
AI as the future of Monte Carlo modelling

February 19th, 2026
EURADOS Annual Meeting
Helsinki, Finland

Chi-Hieu Pham & Julien Bert
LaTIM – INSERM UMR1101
University of Brest



Laboratoire de Traitement
de l'Information Médicale
Laboratory of Medical
Information Processing



Thank EURADOS and UBO DAI for funding of this presentation

Monte Carlo simulation & radiotherapy

Radiation therapy

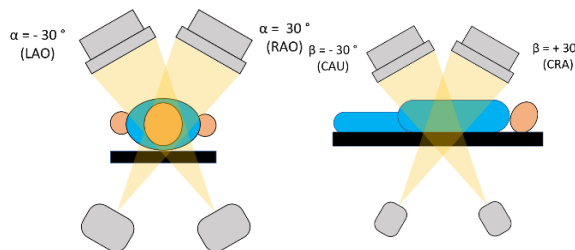
- Treatment : ionizing radiation to control malignant cells
- Minimizing side effects
- **Treatment planning system :**
 - Dose calculation
 - Simulating interactions : radiation beams & tissue

Monte Carlo simulation

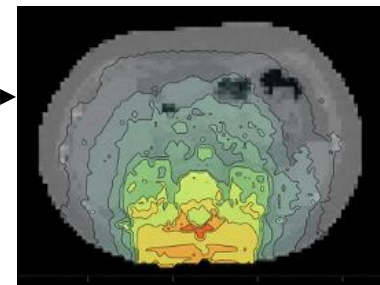
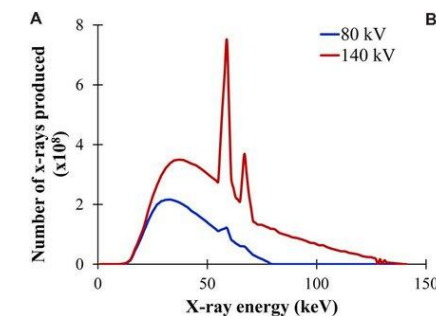
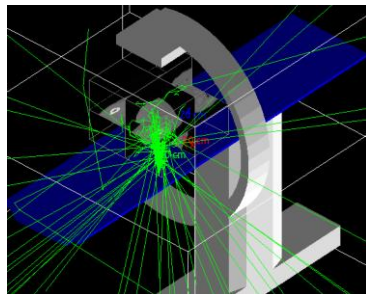
- Simulating particle by particle
- Rules of physics (Geant4)



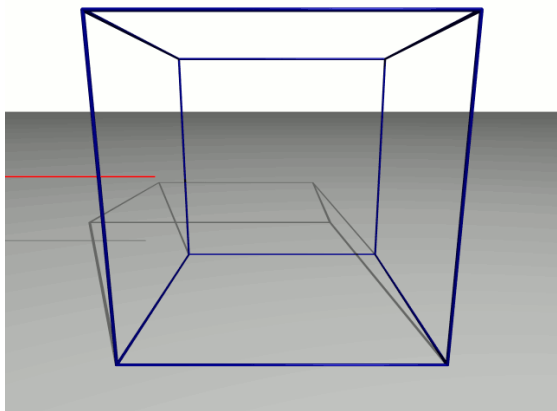
<https://github.com/OpenGATE>



CHU Brest



Dose model : Monte Carlo simulation



Monte Carlo simulation :

- Stochastic model
- Accurate
- Robust (no limited domain)

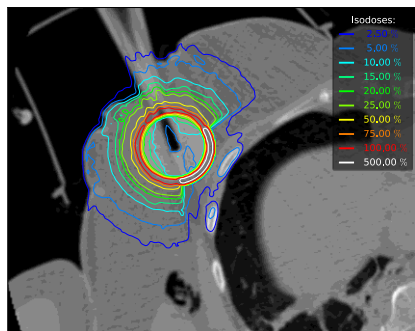
Law of large numbers

$$\epsilon = f\left(\frac{1}{\sqrt{N}}\right)$$

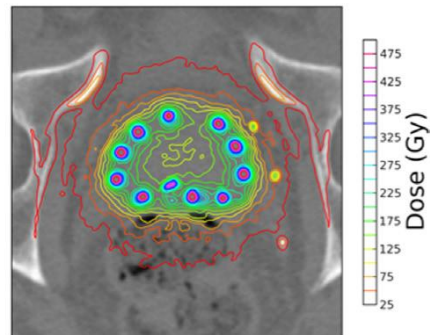
○ Time calculation

- Patient workflow
- Workspace
- Financial burden
- Availability issue

Intra-Operative
Radiation
Therapy (IORT)
20 hours
One CPU core

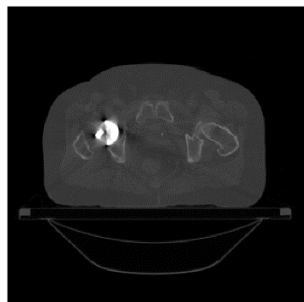


Brachytherapy
15 hours
One CPU core

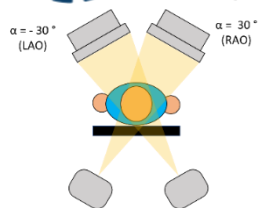
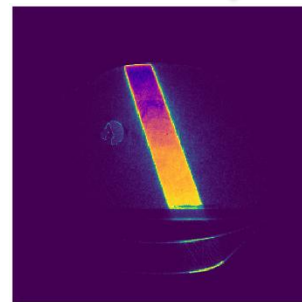
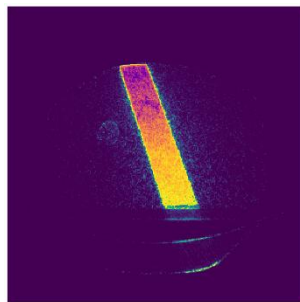
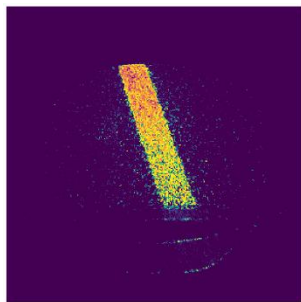


MC simulation for external beam radiation therapy

CT scan
2mm x 2mm x 2mm



Denoising : Low sampling to high sampling
○ *ignore beam parameters*



2×10^5 photons
 ≈ 70 seconds

2×10^6 photons
 ≈ 12 minutes

2×10^7 photons
 ≈ 2 hours

AI regression task :

- Speed
- Same accuracy ?

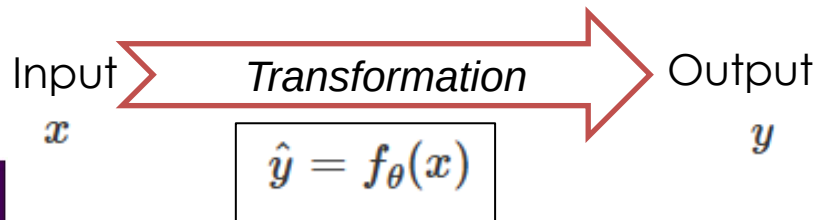


Specifications of
the Novalis
system (VARIAN)

Fast-denoising of MC simulation

2×10^5 photons

2×10^7 photons



Learning method f :

- Training dataset $\mathcal{D} = \{(x^{(i)}, y^{(i)})\}_{i=1}^N$
- Large number of samples N
- “Good transformation”

Loss function : $\mathcal{L}(\hat{y}, y)$

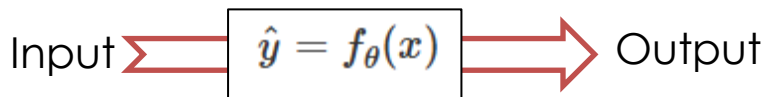
- Minimizing the error
 - Mean squared error MSE
 - Mean absolute error MAE (better)
 - Robust to outliers

$$\hat{y} \approx y$$

$$\text{MSE} = \frac{1}{WHD} \sum_{w=1}^W \sum_{h=1}^H \sum_{d=1}^D (y_{w,h,d} - \hat{y}_{w,h,d})^2$$

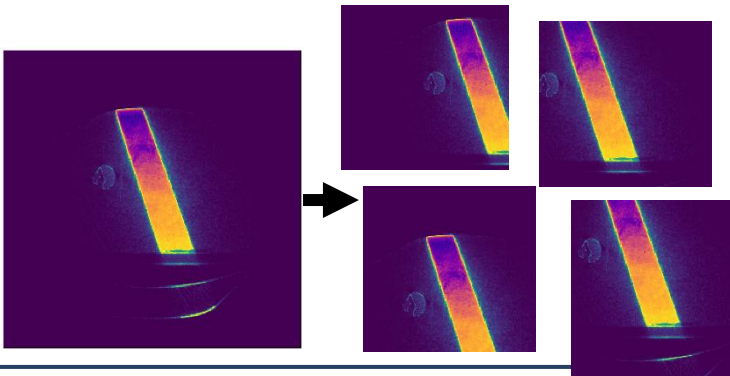
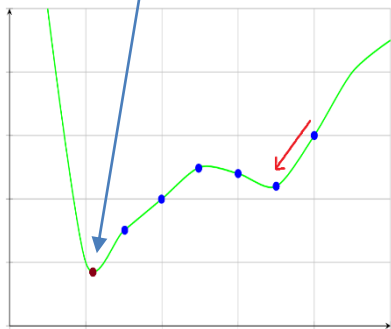
$$\text{MAE} = \frac{1}{WHD} \sum_{w=1}^W \sum_{h=1}^H \sum_{d=1}^D |y_{w,h,d} - \hat{y}_{w,h,d}|$$

Fast-denoising of MC simulation



- Large number of samples N
- “Good transformation”
 - Optimized parameters

$$\hat{\theta} = \arg \min \mathcal{L}(\hat{y}, y)$$



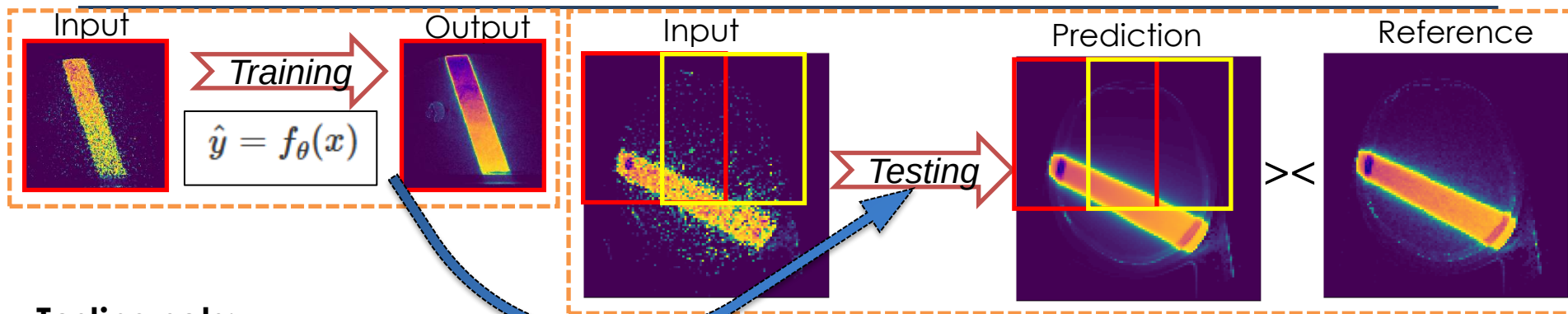
Iterative optimization : $\theta^{(new)} = \theta^{(old)} - \alpha \nabla_{\theta} \mathcal{L}(\theta^{(old)})$

- Stochastic gradient descent SGD
 - First order derivative ∇
 - Learning rate α
- Minibatch SGD (*batch*)
 - One update = iteration
 - Complete dataset = epoch
 - Several epochs
- Variants : Adam, AdamW

Training tricks:

- Patches of 3D images
 - Overcome hardware limits
 - 128x128x128 or 64x64x64, ...
- Data augmentation
 - Rotation, flip, ...
 - Avoiding overfitting

Performance evaluation of learning methods

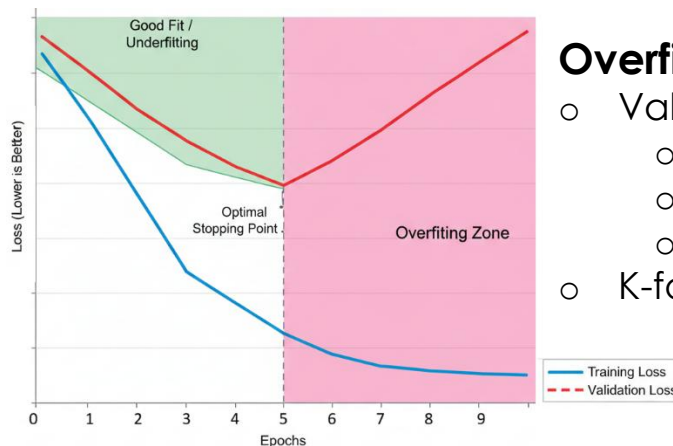


Testing sets:

- Independent to training set
- Sliding window inference

Metrics: Measuring differences

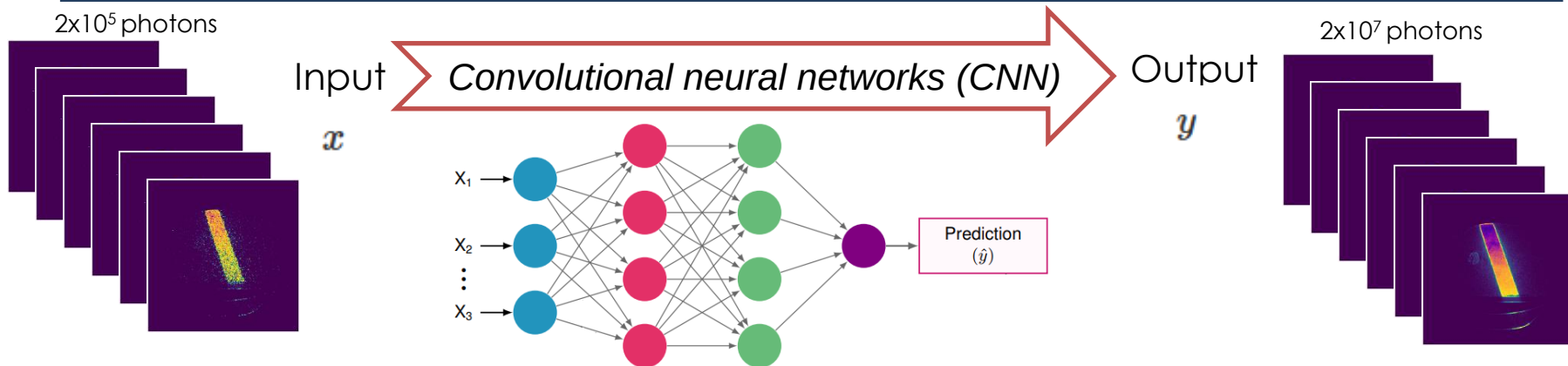
- Root Mean Squared Error (RMSE)
- Peak Signal-to-Noise Ratio (PSNR)
- Gamma Index, Gamma Passing Rate (GPR)
- ...



Overfitting :

- Validation set :
 - Ex. Ratio 7:2:1
 - (Train:Valid:Test)
 - Early stopping !
- K-fold cross-validation

Deep learning for fast-denoising



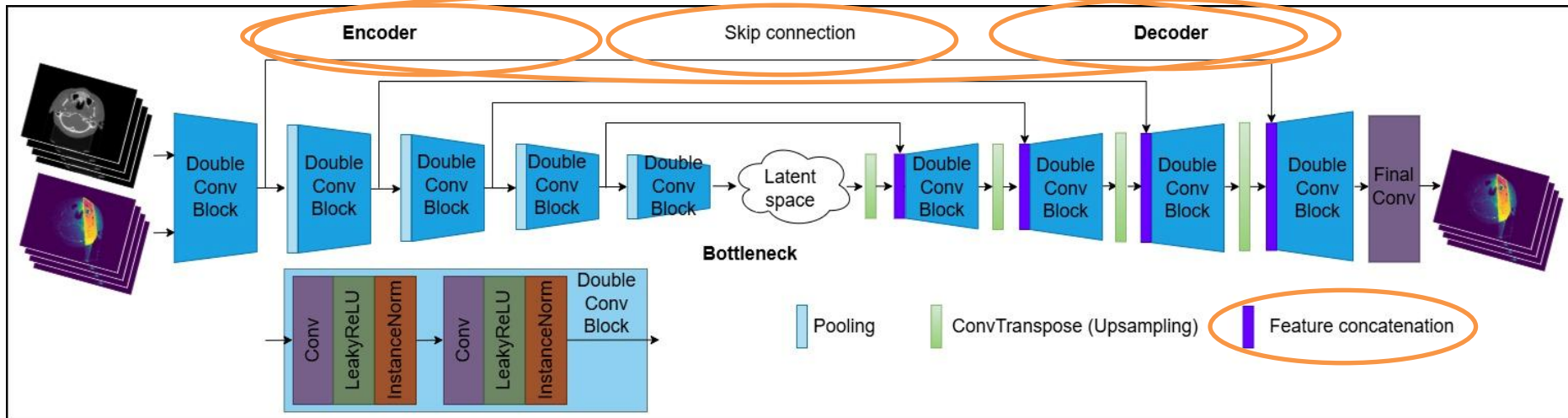
CNN :

- Convolution layers
- Activation functions
 - Rectifier Linear Unit (ReLU)
- Pooling
- Transposed convolution layers
- Normalization layers
- ...

Dimension :

- 2D : lack of inter-slice information
- 3D networks :
 - Spatial context
 - Voxel relationship

Baseline : UNet 3D



Baseline for medical imaging :

- segmentation
- regression
 - CT-to-MRI synthesis,
 - dose prediction

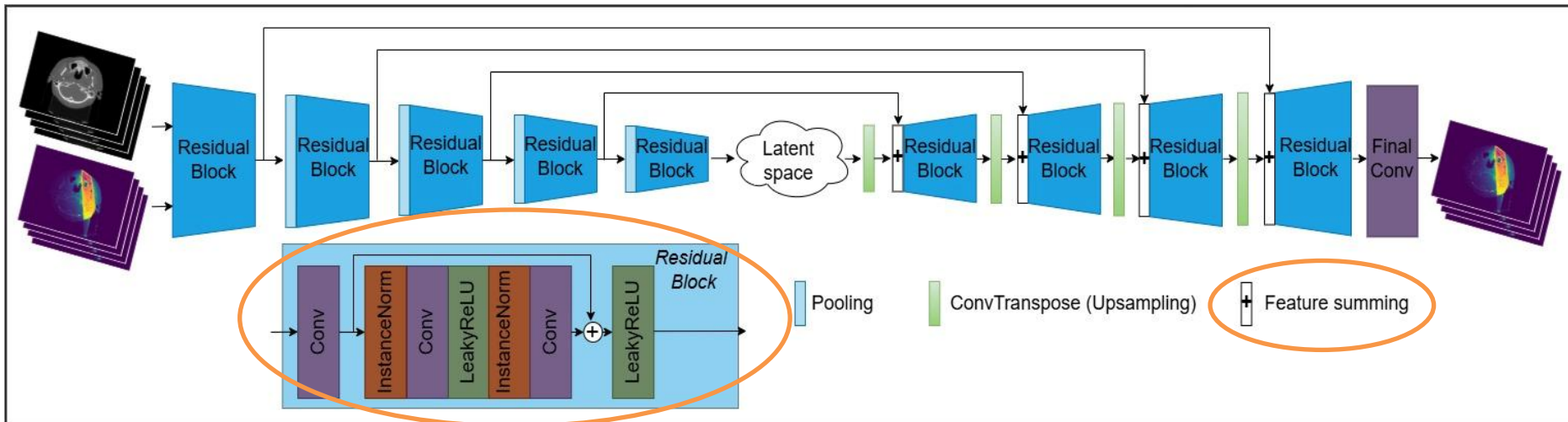
Symmetrical Encoder-Decoder :

- Encoder :
 - feature extractor (*what*)
 - edges, shapes, textures
- Decoder :
 - feature localization (*where*)
 - feature fusion

Skip Connections :

- feature concatenation
- between encoder and decoder
- high-frequency, spatial information
- preventing information loss
 - the sharp edges

Residual block : Residual UNet 3D



- Direct mapping by convolutional blocks :

$$y = F_{direct}(x)$$

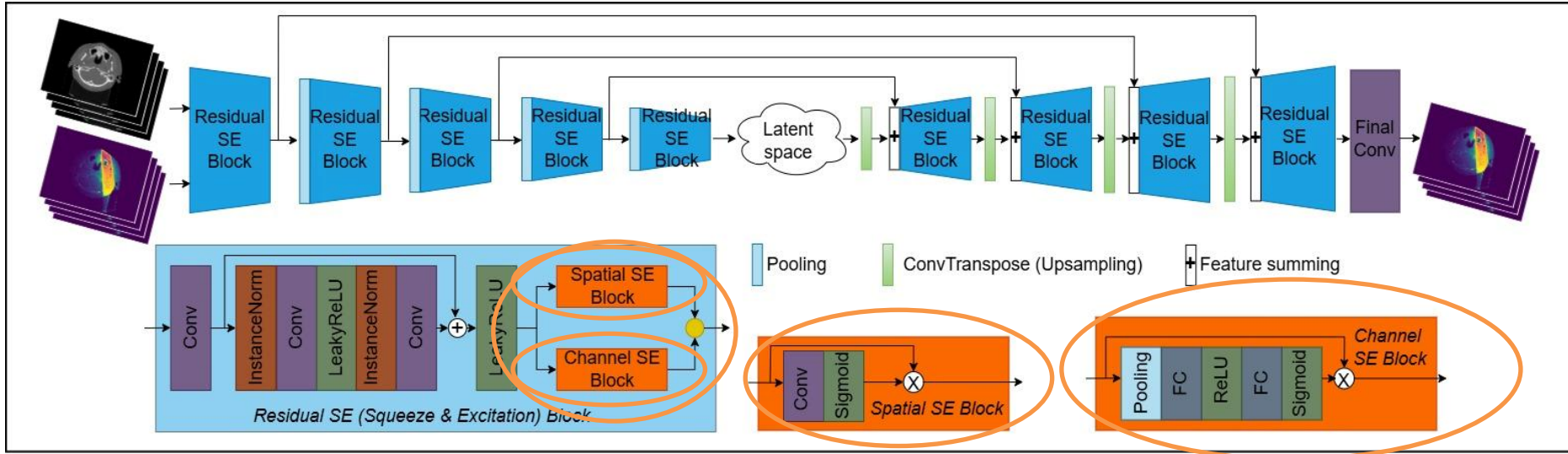
- Replaced by **Residual Blocks** (from ResNet):

$$y = F_{residual}(x) + x$$

Advantages :

- Enhancing gradient flow
 - easier to train network without performance degradation
 - avoiding gradient vanishing
 - faster convergence
- Improving feature propagation
 - features learned in encoder are preserved and refined rather than re-learned

Squeeze-and-Excitation : Residual UNet SE 3D



Channel Squeeze-and-Excitation (cSE) :

- the most useful feature channels (e.g., edges, texture)

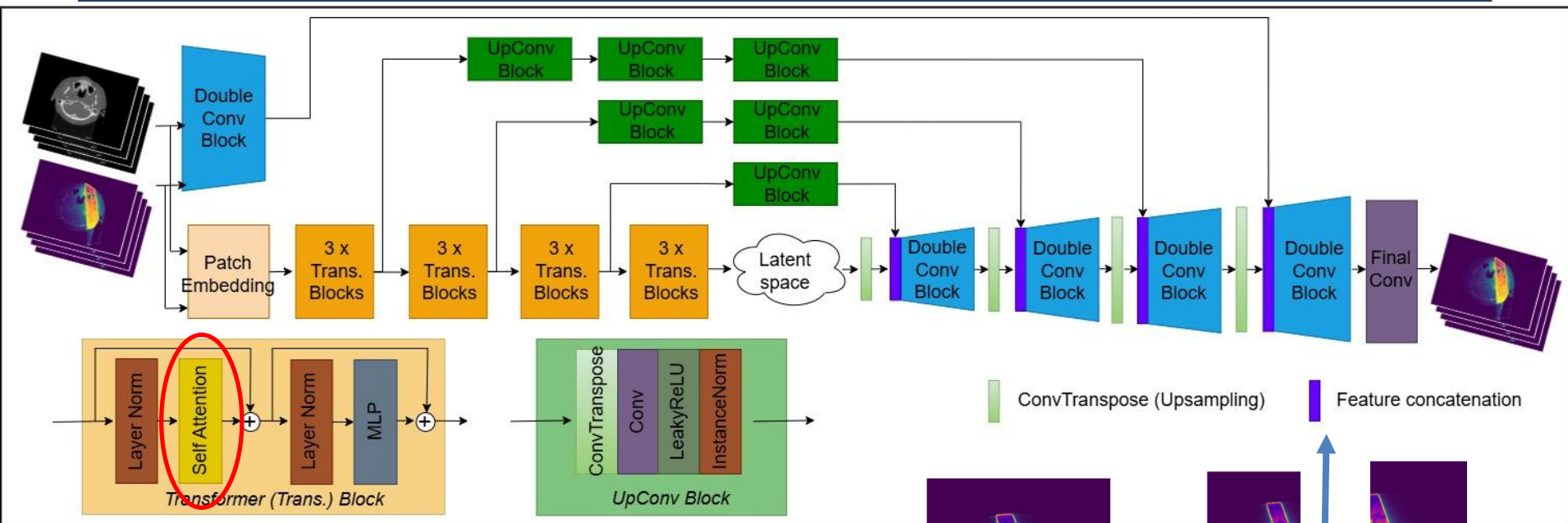
Spatial Squeeze-and-Excitation (sSE) :

- the most important specific voxels in the feature

Spatial Channel SE :

- simultaneously which features are important and where they are located

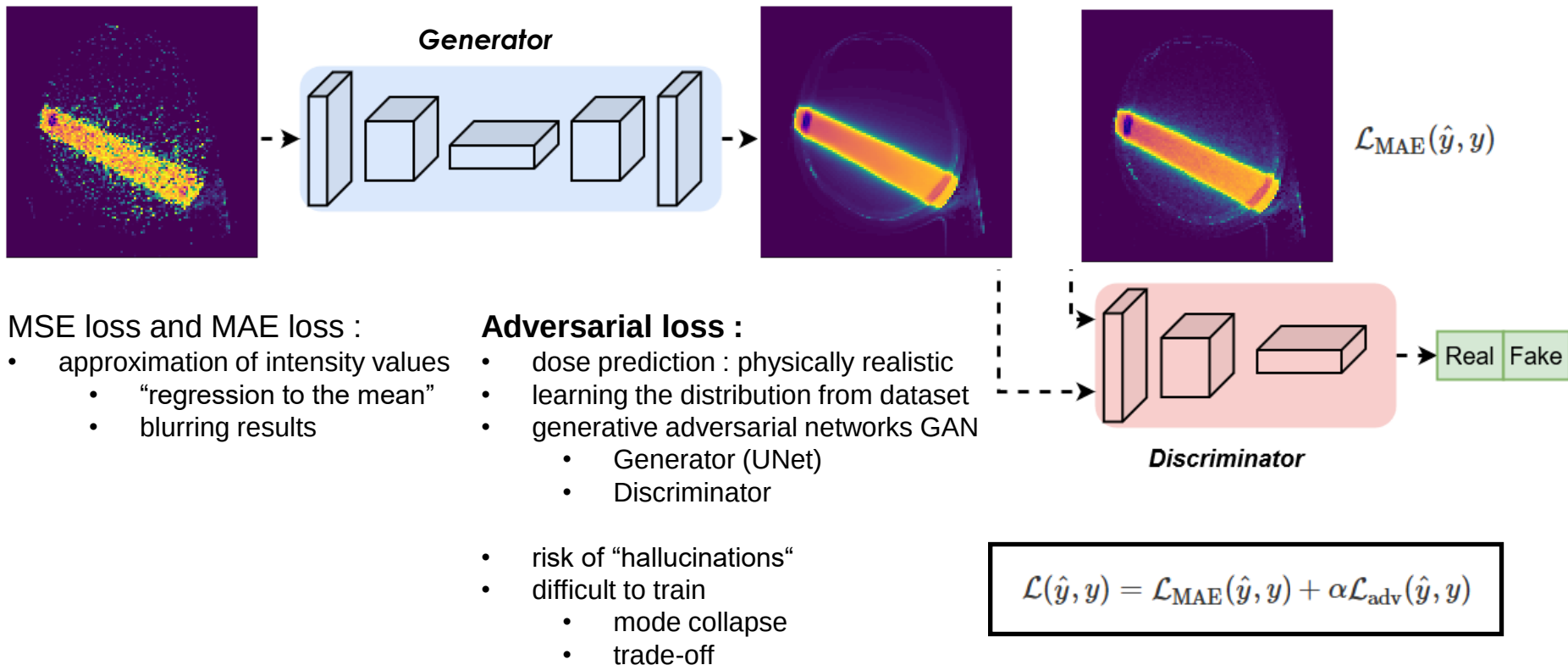
UNETR : Transformer blocks for encoder



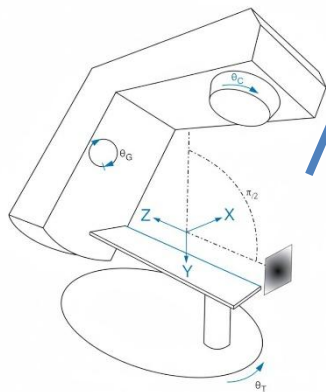
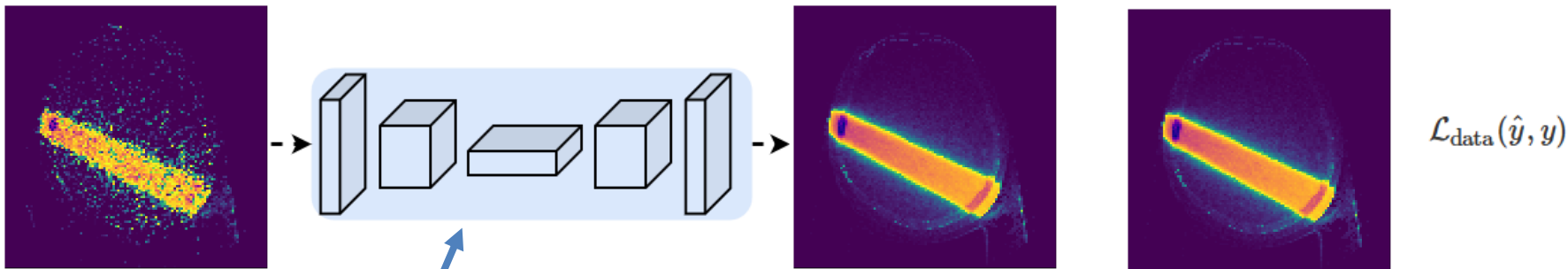
Transformer blocks :

- Self-Attention :
 - used in Large Language Models (LLMs)
 - "global context" relationships between distant parts

Perceptual quality : Adversarial loss



Physics-informed neural networks



Physics-Informed Neural Networks :

- Data loss:
 - dose matching
- Physics loss:
 - specific physical properties of radiation
- Physics-informed priors
 - normalized distance-aware beam plates
 - mass density map

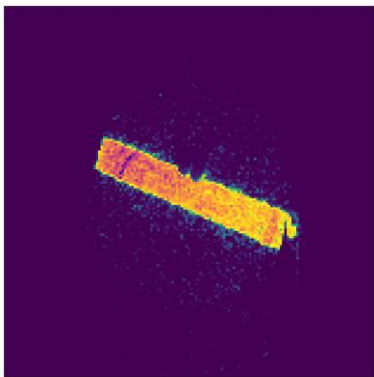
$$\mathcal{L}(\hat{y}, y) = \mathcal{L}_{\text{data}}(\hat{y}, y) + \mathcal{L}_{\text{physics}}(\sigma)$$

Energy-shifting

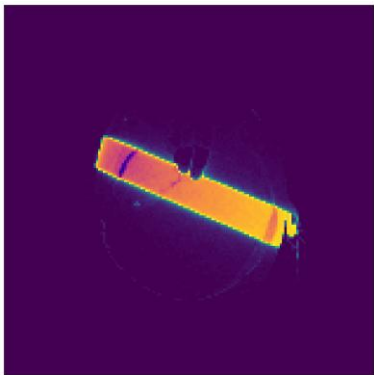
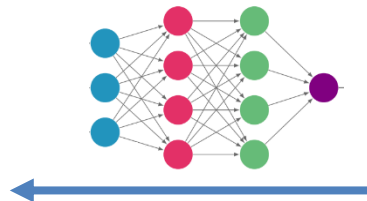
Specifications of
the Novalis
system (VARIAN)



Denoising

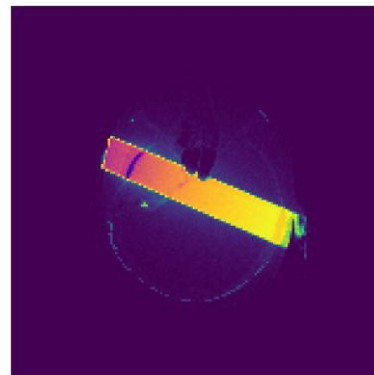


2×10^5 photons
LINAC beam
one CPU
 ≈ 70 seconds



2×10^7 photons
LINAC beam
one CPU
 ≈ 2 hours

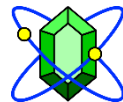
2×10^7 photons
500keV
mono beam
one GPU
 ≈ 5 seconds



Energy-shifting learning model :

- Predict poly-energetic beam from fast mono-energetic
 - Consistency of particles
 - Texture preservation
 - Focus only on intensity change
 - Fast dose calculation

GPU **GE**ant4-based
Monte Carlo **S**imulations
<http://www.ggems.fr>



Results of energy-shifting



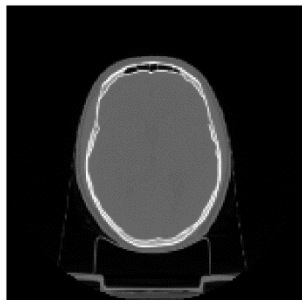
Same models :

- Baseline UNet3D
- MAE loss

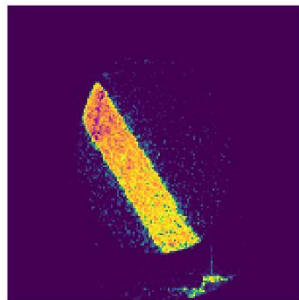
Intra-domain sets :

- Training on Head
- Testing on Head

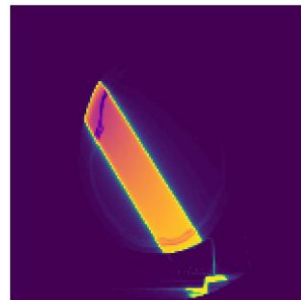
- Voxel size :
- 2mm x 2mm x 2mm



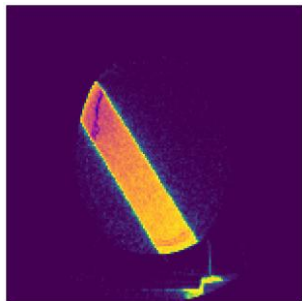
OpenGATE
LINAC dose
on CPU
1,6 minutes
 2×10^5 photons



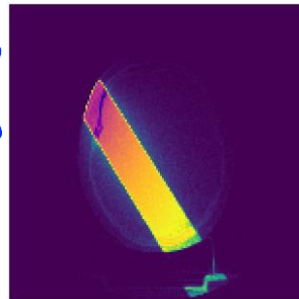
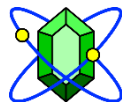
Denoising
on GPU
+ 9 seconds



PSNR : 58.30, GPR: 99.72%

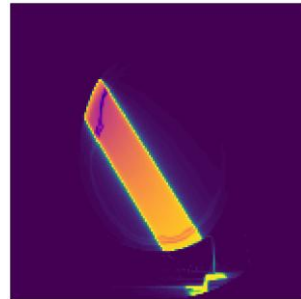


OpenGATE
LINAC dose
on CPU
2,6 hours
 2×10^7 photons



GGEMS
Mono beam
on GPU
5 seconds

Energy-Shifting
on GPU
+ 9 seconds

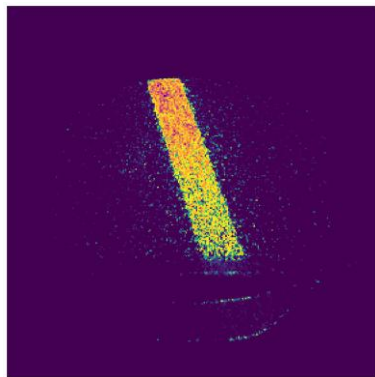


PSNR : 58.93, GPR: 99.83%

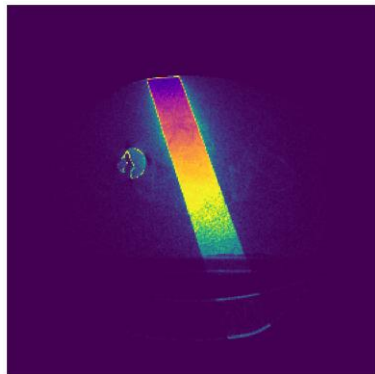
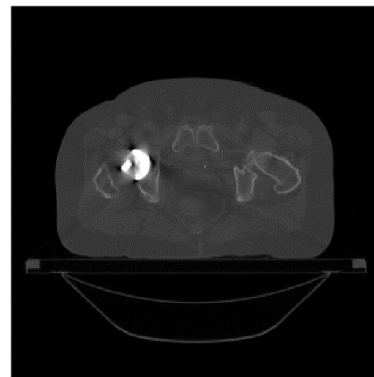
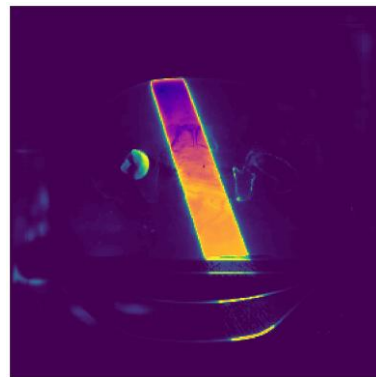
Results of energy-shifting

Extra-domain sets :

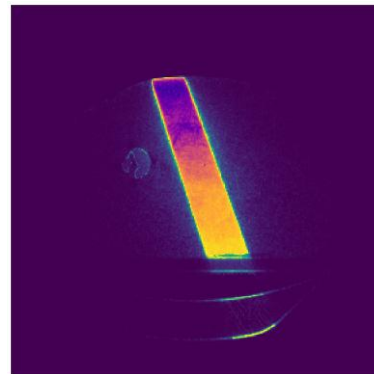
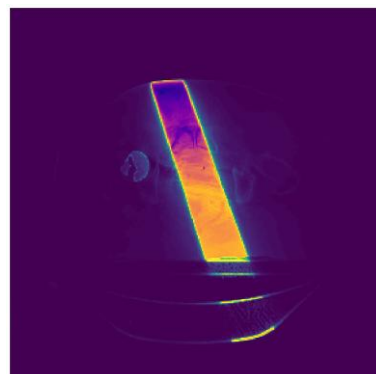
- Training on Head
- Testing on Pelvis
- Energy-shifting more robust on unseen data than denoising



Denoising



Energy-shifting



Perspective

- Application to other types of radiotherapy :
 - Carbon Ion Radiotherapy (CIRT)
- Clinical validation
- Model compression :
 - Inference all on CPU
- Semi-supervised learning or one-shot learning :
 - Transfer learning from large dataset to small dataset
- Generative models :
 - Diffusion models

Reference

- Sarrut, D., Baudier, T., Bert, J., Chatzipapas, K., Favaretto, M., Fuchs, H., ... & Krah, N. (2026). GATE 10 Monte Carlo particle transport simulation: I. Development and new features. *Physics in Medicine & Biology*, 71(1), 015042.
- Bert, J., Perez-Ponce, H., Bitar, Z. E., Jan, S., Boursier, Y., Vintache, D., ... & Visvikis, D. (2013). Geant4-based Monte Carlo simulations on GPU for medical applications. *Physics in medicine and biology*, 58(16), 5593-5611.
- Vazquez-Quino, L. A., Massingill, B., Shi, C., Gutierrez, A., Esquivel, C., Eng, T., ... & Stathakis, S. (2012). Monte Carlo modeling of a Novalis Tx Varian 6 MV with HD-120 multileaf collimator. *Journal of applied clinical medical physics*, 300-313.
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016, October). 3D U-Net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention* (pp. 424-432). Cham: Springer International Publishing.
- Delannoy, Q., Pham, C. H., Cazorla, C., Tor-Díez, C., Dollé, G., Meunier, H., ... & Rousseau, F. (2020). SegSRGAN: Super-resolution and segmentation using generative adversarial networks—Application to neonatal brain MRI. *Computers in Biology and Medicine*, 120, 103755.
- Maier, J., Eulig, E., Dorn, S., Sawall, S., & Kachelrieß, M. (2018, November). Real-time patient-specific CT dose estimation using a deep convolutional neural network. In *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)* (pp. 1-3). IEEE.

Reference

Harb, H., Villa, M., Benoit, D., Pham, C. H., Nasr, B., & Bert, J. (2025). Fast operating room scattered radiation calculation in x-ray guided interventions by using deep learning. *Journal of Radiological Protection*, 45(4), 041524.

Kontaxis, C., Bol, G. H., Lagendijk, J. J. W., & Raaymakers, B. W. (2020). DeepDose: Towards a fast dose calculation engine for radiation therapy using deep learning. *Physics in Medicine & Biology*, 65(7), 075013.

Wolny, A., Cerrone, L., Vijayan, A., Tofanelli, R., Barro, A. V., Louveaux, M., ... & Kreshuk, A. (2020). Accurate and versatile 3D segmentation of plant tissues at cellular resolution. *Elife*, 9, e57613.

Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., ... & Xu, D. (2022). Unetr: Transformers for 3d medical image segmentation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision* (pp. 574-584).

Villa, M., Nasr, B., Benoit, D., Padoy, N., Visvikis, D., & Bert, J. (2023). Fast dose calculation in x-ray guided interventions by using deep learning. *Physics in Medicine & Biology*, 68(16), 164001.

Zhang, X., Zhang, H., Wang, J., Ma, Y., Liu, X., Dai, Z., ... & Li, Q. (2023). Deep learning-based fast denoising of Monte Carlo dose calculation in carbon ion radiotherapy. *Medical Physics*, 50(12), 7314-7323.

Zhang, J., Bousse, A., Pham, C. H., Shi, K., & Bert, J. (2026). Semi-supervised learning for dose prediction in targeted radionuclide therapy: a synthetic data study. *Physics in Medicine and Biology*.