



PAPER

A feasibility study of enhanced prompt gamma imaging for range verification in proton therapy using deep learning

RECEIVED
3 September 2022REVISED
23 January 2023ACCEPTED FOR PUBLICATION
27 February 2023PUBLISHED
20 March 2023Zhuoran Jiang^{1,2}, Jeremy C Polf^{3,*}, Carlos A Barajas⁴, Matthias K Gobbert⁴ and Lei Ren^{1,2,3,*}¹ Medical Physics Graduate Program, Duke University, Durham, NC, 27705, United States of America² Department of Radiation Oncology, Duke University Medical Center, Durham, NC, 27710, United States of America³ Department of Radiation Oncology, University of Maryland School of Medicine, Baltimore, MD, 21201, United States of America⁴ Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, MD, 21250, United States of America

* Authors to whom any correspondence should be addressed.

E-mail: lren@som.umaryland.edu, jpolf@yahoo.com, zhuoran.jiang@duke.edu, 12barajasc@gmail.com and gobbert@umbc.edu**Keywords:** proton range verification, prompt gamma imaging, Compton camera, deep learning, proton therapy

Abstract

Background and objective. Range uncertainty is a major concern affecting the delivery precision in proton therapy. The Compton camera (CC)-based prompt-gamma (PG) imaging is a promising technique to provide 3D *in vivo* range verification. However, the conventional back-projected PG images suffer from severe distortions due to the limited view of the CC, significantly limiting its clinical utility. Deep learning has demonstrated effectiveness in enhancing medical images from limited-view measurements. But different from other medical images with abundant anatomical structures, the PGs emitted along the path of a proton pencil beam take up an extremely low portion of the 3D image space, presenting both the attention and the imbalance challenge for deep learning. To solve these issues, we proposed a two-tier deep learning-based method with a novel weighted axis-projection loss to generate precise 3D PG images to achieve accurate proton range verification. **Materials and methods:** the proposed method consists of two models: first, a localization model is trained to define a region-of-interest (ROI) in the distorted back-projected PG image that contains the proton pencil beam; second, an enhancement model is trained to restore the true PG emissions with additional attention on the ROI. In this study, we simulated 54 proton pencil beams (energy range: 75–125 MeV, dose level: 1×10^9 protons/beam and 3×10^8 protons/beam) delivered at clinical dose rates (20 kMU min^{-1} and 180 kMU min^{-1}) in a tissue-equivalent phantom using Monte-Carlo (MC). PG detection with a CC was simulated using the MC-Plus-Detector-Effects model. Images were reconstructed using the kernel-weighted-back-projection algorithm, and were then enhanced by the proposed method. **Results.** The method effectively restored the 3D shape of the PG images with the proton pencil beam range clearly visible in all testing cases. Range errors were within 2 pixels (4 mm) in all directions in most cases at a higher dose level. The proposed method is fully automatic, and the enhancement takes only ~ 0.26 s. **Significance.** Overall, this preliminary study demonstrated the feasibility of the proposed method to generate accurate 3D PG images using a deep learning framework, providing a powerful tool for high-precision *in vivo* range verification of proton therapy.

1. Introduction

When protons travel through matters, a peak in energy deposition is observed at the end of their range, referred to as ‘Bragg Peak’. Beyond the peak, energy deposition drops off sharply. This characteristic enables the proton therapy to achieve high dose concentrations within the targets while significantly sparing the surrounding healthy tissues. But in turn, the treatment precision highly depends on the proton range accuracy. Patient positioning errors and anatomy motions/changes can cause proton range uncertainties during delivery,

resulting in an underdose of targets and/or overdose of healthy tissues (Paganetti 2012). A safety margin can be added to the target to ensure it receives the prescribed dose. However, this inevitably leads to an increased target volume and consequently escalates the dose to the surrounding healthy tissues, defeating the purpose of using proton therapy to spare healthy tissues. Therefore, there is an urgent clinical need to verify the proton beam range during treatment so that strategies can be designed to correct range delivery errors to minimize their impact.

Various techniques have been developed to detect the secondary signals emitted during proton beam irradiation to achieve non-invasive *in vivo* dose verification (Knopf and Lomax 2013, Parodi and Polf 2018), such as positron emission tomography (PET) (Nishio *et al* 2010), magnetic resonance imaging (MRI) (Gensheimer *et al* 2010), proton-acoustic imaging (Assmann *et al* 2015, Jones *et al* 2015), and prompt gamma (PG) imaging (Richard *et al* 2010, Panthi *et al* 2020, Polf *et al* 2022). In particular, several studies (Min *et al* 2006, Kim 2009, Bom *et al* 2011, Kormoll *et al* 2011, Moteabbed *et al* 2011, Roellinghoff *et al* 2011, Verburg *et al* 2013) have shown a strong correlation between the origins of PG emissions and the dose deposited by proton beams using both measurements and the Monte Carlo (MC) simulations, demonstrating the possibilities of PG imaging for *in vivo* range verification. With the Compton camera (CC), the origins of PG emissions can be reconstructed in 3D volumes (Richard *et al* 2010) using back-projecting algorithms. But the PG image quality is limited by the false, scattered, and mis-ordered PG events detected by the CC. Recently, deep learning-based methods have been proposed to pre-process the CC measured PG data (Zoglauer and Boggs 2007, Munoz *et al* 2021, Polf *et al* 2022), substantially reducing the noise and artifacts within the images by discarding false PG events and correcting mis-ordered PG events. However, due to the limited view of CCs acquisition, the PG images still suffer from severe distortions, significantly limiting their accuracy in range verification.

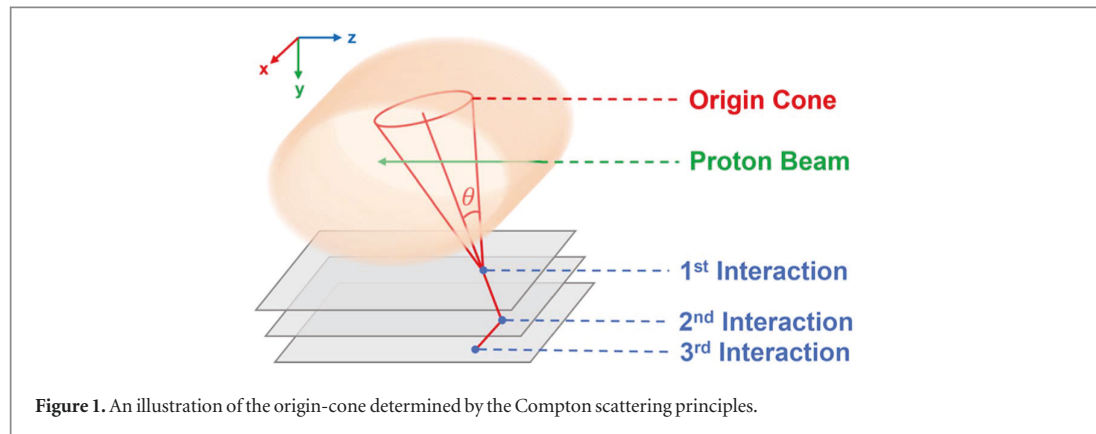
Reconstructing images from limited-view acquisition is essentially an ill-conditioned inverse problem. In recent years, deep learning has emerged as a powerful tool to address such challenges and generate high-quality volumetric images. Yang *et al* (2017) proposed a deep de-aliasing generative adversarial network (DA-GAN) for fast MRI reconstruction and showed superior performance to conventional compressed sensing-based algorithms in terms of both image quality and reconstruction speed. Previously, convolutional neural networks (CNNs) were developed to enhance cone-beam computerized tomography (CBCT) images reconstructed from sparse (Jiang *et al* 2019) and limited-angle (Jiang *et al* 2021) projections, generating high-quality volumes with clear and accurate structures. However, different from other medical images in which patient anatomy and other structures fill a large percentage of the image space, PG signals emitted along a single proton pencil beam path take up an extremely low portion of the total reconstructed 3D image space, presenting both the attention and the imbalance challenge for deep learning (Johnson and Khoshgoftaar 2019a, Bria *et al* 2020, Zhang *et al* 2020).

To address these issues, in this study, we proposed a deep learning-based two-tier scheme to reconstruct high-quality PG images using a Compton camera for proton range verification. First, a localization model is trained to define a region of interest (ROI) in the distorted back-projecting PG images that contains the proton pencil beam. And second, an enhancement model is trained to correct the distortions and restore the true PG signals with additional attention on the defined ROI. This study demonstrates several contributions to the field of CC-based PG imaging. (1) For the first time, we demonstrated the feasibility of deep learning to correct the distortion caused by limited-view acquisition and generate high-quality 3D PG images using a single Compton camera. (2) We proposed a two-tier deep learning framework using a novel weighted axis-projection loss to address the attention and the imbalance challenges in PG image enhancement. Ablation tests were performed to demonstrate its effectiveness in enhancing 3D PG images. (3) Quantitative metrics were used to evaluate the PG image errors in terms of both ranges and intensities. (4) The proposed method is fully automatic and generates accurate PG images in nearly real-time, providing a powerful tool for high-precision *in vivo* range verification in proton therapy.

2. Methods

2.1. Compton camera-based PG imaging

During the proton beam irradiation, nuclei of atoms in tissues are excited due to the proton-nuclear scatter. Characteristic PGs are emitted immediately following the interactions by the nuclei decaying to the ground state. When a PG exits the body and reaches the CC, it can Compton scatter in the different detection stages of the CC. These scatter interactions are recorded by the CC as an event, which contains information about the deposited energies and positions for each scatter interaction of the PG. From this event, an origin-cone of the PG emission can be determined, as shown in figure 1. Specifically, the origin-cone central axis is a vector from the second interaction location to the first interaction location, the origin-cone apex is the first interaction location, and the



origin-cone open-angle is calculated based on the principles of Compton scattering as

$$\cos(\theta) = 1 - m_e c^2 \left(\frac{1}{E_1} - \frac{1}{E_0} \right), \quad (1)$$

where θ is the origin-cone open-angle, E_1 is the residual energy after the first interaction, and E_0 is the incident energy of the PG which is estimated based on all the deposited energies of the event (Kroeger *et al* 2000, Peterson *et al* 2010, 2016). The true origin of the PG emission is restricted to the surface of the origin cone.

Therefore, a 3D image of PG emissions can be reconstructed by back-projecting the origin-cones of all events into the image domain, which is hereafter referred to as ‘back-projected PG’.

2.2. Overall workflow of the proposed PG enhancement

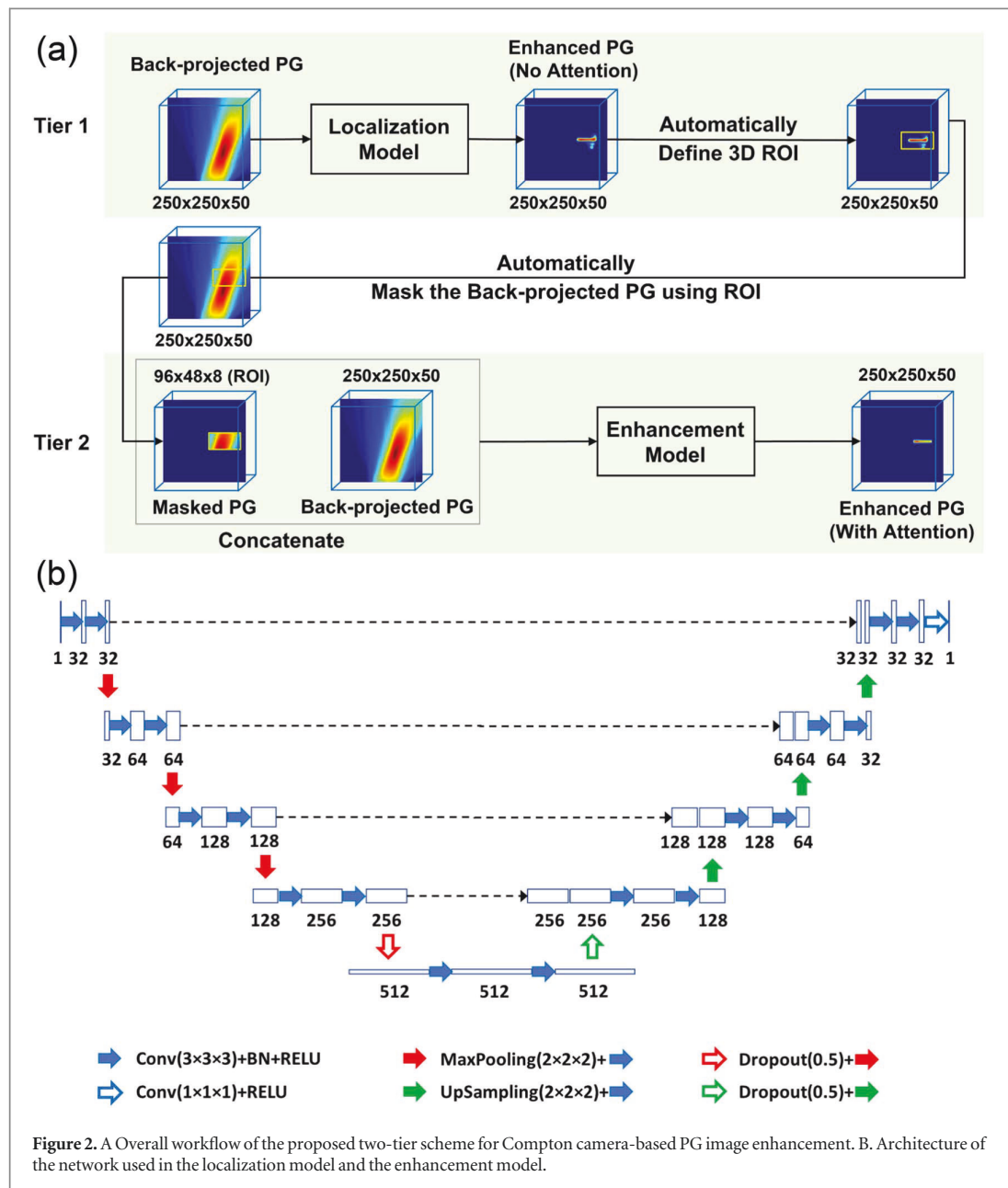
Figure 2(A) shows the overall workflow of the proposed method to enhance CC-based PG image quality. It is a two-tier scheme. First, a deep learning model referred to as the ‘localization model’ is trained to initially enhance the input back-projected PG image to reduce the distortions in the original large 3D space. Then, in the localization model output PG image (referred to as the enhanced PG (no attention)), a 3D ROI is automatically defined in the high-intensity region to direct attention around the path of the proton beam, along which most PGs emit. The back-projected PG image is automatically masked by the ROI. The masked PG and the back-projected PG images are concatenated in the channel dimension, and are then fed into the second deep learning model (referred to as the ‘enhancement model’). The enhancement model is trained to enhance the back-projected PG image with ROI attention to match with the ground truth, generating the final enhanced PG image (referred to as the enhanced PG (with attention)).

The localization model and the enhancement model have the same U-Net architecture, but are trained separately using different datasets. The network structure is illustrated in figure 2(B). More details can be found in Jiang *et al* (2021). This multi-scale architecture with feature concatenation at each scale level has demonstrated effectiveness in various image-related tasks (Ronneberger *et al* 2015, Liao *et al* 2018, Jiang *et al* 2020, Ernst *et al* 2021, Jiang *et al* 2021). Compared to the original U-Net (Ronneberger *et al* 2015), in this study, we made several major modifications to adapt to the current task. (1) 3D convolutional layers are used to extract features in 3D space. (2) Batch normalization layers are used after each convolutional layer to normalize the extracted features. This technique has been widely used to stabilize and accelerate the training process with a higher learning rate. (3) Dropout layers with a dropout rate of 0.5 are used in the U-shape bottom, where very high-dimensional features are extracted. They are used to improve the model’s generalizing capabilities and to avoid overfitting.

2.3. Experiment design

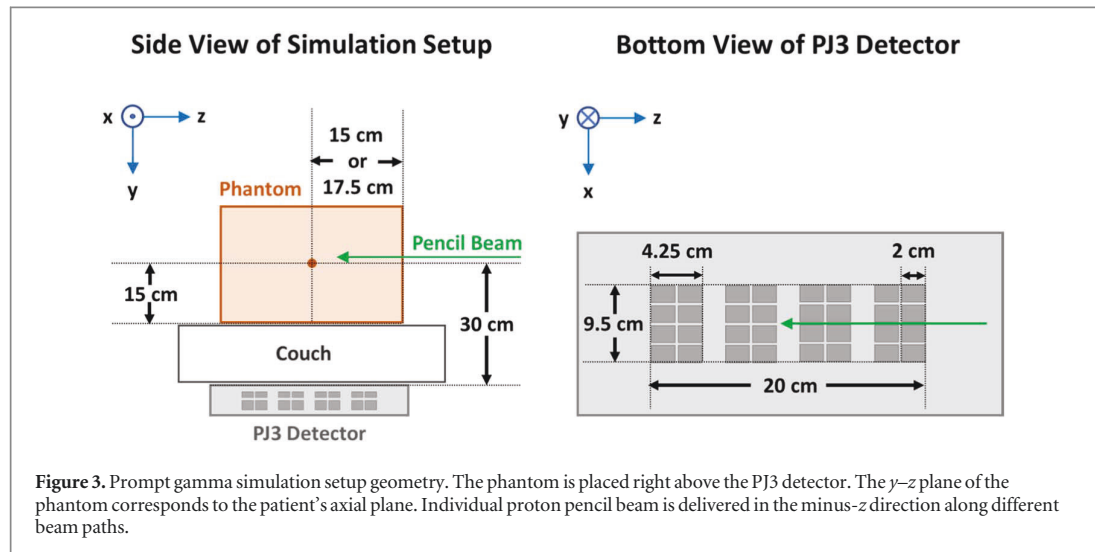
2.3.1. PG simulation and reconstruction

The proposed method aims to enhance the back-projected PG images to restore the true PG emissions. For this purpose, a dataset was built to contain image pairs consisting of the back-projected PG images and the corresponding ground truth PG images. The experiment setup geometry is shown in figure 3. A cuboid homogeneous tissue-equivalent phantom is placed on a couch, whose y - z plane corresponds to the patient’s axial plane. The dimensions of the phantom are set to 30 cm \times 30 cm \times 35 cm and 30 cm \times 30 cm \times 30 cm, in the directions of x -, y -, and z -axis, respectively. A single CC is placed under the couch to detect the PG events. The center of the phantom is right above the detector center.



In this study, we simulated 54 proton pencil beams along multiple beam paths with energies ranging from 75 to 125 MeV delivered at clinical dose rates (20 kMU min^{-1} and 180 kMU min^{-1}) and levels (1×10^9 and 3×10^8 protons per pencil beam) in a tissue-equivalent phantom. More details about this dataset can be found in appendix [appendix](#). (table A1).

The PG data were first simulated by a Monte Carlo model using the Geant4 (v10.3) toolkit (Agostinelli *et al* 2003), and were then transformed by the Monte-Carlo-plus-Detector-Effects (MCDE) model (Maggi *et al* 2020) based on the response and data acquisition characteristics of the PJ3 CC (H3D, Inc., Ann Arbor, MI). More design details are discussed in previous works (Maggi *et al* 2020, Polf *et al* 2021). PG events received by the CC were first cleaned by a previously developed neural network (Polf *et al* 2022) to remove the false events and to correct the mis-ordered events, and were then used to reconstruct PG images using the kernel weighted back projection (KWBP) algorithm (Panthi *et al* 2020) in a 3D space (volume size: $50 \text{ cm} \times 50 \text{ cm} \times 15 \text{ cm}$, grid size: $250 \times 250 \times 50$, voxel size: $2.0 \text{ mm} \times 2.0 \text{ mm} \times 3.0 \text{ mm}$). Ground truth images of the origins of the PGs emitted along the proton beam path were also recorded by the MC model. Intensities of the ground truth PG images are normalized to $[0, 100]$. They have the same dimensions and resolution as the reconstructed PG images.



2.3.2. Model training

The localization model and the enhancement model were trained separately, fulfilling individual tasks as described in section 2.2. The models were trained using the leave-one-energy-out strategy due to the limited size of the dataset. One energy was left out for testing, and the other energies were used for training and validation. Besides, the models were trained for a specific delivery dose rate and a specific dose level. For each model, PG data at the corresponding dose rate and dose level were used for training and testing.

2.3.2.1. Training configuration

In the training process, model weights were optimized by minimizing the loss between the model output and the corresponding ground truth using the 'Adam' optimizer (Kingma and Ba 2014) with a learning rate of 0.001. Batch size was set to 1 for the localization model and 8 for the enhancement model accounting for the memory consumption. Epoch number was empirically set to 1000. The best checkpoint was determined by the validation data.

To address the imbalance challenge for deep learning posed by the PG imaging, we proposed a novel loss of weighted axis-projection mean absolute error (wAP-MAE) to supervise errors in the axis-projections of volumes. The loss is calculated as follows.

$$\begin{aligned}
 &L_{wAP-MAE}(x, y) \\
 &= \lambda_1 \cdot \frac{\sum_{i=1}^I \left| \sum_{j=1}^J \sum_{k=1}^K y_{i,j,k} - \sum_{j=1}^J \sum_{k=1}^K x_{i,j,k} \right|}{I} \\
 &+ \lambda_2 \cdot \frac{\sum_{j=1}^J \left| \sum_{i=1}^I \sum_{k=1}^K y_{i,j,k} - \sum_{i=1}^I \sum_{k=1}^K x_{i,j,k} \right|}{J} \\
 &+ \lambda_3 \cdot \frac{\sum_{k=1}^K \left| \sum_{j=1}^J \sum_{i=1}^I y_{i,j,k} - \sum_{j=1}^J \sum_{i=1}^I x_{i,j,k} \right|}{K} \\
 &+ \lambda_4 \cdot \frac{\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K |y_{i,j,k} - x_{i,j,k}|}{I \cdot J \cdot K}, \tag{2}
 \end{aligned}$$

where x is the model output with dimensions of $I \times J \times K$, y is the corresponding ground truth with the same dimensions as x , λ_1 , λ_2 , λ_3 , and λ_4 are the weighting factors of the errors in the different axes and the space. In this study, λ_1 , λ_2 , λ_3 , and λ_4 were empirically set to 1, 1, 1, 1 for the localization model, and 1, 1, 1, 1000 for the enhancement model according to the value dynamic range. Compared to the conventional MAE, which supervises errors in a voxel-wise way, the proposed wAP-MAE effectively solves the imbalance issue by significantly improving the portion of aimed data in the 3D space.

2.3.2.2. Training of the localization model

The localization model was trained in the full 3D space (dimension: $250 \times 250 \times 50$) to learn to define an ROI in the back-projected PG images that contain the PG emissions along the proton pencil beam. In this step, one energy was left out for testing, and the other energies were used for training and validation.

In the training process, the full 3D back-projected PG images were fed into the localization model to reduce distortions. Weights of the model were optimized by minimizing the wAP-MAE loss between the model output and the corresponding ground truth PG images.

2.3.2.3. Training of the enhancement model

The enhancement model was trained in the full 3D space (dimension: $250 \times 250 \times 50$), with an additional input channel of the 3D ROI attention (masked region dimension: $96 \times 48 \times 8$), to learn to restore the true PG emissions. In this step, one energy was left out for testing. The training dataset was augmented by randomly masking the full PG images using an ROI (dimension: $96 \times 48 \times 8$) around the proton pencil beams. Among the training dataset, 85% of the samples were used for training, and 15% were used for validation.

In the training process, a two-channel input, consisting of an ROI-masked back-projected PG image and its unmasked copy, is fed into the enhancement model to restore the origins of PG emissions. Weights of the model were optimized by minimizing the wAP-MAE loss between the enhanced PG images and the corresponding ground truth.

2.3.3. Model testing

2.3.3.1. Experiment setup

The overall workflow is described in section 2.2. In this section, we will discuss more details of the evaluations performed in this study. PG data from proton beams of 80 MeV, 100 MeV, and 120 MeV were left out of the training dataset to independently evaluate the performance of the proposed method. The back-projected PG image was fed into the localization model to reduce distortions. The output PG image of the localization model was then normalized to its maximum value, and the signal region was determined by an empirically set threshold of 0.4. An ROI (dimension: $96 \times 48 \times 8$) was then automatically placed with its center aligned to the mass center of the signal region. The back-projected PG image was masked by the ROI, concatenated with its unmasked copy, and were then fed into the enhancement model to generate the final enhanced PG images, which were compared against the ground truth PG images for evaluation.

Note that, although the localization and the enhancement are separate models, the entire workflow is fully automatic, requiring no manual operations or parameter tuning.

2.3.3.2. Evaluation of the effect of attention

The proposed method features a two-tier architecture, which poses attention to the signal regions for more accurate enhancement. As described in section 2.2, the localization model enhances the PG images in the whole large 3D space without any attention, and the second model further enhances the PG images with an ROI attention. In this evaluation study, we compared the PG images generated by the first and second models to evaluate the effect of attention. The models were trained following the instructions in section 2.3.2.

2.3.3.3. Evaluation of the effect of loss function

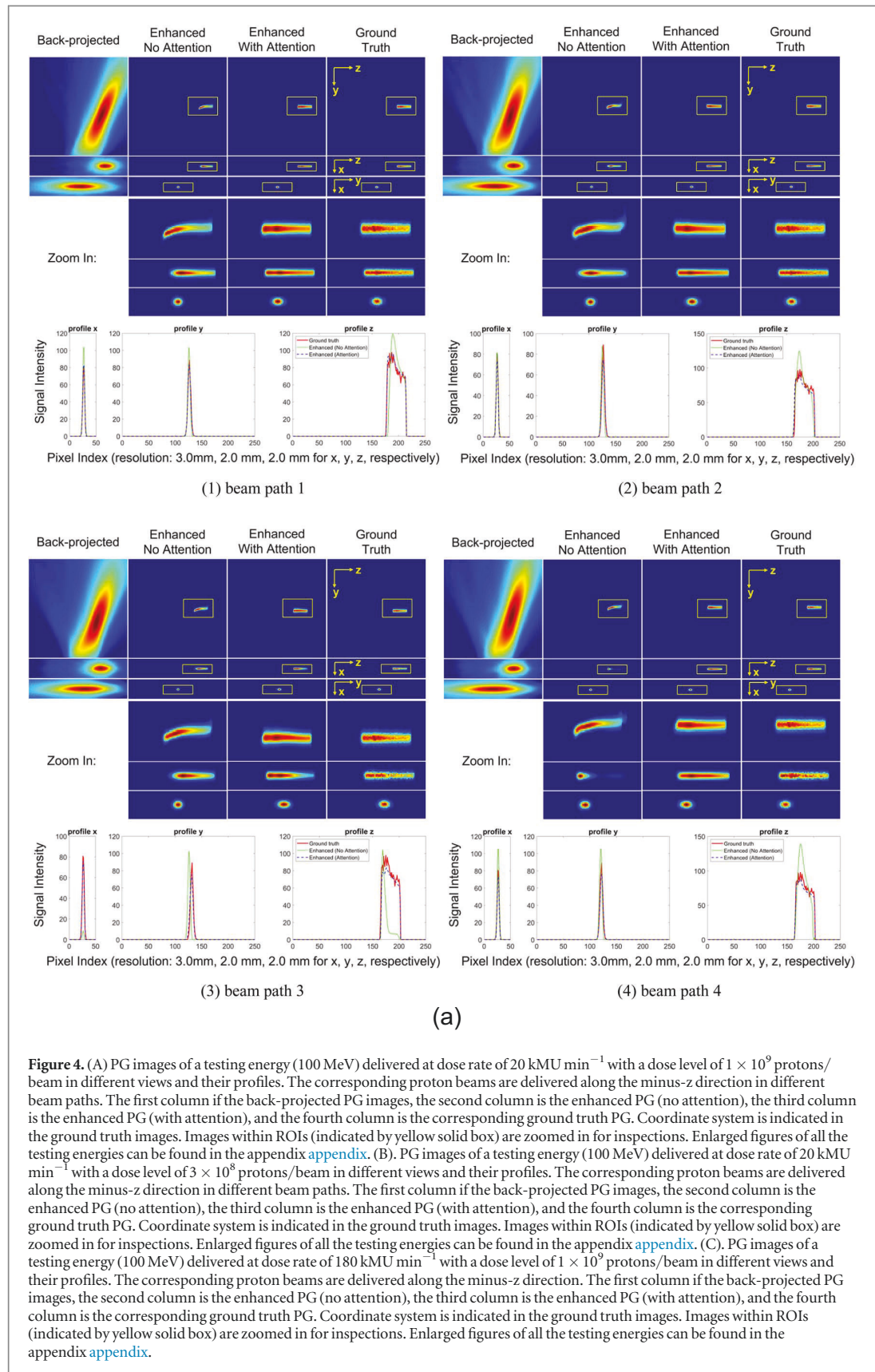
A novel loss function of wAP-MAE was proposed in this study to solve the imbalance challenge in the PG images. In this evaluation study, we trained the proposed method twice using the same datasets and the same training configurations except for the loss function: training with the classical MAE loss for the first time, and training with the proposed wAP-MAE loss for the second time. The final enhanced PG images of the models using different losses were compared to evaluate the effect of the loss functions.

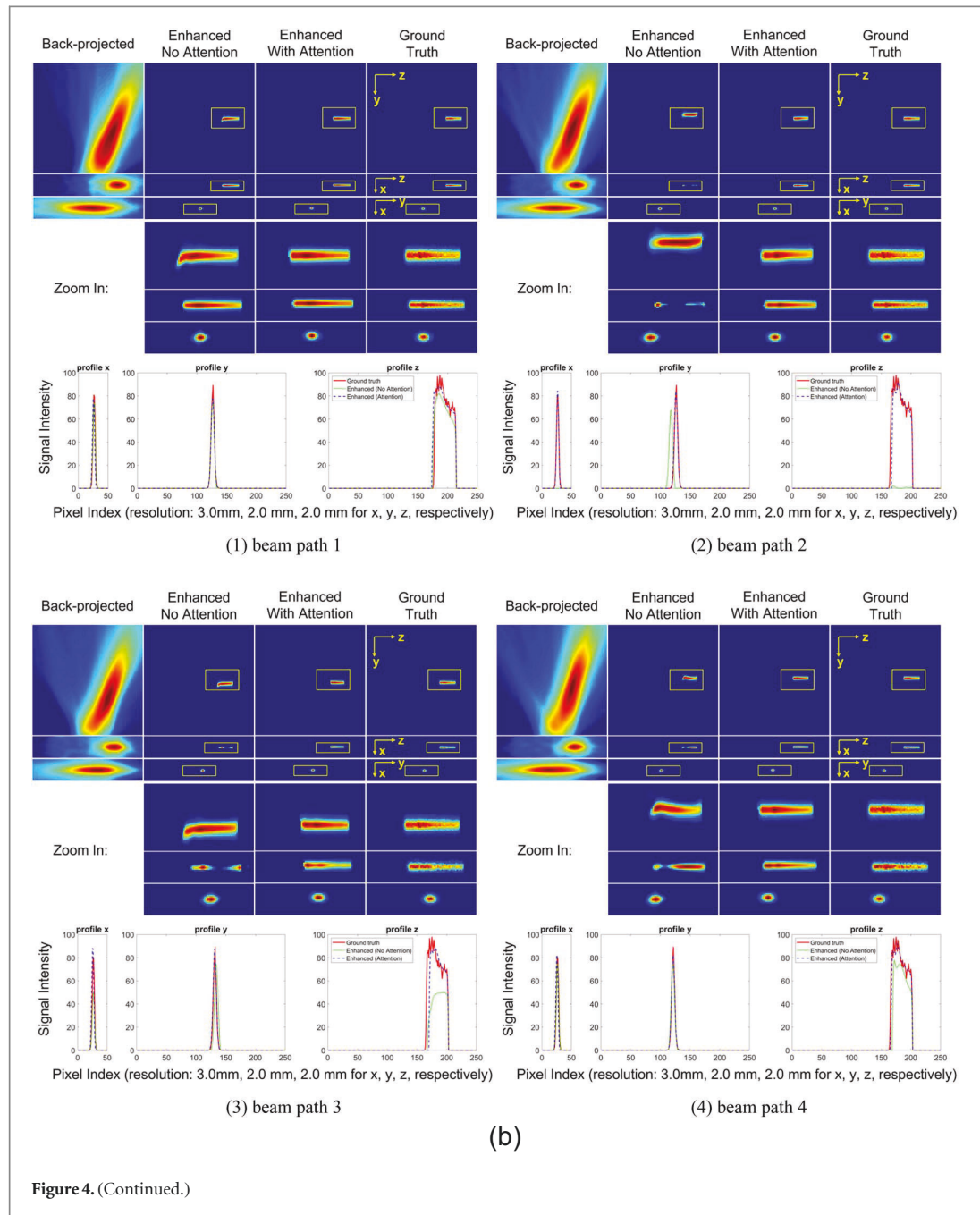
2.3.3.4. Evaluation metrics

The enhanced PG images predicted by the proposed method were evaluated both qualitatively with visual inspections and quantitatively with the PG range error (ΔR) and the PG path direction error in the 3D space, and the registered mean relative absolute errors (RMRAE).

The PG range is determined by the starting point and the ending point along the PG emission path in the 3D space. Both points are detected by a falloff of 50% regarding the maximum global intensity. The ending point is more important as it corresponds to the Bragg Peak of the proton pencil beam. ΔR is the difference between the predicted PG range and the ground truth PG range.

To calculate the RMRAE, the predicted PG image is first rigidly registered to the ground truth PG image by aligning the PG range starting points. And then, relative MAE is calculated within the union region of the predicted signals and the ground truth signals. The signal regions are determined by a threshold of 0. The RMRAE is used to evaluate the intensity errors of the predicted PG signals. With registration, the RMRAE serves as an indicator of intensity accuracy, which is independent of the range errors.





3. Results

3.1. Effect of attention on PG image enhancement

Figure 4 shows the PG images of a representative testing energy (100 MeV) in different views and their profiles along the mass center of the ground truth PG signals. No value normalization was performed in the enhanced PG images. Enlarged figures of all the testing energies can be found in appendix [appendix](#).

The back-projected PG images showed severe distortions from the CC due to the limited view, and the image quality further degraded as the dose level decreased from 1×10^9 protons per beam (figures 4(A)) to 3×10^8 protons per beam (figure 4(B)), and as the delivery dose rate increased from 20 kMU min^{-1} (figure 4(A)) to 180 kMU min^{-1} (figure 4(C)). The enhanced PG (no attention) were much less noisy compared to the back-projected PG images. However, their ranges and intensities showed obvious deviations from the ground truth. For the high dose rate of 180 kMU min^{-1} shown in figure 4(C), the range in the enhanced PG (no attention) can

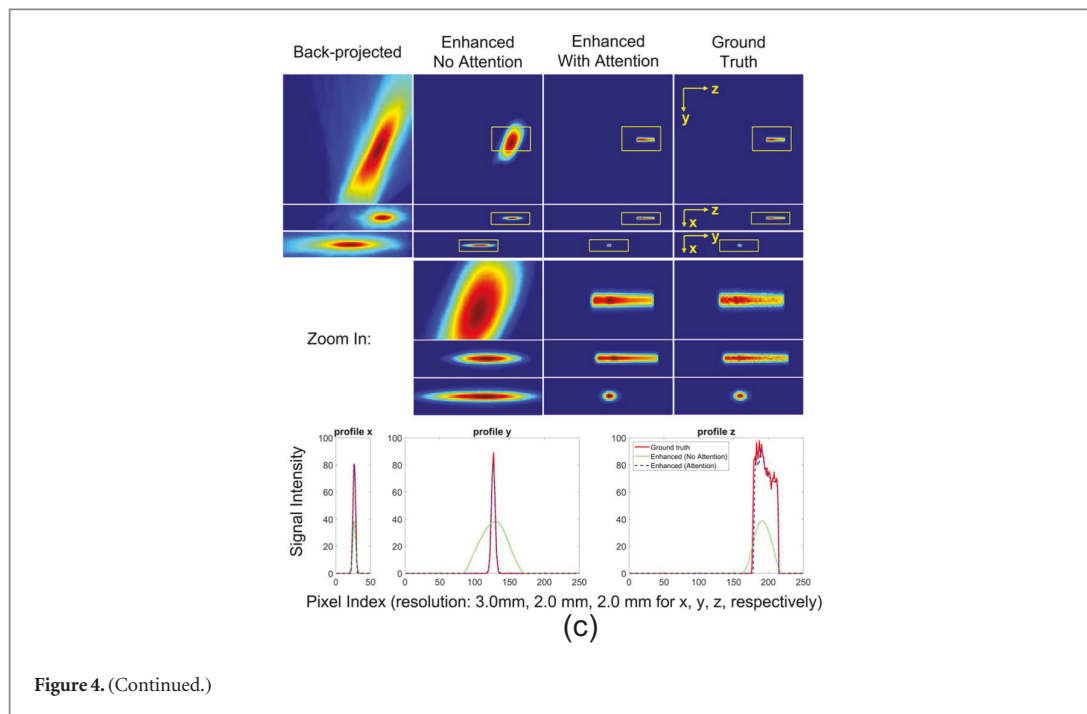


Figure 4. (Continued.)

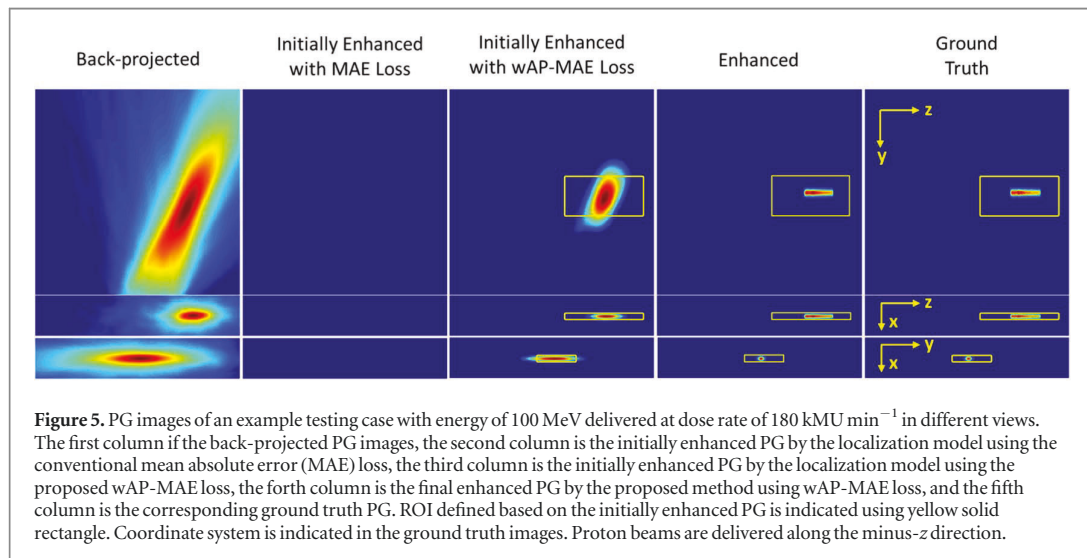


Figure 5. PG images of an example testing case with energy of 100 MeV delivered at dose rate of 180 kMU min^{-1} in different views. The first column is the back-projected PG images, the second column is the initially enhanced PG by the localization model using the conventional mean absolute error (MAE) loss, the third column is the initially enhanced PG by the localization model using the proposed wAP-MAE loss, the fourth column is the final enhanced PG by the proposed method using wAP-MAE loss, and the fifth column is the corresponding ground truth PG. ROI defined based on the initially enhanced PG is indicated using yellow solid rectangle. Coordinate system is indicated in the ground truth images. Proton beams are delivered along the minus-z direction.

hardly be determined. In contrast, the enhanced PG (with attention) demonstrated a highly clear range of the PG emissions, most of which agreed well with the corresponding ground truth.

As demonstrated by the profiles, PG emissions from the proton beams with a dose level of 1×10^9 protons/beam were precisely restored with accurate ranges and intensities with respect to the ground truth. As the dose level decreased, larger range errors and intensity errors can be observed.

3.2. Effect of loss function on PG image enhancement

Using conventional MAE loss, the two-tier scheme failed in the first localization step. Figure 5 shows an example of the testing PG data (energy: 100 MeV, dose rate: 180 kMU min^{-1}). The localization model was unable to generate any signals in the 3D space using MAE loss, yielding only background signals due to the imbalance issue. As a result, no ROI can be defined in the initially enhanced PG in the first step. In contrast, the localization model using the proposed wAP-MAE loss successfully generated an initially enhanced PG image, based on which an ROI was automatically defined following the instructions described in section 2.3.3.1 (indicated by the

Table 1. Quantitative analysis of the PG emissions restored by the proposed method.

Dose rate and dose level	Beam energy (MeV)	Range error						Intensity error RMRAE ^b
		Ending error (pixel ³)			Starting error (pixel ³)			
		<i>x</i>	<i>y</i>	<i>z</i>	<i>x</i>	<i>y</i>	<i>z</i>	
20 kMU min ⁻¹ 1 × 10 ⁹ protons/beam	80	0	0	0	0	0	0	1.1%
		1	0	1	1	0	0	1.3%
		1	0	1	1	0	0	2.1%
		0	0	2	0	0	0	2.3%
		0	0	0	0	0	0	0.9%
		0	0	1	0	0	0	1.3%
	100	0	0	0	0	0	0	1.5%
		0	0	0	0	0	0	1.1%
		0	0	0	0	0	0	1.1%
		1	0	0	1	0	0	0.7%
		0	0	2	0	0	0	1.1%
		0	2	0	0	0	0	1.1%
20 kMU min ⁻¹ 3 × 10 ⁸ protons/beam	80	0	1	0	0	0	0	1.0%
		0	0	1	0	0	0	1.5%
		0	1	1	0	1	0	1.5%
		1	0	0	1	0	0	1.4%
		1	5	2	1	5	0	1.7%
		1	0	3	1	0	0	1.4%
	100	0	0	3	0	1	0	2.2%
		1	0	5	1	0	0	2.1%
		1	0	1	1	0	0	1.2%
		0	0	3	0	0	0	1.3%
		0	0	2	0	0	0	1.1%
		0	0	4	0	0	0	1.6%
180 kMU min ⁻¹ 1 × 10 ⁹ protons/beam	80	0	5	1	0	5	0	1.0%
		0	0	1	0	0	0	1.3%
		0	0	1	0	0	0	1.2%
	100	0	0	1	0	0	0	1.2%
		0	0	1	0	0	0	1.2%
		0	0	3	0	0	0	1.4%

^a Pixel size: 3.0 mm, 2.0 mm, and 2.0 mm for *x*, *y*, and *z* directions, respectively. The *z*-axis is along the horizontal beam path direction, the *y*-axis is vertical to the CC, and the *x*-axis is orthogonal to the *y*-axis and the *z*-axis. The coordinate system can be referred to in figure 3.

^b Registered mean relative absolute error (RMRAE). This metric is independent of the range accuracy. It evaluates the intensity accuracy ignoring the range errors via registration. More details are referred to in section 2.3.3.2.

solid yellow rectangle in figure 5). Results showed that the ROI was correctly defined covering the PG emission regions, which helped further improve the quality and accuracy of PG image enhancement.

3.3. Quantitative analysis of PG image enhancement

Table 1 shows the quantitative analysis of the PG images of testing energies enhanced by the proposed method. For the PG data from high dose level beams (1 × 10⁹ protons/beam), the range errors were within 2 pixels except for only one case from the 120 MeV proton beam delivered at 180 kMU min⁻¹ (hereafter referred to as the 120 MeV@180 kMU min⁻¹ PG). As the dose level decreased to 3 × 10⁸ protons/beam, larger range errors could be observed in the enhanced PG. But all the testing cases demonstrated accurate intensities regarding the ground truth, and the RMRAE is within 3%. The quantitative analysis agreed with the qualitative inspections.

3.4. Runtime

The deep learning models in the proposed method were implemented using the Keras framework with the TensorFlow backend (v2.3.0). Training and testing of the models were conducted on a computer equipped with a GPU of NVIDIA Titan RTX with 24 GB memory and a CPU of Intel Xeon with 32 GB memory. The proposed PG image enhancement was fully automatic, which takes about 0.26 s.

4. Discussion

Range uncertainty is a major concern in proton therapy, significantly limiting its delivery precision. The CC-based PG imaging has been developed as a promising tool to provide *in vivo* 3D range verification. But conventional back-projected PG images suffer from severe distortions due to the limited acquisition view of the

CC. Proton ranges can hardly be distinguished in the distorted PG images, severely limiting its clinical utility. To address this problem, in this study, we proposed a novel two-tier deep learning-based method to restore the 3D PG emissions images acquired with a single CC.

The limited-view measurement of CC poses severe illness in the PG image reconstruction. A common way to address the ill-conditioning is to assume prior knowledge of the underlying reconstruction for additional constraints. One category of the algorithms is based on the compressed sensing theory (Donoho 2006). These methods are effective in reducing image noises and artifacts by assuming the underlying images to have certain features, such as sparsity (Lustig *et al* 2008, Sidky and Pan 2008). But their capabilities in correcting limited-view distortions are limited. Another category is based on deep learning. They assume that there is a common restoring pattern from corrupted images to their high-quality counterparts, and this pattern can be learned from a sample group in a data-driven approach. Previous study Jiang *et al* (2021) demonstrated the effectiveness of deep learning in correcting distortions and restoring volumetric structures from limited-angle CBCT projections. Therefore, we employed deep learning models in the proposed method to restore 3D PG emissions from distorted back-projected images.

As mentioned in the introduction, the PG emissions along a proton pencil beam take up an extremely low portion of the large 3D space, presenting an attention challenge to the deep learning-based method. Generally, attention in deep learning is a technique aiming to mimic cognitive attention. It directs the model to devote more focus to some parts of the inputs, which are believed to be more relevant to the aimed results. Weighting factors of attention can be learned in a data-driven method. Based on this idea, we proposed a two-tier scheme in this study to address this attention challenge. First, a localization model was trained to define a small ROI containing the true PG emissions in the large 3D space. Then, the ROI was used as the attention region, focusing on which, the following enhancement model was trained to restore the PG signals. This scheme enables the enhancement model to have additional attention on the PG signal regions, thus generating more accurate PG images, as demonstrated in our results. Note that the first localization step aims to roughly estimate the PG emissions and define an ROI containing the signals; thus, it does not require a very clean and precise restoration of the PG images. As a result, a relatively small training dataset can serve the training purpose. In the enhancement model training process, random cropping was used to considerably boost the training samples, which accounts for various signal positions within ROIs. This technique enables the enhancement model to deal with the signals deviated from the ROI centers due to the initial restoration errors in the first step, further improving the robustness of the entire workflow. The effectiveness of the proposed two-tier scheme has been demonstrated by the results in section 3.1.

Besides the attention challenge, the severe imbalance between the true PG signals and the background poses additional difficulties for the training of the deep learning models. It has long been a challenge for deep learning to deal with small objects (Wang *et al* 2019). The key is to reduce both the *misses* (in this case, predicting PG signal as background) and the *false alarms* (predicting background as PG signals). Commonly, data-driven learners tend to exhibit bias towards the majority group, and in some extreme cases, ignore the minority group altogether (Johnson and Khoshgoftaar 2019b). In this study, the minority group, the true PG signals, takes up a significantly lower portion than the majority group, the background. Using conventional pixel-wise error (such as MAE) as a loss function, the models ignored the PG signals altogether and predicted the background values for the entire 3D space. Such results agree with the analysis. The training process of the deep learning models is to minimize the loss between the model prediction and the ground truth. Predicting PG signals in the background region can contribute to the *false alarms*, while predicting background in the signal region will contribute to the *misses*. Given the dominant majority of the background, predicting the background altogether is a much easier point for the models to converge to than finding a complicated distortion correction pattern to predict correct PG signals. To address this imbalance challenge, we proposed a novel loss of wAP-MAE. The imbalance between the PG signals and the background is largely solved by projecting volumes onto different axes. A weighted sum of the axis-projection losses and the volume loss can further help improve the spatial precision and the pixel-wise intensity accuracy of the PG signals. The effectiveness of the proposed wAP-MAE loss has been demonstrated by the results in section 3.2.

Overall, the proposed two-tier deep learning scheme with the wAP-MAE loss is effective in restoring PG signals in a large 3D space. The proposed method significantly improved the PG image quality compared to the original images reconstructed by the widely used back-projection algorithm. The signal ranges were very clearly demonstrated in the enhanced PG images, which agreed well with the ground truth PG images. In terms of intensity accuracy, the proposed method yielded a relative MAE within the signal regions of < 3.0% for all the testing cases. And for the PG ranges, all the testing cases with high dose level (1×10^9 protons per beam), except for the 120 MeV@180 kMU min⁻¹ PG case, showed offsets of within 2 pixels in all directions (majority within 1 pixel). When it came to the lower dose level (3×10^8 protons per beam), the testing cases showed an overall poorer range accuracy than the high dose-level cases.

The main reason for the range offset in the 120 MeV@180 kMU min⁻¹ PG case is the increased noises in the input back-projected PG images, which is caused by the decreasing portion of the primary PGs and the increasing mis-ordered and false PG events as the dose rate considerably increases (Panthi *et al* 2020). Although the PG events were preprocessed by a neural network (Polf *et al* 2022) for event cleaning, the back-projected PG image is still noisy due to the residual errors of the preprocessing. As for the cases of lower dose levels, fewer PG events can be detected for reconstruction. Using the back-projected reconstruction, the residual bad events and the scattered PG events are more likely to generate various background artifacts in the PG images without sufficient averaging. As a result, the input back-projected PG images have poorer quality. It is an inherent limitation of the image-enhancing methods that their performance can be compromised by the degradation of the input image quality. Such a phenomenon has also been demonstrated in our previous studies (Jiang *et al* 2019, 2021, 2022).

In this preliminary study, we demonstrated the power of deep learning-based image enhancement to restore the 3D PG images from a single view CC acquisition. Compared to previous studies, the PG images were significantly improved after enhancement in our study, providing valuable information for range verification in proton therapy. Therefore, this pilot study demonstrated the feasibility and potential of PG imaging for 3D *in vivo* proton dose verification, which is urgently needed in proton therapy. However, there are some limitations in this study. Firstly, a relatively small dataset was used in this preliminary study. More PG data from proton pencil beams with more energies and various beam paths are warranted in future studies. Secondly, in the current stage, the proposed method is prone to errors when the input PG image is highly contaminated by noises, such as cases at low dose levels or high-energy cases at high dose rates. We would like to point out that, it is not a fundamental limitation of the proposed method. As a postprocessing model, it mainly aims to correct the severe distortion artifacts due to the limited view of a single CC, which can hardly be addressed in the preprocessing or reconstruction steps. In future studies, more advanced preprocessing techniques can be used to clean the detected PG events by removing the false events, correcting the mis-ordered events, and reducing the scattered events. Besides, deep learning algorithms can be developed to reconstruct PG images directly from PG events to reduce noise. In addition, developments in the hardware, e.g., new detectors with higher detection efficiency, can also help to improve the quality and the amount of the detected PG events, and thus resulting in higher-quality PG images. Thirdly, only simulation studies were performed in this work. Experiments on patient data are warranted in the future to further evaluate the clinical utility of the proposed method. Fourthly, the proposed method only generates PG signals without dose deposition information. Conversions from PG to dose can be studied in future works.

5. Conclusion

A preliminary study presented in this work demonstrated the feasibility of the proposed deep learning-based method to generate high-quality 3D PG images acquired from a single CC in nearly real-time, which substantially improves the clinical utility of PG imaging for range verification in proton therapy. Evaluations on more PG cases from proton beams with higher energies, lower dose levels, and different beam directions, as well as in-patient studies, are warranted in future studies.

Acknowledgments

This work was supported by the National Institutes of Health under Grant Nos. R01-EB028324, R01-EB032680, R01-CA187416, and R01CA279013. The prompt gamma data used in this work was generated using the UMBC High Performance Computing Facility (HPCF). The facility is supported by the U.S. National Science Foundation through the MRI program (Grant Nos. CNS-0821258, CNS-1228778, OAC-1726023, and CNS-1920079) and the SCREMS program (Grant No. DMS-0821311), with additional substantial support from the University of Maryland, Baltimore County (UMBC). See hpcf.umbc.edu for more information on HPCF and the projects using its resources

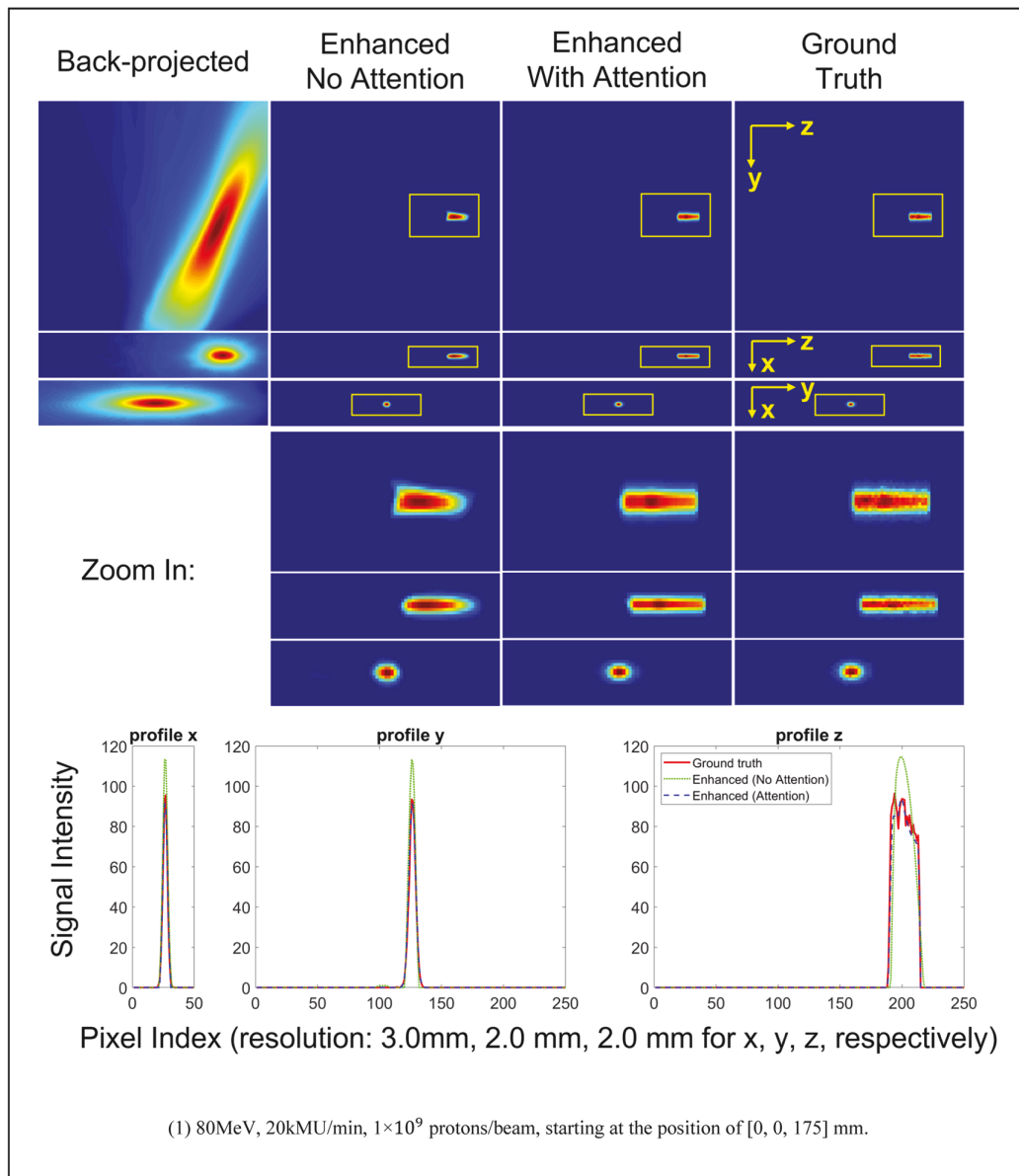
Appendix

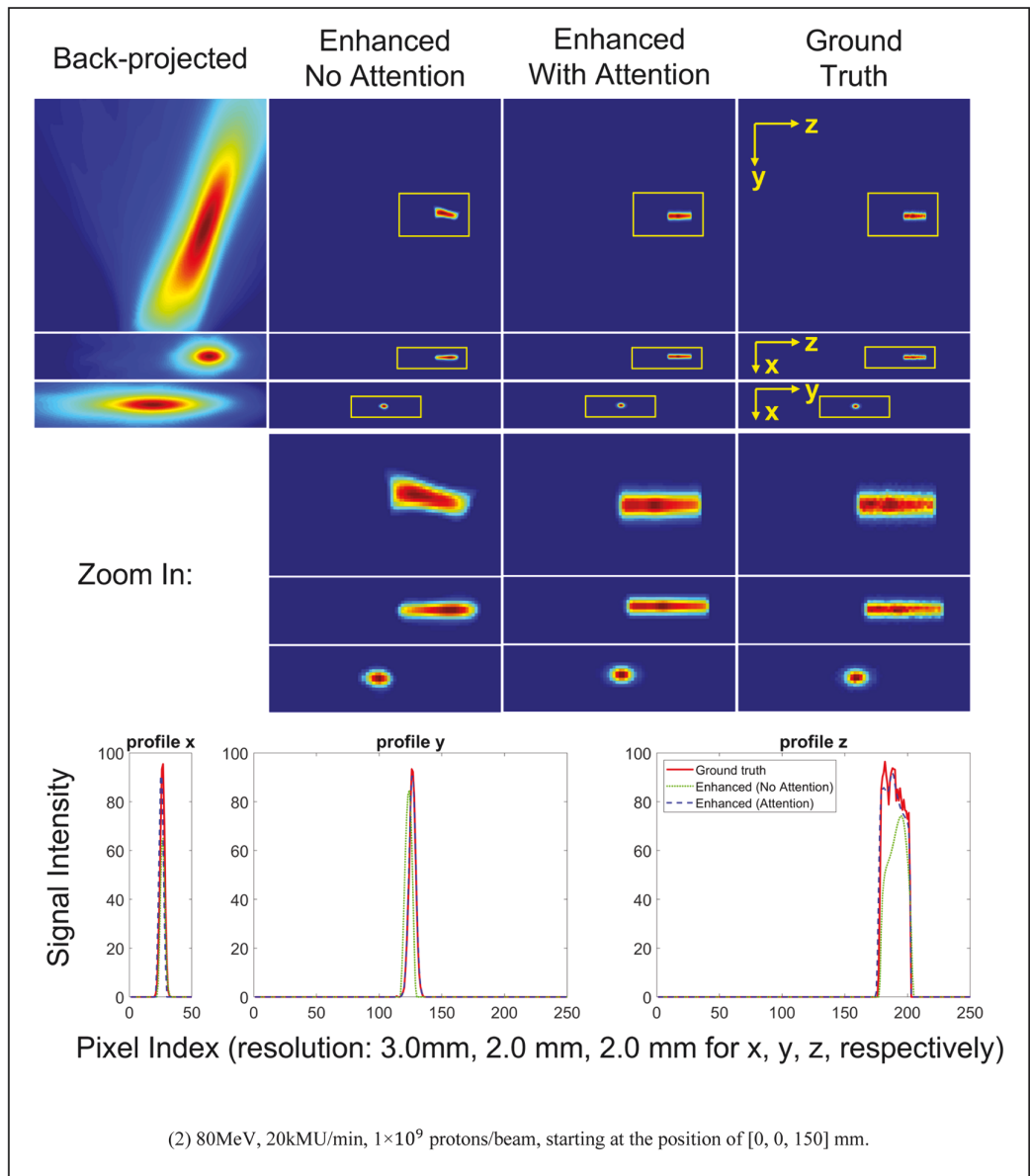
A1. Proton pencil beams simulated in this study

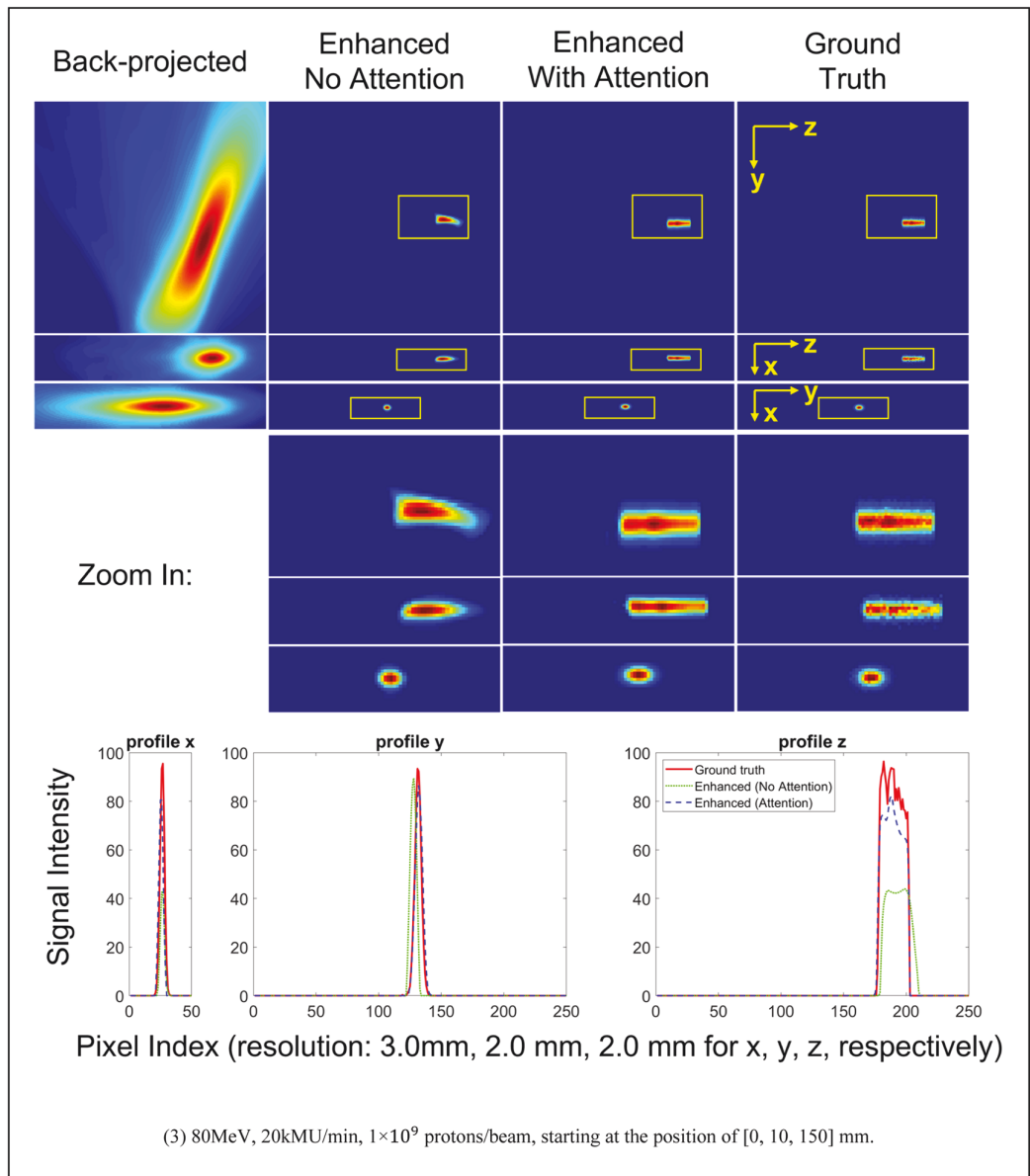
A2. Figures of all the testing PG cases

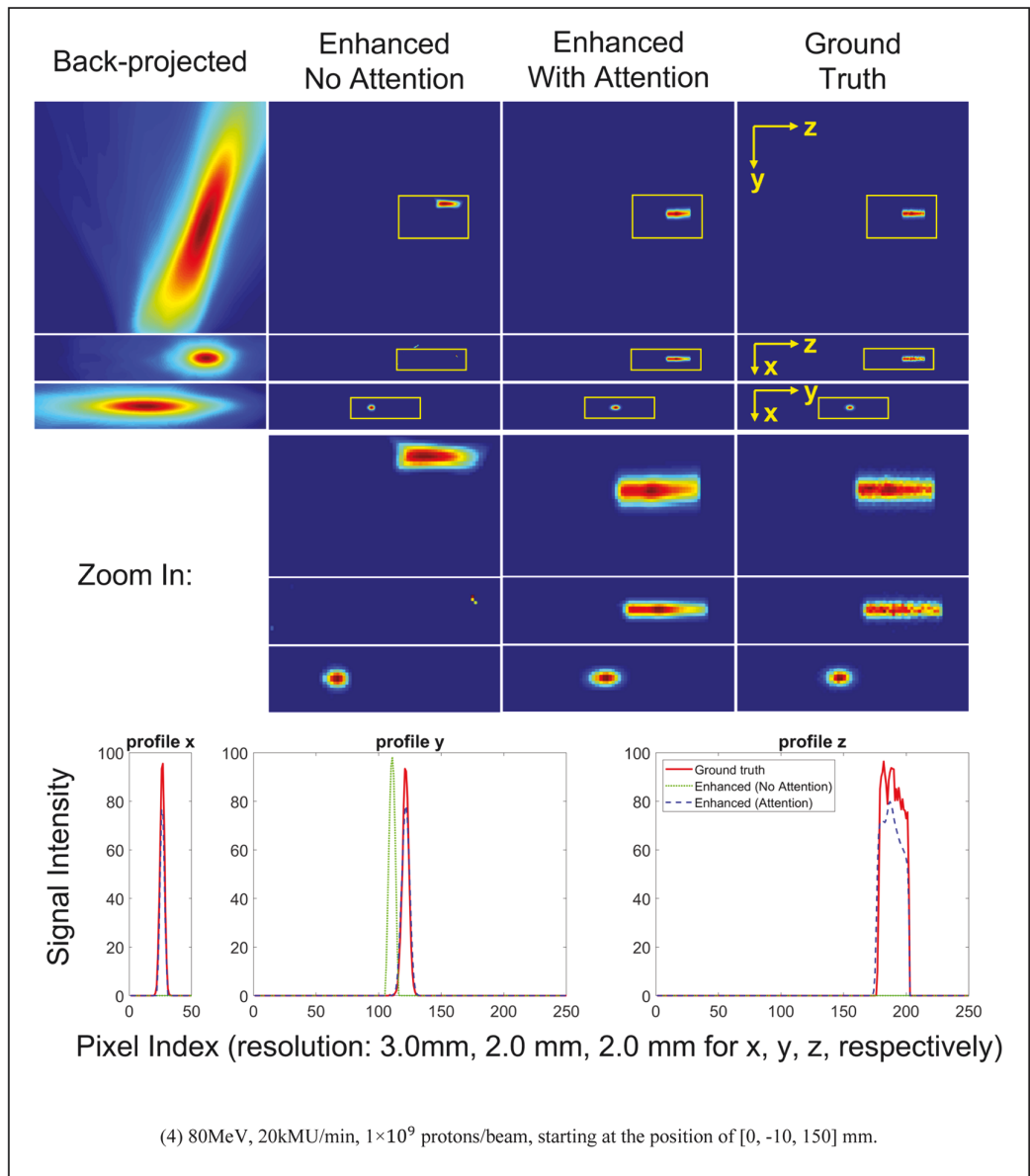
The following figures show the PG images of all the testing energy delivered at various dose rates and dose levels in different slice views and their profiles. The corresponding proton beams are delivered along the minus-z

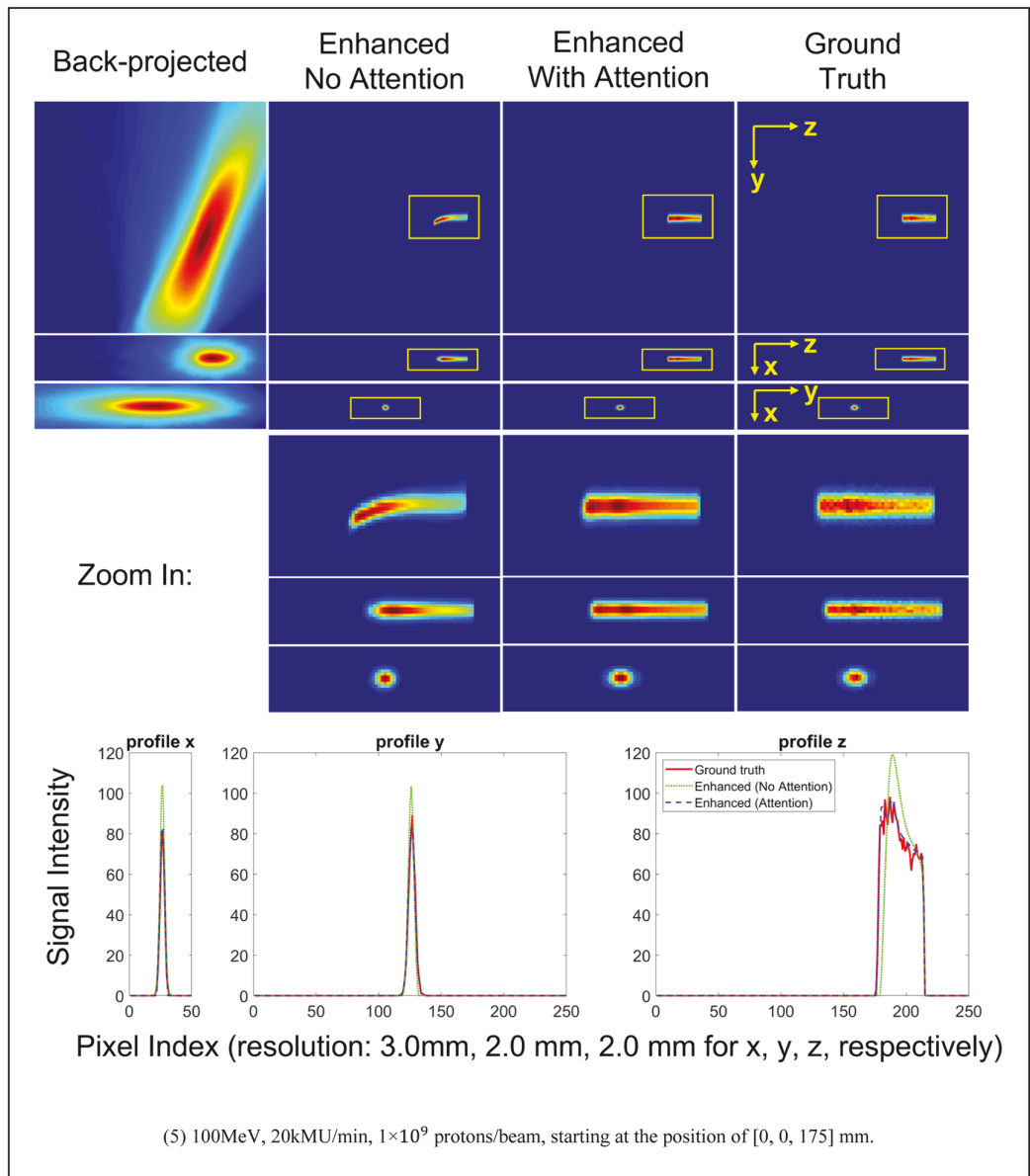
direction in multiple beam paths. The first column is the back-projected PG images, the second column is the enhanced PG (no attention), the third column is the enhanced PG (with attention), and the fourth column is the corresponding ground truth PG. Coordinate system is indicated in the ground truth images. Images within ROIs (indicated by yellow solid box) are zoomed in for inspections.

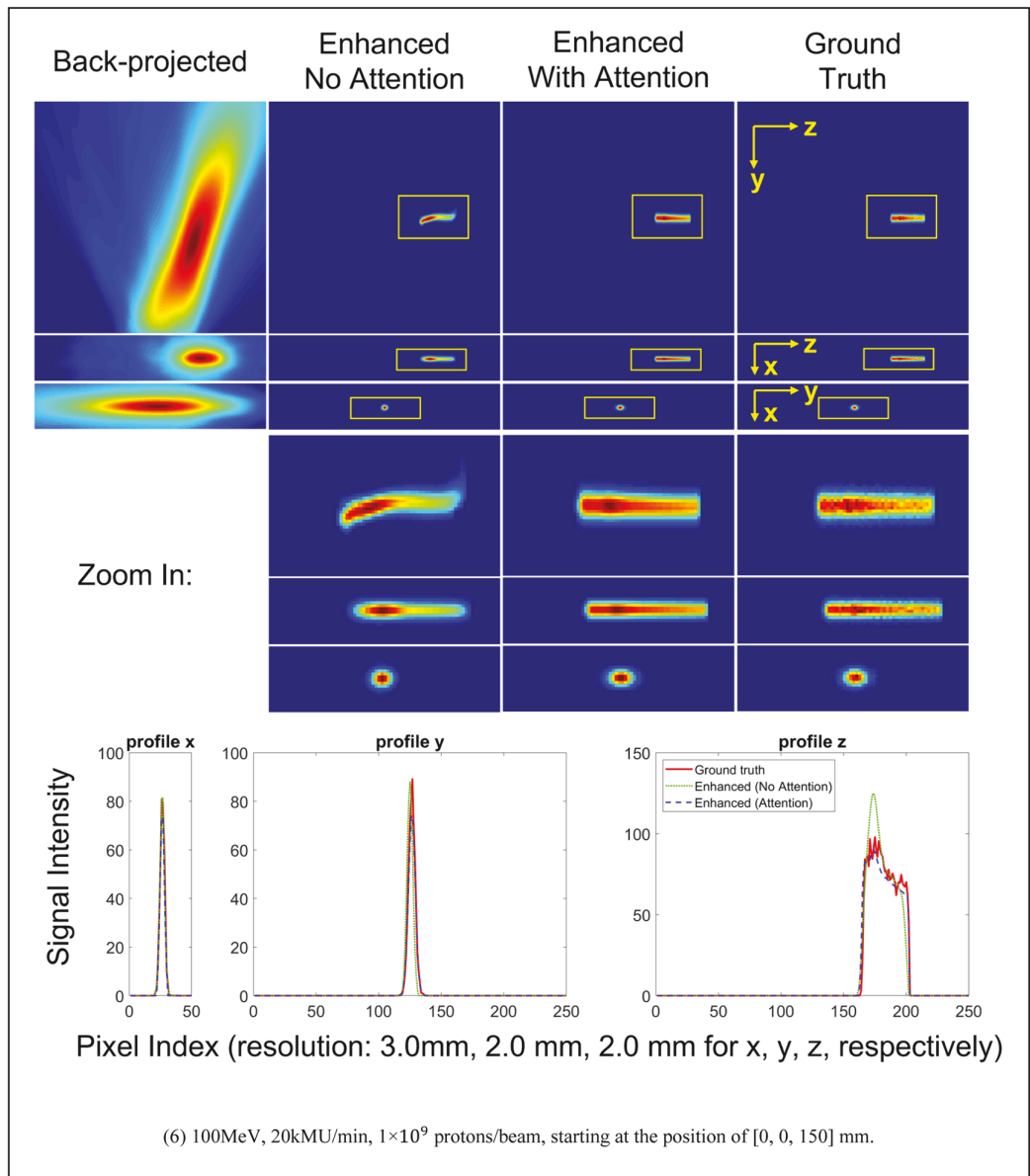


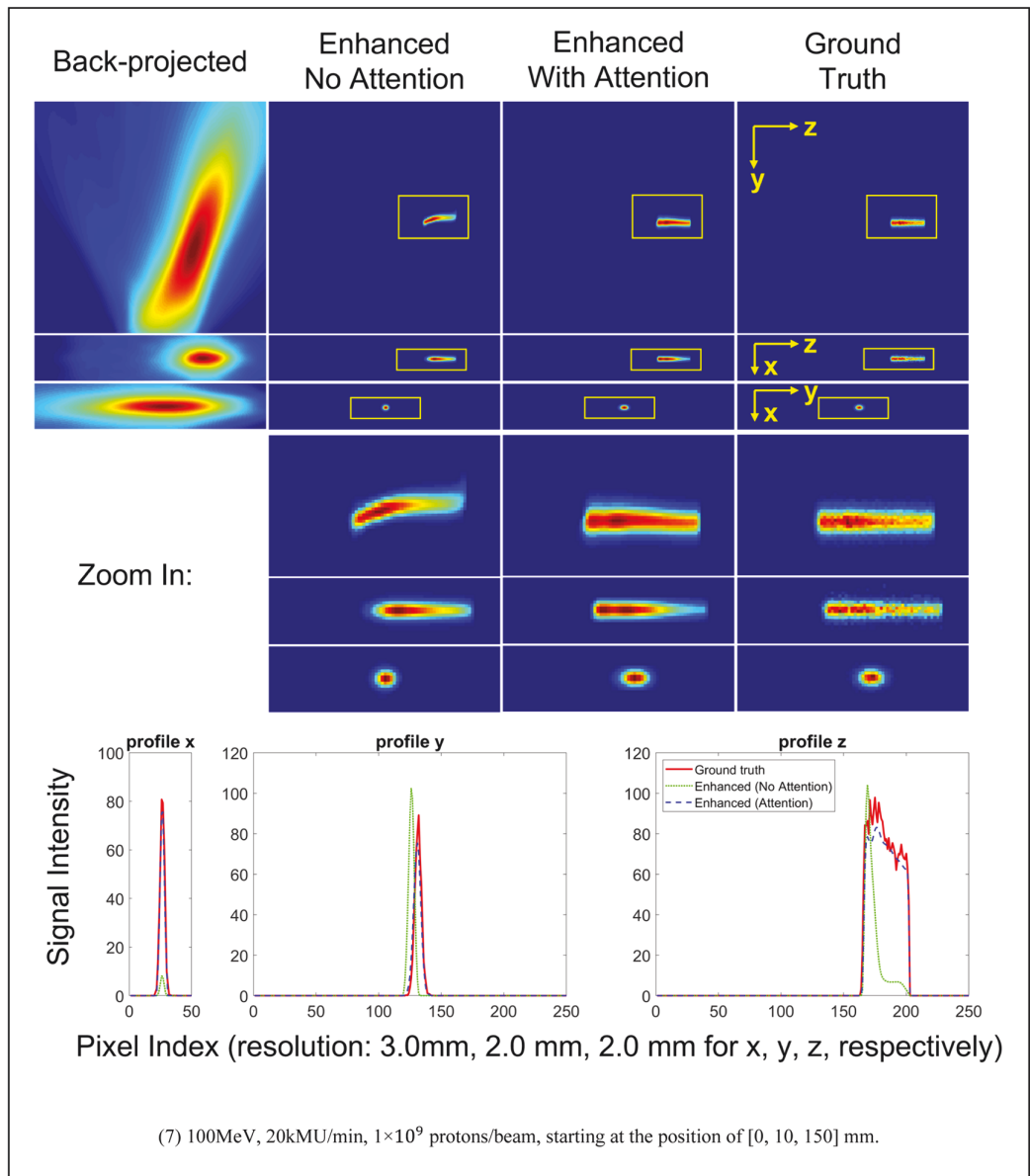


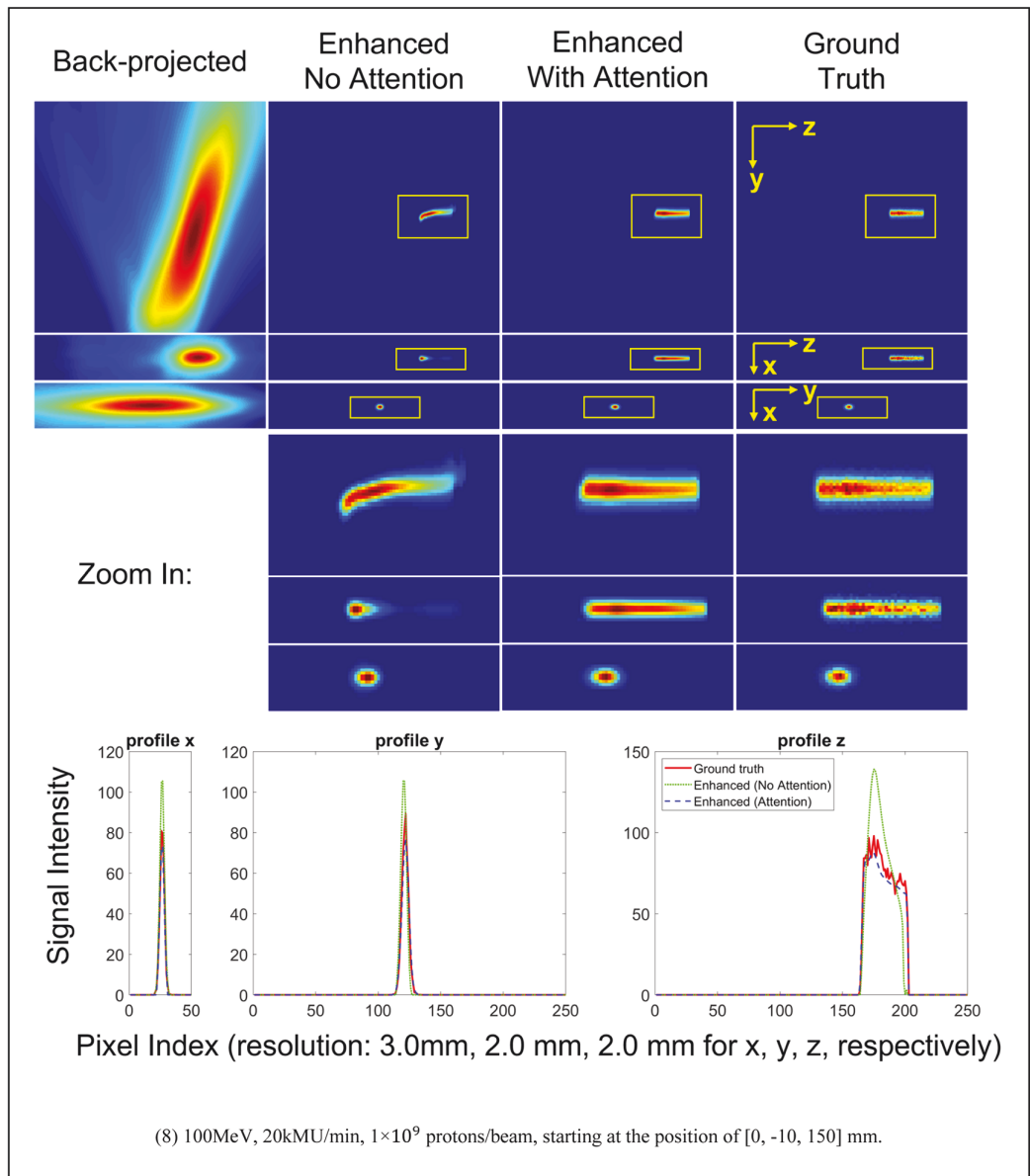


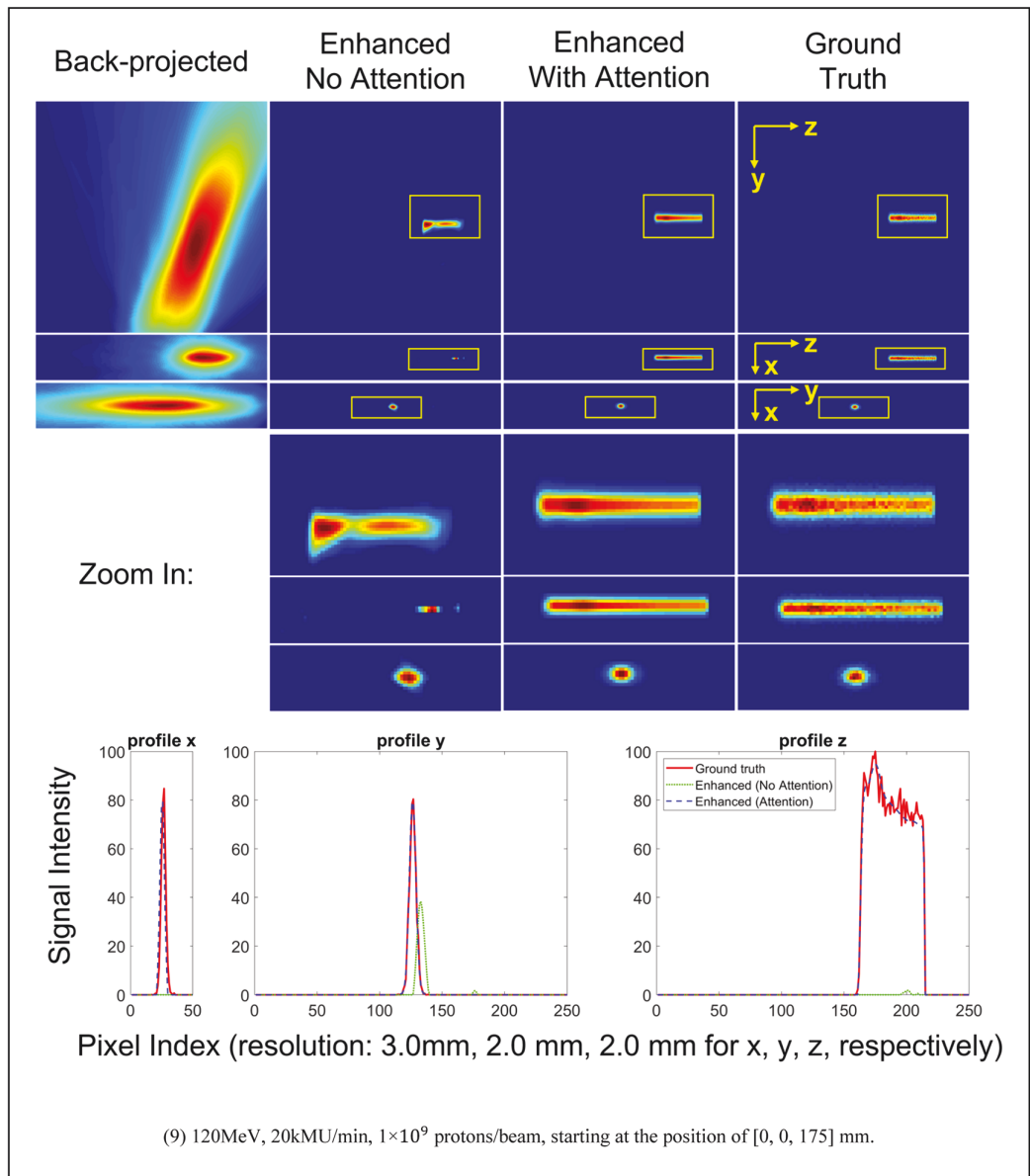


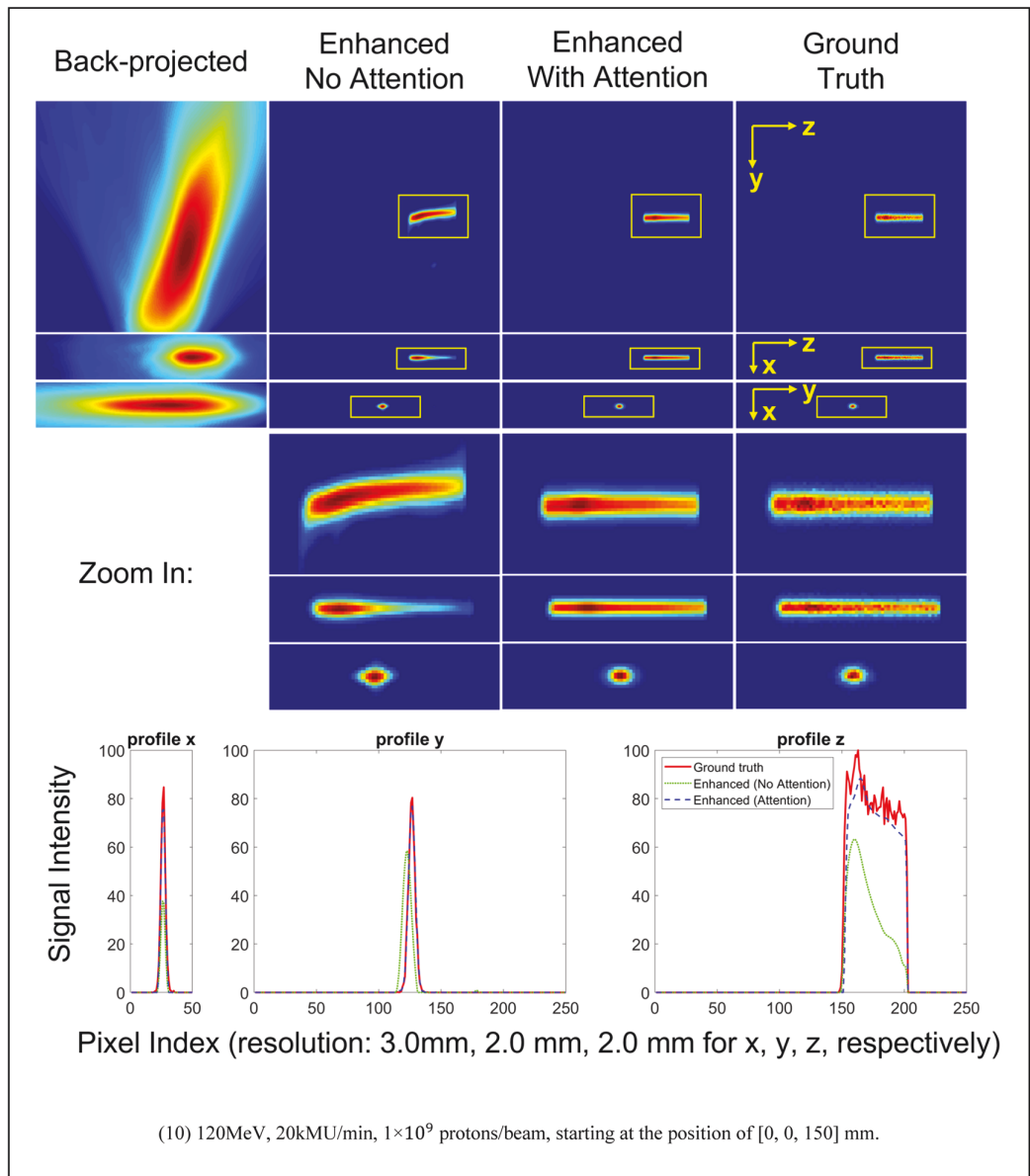


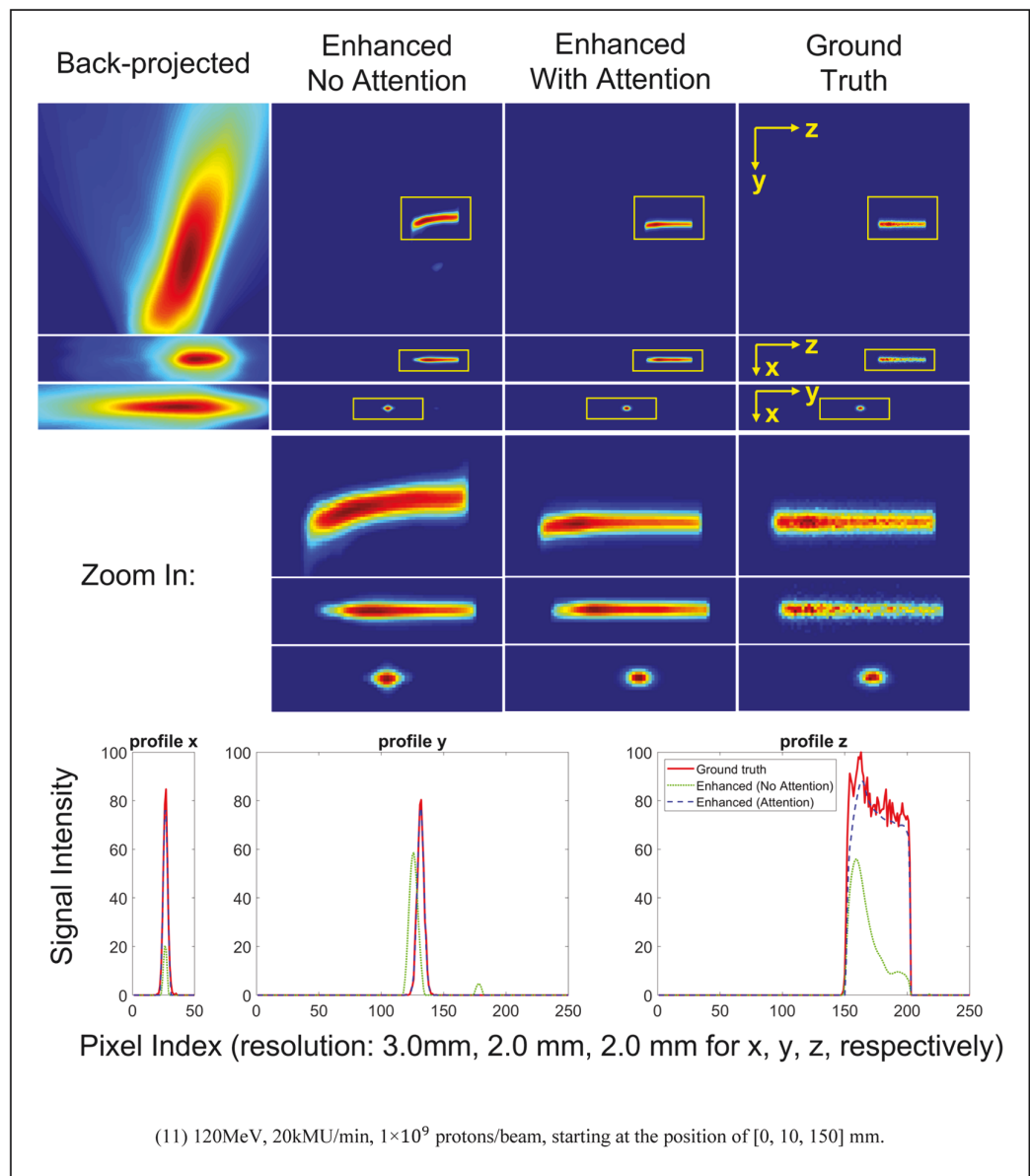


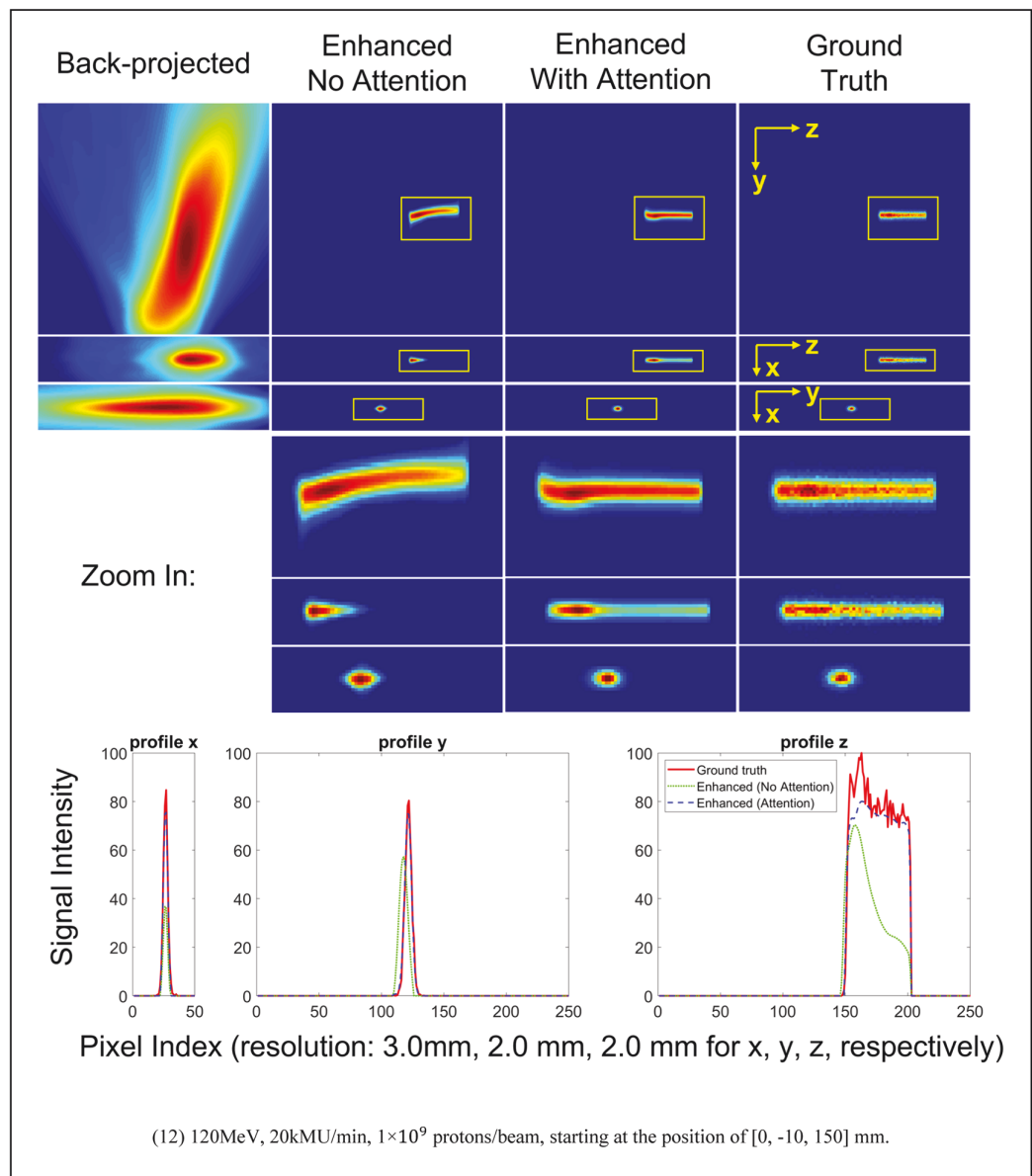


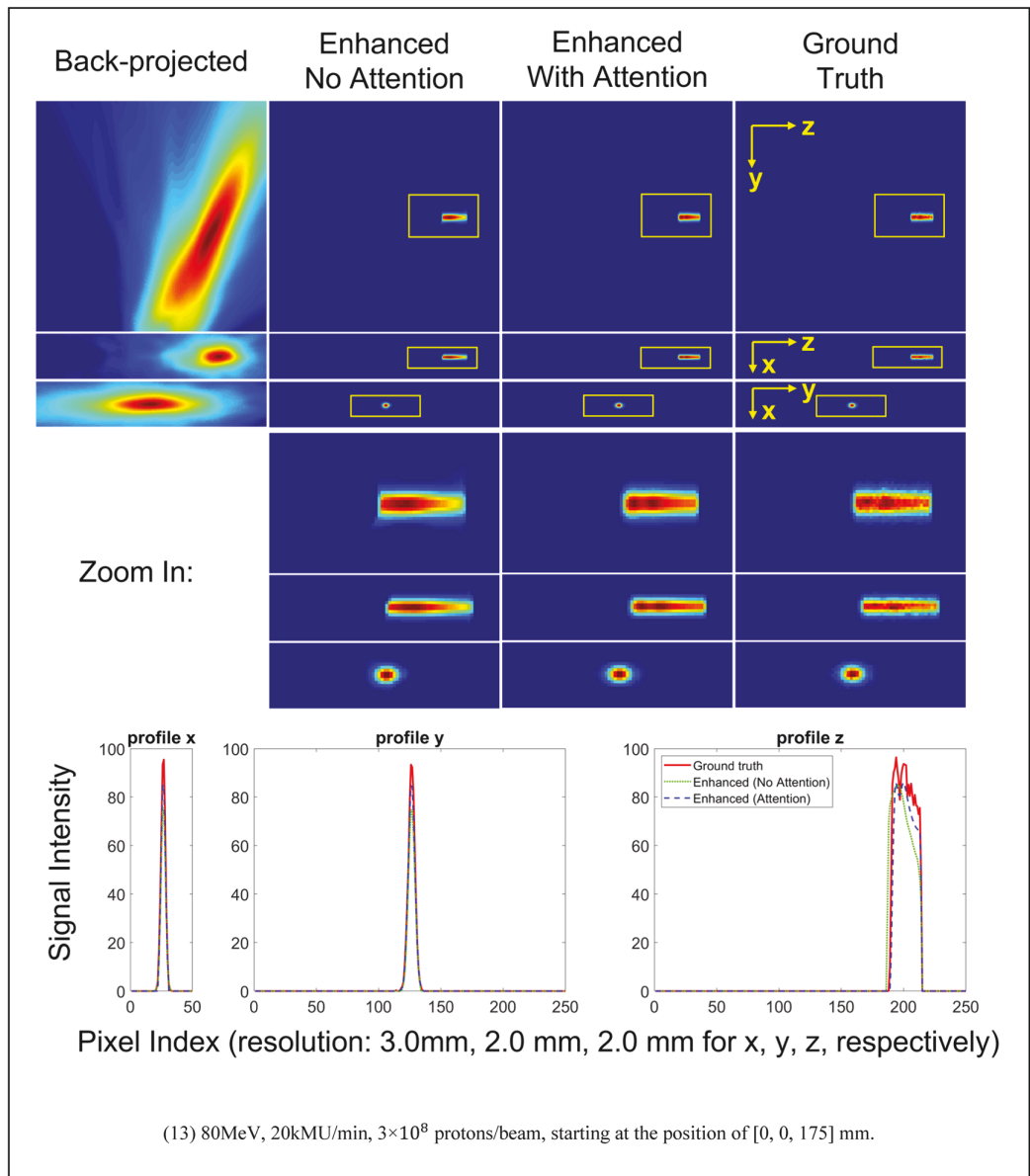


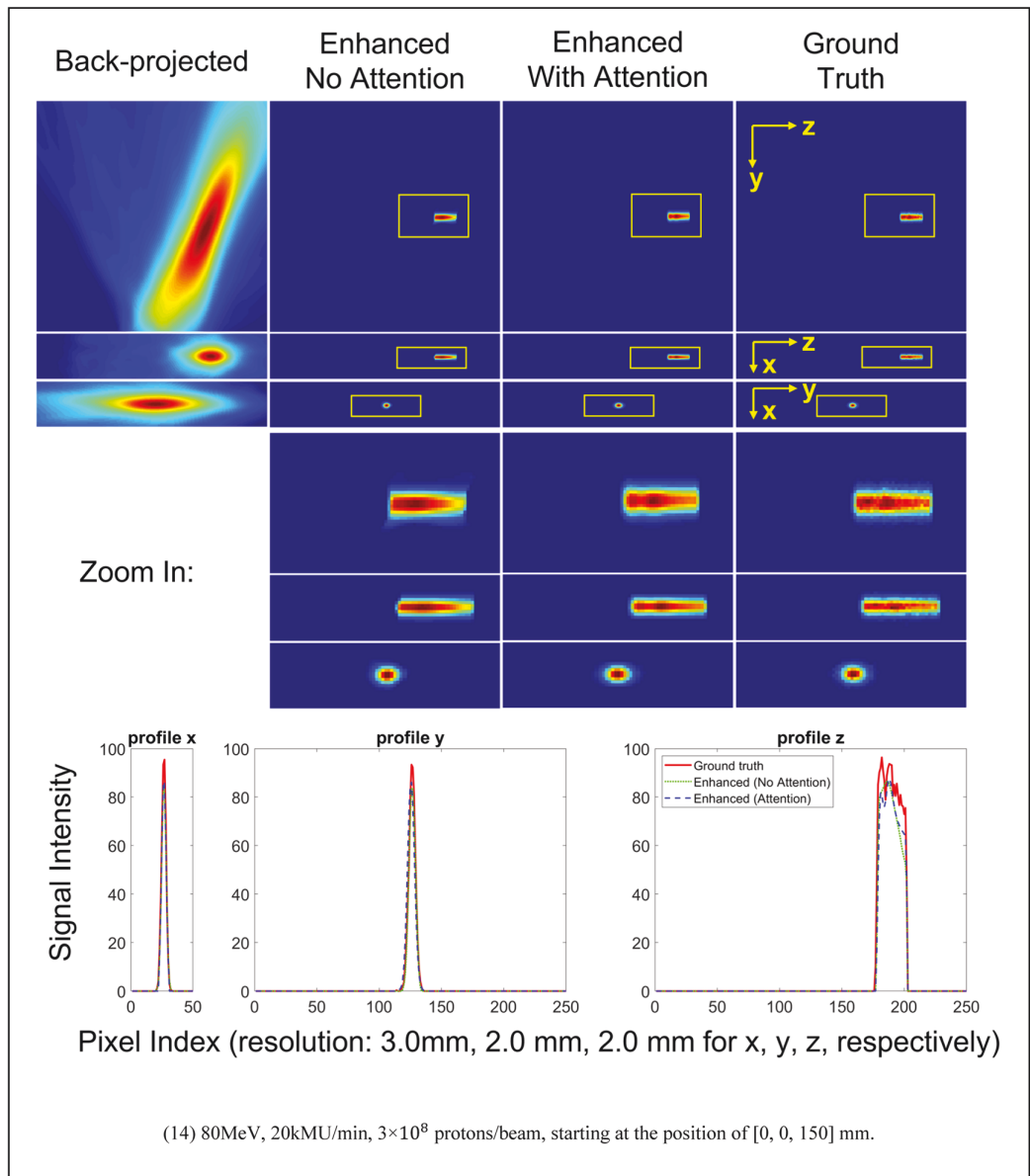


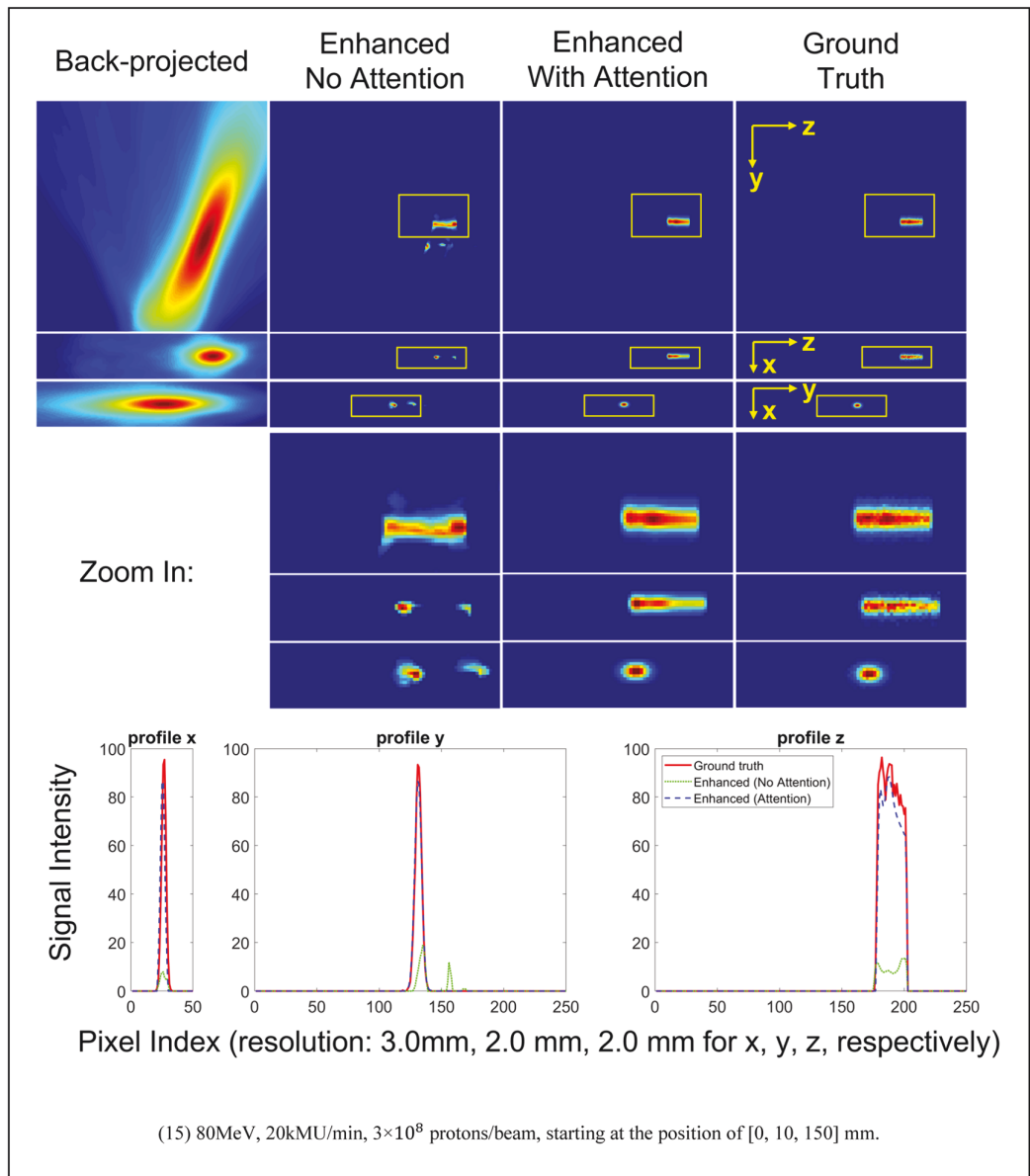


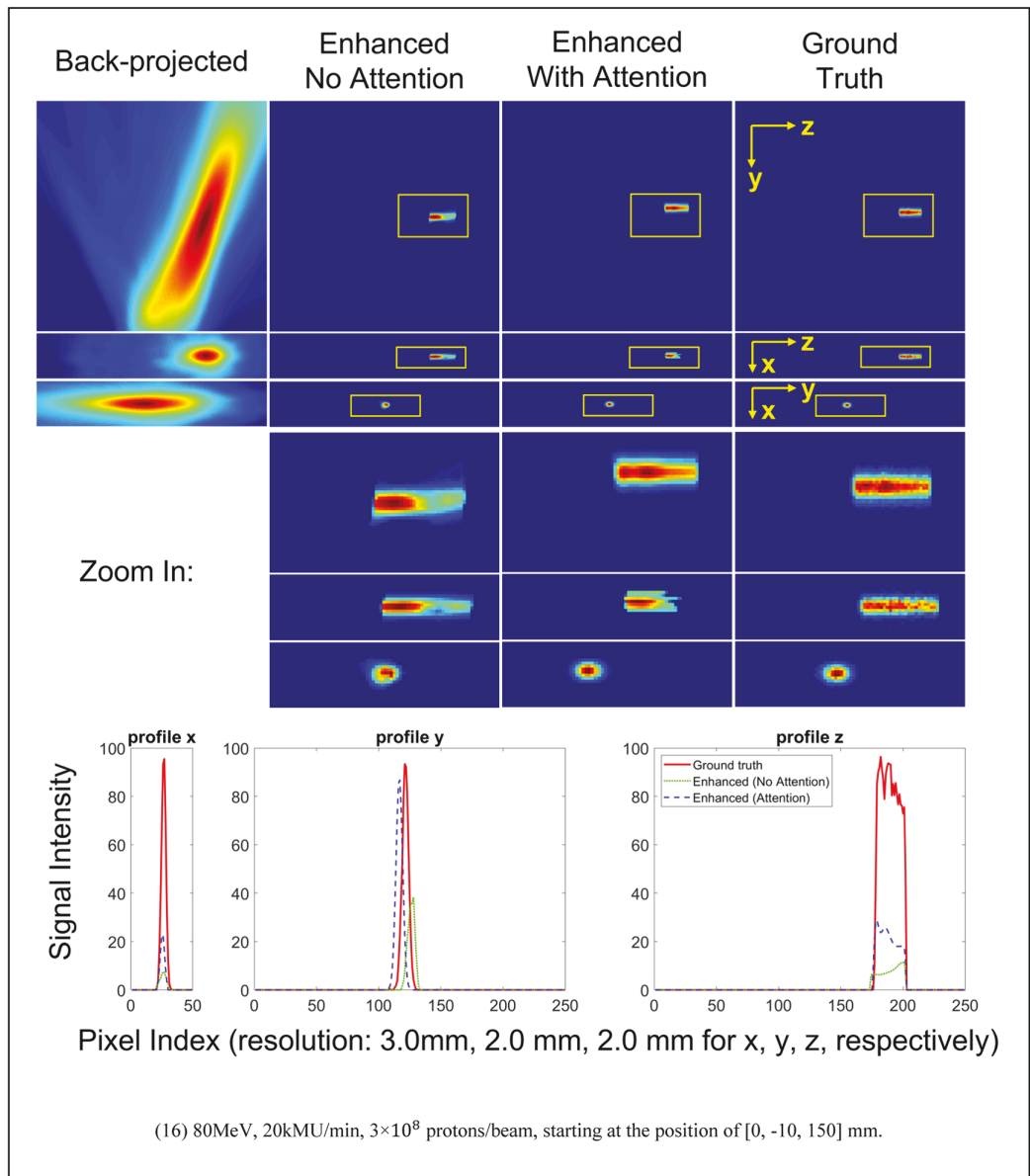


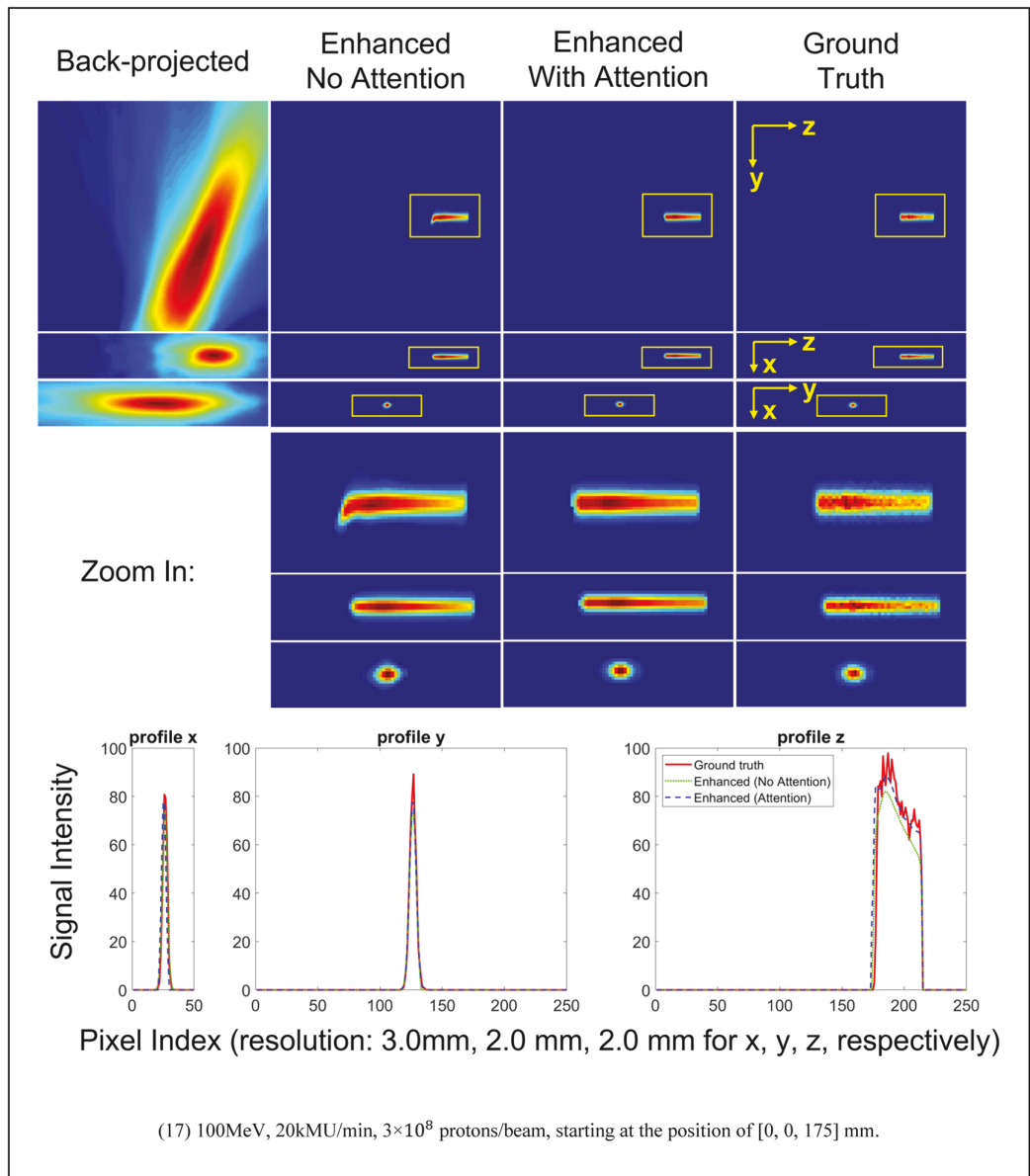


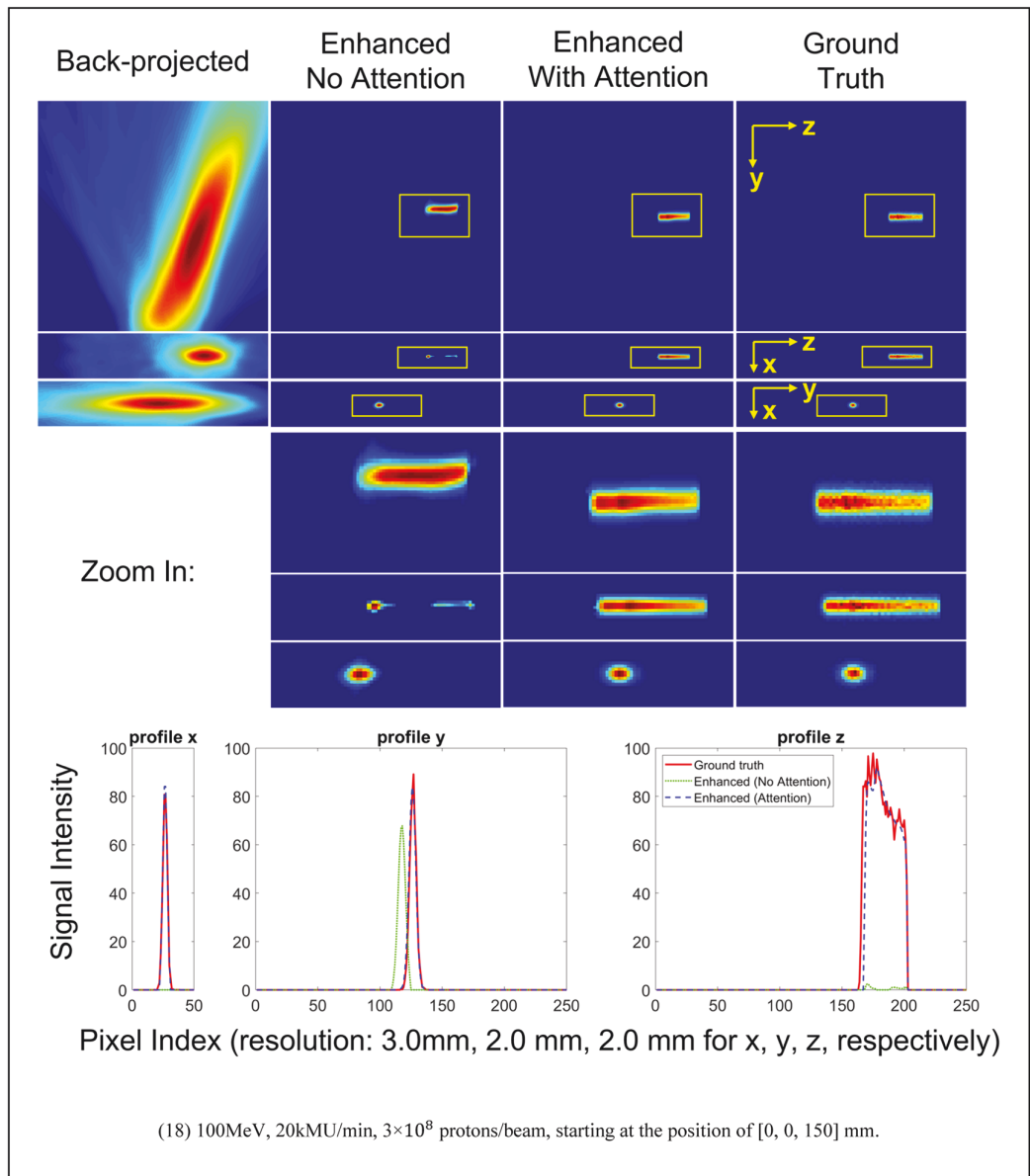


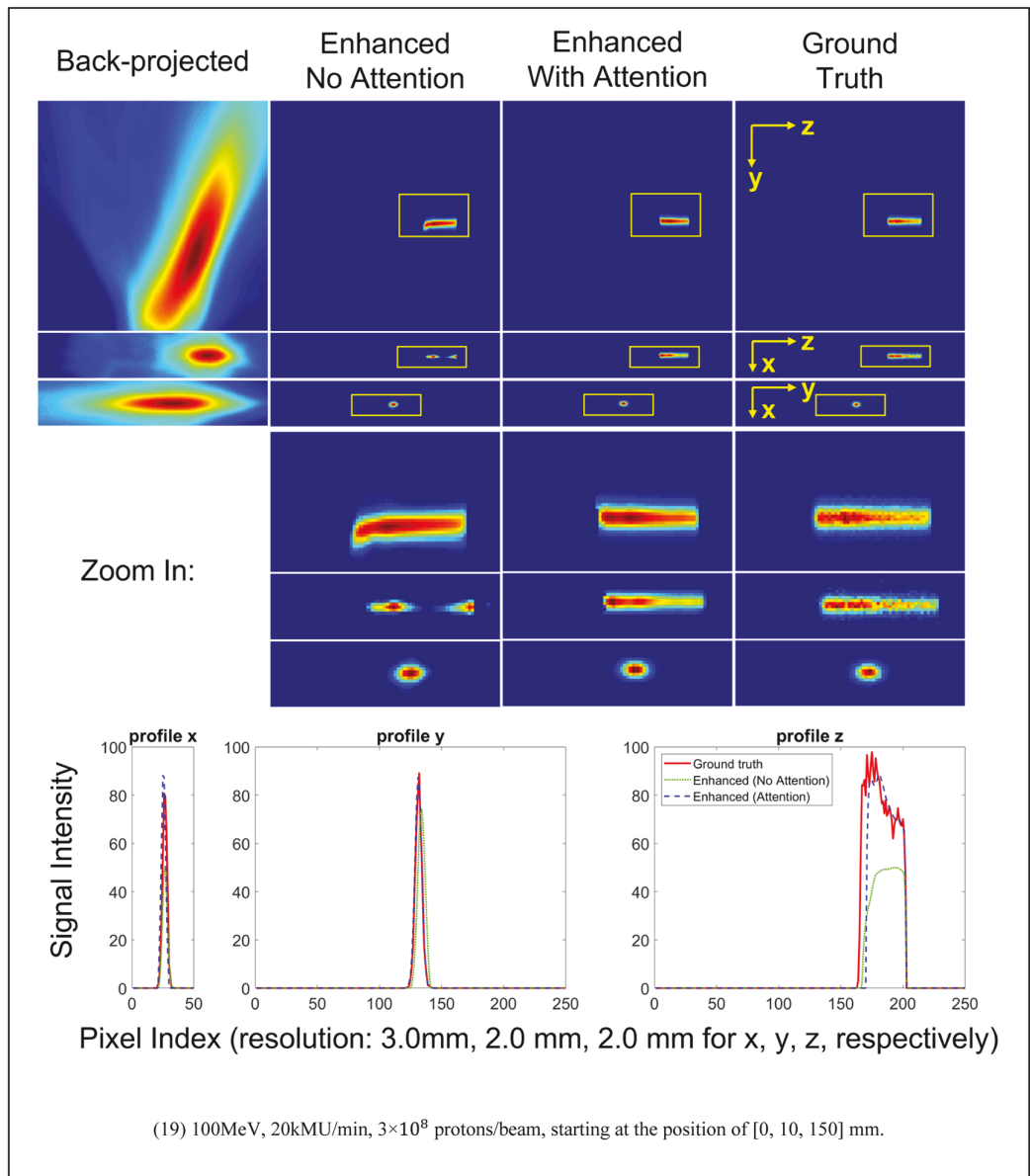


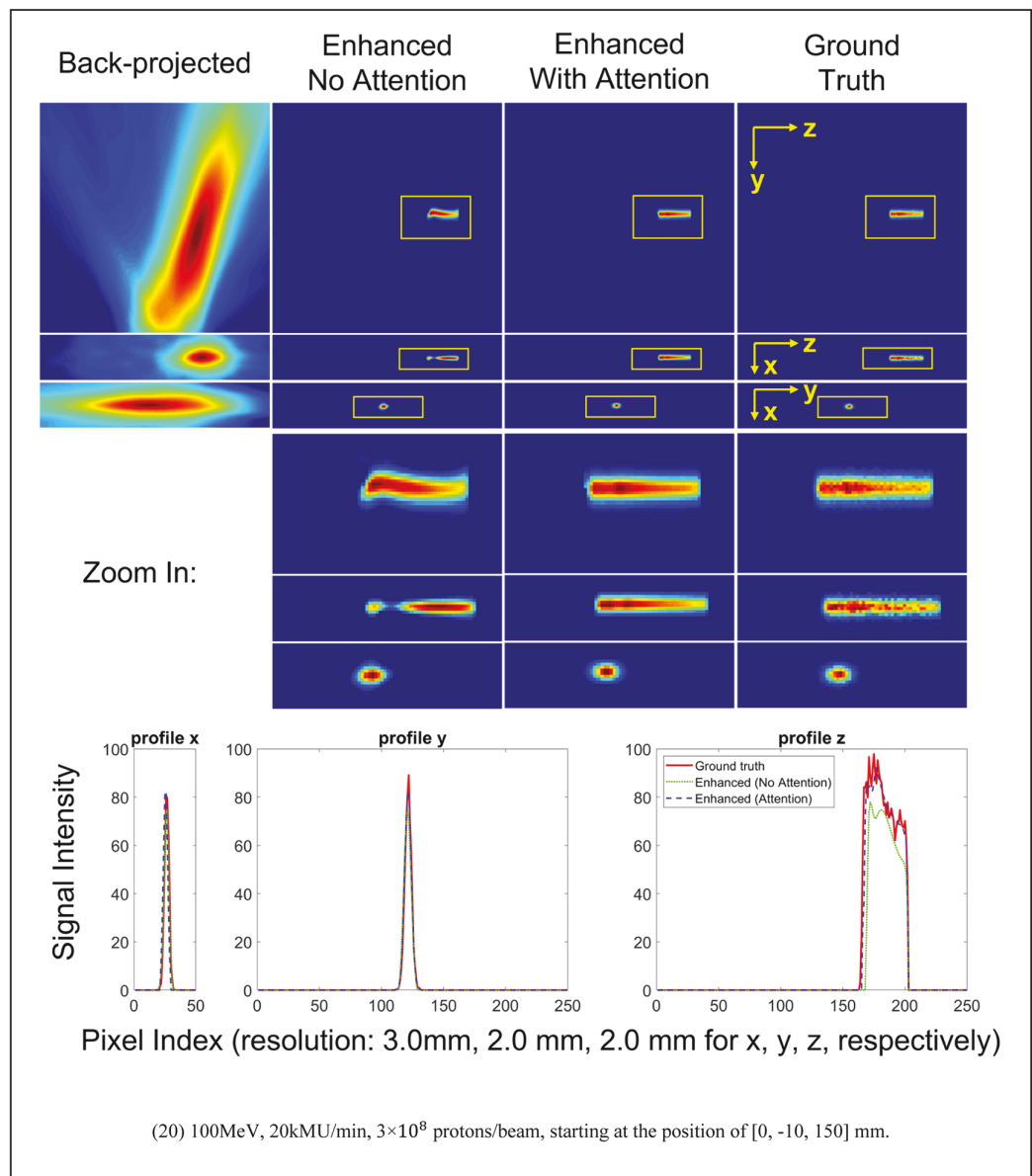


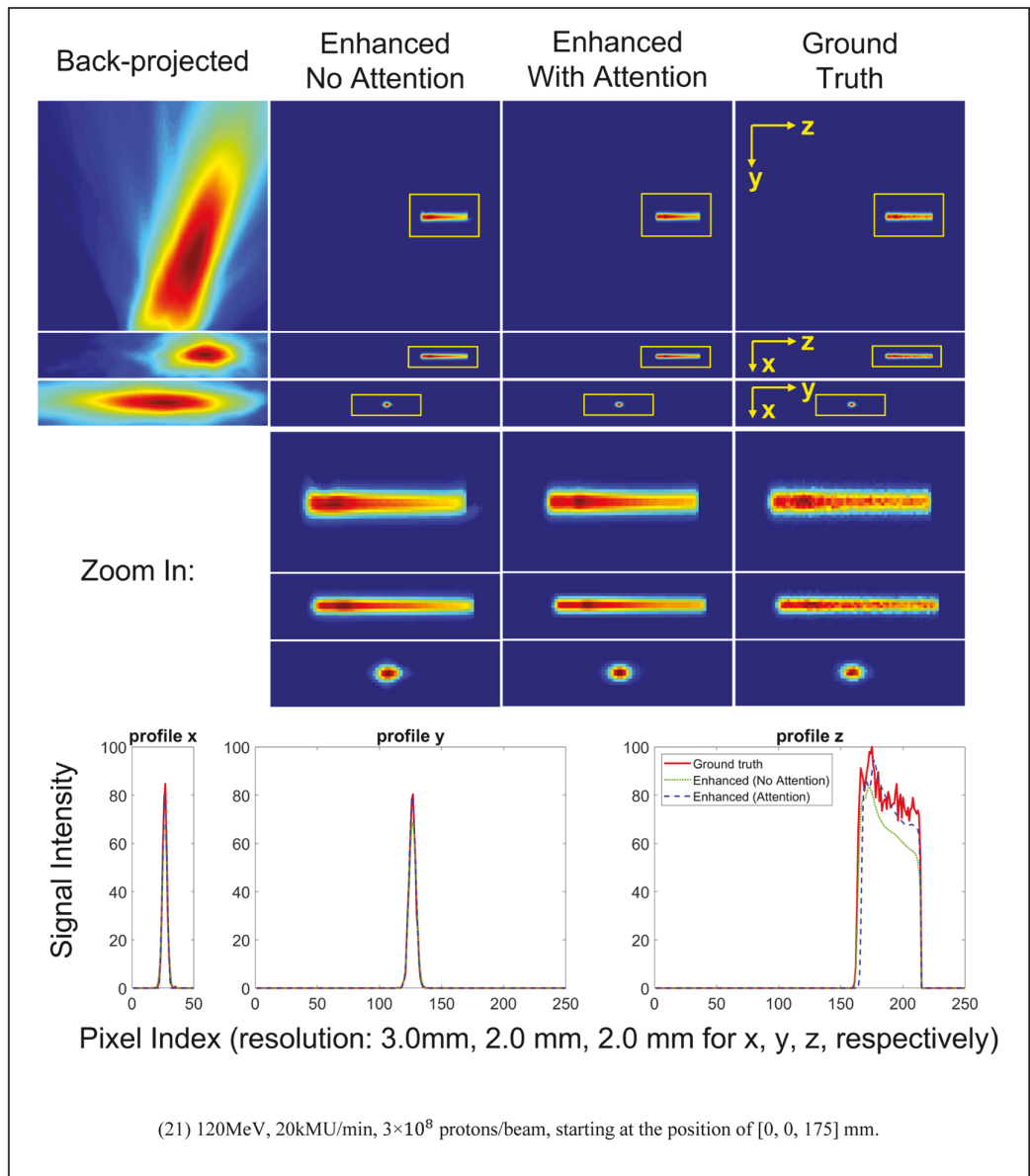


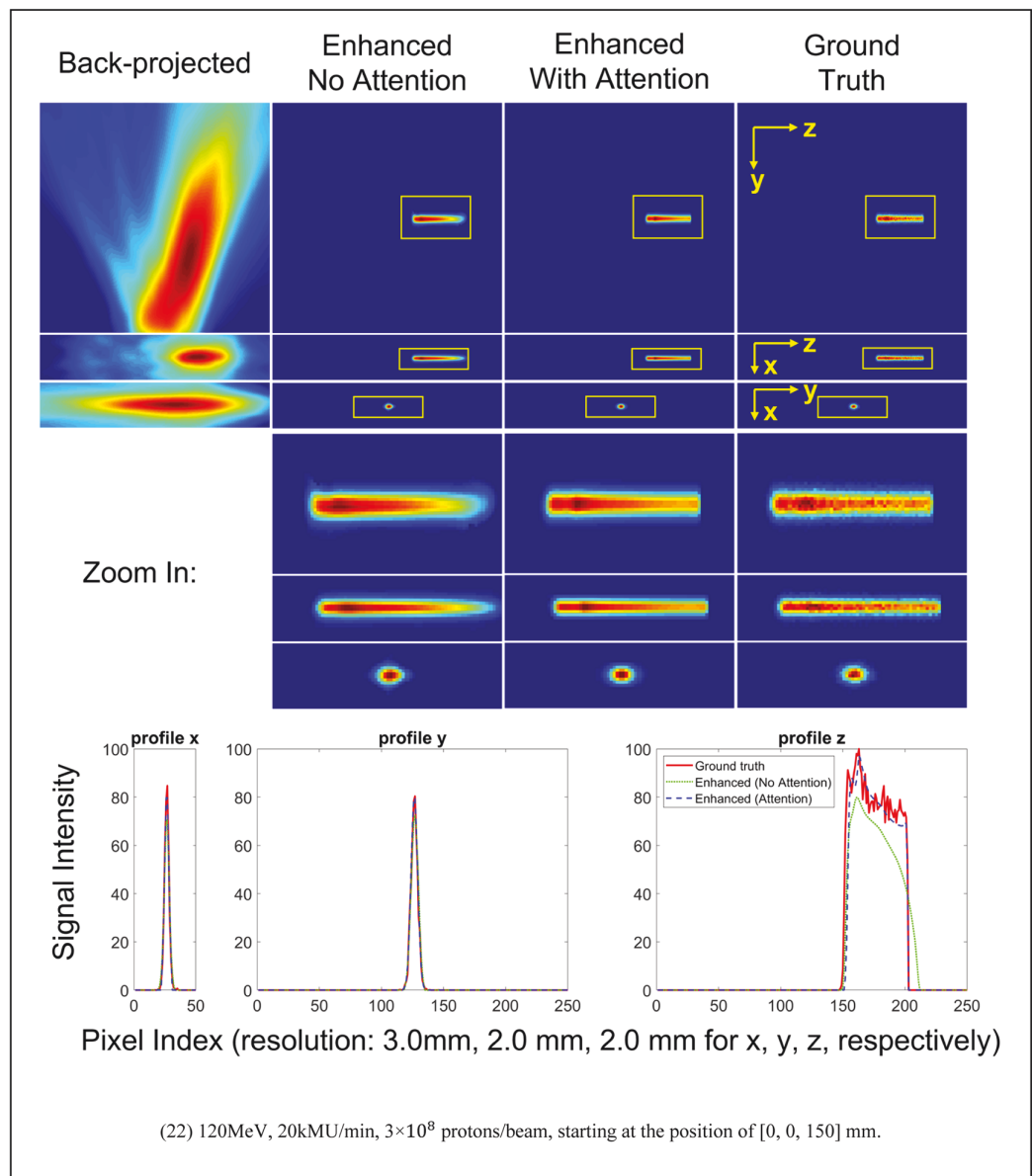


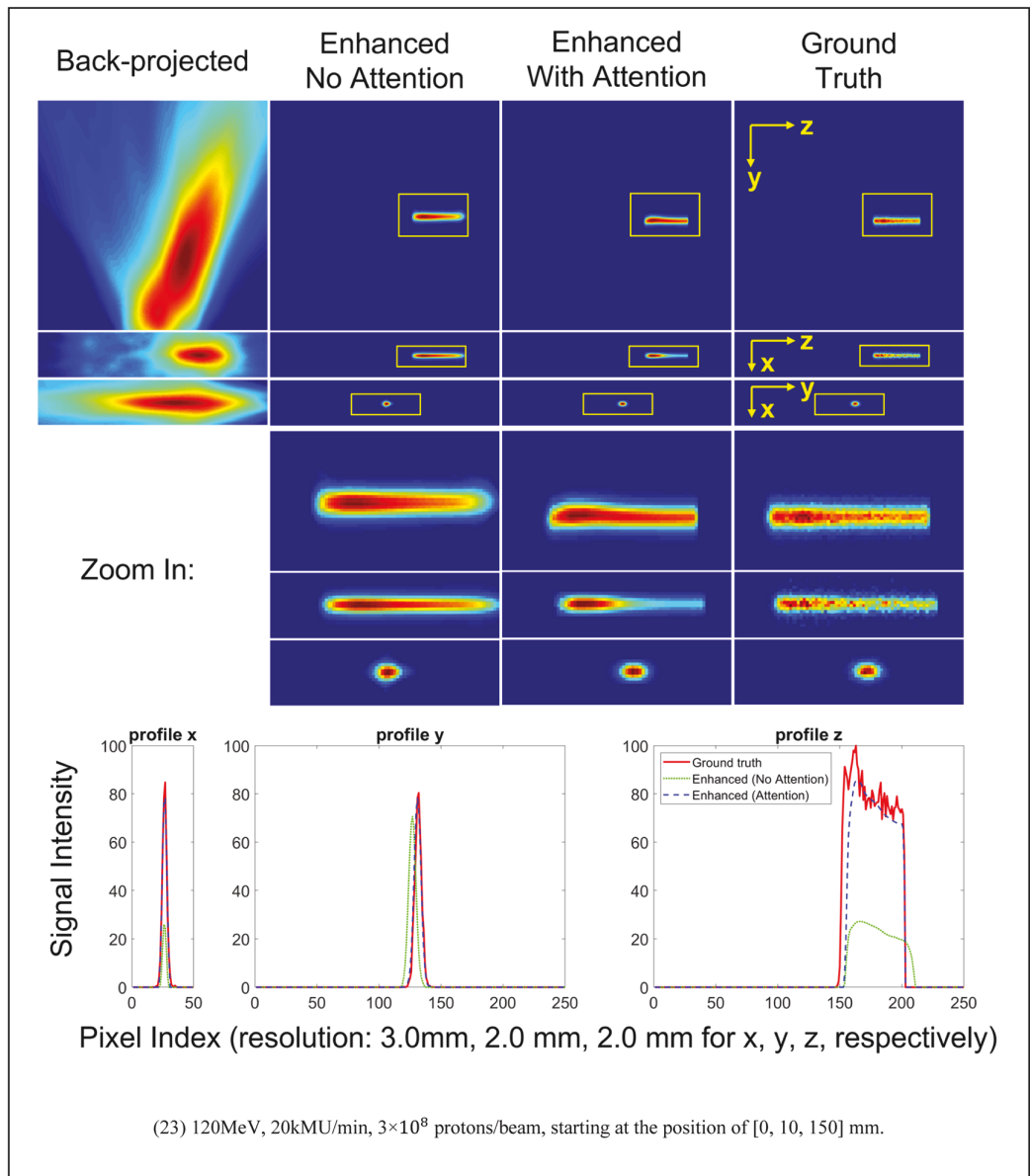


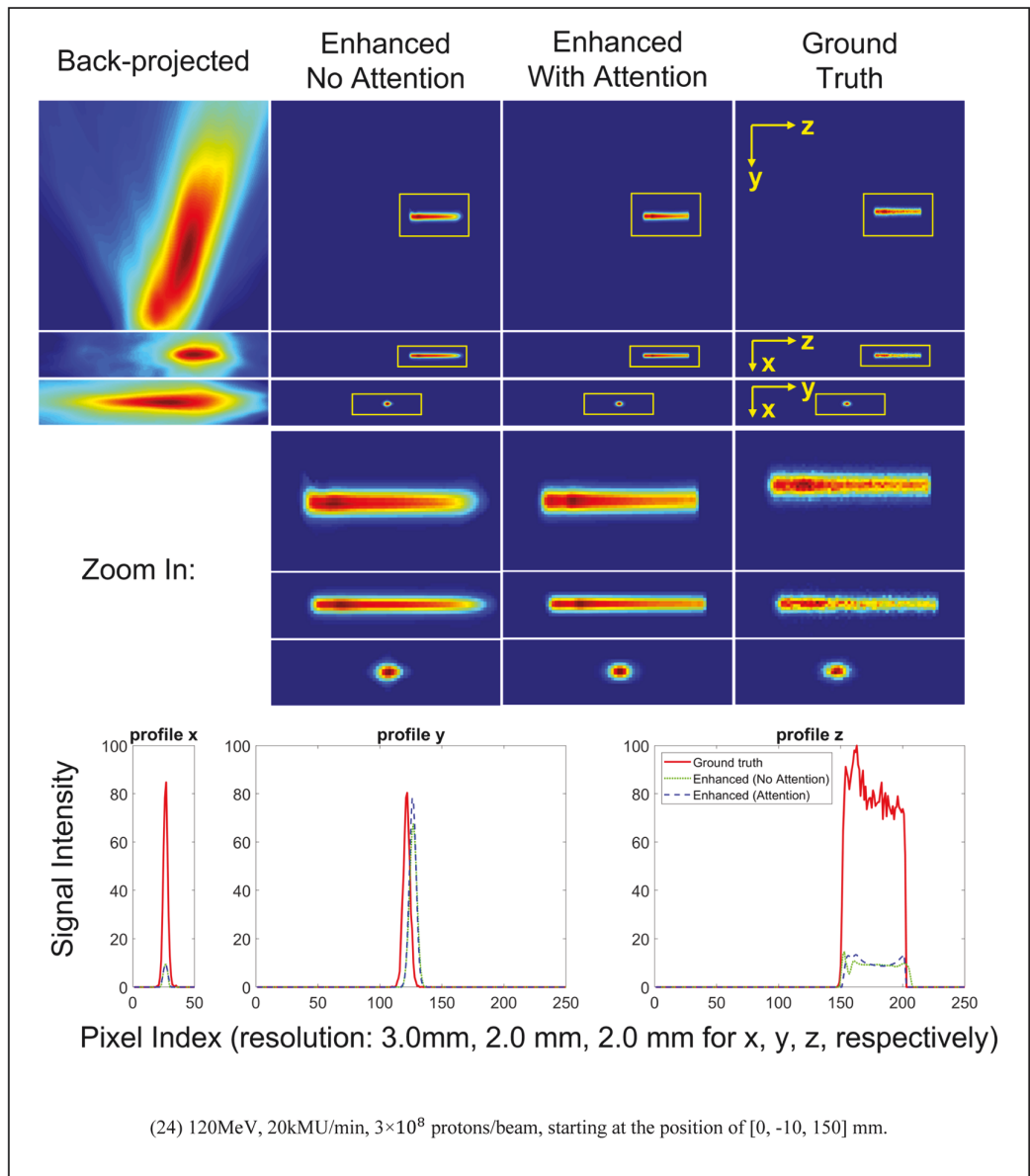


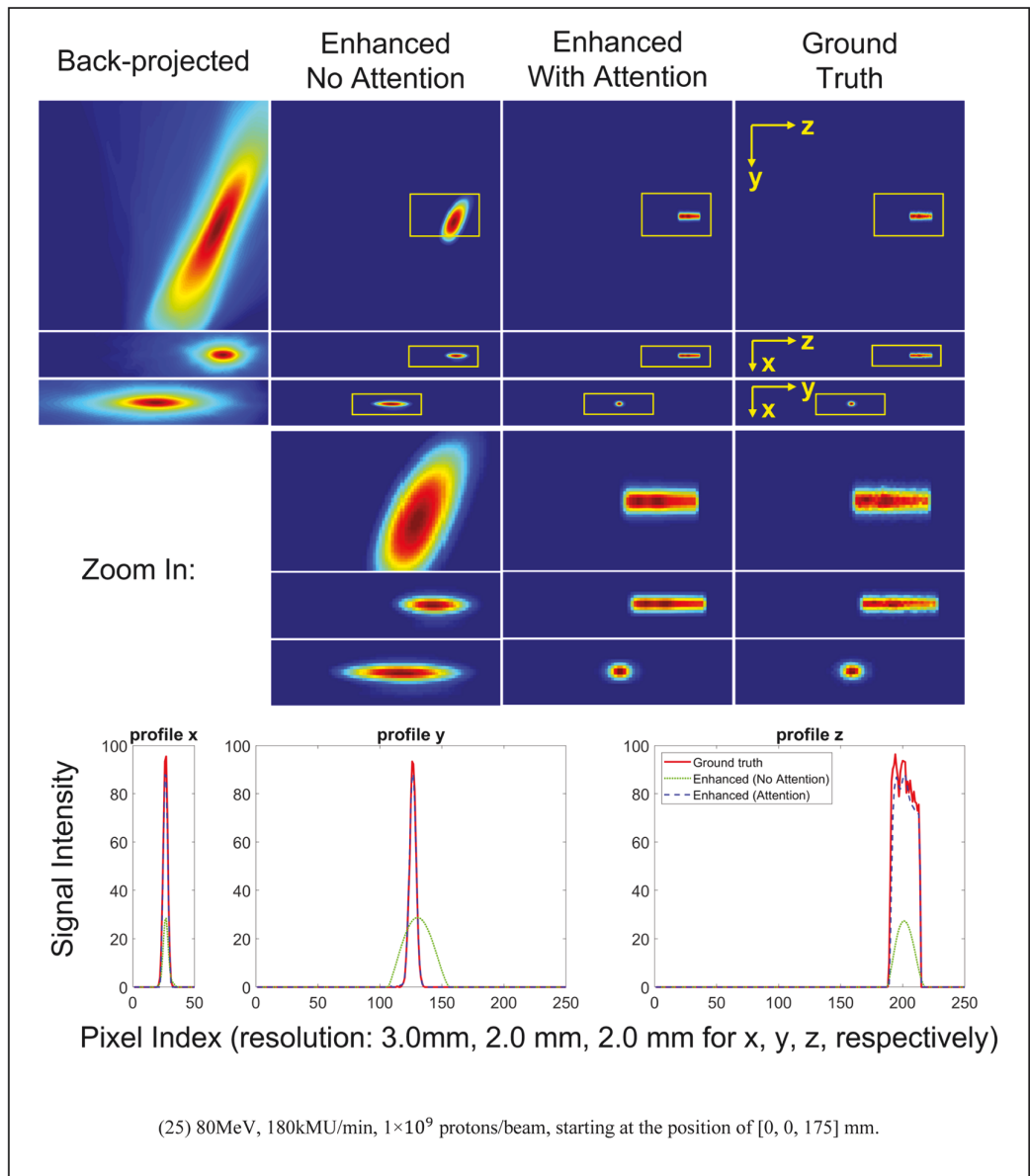


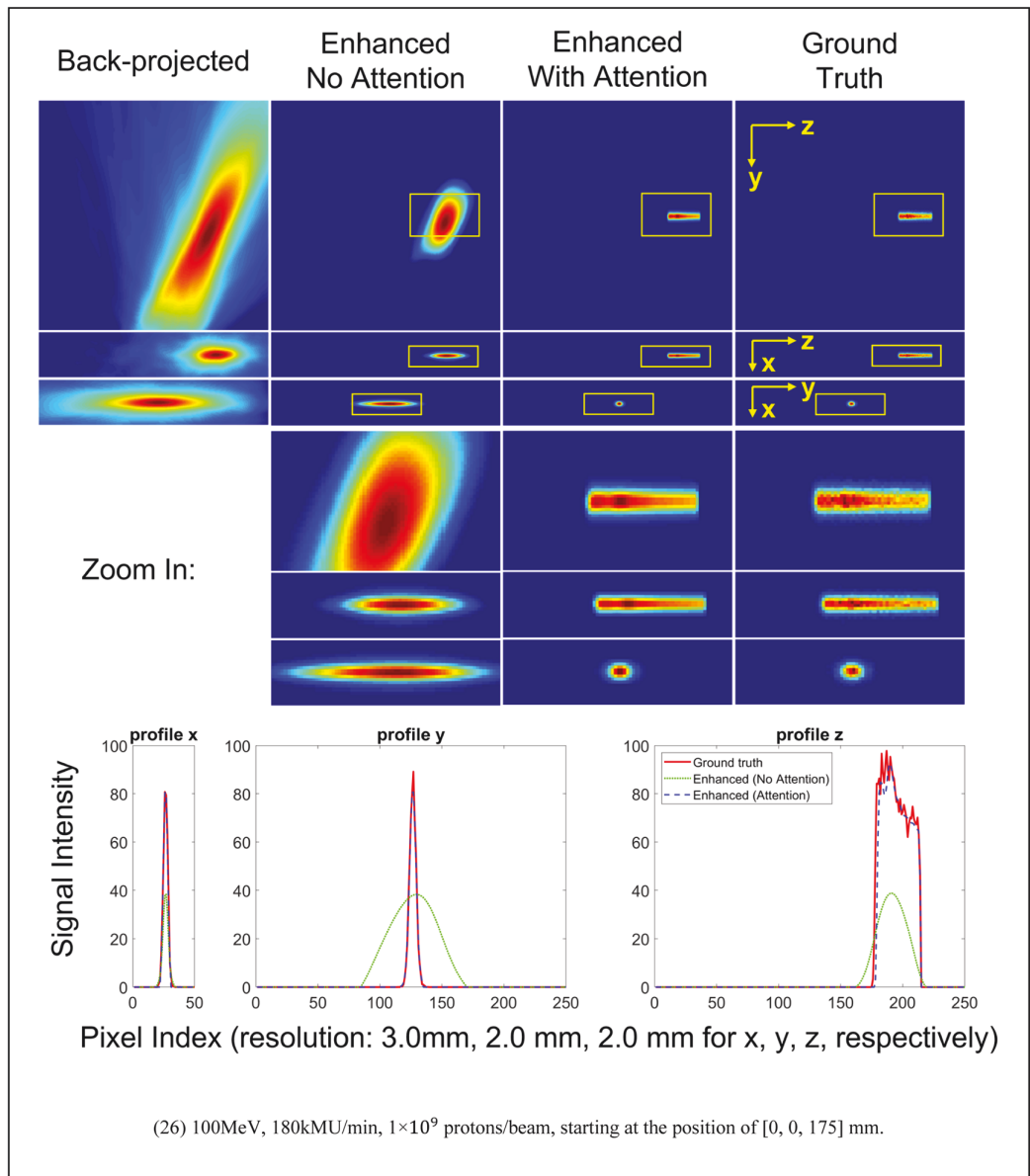












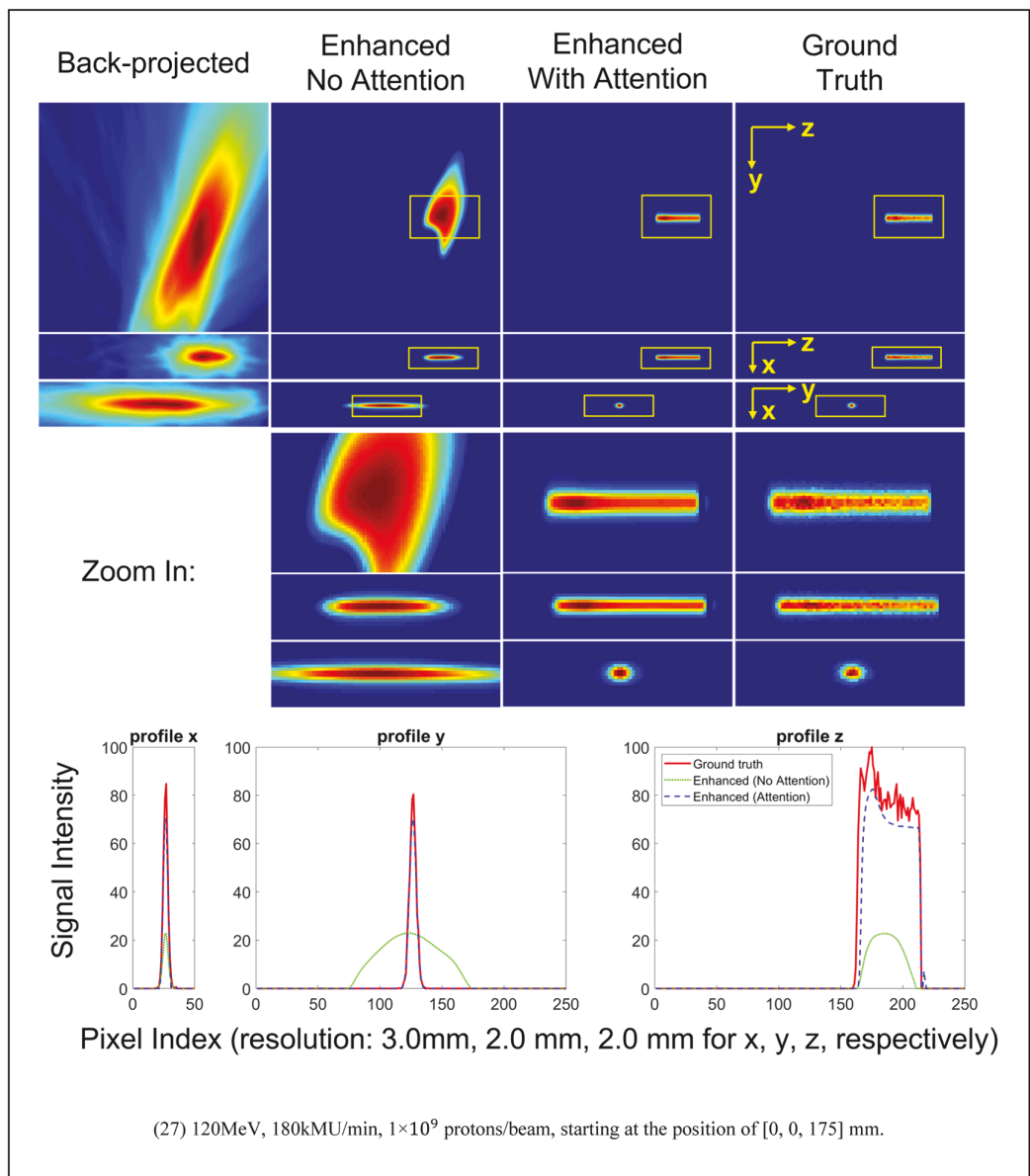


Table A1. Proton pencil beams simulated in this study.

Dose rate (kMU min ⁻¹)	Beam energy (MeV)	Dose level (protons/beam)	Beam starting point [x, y, z] in mm ⁺	Beam direction ⁺	Total numbers
20	75, 85, 95, 105, 115, 125	1 × 10 ⁹	[0, 0, 175]	-z	54
			[0, 0, 150]		
			[0, 20, 150]		
20	80, 90, 100, 110, 120	1 × 10 ⁹	[0, -20, 150]	-z	54
			[0, 0, 175]		
			[0, 0, 150]		
20	77.5, 82.5, 87.5, 92.5, 97.5, 102.5, 107.5, 112.5, 117.5, 122.5	1 × 10 ⁹	[0, 10, 150]	-z	54
			[0, -10, 150]		
			[0, 0, 175]		
20	75, 85, 95, 105, 115, 125	3 × 10 ⁸	[0, 0, 175]	-z	54
			[0, 0, 150]		
			[0, 20, 150]		
20	80, 90, 100, 110, 120	3 × 10 ⁸	[0, -20, 150]	-z	54
			[0, 0, 175]		
			[0, 0, 150]		
20	77.5, 82.5, 87.5, 92.5, 97.5, 102.5, 107.5, 112.5, 117.5, 122.5	3 × 10 ⁸	[0, 10, 150]	-z	54
			[0, -10, 150]		
			[0, 0, 175]		
180	75, 77.5, 80, 82.5, 85, 87.5, 90, 92.5, 95, 97.5, 100, 102.5, 105, 107.5, 110, 112.5, 115, 117.5, 120, 122.5, 125	1 × 10 ⁹	[0, 0, 175]	-z	21

References

- Agostinelli S *et al* 2003 GEANT4—a simulation toolkit *Nucl. Instrum. Methods Phys. Res. A* **506** 250–303
- Assmann W *et al* 2015 Ionoacoustic characterization of the proton Bragg peak with submillimeter accuracy *Med. Phys.* **42** 567–74
- Bom V, Joulaeizadeh L and Beekman F 2011 Real-time prompt gamma monitoring in spot-scanning proton therapy using imaging through a knife-edge-shaped slit *Phys. Med. Biol.* **57** 297
- Bria A, Marrocco C and Tortorella F J C 2020 medicine. Addressing class imbalance in deep learning for small lesion detection on medical images *Computers in biology and medicine* **120** 103735
- Donoho D L 2006 Compressed sensing *IEEE Trans. Inf. Theory* **52** 1289–306
- Ernst P, Chatterjee S, Rose G, Speck O and Nürnberger A 2021 Sinogram upsampling using Primal-Dual UNet for undersampled CT and radial MRI reconstruction arXiv:211213443
- Gensheimer M F *et al* 2010 *In vivo* proton beam range verification using spine MRI changes *Int. J. Radiat. Oncol. * Biol. * Phys.* **78** 268–75
- Jiang Z, Chang Y, Zhang Z, Yin F F and Ren L J M P 2022 Fast four-dimensional cone-beam computed tomography reconstruction using deformable convolutional networks *Med Phys* **49** 6461–76
- Jiang Z, Chen Y, Zhang Y, Ge Y, Yin F-F and Ren L 2019 Augmentation of CBCT reconstructed from under-sampled projections using deep learning *IEEE Trans. Med. Imaging* **38** 2705–15
- Jiang Z, Yin F-F, Ge Y and Ren L 2020 A multi-scale framework with unsupervised joint training of convolutional neural networks for pulmonary deformable image registration *Phys. Med. Biol.* **65** 015011
- Jiang Z, Yin F-F, Ge Y and Ren L 2021 Enhancing digital tomosynthesis (DTS) for lung radiotherapy guidance using patient-specific deep learning model *Phys. Med. Biol.* **66** 035009
- Johnson J M and Khoshgoftaar T M 2019b Survey on deep learning with class imbalance *J. Big Data* **6** 1–54
- Johnson J M and Khoshgoftaar T M J J B D 2019a Survey on deep learning with class imbalance *J. Big Data* **6** 1–54
- Jones K C *et al* 2015 Experimental observation of acoustic emissions generated by a pulsed proton beam from a hospital-based clinical cyclotron *Med. Phys.* **42** 7090–7
- Kim J-W 2009 Pinhole camera measurements of prompt gamma-rays for detection of beam range variation in proton therapy *J. Korean Phys. Soc.* **55** 1673–6
- Kingma D P and Ba J 2014 Adam: a method for stochastic optimization arXiv:1412.6980
- Knopf A-C and Lomax A J P M 2013 Biology. *In vivo* proton range verification: a review *Phys. Med. Biol.* **58** R131
- Kormoll T, Fiedler F, Schöne S, Wüstemann J, Zuber K and Enghardt W 2011 A Compton imager for *in vivo* dosimetry of proton beams—a design study *Nucl. Instrum. Methods Phys. Res. A* **626** 114–9
- Kroeger R, Johnson W, Kurfess J, Philips B and Wulf E 2000 Gamma ray energy measurement using the multiple compton technique *Paper presented at: 2000 IEEE Nuclear Science Symp.. Conf. Record (Cat. No. 00CH37149)*
- Liao H, Huo Z, Sehnert W J, Zhou S K and Luo J 2018 Adversarial sparse-view CBCT artifact reduction *Paper presented at: Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*
- Lustig M, Donoho D L, Santos J M and Pauly J M 2008 Compressed sensing MRI *IEEE Signal Process Mag.* **25** 72–82
- Maggi P *et al* 2020 Computational model for detector timing effects in Compton-camera based prompt-gamma imaging for proton radiotherapy *Phys. Med. Biol.* **65** 125004
- Min C-H, Kim C H, Youn M-Y and Kim J-W 2006 Prompt gamma measurements for locating the dose falloff region in the proton therapy *Appl. Phys. Lett.* **89** 183517

- Moteabbed M, España S and Paganetti H 2011 Monte Carlo patient study on the comparison of prompt gamma and PET imaging for range verification in proton therapy *Phys. Med. Biol.* **56** 1063
- Munoz E et al 2021 Proton range verification with MACACO II Compton camera enhanced by a neural network for event selection *Scientific Reports* **11** 1–12
- Nishio T, Miyatake A, Ogino T, Nakagawa K, Saijo N and Esumi H 2010 The development and clinical use of a beam ON-LINE PET system mounted on a rotating gantry port in proton therapy *Int. J. Radiat. Oncol. * Biol. * Phys* **76** 277–86
- Paganetti H J P M 2012 Biology. Range uncertainties in proton therapy and the role of Monte Carlo simulations *Phys. Med. Biol.* **57** R99
- Panthi R, Maggi P, Peterson S, Mackin D, Polf J and Beddar S 2020 Secondary particle interactions in a Compton camera designed for *in vivo* range verification of proton therapy *IEEE Trans. Radiat. Plasma Med. Sci.* **5** 383–91
- Parodi K and Polf J C J M 2018 *In vivo* range verification in particle therapy **45** e1036–50
- Peterson S, Mackin D, Hillebrand M, Draeger E, Polf J and Beddar S 2016 Tracking electrons produced by compton scatter within a prompt gamma imaging device *Paper presented at: Proc. SAIP*
- Peterson S, Robertson D and Polf J J P M 2010 Biology. Optimizing a three-stage Compton camera for measuring prompt gamma rays emitted during proton radiotherapy *Phys. Med. Biol.* **55** 6841
- Polf J C et al 2022 Applications of machine learning to improve the clinical viability of Compton camera based *in vivo* range verification in proton radiotherapy *Front. Phys.* **10** 284
- Polf J C, Maggi P, Panthi R, Peterson S, Mackin D and Beddar S 2021 The effects of Compton camera data acquisition and readout timing on PG imaging for proton range verification *IEEE Trans. Radiat. Plasma Med. Sci.* **6** 366–73
- Richard M-H et al 2010 Design guidelines for a double scattering compton camera for prompt- γ imaging during ion beam therapy: a monte carlo simulation study *IEEE Trans. Nucl. Sci.* **58** 87–94
- Roellinghoff F et al 2011 Design of a Compton camera for 3D prompt- γ imaging during ion beam therapy *Nucl. Instrum. Methods Phys. Res. A* **648** S20–3
- Ronneberger O, Fischer P and Brox T 2015 U-net: convolutional networks for biomedical image segmentation *Paper presented at: Int. Conf. on Medical Image Computing and Computer-assisted Intervention*
- Sidky E Y and Pan X 2008 Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization *Phys. Med. Biol.* **53** 4777
- Verburg J M, Riley K, Bortfeld T and Seco J 2013 Energy-and time-resolved detection of prompt gamma-rays for proton range verification *Phys. Med. Biol.* **58** L37
- Wang H, Zhou L and Wang L 2019 Miss detection versus false alarm: adversarial learning for small object segmentation in infrared images *Paper presented at: Proc. of the IEEE/CVF Int. Conf. on Computer Vision*
- Yang G et al 2017 DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction *IEEE Trans. Med. Imaging* **37** 1310–21
- Zhang M et al 2020 Deep-learning detection of cancer metastases to the brain on MRI *J Magnetic Resonance Imaging* **52** 1227–36
- Zoglauer A and Boggs S E 2007 Application of neural networks to the identification of the compton interaction sequence in compton imagers *Paper presented at: 2007 IEEE Nuclear Science Symp. Conf. Record*