

Convolutional Neural Network for Material Decomposition in Spectral CT Scans

Suzanne Bussod
CNRS UMR 5220, U1206
Univ. Lyon, INSA-Lyon
Lyon, France
suzanne.bussod@creatis.insa-lyon.fr

Juan F.P.J. Abascal
CNRS UMR 5220, U1206
Univ. Lyon, INSA-Lyon
Lyon, France
juan.abascal@creatis.insa-lyon.fr

Simon Arridge
Department of Computer Science,
University College London
London, United Kingdom
s.arridge@cs.ucl.ac.uk

Andreas Hauptmann
Research Unit of Mathematical Sciences
University of Oulu
Oulu, Finland
Department of Computer Science,
University College London
a.hauptmann@ucl.ac.uk

Christine Chappard
B3OA, UMR CNRS 7052, Inserm U1271
University Paris Diderot
Paris, France
christine.chappard@inserm.fr

Nicolas Ducros
CNRS UMR 5220, U1206
Univ. Lyon, INSA-Lyon
Lyon, France
nicolas.ducros@creatis.insa-lyon.fr

Françoise Peyrin
The European Synchrotron Radiation Facility
Grenoble, France
CNRS UMR 5220, U1206 Univ. Lyon, INSA-Lyon
Lyon, France
peyrin@esrf.fr

Abstract—Spectral computed tomography acquires energy-resolved data that allows recovery of densities of constituents of an object. This can be achieved by decomposing the measured spectral projection into material projections, and passing these decomposed projections through a tomographic reconstruction algorithm, to get the volumetric mass density of each material. Material decomposition is a nonlinear inverse problem that has been traditionally solved using model-based material decomposition algorithms. However, the forward model is difficult to estimate in real prototypes. Moreover, the traditional regularizers used to stabilized inversions are not fully relevant in the projection domain.

In this study, we propose a deep-learning method for material decomposition in the projection domain. We validate our methodology with numerical phantoms of human knees that are created from synchrotron CT scans. We consider four different scans for training, and one for validation. The measurements are corrupted by Poisson noise, assuming that at most 10^5 photons hit the detector. Compared to a regularized Gauss-Newton algorithm, the proposed deep-learning approach provides a compromise between noise and resolution, which reduces the computation time by a factor of 100.

Index Terms—Spectral computed tomography, Convolutional neural network, Material decomposition, Knee phantom

We acknowledge the support of the ANR project SALTO (ANR-17-CE19-0011-01). This study was also performed within the framework of LabEx PRIMES (ANR-11-LABX-0063) of the Université de Lyon. The project also received funding from the European Union Horizon 2020 Research and Innovation Programme under both grant agreement N° 668142 and Marie Skłodowska-Curie grant agreement N° 701915. This study was also carried out in the context of France Life Imaging (grant ANR-11-INBS-0006). A PhD is funded by the ED EEA of Université de Lyon.

I. INTRODUCTION

Spectral computed tomography (SCT) acquires energy-resolved data based on photon-counting detectors, which allows decomposition of a sample into maps of its different materials [1], [2]. Among other applications [3], SCT has been proposed for detection of atherosclerosis in cardiovascular disease [4], [5], and more recently for osteoarthritis, with iodine used for contrast [6]. With SCT, the contrast between the different types of soft tissues is improved.

Material decomposition is a nonlinear, nonconvex, ill-posed inverse problem [7]. Regularized variational methods are state-of-the-art for material decomposition [7], [8]. The regularized Gauss-Newton (RGN) method was shown to outperform the commonly used maximum likelihood estimation. In our previous study, we investigated the use of the RGN approach for application to osteoarthritis [9]. However, such model-based material decomposition methods require knowledge of the detector response function of the scanner, the source spectrum, and the attenuation of each material with the energy, which cannot be accurately known in practice. Moreover, the detector response function suffers from pixel-to-pixel deviations, and can change over time [1], [8].

Deep-learning methods have been proposed to solve inverse problems, including image denoising and image reconstruction, with limited or noisy data [10]–[12]. Recently, deep-learning methods were proposed to solve the material decom-

position problem in the image domain [13]. As decomposition in the image domain can lead to artifacts, decomposition in the projection domain is to be preferred [3]. However, decomposition in the projections has not been addressed using deep-learning approaches. Moreover, in previous studies, deep-learning methods were not compared to variational approaches.

In this study, we propose a deep-learning approach for solving the material decomposition problem of SCT in the projection domain, and we compare this to the RGN method [8], [9]. Our deep-learning approach is based on a U-net convolutional neural network [14], to decompose projections at different energy bins into projections for different materials. Our method is assessed for application to osteoarthritis using unique human knee numerical phantoms created from synchrotron CT acquisitions that were presented in [9]. The methods are then compared in terms of mean squared error (MSE), structural similarity (SSIM), and noise, computed as the standard deviation within a homogeneous region.

II. SPECTRAL COMPUTED TOMOGRAPHY

A. Forward model

We assume an object that is composed of M materials with V voxels scanned with a sensor that acquires I energy bins over P pixels and Θ projections. Let

$$\mathbf{s} = (s_{1,1}^1, \dots, s_{i,p}^\theta, \dots, s_{I,P}^\Theta)^\top$$

be the spectral measurement vector, where $s_{i,p}^\theta$ is the data acquired in the i -th energy bin, at the p -th pixel, and for the θ -th projection. Similarly, let

$$\boldsymbol{\rho} = (\rho_{1,1}, \dots, \rho_{m,v}, \dots, \rho_{M,V})^\top$$

be the unknown mass densities vector, where $\rho_{m,v}$ is the mass density of the m -th material at the v -th voxel of the sample. We note

$$\mathbf{s} = G(\boldsymbol{\rho}) \quad (1)$$

where G is the full forward model that maps $\boldsymbol{\rho}$ onto \mathbf{s} . The goal of SCT is to invert (1).

B. Two-step inverse problem

The forward model G can be seen as the composition of the X-ray transform X and a spectral mixing operator F . The X-ray transform X applies to each material independently. If we let $\mathbf{a}_m = (a_{m,1}^1, \dots, a_{m,p}^\theta, \dots, a_{m,P}^\Theta)^\top$ be the projected mass density of the m -th material, we have

$$\mathbf{a}_m = X(\boldsymbol{\rho}_m), \quad 1 \leq m \leq M, \quad (2)$$

where $\boldsymbol{\rho}_m = (\rho_{m,1}, \dots, \rho_{m,v}, \dots, \rho_{m,V})^\top$ is the mass density of the m -th material.

Spectral mixing applies to each projection angle independently; i.e.,

$$\mathbf{s}^\theta = F(\mathbf{a}^\theta), \quad 1 \leq \theta \leq \Theta, \quad (3)$$

where $\mathbf{s}^\theta = (s_{1,1}^\theta, \dots, s_{i,p}^\theta, \dots, s_{I,P}^\theta)^\top$ is the spectral measurement for the θ -th projection angle, and $\mathbf{a}^\theta =$

$(a_{1,1}^\theta, \dots, a_{m,p}^\theta, \dots, a_{M,P}^\theta)^\top$ are the mass densities projected at the θ -th projection angle. We consider the standard nonlinear mixing used in [1], [8]; i.e.,

$$s_{i,p}^\theta = \int_{\mathcal{E}} n_0(E) d_i(E) \exp \left[- \sum_{m=1}^M a_{m,p}^\theta \tau_m(E) \right] dE \quad (4)$$

where \mathcal{E} is the energy range of the source, n_0 is the source energy spectrum, d_i is the detector response function of the i -th bin, and τ_m is the m -th material mass attenuation coefficient.

Equations (2) and (3) suggest a two-step approach where the spectral projections are first decomposed into material projections, and then a standard CT algorithm is used to reconstruct each material density volume from the decomposed material projections. The first step, which is referred to as material decomposition, is the main focus of this work.

C. Model-based material decomposition

Material decomposition aims to invert (3). Assuming the sample is composed of $M = 2$ materials, as bone and soft tissue here, we seek the solution by minimizing the cost function

$$C(\mathbf{a}^\theta) = \frac{1}{2} \|\mathbf{F}(\mathbf{a}^\theta) - \mathbf{s}^\theta\|_{\mathbf{W}}^2 + \alpha \mathcal{R}(\mathbf{a}_{\text{soft}}^\theta, \mathbf{a}_{\text{bone}}^\theta), \quad (5)$$

where α is a regularization parameter, \mathcal{R} is a regularization functional, and $\mathbf{a}_{\text{soft}}^\theta$ and $\mathbf{a}_{\text{bone}}^\theta$ are the projected mass densities of soft tissues and bone, respectively. The first term of the cost function is a weighted least squares data fidelity term, with $\mathbf{W} = \mathbf{Diag}(1/s^\theta)$, which is chosen to handle data corrupted by Poisson noise [15]. The second term is the regularization term that conveys prior knowledge of the solution. As in [8], regularization consists of first-order and second-order Tikhonov regularization for bone and soft tissue, respectively, which has a denoising effect.

To minimize (5), we use the Gauss-Newton algorithm implemented in the Matlab toolbox SPRAY [8], [16].

III. DEEP LEARNING FOR MATERIAL DECOMPOSITION

A. Proposed approach

In this study, we propose to learn mapping from the measured spectral projection to the projected material maps; i.e.,

$$h_{\boldsymbol{\beta}}: \hat{\mathbf{s}}^\theta \mapsto \mathbf{a}^\theta \quad (6)$$

where $\hat{\mathbf{s}}$ is the normalized data with $\hat{\mathbf{s}}^\theta = \ln(\bar{\mathbf{s}}^\theta / \mathbf{s}^\theta)$, where $\bar{\mathbf{s}}^\theta$ represents the measurement in the absence of the object, i.e., $\bar{\mathbf{s}}^\theta = F(0)$, and $\boldsymbol{\beta}$ represents the parameters of the mapping. The log allows the data to be linearized, as the CT model is exponential [17], and the data space to be restricted [18]. As shown in Fig. 1, we consider a convolutional neural network with a U-net architecture.

Given N training input-output pairs $\{\hat{\mathbf{s}}^{(n)}, \mathbf{a}^{(n)}\}$, $1 \leq n \leq N$, where N is the number of samples multiplied by the

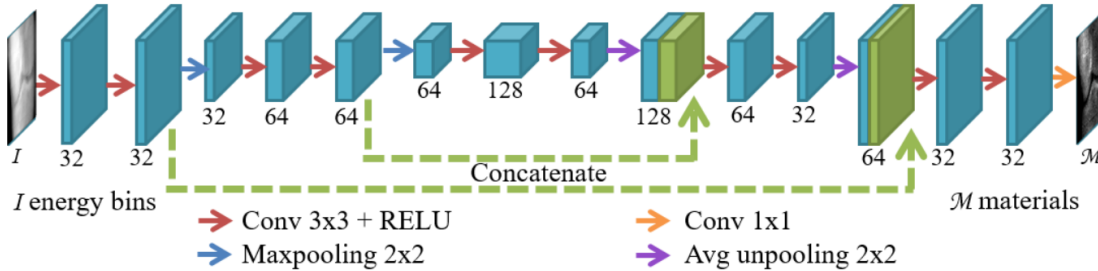


Fig. 1: U-Net architecture used to recover M material images from I spectral images at a given projection angle.

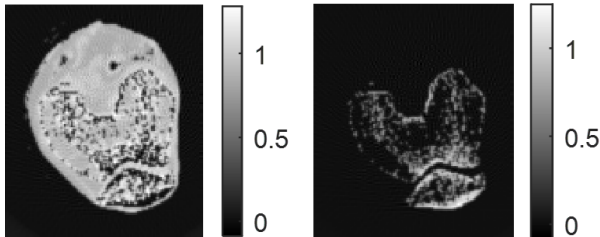


Fig. 2: Representative numerical phantom of the material mass densities obtained from the synchrotron CT scans. Single slices are shown for soft tissue (left) and bone tissue (right).

number of projections Θ , we consider the following loss function

$$L(\beta) = \sum_{n=1}^N \|\mathbf{h}_{\beta}(\hat{\mathbf{s}}^{(n)}) - \gamma \mathbf{a}^{(n)}\|^2. \quad (7)$$

where γ is a normalization parameter to put soft tissue and bone at the same level. This provides weighting of 8 for the soft tissue image, and 3 for the bone image. This weighting corresponds to the average maximum level of each material map. The inputs of the networks are $I = 4$ energy channels, each of size 51×161 , and the outputs are $M = 2$ material channels, each of size 51×161 . We minimize (7) using the adaptive moment estimation, or Adam algorithm [19], using PyTorch [20] running on a GeForce NVIDIA RTX 2080 Ti graphics card. The learning rate is set to 10^{-3} , and the batch size to 15. We use an early stopping criterion of 200 epochs.

B. Data

1) *Knee acquisitions*: Scans of eight knee samples at different stages of osteoarthritis were provided by the Institut d'Anatomie Paris Descartes. All of the knees were scanned using the synchrotron CT at the European Synchrotron Radiation Facility (Grenoble, France) [21]. The synchrotron CT scans were performed with a voxel size of $50 \mu\text{m}$ and at 55 keV, an energy that provides improved contrast for cartilage visualization. The study was approved by the Ethics Committee of Descartes University, Paris.

2) *Generation of material mass density phantoms*: We create eight realistic numerical phantoms from the synchrotron CT scans. All of the synchrotron volumes are subsampled by a

factor of 5 to reach a resolution of $250 \mu\text{m}$, which corresponds to the resolution of current SCT prototypes. For each knee, we automatically segment soft tissue and bone using a K-means algorithm [22]. After segmentation, the density of each material is taken from the synchrotron volume, assuming that only one material is present in each voxel. Fig. 2 shows a representative slice of the numerical phantoms obtained.

3) *Training data*: The material mass density phantoms are used to simulate SCT data according to (2) and (3). First, we again subsample the volumes, here by a factor of 4, to obtain a final resolution of 1 mm. We implement data augmentation by rotating the volumes around an axis perpendicular to the detector plane (-10° to 10°). We also perform two-dimensional scaling of the volumes using a factor ranging from 0.55 to 1.45. Then, the material density volumes are projected according to (2). We use the Matlab function `radon` and consider $\Theta = 720$ projections over 180° . Each volume is augmented to obtain 156 volumes, which leads to $N = 156 \times 720 = 112,320$ projections of size 51×161 . Finally, spectral measurements are simulated according to (4) using the Matlab toolbox `SPRAY` [16]. We consider the same detector response function as in [1] ($I = 4$ energy bins) and a conventional 120 kVp X-ray source. Spectral measurements are corrupted by Poisson noise; the maximum number of photons is set to 10^5 .

4) *Training procedure*: The knee phantoms are split as follows: four for training (449,280 projections), one for validation (720 projections), and three for testing (2,160 projections). We use the training dataset to update the weights of the network, the validation dataset to control for overfitting during training, and the test dataset to determine whether our algorithm can be generalized to unseen samples.

After material decomposition, we reconstruct the three-dimensional density volumes using filtered back projection (Matlab function `iradon`).

IV. RESULTS

Fig. 3 shows the material projections decomposed using the model-based RGN algorithm and the convolutional neural network methods, together with the decomposition error. For the RGN method, we display the results for two regularization parameters ($\alpha = 0.03$, $\alpha = 3$). Table I gives the MSE, SSIM, and noise values for both of the decomposed material

projections at 90° (as it is consistent along all projections), and the reconstructed material maps. We evaluate noise as the standard deviation in a (homogeneous) region of interest. The SSIM is also computed in a homogeneous region of interest, and the MSE is computed for the whole projection. The RGN algorithm with the lowest regularization parameter has higher MSE and SSIM, but also higher noise, compared to the higher regularization parameters. U-Net leads to low-noise decompositions with good compromise for the MSE and SSIM. In Fig. 3, it can be seen that for the RGN method, a low regularization parameter gives noisy decomposition and a high regularization parameter gives blurry images. Visually, U-net gives better results, as the images have both low noise and sharp detail. Moreover, U-net decomposes a whole knee, i.e., 720 projections, in 40 s, while the RGN method needs 1 h for central processing units. Using graphic processing units, the U-net decomposition of the whole knee is reduced to only 2 s.

V. DISCUSSION

Here, we propose a convolutional neural network approach based on the U-net architecture for learning material decomposition for SCT, and we compare this to the RGN method. The algorithms are assessed on numerical phantoms created from human knee synchrotron CT data [9]. A similar approach has already been applied for material decomposition in the image domain [13]. Here, we applied the projection domain, and we show that this gives comparable results to those for the model-based RGN method. Moreover, the RGN method decomposes materials under the assumption of ideal energy responses of the scanner, which is not realistic in practice. Thus, we expect model-based approaches to underperform learning approaches in nonideal situations.

The RGN method with low regularization leads to the best results in terms of the MSE and SSIM, but the images are highly corrupted by noise. The RGN method with high regularization is better in terms of noise, but the images are very blurred. U-net gives a good compromise between noise and recovery of detail. This can be explained because

TABLE I: Quantitative data for the decomposed projections and the reconstructed materials.

		Projected soft tissue	Projected bone	Soft tissue	Bone
MSE	RGN	0.020	0.009	0.0013	0.0006
	$\alpha = 0.03$				
	RGN $\alpha = 3$	0.021	0.010	0.0038	0.0018
	U-net	0.038	0.015	0.0032	0.0015
SSIM	RGN	0.89	0.84	0.96	0.96
	$\alpha = 0.03$				
	RGN $\alpha = 3$	0.91	0.67	0.51	0.52
	U-net	0.91	0.72	0.69	0.69
Noise	RGN	0.76	0.43	0.32	0.26
	$\alpha = 0.03$				
	RGN $\alpha = 3$	0.74	0.39	0.18	0.18
	U-net	0.71	0.46	0.21	0.19

U-net implicitly learns the prior probability density of the decomposed materials, while this is assumed for the RGN method. Moreover, deep learning allows the computation time to be reduced (after training) by a factor of 100 compared to the RGN method.

Our methodology has some limitations. The low resolution of the phantom used does not allow us to visualize all of the detail in the cartilage, and thus to assess possible improvements in cartilage characterization. The level of noise used can also be considered to be low in comparison to realistic acquisition conditions. In addition, the simulations used are phantoms with only two materials.

For further studies, we will assess the potential applications of the proposed approach for early detection of cartilage deterioration in osteoarthritis. For this, we will use high-resolution images and lower numbers of photons, which corresponds to more realistic conditions. We expect that the convolutional neural network approach will be more robust to noise, while the RGN method will be highly compromised. To visualize the finer structures, such as the cartilage in the knee joint, we will need to work at the same resolution as the SCT (i.e., $250\mu\text{m}$). In addition, this method should be validated on experimental data. Finally, a possible improvement to the proposed method will be to decompose the images into three materials within the cartilage, which would lead to automatic segmentation of the cartilage.

VI. CONCLUSION

In this study, we propose a U-net approach for material decomposition in the projection domain for application to knee osteoarthritis. We compare this to a RGN method. Our neural network provides a compromise between image noise and recovery of detail, which can be essential for detection of cartilage deterioration in osteoarthritis. In addition, our approach is rapid and does not require knowledge of the energy response of the scanner, which makes it a good candidate for future scanners.

ACKNOWLEDGMENT

The authors thank the European Synchrotron Radiation Facility for the allocated beam time within the MD1045 experiment at the ID17 beamline, and we are grateful to Alberto Bravin, Simon Rit, Claude Goubet, Tom Hohweiller, Cyril Mory, Cecile Olivier, Odran Pivot, Adrien Halty, Gaelle Mitton, Olga Kochebina, and Herwig Requardt for assistance in the acquisition of the synchrotron CT images. We thank Cecile Olivier for processing the European Synchrotron Radiation Facility data.

REFERENCES

- [1] J. Schlomka, E. Roessl, R. Dorschied, S. Dill, G. Martens, T. Istel, C. Bäumer, C. Herrmann, S. R. G. Zeitler, A. Livne, and R. Proksa, "Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography," *Physics in Medicine and Biology*, vol. 53, no. 15, pp. 4031–4047, Aug. 2008.
- [2] K. Taguchi and J. Iwanczyk, "Vision 20/20: Single photon counting x-ray detectors in medical imaging," *Medical Physics*, vol. 40, no. 10, p. 100901, Oct. 2013.

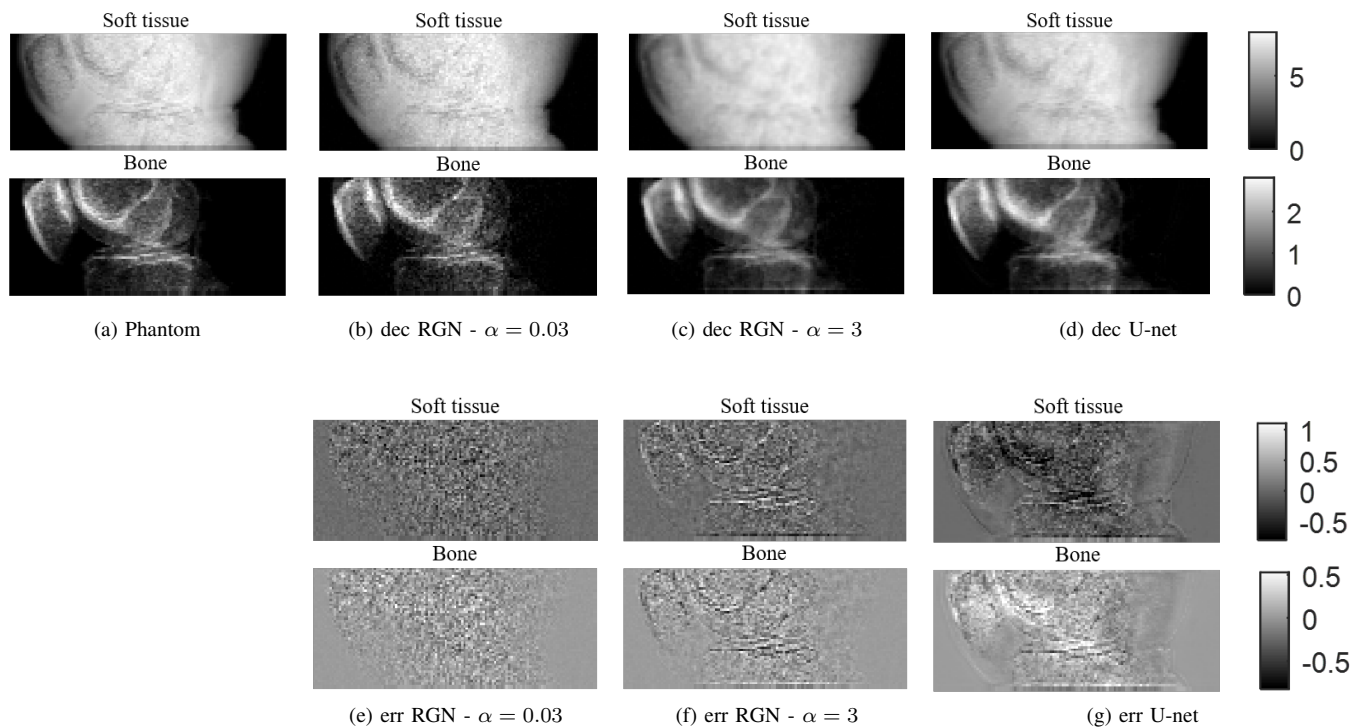


Fig. 3: Decomposed projections for an angle of 90° and a maximum number of photons of 10^5 . (a) Ground truth for the decomposition. (b, c) Decomposed materials with the Gauss-Newton method with regularization parameters of $\alpha = 0.03$ (b) and $\alpha = 3$ (c). (d) Decomposed materials with U-net. (e, f) Errors between the decomposed materials with the Gauss-Newton method and the ground truth, with the regularization parameters of $\alpha = 0.03$ (e) and $\alpha = 3$ (f). (g) Error between the decomposed materials with U-net and the ground truth.

- [3] B. J. Heismann, B. T. Schmidt, and T. G. Flohr, *Spectral Computed Tomography*. SPIE Press, 2012.
- [4] D. Cormode, E. Roessl, A. Thran, T. Skajaa, G. R.E., J. Schlomka, V. Fuster, E. Fisher, W. Mulder, R. Proksa, and Z. Fayad, "Atherosclerotic plaque composition: analysis with multicolor CT and targeted gold nanoparticles," *Radiology*, vol. 256, no. 3, pp. 774–782, Sep. 2010.
- [5] S. Si-Mohamed, A. Thivolet, P.-E. Bonnot, D. Bar-Ness, V. Képénékian, D. P. Cormode, P. Douek, and P. Rousset, "Improved Peritoneal Cavity and Abdominal Organ Imaging Using a Biphasic Contrast Agent Protocol and Spectral Photon Counting Computed Tomography K-Edge Imaging," *Investigative Radiology*, vol. 53, no. 10, pp. 629–639, 2018.
- [6] K. Rajendran, C. Löbker, S. B.S., C. Bateman, R. Younis, N. de Ruiter, A. Chernoglazov, M. Ramyar, G. Hooper, A. Butler, T. Woodfield, and N. Anderson, "Quantitative imaging of excised osteoarthritic cartilage using spectral CT," *European Radiology*, vol. 27, May 2016.
- [7] J. F. Abascal, F. Peyrin, and N. Ducros, "Nonlinear material decomposition using a regularized iterative scheme based on the bregman distance," vol. 34, no. 12, p. 124003, 2018.
- [8] N. Ducros, J. F. Abascal, B. Sixou, S. Rit, and F. Peyrin, "Regularization of Nonlinear Decomposition of Spectral X-ray Projection Images," *Medical Physics*, vol. 44, no. 9, pp. e174–e187, 2017.
- [9] S. Bussod, J. F. Abascal, N. Ducros, C. Olivier, S. Si-Mohamed, P. Douek, C. Chappard, and F. Peyrin, "Human Knee Phantom for Spectral CT: Validation of a Material Decomposition Algorithm," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, Apr. 2019, pp. 1170–1173.
- [10] M. McCann *et al.*, "Convolutional Neural Networks for Inverse Problems in Imaging: A Review," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 85–95, 2017.
- [11] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, and D. Firmin, "DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction," *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1310–1321, 2018.
- [12] A. Hauptmann, F. Lucka, M. Betcke, N. Huynh, J. Adler, B. Cox, P. Beard, S. Ourselin, and S. Arridge, "Model-Based Learning for Accelerated, Limited-View 3-D Photoacoustic Tomography," *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1382–1393, 2018.
- [13] D. Clark, M. Holbrook, and C. T. Badea, "Multi-energy CT decomposition using convolutional neural networks," in *Medical Imaging 2018: Physics of Medical Imaging*, vol. 10573. SPIE, 2018, p. 59.
- [14] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation." Springer, C., 2015, pp. 234–241.
- [15] J. A. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE transactions on medical imaging*, vol. 13, no. 2, pp. 290–300, 1994.
- [16] SPRAY, "Spectral X-ray image reconstruction toolbox," <https://www.creatis.insa-lyon.fr/ducros/WebPage/spray.html>.
- [17] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*. SPIE Press, 2003, google-Books-ID: JX__ILLXFHkC.
- [18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.
- [19] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization." ICLR, San Diego, 2015.
- [20] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," in *NIPS Autodiff Workshop*, 2017.
- [21] M. Salomé, F. Peyrin, P. Cloetens, C. Odet, A. M. Laval-Jeantet, J. Baruchel, and P. Spanne, "A synchrotron radiation microtomography system for the analysis of trabecular bone samples," *Medical Physics*, vol. 26, no. 10, pp. 2194–2204, Oct. 1999.
- [22] S. Lloyd, "Least Squares Quantization in PCM," pp. 129–137, 1982.