

# Décomposition en matériaux de base par apprentissage profond pour la tomographie spectrale

Kannara MOM<sup>1</sup> Jérôme LESAIN<sup>2</sup> Nicolas DUCROS<sup>1,3</sup> Bruno SIXOU<sup>1</sup> Max LANGER<sup>4</sup>

<sup>1</sup>Univ Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Inserm, CREATIS UMR 5220, U1206, F-69621 Villeurbanne, France

<sup>2</sup>Univ. Grenoble Alpes, Inserm UA7 Strobe, Grenoble, France.

<sup>3</sup>Institut Universitaire de France (IUF), France.

<sup>4</sup>Univ. Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, TIMC, F-38000 Grenoble, France

**Résumé** – Dans ce travail, nous nous intéressons au problème de décomposition en matériaux de base dans l’espace des projections en tomographie spectrale. Nous proposons une nouvelle méthode d’apprentissage profond pour résoudre ce problème inverse non linéaire. La méthode proposée s’appuie sur le déroulement d’une descente de gradient. Nous montrons que l’approche développée permet d’obtenir de meilleurs résultats de décomposition que la version itérative basée sur une approche variationnelle et qu’un réseau de neurones U-Net qui n’utilise pas la physique du problème direct. Le temps de calcul nécessaire est aussi réduit.

**Abstract** – In this work, we investigate the material decomposition in the projection space for spectral computed tomography. We propose a new deep learning method to solve this nonlinear inverse problem. The method is based on the unrolling of a gradient descent. We show that the proposed approach improves the material decomposition compared to both the classical iterative variational method and a U-Net architecture which does not take into account the physics of the forward problem. The calculation time is also reduced.

## 1 Introduction

En tomographie par rayons X classique, l’objet d’intérêt est modélisé par son coefficient d’atténuation linéaire. Celui-ci est reconstruit à partir d’une séquence de radiographies acquises sous plusieurs angles du vue. En tomographie spectrale, les mesures sont résolues en énergie, permettant ainsi de décomposer le coefficient d’atténuation linéaire sur une base de matériaux [4] (typiquement eau/os en imagerie médicale). Les deux étapes du problème – reconstruction tomographique et décomposition en matériaux de base (DMB) – peuvent être réalisées conjointement ou consécutivement. Les méthodes conjointes, appelées *one-step* dans la littérature, inversent directement le modèle global de formation des mesures. Bien que théoriquement optimales en termes de compromis bruit/résolution spatiale des reconstructions, ces méthodes sont extrêmement coûteuses en temps de calcul. Les méthodes séquentielles (dites *two-step*) séparent le problème de décomposition en matériaux du problème de reconstruction tomographique. L’étape de DMB peut être réalisée dans l’espace des images [6] (i.e. après l’étape de reconstruction tomographique) ou dans l’espace des projections [7] (avant la reconstruction). La DMB est un problème non-linéaire et mal posé (voir ci-dessous). Cet article propose une méthode de DMB dans l’espace des projections. De nombreuses méthodes existent : les premières méthodes proposaient d’approcher l’opérateur inverse par un modèle paramétrique plus simple (e.g. polynomial [4]). Les approches variationnelles estiment un maximum de vraisemblance poissonien ou gaussien [7]. Ces dernières années, de nombreuses approches basées sur les techniques d’apprentissage profond (*deep learning* ou DL) ont été proposées pour résoudre des problèmes inverses [5]. Plus récemment, des méthodes à base de DL tentent d’intégrer la physique sous-jacente,

à des degrés divers [1, 8].

Dans cet article, nous proposons une méthode qui s’appuie sur le déroulement (dit *unrolling*) d’une descente de gradient afin d’améliorer les résultats de la décomposition obtenus avec les méthodes variationnelles classiques.

## 2 Matériel et méthodes

### 2.1 Modèle de formation des mesures

Le coefficient d’atténuation linéaire  $\mu$  d’un objet d’étude dépend de l’énergie  $E$  du rayonnement incident et peut être écrit comme combinaison linéaire des coefficients d’atténuation  $f_m$  d’un faible nombre  $M$  (typiquement  $M = 2$ ) de matériaux [3]

$$\mu(\mathbf{x}, E) = \sum_{m=1}^M f_m(E) \tau_m(\mathbf{x}). \quad (1)$$

Lorsque l’objet est soumis à un rayonnement polychromatique, la mesure en un pixel  $\mathbf{u}$  est modélisée par la loi de Beer-Lambert

$$\bar{s}_b(\mathbf{u}) = \int_0^\infty S_b^{\text{eff}}(E) \exp\left(-\sum_{m=1}^M f_m(E) a_m(\mathbf{u})\right) dE, \quad (2)$$

où  $S_b^{\text{eff}}(E)$  est le spectre effectif (i.e. tenant compte de la réponse du détecteur) dans le  $b^e$  canal d’énergie du détecteur et  $a_m(\mathbf{u})$  est l’épaisseur projetée des cartes spatiales de matériaux  $\tau_m$ , i.e. l’intégrale des fonctions  $\tau_m(\mathbf{x})$  le long du rayon X (depuis la source jusqu’au pixel  $\mathbf{u}$  du détecteur)

$$a_m(\mathbf{u}) = \int_{\mathcal{L}(\mathbf{u})} \tau_m(\mathbf{x}) d\mathbf{x}. \quad (3)$$

La quantité  $\bar{s}_b(\mathbf{u})$  est en réalité le nombre de photons moyens. Le nombre de photons réels  $s_b$  suit une loi de Poisson

$$s_b(\mathbf{u}) \sim \mathcal{P}(\bar{s}_b(\mathbf{u})). \quad (4)$$

Par analogie à la tomographie classique, nous choisissons de travailler sur les atténuations

$$\hat{s}_b(\mathbf{u}) = -\log \frac{s_b(\mathbf{u})}{s_b^0(\mathbf{u})} \quad (5)$$

où  $s_b^0(\mathbf{u}) = \int_0^\infty S_b^{\text{eff}}(E) dE$  est le nombre total de photons mesurés en l'absence d'objet dans le faisceau incident.

En pratique, le détecteur numérique fournit une image discrétisée de  $P = P_x \times P_y$  pixels  $\mathbf{u}_{x,y}$ . Une mesure est donc la donnée de  $B$  images  $\hat{\mathbf{s}} = (\hat{\mathbf{s}}_1, \dots, \hat{\mathbf{s}}_B) \in \mathbb{R}^{P \times B}$  où  $\hat{\mathbf{s}}_b = (\hat{s}_b(\mathbf{u}_{x,y}))_{1 \leq x \leq P_x, 1 \leq y \leq P_y}$  est une image d'atténuation. Les grandeurs  $\mathbf{a}_m$  et  $\mathbf{s}_b$  sont toutes deux des images bi-dimensionnelles. Pour les distinguer, le terme de "projection" sera réservé aux images  $\mathbf{a}_m$ , qui désignent les longueurs projetées des cartes de matériaux  $\tau_m$  (voir (3)). La DMB dans le domaine des projections consiste à retrouver les  $M$  projections  $\mathbf{a} = (\mathbf{a}_1, \dots, \mathbf{a}_M) \in \mathbb{R}^{P \times M}$  à partir de  $B$  images d'atténuation  $\hat{\mathbf{s}}$ . Typiquement,  $B = 2$  : l'atténuation est mesurée à deux niveaux d'énergie différents. En notant  $\mathcal{F}$  le modèle direct, non-linéaire et discrétisé des données, on a

$$\mathcal{F} : \begin{array}{ccc} \mathbb{R}^{P \times M} & \rightarrow & \mathbb{R}^{P \times B} \\ \mathbf{a} & \mapsto & \hat{\mathbf{s}} \end{array} \quad (6)$$

La DMB consiste à inverser l'opérateur  $\mathcal{F}$ . Estimer l'inverse de l'opérateur  $\mathcal{F}$  peut se faire par approche variationnelle, qui minimise une énergie définie comme la somme d'un terme de fidélité aux données et d'une régularisation  $\mathcal{R}$ . Cela revient à minimiser, par rapport à la variable  $\mathbf{a}$ , la fonctionnelle suivante

$$F(\mathbf{a}, \hat{\mathbf{s}}) := \frac{1}{2} \|\mathcal{F}(\mathbf{a}) - \hat{\mathbf{s}}\|_2^2 + \mathcal{R}(\mathbf{a}) \quad (7)$$

Bien que le bruit photonique soit poissonnien (Équation 4), les niveaux de flux typiques en tomographie spectrale rendent légitime l'utilisation de la norme  $L_2$  du résidu comme critère d'attache aux données. Des méthodes itératives basées sur une approximation de Gauss-Newton ont été proposées dans [7] mais le choix de la régularisation  $\mathcal{R}$  et des hyperparamètres (spécifiques à un niveau de bruit) est délicat. Nous proposons ici un schéma itératif appris, l'algorithme de Deep Gradient Descent (DGD), obtenu en déroulant une descente de gradient. En utilisant les données d'entraînement, les méthodes basées de DL apprennent une décomposition optimale  $\Lambda_\theta : \mathbb{R}^{P \times B} \rightarrow \mathbb{R}^{P \times M}$  associée à un réseau de neurones paramétré par un ensemble de paramètres  $\theta$  à optimiser. Parmi les méthodes d'apprentissage supervisés, les méthodes *unrolling* s'appuient sur un schéma itératif pour minimiser le terme d'attache aux données dans l'Équation 7. Le schéma itératif est tronqué après un nombre fixé d'itérations  $K$  et certaines mises à jour sont modifiées en ajoutant des réseaux de neurones minces. Le réseau correspondant à  $\Lambda_\theta$  combine donc des réseaux de neurones avec les opérateurs direct et adjoint associés à la physique du problème [5, 2, 12]. Récemment nous avons montré que ce type d'approche peut être utilisé pour un problème inverse non linéaire en utilisant la dérivée de l'opérateur et son adjoint [11]. Nous considérons ici la méthode *unrolling* obtenue en adaptant une descente de gradient. Si  $\mathcal{R}$

est suffisamment régulière, le schéma de descente de gradient pour résoudre le problème (7) peut s'écrire, à l'itération  $k$

$$\mathbf{a}^{(k+1)} = \mathbf{a}^{(k)} - \lambda_k \nabla F(\mathbf{a}^{(k)}, \hat{\mathbf{s}}). \quad (8)$$

Si l'on arrête l'algorithme de descente après  $K$  itérations, on obtient alors

$$\mathbf{a}^{(K)} = (G_{\lambda_K} \circ \dots \circ G_{\lambda_1})(\mathbf{a}^{(0)}, \hat{\mathbf{s}}), \quad (9)$$

où  $G_{\lambda_k}(\cdot, \hat{\mathbf{s}}) := \text{Id} - \lambda_k \nabla F(\cdot, \hat{\mathbf{s}})$  pour  $1 \leq k \leq K$ . Dans l'approche *unrolling*, l'idée est de remplacer les fonctions  $G_{\lambda_k}$ , représentant une itération du schéma (8), par un réseau de neurones convolutif  $\Gamma_{\lambda_k}^\theta$ . Les réseaux  $\Gamma_{\lambda_k}^\theta$  partagent un ensemble de paramètres  $\theta$  qui sera optimisé par apprentissage. Le pas de descente  $\lambda_k$  sera lui aussi optimisé mais spécifique à chaque itération. Un réseau  $\Gamma_{\lambda_k}^\theta$  prend en entrée l'itérée  $\mathbf{a}^{(k)}$  ainsi que les atténuations mesurées  $\hat{\mathbf{s}}$ . À partir de celles-ci, le gradient du terme d'attache aux données est calculé

$$\nabla \left( \frac{1}{2} \|\mathcal{F}(\mathbf{a}^{(k)}) - \hat{\mathbf{s}}\|_2^2 \right) = \mathbf{J}^{(k)\top} [\mathcal{F}(\mathbf{a}^{(k)}) - \hat{\mathbf{s}}] \quad (10)$$

où  $\mathbf{J}^{(k)}$  représente la jacobienne de l'opérateur  $\mathcal{F}$  au point  $\mathbf{a}^{(k)}$ . L'architecture du réseau  $\Gamma_{\lambda_k}^\theta$  est représentée dans la Figure 1. L'architecture est inspirée de [9], elle a été adaptée au problème de la DMB. Contrairement à [9], les filtres sont de tailles  $3 \times 3$  et les fonctions d'activation utilisées sont des Leaky ReLU, définie par  $\text{LReLU}_\alpha(x) = \max(x, \alpha x)$ ,  $\alpha > 0$ . De plus, les pas  $\lambda_k \in \mathbb{R}^M$  sont différents à chaque itération, avec un pas de descente pour chacun des canaux. Cette structure de réseau, similaire à un algorithme de descente de gradient, est appelé Deep Gradient Descent (DGD). On définit ainsi un réseau de neurones convolutif  $\Lambda_\theta = \Gamma_{\lambda_K}^\theta \circ \dots \circ \Gamma_{\lambda_1}^\theta$  représentant  $K$  itérations et  $(\theta, \lambda = \{\lambda_k\}_{k=1}^K)$  représentent les paramètres de ce réseau.

## 3 Protocole expérimental

### 3.1 Données

Le jeu de données utilisé est construit sur une base de 1600 fantômes numériques tridimensionnels. Chaque fantôme comporte entre 4 et 7 ellipsoïdes, eux-mêmes constitués d'eau ou d'os (les deux matériaux de base utilisés pour la DMB). Chaque fantôme est ainsi constitué d'une paire  $(\tau_1, \tau_2)$  de carte matériaux 3D (voir §2.1). Taille, orientation et position des ellipsoïdes sont choisis aléatoirement dans des limites qui garantissent d'une part que le fantôme est entièrement contenu dans une boîte englobante de  $200 \times 200 \times 200 \text{ mm}^3$  et d'autre part une variabilité suffisante. Chaque fantôme  $(\tau_1, \tau_2)$  est ensuite projeté sous 10 angles de vue différents pour donner les

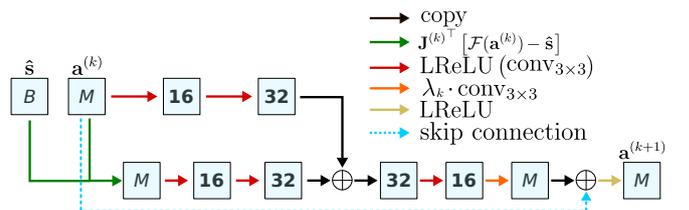


FIGURE 1 : Architecture du réseau  $\Gamma_{\lambda_k}^\theta$  qui représente une itération de l'approche Deep Gradient Descent.

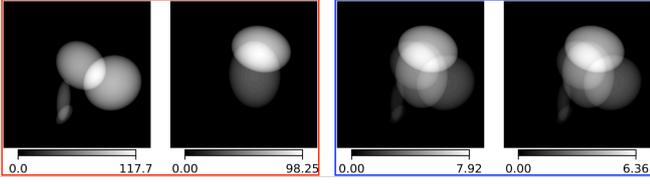


FIGURE 2 : Un exemple de paire  $(\mathbf{a}, \hat{\mathbf{s}})$  du jeu de données. *Encadré rouge* : Projection de référence d'eau (*gauche*) et d'os (*droite*). *Encadré bleu* : Les mesures basse énergie (*gauche*) et haute énergie (*droite*).

projections de référence  $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2)$ . Puis le modèle spectral complet décrit dans §2.1 est appliqué avec deux spectres effectifs différents  $S_1^{\text{eff}}$  et  $S_2^{\text{eff}}$  correspondant à des tensions de tube X de 80 kVp et 120 kVp respectivement. Les images ainsi obtenues  $\hat{\mathbf{s}} = (\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2)$  constituent les paires d'images en entrée des algorithmes de DMB. Au total, le jeu de données est ainsi composé de 16000 (1600 fantômes pris sous 10 angles de vue différents) paires d'images  $\hat{\mathbf{s}}^n$  ( $1 \leq n \leq 16000$ ). Chaque paire est associée à la paire de projections de référence correspondante  $\mathbf{a}^n$ . La Figure 2 illustre un exemple de paire du jeu de données. Le jeu de données est ensuite découpé en 14000 paires pour l'entraînement des réseaux, 1000 paires pour la validation et 1000 paires pour le test (toutes issues de fantômes non utilisés pour l'entraînement).

Quatre jeux de données  $\mathcal{T}(\varepsilon) = \{(\mathbf{a}^n, \hat{\mathbf{s}}^n), n \leq 16000\}$  ont été générés, correspondant à des niveaux de bruit distincts. Par convention,  $\mathcal{T}(0)$  est un jeu de données sans bruit : l'équation 4 n'est pas appliquée. Les valeurs  $\varepsilon = 10^4, 10^5$  et  $10^6$  correspondent à des nombres de photons incidents (le nombre total de photons émis par les spectres d'émission  $S_1^{\text{eff}}$  et  $S_2^{\text{eff}}$ ), qui déterminent le niveau de bruit dans les images  $\hat{\mathbf{s}} = (\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2)$ . Plus le nombre de photons incidents  $\varepsilon$  est élevé, moins il y a de bruit dans les données.

### 3.2 Entraînement et implémentation

Notre stratégie d'entraînement consiste à utiliser une fonction de coût régularisée avec des poids exponentiels pour optimiser les paramètres du réseau [13]

$$\min_{\theta} \mathbb{E}_{(\hat{\mathbf{s}}, \mathbf{a}^*)} \sum_{k=1}^K e^{-\tau(K-k)} \|\mathbf{a}^{(k)} - \mathbf{a}^*\|_2^2 \quad (11)$$

où  $K$  désigne le nombre d'itérations du réseau déroulé,  $\mathbf{a}^*$  est la vérité terrain,  $\mathbf{a}^{(k)} = (\Gamma_{\lambda_k}^{\theta} \circ \dots \circ \Gamma_{\lambda_1}^{\theta})(\mathbf{a}^{(0)}, \hat{\mathbf{s}})$ , et  $\tau$  est un paramètre qui contrôle la pénalisation des couches au cours des itérations. Pénaliser la fonction de coût force donc les sorties des itérations intermédiaires à être proches de la vérité terrain, ce qui a pour effet de changer la fonction du réseau. Lorsque  $\tau \rightarrow +\infty$ , seule la sortie du réseau  $\mathbf{a}^{(K)}$  est optimisée, ce qui améliore l'ajustement sur les données d'apprentissage.

Le nombre d'itérations est fixé à  $K = 10$  : à chaque *batch*, l'opérateur  $\mathcal{F}$  et la jacobienne  $\mathbf{J}$  sont évalués dix fois. Le choix du paramètre  $\tau$  a été fixé à  $10^{-1}$ , ce qui correspond à la meilleure erreur de validation pour  $\tau \in \{10^{-2}, 10^{-1}, 1, 10\}$ . Le nombre total de paramètres appris par le réseau  $\Lambda_{\lambda}^{\theta}$  est d'environ  $28 \times 10^3$ . Le réseau est entraîné sur 100 époques, avec une taille de *batch* fixé à 10, un optimiseur d'Adam, un pas d'apprentissage initialisé à  $5 \times 10^{-4}$  avec une décroissance en cosinus. La valeur du paramètre  $\alpha$  de la fonction d'activation LReLU a été laissée par défaut à 0.3.

Nous avons comparé l'algorithme DGD à une méthode de Gauss-Newton régularisée [7], avec pour régularisation Tikhonov du premier et du second ordre pour l'os et l'eau, respectivement, ainsi qu'une contrainte de positivité. Les paramètres de régularisation ont été choisis en suivant la procédure décrite dans [7], qui a donné  $\alpha = 10^{-1.5}$  et  $\beta_{\text{eau}} = \beta_{\text{os}} = 1$ . Un autre réseau, U-Net, a été entraîné à décomposer directement à partir des mesures d'atténuations en entrée. Cette méthode n'inclut pas la physique du modèle. L'architecture utilisée du réseau U-Net correspond à celle de [10]. Elle compte environ  $30 \times 10^6$  paramètres, a été entraînée avec 100 époques, une taille de batch de 10. La fonction de coût est l'erreur quadratique. Les entraînements ont duré environ 20 h et 15 h pour les réseaux DGD et U-Net, respectivement. Les résultats présentés ici pour DGD sont obtenus en considérant des initialisations nulles  $\mathbf{a}^{(0)} = \mathbf{0}_{\mathbb{R}^{P \times M}}$ . Les réseaux ont été entraînés sur chacune des bases de données, ce qui fait un total de huit réseaux entraînés.

### 3.3 Métriques d'évaluation

Pour évaluer la qualité de la décomposition en matériaux obtenus par les réseaux, nous avons utilisé deux métriques : l'erreur quadratique normalisée  $\text{NMSE}(\mathbf{a}, \mathbf{a}^*) = \frac{\|\mathbf{a} - \mathbf{a}^*\|_2^2}{\|\mathbf{a}^*\|_2^2}$ . Plus la NMSE est faible, meilleure est la décomposition.

D'autre part, l'indice de similarité  $\text{SSIM}(\mathbf{a}, \mathbf{a}^*) = \frac{(2m_{\mathbf{a}}m_{\mathbf{a}^*} + C_1)(2\sigma_{\mathbf{a}\mathbf{a}^*} + C_2)}{(m_{\mathbf{a}}^2m_{\mathbf{a}^*}^2 + C_1)(\sigma_{\mathbf{a}}^2\sigma_{\mathbf{a}^*}^2 + C_2)}$  où  $m_{\mathbf{a}}$  et  $\sigma_{\mathbf{a}}$  représentent la moyenne et l'écart-type des pixels de  $\mathbf{a}$ , respectivement, et  $\sigma_{\mathbf{a}\mathbf{a}^*}$  la covariance entre  $\mathbf{a}$  et  $\mathbf{a}^*$  et  $C_1 = 10^{-4}$ ,  $C_2 = 9 \times 10^{-4}$  sont des constantes.

## 4 Résultats

Les résultats obtenus sont résumés dans le Tableau 1. Lorsque le bruit est important, la qualité de la décomposition se dégrade. Dans tous les cas, les performances de la méthode déroulée DGD dépassent celles de U-Net. De plus, bien que les erreurs normalisées des deux méthodes soient à peu près similaires, on observe qu'en terme de similarité, l'approche DGD surpasse U-Net, en ayant un indice presque parfait. Pour la méthode Gauss-Newton et la base de données non bruitées, la NMSE moyenne est de 0.035 (0.12) et 0.022 (0.021) et l'indice SSIM moyen de 0.351 (0.075) et 0.437 (0.072), pour l'eau et l'os respectivement. Pour les bases de données bruitées, la méthode Gauss-Newton donne aussi des résultats moins bons en moyenne avec certains cas qui divergent pour des paramètres de régularisation fixés.

L'évolution des moyennes des métriques ainsi que les écart-types pour DGD sur la base de données test  $\mathcal{T}(10^4)$ , sont représentés dans la Fig. 3. Au niveau du logarithme de la NMSE, les toutes premières itérations permettent d'avoir l'amélioration la plus importante. Cela ne semble plus s'améliorer beaucoup après la cinquième itération. Cependant, l'indice SSIM ne cesse de s'améliorer au fil des itérations. Ce comportement de l'évolution des métriques est dû à la régularisation imposée à la fonction coût (11), qui force les sorties des itérations intermédiaires à être proches de la sortie désirée.

Un exemple de décomposition en eau/os, à partir de la donnée test Figure 2, est donné dans la Fig 4. Bien que les paramètres de régularisation pour Gauss-Newton aient été optimisés, le bruit est trop élevé pour obtenir une décomposition

TABLE 1 : Moyenne (écart-type) des métriques NMSE et SSIM pour les base de données de test à différents niveaux de bruit.

	$\mathcal{T}(0)$		$\mathcal{T}(10^6)$		$\mathcal{T}(10^5)$		$\mathcal{T}(10^4)$	
	Eau	Os	Eau	Os	Eau	Os	Eau	Os
<b>NMSE</b>								
U-Net	0.032 (0.073)	0.016 (0.018)	0.037 (0.062)	0.017 (0.016)	0.056 (0.111)	0.019 (0.019)	0.064 (0.102)	0.029 (0.020)
DGD	<b>0.027 (0.076)</b>	<b>0.014 (0.021)</b>	<b>0.036 (0.127)</b>	<b>0.015 (0.021)</b>	<b>0.034 (0.067)</b>	<b>0.015 (0.018)</b>	<b>0.053 (0.091)</b>	<b>0.019 (0.021)</b>
<b>SSIM</b>								
U-Net	0.367 (0.067)	0.592 (0.064)	0.331 (0.071)	0.552 (0.044)	0.292 (0.067)	0.424 (0.059)	0.239 (0.071)	0.319 (0.062)
DGD	<b>0.988 (0.008)</b>	<b>0.982 (0.011)</b>	<b>0.986 (0.009)</b>	<b>0.974 (0.013)</b>	<b>0.985 (0.010)</b>	<b>0.973 (0.017)</b>	<b>0.965 (0.017)</b>	<b>0.953 (0.028)</b>

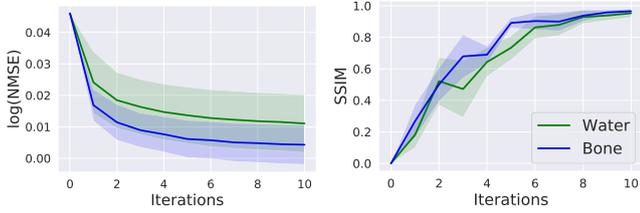


FIGURE 3 : Évolution des moyennes des métriques  $\log(\text{NMSE})$  et  $\text{SSIM}$  au fil des itérations pour la méthode DGD sur la base de données test  $\mathcal{T}(10^4)$ . Les zones transparentes représentent l'écart-type.

satisfaisante. Quand à celles obtenues par les réseaux de neurones, la décomposition ainsi que les NMSE sont similaires, mais DGD présente un indice de similarité plus élevé que U-Net pour les deux matériaux. La durée moyenne pour la décomposition d'un cas a été de 30 s, 0.6 s et 1.7 s pour Gauss-Newton, U-Net et DGD, respectivement. Les réseaux, en déportant la charge de calcul sur la phase d'apprentissage, permettent donc une décomposition 20 à 50 fois plus rapide que la méthode variationnelle GN, qui nécessite ici une dizaine d'itérations pour converger elle aussi.

## 5 Conclusion et discussion

Dans ce travail, nous avons présenté une approche *unrolling* appliquée à une descente de gradient pour le problème inverse non linéaire de la DMB. Cette méthode consiste à dérouler une simple descente de gradient en la modifiant avec un réseau de neurones convolutif. La méthode améliore la qualité des décompositions et raccourcit le temps calcul nécessaire. Elle préserve les propriétés de convergence de la descente de gradient et permet d'apprendre une régularisation adaptée. Une extension du travail proposé serait d'appliquer le même type de schéma à des algorithmes itératifs plus complexes tels que la méthode de Gauss-Newton ou une méthode primale-duale non linéaire.

## Références

[1] J. F. P. J. ABASCAL, N. DUCROS et F. PEYRIN : Nonlinear material decomposition using a regularized iterative scheme based on the Bregman distance. *Inverse Probl.*, 2018.

[2] J. ADLER, A. RINGH, O. ÖKTEM et al. : Learning to solve inverse problems using Wasserstein loss. *arXiv :1710.10898*, 2017.

[3] R. E. ALVAREZ : Near optimal energy selective x-ray imaging system performance with simple detectors. *Med. Phys.*, 2010.

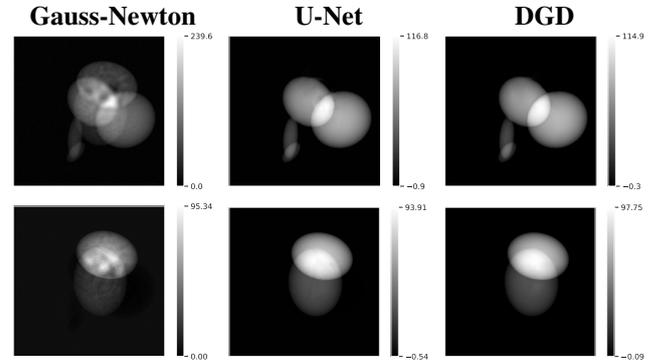


FIGURE 4 : Décomposition obtenue pour la paire test représentée dans la Figure 2 avec  $\epsilon = 10^4$ . La première ligne correspond à l'eau et la seconde à l'os.

[4] R. E. ALVAREZ et A. MACOVSKI : Energy-selective reconstructions in X-ray computerized tomography. *Phys. Med. and Biol.*, 1976.

[5] S. ARRIDGE, P. MAASS, O. ÖKTEM et al. : Solving inverse problems using data-driven models. *Acta Numer.*, 2019.

[6] Q. DING, T. NIU, X. ZHANG et al. : Image-domain multimaterial decomposition for dual-energy CT based on prior information of material images. *Med. Phys.*, 2018.

[7] N. DUCROS, J. F. P. J. ABASCAL, B. SIXOU et al. : Regularization of nonlinear decomposition of spectral x-ray projection images. *Medical Physics*, 2017.

[8] A. EGUIZABAL, O. ÖKTEM et M. U. PERSSON : Deep Learning for Material Decomposition in Photon-Counting CT. *arXiv :2208.03360*, 2022.

[9] A. HAUPTMANN, F. LUCKA, M. BETCKE et al. : Model-based learning for accelerated, limited-view 3-d photoacoustic tomography. *IEEE TMI*, 37(6):1382–1393, 2018.

[10] K. H. JIN, M. T. MCCANN, E. FROUSTEY et al. : Deep convolutional neural network for inverse problems in imaging. *IEEE TMI*, 26:4509–4522, 2017.

[11] K. MOM, M. LANGER et B. SIXOU : Apprentissage de méthodes itératives pour l'imagerie de contraste de phase des rayons x. *In GRETSI 2022*, 2022.

[12] V. MONGA, Y. LI et Y. C. ELДАР : Algorithm unrolling : Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Process. Mag.*, 2021.

[13] V. VISHNEVSKIY, J. WALHEIM et S. KOZERKE : Deep variational network for rapid 4D flow MRI reconstruction. *Nat. Mach. Intell.*, 2020.