RAPTOR+ Project 16 Proposal Scaffold: AI-driven Anatomical and Response Adapted Proton Therapy

Role: Senior Medical Physicist & Deep Learning Researcher (Adaptive Proton Therapy)

January 15, 2026

Executive Synopsis

Current adaptive radiotherapy largely compensates for geometric/anatomical change (tumor shrinkage, weight loss) but not biological response (tissue composition change, early necrosis, hypoxia). We outline a physics-informed, multi-stream deep learning pipeline to: (i) generate proton-ready synthetic CTs (sCT) from CBCT; (ii) model expected anatomy via inverse-consistent deformable registration of the planning CT; and (iii) learn residuals and texture/functional signatures that indicate biology-driven change, validated with dose-impact analyses using Monte Carlo in proton therapy.

1 SECTION 1: Academic State-of-the-Art (2020–2025 focus)

A. sCT from CBCT for proton therapy and adaptive RT

- Deep learning CBCT-to-CT correction for dose recalculation in adaptive RT (unpaired and paired translation). Representative studies demonstrate clinically usable HU fidelity and dosimetric agreement for head-and-neck within $\sim$ 1–2% dose differences and < 2 mm range error:
  - Maspero et al. (2020): deep-learning sCT generation for on-line ART with robustness to scanner/vendor differences; shows improved HU accuracy and downstream dose agreement in photon RT; methods translate to protons with RSP-aware losses.
  - Kurz et al. (2020): CBCT correction pipeline for online ART; demonstrated improved CBCT HU and dose fidelity; discussed extension to proton settings.
  - Thummerer et al. (2020, 2022): head-and-neck CBCT-to-CT synthesis with GANs evaluated for adaptive proton dose recalculation; showed small deviations in D_{95}/D_{mean} for targets and OARs when using sCT.
  - Liu et al. (2021): CycleGAN-based CBCT correction with gradient-consistency/information losses; significant reduction in HU bias and improved plan recalculation metrics.
  - Jiang et al. (2021): RegGAN/registration-regularized GAN for CBCT$\rightarrow$CT with reduced misalignment artifacts; improved dosimetric agreement vs. vanilla CycleGAN.
  - Veiga et al. (2016): foundational CBCT-based proton dose recalculation with physics corrections; important baseline for DL improvements in later works.
- Generative trends: while CycleGAN/Pix2Pix dominate 2020–2022, diffusion models (Ho et al., 2020) have gained traction for HU-faithful CT restoration, producing fewer hallucinations and better calibrated intensity distributions—crucial for proton stopping-power mapping.

B. Distinguishing anatomical vs. biological change (DIR residuals, radiomics, multi-modality)

- DIR residuals as biomarkers: Regions where inverse-consistent, diffeomorphic registration (Avants et al., 2008; Klein et al., 2010) fails to reconcile intensity/shape changes often correlate with biological processes, not just geometrical displacement. Quantities like the Jacobian determinant, inverse-consistency error, and local image dissimilarity (e.g., MIND/NGF) highlight non-deformable changes.
- Delta-radiomics and response: Temporal texture/intensity features on CT/CBCT/MRI/PET have been associated with early response (Traverso et al., 2019). Fusing DIR-derived mechanics (e.g., strain, bending energy) with delta-radiomics improves the discrimination of biology vs. anatomy.
- Biomechanical/DIR-informed modeling: Using biomechanical regularizers or patient-specific finite-element approximations during registration yields deformation fields that better capture pure anatomy; residual intensity patterns after warping are enriched for biology-driven change.

C. Dosimetric impact: sCT versus planning CT for protons

- Studies report that DL-corrected sCTs can support proton dose recalculation with median target coverage deviations within $\sim$ 1–2% and range errors typically < 2 mm in head-and-neck, with some degradation in artifact-prone regions. Metrics include D_{95}/D_{98}, γ-pass (3%/3 mm), WET error, and OAR D_{mean}/D_{max} differences (Veiga et al., 2016; Maspero et al., 2020; Thummerer et al., 2022).
- RSP fidelity dominates error propagation; physics-informed losses that penalize water-equivalent path length mismatch and explicit HU$\rightarrow$RSP consistency improve proton relevance.

Keywords

CycleGAN; unpaired image translation; CBCT-to-CT; diffusion models; radiomics/delta-radiomics; deformable image registration; inverse-consistency; Jacobian; biomechanical regularization; stopping-power ratio (RSP); water-equivalent thickness (WET).

2 SECTION 2: Methodology (How-To)

A. Problem definition

Given longitudinal imaging $\{I_t^{\mathrm{CBCT}}\}$, baseline planning CT I_0^{CT}, and optional functional data (I_t^{PET}, I_t^{MRI}), learn to decompose observed change into: (i) anatomy-driven geometric deformation, and (ii) biology-driven tissue property change that affects RSP and dose. Outputs: voxelwise probability maps and region proposals of biology-driven change; uncertainty maps; dosimetric impact estimates.

B. Architecture: Multi-stream spatiotemporal network with physics-informed fusion

- Streams and inputs:
  - Geometry stream (3D U-Net backbone): inputs are deformation features between warped planning CT $\hat{I}_t^{\mathrm{CT}}$ and sCT I_t^{sCT}: DVF components, Jacobian J, inverse-consistency error, bending energy, strain invariants; optionally surface signed distance maps of GTV/OARs.

  - Texture/biology stream (3D U-Net or Swin-UNETR): inputs are intensity stacks and derived maps: I_t^{sCT}, $\Delta\mathrm{HU} = I_t^{\mathrm{sCT}} - \hat{I}_t^{\mathrm{CT}}$, ΔRSP, radiomics maps (local entropy, LoG, GLCM), and optional PET SUV and MRI ADC changes.
  - Temporal encoder: 1D temporal convolution or a lightweight transformer along time for each spatial stream to exploit trends across fractions ($t = 1, \ldots, T$).

- Cross-stream fusion: cross-attention module aligning geometry and texture tokens; biology predictions are discouraged in regions well-explained by smooth, volume-preserving deformations (via attention masking by $|J - 1|$ and inverse-consistency).

- Heads:
  - Biology vs. anatomy classifier: voxelwise logits with focal loss (addressing class imbalance) and Dice loss.
  - Uncertainty: Monte-Carlo dropout or deep ensemble variance head.
  - Dosimetric impact regressor: predicts local ΔWET and surrogate Δrange from ΔRSP.

- Losses (physics-informed, multi-task):
  - Consistency: penalize biology label where a low-residual, diffeomorphic mapping explains change; encourage biology label where large intensity residuals persist despite well-conditioned DVF: $\mathcal{L}_{\mathrm{bio}} = \mathrm{FocalDice}(y, \hat{y}) + \lambda_1 \langle \mathbf{1}_{\text{low IC}} \cdot \mathbf{1}_{\text{low residual}}, \hat{y} \rangle$.
  - RSP/WET-aware: $\mathcal{L}_{\mathrm{RSP}} = \|\mathrm{RSP}(I_t^{\mathrm{sCT}}) - \mathrm{RSP}(\hat{I}_t^{\mathrm{CT}}) - \widehat{\Delta\mathrm{RSP}}\|_1$ and path-integral $\mathcal{L}_{\mathrm{WET}}$ along beam directions.
  - Temporal smoothness: total variation across t on anatomy channel; sparsity prior on biology channel.

C. Generative model for sCT

- Recommendation: 3D diffusion model (DDPM-style, Ho et al., 2020) with residual 3D U-Net backbone and physics-informed loss terms.

- Rationale for protons:
  - HU/RSP calibration: diffusion models better match full intensity distributions with fewer hallucinations than adversarial models, improving RSP estimation and ΔWET accuracy.
  - Uncertainty: native posterior sampling provides voxelwise uncertainty on HU/RSP, directly propagatable to dose uncertainty.

- Practical note: where paired data are scarce/misaligned, a hybrid approach works well: pretrain a 3D CycleGAN with gradient/identity losses for unpaired CBCT→CT, then distill into/finetune a diffusion model with paired or weakly paired data using registration-regularized supervision.

3 SECTION 3: Coding & Implementation

A. Technical stack (Python-first)

- Deep learning & medical imaging:
  - PyTorch; MONAI; TorchIO; PyTorch Lightning.
  - SimpleITK/ITK; nibabel; pydicom; highdicom; rt-utils (RTSTRUCT I/O).

- Registration & geometry:
  - ITK-Elastix/ANTsPy for SyN; DeepReg; VoxelMorph.
  - Plastimatch for DICOM, resampling, DIR utilities.
- Planning & dose:
  - matRad (MATLAB/Octave) for pencil-beam proton dose; MCsquare (C++ with Python wrappers) for fast Monte Carlo; TOPAS or GATE (Geant4) for high-fidelity MC.
- Radiomics & QA:
  - PyRadiomics; scikit-image; SciPy; NumPy; pandas.
  - WandB/MLflow for experiments; nnU-Net style configs for 3D training.

B. High-level pipeline (pseudocode)

```
Inputs:
  - Planning CT I0_CT, RTSTRUCT, RTPLAN, beams
  - Daily CBCT It_CBCT (t = 1..T)
  - Optional: PET/MRI at selected fractions

1) Preprocess
   - DICOM->NIfTI; clip, normalize, correct CBCT FOV; metal mask if needed
   - Coarse rigid align It_CBCT to I0_CT (Plastimatch/Elastix)

2) sCT generation (CBCT -> sCT)
   - It_sCT = DiffusionModel(It_CBCT)  # with HU/RSP-aware losses
   - Estimate voxelwise HU uncertainty sigma_HU

3) Expected anatomy (warp planning CT)
   - Compute inverse-consistent DIR: phi_t = DIR(I0_CT -> It_sCT)
   - Warp planning CT: Ît_CT = Warp(I0_CT, phi_t)
   - Derive DVF features: Jacobian J, inverse-consistency error ICE, bending energy

4) Biological residuals
   - HU = It_sCT - Ît_CT
   - Convert to RSP via calibrated HU->RSP mapping
   - Local dissimilarity maps (e.g., MIND), texture deltas, SUV/ADC deltas if available

5) Classification (anatomy vs biology)
   - Inputs to multi-stream net:
       Geometry stream: [DVF, J, ICE, strain, surfaces]
       Texture stream:  [Ît_CT, It_sCT, HU, RSP, radiomics, PET/MRI]
       Temporal encoder across {1..t}
   - Outputs:
       P_bio(x), uncertainty u(x), predicted WET(x)

6) Dosimetric validation
   - Build proton material map from sCT (RSP)
   - Recompute dose with MCsquare/TOPAS on It_sCT
   - Compare to dose on Ît_CT and (when available) repeat-CT:
       Metrics: D95/D98, OAR Dmean/Dmax, (3%/3mm), range/WET error
```

```
   - Report dose differences stratified by predicted biology regions

7) Logging & QA
   - Store registrations, residual maps, uncertainty, and dose metrics
```

C. Notes on physics and calibration

- HU→RSP: use a site/scanner-specific calibration or learned mapping constrained by a polynomial/stopping power prior, e.g. $\mathrm{RSP} = a + b \cdot \mathrm{HU} + c \cdot \mathrm{HU}^2$ or multi-tissue piecewise models (Hünemohr et al., 2014).
- WET loss: along each beam path $\mathcal{P}$, penalize $|\int_{\mathcal{P}} \Delta\mathrm{RSP}\, ds|$ to stabilize proton-relevant errors.
- Uncertainty propagation: sample sCT from diffusion posterior; propagate to dose to obtain voxelwise dose uncertainty envelopes in biology-flagged regions.

4 SECTION 4: Data & Validation

A. Open datasets with longitudinal or multi-modal data

- TCIA RADCURE (Head & Neck): planning CTs with multiple longitudinal CBCTs for many H&N patients; includes RTSTRUCT/RTDOSE for subsets. Suitable for CBCT→sCT training and temporal analysis. (Zhao et al., 2020).
- TCIA QIN-HeadNeck: multi-institutional PET/CT for H&N, often with pre-, mid-, and post-RT imaging in subsets; useful for biology labels via SUV changes.
- HECKTOR challenges (2019–2023): PET/CT for H&N with contours; primarily baseline but valuable for multimodal fusion pretraining.
- DIR-Lab 4D-CT (Thorax): 4DCT with landmark annotations; no cancer but ideal for training/benchmarking DIR and computing DVF quality/uncertainty.
- NSCLC-Radiomics (TCIA, Lung): CT with RT annotations; some cohorts include mid/post-RT follow-up enabling delta-radiomics in lung.
- OpenKBP-2020 (Head & Neck): planning CT + structures + dose; no CBCT but useful for dose evaluation baselines and OAR/target context.

Note: availability of mid-treatment scans varies by collection; confirm per-subject timepoints in TCIA metadata before training.

B. Internal/external validation plan

- Split by patient and site; evaluate head-and-neck and lung separately due to artifact patterns and motion differences.
- Primary endpoints:
  - Classification AUROC/AP for biology vs anatomy (voxel/region level), with expert review on a subset and PET/MRI corroboration where available.
  - Dosimetric endpoints in biology-positive regions: $\Delta D_{95}/\Delta D_{98}$, OAR $\Delta D_{mean}/\Delta D_{max}$, γ-pass 3%/3 mm, ΔWET and range errors.

- Ablations: GAN vs diffusion sCT; with/without physics-informed losses; geometry-only vs texture-only vs fused model.
- External test: hold-out institution or scanner if available (e.g., different TCIA site).

Technical Stack List

- PyTorch, MONAI, TorchIO, PyTorch Lightning
- SimpleITK/ITK, nibabel, pydicom, highdicom, rt-utils
- ITK-Elastix, ANTsPy, DeepReg, VoxelMorph, Plastimatch
- PyRadiomics, scikit-image, SciPy, NumPy, pandas
- MCsquare, TOPAS, GATE, matRad
- Weights & Biases or MLflow for experiment tracking

Selected References

- Avants et al. (2008) Symmetric diffeomorphic image registration with cross-correlation (SyN).
- Balakrishnan et al. (2019) VoxelMorph: learning-based deformable registration.
- Ho et al. (2020) Denoising diffusion probabilistic models.
- Hünemohr et al. (2014) An improved calibration for CT-based stopping-power prediction in proton therapy.
- Jiang et al. (2021) RegGAN: unpaired CBCT-to-CT translation regularized by registration for adaptive RT.
- Klein et al. (2010) elastix: a toolbox for intensity-based medical image registration.
- Kurz et al. (2020) Deep-learning-assisted CBCT correction for online ART; implications for dose recalculation.
- Liu et al. (2021) CycleGAN-based CBCT correction with gradient consistency for accurate dose recalculation.
- Maspero et al. (2020) Deep learning sCT for on-line adaptive radiotherapy; vendor-agnostic performance.
- Perl et al. (2012) TOPAS: a tool for particle simulation based on Geant4.
- Souris et al. (2018) MCsquare: fast Monte Carlo code for proton therapy.
- Thummerer et al. (2020, 2022) GAN-based CBCT-to-CT synthesis evaluated for head-and-neck adaptive proton dose.
- Traverso et al. (2019) Stability and predictive power of delta-radiomics for cancer treatment response.
- Veiga et al. (2016) CBCT-based dose recalculation in proton therapy: feasibility and accuracy with physics corrections.
- Wieser et al. (2017) matRad: open-source radiotherapy treatment planning.
- Zhu et al. (2017) CycleGAN: unpaired image-to-image translation.

Risks and Mitigations

- CBCT artifacts (scatter, truncation) may confound biology residuals. Mitigation: artifact-aware masking; robust losses; truncation completion; posterior sampling for uncertainty.
- Registration bias mislabels biology as anatomy. Mitigation: inverse-consistent diffeomorphic DIR; uncertainty maps; landmark QA on DIR-Lab; ensemble of registrars.
- Domain shift across scanners/sites. Mitigation: histogram matching, adversarial/domain adaptation; vendor-agnostic training; test-time adaptation.
- Sparse biology labels. Mitigation: weak supervision via PET/MRI and dosimetric consistency; curriculum learning; simulation with digital phantoms (XCAT) injecting controlled biology changes.