Deep learning methods for virtual monoenergetic imaging from spectral CT

Suzanne Bussod, Nicolas Ducros, Juan F P J Abascal, Philippe Douek, Christine Chappard, Françoise Peyrin

Deep Learning Methods for Virtual Mono Energetic Imaging from Spectral CT

Bussod Suzanne[⋆], Ducros Nicolas[⋆], Abascal Juan FPJ[⋆], Douek Philippe[⋆], Chappard Christine[†], Peyrin Françoise[⋆,◇]

[⋆] Univ Lyon, INSA-Lyon, CNRS, Inserm, CREATIS UMR 5220, U1294, Lyon, France
[†] Inserm U1271, UMR CNRS 7052, Univ Paris B3OA Lab, France
[◇] European Synchrotron Radiation Facility (ESRF), Grenoble, France

***Abstract*—Spectral photon counting computed tomography (CT) is a X-ray imaging modality that acquires energy-resolved data thanks to photon counting detectors that sort photons depending on their energy. This allows to decompose the object into its material constituents or to reconstruct virtual monoenergetic images. In this paper, we address for the first time the reconstruction of virtual monoenergetic images from spectral CT measurements, which is a non linear inverse problem, focusing on the application to knee osteoarthrisis.**

While traditional methods are based on the inversion of a physical model, deep learning methods have recently demonstrated their ability to solve inverse problems. In this work, we propose several physics-informed deep learning strategies for virtual monoenergetic image reconstruction. We consider four different reconstruction algorithms for the recovery of virtual monoenergetic images in the projection and in the image domain. All of our algorithms include a variant of the U-net convolutional neural network. The proposed algorithms were trained and evaluated on the spectral CT data simulated from realistic knee phantoms generated from synchrotron radiation CT. They were also compared to a Gauss-Newton algorithm that minimized a cost function with a hand-crafted regularization term. Finally, our algorithms were applied to an experimental knee data set acquired on a clinical spectral CT scanner.

We found that the proposed approaches provide virtual monoenergetic images with improved mean squared errors and structural similarities, compared to the Gauss-Newton method. Moreover, the image-domain network improved the mean squared error by a factor of two, compared to the projection-domain network. In both simulated and experimental data of osteoarthritis knees, we found that the cartilage was visible with naked eye on the virtual monoenergetic images reconstructed by our methods.

The proposed deep learning networks outperformed the Gauss-Newton algorithm in the projection domain. Among deep reconstruction strategies, we found that the image-domain direct virtual monoenergetic reconstruction performs the best. They also allow for the direct visualization of the cartilage, which is essential for the assessment of cartilage integrity.

***Index Terms*—spectral CT, VMI, Osteoarthritis**

I. Introduction

Spectral photon counting computed tomography (Spectral CT) scanner is an emerging generation of X-ray CT scanners that allows acquiring energy-resolved data thanks to photon counting detectors [1]. Contrary to conventional detectors, photon counting detectors dispatch photons into different energy bins, allowing to obtain energy-resolved measurements. Spectral CT scanners extend the capabilities of dual energy CT introduced in clinical practice [2]. Spectral CT is able to track and monitor the biodistribution of gold nanoparticles in vivo [3], to determine contrast agent concentrations in liver [4], [5], and to evaluate the risk of breast cancer [6]. Spectral CT also allows for calcium content measurement in different locations (e.g., bone, teeth, kidney stone, or coronary plaques) and the discrimination of calcium pyrophosphate hydroxyapatite crystals encountered in joint diseases [7], [8] but also identify gout signature [8]. In addition, spectral CT scanners offer high spatial resolution and reduce electronic noise [9].

In this work, we will focus on knee osteoarthritis (OA), which is a common joint disorder, suspected to affect more than 40% of the people aged of more than 60 [10]. OA is a whole joint disease affecting cartilage, bone, synovium and surrounding tissues [11]. Progressive cartilage breakdown is a major therapeutic target for putative disease-modifying OA drugs [12]. The possibility to diagnose OA at an early stage could allow the development of specific treatments to stabilize or reverse the articular cartilage degradation. While three-dimensional imaging methods are already used, as computed tomography and magnetic resonance imaging, none of them allows the simultaneous assessment of bone and cartilage in details. In vitro studies with spectral CT have shown the capability to quantify the proteoglycan content of cartilage in knee specimens injected with iodine [13], and the glycosaminoglycan content using multiple contrast agents [14]. More recently, the concentration of cationic iodinated CA4+ (proteoglygans content) and non-ionic based gadolinium based gadoteridol (water content) was estimated in small samples using a calibration-based method [15]. While the use of a contrast agent in spectral CT is an effective way to quantify its concentration, it is associated to patient safety questions. The potential of spectral CT to detect knee OA without contrast agent is a new application that has been investigated in a previous work group [16].

Spectral CT imaging requires dedicated algorithms. Most of them rely on material decomposition methods related to the linear attenuation coefficient of the object in some material basis [17]–[19]. The first decomposition methods were based on polynomial calibration using step-wedge phantoms but no

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Descartes University, Paris.

effective, for more than two materials. Alternatively,from a physical model of the source and the detector, it is possible to solve an inverse problem. The decomposition problem can be formalized as the minimization of a data fidelity term [20] and can benefit from prior knowledge about the materials [21]. While material decomposition can be achieved on projections, one-step approach allows for exploiting priors in the image domain, which is more natural but it leads to an increased computational cost that requires efficient optimization algorithms [22], [23].

More recently, methods based on deep learning (DL) have been proposed for material decomposition [24]–[28]. The idea is to learn a mapping between the spectral measurements and the material images. The mapping is achieved through the use of well-known network architectures, such as VGG 16 and U-net [24], [25] or more elaborated architecture [26], [27]. Authors in [28] show that the combination of inception modules allows to outperform previous networks. While most deep methods work in the image domain, a feed forward multilayer perceptron was proposed for decomposition in the projection domain, where convolutional architecture are not relevant, and trained using measurement in a step wedge phantom [29]. In [30], we introduced a U-net based approach for material decomposition in either the projection or image domain. This approach was successfully applied to the raw data from a phantom acquired by a spectral CT Philips prototype [31].

Spectral CT also allows to generate virtual monoenergetic (VM) images at different energies [32]. VM imaging (VMI) has the advantage to provide Hounsfield unit-like images, which can be more easily interpreted by radiologists [33]. VM images can be recovered from material decompositions, with an added-value compared to conventional CT images. Recently, DL was also explored to decompose directly VM images. In [34], the authors proposed a fully connected neural network for VMI, which was shown to improve results compared to a polynomial calibration. It has been also shown that VM images can be recovered from a single spectrum energy images by using a residual neural network [35].

The purpose of this work is to assess DL algorithms for VMI, targeting the specific application of cartilage assessment in OA. We consider four different reconstruction algorithms for the recovery of VMI in the projection and in the image domain. All of our algorithms include a variant of the U-net convolutional neural network. The proposed algorithms were trained and evaluated on the spectral CT data simulated from realistic knee phantoms generated from synchrotron radiation CT [36].

The paper is organized as follows. In section 2, we present the theoretical background of spectral CT acquisition and material decomposition. We describe four DL-based methods for the recovery of VM images either after material decomposition or directly from the raw measurements, and either in the projection or image domain. Finally, we detail our numerical experiments, based on a synthetic knee phantom and used for the training our methods, as well as our experimental measurements. In section 3, we first present our VMI results in simulations, and compare them to that obtained from traditional material decomposition [21]. Finally, we test our methods on experimental data from a clinical prototype [31].

II. Material and methods

A. Position of the problem

1) Forward model: In ideal X-ray CT, the measurement at position p under the projection angle θ corresponds to the attenuation of the X-ray beam path

$$s_p^\theta(E) = n_p^0(E) \exp\left[-\int\limits_{D_p^\theta} \mu(v,E)\mathrm{d}v\right], \tag{1}$$

where $n_p^0(E)$ is the source spectrum at position p, $\mu(v,E)$ is the linear attenuation coefficient of the object at position v and energy E, and D_p^θ is the line joining the source to the detector at position p. In this case, the log normalized measurements are line integrals of the linear attenuation, i.e.,

$$\hat{s}_p^\theta(E) := \ln\left(\frac{s_p^\theta(E)}{n_p^0(E)}\right) = -\int\limits_{D_p^\theta} \mu(v,E)\mathrm{d}v. \tag{2}$$

In practice, however, standard CT measurements are integrated over the full source spectrum, while spectral CT measurements are integrated in several energy bins. Denoting by $s_{i,p}^\theta$ the photon count in the i-th energy bin at position p on the detector for a projection view θ, equation (1) can be rewritten as

$$s_{i,p}^\theta = \int\limits_{\mathcal{E}} n_p^0(E) d_i(E) \exp\left[-\int\limits_{D_p^\theta} \mu(v,E)\mathrm{d}v\right] \mathrm{d}E, \tag{3}$$

where $d_i(E)$ is the detector response function that is the probability of a photon with an energy E to be sorted in the i-th bin. By analogy with (2), we can define log normalized spectral measurements as

$$\hat{s}_{i,p}^\theta = \ln\left(\frac{s_{i,p}^0}{s_{i,p}^\theta}\right), \quad \text{where } s_{i,p}^0 = \int\limits_{\mathcal{E}} n_p^0(E) d_i(E)\mathrm{d}E \tag{4}$$

For an ideal photon counting detector that can be modelled as $d_i(E) = \delta(E - E_i)$, the log normalization step linearizes the measurements, i.e., $\hat{s}_{i,p}^\theta = \int\limits_{D_p^\theta} \mu(v,E_i)\mathrm{d}v$. However, for an actual photon counting detector $\hat{s}_{i,p}^\theta \neq \int\limits_{D_p^\theta} \mu(v,E_i)\mathrm{d}v$ and a non linear inverse problem must be solved to recover the linear attenuation from the raw measurements.

Spectral CT classically assumes that the object is the linear combination of M materials, i.e.,

$$\mu(v,E) = \sum_{m=1}^{M} \rho_m(v)\tau_m(E), \tag{5}$$

where $\rho_m(v)$ is the mass density of the m-th material and $\tau_m(E)$ the material mass attenuation coefficient of the m-th material. By integration of Eq. (5), we obtain the projected linear attenuation as

$$\int\limits_{D_p^\theta} \mu(v,E)\mathrm{d}v = \sum_{m=1}^{M} a_{m,p}^\theta \tau_m(E), \tag{6}$$

where

$$a_{m,p}^{\theta} = \int_{D_p^{\theta}} \rho_m(v)\mathrm{d}v \tag{7}$$

represents the projected mass density. By taking into account this object model, the spectral measurements can then be expressed as [20]

$$s_{i,p}^{\theta} = \int_{\mathcal{E}} n_p^0(E) d_i(E) \ \exp\left[-\sum_{m=1}^{M} a_{m,p}^{\theta}\tau_m(E)\right]\mathrm{d}E. \tag{8}$$

Assuming discretization with P detector pixels, V object voxels, Θ projection views, and I energy bins, the forward spectral CT can be modelled as

$$\boldsymbol{s} = G(\boldsymbol{\rho}), \tag{9}$$

where $\boldsymbol{s} = (s_{1,1}^1, \ldots, s_{i,p}^{\theta}, \ldots, s_{I,P}^{\Theta})^\top$ represents the spectral photon counts and $\boldsymbol{\rho} = (\rho_{1,1}, \ldots, \rho_{m,v}, \ldots, \rho_{M,V})^\top$ represents material density maps. The full spectral CT mapping G can be seen as the composition of the X-ray transform and the spectral mixing [30]. The X-ray transform applies to each material independently

$$\boldsymbol{a}_m = X(\boldsymbol{\rho}_m), \quad 1 \leq m \leq M, \tag{10}$$

where $\boldsymbol{a}_m = (\boldsymbol{a}_{m,1}^1, \ldots, \boldsymbol{a}_{m,P}^{\Theta})^\top$ is the projected mass density of the m-th material, $\boldsymbol{\rho}_m = (\boldsymbol{\rho}_{m,1}, \ldots, \boldsymbol{\rho}_{m,V})^\top$ is the mass density of the m-th material, and X is the operator defined by (7). In simple cases as for simulations, the X-ray operator is the Radon transform. In real cases, the X-ray operator depends on the scanner geometry and the type of data acquisition. The spectral mixing applies to each angle independently as

$$\boldsymbol{s}^{\theta} = F(\boldsymbol{a}^{\theta}), \quad 1 \leq \theta \leq \Theta, \tag{11}$$

where $\boldsymbol{s}^{\theta} = (\boldsymbol{s}_{1,1}^{\theta}, \ldots, \boldsymbol{s}_{I,P}^{\theta})^\top$ represents the raw data under the θ-th projection view, $\boldsymbol{a}^{\theta} = (\boldsymbol{a}_{1,1}^{\theta}, \ldots, \boldsymbol{a}_{M,P}^{\theta})^\top$ the projected mass density under the θ-th projection view, and F is the spectral mixing operator defined by (8).

2) Material decomposition: The classical inverse problem in spectral CT aims to recover the material maps $\boldsymbol{\rho}$ from spectral measurements $\boldsymbol{s}$. The direct approach consists in inverting (9). To alleviate the computational burden, this problem can also be solved in the projection or in the image domain.

Projection domain approaches start by decomposing each projection angle independently, followed by tomographic reconstructions. Considering the Gauss-Newton approach described in [21], this two-step approach reads

$$\begin{cases} \boldsymbol{a}^{\theta} \in \arg\min \dfrac{1}{2}||\boldsymbol{s}^{\theta} - F(\boldsymbol{a}^{\theta})||_{\boldsymbol{W}^{\theta}}^2 + \alpha R(\boldsymbol{a}^{\theta}), & 1 \leq \theta \leq \Theta, \quad \text{(12a)} \\ \boldsymbol{\rho}_m = X^{-1}(\boldsymbol{a}_m), & 1 \leq m \leq M, \quad \text{(12b)} \end{cases}$$

where $\boldsymbol{W}^{\theta} = \mathrm{Diag}(\frac{1}{\boldsymbol{s}^{\theta}})$ is a quadratic weight and R is a regularization function. Alternatively, as in [30], Eq. (12a) can be replaced by a decomposition mapping (e.g., neural network) $H\colon \boldsymbol{s}^{\theta} \mapsto \boldsymbol{a}^{\theta}$.

Image domain approaches start with tomographic reconstructions, in each bin independently, followed by material decomposition. The processing pipeline can be summarized as

$$\begin{cases} \boldsymbol{r}_i = X^{-1}(\boldsymbol{a}_i), \quad 1 \leq i \leq I, & \text{(13a)} \\ \boldsymbol{\rho} = H(\boldsymbol{r}), & \text{(13b)} \end{cases}$$

where $H\colon \boldsymbol{r} \mapsto \boldsymbol{\rho}$ represents the decomposition mapping (e.g., neural network) in the image domain.

B. Virtual monoenergetic image reconstruction

Monoenergetic imaging consists in the recovery of the linear attenuation at a single specific energy E^*. This can be done by synchrotron radiation CT where a monochromatic X-ray beam is used [37]. Dual energy and spectral CT make possible *virtual* monoenergetic imaging. In this case, the virtual monoenergetic images are recovered from the spectral measurements.

In this work, we consider the recovery of the VM images either after material decomposition (see Sec II-B1) or directly from the raw measurements (see Sec II-B2) as illustrated in Figure 1. For each case, we adopt a DL approach and define four variants of the U-net architecture. Motivated by the OA application, we consider here the decomposition in two materials, bone and soft tissue, and the recovery of VMI at $E^* = 70$ keV, which has been shown to be the VMI that provides optimal SNR [38], [39].

1) VM image reconstruction via material decomposition: After material decomposition, we can obtain the VM images at energy E^* using Eq. (5), i.e., computing

$$\boldsymbol{\mu} = \sum_{m=1}^{M} \boldsymbol{\rho}_m \tau_m(E^*), \tag{14}$$

where $\boldsymbol{\rho}_m$, $1 \leq m \leq M$, represent the material density maps that are obtained by material decomposition, either in the projection or image domain.

In the projection domain, we consider the three-step approach

$$\text{MD-P method:} \begin{cases} \boldsymbol{a}^{\theta} = H_{\boldsymbol{\omega}}^{\mathrm{MD-P}}(\hat{\boldsymbol{s}}^{\theta}), \quad 1 \leq \theta \leq \Theta, & \text{(15a)} \\ \boldsymbol{\rho}_m = X^{-1}(\boldsymbol{a}_m), \quad 1 \leq m \leq M, & \text{(15b)} \\ \boldsymbol{\mu} = \sum_{m=1}^{M} \boldsymbol{\rho}_m \tau_m(E^*), & \text{(15c)} \end{cases}$$

where $H_{\boldsymbol{\omega}}^{\mathrm{MD-P}}$ is a neural network trained for material decomposition in the projection domain. We optimize the network parameters $\boldsymbol{\omega}$ to minimize the quadratic loss

$$L(\boldsymbol{\omega}) = \frac{1}{N\Theta} \sum_{n=1}^{N} \sum_{\theta=1}^{\Theta} \|H_{\boldsymbol{\omega}}^{\mathrm{MD-P}}(\hat{\boldsymbol{s}}_{(n)}^{\theta}) - \boldsymbol{a}_{(n)}^{\theta}\|^2, \tag{16}$$

where we consider $N\Theta$ measurement pairs $\{\hat{\boldsymbol{s}}_{(n)}^{\theta}, \boldsymbol{a}_{(n)}^{\theta}\}$, with N the number of volumes in the training database.

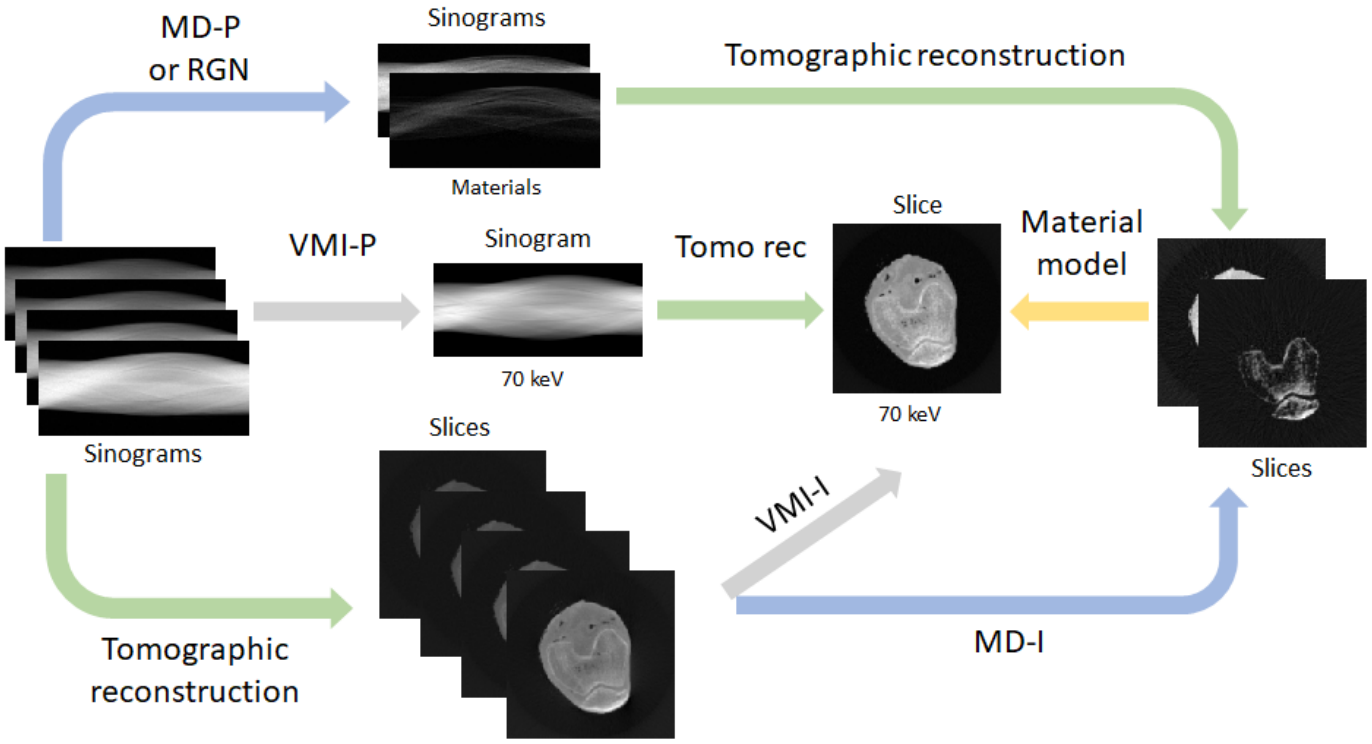


Fig. 1. Illustration of the different approaches to obtain virtual monoenergetic images from spectral measurements.

In the image domain, we consider the three-step approach

$$\text{MD-I method:} \begin{cases} \boldsymbol{r}_i = X^{-1}(\hat{\boldsymbol{s}}_i), \quad 1 \le i \le I, & \text{(17a)} \\ \boldsymbol{\rho}^z = H_{\boldsymbol{\omega}}^{\text{MD-I}}(\boldsymbol{r}^z), \quad 1 \le z \le Z, & \text{(17b)} \\ \boldsymbol{\mu} = \sum_{m=1}^{M} \boldsymbol{\rho}_m \tau_m(E^*), & \text{(17c)} \end{cases}$$

where $H_{\boldsymbol{\omega}}^{\text{MD-I}}$ is a neural network trained for material decomposition in the image domain. We optimize the network parameters $\boldsymbol{\omega}$ to minimize the quadratic loss

$$L(\boldsymbol{\omega}) = \frac{1}{NZ} \sum_{n=1}^{N} \sum_{z=1}^{Z} \| H_{\boldsymbol{\omega}}^{\text{MD-I}}(\boldsymbol{r}_{(n)}^z) - \boldsymbol{\rho}_{(n)}^z \|^2, \qquad (18)$$

where we consider a set of NZ measurement pairs $\{\boldsymbol{r}_{(n)}^z, \boldsymbol{\rho}_{(n)}^z\}$, with N the number of volumes in the training database and Z the number of slices in each volume.

2) Direct VMI decomposition: We can also recover the VM images directly, i.e., without exploiting the object decomposition model given by (5), either in the projection or image domain.

In the projection domain, we consider the following two-step approach

$$\text{VMI-P method:} \begin{cases} \boldsymbol{\nu}^\theta = H_{\boldsymbol{\omega}}^{\text{VMI-P}}(\hat{\boldsymbol{s}}^\theta), \quad 1 \le \theta \le \Theta, & \text{(19a)} \\ \boldsymbol{\mu} = X^{-1}(\boldsymbol{\nu}), & \text{(19b)} \end{cases}$$

where $H_{\boldsymbol{\omega}}^{\text{VMI-P}}$ is a neural network trained for direct VM decomposition in the projection domain. We optimize the network parameters $\boldsymbol{\omega}$ to minimize the quadratic loss

$$L(\boldsymbol{\omega}) = \frac{1}{N\Theta} \sum_{n=1}^{N} \sum_{\theta=1}^{\Theta} \| H_{\boldsymbol{\omega}}^{\text{VMI-P}}(\hat{\boldsymbol{s}}_{(n)}^\theta) - \boldsymbol{\nu}_{(n)}^\theta \|^2, \qquad (20)$$

where we consider $N\Theta$ measurement pairs $\{\hat{\boldsymbol{s}}_{(n)}^\theta, \boldsymbol{\nu}_{(n)}^\theta\}$, with N the number of volumes in the training database.

In the image domain, we consider the two-step approach

$$\text{VMI-I method:} \begin{cases} \boldsymbol{r}_i = X^{-1}(\hat{\boldsymbol{s}}_i), \quad 1 \le i \le I, & \text{(21a)} \\ \boldsymbol{\mu}^z = H_{\boldsymbol{\omega}}^{\text{VMI-I}}(\boldsymbol{r}^z), \quad 1 \le z \le Z & \text{(21b)} \end{cases}$$

where $H_{\boldsymbol{\omega}}^{\text{VMI-I}}$ is a neural network trained for direct VM decomposition in the image domain. We optimize the network parameters $\boldsymbol{\omega}$ to minimize the quadratic loss

$$L(\boldsymbol{\omega}) = \frac{1}{NZ} \sum_{n=1}^{N} \sum_{z=1}^{Z} \| H_{\boldsymbol{\omega}}^{\text{VMI-I}}(\boldsymbol{r}_{(n)}^z) - \boldsymbol{\rho}_{(n)}^z \|^2, \qquad (22)$$

where we consider a set of NZ measurement pairs $\{\boldsymbol{r}_{(n)}^z, \boldsymbol{\rho}_{(n)}^z\}$, with N the number of volumes in the training database and Z the number of slices in each volume. The dimension of $\boldsymbol{r}^z$ is $V_x \times V_y$, such that the number of voxels is $V = X \times Y \times Z$.

C. Simulated data

In this section, we describe the simulated datasets generated from realistic knee phantoms and used to train and evaluate the networks.

1) Knee phantoms: Excised human knees were collected at the Anatomy Institute of Paris Descartes. The study was approved by the Ethics Committee of Descartes University, Paris.Eight knee samples were imaged by Synchrotron Radiation (SR) CT at beamline ID17 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. The samples were scanned at an energy of 55 keV with a cubic voxel size of 45 μm. Due to the limited field of view in height, the samples were scanned in several stacks. The SR CT images were reconstructed using the filtered backprojection algorithm and the different stacks were merged together to get the full knee volumes.

These volumes were then used to generate realistic numerical human knee phantoms. First, the segmentation of bone and soft tissue was performed in all knee specimens, using the K-means algorithm with 3 seeds [40]. Each voxel was then considered to contain only one material, either bone or soft tissue. By combining this information with the gray level volumes, we obtained the two material maps $\boldsymbol{\rho}_m$, respectively for bone and soft tissue. In order to work with data sets closer to the real experimental data, before segmentation, the volumes were undersampled after Gaussian filtering to reach a cubic voxel size of 0.25 mm. The bone and soft tissue maps have a final size of $X \times Y \times Z = 635 \times 635 \times 202$ pixels, where X, Y, and Z represents the number of voxels in the lateral and longitudinal directions, respectively.

2) Simulated spectral CT knee data: The spectral CT data were simulated from the two material maps using the forward model described in Section II.A.1. The projections were generated from the 2D Radon transform, slice by slice. We generated $\Theta = 720$ projections evenly distributed over $180°$. Then, we applied the spectral mixing transform defined in Equation (8), where we consider the source spectra and detector response function for $I = 5$ energy bins given in [31]. The irradiation of the detector decreases radially, with more photons hitting the center of the detector. The spectral mixing model correspond to typical exams performed using the spectral CT scanner described in Section II-D. However, the model does not include pulse pile-up and charge sharing effects. After applying the forward model to the eight knee volumes, we obtained eight energy-resolved projection data sets with a size of $\Theta \times P_x \times Z \times I = 720 \times 643 \times 202 \times 5$.

D. Experimental spectral CT knee data

A raw measurement dataset was acquired for one knee using the clinical prototype of spectral CT scanner available at the CERMEP (Lyon, France) [31]. The acquisition was performed with a tube voltage of 120 kVp and a tube current of 226 mA. The photon counting cells of the detector are 2-mm-thick cadnium zinc telluride. Each cell sorts photons in five different energy bins (30–51, 51–62, 62–72, 72–81 and 81–120 keV). The acquisition geometry is helical with scans over $360°$ with $\Lambda = 2400$ projections per rotation. The detector pixel size is 250 μm. The scan field of view is 500 mm in the lateral direction with a longitudinal coverage of 80 mm. Data were binned to obtain a dataset of total size of $P_\rho \times P_z \times \Theta \times I = 924 \times 8 \times 24912 \times 5$, where P_ρ and P_z represent the lateral and longitudinal dimension of the detector, respectively, such that the number of detector pixels is $P = P_\rho \times P_z$ and Θ corresponds to the number of projections. All tomographic reconstructions were performed using the Feldkamp algorithm [41] implemented within the RTK package [42]. We reconstructed a region of interest of 180×100 mm with a a voxel size of $0.25 \times 0.25 \times 1$ mm.

E. Training and evaluation

The loss functions of the four networks, as given by equations 16, 18, 20, and 22, were minimized using the ADAM optimizer [43] in Pytorch [44]. Our different neural networks are all based on the U-net architecture [45] with 11 convolutional layers and two undersampling levels. The architecture of our networks only differ in their first and last layers. For both material decomposition networks $H_{\boldsymbol{\omega}}^{\mathrm{MD-P}}$ and $H_{\boldsymbol{\omega}}^{\text{MD-I}}$, we consider the U-net architecture depicted in Fig. 2(a). For both direct neural networks $H_{\boldsymbol{\omega}}^{\text{VMI-P}}$ and $H_{\boldsymbol{\omega}}^{\text{VMI-I}}$, we consider the U-net architecture as represented in Fig. 2(b). We adopt patchwise training where each input/output image pair is randomly cropped to obtain a 64×64 patch. We use a learning scheduler that multiply the learning rate by 0.3 every 30 epochs and an early stopping criterion to halt training at 100 epochs. We considered different initial learning rate for the different networks. We used 10^{-3} for MD-P, 10^{-2} for MD-I, 10^{-1} for VMI-P and 10^{-1} for VMI-I, which are the best learning rate considering validation data.

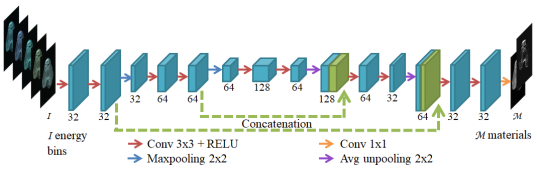
All networks were trained using a NVidia Tesla V100. In the projection domain, the training phase took about 6 hours considering $\Theta = 2,880$ projections and $P = 129,886$ pixels. The test takes about 3 minutes. In the image domain, the training took around 4 hours for $Z = 808$ slices and $V = 403'225$ pixels. The test took only 1 minute on GPU. The Gauss-Newton method took 10 s by projection, i.e., 2 hour for the test. In the image domain, the tomographic reconstruction is performed before the training phase so it is applied once, reducing the computational burden. Tomographic reconstructions took 46 seconds by bin, i.e., 3 minutes 50 seconds for each simulate volume. For experimental data, the tomographic reconstruction took 22 minutes by bin, i.e., 1 hour 50 minutes, as the algorithm exploits more projections.

In simulations, we evaluated all approaches in the image domain using the mean square error (MSE) and the structure similarity index (SSIM). Each of the four different networks was tested with unknown knees. For comparison, we also consider the Gauss-Newton material decomposition approach given by Equation 12. For experimental data where no ground-truth is available, we computed the standard deviation in homogeneous regions of the soft tissue map. We also assessed the images qualitatively considering the presence of artifacts or blurring of details. We also assessed our ability to identify the presence of the cartilage with the naked eye.

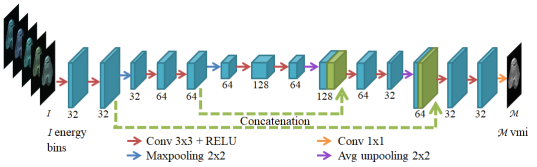
III. Results

A. Simulated data

1) Material decomposition: The two networks MDP and MDI were trained for different initial learning rates. We



(a) MD-P and MD-I



(b) VMI-P and VMI-I

Fig. 2. Architecture of our U-net variants for (a) material decomposition and (b) direct VMI decomposition. Note that only the last layer of MD and VMI networks differs. The illustration corresponds to image domain decomposition but inputs can be replaced by sinograms, leading to projection domain networks.

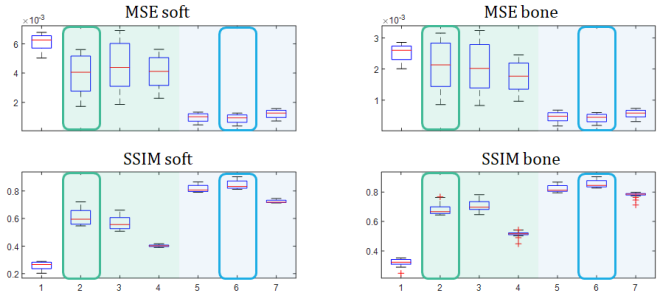


Fig. 3. Quantitative boxplots for the MSE and SSIM quality metrics on the soft tissue and bone maps, for the two material decomposition networks with the three learning rates, compared to RGN. 1 : RGN; 2, 3 and 4: MDP (green) for the learning rates 10^{-3}, 10^{-2} and 10^{-1}; 5, 6 and 7: MDI (blue) for the learning rates 10^{-3}, 10^{-2} and 10^{-1}.

calculated the MSE and the SSIM of bone and soft tissue for the initial learning rates of 10^{-3}, 10^{-2} and 10^{-1}. Figure 3 shows MSE and SSIM for bone and soft tissue, as box plots for the different methods, while the mean and standard deviation of the quality metrics are reported on Table I.

We found that the initial learning rate at 10^{-3} for training MDP (green domain on Figure 3) gave the best compromise between MSE and SSIM considering both the soft tissue and the bone. The initial learning rate of 10^{-2} for the material decomposition method with MDI (blue domain on Figure 3) gives the best MSE and SSIM for both soft tissue and bone with values of $9{\cdot}10^{-4}$ and $4{\cdot}10^{-4}$ for the MSE and 0.8450 and 0.8552 for the SSIM . Moreover we can see that both MDP and MDI give better results than our reference method RGN. MDI is better than MDP considering these metrics. Table I, shows that adopting RGN as reference, MDI improves the MSE by a factor less than 2 in the projection domain and by 5 in the image domain. It also improve the SSIM by around 2.5 with a bigger improvement in the image domain than in the projection domain, compared to RGN.

The Figure 4 shows for a given slice, the reconstructed soft tissue and bone material maps obtained with MDP, MDI and RGN as well as the error images with respect to the ground truth. We can see again that MDI is better than MDP. Visually, MDP tends to blur the image where MDI gives a sharper image with less noise than the state-of-the-art RGN method. Both networks deals better with noise.

	Soft tissue		Bone	
	MSE	SSIM	MSE	SSIM
RGN	0.0061± 0.0005	0.2589± 0.0249	0.0025± 0.0003	0.3252± 0.0171
MDP $\lambda = 10^{-3}$	**0.0039**± 0.0013	**0.6100**± 0.0562	0.0021± 0.0007	0.6790± 0.0323
MDP $\lambda = 10^{-2}$	0.0045± 0.0016	0.5682± 0.0465	0.0021± 0.0008	**0.7091**± 0.0343
MDP $\lambda = 10^{-1}$	0.0041± 0.0010	0.4024± 0.0063	**0.0018**± 0.0005	0.5175± 0.0103
MDI $\lambda = 10^{-3}$	0.0010± 0.0003	0.8172± 0.0248	0.0005± 0.0001	0.8223± 0.0234
MDI $\lambda = 10^{-2}$	**0.0009**± 0.0003	**0.8450**± 0.0297	**0.0004**± 0.0001	**0.8552**± 0.0248
MDI $\lambda = 10^{-1}$	0.0012± 0.0003	0.7237± 0.0099	0.0006± 0.0001	0.7844± 0.0086

TABLE I

QUANTITATIVE MEAN AND STANDARD DEVIATION FOR THE MSE AND SSIM FOR THE DIFFERENT MATERIAL DECOMPOSITION MEHODS ON THE TEST DATA SETS USING DIFFERENT LEARNING RATES.

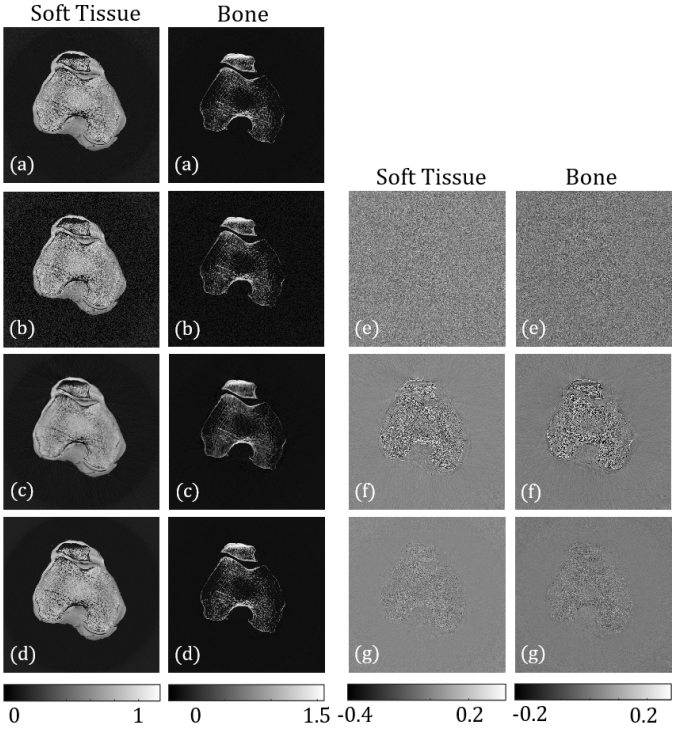


Fig. 4. Decomposed soft tissue and bone material maps for the different methods and error maps computed as the differences with respect to the ground truth (a) Materials of the original phantom. (b), (e): Results obtained from RGN and error maps. (c), (f): Results of MDP and error maps. (d), (g): Results of MDI and error maps.

2) VMI decomposition: The direct VMI neural networks VMP and VMI were also trained for the different initial learning rates of 10^{-3}, 10^{-2} and 10^{-1}. Figure 5 synthesizes the boxplots for the MSE and the SSIM for the different learning rates for these two networks, as well as those the VMI reconstructed from the material decomposition networks

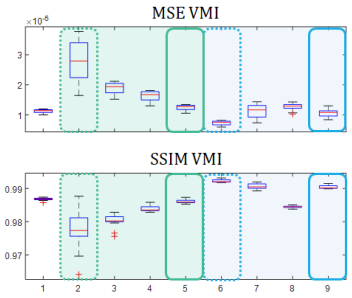


Fig. 5. Quantitative boxplots for the MSE and SSIM quality metrics for the reconstructed VMI at 70 keV for the different methods and different learning rates. 1 : RGN ; 2 : VMI from MDP for the selected learning rate 10^{-3} ; 3, 4 and 5 : VMP for the learning rates 10^{-3}, 10^{-2} and 10^{-1} ; 6 : VMI from MDI for the selected learning rate 10^{-2} ; 7, 8 and 9 : VMI for the learning rates 10^{-3}, 10^{-2} and 10^{-1}.

	MSE	SSIM
RGN	**1.12** ± 0.06	0.9869 ± 0.0003
MDP $\lambda = 10^{-3}$	2.80±0.61	0.9783 ± 0.0035
VMP $\lambda = 10^{-3}$	1.89±0.19	0.9807 ± 0.0011
VMP $\lambda = 10^{-2}$	1.63±0.16	0.9839 ± 0.0008
VMP $\lambda = 10^{-1}$	**1.25** ± 0.08	**0.9861** ± 0.0005
MDI $\lambda = 10^{-2}$	**0.73** ± 0.07	**0.9923** ± 0.0004
VMI $\lambda = 10^{-3}$	1.12±0.20	**0.9907** ± 0.0008
VMI $\lambda = 10^{-2}$	1.28±0.08	0.9845 ± 0.0003
VMI $\lambda = 10^{-1}$	**1.05** ± 0.10	0.9904 ± 0.0005

TABLE II
QUANTITATIVE RESULTS FOR VMI DECOMPOSITION ON THE TESTSET USING DIFFERENT LEARNING RATES.

MDP and MDI from the last section. The mean and standard deviation of the metrics are reported in Table II.

We selected an initial learning rate of 10^{-1} for the VMI decomposition method with VMP (green domain in Figure 5) and VMI (blue domain in Figure5) that gives the best MSE and SSIM. We can also see that the VMI reconstructed from the decomposed materials with MDI gives the best results for all metrics. We can sort them as VMI from decomposed materials with MDI better than VMI with VMI better than VMI with VMP better than MDP. Table II shows that the decomposition in the image domain provides improved metrics compared to decomposition obtained from material decomposition or directly, is improved compared to RGN and the results from the projection domain.

The Figure 6 shows the VMI from decomposed materials with MDP and MDI and directly with VMP and VMI. We see the same order of quality as with the box plots. The VMI from decomposed materials with MDP exhibits errors at the border of the knee that appears as artifacts in the border of materials having strong attenuation. For VMI, we can also see errors at the top of the patella which also has the largest attenuation.

Figure 7 shows zooms on a cartilage region of the posterior part of the tibia. An interesting point is that the cartilage

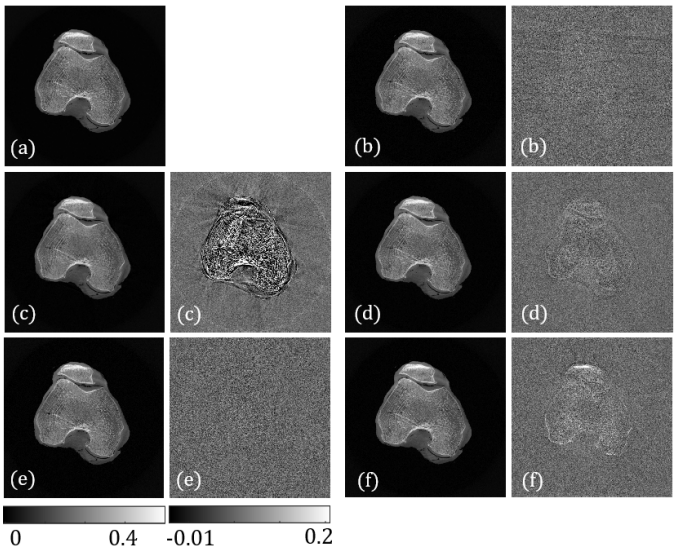


Fig. 6. Reconstructed VMIs at 70 KeV from the different methods and their corresponding error maps with respect to ground truth. (a) VMI of the original phantom (ground truth). (b) RGN. (c) MDP. (d) MDI. (e) VMP. (f) VMI.

is visible with naked eye with similar image quality for all methods.

B. First results on experimental data

Finally, the four networks, trained on the simulated phantoms and with the optimal learning rates selected previously, were applied to the experimental data set acquired on the Philips spectral CT prototype. The results are shown in Figure 8. The VMI could be reconstructed with all methods with a coherent gray scale. A zoom in the femoro-patella region shows that we can see a bright part just beneath the bone corresponding to cartilage. We can visualize the cartilage with naked eye in experimental data, which is one of the advantages of using spectral CT compared to standard CT. We also computed the standard deviation in an homogeneous area in the soft tissue. The values show us that the decomposition in the image domain gives the noisiest results with a std in a homogeneous region of 0.0164. The VMI obtained from material decomposition in the projection domain gives us the lowest level of noise with a standard deviation of 0.0105. The VMI decomposition in the projection domain and the VMI from material decomposition with RGN gives similar results in term of noise.

IV. DISCUSSION & CONCLUSION

In this work, we focused on a novel DL approach to obtain VMI from spectral CT data. VMI have a great potential for application as OA by providing HU-like CT images with increased resolution, contrast and SNR. In addition, we found in a preliminary work that experimental VMI at 70 keV acquired from a spectral CT prototype were particularly suited for the analysis of excised knees in OA and normal subject [38].

To visualize on a same image bone and cartilage, our networks were trained on simulations from a realistic phantom

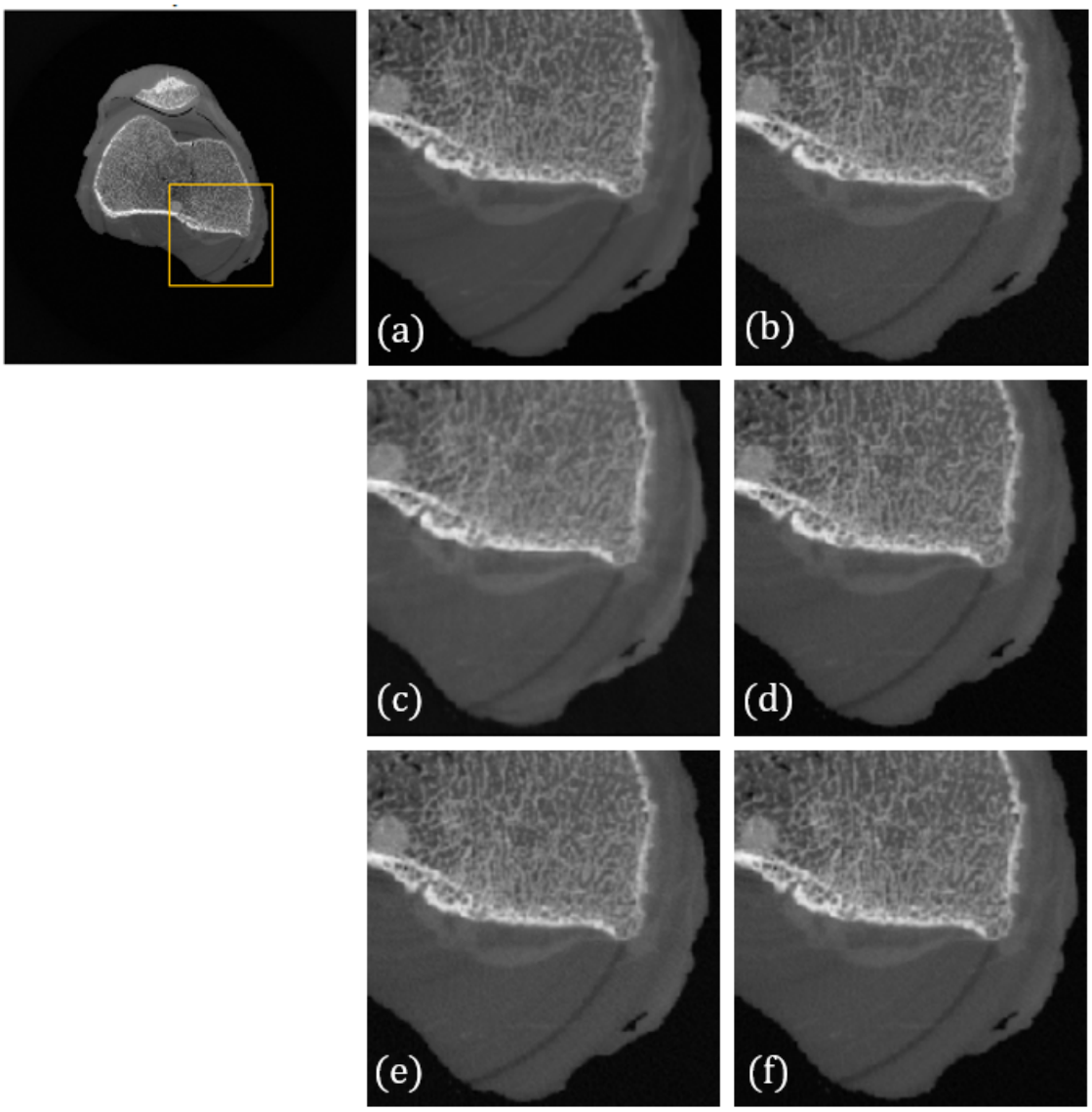


Fig. 7. Zoom on the condyle cartilage (yellow ROI) in a VMI slice at 70 KeV obtained from the different methods. (a) Original phantom. (b) RGN. (c) VMDP. (d) MDI. (e) VMP. (f) VMI.

database based on the 8 knee specimens imaged by synchrotron CT downscaled to have the same spatial resolution as the spectral CT prototype. We chose to learn using patches to boost the number of data loaded by the networks. We assessed the different methods on test data and also with an experimental data set from this prototype. We evaluate our results with respect to the ground truth visually with both the MSE to assess the gray level content and SSIM to assess the preservation of the structures.

Most DL methods dedicated to spectral CT have been devoted to material decomposition.In [30] we found that U-net based networks for material decomposition gave better results than the state-of-the-art optimization method RGN. Moreover, neural network methods allow to bypass the knowledge of the forward model necessary with an optimization method. The results of the networks trained on the KiTS19 database were better when working in the image domain rather than in the projection domain. The present study dedicated to knee data confirms that the material decomposition networks achieve better results than the RGN method and that the results are better when materials are decomposed in the image domain. Our MDP leads to a blurred decomposition but with less noise in the background than RGN. This problem does not appear with MDI which provides a sharp decomposition while keeping a small level of noise.

The VMI at a given energy are generally generated by combining the material maps obtained either by a standard method or a DL network. Here we designed DL networks to recover the VMI, either by recombining the material maps or recovering directly the VMI at a given energy. In a previous work [34], material decomposition neural networks have been extended to decompose VMI at a random energy present in the training set. In our work, we found that the direct VMI decomposition with U-net in the projection domain provided better results than recovering the VMI from the decomposed materials. The direct VMI decomposition in the image domain

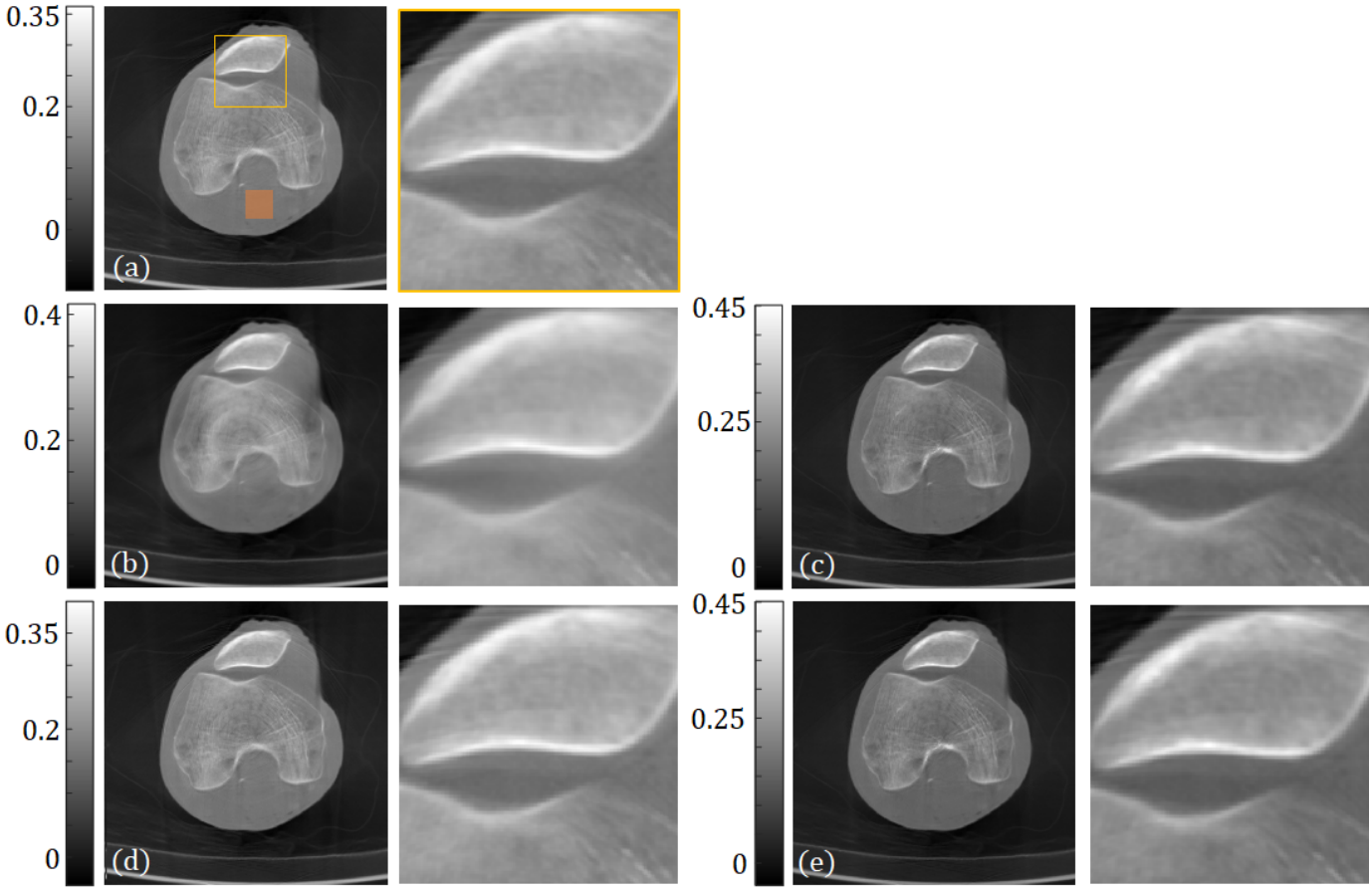


Fig. 8. VMI reconstructed from the experimental spectral data and zoom on the patella cartilage (yellow ROI) from the different methods. The standard deviation in the orange square is reported. (a) RGN, std = 0.0149. (b) MDP, std = 0.0105. (c) MDI, std = 0.0164. (d) VMP, std = 0.0143. (e) VMI, std = 0.0164.

is always better than in projection domain, but, in this case, the VMI from decomposed material was found better than the direct VMI decomposition.

Our networks trained on simulated phantoms tailed to the OA application performed reasonably well on new test data as well as on an experimental knee data set. Indeed, we were able to reconstruct good quality VMI in which the cartilage was visible with naked eye. In the experimental data set, we can even see some inhomogeneities in cartilage probably related to OA.

Considering computation time, an important aspect is that DL networks are 30 to 40 times faster than a model-based state-of-the-art material decomposition method. In addition, working in the image domain implies faster training and requires to perform tomographic reconstruction only once, which reduce burden for training and hyperparameters tuning. Despite of this, at test time, the tomographic reconstruction has to be performed for each energy channel and so it is 5 times longer compared to the projection domain. However, we save time in the image domain compared to the projection domain from 5 test, which is the foreseen case. This time saving is important for application to experimental data.

Direct VMI decomposition networks present advantages in terms of computing time and memory since they do not require an explicit material decomposition step. Then, even if the recovery of VMI from material maps is slightly better than the direct VMI decomposition in the image domain, direct VMI decomposition may be valuable. Note that even if we restricted ourselves to the reconstruction of VMI at 70keV, any other energy in the energy range of the spectral CT device could be obtained after retraining the DL VMI networks.

Nevertheless, this study is subject to few limitations. Our training strategy, although giving promising results, it would be necessary to enrich the phantoms database with additional ones to create more variability in the training set and then increasing the robustness of the network. Moreover, the trained networks should be tested on additional experimental data sets, which should be possible in future with the increasing availability of the spectral CT devices. This could allow to apply transfer learning to fine tune on experimental data sets.

In conclusion, DL networks are attractive for the recovery of VMI from spectral CT data. The technique appears promising for the investigation of bone and cartilage simultaneously for the early detection of OA which could help for a better characterization of cartilage with a good resolution. Additional post-processing could be applied to the VMI to further exploit these images and compute biomarkers on the cartilage and bone mesh in order to have new imaging methods to characterize bone and cartilage in OA for better phenotyping. Moreover, material decomposition images could also be interesting in

future to characterize pathological calcification of joints.

ACKNOWLEDGEMENTS

This work was supported by the French ANR project SALTO under Grant ANR-17-CE19-0011-01 and performed in the framework of the LabEx PRIMES of the University de Lyon under Grant ANR-11-LABX-0063. It is also part of France Life Imaging from the French Investissements d'Avenir under Grant ANR-11-INBS-0006.

The authors acknoledge the European Synchrotron Radiation Facility (ESRF) for proving beamtime for the experiment MD-1142 and thanks the ID17 staff for kind assistance and support for SR-CT data acquisition.

All authors declare that they have no known conflicts of interest in terms of competing financial interests or personal relationships that could have an influence or are relevant to the work reported in this paper.

REFERENCES

[1] 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.

[2] C. H. McCollough, S. Leng, L. Yu, and J. G. Fletcher, "Dual- and Multi-Energy CT: Principles, Technical Approaches, and Clinical Applications," *Radiology*, vol. 276, no. 3, pp. 637–653, Sep. 2015, publisher: Radiological Society of North America.

[3] S. Si-Mohamed, D. P. Cormode, D. Bar-Ness, M. Sigovan, P. C. Naha, J.-B. Langlois, L. Chalabreysse, P. Coulon, I. Blevis, E. Roessl, K. Erhard, L. Boussel, and P. Douek, "Evaluation of spectral photon counting computed tomography K-edge imaging for determination of gold nanoparticle biodistribution in vivo," *Nanoscale*, vol. 9, no. 46, pp. 18 246–18 257, Nov. 2017, publisher: The Royal Society of Chemistry.

[4] P. Baturin, Y. Alivov, and S. Molloi, "Spectral CT imaging of vulnerable plaque with two independent biomarkers," *Physics in Medicine and Biology*, vol. 57, no. 13, pp. 4117–4138, Jun. 2012, publisher: IOP Publishing.

[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] H. Ding, B. Zhao, P. Baturin, F. Behroozi, and S. Molloi, "Breast tissue decomposition with spectral distortion correction: A postmortem study," *Medical Physics*, vol. 41, no. 10, p. 101901, 2014.

[7] T. E. Kirkbride, A. Y. Raja, K. Müller, C. J. Bateman, F. Becce, and N. G. Anderson, "Discrimination Between Calcium Hydroxyapatite and Calcium Oxalate Using Multienergy Spectral Photon-Counting CT," *American Journal of Roentgenology*, vol. 209, no. 5, pp. 1088–1092, Nov. 2017.

[8] L. K. Stamp, N. G. Anderson, F. Becce, M. Rajeswari, M. Polson, O. Guyen, A. Viry, C. Choi, T. E. Kirkbride, and A. Y. Raja, "Clinical Utility of Multi-Energy Spectral Photon-Counting Computed Tomography in Crystal Arthritis," *Arthritis & Rheumatology*, vol. 71, no. 7, pp. 1158–1162, 2019.

[9] S. S. Hsieh, S. Leng, K. Rajendran, S. Tao, and C. H. McCollough, "Photon Counting CT: Clinical Applications and Future Developments," *IEEE transactions on radiation and plasma medical sciences*, vol. 5, no. 4, pp. 441–452, Jul. 2021.

[10] E. H. Oei, J. Tiel, W. Robinson, and G. Gold, "Quantitative Radiologic Imaging Techniques for Articular Cartilage Composition: Toward Early Diagnosis and Development of Disease-Modifying Therapeutics for Osteoarthritis," *Arthritis Care & Research*, vol. 66, no. 8, pp. 1129–1141, 2014.

[11] J. Kellgren and J. Lawrence, "Radiological assessment of osteo-arthrosis," *Annals of the Rheumatic Diseases*, vol. 16, no. 4, pp. 494–502, Dec. 1957.

[12] H. Braun and G. Gold, "Diagnosis of Osteoarthritis: Imaging," *Bone*, vol. 51, no. 2, pp. 278–288, Aug. 2012.

[13] K. Rajendran, C. Löbker, B. Schon, C. Bateman, R. Aamir, N. J. A. de Ruiter, A. I. 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.

[14] K. Baer, S. Kieser, B. Schon, K. Rajendran, T. ten Harkel, M. Ramyar, C. Löbker, C. Bateman, A. Butler, A. Raja, G. Hooper, N. Anderson, and T. Woodfield, "Spectral CT imaging of human osteoarthritic cartilage via quantitative assessment of glycosaminoglycan content using multiple contrast agents," *APL Bioengineering*, vol. 5, no. 2, p. 026101, 2021.

[15] P. Paakkari, S. I. Inkinen, M. K. M. Honkanen, M. Prakash, R. Shaikh, M. T. Nieminen, M. W. Grinstaff, J. T. A. Mäkelä, J. Töyräs, and J. T. J. Honkanen, "Quantitative dual contrast photon-counting computed tomography for assessment of articular cartilage health," *Scientific Reports*, vol. 11, no. 1, p. 5556, Mar. 2021.

[16] C. Chappard, J. Abascal, S. Bussod, S. Uk, S. Si-Mohamed, P. Douek, and F. Peyrin, "Feasibility of spectral computed tomography to assess knee cartilage," *Osteoarthritis and Cartilage*, vol. 28, pp. S277–S278, Apr. 2020.

[17] R. Alvarez and A. Macovski, "Energy-selective reconstructions in X-ray computerised tomography," *Physics in Medicine & Biology*, vol. 21, no. 5, p. 733, 1976.

[18] L. A. Lehmann, R. E. Alvarez, A. Macovski, W. R. Brody, N. J. Pelc, S. J. Riederer, and A. L. Hall, "Generalized image combinations in dual KVP digital radiography," *Medical Physics*, vol. 8, no. 5, pp. 659–667, Oct. 1981.

[19] D. Cormode, E. Roessl, A. Thran, T. Skajaa, R. Gordon, J.-P. 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.

[20] J. Schlomka, E. Roessl, R. Dorscheid, S. Dill, G. Martens, T. Istel, C. Bäumer, C. Herrmann, R. Steadman, 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.

[21] N. Ducros, J. 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.

[22] S. Tairi, S. Anthoine, y. Boursier, M. Dupont, and C. Morel, "ProMeSCT: A Proximal Metric Algorithm for Spectral CT," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, pp. 548–558, 2021, publisher: IEEE.

[23] C. Mory, B. Sixou, S. Si-Mohamed, L. Boussel, and S. Rit, "Comparison of five one-step reconstruction algorithms for spectral CT," *Physics in Medicine and Biology*, vol. 63, no. 23, p. 235001, Nov. 2018.

[24] Z. Chen and L. Li, "Preliminary Research on Multi-Material Decomposition of Spectral CT Using Deep Learning," in *The 14th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2017, pp. 52–526.

[25] D. Clark, M. Holbrook, and C. Badea, "Multi-energy CT decomposition using convolutional neural networks," in *Medical Imaging 2018: Physics of Medical Imaging*, vol. 10573. SPIE, 2018, p. 59.

[26] W. Zhang, H. Zhang, L. Wang, X. Wang, X. Hu, A. Cai, L. Li, T. Niu, and B. Yan, "Image domain dual material decomposition for dual-energy CT using butterfly network," *Medical Physics*, vol. 46, no. 5, pp. 2037–2051, 2019.

[27] Z. Li, I. Chun, and Y. Long, "Image-Domain Material Decomposition Using an Iterative Neural Network for Dual-Energy CT," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, Apr. 2020, pp. 651–655, iSSN: 1945-8452.

[28] H. Gong, S. Tao, K. Rajendran, W. Zhou, C. McCollough, and S. Leng, "Deep-learning-based direct inversion for material decomposition," *Medical physics*, vol. 47, no. 12, pp. 6294–6309, Dec. 2020.

[29] K. Zimmerman, G. Sharma, A. Parchur, A. Joshi, and T. Schmidt, "Experimental investigation of neural network estimator and transfer learning techniques for K-edge spectral CT imaging," *Medical Physics*, vol. 47, no. 2, pp. 541–551, Feb. 2020.

[30] J. Abascal, N. Ducros, V. Pronina, S. Rit, P.-A. Rodesch, T. Broussaud, S. Bussod, P. Douek, A. Hauptmann, S. Arridge, and F. Peyrin, "Material Decomposition in Spectral CT Using Deep Learning: A Sim2Real Transfer Approach," *IEEE Access*, vol. 9, pp. 25 632–25 647, 2021.

[31] S. Si-Mohamed, D. Bar-Ness, M. Sigovan, D. Cormode, P. Coulon, E. Coche, A. Vlassenbroek, G. Normand, L. Boussel, and P. Douek, "Review of an initial experience with an experimental spectral photon-counting computed tomography system," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 873, pp. 27–35, Nov. 2017.

[32] L. Yu, J. A. Christner, S. Leng, J. Wang, J. G. Fletcher, and C. H. McCollough, "Virtual monochromatic imaging in dual-source dual-energy CT: Radiation dose and image quality," *Medical Physics*, vol. 38, no. 12, pp. 6371–6379, 2011.

[33] M. H. Albrecht, T. J. Vogl, S. S. Martin, J. W. Nance, T. M. Duguay, J. L. Wichmann, C. N. De Cecco, A. Varga-Szemes, M. van Assen, C. Tesche, and U. J. Schoepf, "Review of Clinical Applications for Virtual Monoenergetic Dual-Energy CT," *Radiology*, vol. 293, no. 2, pp. 260–271, Nov. 2019, publisher: Radiological Society of North America.

[34] C. Feng, K. Kang, and Y. Xing, "Fully connected neural network for virtual monochromatic imaging in spectral computed tomography," *Journal of Medical Imaging (Bellingham, Wash.)*, vol. 6, no. 1, p. 011006, Jan. 2019.

[35] W. Cong, Y. Xi, P. Fitzgerald, B. De Man, and G. Wang, "Virtual Monoenergetic CT Imaging via Deep Learning," *Patterns (New York, N.Y.)*, vol. 1, no. 8, p. 100128, Nov. 2020.

[36] S. Bussod, J. Abascal, S. Arridge, A. Hauptmann, C. Chappard, N. Ducros, and F. Peyrin, "Convolutional Neural Network for Material Decomposition in Spectral CT Scans," in *2020 28th European Signal Processing Conference (EUSIPCO)*, Jan. 2021, pp. 1259–1263.

[37] 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, 1999.

[38] C. Chappard, J. Abascal, C. Olivier, S. Si-Mohamed, L. Boussel, J. B. Piala, P. Douek, and F. Peyrin, "Virtual monoenergetic images from photon-counting spectral computed tomography to assess knee osteoarthritis," *European Radiology Experimental*, vol. 6, no. 1, p. 10, Feb. 2022. [Online]. Available: https://doi.org/10.1186/s41747-021-00261-x

[39] C. Garcelon, J. Abascal, C. Olivier, S. Uk, S. Si-Mohamed, H.-K. Ea, P. Douek, F. Peyrin, and C. Chappard, "Quantification of cartilage and subchondral bone cysts on knee specimens based on a spectral photon-counting computed tomography," *Scientific Reports*, vol. 13, no. 1, p. 11080, Jul. 2023, number: 1 Publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/s41598-023-38238-y

[40] S. Lloyd, "Least Squares Quantization in PCM," *IEEE Transaction on Information Theory*, pp. 129–137, 1982.

[41] L. Feldkamp, L. Davis, and J. Kress, "Practical cone-beam algorithm," *JOSA A*, vol. 1, no. 6, pp. 612–619, Jun. 1984.

[42] S. Rit, M. Oliva, S. Brousmiche, R. Labarbe, D. Sarrut, and G. Sharp, "The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)," *Journal of Physics: Conference Series*, vol. 489, p. 012079, Mar. 2014.

[43] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization." ICLR, San Diego, 2015.

[44] 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.

[45] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation." Springer, C., 2015, pp. 234–241.