bone, blood, and air.



**Fig. 14** Decomposition results of EDI-MMD algorithm. The rows from top to bottom are the results using spectrum (a), (b), (c), and (d) of Fig. 8, respectively. The columns from left to right are the concentrations of soft tissue, lung, bone, blood, and air.

images at the 15 energy-bins of Figs. 8(a)–8(c). Then, from each of these 15 channels of patch, five channels were randomly selected to form the input of CNN. That is, the input tensor still had  $32 \times 32 \times 5$  voxels. The other 15 phantoms were randomly picked to generate the validation dataset in exactly the same way. We use the same learning rate decrease policy as in the first simulation to train the CNNs. As the test dataset, the reconstructed images of the rest 15 phantoms in all of 20 energy-bins of Fig. 8(a)–8(d) were used to test the decomposition results of the CMD algorithm.

The results are shown in Fig. 15. For the first three spectra, the average MSE on 15 testing phantoms are  $1.3352e-3 \text{ mm}^{-2}$  for VGG-CMD and  $6.6640e-4 \text{ mm}^{-2}$  for DRN-CMD, respectively. And for the fourth spectra, which was not used in training, the MSE of the above two CMD algorithms are  $9.7165e-4$  and  $5.3719e-4 \text{ mm}^{-2}$ , respectively. As shown in Fig. 15, the CMDs' performance did not turn worse when the x-ray spectrum was changed from the one calibrating the basis materials. This result means that we can use the proposed CMD algorithm to solve the MMD



**Fig. 15** MSE on 15 testing phantoms with four different spectra.

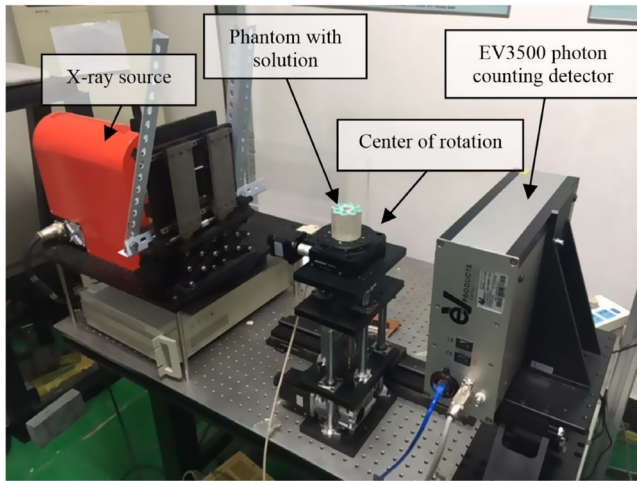


Fig. 16 The experimental spectral system.

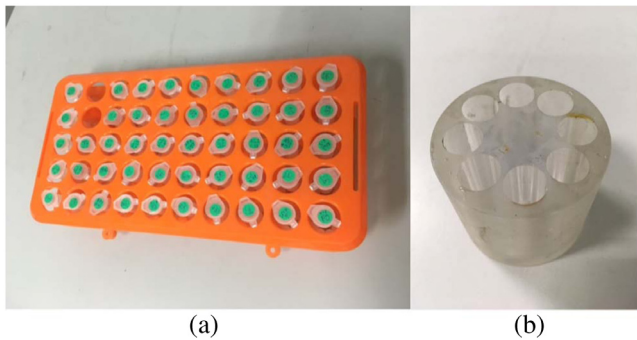


Fig. 17 (a) All of the 48 centrifugal tubes filled with different concentration solutions. (b) The resinic cylinder phantom.

problems in different spectral CT systems with different hardware configurations even if the spectrum of a CT system is unknown.

### 3.2 Physical Experiment Results

In the physical experiment, we used the reconstructed images of our spectral CT system to validate CMD algorithm and compare it with the EDI-MMD algorithm. The experiments were completed on our spectral CT system as shown in Fig. 16. It uses a conventional Hamamatsu L12161-07 x-ray tube and a linear photon-counting detector array (eV3500,

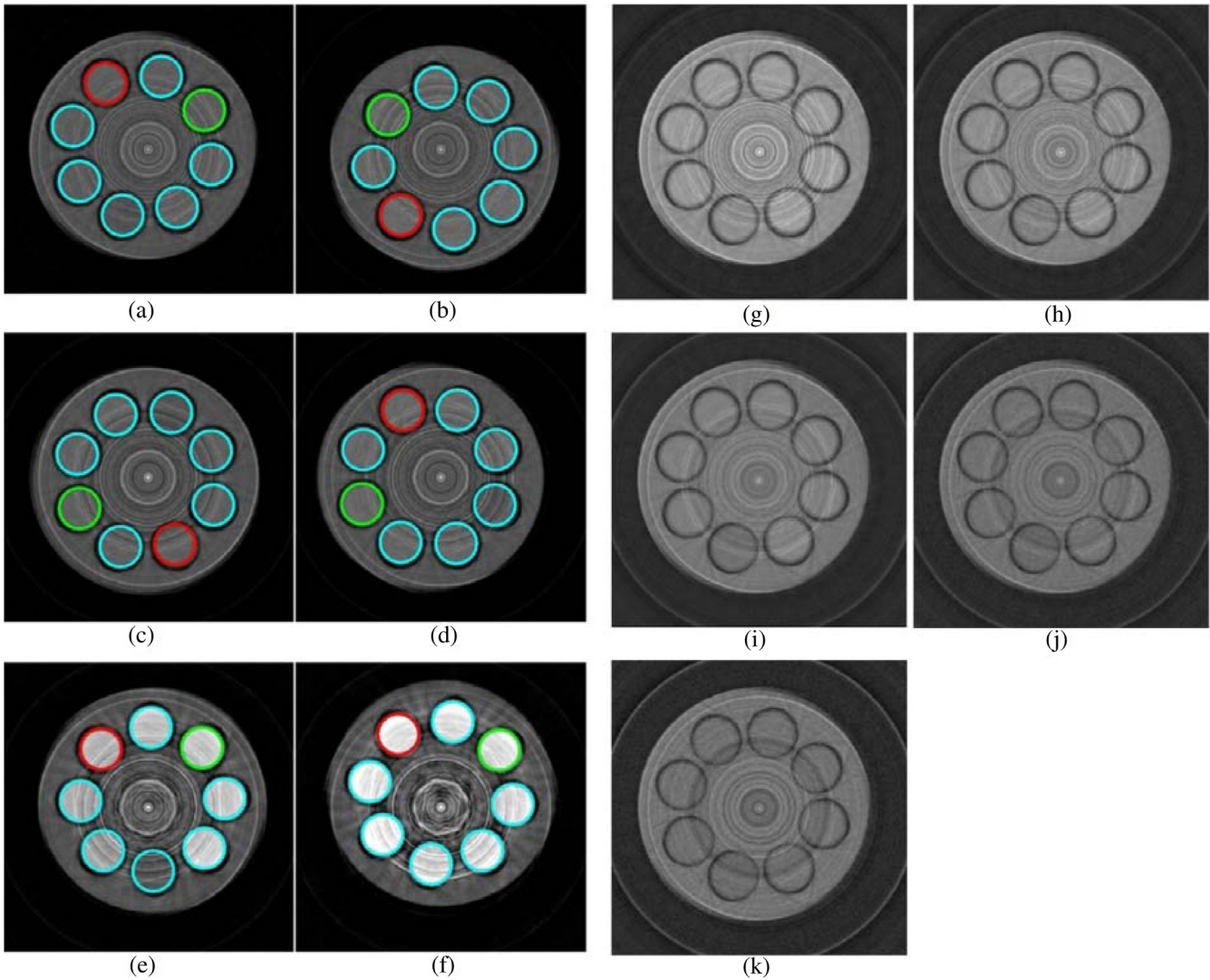
eV PRODUCTS, Saxonburg, Pennsylvania).<sup>46,48</sup> The detector array has 256 elements and 5 energy-bins of each one. The energy-bins were set at [26, 33], [33, 40], [40, 50], [50, 60], and [60, 80] keV. The size of each element is 0.5 mm × 2 mm. The x-ray tube's max energy was set 75 kVp, and the tube current 20  $\mu$ A. Over a 360-deg scan, 360 projections were acquired, and every projection radiated for 5 s. The distance between the x-ray source and the detector was 70 cm, and the distance between the x-ray source and the center of rotation was 44 cm. The reconstructed image had 256 × 256 pixels whose size was 0.025 mm × 0.025 mm. The image was reconstructed by the filtered-backprojection algorithm for each energy-bin.

The phantom used in the experiments was a 50-mm diameter resinic cylinder with eight 11-mm-diameter round holes to insert centrifugal tubes, as shown in Fig. 17. We used three types of solutions, namely, NaI, Gd(NO<sub>3</sub>)<sub>3</sub>, and CaCl<sub>2</sub>, with different concentrations to test the performance of the CMD algorithm and compare it with EDI-MMD algorithm. Four basis materials were used. NaI and Gd(NO<sub>3</sub>)<sub>3</sub> represented the typical components of the contrast agents used in clinics, CaCl<sub>2</sub> and H<sub>2</sub>O represented the typical components of the human body. According to the concentrations in Refs. 11 and 49, all the concentrations (weight/volume) of the three solutes we used are listed in Table 5. In each spectral CT scanning, we used eight solutions with different concentrations in one row of Table 5. After six scans, we obtained 6 × 5 = 30 reconstructed images of different solutions and energy-bins.

Figures 18(a)–18(f) show the six reconstructed images in the first energy-bin of [26, 33] keV, and Figs. 18(g)–18(k) are the reconstructed images of the first phantom in all five energy-bins, whose display window are [−0.15, 0.7]. It shows that the images at higher energy-bins are darker than the ones at lower energy-bins, which means high energy x-ray photons were less absorbed than low energy x-ray photons. The diameter of centrifuge tubes contains 36 voxels in reconstructed images. Note that there are some ring artifacts in the images which were caused by the different detection response functions element-by-element, especially the abnormally high detection efficiency of the elements on the edges of the 32-element modules that comprised the whole detector array. Moreover, these elements suffered from severer pile-up distortion compared with other elements, resulting in further deviation from the Beer-Lambert law.

Table 5 Concentrations in centrifugal tubes.

Solute	Concentrations in the centrifugal tubes (mg/mL)							
NaI	0	2.9	3.2	3.5	3.8	4.1	4.4	4.7
NaI	5.0	5.3	5.6	5.9	6.2	6.5	6.8	7.1
Gd(NO <sub>3</sub> ) <sub>3</sub> · 6H <sub>2</sub> O	0	7.1	7.8	8.5	9.2	9.9	10.6	11.3
Gd(NO <sub>3</sub> ) <sub>3</sub> · 6H <sub>2</sub> O	12.0	12.7	13.4	14.1	14.8	15.5	16.2	16.9
CaCl <sub>2</sub> · 2H <sub>2</sub> O	0	200	240	280	320	360	400	440
CaCl <sub>2</sub> · 2H <sub>2</sub> O	480	520	560	600	640	680	720	760



**Fig. 18** (a)–(f) The spectral CT reconstructions with the first energy-bin [26, 33] keV. (a), (b) The reconstructed image of the NaI solutions, (c), (d) for  $Gd(NO_3)_3$  solutions, and (e), (f) for  $CaCl_2$  solutions. The display window is [0, 1]. (g)–(k) The reconstructed images of the NaI solutions in all five energy-bins. The display window is [−0.15, 0.7].

As shown in Fig. 18, the concentrations within the red circles were used to generate the validation dataset. Those within the green circles were used to generate the testing dataset. The rest were used for the training dataset. Because it was difficult to exactly obtain the effective spectrum of the x-ray beam used in experiment, we could not get the exact linear attenuation coefficients of every solutes. Thus, we used the maximum concentration solutions of three solutes and water as the basis materials to validate the decomposition algorithms.

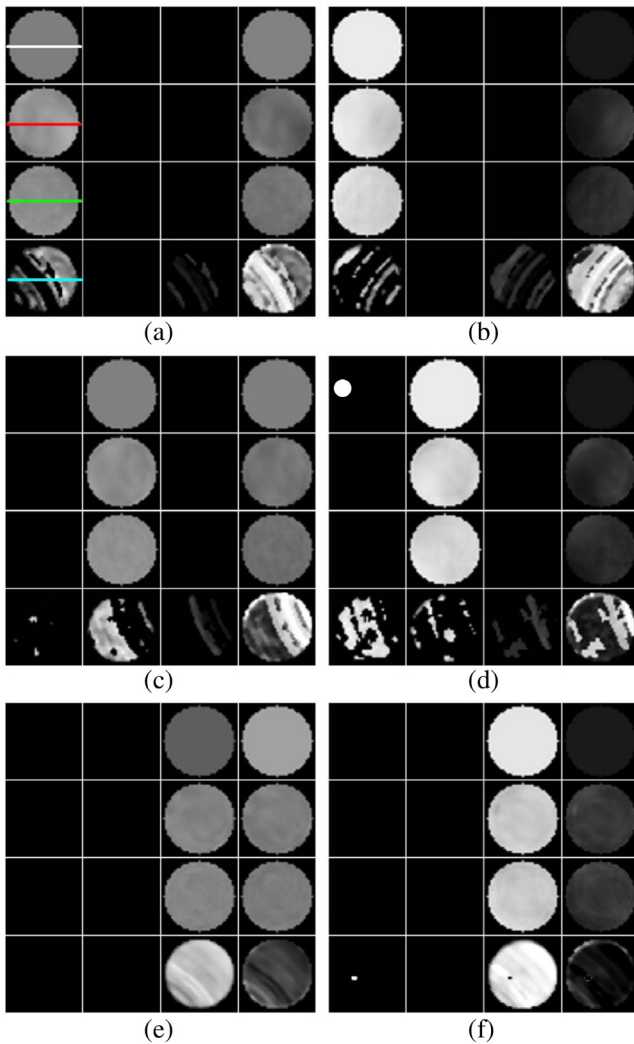
The regions within the centrifugal tubes were extracted to validate our CMD algorithm and compare it with EDI-MMD algorithm. To improve the generalization of the CNNs with limited training data, we randomly overturned the input patches.<sup>35</sup> We obtained 36,324 samples for the training dataset, and 6054 samples for the validation and testing datasets, respectively. Because there always had some noise and artifacts, we used a smaller learning rate than that used in the simulation. We set  $LR = 0.01$ ,  $P = 2.5$ ,  $P\_STEP = 0.5$ ,  $VAL\_NUM = 6054$ ,  $BATCH\_NUM = 128$ ,  $ITER\_$

$NUM = 64$ ,  $\epsilon = 1e-5$ , and  $\rho = 0.9$ . We also calculated the decomposition results by EDI-MMD algorithm as a contrast. The triplets set for EDI-MMD is in Table 6.

Figure 19 shows the decomposition results of six different concentration solutions highlighted in italics in Table 5. Figures 19(a)–19(f) are the results of 4.1 mg/mL NaI, 5.9 mg/mL NaI, 9.9 mg/mL  $Gd(NO_3)_3 \cdot 6H_2O$ , 14.1 mg/mL  $Gd(NO_3)_3 \cdot 6H_2O$ , 360 mg/mL  $CaCl_2 \cdot 2H_2O$ , and 600 mg/mL  $CaCl_2 \cdot 2H_2O$ , respectively. In

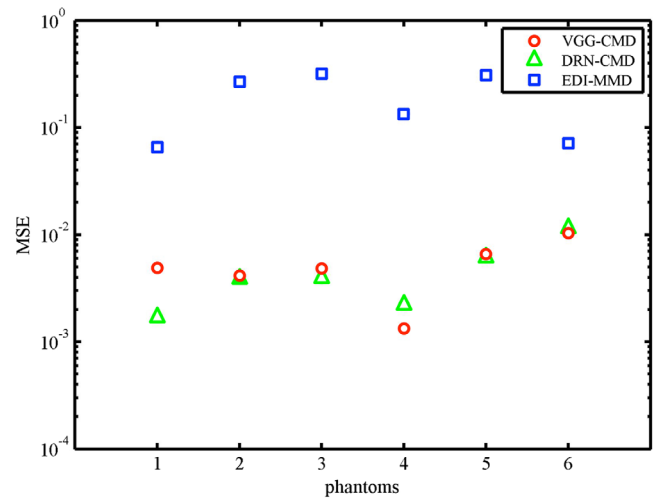
**Table 6** Triplets set for phantom in experiment.

Triplet ID	NaI (7.1 mg/mL)	$Gd(NO_3)_3 \cdot 6H_2O$ (16.9 mg/mL)	$CaCl_2 \cdot 2H_2O$ (760 mg/mL)	$H_2O$
1	✓			✓
2		✓		✓
3			✓	✓

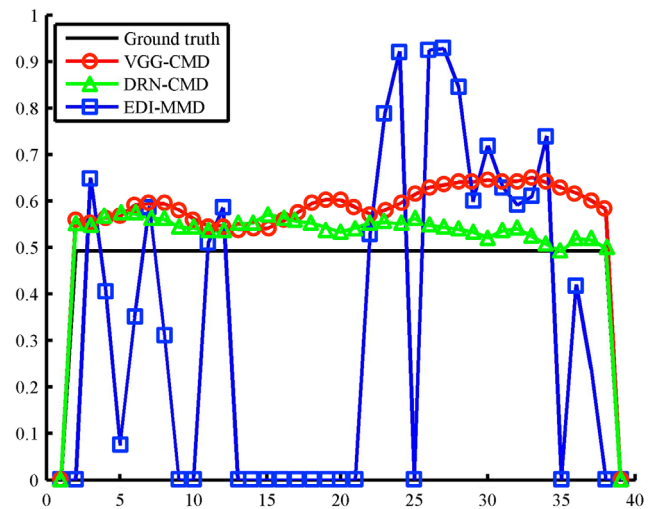


**Fig. 19** Decomposition results of italic cells in Table 5, namely, (a) 4.1 mg/mL NaI, (b) 5.9 mg/mL NaI, (c) 9.9 mg/mL  $\text{Gd}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$ , (d) 14.1 mg/mL  $\text{Gd}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$ , (e) 360 mg/mL  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , and (f) 600 mg/mL  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ . In each figure of (a–f), there are  $4 \times 4$  images. The rows from top to bottom are the ground truth, the decomposition results of VGG-CMD, DRN-CMD, and EDI-MMD algorithm. The columns from left to right are the results of NaI,  $\text{Gd}(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$ ,  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , and  $\text{H}_2\text{O}$ .

each figures [Figs. 19(a)–19(f)], the columns from left to right are the results of four basis materials, NaI,  $\text{Gd}(\text{NO}_3)_3$ ,  $\text{CaCl}_2$ , and  $\text{H}_2\text{O}$ , respectively. The rows from top to bottom are the ground truth, decomposition results of VGG-CMD, DRN-CMD, and EDI-MMD algorithms, respectively. We can find that both the VGG-CMD and DRN-CMD algorithms can provide good decomposition results. They are much better than the results of EDI-MMD algorithm which are seriously suffered from the ring artifacts existing in the spectral CT images. As shown in the last rows of Figs. 19(a)–19(d), the decomposition results are far from desire for the low concentration solutions. Moreover, the deeper network used in DRN-CMD algorithm can provide a smoother and more uniform results. In addition, compared with Ref. 37, the decomposition of traditional materials cannot achieve good results even if there is a clear prior knowledge (the solution consists of two basic materials: water and solvent) in the case of serious ring artifacts.



**Fig. 20** MSEs of the six test images in Fig. 19.



**Fig. 21** The profiles of the decomposition results of NaI locating at the lines in Fig. 19(a).

Figure 20 shows the decomposition MSE of three algorithms. The average MSE of six testing images are  $5.3719 \times 10^{-3} \text{ mm}^{-2}$  for VGG-CMD,  $4.8768 \times 10^{-3} \text{ mm}^{-2}$  for DRN-CMD, and  $1.9544 \times 10^{-1} \text{ mm}^{-2}$  for EDI-MMD. In general, the MSE of CMD algorithms are reduced more than one order comparing to the MSE of EDI-MMD. Figure 21 shows the profiles of the decomposition results of NaI locating at the colorful lines in Fig. 19(a). MSE of the decomposition section line are  $1.0645 \times 10^{-2}$ ,  $2.7356 \times 10^{-3}$ , and  $1.2634 \times 10^{-1} \text{ mm}^{-2}$  for three methods. The conclusion is similar to the numerical simulations. Although the CMD algorithm has never been trained by the testing solutions, the generalization of the CNNs guarantees its accuracy. Compared with the direct inversion decomposition method, CMD can resist the effects of ring artifacts that provide much more accurate and uniform decomposition results.

#### 4 Discussion

Compared with traditional direct inversion algorithms, the proposed algorithm improves the decomposition results

significantly especially for the cases that artifacts, e.g., ring and beam-hardening artifacts, existing in the CT images or the spectrum of x-ray source is distorted. For example, in current spectral CT systems using photon-counting detectors, the ring artifacts often appear due to the property inconsistency among the detector units (e.g., detection response function, dead time, pulse pile-up, etc.). As an image-domain method, no matter how the images are reconstructed, the proposed CMD algorithm can provide better decomposition results than direct inversion methods as shown in Fig. 19. It works well even the linear attenuation coefficients of the basis materials are close. Due to the robustness of CMD, it is also expected to be used in other incomplete-data or sparse-data CT cases in the future, e.g., limited-angle CT, few-view CT, and low-dose CT.<sup>46-48,49-52</sup>

Because the forward-propagation process in the CNN can be configured on GPU, the CMD algorithm can complete the material decompositions in a few seconds once it has been well trained. However, CMD also has some disadvantages. Traditional MMD algorithms in image domain only need the linear attenuation coefficients of the basis materials. When environment changes, we can update the algorithms easily. However, CMD method needs to gather training data and be retrained. The simulation 2 shows that the trained neural network is compatible with different imaging conditions, but it is almost impossible to generate every data for all the situations. For CMD, the biggest problem is that it is difficult to get the ground truth in practical situations for training. Even though we know all the material distributions of the designed phantoms, it is still difficult to ensure the diversity of data. We hope to train the neural network with simulation data as real as possible, and then decompose the materials directly on the real samples in our further experiments. Incorporating the models of the material decomposition and the physics of spectral CT, the direct iterative reconstruction and decomposition method sometimes provide good results. In the future, it is worth combining CMD algorithm with iterative reconstruction algorithm.

## 5 Conclusion

In this study, we proposed a deep-learning-based CMD algorithm to solve the MMD problem of spectral CT in image-domain. We redesigned two CNNs that worked well in the image classification field. Given the property of the Softmax layer, the output of the CMD can satisfy the mass and volume conservation constraints automatically. Simulation and experimental results proved the CMD algorithm can robustly solve the MMD problem of spectral CT.

## 6 Appendix

### 6.1 Layers Used in This Work

#### 6.1.1 Convolutional layer

The convolutional layer is the most important layer in a CNN. The advanced features in an image are detected by convolutional layers. A convolutional layer receives a three-rank tensor and outputs a three-rank tensor as well. A convolutional layer can be divided into the weight tensor  $\mathbf{W}(c_x, c_y, k, l)$  and the bias matrix  $\mathbf{b}(l)$ . Receiving a tensor  $\mathbf{T}(x, y, k)$ , the output tensor is

$$\mathbf{U}(x, y, l) = \left\{ \sum_{c_x} \sum_{c_y} \sum_k [\mathbf{T}(x + c_x, y + c_y, k) \times \mathbf{W}(c_x, c_y, k, l)] \right\} + \mathbf{b}(l). \quad (11)$$

Two types of convolutional operations are available for the edge of an image, i.e., using Eq. (11) to directly lead the size reduction or supplementing 0 around the input tensor to keep the size unchanged.

#### 6.1.2 Activation layer

As shown above, no matter how many convolutional layers are piled up, the effect is the same as the case of only one convolutional layer because the operation in the convolutional layer is linear.<sup>38</sup> To enhance the capability of the CNN, nonlinearity should be introduced. We explain two types of activation layers, each of which has its own features.

Rectified linear unit (ReLU) is a simple activation function that effectively introduces nonlinearity into CNNs and has thus been widely used since its discovery. ReLU sets the negative value of the output tensor of a convolutional layer to 0, that is,

$$\mathbf{U}(x, y, k) = \begin{cases} \mathbf{T}(x, y, k), & \mathbf{T}(x, y, k) > 0 \\ 0, & \mathbf{T}(x, y, k) \leq 0 \end{cases}. \quad (12)$$

ReLU simplifies the process of derivation and lightens the vanishing gradient problem of traditional artificial neural networks, resulting in the rapid convergence of networks.

Softmax is the last layer of a CNN. It receives a one-rank tensor and outputs a one-rank tensor:<sup>38</sup>

$$\mathbf{U}(x) = \frac{e^{\mathbf{T}(x)}}{\sum_{x'} e^{\mathbf{T}(x')}}. \quad (13)$$

Obviously, the output of the Softmax activation layer satisfies the constraints of Eq. (2). In image classification, the output of Softmax (in common, the Softmax activation layer is the last layer of the CNN, so it is also the output of a CNN) is considered as the possibility of each image type predicted by CNN. However, in the material decomposition problem, we train the output as the volume fraction of each material of the central pixel of the input patch.

#### 6.1.3 Batch-normalization layer

Similar to the ordered subset used in CT reconstruction, which can accelerate convergence, we send a certain number of samples to update the CNN at a time during training. This set of samples is called a minibatch. BN is set between the convolutional and activation layers to scale and translate the values in a batch uniformly and to improve the following activation layers.<sup>38,53</sup> BN accelerates convergence, relaxes the requirements of the initial values of networks, and makes very deep networks possible.

#### 6.1.4 Max pooling layer

The pooling layer is a screening operation of the input three-rank tensor. Given the strides of the  $x$  and  $y$  directions  $s_x$  and  $s_y$ , the output is



$$\mathbf{U}(x, y, k) = \max_{\substack{0 \leq i \leq s_x - 1 \\ 0 \leq j \leq s_y - 1}} \{\mathbf{T}[(s_x - 1)x + i, (s_y - 1)y + j, k]\}. \quad (14)$$

As shown in Eq. (14), the max pooling layer shrinks the size of the image and introduces nonlinearity into the CNN. The max pooling layer also preserves only one value of a region of the image. Therefore, small translations in the image will be ignored by the CNN. In another word, the translation invariance is improved.

### 6.1.5 Fully connected layer

An inchoate neural network consists of fully connected layers and activation layers. This composition is ineffective.<sup>38</sup> After extracting features through the convolutional layer, the fully connected layers can work effectively.

The fully connected layer receives a 1-rank tensor  $\mathbf{T}(x)$  and outputs a one-rank tensor:

$$\mathbf{U}(y) = \sum_x \mathbf{W}(x, y) \mathbf{T}(x) + \mathbf{b}(y). \quad (15)$$

In a CNN, the first two parameters of the image's size,  $x$  and  $y$ , will shrink when flowing through the convolutional and pooling layers. The third parameter  $k$  enlarges when operated by the convolutional layer. Therefore, the image becomes increasingly small while the number of its channels increases. Finally, we ignore the difference in the first two parameters, flatten the three-rank tensor image into a one-rank tensor, and send it into the fully connected layer.

## 6.2 Codes Available

Our trained codes, some synthetic and real datasets used in this paper can be downloaded from Github repository available at: [https://github.com/ZhengyangChen/Convolutional\\_material\\_decomposition](https://github.com/ZhengyangChen/Convolutional_material_decomposition).

### Acknowledgments

This work was supported in part by NSFC 11775124, 61571256, 11525521, 81427803, and the National Key Research and Development Program of China, 2017YFC0109103 and 2018YFC0115502.

### References

1. Y. Long and J. A. Fessler, "Multi-material decomposition using statistical image reconstruction for spectral CT," *IEEE Trans. Med. Imaging* **33**(8), 1614–1626 (2014).
2. M. Patino et al., "Material separation using dual-energy CT: current and emerging applications," *Radiographics* **36**(4), 1087–1105 (2016).
3. J. Chu et al., "Combination of current-integrating/photon-counting detector modules for spectral CT," *Phys. Med. Biol.* **58**(19), 7009–7024 (2013).
4. P. R. Mendonça, P. Lamb, and D. V. Sahani, "A flexible method for multi-material decomposition of dual-energy CT images," *IEEE Trans. Med. Imaging* **33**(1), 99–116 (2014).
5. L. Li et al., "Spectral CT modeling and reconstruction with hybrid detectors in dynamic-threshold-based counting and integrating modes," *IEEE Trans. Med. Imaging* **34**(3), 716–728 (2015).
6. Y. Xue et al., "Statistical image-domain multimaterial decomposition for dual-energy CT," *Med. Phys.* **44**(3), 886–901 (2017).
7. R. E. Alvarez and A. Macovski, "Energy-selective reconstructions in x-ray computerised tomography," *Phys. Med. Biol.* **21**(5), 733–744 (1976).
8. L. Lehmann et al., "Generalized image combinations in dual KVP digital radiography," *Med. Phys.* **8**(5), 659–667 (1981).
9. H. Xue et al., "A correction method for dual energy liquid CT image reconstruction with metallic containers," *J. X-Ray Sci. Technol.* **20**(3), 301–316 (2012).
10. L. Li et al., "A dynamic material discrimination algorithm for dual MV energy x-ray digital radiography," *Appl. Radiat. Isot.* **114**, 188–195 (2016).
11. P. Sukovic and N. H. Clinthorne, "Penalized weighted least-squares image reconstruction for dual energy x-ray transmission tomography," *IEEE Trans. Med. Imaging* **19**(11), 1075–1081 (2000).
12. K. Taguchi et al., "Image-domain material decomposition using photon-counting CT," *Proc. SPIE* **6510**, 651008 (2007).
13. Y. Zou and M. D. Silver, "Analysis of fast kV-switching in dual energy CT using a pre-reconstruction decomposition technique," *Proc. SPIE* **6913**, 691313 (2008).
14. C. Maaß, M. Baer, and M. Kachelrieß, "Image-based dual energy CT using optimized precorrection functions: a practical new approach of material decomposition in image domain," *Med. Phys.* **36**(8), 3818–3829 (2009).
15. T. Niu et al., "Iterative image-domain decomposition for dual-energy CT," *Med. Phys.* **41**(4), 041901 (2014).
16. M. Petrongolo, X. Dong, and L. Zhu, "A general framework of noise suppression in material decomposition for dual-energy CT," *Med. Phys.* **42**(8), 4848–4862 (2015).
17. R. F. Barber et al., "An algorithm for constrained one-step inversion of spectral CT data," *Phys. Med. Biol.* **61**(10), 3784–3818 (2016).
18. Y. Zhao, X. Zhao, and P. Zhang, "An extended algebraic reconstruction technique (E-ART) for dual spectral CT," *IEEE Trans. Med. Imaging* **34**(3), 761–768 (2015).
19. T. Zhao, L. Li, and Z. Chen, "K-edge eliminated material decomposition method for dual-energy x-ray CT," *Appl. Radiat. Isot.* **127**, 231–236 (2017).
20. J. W. Stayman and S. Tilley, "Model-based multi-material decomposition using spatial-spectral CT filters," [http://ai.jhu.edu/papers/CT2018\\_stayman.pdf](http://ai.jhu.edu/papers/CT2018_stayman.pdf) (2018).
21. A. M. Alessio and L. R. MacDonald, "Quantitative material characterization from multi-energy photon counting CT," *Med. Phys.* **40**(3), 031108 (2013).
22. P. R. Mendonça et al., "Multi-material decomposition of spectral CT images," *Proc. SPIE* **7622**, 76221W (2010).
23. G. Wang, "A perspective on deep imaging," *IEEE Access* **4**, 8914–8924 (2016).
24. Z. Chen and L. Li, "Preliminary research on multi-material decomposition of spectral CT using deep learning," in *Fully3D*, pp. 523–526 (2017).
25. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556 (2014).
26. J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural Networks* **61**, 85–117 (2015).
27. K. He et al., "Deep residual learning for image recognition," in *Computer Vision and Pattern Recognition*, pp. 770–778 (2016).
28. H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique," *IEEE Trans. Med. Imaging* **35**(5), 1153–1159 (2016).
29. P. Hu et al., "Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution," *Phys. Med. Biol.* **61**(24), 8676–8698 (2016).
30. P. Moeskops et al., "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Trans. Med. Imaging* **35**(5), 1252–1261 (2016).
31. M. Anthimopoulos et al., "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," *IEEE Trans. Med. Imaging* **35**(5), 1207–1216 (2016).
32. H.-C. Shin et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imaging* **35**(5), 1285–1298 (2016).
33. Q. Dou et al., "Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks," *IEEE Trans. Med. Imaging* **35**(5), 1182–1195 (2016).
34. J. M. Wolterink et al., "Generative adversarial networks for noise reduction in low-dose CT," *IEEE Trans. Med. Imaging* **36**(12), 2536–2545 (2017).
35. H. Chen et al., "Low-dose CT via convolutional neural network," *Biomed. Opt. Express* **8**(2), 679–694 (2017).
36. Y. Han, J. Kang, and J. C. Ye, "Deep learning reconstruction for 9-view dual energy CT baggage scanner," arXiv:1801.01258 (2018).
37. S. Dorn et al., "Prior-based multi material decomposition for dual energy CT," [https://www.dkfz.de/en/roentgenbilgebung/ct\\_ct\\_conference\\_contributions/Prior-Based-Multi-Material-Decomposition-for-DECT\\_SabrinaDorn.pdf?m=1527056293](https://www.dkfz.de/en/roentgenbilgebung/ct_ct_conference_contributions/Prior-Based-Multi-Material-Decomposition-for-DECT_SabrinaDorn.pdf?m=1527056293) (2018).
38. I. Goodfellow et al., *Deep Learning*, MIT Press, Cambridge (2016).
39. L. Li et al., "A tensor PRISM algorithm for multi-energy CT reconstruction and comparative studies," *J. X-Ray Sci. Technol.* **22**(2), 147–163 (2014).
40. J. Zhang, I. Mitliagkas, and C. Ré, "Yellowfin and the art of momentum tuning," arXiv:1706.03471 (2017).
41. A. Choromanska et al., "The loss surfaces of multilayer networks," in *AISTATS*, pp. 192–204 (2015).

42. S. Hochreiter and J. Schmidhuber, "Flat minima," *Neural Comput.* **9**(1), 1–42 (1997).
43. D. Soudry and E. Hoffer, "Exponentially vanishing sub-optimal local minima in multilayer neural networks," arXiv:1702.05777 (2017).
44. Keras, <https://keras.io/>.
45. Siemens, [https://bps-healthcare.siemens.com/cv\\_oem/radfn.asp](https://bps-healthcare.siemens.com/cv_oem/radfn.asp).
46. R. Li, L. Li, and Z. Chen, "Spectrum reconstruction method based on the detector response model calibrated by x-ray fluorescence," *Phys. Med. Biol.* **62**(3), 1032–1045 (2017).
47. E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.* **53**(17), 4777–4807 (2008).
48. L. Li et al., "Full-field fan-beam x-ray fluorescence computed tomography with a conventional x-ray tube and photon-counting detectors for fast nanoparticle bioimaging," *Opt. Eng.* **56**(4), 043106 (2017).
49. H. Q. Le and S. Molloy, "Least squares parameter estimation methods for material decomposition with energy discriminating detectors," *Med. Phys.* **38**(1), 245–255 (2011).
50. Z. Chen et al., "A limited-angle CT reconstruction method based on anisotropic TV minimization," *Phys. Med. Biol.* **58**(7), 2119–2141 (2013).
51. M. Chang et al., "A few-view reweighted sparsity hunting (FRESH) method for CT image reconstruction," *J. X-Ray Sci. Technol.* **21**(2), 161–176 (2013).
52. J. Wang et al., "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose x-ray computed tomography," *IEEE Trans. Med. Imaging* **25**(10), 1272–1283 (2006).
53. S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," arXiv:1502.03167 (2015).

Biographies of the authors are not available.