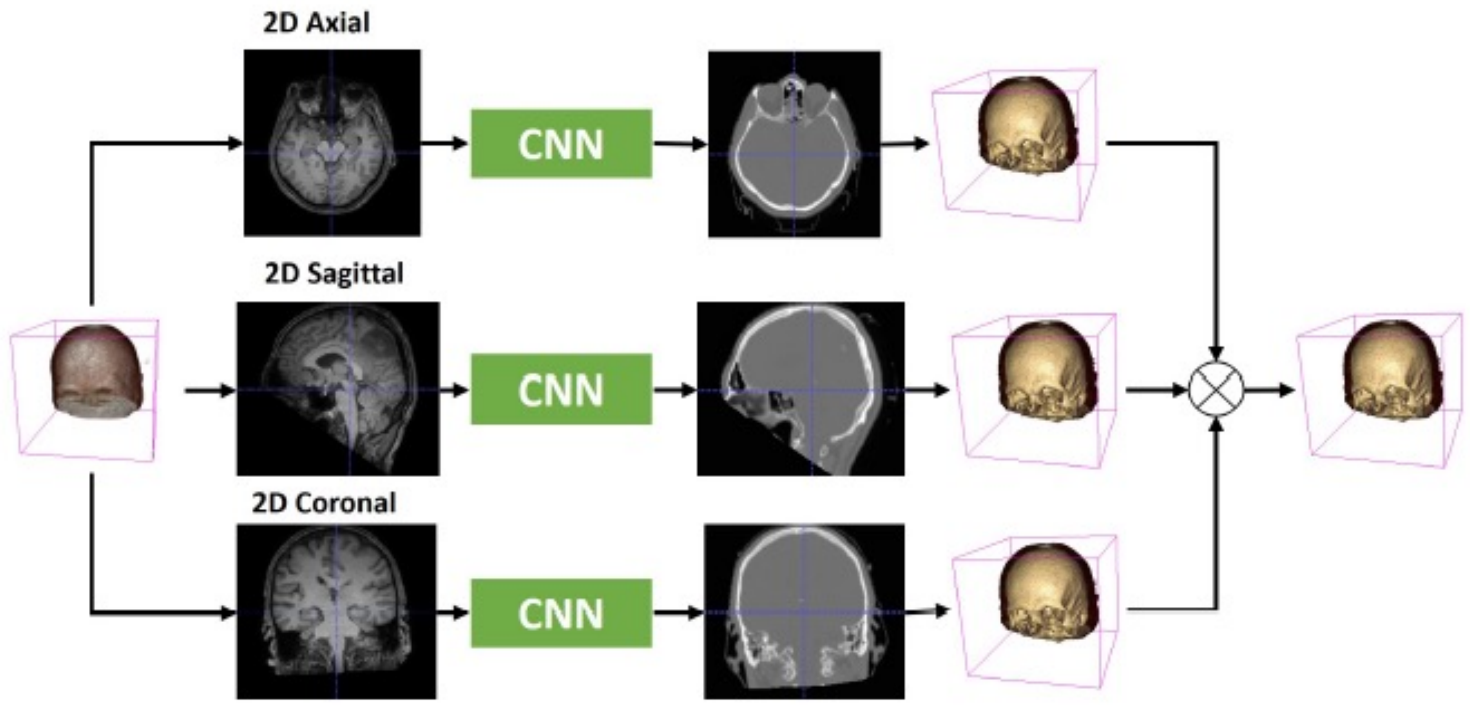
Multi-Planar Convolutional Neural Networks for MRI and CBCT to CT Translation

Gustav Müller-Franzes[1,], Firas Khader[1,*], Daniel Truhn[1]*

[1]Department of Diagnostic and Interventional Radiology, University Hospital Aachen, Aachen, Germany
* equal contribution

Introduction

- Computed tomography (CT) is important for radiotherapy due to its ability to provide accurate dose calculations.
- Cone Beam CT (CBCT) is gaining popularity in radiotherapy for image-guided adaptive radiation therapy (IGART) but suffers from image quality issues
- Magnetic resonance imaging (MRI) offers better soft-tissue contrast without radiation exposure, making it promising for tumor delineation.
- The SynthRAD 2023 challenge aims to provide datasets for researchers to develop machine learning models that convert MRI to CT images for MRI-only radiation therapy (Task 1) and CBCT to CT images for CBCT-only IGART (Task 2).



Multi-Planar CNN: 3D Input is separated along the axial, sagittal and coronal axes into 2D slices. Each axis and slice is feed through a convolutional neural network. Predictions are stacked as 3D volumes and averaged among between all three planes.

Material & Methods

- **Dataset**
  - Task 1 included 180 brain and 180 pelvis MR-CT paired images.
  - Task 2 included 180 brain and 180 pelvis CBCT-CT paired images.
  - Each MR-CT and CBCT-CT pair was accompanied by a binary mask outlining the brain or pelvis location.
- **Preprocessing**
  - Signal intensity of MR images was clipped to the 99th percentile within the binary mask area and normalized to a range between -1 and 1
  - The signal intensity of the CBCT images was subtracted by minimum, clipped between 0 and 3000, and scaled from -1 to 1.
  - CT images were clipped between -1024 and 3000 and rescaled to [-1, 1].
  - All 3D images were converted to 2D images stored in the axial, coronal, and sagittal planes.
- **Training**
  - Training was conducted separately for both tasks and for predicting brain and pelvis regions.
  - Each task's dataset (n=180 images) was divided into a training set (n=162) and validation set (n=18) with a 9-to-1 ratio.
  - Images were padded to a size divisible by 8.
  - The sum of Mean Absolute Error (MAE) and Structural Similarity Index Measure (SSIM) was used as loss function
  - AdamW optimizer with a learning rate of $\alpha = 10^{-4}$ was used, and early stopping was employed based on the loss
- **Inference**
  - 3D input images were divided into individual slices along axial, coronal, and sagittal planes.
  - Test-time augmentation (flipping) was applied and resulting predictions were averaged.
  - Model weights corresponding to the best three checkpoints with the lowest validation loss were used in an ensemble, where predictions of the models were averaged.
  - 3D volumes corresponding to axial, sagittal, and coronal predictions were averaged to obtain the final prediction for synthetic CT.

Architecture (CNN)

- The multi-planar convolutional neural network (CNN) consisted of three (axial, sagittal, coronal) identical, fully convolutional, neural networks with a symmetric encoder-bottleneck-decoder design.
- The encoder had a 64-channel input layer and three convolutional downsampling layers, each halving the image size.
- The bottleneck layer had nine convolutional layers with 512 channels.
- The decoder consisted of three transposed convolutional upsampling layers, followed by an output layer.

Results – Intern

- **Single plane Performance**
  - The network's performance in coronal and sagittal plane was notably lower compared to the axial plane for the pelvis region, which was probably attributed to anisotropic voxel spacing (1 × 1 × 2.5mm).
  - The network's performance for the brain region showed comparable performance for all three planes, which was probability attributed to isotropic voxel spacing (1 × 1 × 1mm).

	Task 1				Task 2			
	Brain		Pelvis		Brain		Pelvis	
	MAE	SSIM	MAE	SSIM	MAE	SSIM	MAE	SSIM
2D Axial	73.9	0.92	54.1	0.87	55.9	0.94	59.9	0.85
2D Coronal	75.4	0.92	62.3	0.84	58.5	0.93	NA	NA
2D Sagittal	72.1	0.92	62.9	0.84	59.2	0.93	NA	NA

- **Test-time augmentation & checkpoint ensembling**
  - Test-time augmentation and ensembling improved the performance across all planes, tasks, and regions by approximately -4 MAE

	Task 1				Task 2			
	Brain		Pelvis		Brain		Pelvis	
	MAE	SSIM	MAE	SSIM	MAE	SSIM	MAE	SSIM
2D Axial	69.5	0.926	51.4	0.867	52.6	0.941	57.4	0.860
2D Coronal	68.2	0.928	NA	NA	55.3	0.938	NA	NA
2D Sagittal	67.4	0.929	NA	NA	56.0	0.937	NA	NA

- **Multi-planar ensemble**
  - The ensembling of the predictions of all three planes increase the performance further by approximately -3 MAE for the brain region but not for the pelvis (* without test-time augmentation)

	Task 1				Task 2			
	Brain		Pelvis		Brain		Pelvis	
	MAE	SSIM	MAE	SSIM	MAE	SSIM	MAE	SSIM
Multi-planar	64.4	0.933	53.6(*)	0.862(*)	51.4	0.944	NA	NA

Results – Challenge

- The multi-planar approach ranked 1 to 4 in for MAE, SSIM, and Peak Signal-to-Noise Ratio (PSNR) in Task 1 and 2 on the Post-SynthRAD challenge (accessed Sept. 24, 2023).

	Task 1			Task 2		
	MAE	PSNR	SSIM	MAE	PSNR	SSIM
Multi-planar	60.7 ± 13.6 (2)	29.4 (2)	0.883 (2)	51.3 ± 11.7 (4)	31.0 (1)	0.910 (1)

Conclusion

- A simple encoder-decoder 2D CNN provides highly competitive results.
- Multi-planar ensembling improves the performance if the spacing is isotropic

UNIKLINIK RWTHAACHEN

MICCAI 2023 Vancouver CANADA

SynthRAD2023 Synthesizing computed tomography for radiotherapy