Residual W-shape Network (ResWnet) for Dual-energy Cone-beam CT Imaging

Xiao Jiang[a], Hehe Cui[a], Zihao Liu[a], Lei Zhu[a*]
[a]Department of Engineering and Applied Physics, University of Science and Technology of China, Hefei, Anhui, 230026 China

ABSTRACT

Deep learning has achieved great success in many medical imaging tasks without explicit solutions. In this work, learning method was applied to dual-energy cone-beam CT imaging. We proposed a Residual W-shape Network (ResWnet). ResWnet consists of three modules: scatter correction module $\mathcal{S}$, material decomposition module $\mathcal{M}$, decomposition denoising module $\mathcal{D}$. Both $\mathcal{S}$ and $\mathcal{D}$ use ResUnet architecture, and this lightweight model fuses multi-level features, achieving satisfied performance with a small number of parameters. $\mathcal{S}$ acts on dual-energy attenuation projections to reduce the scatter contaminations, and $\mathcal{D}$ acts on material composition projections to suppress the noise. $\mathcal{M}$ links the modules $\mathcal{S}$ and $\mathcal{D}$, and is used for domain transform from attenuation projections to material projections. This process could be approximated by polynomials with pre-calibrated parameters, that is, $\mathcal{M}$ is a known operator in proposed network with no trainable parameters. This helps to reduce model parameters and improve the performance with small training dataset. Using public head CT dataset, we simulated dual-energy cone-beam CT projections and material projections. Proposed ResWnet was trained, validated and tested on this simulated dataset, verifying its effectiveness in projection-domain scatter correction and low-noise decomposition.

Keywords: Cone-beam CT, Dual Energy CT, Resnet, Deep Learning

1. INTRODUCTION

Conventional CT measures the spatial distribution of x-ray linear attenuation coefficient (LAC) [1]. Dual-energy CT (DECT) [2], which scans object with two different x-ray spectrums, extends the measurement to the energy dimension, and the quantitative information provided by DECT facilitates various new applications, including but not limited to electron density/stopping power calculation [3], synthesis of monochromatic images [4], virtual-non-enhanced images [5]. After several decades development, DECT has become a powerful tool in clinical diagnosis [6]. Another widely used CT is cone-beam CT (CBCT) [7]. Taking advantages of high spatial resolution, large volume coverage and open structure, CBCT provides flexible image guidance in image-guided radiotherapy [8] and image-guided intervention [9], and the flexible geometry is also well suitable for some dedicated clinical tasks, such as breast CT, extremity CT and dental CT [10].

Recently, some groups investigated the feasibility of dual-energy cone-beam CT [11-13] that combines the advantages of DECT and CBCT. Our group implemented a rotation filter [14] configuration to acquired dual-energy data within single rotation. We further proposed a joint bilateral filtering-based algorithm to suppress the image streaks and amplified decomposition noise [15]. However, photon scatter, a major issue in CBCT imaging [16], was not taken into consideration in our previous research. The scatter contamination could severely degrade the imaging accuracy, hampering quantitative dual-energy imaging. Moreover, according to the dual-energy imaging theory [17], image-domain decomposition cannot provide accurate material composition, as well as cannot eliminate beam-hardening effects. Polynomial fitting-based projection-domain decomposition [18] tackles this issue but it is sensitive to projection noise [19].

This work aims to perform projection-domain scatter correction and material decomposition. To this end, we designed a Residual W-shape Network (ResWnet), which consists of two cascade ResUnets. In order to reduce the model parameters and achieve satisfied performance using small training dataset, a known decomposition operator was used to link the two ResUnet.

7th International Conference on Image Formation in X-Ray Computed Tomography,
edited by Joseph Webster Stayman, Proc. of SPIE Vol. 12304, 1230427
 · 0277-786X · doi: 10.1117/12.2646505

2. METHODOLOGY

2.1 Dual-Energy Cone-Beam Projection Model

The polyenergetic forward model for dual-energy cone-beam projections is written as:

$$\boldsymbol{p}_{l/h} = \boldsymbol{I}_{l/h} \int \omega_{l/h}(E) \exp\left(-\int \boldsymbol{\mu}(\boldsymbol{x}, E)\mathrm{d}l\right)\mathrm{d}E \tag{1}$$

where l/h means physical quantities under low/high incident x-ray spectrum, $\boldsymbol{p}$ represents transmission photons, ω represents spectrum, $\boldsymbol{\mu}$ stands for LAC. Due to low scanning dose and large volume coverage, CBCT projections suffer high noise level ($\boldsymbol{n}_{l/h}$) and severe scatter contamination ($\boldsymbol{s}_{l/h}$). Consequently, the actual measurements are modeled as:

$$\boldsymbol{p}_{l/h}^{ns} = \boldsymbol{p}_{l/h} + \boldsymbol{n}_{l/h} + \boldsymbol{s}_{l/h} \tag{2}$$

The superscript n, s represent noise and scatter, respectively. After log normalization, the line integrals with and with scatter and noise are:

$$\boldsymbol{q}_{l/h} = -\log\left(\int \omega_{l/h}(E) \exp\left(-\int \boldsymbol{\mu}(\boldsymbol{x}, E)\mathrm{d}l\right)\mathrm{d}E\right) \tag{3}$$

$$\boldsymbol{q}_{l/h}^{ns} = -\log\left(\int \omega_{l/h}(E) \exp\left(-\int \boldsymbol{\mu}(\boldsymbol{x}, E)\mathrm{d}l\right)\mathrm{d}E + \frac{\boldsymbol{n}_{l/h} + \boldsymbol{s}_{l/h}}{\boldsymbol{I}_{l/h}}\right) \tag{4}$$

2.2 Projection-Domain Material Decomposition

According to the dual-energy CT theory [17], the LAC could be decomposed as:

$$\boldsymbol{\mu}(\boldsymbol{x}, E) = \boldsymbol{\rho}_1(\boldsymbol{x})\phi_1(E) + \boldsymbol{\rho}_2(\boldsymbol{x})\phi_2(E) \tag{5}$$

$\phi_{1/2}$ is energy-dependent basis function, which could be interpreted as mass attenuation coefficient of two basis materials, then $\boldsymbol{\rho}_{1/2}$ is the density of the basis material correspondingly. Plug Eq. (5) into Eq. (3), we obtain:

$$\boldsymbol{q}_{l/h} = -\log\left(\int \omega_{l/h}(E) \exp(-\phi_1(E)\boldsymbol{q}_1 - \phi_2(E)\boldsymbol{q}_2)\,\mathrm{d}E\right) \tag{6}$$

where $\boldsymbol{q}_1 = \int \boldsymbol{\rho}_1(\boldsymbol{x})dl$, $\boldsymbol{q}_2 = \int \boldsymbol{\rho}_2(\boldsymbol{x})dl$. Given incident spectrum $\omega_{l/h}$ and two basis functions $\phi_{1/2}$, the material decomposition is an inverse problem which recovers $\boldsymbol{q}_{1/2}$ from $\boldsymbol{q}_{l/h}$. Since no explicit expression for this inversion, A commonly used analytical decomposition method is polynomial approximation [18]:

$$\boldsymbol{q}_1 = \sum_{i,j=0}^{N} \alpha_{i,j}\, \boldsymbol{q}_l^i \boldsymbol{q}_h^j, \qquad \boldsymbol{q}_2 = \sum_{i,j=0}^{N} \beta_{i,j}\, \boldsymbol{q}_l^i \boldsymbol{q}_h^j \tag{7}$$

where $\alpha_{i,j}$, $\beta_{i,j}$ are pre-calibrated coefficients using calibration phantom. Although this method could achieve high accuracy via increasing the polynomial order N, previous research revealed that this decomposition is sensitive to projection noise [19].

2.3 Residual W-shape Network

Figure 1 presents the ResWnet Architecture, which consists of scatter correction module, material decomposition module and decomposition denoising module. For simplicity, projections represent log-normalized projections hereafter.

Scatter Correction Module ($\mathcal{S}: \boldsymbol{q}_{l/h}^{ns} \mapsto \boldsymbol{q}_{l/h}$)

Scatter correction module aims to remove the noise and scatter signal in projections. In this work, scatter correction module $\mathcal{S}$ employs a Residual U-shape network (ResUnet) [20]. The paired $\boldsymbol{q}_l^{ns}, \boldsymbol{q}_h^{ns}$ are concatenated as two-channel input, which then passes six encoder blocks and five decoder blocks successively. To avoid gradient vanishing and exploding [21] in very deep networks, each encoder and decoder block adopts residual architecture [22] with a shortcut connection from input to output. Considering that $\boldsymbol{q}_{l/h}^{ns}$ and $\boldsymbol{q}_{l/h}$ share the same structure with only numerical difference, a global shortcut connection is applied to directly add raw projections to the output projections. Different from the conventional Unet, the last two encoder blocks in $\mathcal{S}$ keep the same number of features to reduce model parameters without compromising the performance.

Material Decomposition Module ($\mathcal{M}: \boldsymbol{q}_{l/h} \mapsto \boldsymbol{q}_{1/2}$)

Material decomposition module $\mathcal{M}$ aims to recover the material composition projections $\boldsymbol{q}_{1/2}$ from the dual-energy projections $\boldsymbol{q}_{l/h}$. As discussed in Sec.II.B, this inversion could be approximated by polynomials functions, and the parameters $\alpha_{i,j}, \beta_{i,j}$ could be determined by pre-calibration. Thus, there is no need to train a sub network for decomposition. In proposed network, $\mathcal{M}$ acts as depicted in Eqs. (7). N was set to 4 and there is no trainable parameter in this module.

Decomposition Denoising Module ($\mathcal{D}$: $\boldsymbol{q}_{1/2}^{s} \mapsto \boldsymbol{q}_{1/2}$)

Since the module $\mathcal{S}$ cannot reduce the noise level to zero and the polynomial-based decomposition is sensitive to noise, the output of $\mathcal{M}$ would be noisy. Decomposition denoising module $\mathcal{D}$ was used for further suppress the noise. Same as the scatter correction module, $\mathcal{D}$ also employs the ResUnet architecture

2.4 Data Generation and Network Training

To obtain the training, validation and testing datasets, we first download 22 head CT scans from public dataset in https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageid=39879146. These CT images were decomposed into four tissues: fat, muscle, 200mg/cc bone and 800mg/cc bone using multi thresholds segmentation. We performed polyenergetic forward projection of each volume using two spectrums generated by Spektr [23]. These are the label projections $\boldsymbol{q}_{l/h}$ for scatter correction. The input projections $\boldsymbol{q}_{l/h}^{ns}$ were then obtained via adding scatter and noise signals generated by Monte Carlo simulation. Muscle and 800mg/cc bone were selected as basis materials in this work, and fat and 200mg/cc bone were decomposed onto these two basis. The label projections of material decomposition $\boldsymbol{q}_{1/2}$ were generated by forward projecting the muscle and 800mg/cc composition images. Each head produced 100 projections, and the projection angles equally distributed between 0 and 2π.

During the model training, network $\mathcal{S}$ were firstly optimized by:

$$\mathcal{S}^{*} = \underset{\mathcal{S}}{\operatorname{argmin}}\, \mathrm{w}_1 \left\| \mathcal{S}\left(\boldsymbol{q}_{l/h}^{ns}\right) - \boldsymbol{q}_{l/h} \right\|_1 + \mathrm{w}_2 \left\| \mathcal{M}\left(\mathcal{S}\left(\boldsymbol{q}_{l/h}^{ns}\right)\right) - \mathcal{M}\left(\boldsymbol{q}_{l/h}\right) \right\|_1 \quad \#(8)$$

The second term was added because we hope $\mathcal{S}$ could not only reduce the scatter signal, but also produce a projection noise distribution that minimizes the decomposition noise. In this work, $\mathrm{w}_1, \mathrm{w}_2$ were set to 0.9 and 0.1, respectively. Using the trained network $\mathcal{S}^{*}$, the network $\mathcal{D}$ was finally trained by:

$$\mathcal{D}^{*} = \underset{\mathcal{D}}{\operatorname{argmin}} \left\| \mathcal{D}\left(\mathcal{M}\left(\mathcal{S}^{*}\left(\boldsymbol{q}_{l/h}^{ns}\right)\right)\right) - \boldsymbol{q}_{1/2} \right\|_1 \quad (9)$$

Parameters of both $\mathcal{D}$ and $\mathcal{S}$ were optimized by Adam optimizer with an initial learning rate of 0.0004 which decay 8% after each epoch. Batch size was set to 4 and training stopped after 100 epochs.

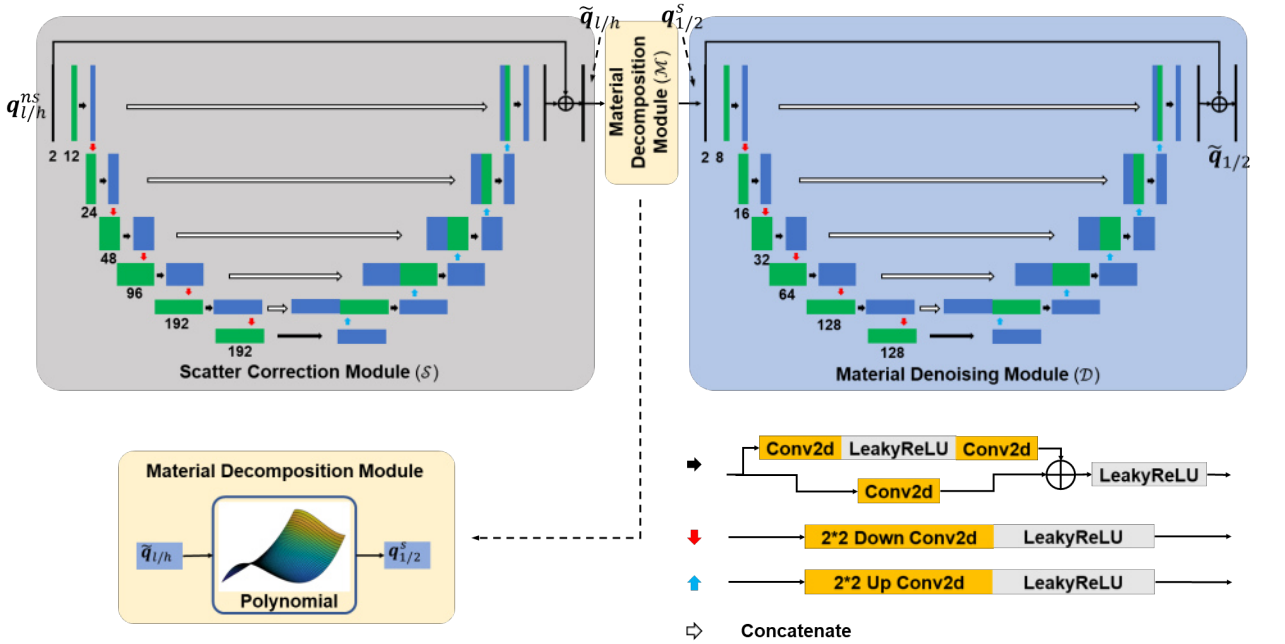

Figure 1. ResWnet architecture. Gray, blue and yellow boxes represent scatter correction, material decomposition, decomposition denoising module, respectively.

3. RESULTS

Corrected projections and decomposed material projections are obtained after the simulated cone-beam projections pass the first and the second ResUnet, respectively. Tomographic images are reconstructed via conventional FDK algorithm [24].

Figure 1 displays the dual-energy CT images. As in the left column, scatter contamination leads to obvious shading artifacts on soft tissue. The bone tissue, although clearly visualized, has a numerical error more than 10%. Proposed network successfully removes the image shading, with preservation of small bones and details in intracranial soft tissues. On both soft tissues and bone tissues, proposed network reduced the error to less than 0.5%.

Figure 2 displays the material composition images. As discussed above, polynomial-fitting based decomposition in the left column suffers terrible noise, especially on the soft tissue images. Without significantly compromising the spatial resolution, proposed network suppresses the noise of bone composition and muscle composition images by 32.7% and 65.6%, respectively. Subtle details around nasal cavity are faithfully recovered by proposed method as well.

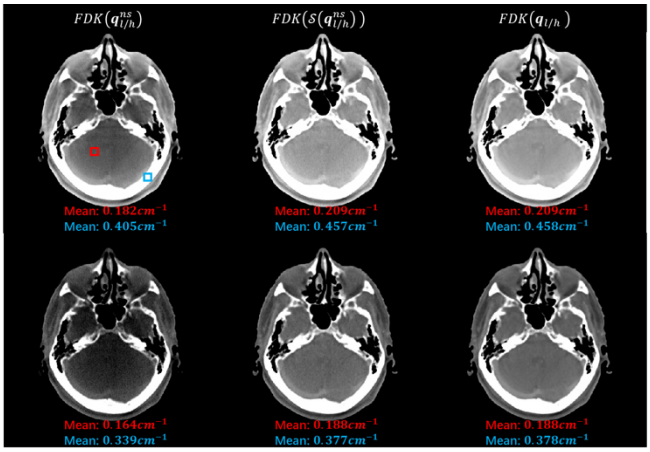

Figure 2. Dual-energy CT images. Top and bottom rows are low-energy and high-energy CT images, respectively. Mean values of two boxed areas are listed below each image. Display window: [0.15,0.25]cm^{-1}

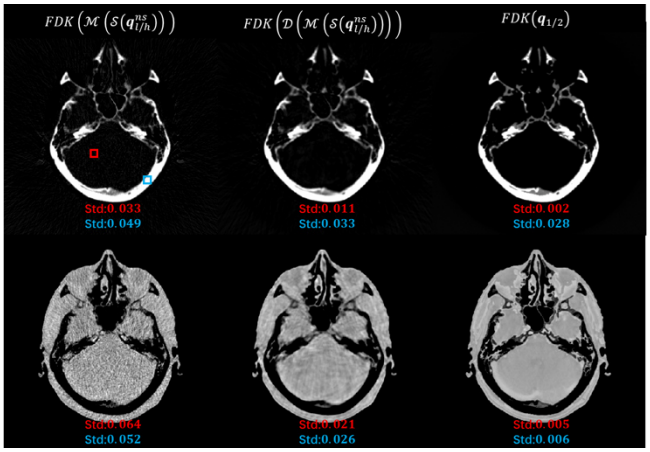

Figure 3. Material composition images. Top and bottom rows are bone and muscle images, respectively. Std values of two boxed areas are listed below each image. Display window: bone: [0,0.8], muscle: [0.6 1.2]

4. CONCLUSION AND DISCUSSION

In this work, we proposed a ResWnet for dual-energy cone-beam CT imaging. Three modules were designed for scatter correction, material decomposition and decomposition denoising, respectively. To reduce the model complexity and optimize the performance using small training dataset, both two trainable modules perform transform between the same domain, and the domain transform from attenuation projection to material composition projection is achieved by a known operator, i.e., polynomial-based decomposition. Simulation study preliminarily demonstrated the performance of proposed methods on scatter correction and low-noise decomposition.

Furthermore, some details lost in the material images, especially on the soft tissue images, sharp bone edges were also blurred to some extent. More complex model is needed to improve the performance. However, current study only used 2000 projections for training and validation, which is not capable of training more complex models. More cone beam projections will be simulated in the next step. Another issue is that current simulated projections only account for the primary signal, photon scatter and noise., and other factors such as off focus and detector glare are not considered. More accurate forward projector is needed to simulate real projections, then trained model could be applied to physical measurements acquired in the real system.

REFERENCES

[1] J. Hsieh, Computed tomography: principles, design, artifacts, and recent advances. SPIE press, 2003.
[2] A. Graser, T. R. Johnson, H. Chandarana, and M. Macari, "Dual energy CT: preliminary observations and potential clinical applications in the abdomen," European radiology, vol. 19, no. 1, p. 13, 2009.
[3] L. I. R. Garcia, J. F. P. Azorin, and J. F. Almansa, "A new method to measure electron density and effective atomic number using dual-energy CT images," Physics in Medicine & Biology, vol. 61, no. 1, p. 265, 2015.
[4] L. Yu, S. Leng, and C. H. McCollough, "Dual-energy CT–based monochromatic imaging," American journal of Roentgenology, vol. 199, no. 5_supplement, pp. S9-S15, 2012.
[5] L.-J. Zhang et al., "Liver virtual non-enhanced CT with dual-source, dual-energy CT: a preliminary study," European radiology, vol. 20, no. 9, pp. 2257-2264, 2010.
[6] M.-J. Kang, C. M. Park, C.-H. Lee, J. M. Goo, and H. J. Lee, "Dual-energy CT: clinical applications in various pulmonary diseases," Radiographics, vol. 30, no. 3, pp. 685-698, 2010.
[7] D. Jaffray and J. Siewerdsen, "Cone-beam computed tomography with a flat-panel imager: initial performance characterization," Medical physics, vol. 27, no. 6, pp. 1311-1323, 2000.
[8] X. Liang, Y. Jiang, and T. Niu, "Quantitative cone-beam CT imaging in radiotherapy: Parallel computation and comprehensive evaluation on the TrueBeam system," IEEE Access, pp. 2169-3536, 2019.
[9] M. Maybody, C. Stevenson, and S. B. Solomon, "Overview of navigation systems in image-guided interventions," Techniques in vascular and interventional radiology, vol. 16, no. 3, pp. 136-143, 2013.
[10] R. Fahrig, D. A. Jaffray, I. Sechopoulos, and J. W. Stayman, "Flat-panel conebeam CT in the clinic: history and current state," Journal of Medical Imaging, vol. 8, no. 5, p. 052115, 2021.
[11] R. Cassetta et al., "Fast-switching dual energy cone beam computed tomography using the on-board imager of a commercial linear accelerator," Physics in Medicine & Biology, vol. 65, no. 1, p. 015013, 2020.
[12] L. E. Schyns et al., "Optimizing dual energy cone beam CT protocols for preclinical imaging and radiation research," The British Journal of Radiology, vol. 90, no. 1069, p. 20160480, 2017.
[13] L. Shi et al., "Characterization and potential applications of a dual-layer flat-panel detector," Medical physics, vol. 47, no. 8, pp. 3332-3343, 2020.
[14] C. Fang, G. Xu, and L. Zhu, "Single scan dual energy cone beam CT using a rotating filter," in Medical Imaging 2020: Physics of Medical Imaging, 2020, vol. 11312, p. 113123S: International Society for Optics and Photonics.
[15] X. Jiang, C. Fang, P. Hu, H. Cui, L. Zhu, and Y. Yang, "Fast and effective single-scan dual-energy cone-beam CT reconstruction and decomposition denoising based on dual-energy vectorization," Medical Physics, vol. 48, no. 9, pp. 4843-4856, 2021.
[16] L. Zhu, Y. Xie, J. Wang, and L. Xing, "Scatter correction for cone-beam CT in radiation therapy," Medical physics, vol. 36, no. 6Part1, pp. 2258-2268, 2009.
[17] R. E. Alvarez and A. Macovski, "Energy-selective reconstructions in x-ray computerised tomography," Physics in Medicine & Biology, vol. 21, no. 5, p. 733, 1976.
[18] P. Stenner, T. Berkus, and M. Kachelriess, "Empirical dual energy calibration (EDEC) for cone-beam computed tomography," Medical physics, vol. 34, no. 9, pp. 3630-3641, 2007.
[19] M. Petrongolo, X. Dong, and L. Zhu, "A general framework of noise suppression in material decomposition for dual-energy CT," Medical physics, vol. 42, no. 8, pp. 4848-4862, 2015.
[20] Z. Liu and H. Yuan, "An Res-Unet Method for Pulmonary Artery Segmentation of CT Images," Journal of Physics: Conference Series, vol. 1924, no. 1, p. 012018 (6pp), 2021.
[21] S. Al-Abri, T. Lin, M. Tao, and F. Zhang, "A Derivative-Free Optimization Method With Application to Functions With Exploding and Vanishing Gradients," IEEE Control Systems Letters, vol. PP, no. 99, pp. 1-1, 2020.
[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.
[23] J. Punnoose, J. Xu, A. Sisniega, W. Zbijewski, and J. Siewerdsen, "spektr 3.0—A computational tool for x-ray spectrum modeling and analysis," Medical physics, vol. 43, no. 8Part1, pp. 4711-4717, 2016.
[24] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," Journal of the Optical Society of America A, vol. 1, no. 6, pp. 612-619, 1984.