

Deep Learning-Based 3D Dose Reconstruction and Deviation Detection from Positron-Emitter Activations in Proton Therapy

Pablo Cabrales

*Group of Nuclear Physics
Complutense University of Madrid
Madrid, Spain
pcabrales@ucm.es*

Víctor V. Onecha

*Massachusetts General Hospital
& Harvard Medical School
Boston, US
vonecha@mgh.harvard.edu*

José Manuel Udías

*Group of Nuclear Physics & IPARCOS
Complutense University of Madrid
Madrid, Spain
jmudiasm@ucm.es*

David Izquierdo-García

*Massachusetts General Hospital
& Harvard Medical School
Boston, US
dizquierdogarcia@mgh.harvard.edu*

Joaquín L. Herraiz

*Group of Nuclear Physics & IPARCOS
Complutense University of Madrid
Madrid, Spain
jlopezhe@ucm.es*

Abstract—Proton therapy has become an established treatment option for various cancers, including pediatric, brain, and deep-seated tumors. Despite its benefits, its use is limited by the uncertainties in radiation dose deposition. This can lead to healthy tissue overdosage or tumor undertreatment. In our work, we present a deep learning-based workflow to detect deviations in 3D dose depositions from the isotope activations induced by the proton beams. We train deep learning models to detect clinically relevant patient positioning and physical parameter deviations on a reference prostate treatment plan. Three widely used convolutional neural networks are compared: SegResNet, UNet, and SwinUNETR. We observe that all models accurately predict the dose deposition deviations with over a 99.99% Gamma Passing Rate and that the SwinUNETR model, which incorporates vision transformer blocks, provides the best results. Moreover, we implement a Deviation Predicting Branch (DPB) of fully connected layers stemming from the penultimate encoder layer. The DPB is shown to accurately identify deviations in patient positioning within a tenth of a millimeter and a tenth of a degree. In total, the data generation and the model training can be completed with a GPU in under ninety minutes. The developed workflow can be used to identify deviations from the treatment plan which can be corrected online or in the following treatment sessions.

Index Terms—Bragg Peak (BP), Spread-Out Bragg Peak (SOBP), Convolutional Neural Network (CNN), Deviation-Predicting Branch (DPB), Mean Absolute Error (MAE), Mean Squared Error (MSE), Distance-To-Agreement (DTA)

I. INTRODUCTION

Proton therapy is a radiation treatment for oncologic patients that conforms a more precise dose in tumors than traditional radiotherapy. This is because protons have a finite range and deposit most of their energy near the end of their path, in a region known as the Bragg Peak (BP) [1]. This property enables the reduction of irradiated healthy tissue and neigh-

boring organs at risk, and therefore, those toxicities derived from their irradiation [2]–[4].

Despite these advantages, dose deposition uncertainties present a major challenge in proton therapy treatment planning. Patient motion, physiological changes between treatment sessions (such as inflammation, gas regions, organ movement, variations in patient weight, etc.), and uncertainties in physical parameters can lead to dose deposition differences between the expected and actual treatment [5]–[8]. Because of these limitations, safety margins are applied during treatment planning to ensure the full irradiation of the tumor [9].

Several methods have been proposed to verify dose deposition and reduce these uncertainties. They rely on the detection of secondary emissions induced by the protons within the patient's body [10]–[13]. Positron Emission Tomography (PET) based verification [14] is one of these approaches and relies on the activation of positron-emitting isotopes, such as ^{11}C , ^{13}N , and ^{15}O . These positron-emitter activations can be imaged non-invasively using a PET scanner. However, because isotope activation and dose deposition correspond to different physical processes, the activation image needs to be translated to a dose image to detect dose deposition deviations.

To this effect, several studies have been carried out. In [15], individual proton beam doses were reconstructed from a dictionary of simulated activity and dose image pairs. However, the reconstruction of a single beam might take more than two seconds, which makes real-time, in-beam verification unfeasible for the hundreds of beams applied in each session. Moreover, if the detected activation falls between two dictionary beams, the method will not be able to generalize and reconstruct the dose. To generate dose profiles in milliseconds and generalize beyond simulated treatment cases, deep learning approaches have also been proposed [16]–

[19]. However, in these studies, only one or two-dimensional dose profiles were considered. In addition, the input activations are either individual proton beams [19], which differ from the expected cumulative activations of many beams, or Spread-Out Bragg Peaks (SOBPs) at different locations of a single CT [18], which do not yield realistic deviations from the treatment plan. Moreover, the SOBPs are CPU-simulated, which may require several days to complete. Finally, both [18], [19] are trained with GAN-based deep learning models, which have been proven to be challenging to train and deploy [20].

To overcome these limitations, in this work, we have evaluated the performance of three convolutional-based deep learning models (UNet [21], SegResNet [22], and SwinUNETR [23]) to predict the 3D dose deposition images of an SOBP from the corresponding activation images. They were trained with realistic simulations obtained using the Monte Carlo GPU-based simulation tool FRED [24]. These simulations include variations in tissue density, material excitation energies, and patient positioning.

Furthermore, we modified the SwinUNETR network to include a branch of fully connected layers stemming from the penultimate encoder layer that outputs the deviations in SOBP positioning. We call this branch a Deviation-Predicting Branch (DPB), and it gives the user insight into why the treatment session doses deviated from the treatment plan.

Our goal has been to develop a complete workflow that can be readily implemented within the time window of 24 hours between the planning CT and the treatment session. This workflow would begin with the simulation of a dataset of realistic deviations from the CT treatment plan. Then, the deep learning model would be trained with the dataset to predict these deviations. Finally, the model would be ready to detect realistic 3D dose deposition deviations in milliseconds from PET scans obtained during, or immediately after, the treatment session.

II. METHODS

A. Training Dataset of SOBP Deviations

A reference SOBP treatment plan comprised of 311 parallel proton beams on a prostate CT (Figure 1) was obtained using MatRad [25]. The three-dimensional isotope activations and dose deposition distributions are generated with the GPU Monte Carlo simulation tool FRED [26]. The images have a size of 160x32x32 voxels and a resolution of 1x2x2mm. The first direction is parallel to the proton beam path and has a higher resolution to detect range deviations more precisely. Each beam is obtained by simulating 100,000 protons and the resulting images are scaled to 2.5 Gy (see Figure 1).

The dataset to train the neural networks was created considering this SOBP as the reference and generating secondary, additional scenarios by assuming plausible clinical deviations, both in the CT and the plan itself. On the one hand, we assumed slight rotations on the SOBP's beam angles up to $\Delta\theta = \pm 2^\circ$ with respect to the reference beam direction (x -axis), as well as displacements of up to $\Delta y = \pm 2mm$, $\Delta z = \pm 2mm$ in the two directions perpendicular to the beam.

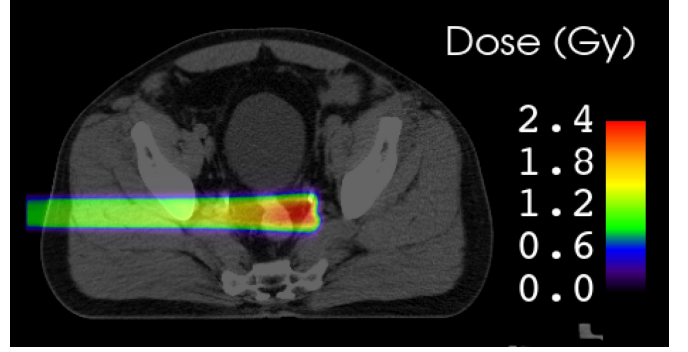


Fig. 1. Dose deposited by the reference SOBP overlaid on the prostate CT

This could reflect subtle patient movements that may take place within the treatment session [27], [28]. On the other hand, we introduce variations of up to 10% in the mean excitation energies and densities of the patient's tissues, which correspond to reasonable uncertainties in these parameters [29]. 300 deviated SOBP were simulated, generating a dataset of 300 dose-activation image pairs with a size of 379 MB.

The presence of clinically relevant differences between the expected reference treatment plan dose and the deviated dose which results from the spatial and numerical uncertainties, in three sample cases, is shown in Figure 2.

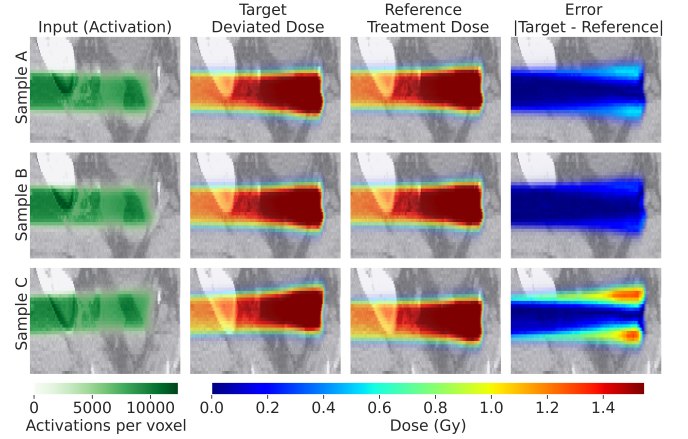


Fig. 2. For three sample cases, the deviated doses are compared to the reference treatment plan doses. The color map is truncated at 1.4 Gy to showcase these differences.

B. Neural Networks, Training Strategies, and Validation Metrics

The purpose of the deep learning models is to accurately predict the dose deposition images (output) from the activation images (input). Convolutional neural networks (CNNs) apply convolutional filters to extract image features [30] and are widely used in computer vision and medical imaging for their ability to identify complex patterns and structures [31]. In particular, the UNet, a CNN based on an encoder-decoder structure with skip-connection blocks [21], has become the architectural backbone for nearly all image translation models

[32]. Image translation tasks involve translating input images of a specific modality to a different output modality, and the UNet excels at capturing both the high-resolution details and low-level features of the translated images. For this reason, we choose the UNet as a baseline model to translate activation into dose images. More specifically, we consider a simple four-layer UNet with 16, 32, 64, and 128 channels per layer, a stride of two, and a kernel size of three.

To study the effect of different network components and to identify the best possible model for our problem, two other UNet-based models are considered. First, the SegResNet model [22], which incorporates residual connections to a UNet architecture, can show improved convergence, generalization, and training time by skipping specific layer connections [33]. Second, the SwinUNETR model [23] incorporates sliding-window vision transformers as encoder layers and better captures the relationship between distant parts of the image [34], which might be useful in images generated from proton beams. This is because protons interact with tissue in their path across the entire image space. By including vision transformer blocks, the resulting SwinUNETR model is larger and more complex than the proposed UNet and the SegResNet. The selected models constitute a relevant display of state-of-the-art image translation models and incorporate architectural choices that are anticipated to yield the best possible outcomes [35]. Although diffusion-based models could also be considered, long inference times would hinder their applicability to our problem [36].

Furthermore, we introduce a Deviation-Predicting Branch (DPB) that originates from the second-to-last encoder layer of the SwinUNETR architecture (see Figure 3). It includes a flattening layer to process the encoder output dimensions of $96 \times 20 \times 4 \times 4$, followed by four densely connected layers with LeakyReLU activation functions [37]. Sequentially, the dimensions of these dense layers are 1024, 512, 256, and 4. The final layer outputs the four values corresponding to the spatial deviations from the reference treatment plan: Δy , Δz , $\Delta \theta$, and $\Delta \phi$. Note that the angle between the deviated and the reference beam, which corresponds to the polar angle θ , is varied within $\pm 2^\circ$ whereas the azimuthal angle ϕ is varied from 0° to 180° . These outputs are compared to the simulation target values and the Mean Squared Error (MSE) is used to train the model. All three models were adapted from MONAI [38] for PyTorch [39].

The models were trained for a maximum of 200 epochs using a batch size of 8 and a Quadro GV100 GPU. An early stopping criterion was implemented to stop model training if the validation loss did not improve in the last 20 epochs. The learning rate was set to 0.001 and the optimizer used was AdamW [40]. As standard scaling may be susceptible to outlier pixel values, activation and dose images were robust-scaled [41]. The training, validation, and testing split was 80% (240 images), 10% (30 images), and 10% (30 images), respectively. The metric used to train the model is a weighted combination of the Mean Absolute Error (MAE), to avoid the over-smoothing caused by the Mean Squared Error (MSE),

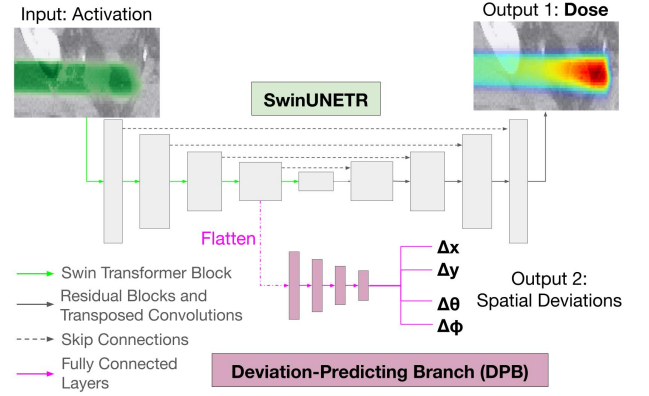


Fig. 3. SwinUNETR model with our Deviation-Predicting Branch (DPB) to output both the dose deposition and the spatial deviation parameters.

and a differentiable gamma index loss [42]. To test the model, we used the MSE loss, given its widespread presence in computer vision problems, and the mean gamma index (Γ) and Gamma Passing Rate (GPR) [43], commonly used for assessing radiotherapeutic dose differences and dosimetric model validations. The gamma index was computed using open-source code from [44]. The gamma index $\Gamma(\vec{r}_0)$ for an image point \vec{r}_0 is calculated by considering its distance to other points in the reference image \vec{r}_f , weighted by a Distance-To-Agreement (DTA), and their dose difference, weighted by a maximum dose difference of ΔD , as shown in Equation 1. For our study, we used clinically standard values of $DTA = 1mm$ and $\Delta D = 3\%$.

$$\Gamma(\vec{r}_0) = \min_{\vec{r}_f} \sqrt{\left(\frac{|\vec{r}_f - \vec{r}_0|}{DTA}\right)^2 + \left(\frac{|D(\vec{r}_f) - D(\vec{r}_0)|}{\Delta D}\right)^2} \quad (1)$$

Accordingly, for voxels that are spatially close in both images (low $|\vec{r}_f - \vec{r}_0|$) and that have similar dose values (low $|D(\vec{r}_f) - D(\vec{r}_0)|$), the gamma index value will be low and both dose deposition images will be in agreement. The GPR is found by finding the fraction of image pixels that meet the passing criterion in Equation 2 and have a dose greater than 10% of the maximum image dose.

$$\Gamma(\vec{r}_0) \leq 1 \quad (2)$$

III. RESULTS

A. Model validation

Table I shows the performances of the four deep learning models considered in this work: SegResNet, UNet, and SwinUNETR, as well as the proposed SwinUNETR with the DPB.

The best-performing model is further validated by showing its ability to identify the deviations in the dose deposition from the input activations, in Figure 4. This is shown for the same three samples from Figure 2.

The mean and standard deviations for the difference between the output and target displacements, $\Delta y, z_{out}$, and

TABLE I
COMPARISON OF MODEL PERFORMANCES IN TERMS OF MSE LOSS, GAMMA INDEX, AND GAMMA PASSING RATE. IN BOLD, THE BEST RESULTS. A LOWER LOSS AND GAMMA INDEX, AND A HIGHER PASSING RATE ARE CONSIDERED BETTER. THE BASELINE IS THE AVERAGE DIFFERENCE BETWEEN A DEVIATED SOBP DOSE AND THE REFERENCE TREATMENT PLAN DOSE.

Model	MSE Loss ($\times 10^{-5}$ Gy ²)	Gamma Index	Gamma Pass. Rate (%)
Baseline (No Correction)	2606	1.4	41
SegResNet	4.4	0.094	99.995
UNet	3.8	0.093	99.997
SwinUNETR	3.6	0.091	99.997
SwinUNETR with DPB	3.5	0.090	99.998

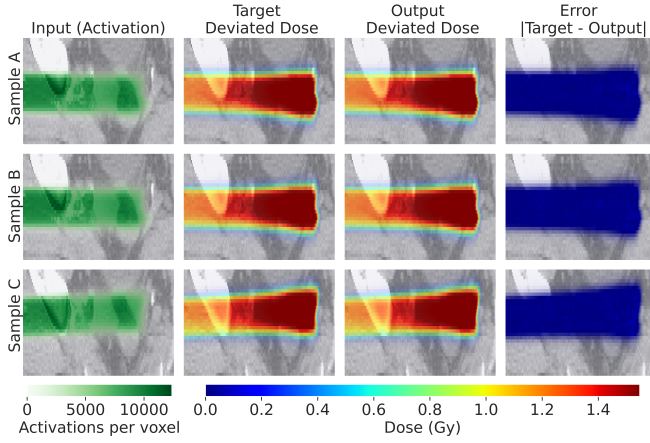


Fig. 4. For each detected activation, the model’s output deviated SOBP dose and the expected, or target, deviated SOBP dose are compared. These samples correspond to the testing set (i.e. the model has not been trained with them and has not “seen” them yet). The color map is truncated at 1.4 Gy for consistency with Figure 2.

$\Delta y, z_{tar}$, respectively, are shown in Equation 3. We aggregate the errors in y and z because they are similar in magnitude. The difference between the output and target angle deviations, $\Delta\theta_{out}$, $\Delta\phi_{out}$, and $\Delta\theta_{tar}$, $\Delta\phi_{tar}$, respectively, are shown in Equations 4 and 5.

$$|\Delta y, z_{out} - \Delta y, z_{tar}| = 0.1 \pm 0.1 mm \quad (3)$$

$$|\Delta\theta_{out} - \Delta\theta_{tar}| = 0.06 \pm 0.06^\circ \quad (4)$$

$$|\Delta\phi_{out} - \Delta\phi_{tar}| = 4 \pm 5^\circ \quad (5)$$

B. Model complexity

In Table II, the model complexities, training times, and inference times are compared. The GFLOPs column represents the computational complexity in giga-floating point operations, which indicates the amount of computational work the model requires for a single inference. The Parameters column shows the total number of trainable parameters within the model, measured in millions (M). Fewer parameters typically mean

a model is simpler and might require less storage memory. The Training column records the time necessary to train the model from scratch until convergence, measured in minutes. Lastly, the Inference column lists the time required for the model to process an input activation and provide an output dose, measured in milliseconds (ms). Faster inference times are preferred for applications requiring real-time predictions, as is the case for online dose deposition verification.

TABLE II
COMPARISON OF MODEL COMPLEXITIES, TRAINING TIMES, AND INFERENCE TIMES. IN BOLD, THE BEST RESULTS. LESS COMPLEX, FASTER MODELS ARE CONSIDERED BETTER.

Model	GFLOPs	Parameters	Training	Inference
SegResNet	5.79	1.18 M	14 min	7 ms
UNet	3.89	1.19 M	17 min	13 ms
SwinUNETR	31.1	15.7 M	24 min	30 ms
SwinUNETR with DPB	31.2	47.5 M	28 min	36 ms

IV. DISCUSSION

The relevance of seemingly slight deviations in positioning and physical parameters is apparent in the large dose deposition deviations recorded. This is shown in the second row of Table I with a critically low gamma passing rate between the reference treatment dose and the deviated dose of 41%, as well as in Figure 2, where specific regions receive doses up to 1.5 Gy greater, or lower than expected. Although we simulated proton beam rotations and displacements, this is qualitatively analogous to deviations in patient positioning.

The deep learning models excel at predicting the dose deposition deviations from the activation images. Although the SwinUNETR model outperformed the SegResNet and UNet models, the three models show nearly perfect gamma passing rates and an almost thousand-fold improvement in MSE loss, as recorded in Table I. This highlights the ability of the deep learning models to translate the activation images to the corresponding dose images. This is not a trivial task because dose deposition and activation correspond to different physical processes, and dose can be deposited at proton energies at which few isotopes are generated. This is seen at the distal end of the SOBP, where the dose deposition is maximal but the activation is comparatively low. The model hyperparameters were selected based on preliminary testing, although a comprehensive optimization process should be undertaken in future works.

Moreover, the SwinUNETR model with the DPB slightly outperforms the original one. This indicates that the additional tuning of the first four encoder blocks from the backpropagation of the DPB’s losses might help our model reconstruct the volumetric dose deposition. We observed that the best DPB results were obtained when branching from the fourth, or second-to-last, encoder layer. The last layer, commonly referred to as the bottleneck, might be over-compressed spatially

to extract spatial deviations and early encoder layers might not be sufficiently downsampled to efficiently extract information from them. Although we also introduced deviations in the excitation energies and densities of the tissues, we did not try to predict them. This is because, since they are inherent to the conversion process, they are harder to correct in a clinical setting. Moreover, a deviation is introduced for many tissue types, which results in a total number of parameters that are likely too large for the model to accurately predict.

As seen in Equations 3, 4, and 5, even though the spatial and angular deviations introduced are small, the DPB can detect them within a tenth of a millimeter and a tenth of a degree in the polar angle Θ . The error is greater in ϕ because the azimuthal angle leads to small beam deviations if Θ is small, and because it varies in a wider range from 0° to 180° .

The size and complexity of the models are an order of magnitude larger for SwinUNETR than for the UNet and SegResNet. The training time is approximately twice as long for SwinUNETR, and the inference time is approximately three times as long. Depending on the specific application, it might be sensible to trade off a minor loss in model performance to benefit from the computation efficiency of the smaller models. The SwinUNETR with the DPB did not introduce longer training times or increase model complexity, but it increased the model size three-fold compared to the original SwinUNETR. Although these times may vary for different image resolutions, it is worth noting that model training spans less than 30 minutes, and dose deposition images are inferred from the training images in less than 40 milliseconds.

As the training dataset with all considered cases is generated in around one hour and the model is trained in less than 30 minutes, we can readily produce a patient-specific, trained model in the approximately 24 hours between the CT acquisition, which is needed to simulate the dataset and train the model, and the first treatment session, when the model would be deployed to detect the dose deposition deviations.

Finally, there are some limitations to our study. Mainly, the input fed to the model is the total isotope activation induced by the beam instead of a PET-acquired activity image, which are the images that would be obtained in a clinical setting. In an activity image, the decay time of the different isotopes, the noising effect of the reconstruction, and the washout need to be accounted for. We plan to address these limitations in future work.

V. CONCLUSION

To summarize, we have trained three-dimensional convolutional neural networks to detect dose deposition deviations in a prostate treatment plan from the isotope activation induced by the proton beams. This work highlights the ability of deep neural networks to capture the underlying physics and the relationship between dose deposition and isotope activation. These images present an additional challenge because protons interact with tissue at different points throughout their path,

which results in correlated effects at different image regions. Moreover, we have implemented a Deviation Predicting Branch (DPB) that accurately predicts the spatial deviations that lead to deviations in the dose depositions.

The complete workflow could be executed between the patient CT acquisition and the first treatment session, and the model would be ready for deployment during, or after, the treatment session. By enabling proton dose reconstruction, it can diminish the risk of tumor undertreatment and surrounding tissue damage in proton therapy.

ACKNOWLEDGMENTS

This work was supported in part by the Prototwin Project (TED2021-130592B-I00), funded by the MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR. Additionally, Pablo Cabañes received support from the Complutense University Predoctoral Fellowship. We also extended our gratitude to FRED contributor Keegan McNamara for his guidance in the utilization of FRED and the implementation of the activation scorers

REFERENCES

- [1] H. Paganetti, "Proton therapy physics," *Proton Therapy Physics. Series: Series in Medical Physics and Biomedical Engineering*, ISBN: 978-1-4398-3644-6. CRC Press, Edited by Harald Paganetti, 12 2011.
- [2] R. Mohan and D. Grosshans, "Proton therapy - present and future." *Advanced drug delivery reviews*, vol. 109, pp. 26–44, 2017.
- [3] C.-L. Chang, K.-C. Lin, W.-M. Chen, B.-C. Shia, and S.-Y. Wu, "Comparing the oncologic outcomes of proton therapy and intensity-modulated radiation therapy for head and neck squamous cell carcinoma," *Radiotherapy and Oncology*, vol. 190, p. 109971, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167814023898656>
- [4] M. L. Bonù, L. Nicosia, A. Turkaj, E. Pastorello, P. Vitali, F. Frassine, C. Toraci, L. Spiazzi, M. Lechiara, B. Frittoli *et al.*, "High dose proton and photon-based radiation therapy for 213 liver lesions: a multi-institutional dosimetric comparison with a clinical perspective," *La radiologia medica*, pp. 1–10, 2024.
- [5] C. Chui, E. Yorke, and L. Hong, "The effects of intra-fraction organ motion on the delivery of intensity-modulated field with a multileaf collimator," *Medical physics*, vol. 30 7, pp. 1736–46, 2003.
- [6] O. Ates, F. Pirlpesov, J. Uh, C.-h. Hua, T. E. Merchant, A. Boria, A. M. Davidoff, D. E. Graetz, and M. J. Krasin, "Evaluating the impact of bowel gas variations for wilms' tumor in pediatric proton therapy," *Cancers*, vol. 16, no. 3, 2024. [Online]. Available: <https://www.mdpi.com/2072-6694/16/3/642>
- [7] H. Paganetti, "Range uncertainties in proton therapy and the role of monte carlo simulations," *Physics in Medicine and Biology*, vol. 57, pp. R99 – R117, 2012.
- [8] Y. Zhang, D. Boye, C. Tanner, A. J. Lomax, and A. Knopf, "Respiratory liver motion estimation and its effect on scanned proton beam therapy," *Physics in Medicine and Biology*, vol. 57, pp. 1779 – 1795, 2012.
- [9] W. Liu, S. J. Frank, X. Li, Y. Li, P. C. Park, L. Dong, X. Ronald Zhu, and R. Mohan, "Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers," *Medical physics*, vol. 40, no. 5, p. 051711, 2013.
- [10] C. Freijo, J. Herraiz, D. Sánchez-Parcerisa, and J. Udías, "Dictionary-based protoacoustic dose map imaging for proton range verification," *Photoacoustics*, vol. 21, 2021.
- [11] M. Ahmad, L. Xiang, S. Yousefi, and L. Xing, "Theoretical detection threshold of the proton-acoustic range verification technique," *Medical physics*, vol. 42 10, pp. 5735–44, 2015.

- [12] C. Richter, G. Pausch, S. Barczyk, M. Priegnitz, I. Keitz, J. Thiele, J. Smeets, F. Stappen, L. Bombelli, C. Fiorini, L. Hotoiu, I. Perali, D. Prieels, W. Enghardt, and M. Baumann, "First clinical application of a prompt gamma based in vivo proton range verification system," *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, vol. 118 2, pp. 232–7, 2016.
- [13] S. España, X. Zhu, J. Daartz, N. Liebsch, G. Fakhri, T. Bortfeld, and H. Paganetti, "Su-gg-j-149: Feasibility of in-room pet imaging for in vivo proton beam range verification," *Medical Physics*, vol. 37, pp. 3180–3180, 2010.
- [14] K. Parodi, H. Paganetti, H. Shih, S. Michaud, J. Loeffler, T. Delaney, N. Liebsch, J. Munzenrider, A. Fischman, A. Knopf, and T. Bortfeld, "Patient study of in vivo verification of beam delivery and range, using positron emission tomography and computed tomography imaging after proton therapy," *International journal of radiation oncology, biology, physics*, vol. 68 3, pp. 920–34, 2007.
- [15] V. Onecha, P. Galve, P. Ibáñez, C. Freijo, F. Arias-Valcayo, D. Sanchez-Parcerisa, S. España, L. Fraile, and J. Urdás, "Dictionary-based software for proton dose reconstruction and submillimetric range verification," *Physics in Medicine & Biology*, vol. 67, no. 4, p. 045002, 2022.
- [16] C. Liu, Z. Li, W. Hu, L. Xing, and H. Peng, "Range and dose verification in proton therapy using proton-induced positron emitters and recurrent neural networks (rnn)," *Physics in Medicine & Biology*, vol. 64, no. 17, p. 175009, 2019.
- [17] Z. Li, Y. Wang, Y. Yu, K. Fan, L. Xing, and H. Peng, "Machine learning approaches for range and dose verification in proton therapy using proton-induced positron emitters," *Medical physics*, vol. 46, no. 12, pp. 5748–5757, 2019.
- [18] A. U. Rahman, M. V. Nemallapudi, C.-Y. Chou, C.-H. Lin, and S.-C. Lee, "Direct mapping from pet coincidence data to proton-dose and positron activity using a deep learning approach," *Physics in Medicine & Biology*, vol. 67, no. 18, p. 185010, 2022.
- [19] X. Zhang, Z. Hu, G. Zhang, Y. Zhuang, Y. Wang, and H. Peng, "Dose calculation in proton therapy using a discovery cross-domain generative adversarial network (discogan)," *Medical physics*, vol. 48, no. 5, pp. 2646–2660, 2021.
- [20] Z. Wang, Q. She, and T. E. Ward, "Generative adversarial networks in computer vision: A survey and taxonomy," *ACM Computing Surveys (CSUR)*, vol. 54, no. 2, pp. 1–38, 2021.
- [21] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III* 18. Springer, 2015, pp. 234–241.
- [22] A. Myronenko, "3d mri brain tumor segmentation using autoencoder regularization," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II* 4. Springer, 2019, pp. 311–320.
- [23] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. Roth, and D. Xu, "Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images," 2022.
- [24] A. Schiavi, M. Senzacqua, S. Pioli, A. Mairani, G. Magro, S. Molinelli, M. Ciocca, G. Battistoni, and V. Patera, "Fred: a gpu-accelerated fast-monte carlo code for rapid treatment plan recalculation in ion beam therapy," *Physics in Medicine & Biology*, vol. 62, no. 18, p. 7482, 2017.
- [25] H.-P. Wieser, E. Cisternas, N. Wahl, S. Ulrich, A. Stadler, H. Mescher, L.-R. Müller, T. Klinge, H. Gabrys, L. Burigo *et al.*, "Development of the open-source dose calculation and optimization toolkit matrad," *Medical physics*, vol. 44, no. 6, pp. 2556–2568, 2017.
- [26] K. McNamara, A. Schiavi, D. Borys, K. Brzezinski, J. Gajewski, R. Kopeć, A. Rucinski, T. Skóra, S. Makkar, J. Hrbacek *et al.*, "Gpu accelerated monte carlo scoring of positron emitting isotopes produced during proton therapy for pet verification," *Physics in Medicine & Biology*, vol. 67, no. 24, p. 244001, 2022.
- [27] Y. Zhang, D. Boye, C. Tanner, A. J. Lomax, and A. Knopf, "Respiratory liver motion estimation and its effect on scanned proton beam therapy," *Physics in Medicine & Biology*, vol. 57, no. 7, p. 1779, 2012.
- [28] M. Moglioni, A. C. Kraan, G. Baroni, G. Battistoni, N. Belcari, A. Berti, P. Carra, P. Cerello, M. Ciocca, A. De Gregorio *et al.*, "In-vivo range verification analysis with in-beam pet data for patients treated with proton therapy at cnao," *Frontiers in Oncology*, vol. 12, p. 929949, 2022.
- [29] P. Andreo, "On the clinical spatial resolution achievable with protons and heavier charged particle radiotherapy beams," *Physics in Medicine & Biology*, vol. 54, no. 11, p. N205, 2009.
- [30] Y. LeCun, Y. Bengio *et al.*, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.
- [31] J. Moorthy and U. D. Gandhi, "A survey on medical image segmentation based on deep learning techniques," *Big Data and Cognitive Computing*, vol. 6, no. 4, p. 117, 2022.
- [32] R. Azad, E. K. Aghdam, A. Rauland, Y. Jia, A. H. Avval, A. Bozorgpour, S. Karimijafarbigloo, J. P. Cohen, E. Adeli, and D. Merhof, "Medical image segmentation review: The success of u-net," *arXiv preprint arXiv:2211.14830*, 2022.
- [33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [34] E. U. Henry, O. Emebob, and C. A. Omonhinmin, "Vision transformers in medical imaging: A review," *arXiv preprint arXiv:2211.10043*, 2022.
- [35] V. Braunstein. (2021, Sep.) Nvidia data scientists take top spots in miccai 2021 brain tumor segmentation challenge. NVIDIA. Accessed: 2024-04-15. [Online]. Available: <https://developer.nvidia.com/blog/nvidia-data-scientists-take-top-spots-in-miccai-2021-brain-tumor-segmentation-challenge/>
- [36] F.-A. Croitoru, V. Hondru, R. T. Ionescu, and M. Shah, "Diffusion models in vision: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.
- [37] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," *arXiv preprint arXiv:1505.00853*, 2015.
- [38] M. J. Cardoso, W. Li, R. Brown, N. Ma, E. Kerfoot, Y. Wang, B. Murrey, A. Myronenko, C. Zhao, D. Yang *et al.*, "Monai: An open-source framework for deep learning in healthcare," *arXiv preprint arXiv:2211.02701*, 2022.
- [39] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems*, vol. 32, 2019.
- [40] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.
- [41] L. B. de Amorim, G. D. Cavalcanti, and R. M. Cruz, "The choice of scaling technique matters for classification performance," *Applied Soft Computing*, vol. 133, p. 109924, 2023.
- [42] S. Martinot, N. Komodakis, M. Vakalopoulou, N. Bus, C. Robert, E. Deutsch, and N. Paragios, "Differentiable gamma index-based loss functions: Accelerating monte-carlo radiotherapy dose simulation," in *International Conference on Information Processing in Medical Imaging*. Springer, 2023, pp. 485–496.
- [43] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions," *Medical physics*, vol. 25, no. 5, pp. 656–661, 1998.
- [44] S. Biggs, M. Jennings, S. Swerdloff, P. Chlap, D. Lane, J. Rembish, J. McAloney, P. King, R. Ayala, F. Guan *et al.*, "Pymedphys: A community effort to develop an open, python-based standard library for medical physics applications," *Journal of Open Source Software*, vol. 7, no. 78, p. 4555, 2022.