



OPEN Deep learning estimation of effective atomic number for HU to RED calibration in dual energy photon counting CT

Kihong Son¹ & Daehong Kim²✉

This study introduces a deep learning (DL) approach to improve relative electron density (RED) estimation in dual-energy CT (DECT) through precise effective atomic number (EAN) prediction. A benchtop photon-counting detector CT acquired spectral images of a tissue-equivalent phantom, and a modified U-Net was trained on synthetic data to directly estimate EAN. Predicted EANs were converted to RED using a physics-based model. Performance was assessed on eight materials using mean absolute error (MAE), relative error, and residuals, with results compared to conventional Rutherford and stoichiometric methods. The DL model achieved an MAE of 0.08% for EAN, outperforming Rutherford (1.59%) and stoichiometric (1.54%) approaches. For RED estimation, DL reached a MAE of 0.62%, with residuals ranging from -0.01 to 0.02 of theoretical values. Calibration analysis showed high linearity for conventional Δ HU-RED curves ($R^2 \approx 0.9995$), but the DL-based method showed slightly higher linearity ($R^2 = 0.9998$). Overall, in this phantom study, the DL framework improved RED accuracy and HU-RED linearity compared with analytical methods. By providing a highly linear Δ HU-RED calibration curve in the phantom experiments conducted in this study, it enhances material differentiation and has the potential to support more precise dose calculations in future clinical applications.

Keywords Dual-energy CT, Deep learning, Effective atomic number, Relative electron density, HU-RED calibration

Radiation therapy aims to maximize treatment effect by delivering highly precise radiation doses to tumors while minimizing unnecessary exposure to surrounding healthy tissues. To achieve this, radiation treatment planning systems (RTPS) rely on accurate dose calculations based on the patient's anatomical information. One of the critical factors influencing dose calculation accuracy is the precise estimation of relative electron density (RED).

Currently, most clinical practices estimate RED using HU-RED calibration curves derived from single-energy CT (SECT) images. These curves are typically constructed based on predefined tissue-equivalent materials. However, SECT has intrinsic limitations—materials with different chemical compositions can yield the same Hounsfield Unit (HU) values. This ambiguity is particularly problematic for high-density materials or heterogeneous tissues, leading to reduced RED estimation accuracy^{1–3}. Such uncertainties can propagate into significant dose calculation errors, thereby compromising the precision of radiation therapy. To address these limitations, considerable effort has been devoted to refining HU-RED calibration techniques based on dual-energy CT (DECT). DECT acquires images using two distinct energy spectra, allowing for the quantitative estimation of physical parameters such as the effective atomic number (EAN) and RED. This approach enhances the differentiation of tissue-specific X-ray attenuation characteristics, improving both material decomposition and dose calculation accuracy.

Among analytical models for DECT-based RED estimation, the Rutherford model⁴ and stoichiometric calibration⁵ are widely recognized. The Rutherford model models the linear attenuation coefficient as a combination of contributions from Compton scattering, Rayleigh scattering, and the photoelectric effect, capturing attenuation behavior across energy levels. In contrast, the stoichiometric method calculates theoretical

¹Medical Information Research Section, Electronics and Telecommunications Research Institute, Daejeon, Republic of Korea. ²Department of Radiological Science, Eulji University, Seongnam, Republic of Korea. ✉email: goldcollar011@gmail.com

attenuation coefficients based on known chemical compositions, offering higher accuracy in some scenarios. Both models use EAN as an intermediate variable to indirectly compute RED.

Bazalova et al. demonstrated that EAN and RED can be estimated with mean errors of 2.8% and 1.8%, respectively, using a DECT-based linear attenuation model². Alvarez et al. proposed an approach that expresses linear attenuation coefficients as functions of atomic number and mass density, enabling EAN estimation via dual-energy Eq⁶. While these methods validated the analytical potential of DECT, they are limited by dependencies on spectral correction accuracy, prior knowledge of material composition, and sensitivity to image noise⁷. These analytical models have also been applied to photon-counting detector CT (PCD-CT). Son et al. used PCD-CT images separated into two energy bins and applied stoichiometric calibration to quantitatively estimate EAN and RED⁸. Although their study highlighted the potential of EAN/RED estimation using high-resolution spectral data, challenges remain in spectral decomposition accuracy, image noise, and the limited predictability in high-density tissues. These limitations can ultimately result in dose calculation errors during radiation therapy planning⁷.

Recently, deep learning (DL) has emerged as an effective approach for solving quantitative analysis and material decomposition problems in medical imaging. In particular, convolutional neural networks (CNNs) are well-suited to learn the complex nonlinear relationships between CT spectral data and physical parameters such as EAN and RED^{9,10}. DL inherently provides noise suppression and has the potential to generalize across diverse tissue compositions and structural complexities, although such generalization must be validated through further experimental studies. The PCD-CT platform, which provides spectrally resolved data across multiple energy bins, is especially well-suited to leverage the advantages of DL-based approaches. However, to date, no study has directly estimated both EAN and RED using DL methods on PCD-CT-based DECT images. This study aims to address this technical gap.

In this work, we propose a customized deep learning architecture based on the U-Net structure, tailored for quantitative material property prediction. Adjustments were made to the input channels, output format, and loss function to optimize the model for the task. The designed network directly estimates EAN, which is then converted to RED, forming the core component of a novel framework for generating HU-RED calibration curves. In our evaluation, the proposed method achieved better quantitative accuracy and structural consistency than the analytical models. The resulting HU-RED calibration curve also showed highly linear behavior, indicating that the DL-based framework can provide a more reliable foundation for HU-RED calibration in future radiation treatment planning studies.

This study aims to compare the quantitative accuracy of DL-based EAN estimation with conventional analytical methods, evaluate the precision and structural fidelity of the derived RED, and quantitatively assess the linearity and clinical applicability of the resulting HU-RED calibration curves.

Methods

PCD-CT and data acquisition

In this study, a benchtop PCD-CT system was configured to acquire DECT-based imaging data. The system comprises an X-ray tube (RAD-14, Varian Medical Systems, UT), a cadmium telluride (CdTe)-based photon-counting detector (XC-TDI, Varex Imaging, UT), and a rotatable sample stage. The X-ray beam was oriented perpendicular to the detector plane. The detector features a pixel size of 100 μm with an array size of 3584 \times 60 pixels, resulting in an active detection area of 360 \times 6 mm^2 . The detector operates in photon-counting mode with a linear array configuration and is capable of separating the X-ray spectrum into two distinct energy bins.

During scanning, the object was rotated between a stationary X-ray source and detector. The source-to-object distance (SOD) was set to 829.0 mm, and the object-to-detector distance (ODD) was set to 508.4 mm. Image reconstruction was performed using the Feldkamp-Davis-Kress (FDK) algorithm. The experimental subject was a Gammex 472 phantom (Sun Nuclear, FL, USA), consisting of eight tissue-equivalent materials. Each rod simulates the physical properties of human tissues. Key parameters such as EAN, RED, and physical density for each material are summarized in Table 1.

Image acquisition was performed at a tube voltage of 125 kVp. A full-spectrum image was first acquired, and the same spectrum was subsequently divided into low-energy (20–60 keV) and high-energy (60–125 keV) bins to generate the respective DECT images. This binning strategy was designed to balance photon counts between the two energy ranges and thereby enhance the quantitative accuracy of material decomposition. Energy bins below 20 keV were excluded from analysis due to a significant increase in image noise.

No.	Material	EAN	RED	Density (g/cm^3)
1	General adipose	6.44	0.94	0.960
2	Breast	6.76	0.97	0.985
3	Solid water	7.24	1.00	1.023
4	Brain	7.40	1.02	1.052
5	Liver	7.38	1.05	1.081
6	Inner bone	10.01	1.16	1.218
7	CaCO_3 50%	12.02	1.46	1.560
8	Cortical bone	13.29	1.78	1.924

Table 1. EAN and RED of the eight tissue-equivalent material rods located in the Phantom shown in Fig. 1a.

An example of the phantom and corresponding DECT images used in this study is presented in Fig. 1, illustrating the spatial arrangement of the tissue-equivalent materials and the image characteristics across different energy spectra.

Deep learning model for EAN estimation

In this study, we developed a deep learning model for the quantitative estimation of the EAN from DECT data. The model is based on a modified U-Net architecture, structurally optimized for the accurate and continuous prediction of physical quantities. While U-Net is commonly used for segmentation tasks, we redesigned the architecture as a regression model by adapting the output format, loss function, and layer configurations to enable precise prediction of continuous values.

The proposed network follows the encoder–decoder structure of U-Net. Each stage consists of a sequence of 3×3 convolutional layers, batch normalization, and ReLU activation functions. To preserve spatial information and minimize resolution loss during upsampling, bilinear interpolation and skip connections were employed. The final output layer generates a pixel-wise EAN map, with each pixel representing a continuous real-valued prediction.

To avoid excessive complexity, the deepest layers of the U-Net (e.g., those with 1024 filters) were omitted. This design choice improves computational efficiency and reduces the risk of overfitting. Additionally, to further enhance generalization and accurately predict fine structural details and material boundaries, a dropout rate of 0.5 was applied during training.

As illustrated in Fig. 2, the network architecture excludes fully connected layers and is composed entirely of convolutional layers. This design enhances the model's flexibility, enabling input images of varying resolutions without requiring architectural modifications. It also avoids the computational burden and parameter explosion typically associated with fully connected layers, making the model more suitable for efficient deployment in clinical applications across diverse medical imaging datasets.

Unlike conventional loss functions such as mean squared error (MSE), we employed the structural similarity index measure (SSIM) as the sole loss function. SSIM is highly sensitive to local contrast and structural features, and it effectively emphasizes perceptual image fidelity. In this study, SSIM was selected to better reflect structural, luminance, and contrast-related features critical for accurate estimation of continuous physical quantities such as EAN.

To construct the synthetic training dataset, we first selected one representative axial slice from the Gammex 472 phantom. Material-specific ROIs were manually delineated for eight tissue-equivalent inserts (adipose, breast, solid water, brain, liver, inner bone, 50% CaCO_3 , and cortical bone), and patches extracted from these ROIs were resampled to the network input size to form a pool of material-specific patches. A circular rod mask was then nonlinearly deformed using a mesh-based warping technique to generate a diverse set of rod shapes and positions. For each warped mask, the interior region was filled with a randomly selected material patch, while the background region was constructed using solid-water patches extracted from the same representative

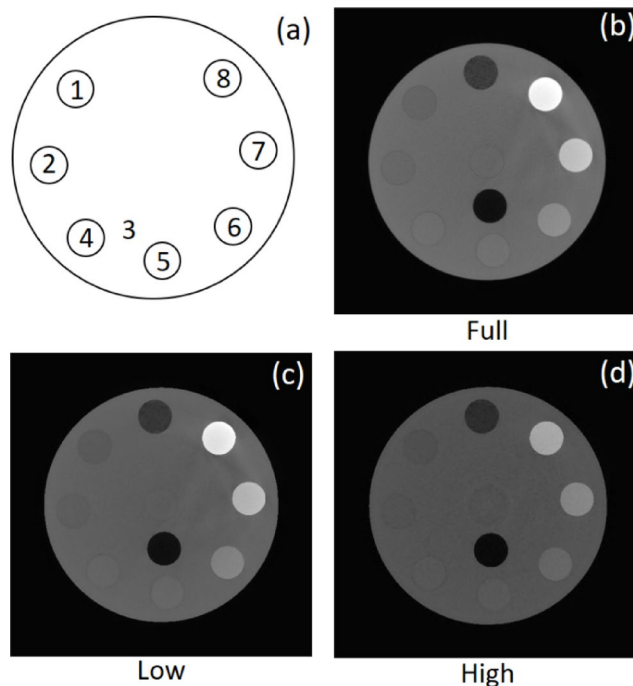


Fig. 1. (a) shows the structure of the phantom along with the numbers and materials of the eight rods. (b) is the full-energy image, (c) and (d) are the low-energy and high-energy images of the phantom in (a), acquired using PCD-CT and displayed in HU.

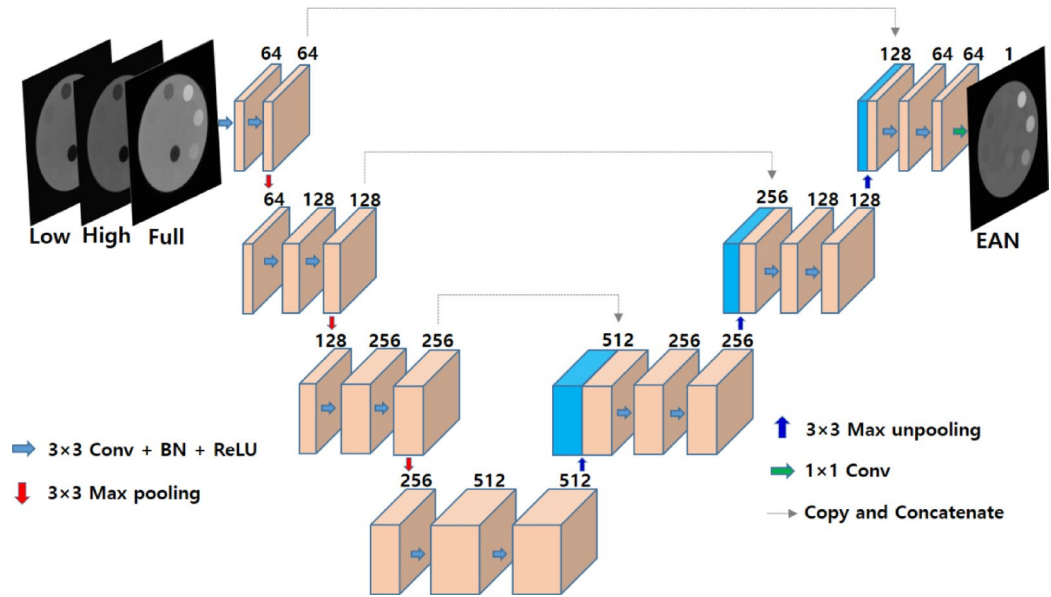


Fig. 2. Architecture of the proposed deep learning model for EAN estimation. The network consists of a modified U-Net with 3 × 3 convolutions, batch normalization, ReLU activation, bilinear upsampling, and skip connections. It was trained for 300 epochs, with SSIM loss stabilizing below 0.005 within 30–40 epochs, and no signs of overfitting observed.

slice. This yielded a multi-channel PCD-CT input sample consisting of full-, low-, and high-energy images. For each synthetic input sample, the corresponding pixel-wise ground-truth EAN map was generated using the manufacturer-provided effective atomic numbers of the Gammex materials. Pixels inside the rod region were assigned the EAN of the selected insert, and background pixels were assigned the EAN of solid water, producing a piecewise-constant but spatially structured label map.

By repeating this process across materials and deformation combinations, a total of 112,000 multi-channel image–label pairs were generated. These synthetic samples were used exclusively for training and validation and were randomly divided at the patch level with an approximately 80/20 split.

Only the single representative Gammex slice described above was used to generate the 112,000 synthetic training and validation samples. The remaining axial slices of the Gammex 472 phantom, including those from repeated acquisitions, were not used in the synthetic data generation or the training/validation process and were reserved exclusively as a fully independent held-out test set. A separate skull phantom acquisition was used only for qualitative evaluation. No synthetic data were used for testing. For the held-out test set, the trained DL model was directly applied to the reconstructed PCD-CT images to generate EAN and RED maps, and all quantitative metrics (MAE, RMSE, SSIM, and R² of the ΔHU–RED curves) were computed from these predictions.

Model performance during training was monitored using the SSIM loss, which measures structural fidelity. Training was performed for 300 epochs, during which the SSIM loss decreased rapidly over the first 30–40 epochs and subsequently stabilized below 0.005 without indications of overfitting. This behavior indicates that the DL model effectively learns both quantitative accuracy and structural preservation from the synthetic training data.

Analytical method for estimation of EAN

For comparison with the deep learning approach, RED values were also estimated using the Rutherford and stoichiometric methods. RED estimation was performed following the estimation of EAN. The approach proposed by Rutherford et al. allows for EAN estimation using the following Eq. 4.

$$\frac{HU_L + 1000}{HU_H + 1000} = \frac{1 + AZ_{\text{eff}}^{m-1}}{B + CZ_{\text{eff}}^{m-1}} \tag{1}$$

where, HU_L and HU_H represent the CT numbers at low and high energy levels, respectively, and A, B, C, and m are fitting parameters. Additionally, Z_{eff} represents the EAN of the materials. The values of A, B, and C were determined by fitting the known EAN values of the eight materials to the ratio ($\frac{HU_L + 1000}{HU_H + 1000}$) using MATLAB (R2013b, MathWorks, USA). After obtaining the parameters, Eq. (1) was implemented as a function, and an iterative algorithm was used to extract the EAN map from the CT images. The stoichiometric model proposed by Schneider et al. can be expressed by the following equation, which is used to estimate the EAN⁵.

$$\frac{HU_L + 1000}{HU_H + 1000} = \frac{1 + AZ_{\text{eff}}^{1.86} + BZ_{\text{eff}}^{3.62}}{1 + AZ_w^{1.86} + BZ_w^{3.62}} \times \frac{1 + CZ_w^{1.86} + DZ_w^{3.62}}{1 + CZ_{\text{eff}}^{1.86} + DZ_{\text{eff}}^{3.62}}, \tag{2}$$

where, HU_L and HU_H represent the CT numbers at low and high energy levels, respectively, and A, B, C, and D are fitting parameters. Z_{eff} denotes the EAN of water, which was set to 7.24 in this study using solid water as a reference.

Both the Rutherford and stoichiometric methods estimate EAN prior to RED calculation. In this study, the EAN estimation using the Rutherford and stoichiometric methods was implemented in MATLAB R2013b (MathWorks, Natick, MA, USA) and performed on a workstation equipped with a 3.20 GHz CPU and 16.0 GB of RAM using an iterative algorithm to extract the EAN map for a 512×512 image matrix.

Software and hardware environment

The deep learning model was implemented in Python 3.7.10 (Spyder environment) using the TensorFlow–Keras 2.5.0 framework and trained and evaluated on a workstation equipped with an NVIDIA RTX A6000 GPU. Thus, the analytical and DL-based methods were executed in different software and hardware environments.

RED estimation

The RED was estimated using the EAN values obtained from the Rutherford, stoichiometric, and DL methods, based on the following equation proposed by Zhu et al.¹¹.

$$\frac{\left(\frac{HU(E)}{1000} + 1\right)}{\rho_e} = \frac{[1 + A(E) Z_{\text{eff}}^{1.86} + B(E) Z_{\text{eff}}^{3.62}]}{[1 + A(E) Z_w^{1.86} + B(E) Z_w^{3.62}]} \quad (3)$$

Here, Z_{eff} is the estimated EAN, and Z_w is the EAN of solid water. $HU(E)$ represents the CT number of the materials in the low- or high-energy images ($E = \text{low or high}$), and ρ_e is the theoretical RED value listed in Table 2. $A(E)$ and $B(E)$ are fitting parameters. In this study, the HU values from the low-energy image, theoretical RED values from Table 2, and the estimated EANs of the eight materials were used to determine $A(E)$ and $B(E)$ by fitting Eq. (3). The fitted parameters $A(E)$ and $B(E)$ were then substituted into Eq. (3), and the estimated EAN images and low-energy HU images obtained from the three methods were used to compute the corresponding RED maps.

HU-RED conversion

For both DL-based and analytical RED estimations, the corresponding HU values must be recalculated using the weighted subtraction method. This is because DECT provides two HU values at different energy levels, whereas treatment planning systems require a single HU value corresponding to each material's RED. Therefore, the weighted subtraction method described by Saito¹² was applied to estimate ΔHU using both low- and high-energy HU values.

$$\Delta HU \equiv (1 + \alpha) HU_H - \alpha HU_L, \quad (4)$$

Here, HU_H and HU_L represent the CT numbers from the high- and low-energy images, respectively. The parameter α is the weighting factor used for subtraction, and it can be determined using Eq. (5) as proposed by Saito (2012).

No.	Material	EAN				RED			
		Theory	Rutherford (error %)	Stoichiometric (error %)	DL (error %)	Theory	Rutherford (error %)	Stoichiometric (error %)	DL (error %)
1	General adipose	6.44	6.45 (0.20)	6.36 (-1.30)	6.44 (0.00)	0.94	0.96 (1.81)	0.96 (1.81)	0.95 (0.85)
2	Breast	6.76	6.68 (-1.26)	6.64 (-1.75)	6.76 (0.00)	0.97	0.98 (1.34)	0.99 (1.55)	0.98 (0.52)
3	Solid water	7.24	7.47 (3.22)	7.31 (1.01)	7.24 (0.00)	1.00	1.00 (-0.20)	1.00 (0.20)	1.00 (-0.10)
4	Brain	7.40	7.31 (-1.27)	7.21 (-2.57)	7.39 (-0.14)	1.02	1.02 (0.29)	1.02 (0.59)	1.02 (-0.49)
5	Liver	7.38	7.20 (-2.43)	7.17 (-2.83)	7.38 (0.00)	1.05	1.05 (-0.10)	1.04 (0.29)	1.04 (-1.05)
6	Inner bone	10.01	10.04 (0.28)	9.98 (-0.27)	10.01 (0.00)	1.16	1.16 (-0.17)	1.17 (0.00)	1.17 (0.60)
7	CaCO ₃ 50%	12.02	12.19 (1.44)	12.17 (1.26)	12.02 (0.00)	1.46	1.47 (0.89)	1.47 (1.10)	1.47 (0.48)
8	Cortical bone	13.29	13.64 (2.61)	13.47 (1.35)	13.22 (-0.53)	1.78	1.76 (-1.18)	1.77 (-1.12)	1.77 (-0.84)

Table 2. Relative errors (%) of EAN and RED estimated using the Rutherford, stoichiometric, DL methods compared to theoretical values.

$$\alpha = \frac{1}{\frac{\mu_L^w}{\mu_L^w - \rho_e} - 1}, \quad (5)$$

where, μ_k and μ_k^w represent the linear attenuation coefficients of a given material and water, respectively, where $k=L$ and $k=H$ denote low-energy and high-energy, respectively. ρ_e refers to the RED value estimated using the Rutherford, stoichiometric, or DL method.

Results

EAN and RED

Figure 3 presents the EAN and RED maps estimated using the Rutherford, stoichiometric, and DL methods on the held-out Gammex phantom slices. In this study, the DL model directly predicted EAN maps, and the corresponding RED maps were subsequently derived from the predicted EAN using Eq. (3). For convenience, both the EAN and RED results shown in Fig. 3 and discussed throughout this section are collectively referred to as outputs of the Rutherford, stoichiometric, or DL methods. Table 2 summarizes the EAN and RED values measured from each rod in Fig. 3, together with their relative errors (in percent) with respect to the theoretical reference values. The mean absolute error (MAE) of EAN for the DL method was 0.08%, with values closely matching the theoretical ones except for the brain (-0.14%) and cortical bone (-0.53%). In contrast, the MAEs for EAN using the Rutherford and stoichiometric methods were 1.59% and 1.54%, respectively.

For RED estimation, the DL method also showed the lowest MAE of 0.62% across the eight materials, while the Rutherford and stoichiometric methods yielded MAEs of 0.75% and 0.83%, respectively.

Figure 4a and b illustrate the relative errors of EAN and RED for eight materials estimated using the Rutherford, stoichiometric, and DL methods. For the Rutherford method, the highest relative errors in EAN and RED were observed in solid water (3.22%) and general adipose (1.81%), respectively. In the stoichiometric method, liver and general adipose showed the highest relative errors in EAN and RED, at -2.83% and 1.81%, respectively. For the DL method, the largest errors were observed in cortical bone (0.53%) for EAN and in liver (1.05%) for RED.

Figure 4c and d present the residuals of EAN and RED, respectively, for the Rutherford, stoichiometric, and DL-based methods, computed with respect to the theoretical reference values for each material.

Figure 4c shows the residuals of EAN. For EAN estimation, the DL-based method exhibited residuals ranging from -0.07 to 0.00 across the eight tissue-equivalent materials, with the largest negative deviation observed for cortical bone (-0.07) and near-zero residuals for most other materials. In comparison, the Rutherford method showed larger variability, with EAN residuals ranging from -0.35 to 0.18, while the stoichiometric method exhibited residuals spanning -0.18 to +0.21, depending on the material.

Figure 4d shows the residuals of RED. For RED estimation, the DL-based method exhibited residuals ranging from -0.01 to 0.02, while both the Rutherford and stoichiometric methods showed residuals within the range of -0.02 to 0.02.

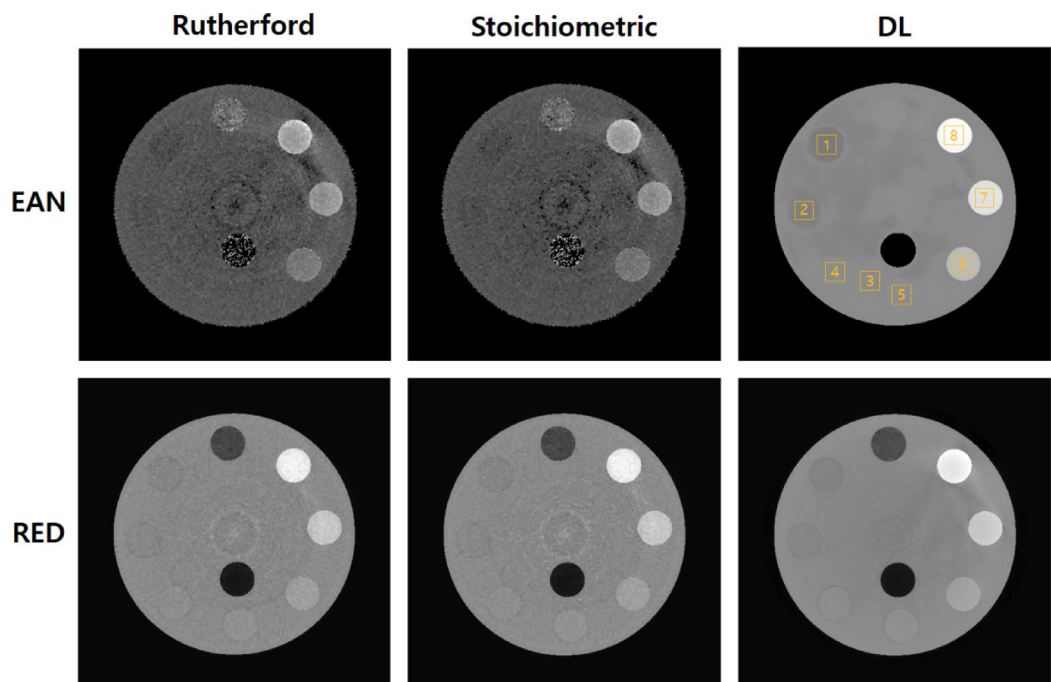


Fig. 3. Estimated images of EAN and RED extracted from Rutherford, stoichiometric, and DL, respectively.

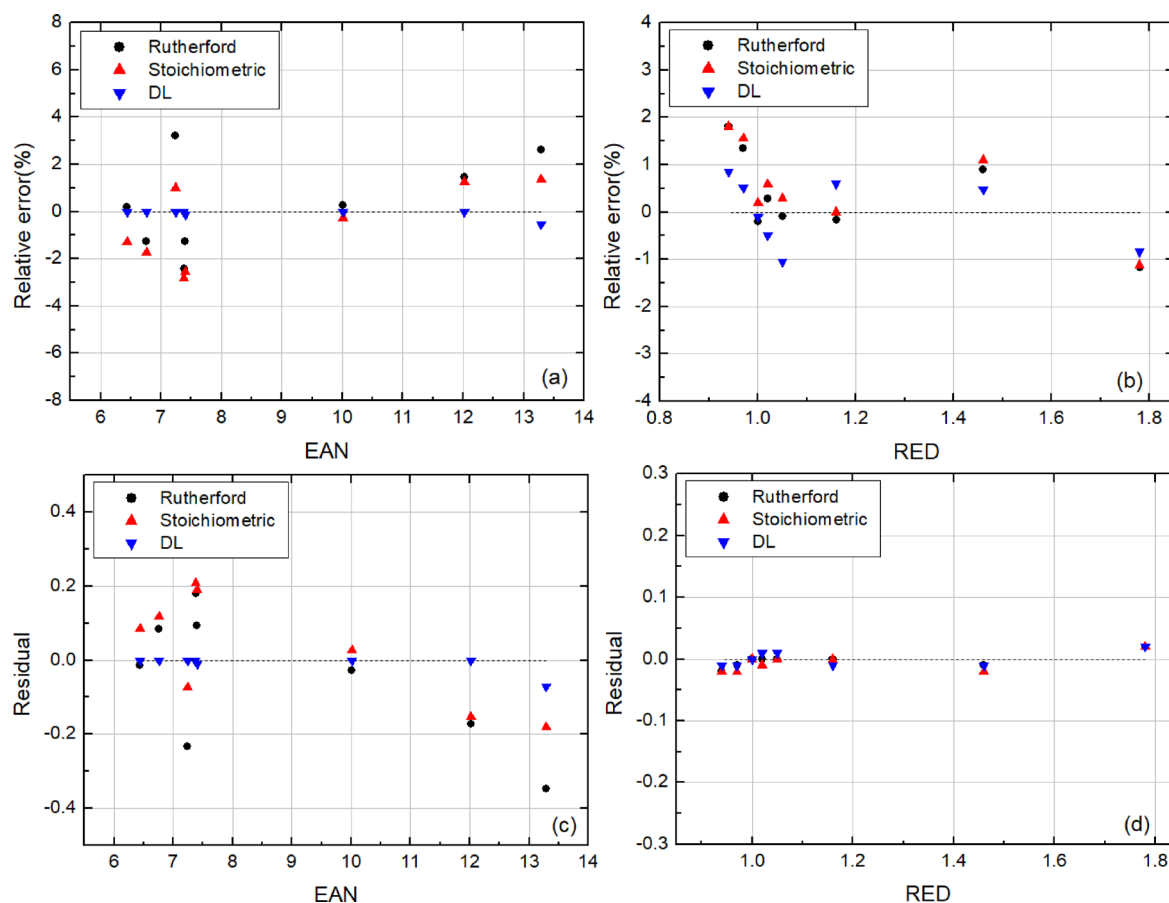


Fig. 4. (a) and (b) is the relative error (%) of EAN and RED for Rutherford, stoichiometric, and DL, respectively. (c) and (d) is their residual of EAN and RED for Rutherford, stoichiometric, and DL, respectively.

Method	SSIM	
	EAN	RED
Rutherford	0.7577	0.8865
Stoichiometric	0.7530	0.8877
DL	0.9575	0.9333

Table 3. SSIM of EAN and RED Phantom images estimated by Rutherford, stoichiometric, and DL methods.

The structural similarity index measure (SSIM) between the EAN and RED maps in Fig. 3 and their corresponding ground truths is summarized in Table 3. The DL-based method showed higher SSIM values, indicating improved structural similarity relative to the analytical methods in this phantom study. For the EAN maps, the Rutherford and stoichiometric methods showed comparable SSIM values of 0.7577 and 0.7530, respectively, while the DL method achieved a substantially higher SSIM of 0.9575, indicating improved structural preservation. Similarly, for the RED maps, the DL method yielded the highest SSIM of 0.9333, outperforming the Rutherford (0.8865) and stoichiometric (0.8877) methods. These findings suggest that the DL-based approach offers more precise structural fidelity in complex material composition estimation than traditional methods.

Table 4 presents the root mean square error (RMSE) values of EAN and RED for the eight materials shown in Fig. 3, calculated by comparing the estimates from the Rutherford, stoichiometric, and deep learning (DL)-based methods with the ground truth images. Overall, the DL-based method demonstrated the lowest RMSEs for both EAN and RED estimation in most tissues, indicating superior quantitative accuracy compared to the conventional methods. Notably, for soft tissues such as general adipose, breast, brain, and liver, the DL model consistently outperformed the Rutherford and stoichiometric methods. For example, in the case of brain tissue, the RMSE of EAN estimation was significantly reduced from 0.7600 (Rutherford) and 0.7522 (stoichiometric) to 0.0128 (DL). Similarly, RED estimation improved from approximately 0.039 to 0.0139. In liver tissue, the DL model also achieved a notably low RMSE of 0.0077 for EAN, more than an order of magnitude lower than the other methods.

No.	Material	RMSE					
		Rutherford		Stoichiometric		DL	
		EAN	RED	EAN	RED	EAN	RED
1	General adipose	0.8793	0.0470	0.8007	0.0493	0.0461	0.0190
2	Breast	0.8420	0.0448	0.8005	0.0444	0.0357	0.0160
3	Solid water	0.7429	0.0375	0.7218	0.0357	0.0954	0.0138
4	Brain	0.7600	0.0396	0.7522	0.0382	0.0128	0.0139
5	Liver	0.7491	0.0401	0.7427	0.0384	0.0077	0.0159
6	Inner bone	0.7656	0.0402	0.7875	0.0425	0.1765	0.0240
7	CaCO ₃ 50%	0.9495	0.0492	0.9426	0.0503	0.1425	0.0357
8	Cortical bone	0.9776	0.0560	0.9118	0.0534	0.0952	0.0388

Table 4. RMSE of the EAN and RED maps estimated using the Rutherford, stoichiometric, and DL-based methods.

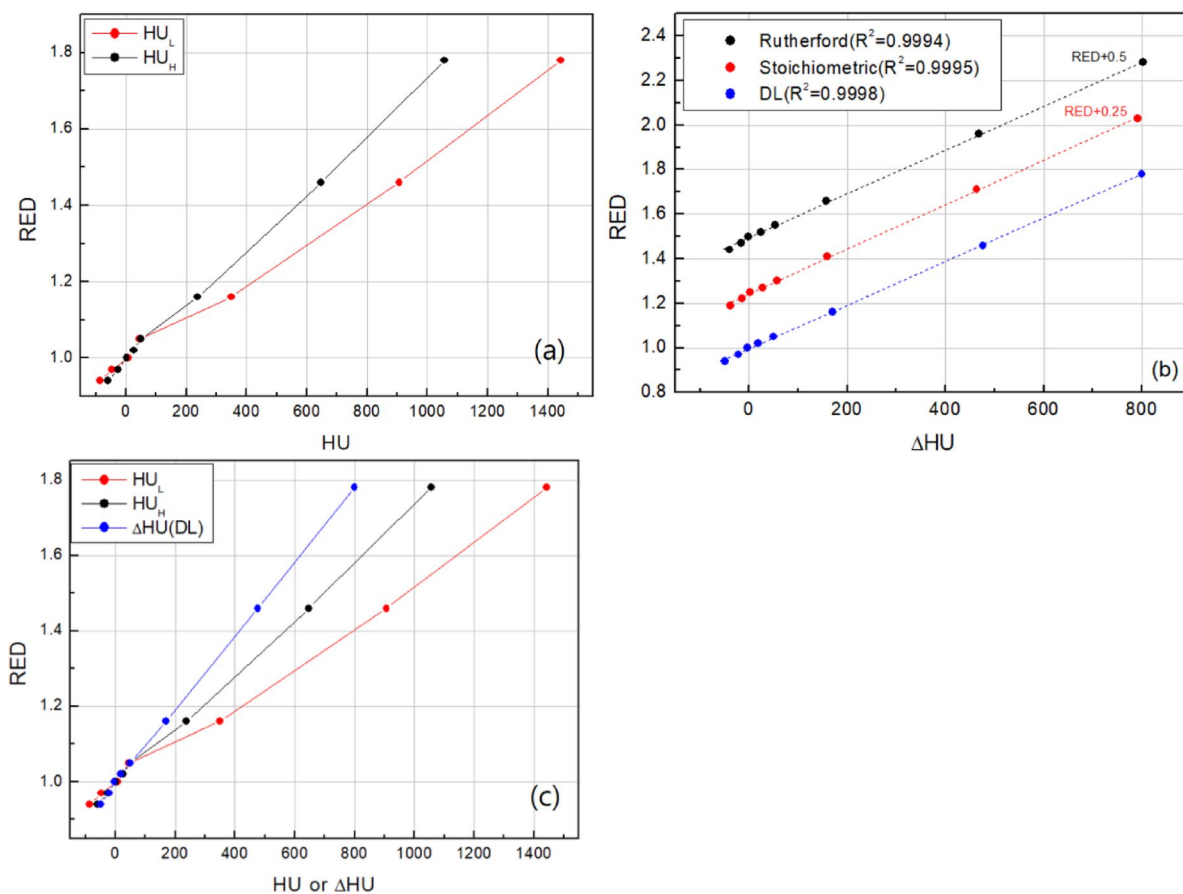


Fig. 5. (a) HU-RED curves at low- and high-energy images. (b) Δ HU-RED curve from Rutherford, stoichiometric, and DL, respectively. (c) comparison between HU-RED and Δ HU-RED (DL-based) curve.

In contrast, for high-density materials such as inner bone and 50% CaCO₃, the DL model showed relatively higher RMSEs in EAN estimation (0.1765 and 0.1425, respectively). Nevertheless, the DL method yielded the lowest RED RMSEs, maintaining stable performance in RED estimation. Across all materials, RMSEs for RED estimation were generally lower than those for EAN, and the DL-based model demonstrated high stability and accuracy, with most RED RMSEs remaining below 0.04. In particular, solid water (0.0138) and breast (0.0160) showed especially low values.

HU-RED curves

Using the RED values estimated by the Rutherford, stoichiometric, and DL-based methods, HU-RED calibration curves were generated. Figure 5a presents the HU-RED curves constructed using the theoretical RED values corresponding to HU_L and HU_H for each material, similar to the conventional calibration procedure in SECT.

Figure 5b illustrates the ΔHU –RED curves, in which ΔHU was calculated using Eqs. (4) and (5) with the RED values estimated by the three methods. To avoid overlapping of the three ΔHU –RED curves, the Rutherford and stoichiometric data points were vertically shifted upward by 0.5 and 0.25 units, respectively. In the ΔHU –RED curves, the dotted lines represent the first-order polynomial fitting curves for each method, and the dots indicate the calculated ΔHU values. Among the three methods, the DL-based approach exhibited the highest agreement with the theoretical values, achieving an R^2 of 0.9998, compared to 0.9994 for the Rutherford method and 0.9995 for the stoichiometric method.

Figure 5c compares the HU–RED curves from HUL and HUH with the DL-based ΔHU –RED calibration curve obtained from DECT. The HU–RED curves tend to show a decreased slope for high-density materials (e.g., inner bone, 50% CaCO_3 , cortical bone) compared to low-density materials (e.g., general adipose, breast, solid water, brain, liver), resulting in a nonlinear behavior across the entire density range. In contrast, the ΔHU –RED curve demonstrates a linear relationship regardless of the material density.

Skull phantom results

Using the skull phantom, the three conventional methods for estimating EAN and RED—Rutherford, stoichiometric, and the proposed deep learning (DL) approach—were applied. Overall, the DL-based method demonstrated superior image quality and structural differentiation performance. The major anatomical structures in the skull phantom, including brain tissue, iodine, and cortical bone, are indicated at the top of Fig. 6. The DL model not only provided clearer delineation between these structures but also significantly reduced image noise. In contrast, while both the Rutherford and stoichiometric methods successfully separated iodine and cortical bone, they exhibited high noise fluctuations in the brain region.

In the RED maps shown at the bottom of Fig. 6, the DL-based estimation also achieved the best performance in terms of structural preservation and noise suppression. In particular, in the brain region, the Rutherford and stoichiometric methods suffered from elevated noise, leading to reduced tissue separability, and iodine could not be distinctly separated from surrounding tissues.

Figure 7 presents the intensity profiles extracted from the same location as in Fig. 6. In the EAN profile shown in Fig. 7a, all three methods successfully distinguished between iodine and cortical bone. However, in the brain region, the DL method maintained a uniform and flat profile, whereas the other two methods exhibited pronounced noise fluctuations. Similarly, in the RED profile shown in Fig. 7b, cortical bone was well differentiated by all methods, but iodine was clearly separable only in the DL results. In the other methods, iodine showed similar intensity to brain tissue, making material discrimination difficult. The DL-based approach also showed the lowest noise amplitude in the brain region, whereas the Rutherford and stoichiometric methods presented higher noise levels and structural instability.

Discussion

This study proposed a HU–RED calibration framework that integrates deep learning–based EAN estimation with RED computation using PCD-CT imaging. The proposed DL method showed better quantitative accuracy

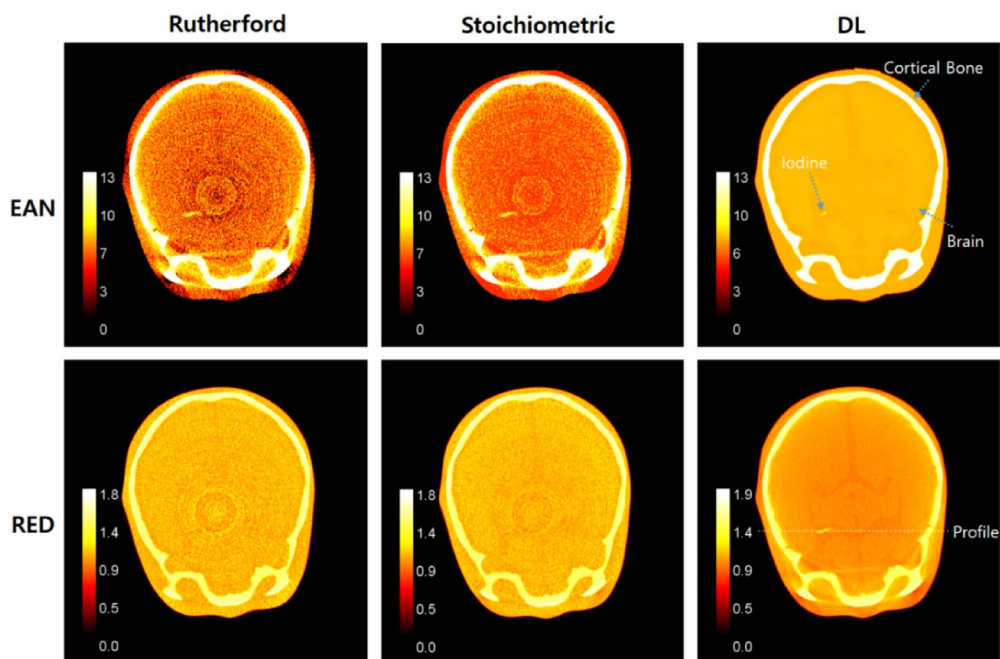


Fig. 6. Estimated EAN and RED in the skull phantom using Rutherford, stoichiometric, and DL-based method. The main structures of the skull phantom include the brain, iodine, and cortical bone, and the profiles of these three materials are presented in Fig. 5.

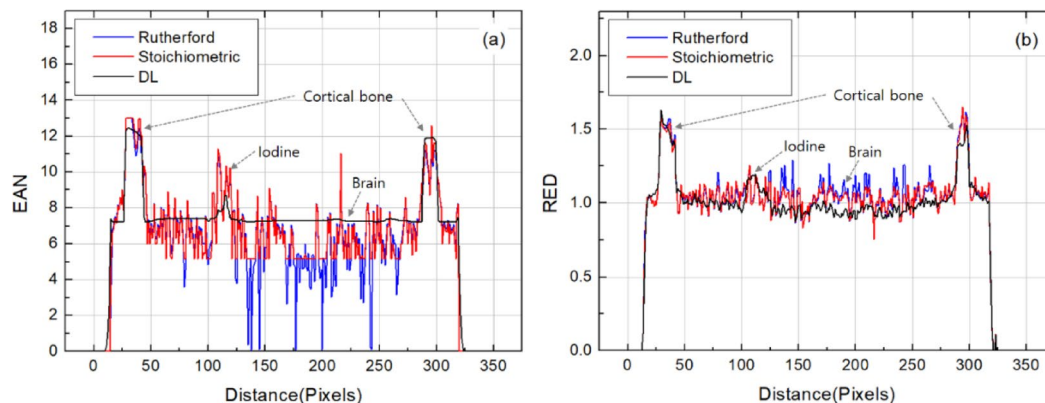


Fig. 7. The profiles include brain, iodine, and cortical bone. **(a)** shows the EAN profile, where the three materials are clearly distinguishable across all methods; however, the Rutherford and stoichiometric methods exhibit large noise fluctuations. **(b)** shows the RED profile, where the DL method clearly separates all three materials, whereas the Rutherford and stoichiometric methods fail to clearly differentiate between brain and iodine.

and structural preservation than the Rutherford and stoichiometric models in our phantom evaluation. These results demonstrate that the DL-based HU–RED calibration framework is technically feasible and can provide a more stable basis for dose-related calculations in future RTPS studies.

In terms of EAN estimation accuracy, the DL model achieved exact agreement with theoretical values for 6 out of 8 tissue-equivalent materials, while maintaining relative errors below 1% for both brain (−0.14%) and cortical bone (−0.53%). This indicates a substantially higher quantitative accuracy compared to conventional methods (Rutherford: 1.59%, Stoichiometric: 1.54%), suggesting that the DL model effectively integrates the complex spectral information from PCD-CT to accurately learn the underlying physical attenuation properties.

The precision of the DL-based estimation was further examined through material-wise residual analysis. For EAN estimation, the DL method yielded residuals ranging from −0.07 to 0.00, indicating close agreement with theoretical values, although a non-zero deviation was observed for cortical bone. In addition, the RED residual results showed that the DL-based method, with residuals ranging from −0.01 to 0.02, produced results closer to the theoretical reference than the Rutherford and stoichiometric methods.

In terms of RED estimation, the DL-based method achieved a MAE of 0.62%, surpassing both Rutherford (0.75%) and Stoichiometric (0.83%) methods. RED estimation accuracy is particularly influenced by the HU values from low-energy CT images, which serve as input variables, making it highly susceptible to image noise and energy-dependent attenuation coefficient variations. Despite these challenges, the DL model maintained overall prediction stability, achieving a residual range of ± 0.01 for all tissues. This not only demonstrates high quantitative accuracy but also suggests significant clinical implications, as the method stays well within the acceptable limits for patient dose calculation^{13,14}.

In addition to quantitative performance, the DL-based method demonstrated significantly superior results in terms of structural preservation. SSIM analysis showed values of 0.9575 and 0.9333 for the EAN and RED maps, respectively, which were notably higher than the SSIM values reported by the analytical methods (EAN: ~ 0.76 , RED: ~ 0.89). The high structural similarity in EAN, in particular, indicates that the DL model effectively captures local contrast characteristics, tissue boundary information, and fine tissue structures between spectra. For RED, a high SSIM was also maintained, emphasizing that the DL model provides stable performance even in RED estimation, which is typically more sensitive to noise. Furthermore, the RMSE of the 8 materials was also evaluated. Overall, the DL-based method demonstrated the lowest RMSE for EAN and RED estimation across most tissues, proving superior quantitative accuracy compared to traditional methods. In particular, the EAN estimation suggests that the DL model can sensitively reflect subtle differences in tissues with lower density. For all materials, the RMSE for RED estimation was generally lower than that for EAN, which reflects the more direct relationship between RED and attenuation coefficients in the clinical DECT energy range. This performance improvement can be attributed to the combined effects of the training data generation strategy, architecture design, and the choice of loss function employed in this study. In particular, the use of synthetic data based on clinical image-derived patches, anatomical diversity achieved through mesh-based nonlinear spatial deformation, and the application of a structure-oriented loss function (SSIM) played a critical role in enhancing the generalization ability and structural preservation characteristics of the DL model.

A notable qualitative difference between the analytical and DL-based EAN maps is the markedly reduced random noise level and the apparent suppression of ring artifacts in the DL results. This behavior does not arise from an explicit post-processing denoising filter, but rather from the formulation of the task as a regression onto physically defined EAN label maps. The U-Net is trained to reproduce pixel-wise Z_{eff} distributions that are piecewise constant for each material and therefore contain no stochastic noise or ring structures. Consequently, the network learns to preserve material boundaries and low-frequency structural components that are consistent with the labels, while high-frequency components—such as random noise and ring artifacts—carry little predictive value and are implicitly attenuated. The SSIM-based loss further promotes structural fidelity,

contributing to the observed combination of edge preservation and noise reduction in the DL-derived EAN maps.

In the evaluation using the skull phantom, the boundaries between the brain, iodine, and cortical bone were most clearly distinguished using the DL method. Profile analysis also revealed that the DL method exhibited the lowest noise fluctuation across all regions, with stable structural preservation. This is an important finding as it demonstrates the ability to perform precise estimation even in regions with low tissue contrast, such as the brain, in high-resolution CT images.

The analysis of the HU–RED calibration curve further supports the improvement of the DL method. The DL-based calibration curve exhibited a slightly higher R^2 value (0.9998) than the analytical methods (0.9994–0.9995). Although the numerical differences in R^2 are small, the DL-based approach maintained a consistent slope across the evaluated material range, which may be beneficial for HU–RED calibration under phantom conditions. In DECT-based HU–RED curves, linearity generally decreases in high-density material regions, and prediction errors of RED accumulate with HU. However, the DL-based curve maintained a consistent slope and high linearity across the entire range, significantly improving the reliability of RED estimation corresponding to CT values. This enhances material segmentation accuracy in RTPS and can serve as a fundamental resource for individualized dose calculation.

From a computational standpoint, the trained network enabled rapid per-slice inference, requiring approximately 0.4–0.5 s for a full 512×512 slice on the RTX A6000 GPU, whereas the MATLAB R2013b/CPU implementations of the analytical methods required longer processing times. These runtimes represent characteristics of our specific implementations and hardware environments and should not be interpreted as a direct benchmark comparison. Rather, they indicate that the proposed DL framework is computationally feasible for use in practical phantom or planning workflows when modern GPU resources are available.

This study has several limitations. First, all evaluations were performed on phantoms (Gammex 472 and a skull phantom), so the reported accuracy of EAN and RED and HU–RED linearity reflect performance under controlled phantom conditions rather than patient environments. Second, the diversity of materials was limited to tissue-equivalent inserts, and clinically important tissues such as lung, blood, and contrast-enhanced mixtures were not assessed. Third, the method was developed and evaluated using a benchtop PCD–CT system; therefore, its applicability to clinical PCD–CT or conventional DECT systems remains to be validated, and large-scale clinical datasets with comparable spectral characteristics are not yet widely available. Finally, no dose calculations or RTPS simulations were performed using the derived HU–RED curves, so potential dosimetric benefits remain to be quantified.

Although the present study was conducted using a benchtop PCD–CT system, the proposed framework is not inherently restricted to this platform. The network architecture accepts multi-energy DECT images as input and can, in principle, be adapted to clinical photon-counting CT systems as well as conventional DECT implementations, such as dual-layer or kVp-switching CT, provided that appropriate spectral input channels and corresponding training data are available. In a clinical workflow, the trained model could be applied after standard image reconstruction to generate EAN and RED maps, followed by HU–RED calibration for use in radiation treatment planning systems. However, differences in spectral separation, detector response, and reconstruction pipelines across CT platforms necessitate system-specific training or fine-tuning. Therefore, further validation using clinical DECT datasets is required before routine clinical integration can be considered.

Future research can be expanded to include the development of a regression model for direct DL-based RED estimation, ensuring generalization through training that incorporates various tissues such as air, lung, fat, and blood. Additionally, automating the generation of patient-specific HU–RED curves and establishing individualized dose calibration systems could be explored. Notably, integrating this work with the patient-specific dose correction model proposed by Vaniqui et al. could lead to the development of an end-to-end framework that is applicable in RTPS¹⁵. This would mark a significant turning point, enhancing both the precision of radiation therapy and its prognostic prediction capabilities.

Conclusions

This study presents a technically feasible framework that integrates DL-based EAN estimation with RED computation and Δ HU–RED calibration using PCD–CT data. In the phantom experiments performed in this work, the proposed method demonstrated improved quantitative accuracy and structural fidelity compared with conventional analytical models (Rutherford and stoichiometric). The DL-based Δ HU–RED calibration curve also showed high linearity across the evaluated materials, providing a more robust basis for HU–RED calibration in future studies.

In the skull phantom evaluation, the DL-based approach provided clearer boundary delineation and reduced noise relative to analytical methods. Once trained, the model enabled rapid per-slice inference, indicating computational feasibility for eventual integration into RTPS workflows following appropriate validation.

However, all findings were obtained from phantoms acquired on a benchtop PCD–CT system, and the applicability of the proposed framework to clinical DECT or PCD–CT systems has not yet been verified. Future work will therefore include validation on clinical DECT/PCD–CT datasets, generation of patient-specific HU–RED calibration curves, and RTPS-based dose calculations to determine the practical clinical impact of the proposed approach.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 29 September 2025; Accepted: 19 December 2025

Published online: 23 December 2025

References

- Demir, H. et al. Investigation of the effect of calibration curves obtained from different computed tomography devices on the dose distribution of tomotherapy plans. *J. Med. Phys.* **49**, 545–550 (2024).
- Bazalova, M. et al. Dual-energy CT-based material extraction for tissue segmentation in Monte Carlo dose calculations. *Phy Med. Biol.* **53**, 2439–2456 (2008).
- Yohannes, I. et al. A formulation of tissue- and water-equivalent materials using the stoichiometric analysis method for CT-number calibration in radiotherapy treatment planning. *Phy Med. Biol.* **57**, 1173–1190 (2012).
- Rutherford, R. A. et al. Measurement of effective atomic number and electron density using an EMI scanner. *Neuroradiology* **11**, 15–21 (1976).
- Schneider, U. et al. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys. Med. Biol.* **41**, 111–124 (1996).
- Alvarez, R. E. et al. Energy-selective reconstruction in X-ray computerized tomography. *Phy Med. Biol.* **21**, 733–744 (1976).
- Möhler, C. et al. Methodological accuracy of image-based electron density assessment using dual-energy computed tomography. *Med. Phys.* **44**, 2429–2437 (2017).
- Son, K. et al. Improving the accuracy of the effective atomic number (EAN) and relative electron density (RED) with stoichiometric calibration on PCD-CT images. *Sensors*. **22** (2022).
- Koetzier, L. R. Deep learning image reconstruction for CT: technical principles and clinical prospects. *Radiology*. **306** (2023).
- Chang, H. Y. et al. Material decomposition using dual-energy CT with unsupervised learning. *Phys. Eng. Sci. Med.* **46**, 1607–1617 (2023).
- Zhu, J. et al. Dosimetric comparison of stopping power calibration with dual-energy CT and single-energy CT in proton therapy treatment planning. *Med. Phys.* **43**, 2845–2854 (2016).
- Saito, M. Potential of dual-energy subtraction for converting CT numbers to electron density based on a single linear relationship. *Med. Phys.* **39**, 2021–2030 (2012).
- Born, E. et al. *Quality Control of Treatment Planning Systems for Teletherapy: Recommendations 7* (Swiss Society for Radiobiology and Medical Physics (SGSMP), 1997).
- Kilby, W. et al. Tolerance levels for quality assurance of electron density values generated from CT in radiotherapy treatment planning. *Phy Med. Biol.* **47**, 1485–1492 (2002).
- Vaniqui, A. et al. The impact of dual energy CT imaging on dose calculations for pre-clinical studies. *Radiat. Oncol.* **12** (2017).

Author contributions

D. Kim and K. Son conducted the experiment. K. Son collected the PCD-CT data and designed the architecture of the AI model. D. Kim implemented the analytical model, analyzed data, and prepared the figures. K. Son and D. Kim wrote the manuscript, and all authors corrected the manuscript. D. Kim is the corresponding author.

Funding

This work was supported by internal fund of Electronics and Telecommunications Research Institute(ETRI) [25YR1610, Next-Generation Mammography: Development of Personalized Precision Breast Cancer Diagnostic Core Technology Utilizing Radiomics, AI, and Targeted Contrast Agents].

Declarations

Ethical statement

This study did not involve human participants or animal experiments. All experiments were performed using phantoms; therefore, ethical approval and informed consent were not required.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to D.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025