Automating Treatment Planning in Radiation Therapy

Lekan Ogunmolu

Perelman School of Medicine

University of Pennsylvania, Philadelphia, PA

February 04, 2021

Talk Outline

- Beam Orientation Optimization (BOO)
  - Monte Carlo Tree Search and Neuro-Dynamic Programming for BOO
  - Column Generation as Pretraining for MCTS for BOO
- Patient Head Motion Correction in External Beam Radiation Therapy
  - Magnetic Resonance Imaging and Linear Accelerator Systems (MRI-LINACs)
  - Intensity-Modulated RT (IMRT): Earlier PhD Work
- Robustness Margins and Robust Deep Policies for Nonlinear Control

Research Significance

2020 → 1.8 million U.S. cancer cases [1] → > 600,000 deaths → 4.2% overall healthcare spending

Cancer Treatment

- Chemotherapy/Radiosurgery.
- Radiation Therapy (RT).

RT Signif.

- Approx. 40% of all curative cancer treatment modality are performed with RT [Baskar et al].
- Cost-effective -- accounting for only 5% of the total cost of cancer care [Ringborg et al.].

Imaging

- CT images in RT = Low soft tissues contrast ⇒ OARs toxicity during treatment planning.
- MRI of brain, and head and neck (H&N) cancers provides exceptional soft tissues contrast and allows greater delineation of cancerous from non-cancerous tissues.

Part I.A: Beam Orientation Optimization (BOO)

- Beam Orientation Optimization (BOO)
  - Monte Carlo Tree Search and Neuro-Dynamic Programming for BOO

Stereotactic Radiosurgery

External Beam RT

- External Beam RT is a widespread method for treating various cancer types.
- High-energy photons from a linear accelerator (LINAC) render cancerous cells necrotic.
- Excessive damage to healthy critical structures minimizes patient's quality of life.
- How to efficiently irradiate tumors while sparing organs-at-risk (OARs)?

BOO Relevant Works

- Sadeghnejad Barkousaraie, Azar, **Olalekan Ogunmolu**, Steve Jiang, and Dan Nguyen. "A fast deep learning approach for beam orientation optimization for prostate cancer treated with intensity-modulated radiation therapy." In *Medical physics: International Journal of Medical Physics Research and Practice,* 47, no. 3 (2020): 880-897.
- **Ogunmolu, Olalekan**, Michael Folkerts, Dan Nguyen, Nicholas Gans, and Steve Jiang. "Deep BOO! Automating Beam Orientation Optimization in Radiation Therapy." In *Algorithm Foundations of Robotics XIII*, Merida, Mexico. Published in *Springer's Proceedings in Advanced Robotics (SPAR) Book*, 2020.
- Barkousaraie, Azar Sadeghnejad, **Olalekan Ogunmolu**, Steve Jiang, and Dan Nguyen. "Using Supervised Learning and Guided Monte Carlo Tree Search for Beam Orientation Optimization in Radiation Therapy." In *Workshop on Artificial Intelligence in Radiation Therapy*, pp. 1-9. Springer, Cham, 2019.
- Azar Sadeghnejad Barkousaraie, **Olalekan Ogunmolu**, Steve Jiang, and Dan Nguyen. "A Fast Deep Learning Approach for Beam Orientation Selection Using Supervised Learning with Column Generation on IMRT Prostate Cancer Patients." *Medical Physics (AAPM)* 46 (6), E237-E237, San Antonio, TX, July 2019.
- **Olalekan Ogunmolu**, Azar Sadeghnejad Barkousaraie, Nicholas Gans, Steve Jiang, and Dan Nguyen. "An Approximate Policy Iteration Scheme for Beam Orientation Selection in Radiation Therapy." *Medical Physics (AAPM)* 46 (6), E386-E386 San Antonio, TX, July 2019.
- Azar Sadeghnejad Barkousaraie, **Olalekan Ogunmolu**, Steve Jiang, and Dan Nguyen. "A Reinforcement Learning Application of Guided Monte Carlo Tree Search Algorithm for Beam Orientation Selection in Radiation Therapy." *Medical Physics (AAPM)* 46 (6), E236-E236, San Antonio, TX, July 2019.

Funding Agencies/Funds

- Cancer Prevention and Research Institute of Texas (CPRIT) (IIRA RP150485): $858,356. PI: Steve Jiang
- CPRIT MIRA RP160661: $4,103,894. PI: Steve Jiang
- NIH R-01 1R01CA237269-01: $490,133. PI: Steve Jiang

IMRT/Beam Orientation Optimization

- Intensity-modulated radiation therapy (IMRT) is one common EBRT method:
  - Delivers geometrically-shaped, high-precision photons
  - From different static beam orientations towards a planning target volume (PTV).
- BOO Problem: Determine best beam angle combinations for delivering radiation to a patient.
  - Essentially a combinatorial optimization problem.
  - Process of determining beamlets' intensities is termed **fluence map optimization** (FMO).

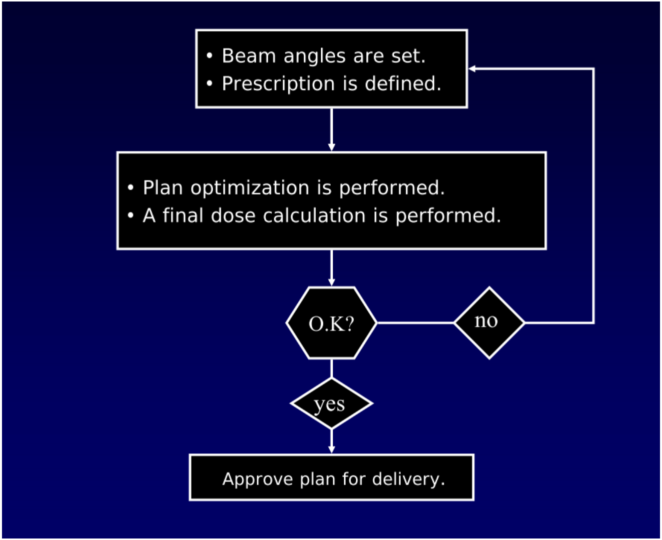
IMRT/BOO Motivation

- In most clinics, beam orientations are still manually chosen or adopted from a standard protocol for clinical use.
- Typical approaches leverage pre-solving the dose influence matrices for each beam orientation.
- Then solve FMO.
- Time consuming (hours for dose fluence), and minutes for (FMO); Still solution is often not optimal.

Current Approaches

- Stochastic optimization approaches: simulated annealing; genetic algorithms and gradient search, or a combination of genetic and gradient search algorithms.
- Mixed-integer programming, branch and cut/bound algorithms, beam angle elimination algorithms.
- Commercial planners use some highly non-convex objective (actual function is proprietary and unknown to public).

BOO Flowchart



Reprinted from "IMRT Optimization Algorithms. David Shepard. Swedish Cancer Institute. AAPM 2007."

Contributions

WAFR '18 Paper

Ogunmolu, Olalekan, Michael Folkerts, Dan Nguyen, Nicholas Gans, and Steve Jiang. "Deep BOO! Automating Beam Orientation Optimization in Radiation Therapy." In *Algorithm Foundations of Robotics XIII*, Merida, Mexico. Published in *Springer's Proceedings in Advanced Robotics (SPAR) Book*, 2020.

- A sparse tree lookout strategy for games with large state spaces guides transition between beam angle sets
- Tree lookout strategy guided by a deep neural network policy

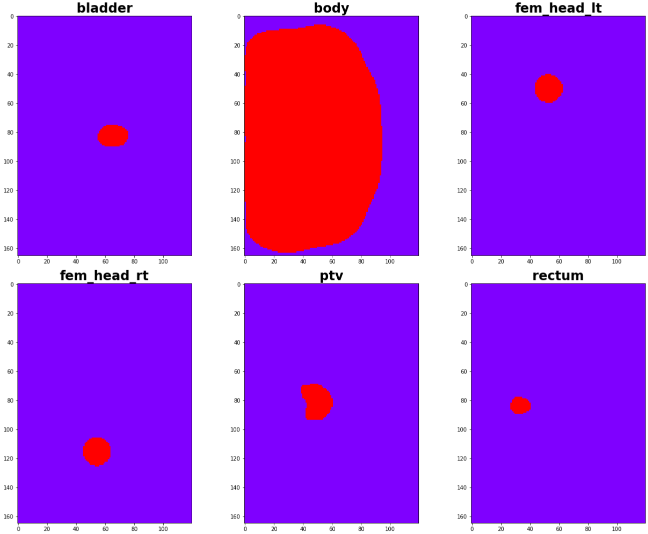
What have we wrought?

- Network generates a policy that guides MCTS simulations for two players in a zero-sum Markov game
  - Produces a *utility (value) function* & a subjective *probability distribution*
- Each player in a two-player Markov game finds an alternating best response to the current player's average strategy
  - driving the neural network policy's weights toward an approximate **saddle equilibrium** [Heinrich et al. (2015)].
  - aids network in finding an *approximately optimal* beam angle candidate set that meets a dosimetric requirements.

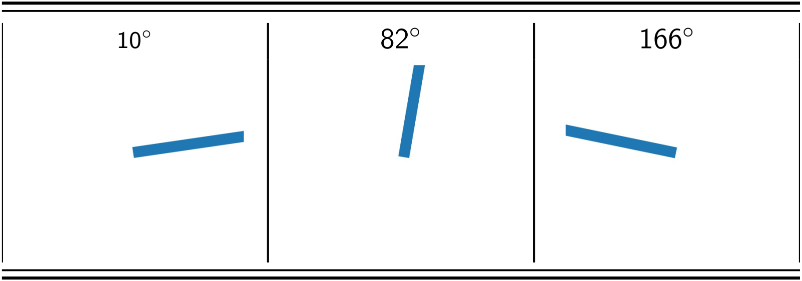
Data Preprocessing

- 77 anonymized patient CT scans, $\boldsymbol{D}$, and their dose influence matrices, $\mathcal{D}_{ij}$
- Scans shaped, $\boldsymbol{D} \times N \times H \times W$ from prostate cases in previous treatment plans
- Each slice resized to 64×64

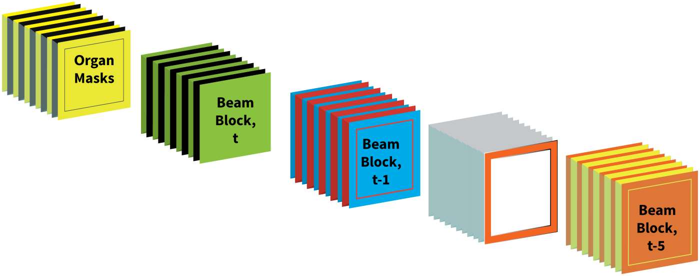
State Representation: Prostate Organ Masks



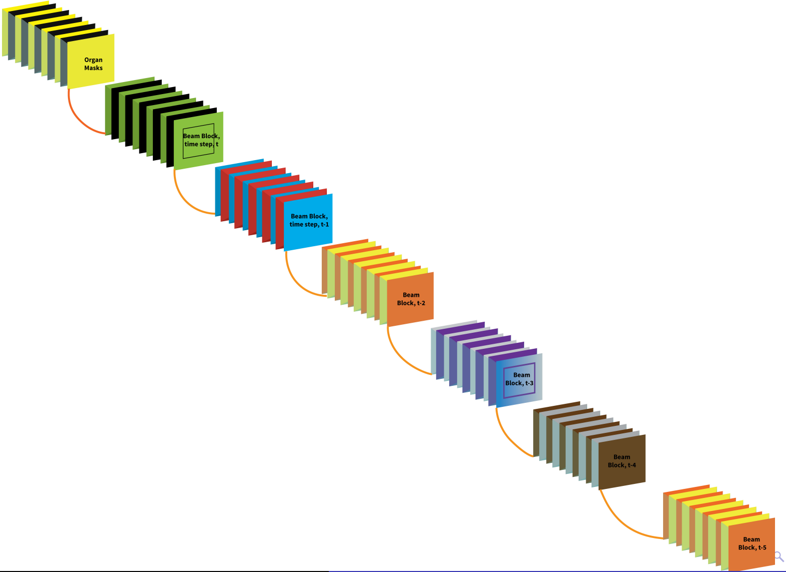
State Representation: Beam Angles



State Representation



State Representation

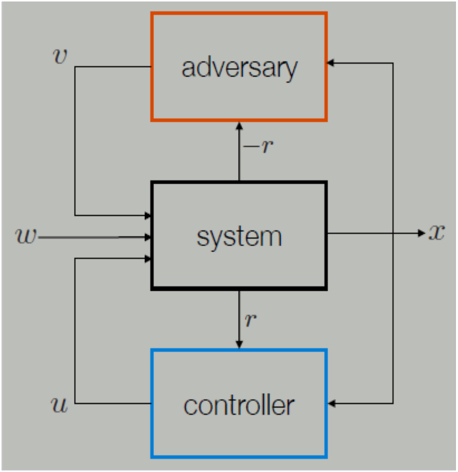


Tree Composition

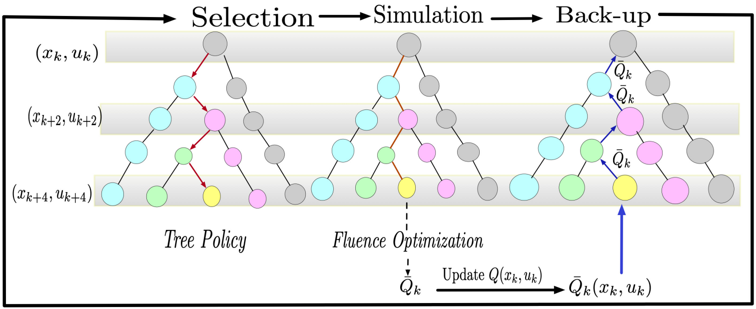
Every **node** of the tree, $\boldsymbol{x}$, has the following fields:

- a pointer to the parent that led to it, $\boldsymbol{x}.p$;
- the beamlets, $\boldsymbol{x}_b$, stored at that node; $b = \{1, \ldots, m\}$;
- a set of move probabilities prior, $p(s, a)$;
- a pointer $\boldsymbol{x}.\boldsymbol{r}$, to the reward r_t, for the state $\boldsymbol{x}_t$;
- a pointer to the state-action value $Q(s, a)$ and its upper confidence bound $U(s, a)$;
- a visit count $N(s, a)$, that indicates the number of times that node was visited; and
- a pointer $\boldsymbol{x}.child_i$ to each of its children nodes.

Two-Player MCTS Training Framework



Game Simulation



Game Tree Simulation

- Network roll-out policy then guides the tree's game toward a *best-first* set of beam angle candidates
- Best-first leaf node encountered is the child node with the highest reward in the tree
- Focuses learning on regions of the state space that are likely to have a good fluence

Mixed Strategies

- Each player, p_1, p_2, bases its decision on a random event's outcome
  - generating a **mixed strategy** determined by **averaging the outcome** of individual plays.
- Both players constitute a two-player **stochastic action selection strategy**: $\pi(s, a) = Pr(a|s) := \{\pi^{p_1}, \pi^{p_2}\}$ that gives the probability of selecting moves in any given state
- Suppose the game simulation starts from an initial condition s_0.

Saddle Point Strategy Formulation

- The **saddle point strategies** for an optimal control sequence pair $\{a_t^{p_1^*}, a_t^{p_2^*}\}$ can be recursively obtained by optimizing a state-action value cost, $\mathcal{J}_t(s, a)$

$$V_t^*(s) = Q_t^*(s_t, \pi_t^{p_1}, \pi_t^{p_2}) = \min_{\pi^{p_1} \in \Pi^{p_1}} \max_{\pi^{p_2} \in \Pi^{p_2}} Q_t^\star(s_t, \pi^{p_1}, \pi^{p_2})$$
$$\forall s_t \in \mathcal{S}, \pi^{p_1} \in \Pi^{p_1}, \pi^{p_2} \in \Pi^{p_2}.$$

such that

$$v_{p_1}^\star \leq v^\star \leq v_{p_2}^\star \quad \forall \{\pi_t^{p_1}, \pi_t^{p_2}\}_{0 \leq t \leq T}.$$

where $v_{p_i}^\star$ are the respective optimal values for each player.

Methods: Search

- Q-value defined as

$$\bar{Q}(s, a) = Q_j(s, a) + c\sqrt{\frac{2 \ln \, n(s)}{N(s, a}}, \tag{1}$$

$$a^\star = \arg\max_a \bar{Q}(s, a) \tag{2}$$

Fluence Map Optimization

- Suppose $\mathcal{X}$ is the total discretized of voxels of interest (VOI's) in a target volume
- Let $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_n \subseteq \mathcal{B}$ represents the partition subset of a beam $\mathcal{B}$,

Methods: FMO problem definition

- The pre-calculated dose term is given by $\boldsymbol{Ax} = \{\sum_s \frac{w_s}{v_s} \mathcal{D}_{ij}^s \boldsymbol{x}_s \,|\, \mathcal{D}_{ij} \in \mathbb{R}^{n \times l}, n > l\}$
- Let $w_s = \{\underline{w}_s, \bar{w}_s\}$ be the respective underdosing and overdosing weights for the OARs and PTVs
- We propose the following cost

$$\frac{1}{w_s} \sum_{s \in \text{OARs}} \|(b_s - \underline{w}_s \mathcal{D}_{ij}^s \boldsymbol{x}_s)_+\|_2^2 + \frac{1}{v_s} \sum_{s \in \text{PTVs}} \|(\bar{w}_s \mathcal{D}_{ij}^s \boldsymbol{x}_s - b_s)_+\|_2^2 \tag{3}$$

Methods: FMO

- Rewriting the objective, subject to nonnegative bixel intensity constraints, we have the minimization problem

$$\min \frac{1}{2}\|A\mathbf{x} - \mathbf{b}\|_2^2 \quad \text{subject to } x \geq 0.$$

- The Lagrangian becomes

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = \min \frac{1}{2}\|A\mathbf{x} - \mathbf{b}\|_2^2 - \boldsymbol{\lambda}^T \boldsymbol{x}.$$

- Introducing an auxiliary variable $\mathbf{z}$, we have

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \mathbf{b}\|_2^2, \quad \text{subject to } \boldsymbol{z} = \boldsymbol{x}, \;\; \boldsymbol{z} \geq 0,$$

Methods: FMO by way of ADMM

- Solving either the $\boldsymbol{x}$ and $\boldsymbol{z}$ sub-problems, we have

$$\boldsymbol{x}^{k+1} = \left(\boldsymbol{A}^T\boldsymbol{A} + \rho\boldsymbol{I}\right)^{-1}\left(\boldsymbol{A}^T\boldsymbol{b} + \rho\boldsymbol{z}^k - \boldsymbol{\lambda}^k\right). \qquad (4)$$

- And using the soft-thresholding operator, $S_{\boldsymbol{\lambda}/\rho}$, we find that

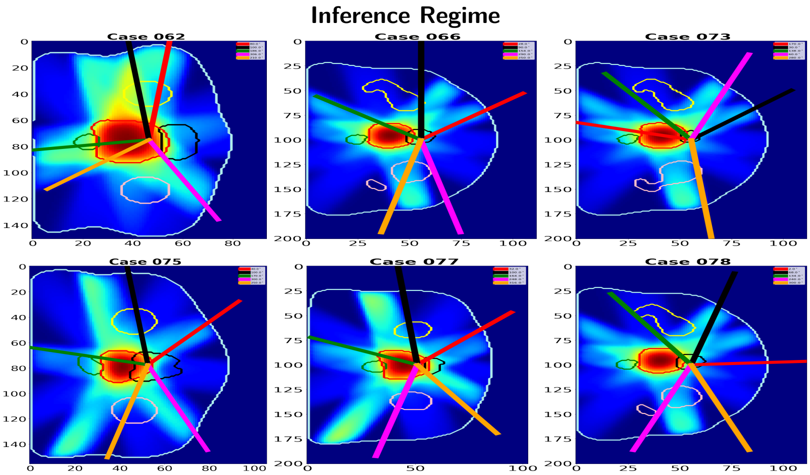
$$\boldsymbol{z}^{k+1} = S_{\boldsymbol{\lambda}/\rho}\left(\boldsymbol{x}^{k+1} + \boldsymbol{\lambda}^k\right), \qquad (5)$$

where $S_{\boldsymbol{\lambda}/\rho}(\tau) = (x - \boldsymbol{\lambda}/\rho)_+ - (-\tau - \boldsymbol{\lambda}/\rho)_+$. $\boldsymbol{\lambda}$ is updated as

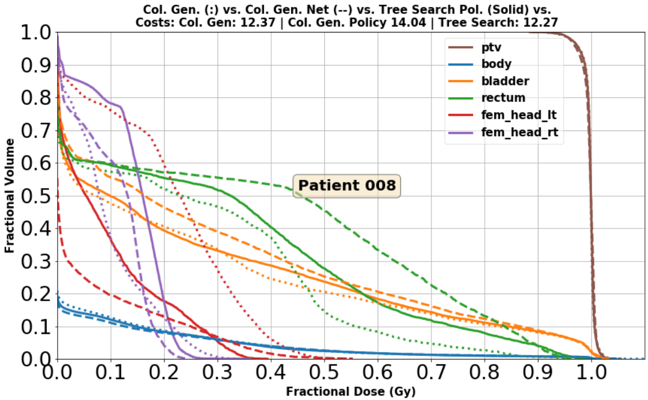
$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k - \gamma(\boldsymbol{z}^{k+1} - \boldsymbol{x}^{k+1}), \qquad (6)$$

where γ is a parameter that controls the step length.

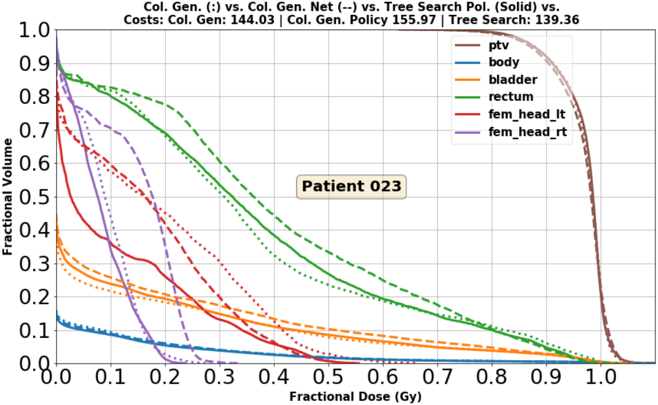
Dose washes for select patients during testing of self-play network



Dose Volume Histograms



Dose Volume Histograms



Discussions

- The policy selects fairly equidistant beams
- Yields wash plots that provide good dosimetric concentration on the tumor
- Gives sharp gradients at transition between tumors and OARs
- Largely avoids strong dose to OARs
- Finding the good beam angle candidates is orders of magnitude faster than the current approaches
- Beam angles prediction now takes between 2-3 minutes before we settle on a good candidate beam angle set.

Conclusions

- Finding the good beam angle candidates is orders of magnitude faster than the current approaches
  - Based on a neural network generative model of an MDP
  - Sparse lookahead search builds tree with nodes labeled by state-action pairs in an alternating manner (2-3 minutes).
  - Tree built stagewise from root to nodes has fixed depth; sample rewards stored on edges connecting state-action with state nodes
- Beam angles prediction now takes between 2-3 minutes before we settle on a good candidate beam angle set.

Part I.B: Supervised Column Generation Pretraining for BOO

- Beam Orientation Optimization (BOO)
  - Monte Carlo Tree Search and Neuro-Dynamic Programming for BOO
  - $\rightarrow$ Column Generation as Pretraining for Deep Neural Network BOO

Funding Agencies/Funds

- Cancer Prevention and Research Institute of Texas (CPRIT) (IIRA RP150485): $858,356. PI: Steve Jiang
- CPRIT MIRA RP160661: $4,103,894. PI: Steve Jiang
- NIH R-01 1R01CA237269-01: $490,133. PI: Steve Jiang

Relevant Publications

- Sadeghnejad Barkousaraie, Azar, **Olalekan Ogunmolu**, Steve Jiang, and Dan Nguyen. "A fast deep learning approach for beam orientation optimization for prostate cancer treated with intensity-modulated radiation therapy." In *Medical physics: International Journal of Medical Physics Research and Practice,* 47, no. 3 (2020): 880-897.
- Barkousaraie, Azar Sadeghnejad, **Olalekan Ogunmolu**, Steve Jiang, and Dan Nguyen. "Using Supervised Learning and Guided Monte Carlo Tree Search for Beam Orientation Optimization in Radiation Therapy." In *Workshop on Artificial Intelligence in Radiation Therapy*, pp. 1-9. Springer, Cham, 2019.
- Azar Sadeghnejad Barkousaraie, **Olalekan Ogunmolu**, Steve Jiang, and Dan Nguyen. "A Fast Deep Learning Approach for Beam Orientation Selection Using Supervised Learning with Column Generation on IMRT Prostate Cancer Patients." *Medical Physics (AAPM)* 46 (6), E237-E237, San Antonio, TX, July 2019.

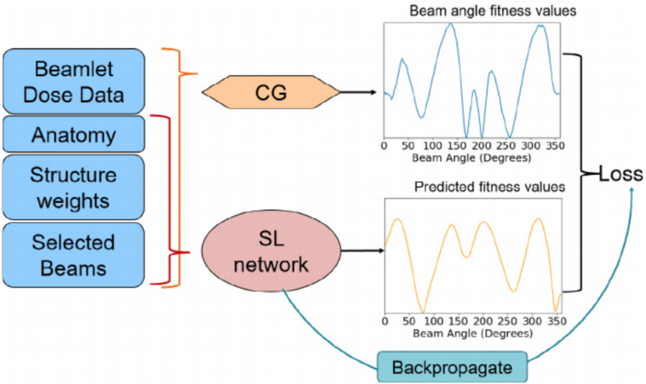
Goals

- Fast+Flexible DNN BOO Solution in seconds
- Application to clinical routines for treatment planning acceleration
- Patient's anatomy $\Leftrightarrow$ Optimal Beam Orientations $\Leftrightarrow$ Optimization algorithm
- Data Preparation as WAFR '18 paper + augmentation + random OAR weight generation $= 30,800$ samples for training, validation and testing proposed DNN algorithm.
- Greedy iterative column generation algorithm for finding optimal beam set.

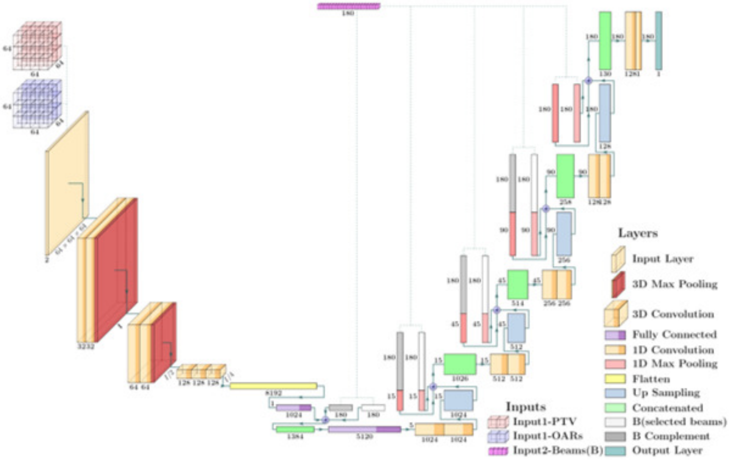
Training Overview

- From an empty beam set
  - Iteratively add a beam with the greatest likelihood to improve the current FMO solution
  - FMO leverages Chambolle-Pock first-order primal-dual proximal operator on GPU
  - DNN then trained to learn beam orientation reasoning of CG
- DNN essentially internalizes the FMO solution via CG.s

Training Schematic



Network Structure

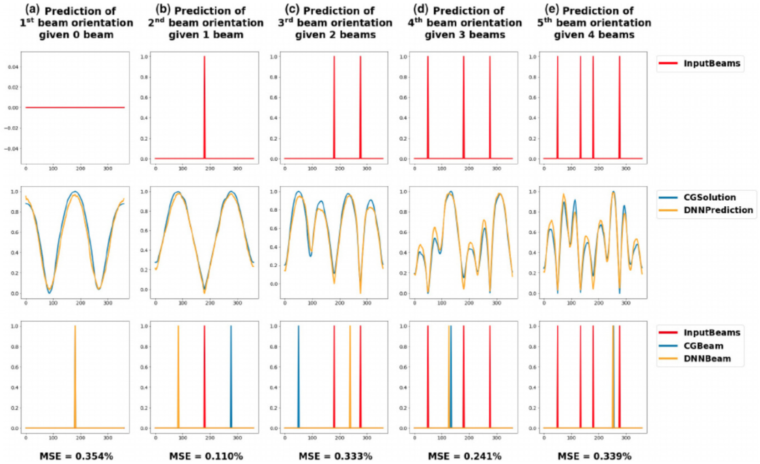


Loss Table

Fold #	1	2	3	4	5	6	Total [a]
Train best [b]	0.50%	0.63%	0.62%	0.68%	0.53%	0.78%	0.62 ± 0.09%
Validation best [c]	0.91%	1.04%	1.06%	1.06%	1.12%	1.06%	1.04 ± 0.06%
Test best [d]	1.39%	1.44%	1.50%	1.30%	1.39%	1.64%	1.44 ± 0.11%
Training last [e]	0.62%	0.52%	0.63%	0.63%	0.53%	0.50%	0.57 ± 0.06%
Validation last [f]	1.92%	1.17%	2.54%	2.49%	2.62%	2.18%	2.15 ± 0.50%
Test last [g]	1.59%	1.54%	1.73%	1.79%	1.68%	1.65%	1.66 ± 0.08%
Best epoch number	374	303	189	272	387	165	

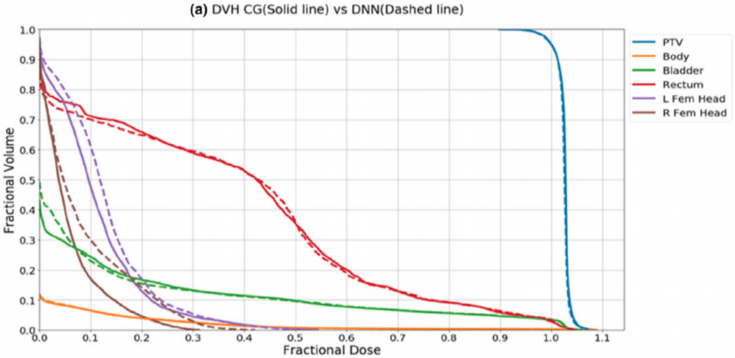
Average training, validation, and test loss functions (MSE%) at different epochs for the full network.

Inference: Column Generation vs U-Net



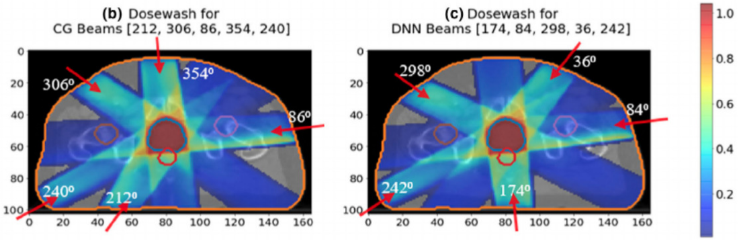
(a) Prediction of 1st beam orientation given no beam. (b) Prediction of 2nd beam orientation given 1 beam. (c) Prediction of 3rd beam orientation given 2 beams. (d) Prediction of 4th beam orientation given 3 beams. (e) Prediction of 5th beam orientation given 4 beams.

DVH of Column Generation vs Neural Network



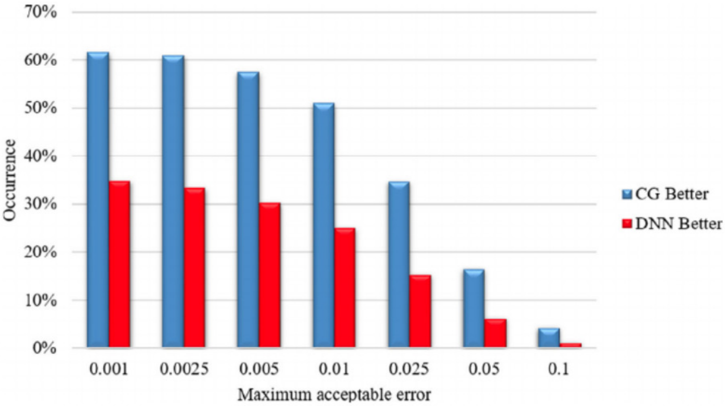
Dose-Volume Histogram of CG vs DNN architectures

Dose Washes of Column Generation vs Neural Network



Dose-Volume Histogram of CG vs DNN architectures

FMO Costs: Column Generation vs Neural Network



Dose-Volume Histogram of CG vs DNN architectures

Conclusions

- A Sparse Lookout Tree Strategy for guiding beam angle transitions while training a neural network
- A supervised deep neural network that learns from a CG algorithm
- Both methods leverage the convex FMO in learning the *optimal* set of beam angles
  - Alternating Direction Method of Multipliers
  - Chambolle-Pock algorithm
- Trade-offs in solutions generated by either approach allows flexibility for treatment planners
- Grants

Part II.A: Head Stabilization in RT

- Head Stabilization in Cancer Radiation Therapy
  - Magnetic Resonance Imaging-Linear Accelerator Systems (MRI-LINACs)
- **Funding Sources**
  - NIH-R01. PI: Rodney Wiersma
  - UT Southwestern Medical Center: PI: Steve Jiang and Nick Gans

Robotic Radiosurgery

Penn

UNIVERSITY *of* PENNSYLVANIA

A Patient Head Motion-Correction Mechanism for MRI-LINAC RT

OLALEKAN OGUNMOLU

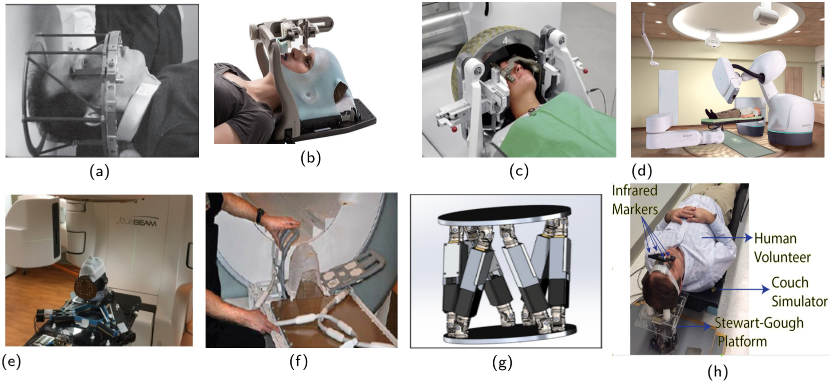
DEPARTMENT OF RADIATION ONCOLOGY, PENN SCHOOL OF MEDICINE

- Current Collaborators: Rodney Wiersma & Xinmin Liu (UChicago → UPenn)
- Past Collaborators: Steve Jiang, Xuejun Gu, (UT Southwestern); Nick Gans (UT Dallas, UT Arlington)

Relevant Publications

- Ogunmolu, Olalekan, Xinmin Liu, Nicholas Gans, and Rodney D. Wiersma. "Mechanism and Model of a Soft Robot for Head Stabilization in Cancer Radiation Therapy." In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4609-4615. IEEE, 2020.
- Ogunmolu, Olalekan, and Rodney D. Wiersma. "Kinematics and Kinects of a Continuum Parallel Soft Robot for MRI-LINAC Motion Correction." Working Paper, In *IEEE Transactions on Robotics*, IEEE 2020.
- Ogunmolu, Olalekan and Rodney D. Wiersma. "A Real-Time Patient Head Motion Correction Mechanism for MRI-Linac Systems." In *2020 Virtual Joint AAPM|COMP Meeting*,AAPM 2020.
- Ogunmolu, Olalekan, Xinmin Liu, Rodney D. Wiersma. "Auto-Determination of the Dextrous WorkSpace in Robotic Stereotactic Radiosurgery." In *2020 Virtual Joint AAPM|COMP Meeting*, AAPM 2020.

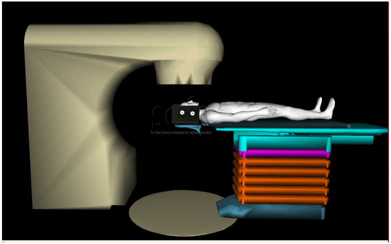
Correcting Head Motion: Existing Techniques

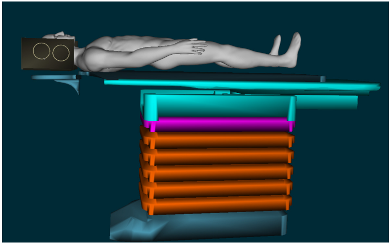


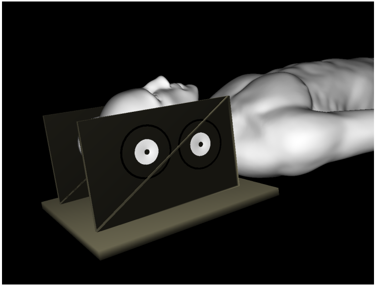
(a) The BRW SRS Frame [Chelvarajah et al. (2004)], (b) A thermoplastic face mask (c) Gamma Knife, (d) Accuray© Inc's Cyberknife (e) The Ostyn robot (f) Frame With MRI Coils [Courtesy of PSOM], (g) Wiersma Platform [Belcher (2017)], (h) Motion Compensation with the Wiersma Stewart-Gough Platform [Ogunmolu and Wiersma (2020)].

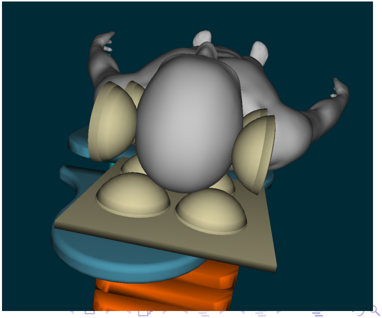
MRI-LINAC

Proposal: External Beam RT with Soft Robots

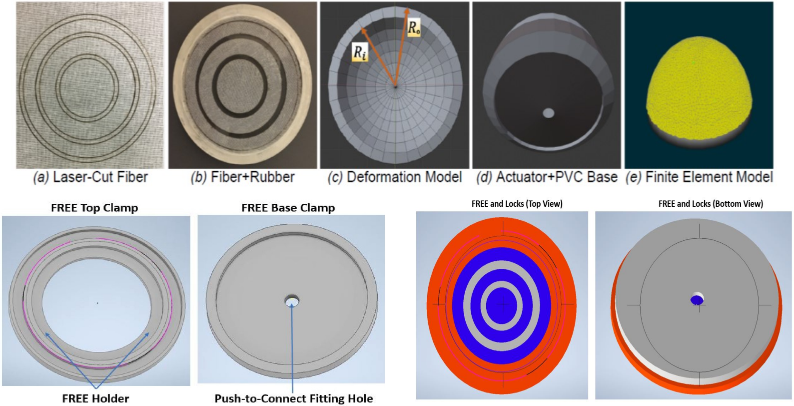








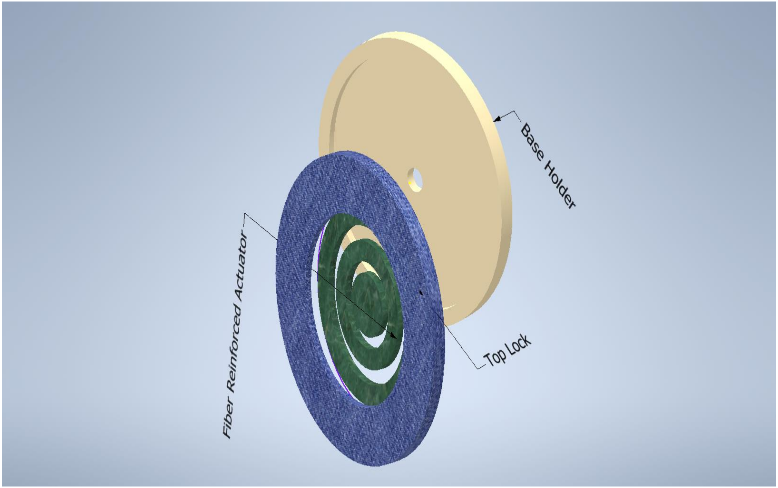
Cephalopods-inspired CCOARSE Actuator Design



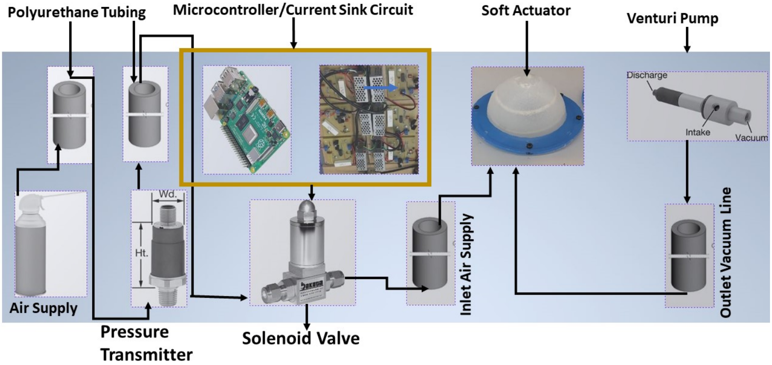
(a) Laser-Cut Fiber (b) Fiber+Rubber (c) Deformation Model (d) Actuator+PVC Base (e) Finite Element Model

(f,g) (h)

Cephalopods-inspired CCOARSE Actuator Design

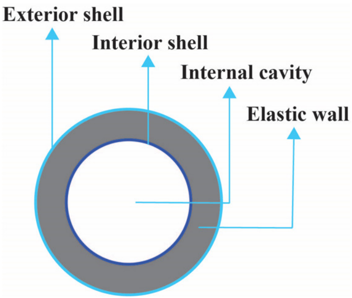


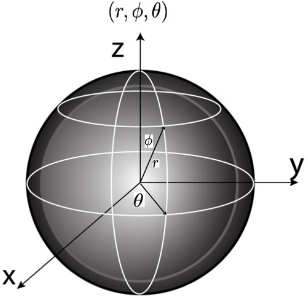
Pneumatic Actuation/Control Scheme



Nonlinear Elastic Deformation Analysis

IAB SHELLS AND AIR CAVITY/DEFORMATION ANALYSIS



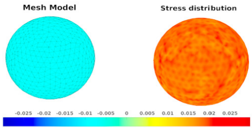


$$r_i \leq r \leq r_o, \qquad 0 \leq \theta \leq 2\pi, \qquad 0 \leq \phi \leq \pi$$

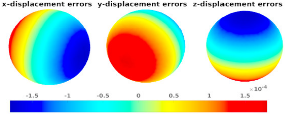
Soft IK via Boundary Value Problem

- Using the following boundary conditions for the Cauchy Stress,
  - $\sigma_{rr|R=R_0} = -P_{atm}, \sigma_{rr|R=R_i} = -P_{atm} - P$
- And together with Cauchy's first law, we find that
  - $\sigma_{rr}(r) = -\int_{r_i}^{r_o}[2C_1(\frac{r}{R^2} - \frac{R^4}{r^5}) + 2\,C_2\left(\frac{r^3}{R^4} - \frac{R^2}{r^3}\right)]dr$
  - $\sigma_{rr}(r) = \int_{R_i}^{R_o}[2C_1\left(\frac{1}{r} - \frac{R^6}{r^7}\right) - 2C_2\left(\frac{R^4}{r^5} - \frac{r}{R^2}\right)]\,dr$
- With $\sigma_{rr|R=R_i} = -P_{atm} - P$ and setting $P_{atm} = 0$, we find
  - $P(r) = \int_{r_i}^{r_o}[2C_1(\frac{r}{R^2} - \frac{R^4}{r^5}) + 2\,C_2\left(\frac{r^3}{R^4} - \frac{R^2}{r^3}\right)]dr$
  - $P(r) = \int_{R_i}^{R_o}[2C_1\left(\frac{1}{r} - \frac{R^6}{r^7}\right) - 2C_2\left(\frac{R^4}{r^5} - \frac{r}{R^2}\right)]\,dr$
  - $r^3 = R^3 + r_i^3 - R_i^3$ and $r_o^3 = R_0^3 + r_i^3 - R_i^3$

Volumetric Deformation Results (Simulation)



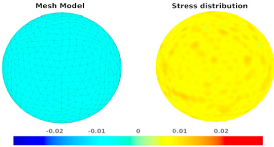
(a) **Left**: Mesh model. **Right**: Stress distribution on outer skin.



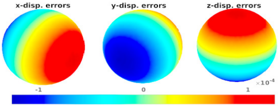
(b) Displacement errors along x, y, z coordinates.

Inputs					Outputs		
C_1	C_2	R_i	r_i	R_o	r_o	P	ΔV
$1.1e4$	$2.2e4$	$.027$	$.03$	$.03$	$.033$	$.76$	≈ 0

Fig. 6: Volumetric Deformation (Expansion).



(a) **Left**: Mesh model. **Right**: Stress distribution on outer skin.

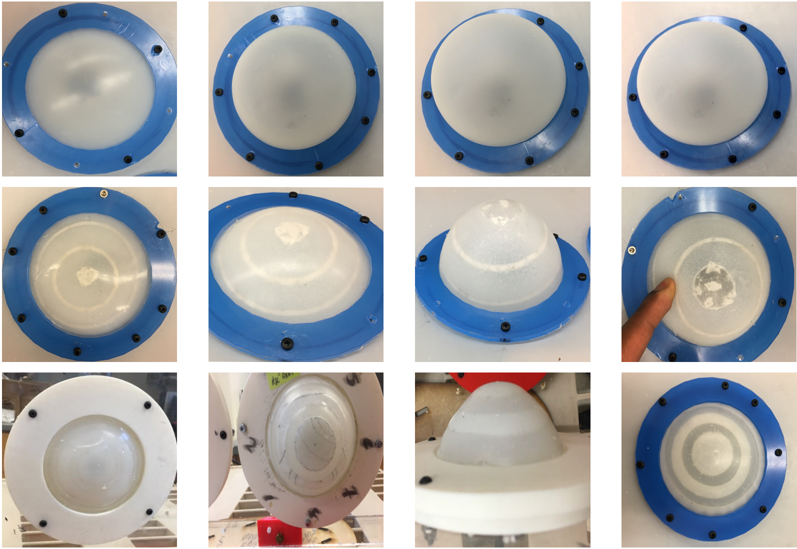


(b) Displacement errors along x, y, z coordinates.

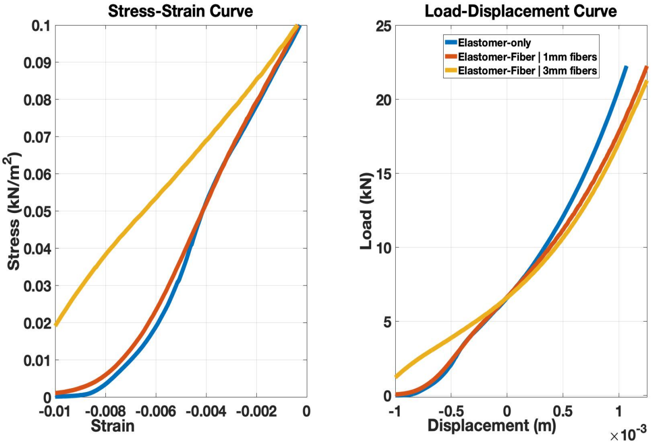
Inputs					Outputs		
C_1	C_2	R_i	r_i	R_o	r_o	P	ΔV
$1.1e4$	$2.2e4$	$.025$	$.03$	$.03$	$.028$	$-.34$	≈ 0

Fig. 7: Volumetric Deformation (Compression).

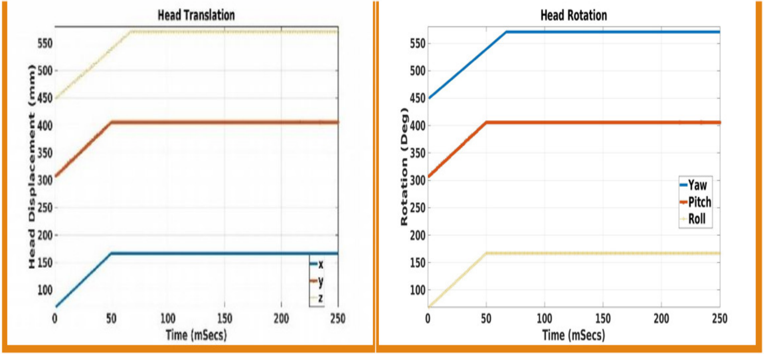
Volumetric Deformation Results (Actual)



Tension/Compression Analysis



Head Motion Control (Independent Actuation)

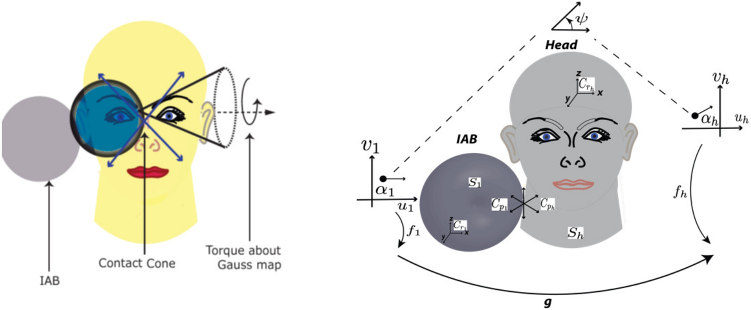


Head Translation along x, y, z for a task of raising the head by a certain threshold above the table

Head rotation in Euler angles for a task of tilting the head about the x, y, z axes on the treatment table.

Ongoing/Future work

Continuum Mechanical Model Validation/Differential Geometry/Newton-Euler Dynamics



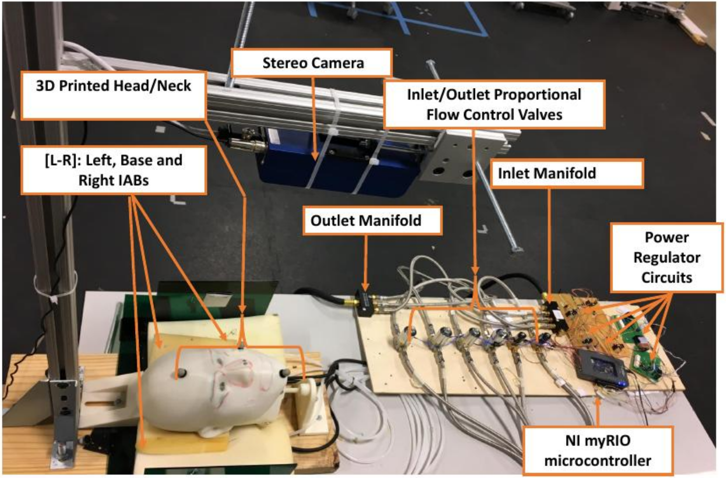
Part II.B: Head Stabilization in Radiation Therapy (RT)

- Head Stabilization in Cancer Radiation Therapy
  - Intensity-Modulated RT (IMRT): Earlier PhD Work

Relevant Papers

- Ogunmolu, O., N. Gans, S. Jiang, and X. Gu. "SU-E-J-12: An Image-Guided Soft Robotic Patient Positioning System for Maskless Head-And-Neck Cancer Radiotherapy: A Proof-Of-Concept Study." *Medical Physics* 42, no. 6Part7 (2015): 3266-3266.
- Ogunmolu, Olalekan P., Xuejun Gu, Steve Jiang, and Nicholas R. Gans. "A real-time, soft robotic patient positioning system for maskless head-and-neck cancer radiotherapy: an initial investigation." In *2015 IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 1539-1545. IEEE, 2015.
- Ogunmolu, Olalekan P., Xuejun Gu, Steve Jiang, and Nicholas R. Gans. "Vision-based control of a soft robot for maskless head and neck cancer radiotherapy," In *2016 IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 180-187.IEEE, 2016.
- Ogunmolu, Olalekan, Adwait Kulkarni, Yonas Tadesse, Xuejun Gu, Steve Jiang, and Nicholas Gans. "Soft-neuroadapt: A 3-dof neuro-adaptive patient pose correction system for frameless and maskless cancer radiotherapy." In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3661-3668. IEEE, 2017.
- Almubarak, Yara, Aniket Joshi, Olalekan Ogunmolu, Xuejun Gu, Steve Jiang, Nicholas Gans, and Yonas Tadesse. "Design and development of soft robot for head and neck cancer radiotherapy." In *Electroactive Polymer Actuators and Devices (EAPAD) XX*, vol. 10594, p. 1059418. International Society for Optics and Photonics, 2018.

Simulation Testbed



Hardware Description

Control Proposals

- Solve a state feedback and feedforward regulation problem
- An adaptation model based on past states and controls:
  - $Z^N = \{u(k), u(k-1), \ldots u(k-n_u), y(k), \ldots, y(k-n_y)\}$
- Design Goal:
  - Stabilize states, $\mathbf{y} = \left[z, \theta, \phi\right]^T$
  - *i.e.* z, pitch, and roll.

Model Reference Adaptive Control

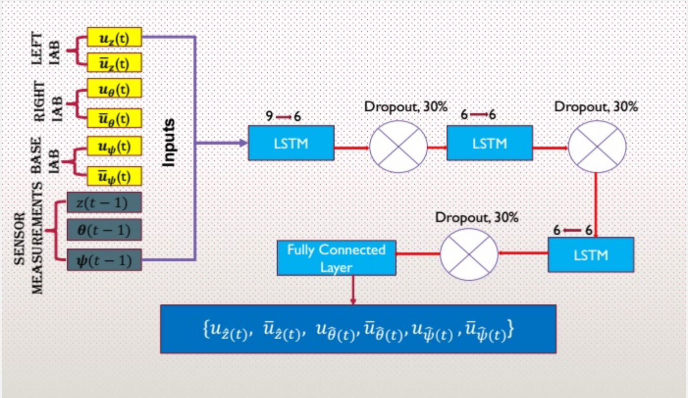
- Model head and bladder dynamics
  - $\dot{\mathbf{y}} = \mathbf{A}\mathbf{y} + \mathbf{B}\boldsymbol{\Lambda}\left(\mathbf{u} - f(\mathbf{y}, \mathbf{u})\right) + \mathbf{w}(k)$
    - $\mathbf{A}$, $\boldsymbol{\Lambda}$ unknown, $\mathbf{B}$, sgn$\boldsymbol{\Lambda}$ known
- Approximate $f(\mathbf{y}, \mathbf{u})$ by a neural network with continuous memory states
  - $\hat{f}(\mathbf{y}(k), \mathbf{u}(k-d))$ is realized with a *long-short term memory* cell (Horchreiter and Schmidhuber, '91, '97)
  - **purpose**: remember good adaptation gains

Adaptive Neuro-Control Scheme

- Derive adaptive adjustment mechanism from Lyapunov analysis for Adaptive Control (Parks, P., 1966)
- $\mathbf{u} = \underbrace{\hat{\mathbf{K}}_y^{\mathbf{T}}\mathbf{y}}_{\text{state feedback}} + \underbrace{\hat{\mathbf{K}}_r^{\mathbf{T}}\mathbf{r}}_{\text{optimal regulator}} + \underbrace{\hat{f}(\mathbf{y}, \mathbf{u})}_{\text{approximator}}$

Neural Network Model

Neural Net Architecture



Lyapunov Redesign: Theorem

- Given correct choice of adaptive gains $\hat{\mathbf{K}}_y$ and $\hat{\mathbf{K}}_r$, the error state vector, $\mathbf{e}(k)$ with closed loop time derivative $\dot{\mathbf{e}}$, is ***uniformly ultimately bounded***, and the state $\mathbf{y}$ will converge to a neighborhood of $\mathbf{r}$.
- Choose a Lyapunov function candidate $\mathbf{V}$ in terms of the generalized error state space $\mathbf{e}$, gains, $\tilde{\mathbf{K}}_y^T$, $\tilde{\mathbf{K}}_r^T$, and parameter error $\varepsilon_f(\mathbf{y}(k))$ space

$$\mathbf{V}(\mathbf{e}, \tilde{\mathbf{K}}_y, \tilde{\mathbf{K}}_r^T) = \mathbf{e}^T\mathbf{P}\mathbf{e} + \mathbf{tr}(\tilde{\mathbf{K}}_y^T \boldsymbol{\Gamma}_y^{-1} \tilde{\mathbf{K}}_y^T |\boldsymbol{\Lambda}|) + \mathbf{tr}(\tilde{\mathbf{K}}_r^T \Gamma_r^{-1} \tilde{\mathbf{K}}_r |\boldsymbol{\Lambda}|)$$

Stability Results

See Proof in *Ogunmolu et al. IROS 2018*. In the end, we have

$$\begin{aligned}\dot{\mathbf{V}}(\cdot) &= -\mathbf{e}^T\mathbf{Q}\mathbf{e} - 2\mathbf{e}^T\mathbf{P}\mathbf{B}\boldsymbol{\Lambda}\varepsilon_f \\ &\leq -\lambda_{low}\|\mathbf{e}\|^2 + 2\|\mathbf{e}\|\,\|\mathbf{P}\mathbf{B}\|\lambda_{high}(\boldsymbol{\Lambda})\varepsilon_{max}\end{aligned}$$

- $\lambda_{low}, \lambda_{high} \equiv$ minimum and maximum eigenvalues of Q and $\boldsymbol{\Lambda}$ respectively.
- $\dot{\mathbf{V}}(\cdot)$ is thus negative definite outside the compact set
- $\chi = \left(\mathbf{e} : \|\mathbf{e}\| \leq \dfrac{2\|\mathbf{P}\mathbf{B}\|\lambda_{high}(\boldsymbol{\Lambda})\varepsilon_{max}(\mathbf{y})}{\lambda_{low}(\mathbf{Q})}\right)$
  - *i.e.* the error $\mathbf{e}$ is uniformly ultimately bounded, or $\mathbf{y}(t) \to 0$ as $t \to \infty$.

Results

- Solving the general form of the lyapunov equation, we have

$$\mathbf{P} = \begin{bmatrix} -\frac{170500}{2668} & 0 & 0 \\ 0 & -\frac{170500}{2668} & 0 \\ 0 & 0 & -\frac{170500}{2668} \end{bmatrix}$$

- Solenoid valves operate in pairs
  - set

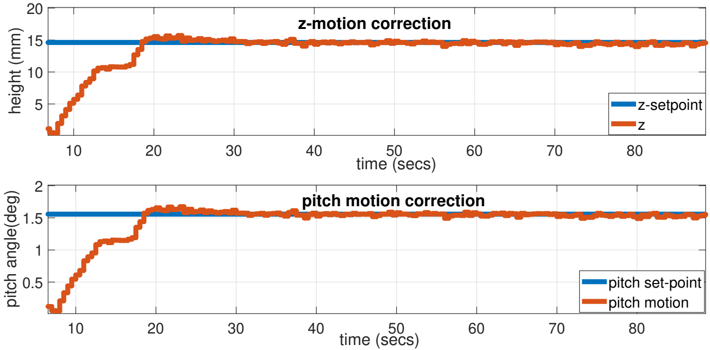
$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

- B maps to the 3-axes controllers

$$\begin{bmatrix} u_z & u_\theta & u_\psi \end{bmatrix}^T$$

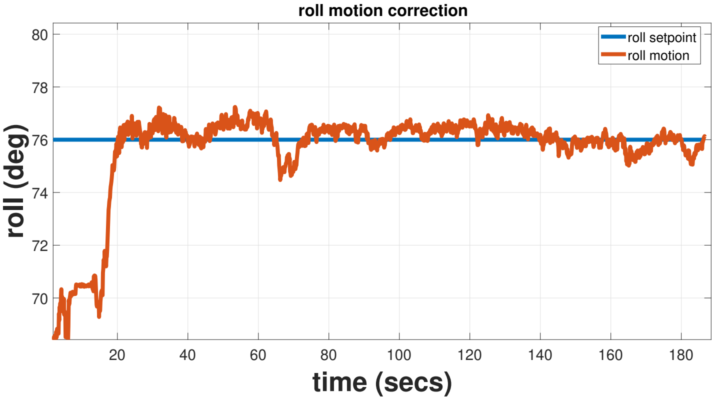
- non-zero terms are the max. duty-cycle to valves

Results: Z and Pitch Motions



Goal command: $(z, \theta, \phi) = (2.5mm, 0.25^\circ, 35^\circ)$ to $(14mm, 1.6^\circ, 45^\circ)^T$.

Results: Roll Motion



Head roll tracking.

Part III: Robustness Margins and Robust Deep Policies

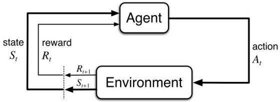
- Robustness Margins and Robust Deep Policies for Nonlinear Control

Relevant Papers

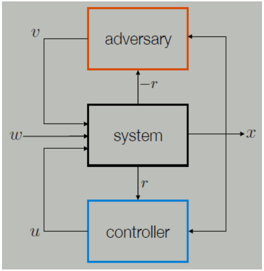
- Ogunmolu, Olalekan, Nicholas Gans, and Tyler Summers. "Minimax iterative dynamic game: Application to nonlinear robot control tasks." In 2018 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6919-6925. IEEE, 2018.
- Ogunmolu, Olalekan, Nicholas Gans, and Tyler Summers. "Robust zero-sum deep reinforcement learning." *arXiv preprint arXiv:1710.00491* (2017).
- Summers, Tyler, Olalekan Ogunmolu, Nicholas Gans. "Robustness Margins and Robust Guided Policy Search for Deep Reinforcement Learning." In *IEEE/RSJ International Conference on Robots and Intelligent Systems,(Abstract Only Track)*, vol. 8. 2017.

The robustness conundrum

- How to know *a priori* a policy's robustness limits?



- How to inculcate robustness into multistage decision policies?



Problem Setup

- To quantify the brittleness, we optimize the stage cost

$$\max_{\mathbf{v}_t \sim \psi \in \Psi} \left[\sum_{t=0}^{T} \underbrace{c(\boldsymbol{x}_t, \mathbf{u}_t)}_{\text{nominal}} - \gamma \underbrace{g(\mathbf{v}_t)}_{\text{adversarial}} \right]$$

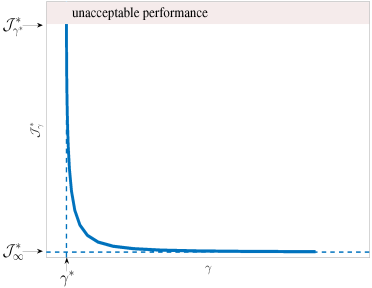
- To mitigate lack of robustness, we optimize the *cost-to-go*

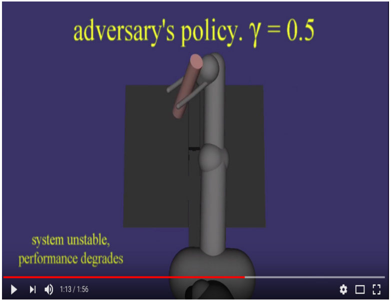
$$\mathcal{J}_t(\boldsymbol{x}_t, \pi, \psi) = \min_{\mathbf{u}_t \sim \boldsymbol{\pi}} \max_{\mathbf{v}_t \sim \psi} \left(\sum_{t=0}^{T-1} \ell_t(\boldsymbol{x}_t, \mathbf{u}_t, \mathbf{v}_t) + L_T(\boldsymbol{x}_T) \right),$$

- and seek a saddle point equilibrium policy that satisfies

$$\mathcal{J}_t(\boldsymbol{x}_t, \pi^*, \psi) \le \mathcal{J}_t(\boldsymbol{x}_t, \pi^*, \psi^*) \le \mathcal{J}_t(\boldsymbol{x}_t, \pi, \psi^*),$$

Results: Brittleness Quantification





Results: Iterative Dynamic Game

$\mathbf{y}_1^\star$

$\mathbf{y}_2^\star$

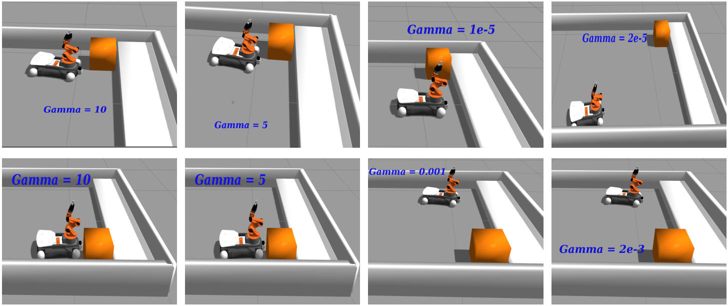


Table: *

End pose of the KUKA platform with our iDG formulation given different goal states and γ-values

End of Slides

Thank you!

Future Work: MRI/RT Immobilization

- Explore multiple parallel robot mechanisms for head motion correction
- Adopt best practices from one of the SOTA modeling approaches viz
  - *constant curvature approach* [Godage et al. (2016)],
  - the *continuous Cosserat approach* [Renda et al. (2014)], and
  - the *multi-body hyper-redundant model* [Kang et al. (2012)].

Future Work

- Robust Stabilization via inverse optimality
  - Imbibing robust Lyapunov stability into deep optimal controllers
    - Build on Freeman and Kokotovic's point-wise min-norm robust control lyapunov function to realize a meaningful value function (Int. Journal of Optimal Control, 1996)
    - Simplify algorithm for real-time control
  - Carry out numerical and and experimental verification and validation
  - Important in multistage decision policies, reinforcement learning controllers

Johannes Heinrich, Marc Lanctot, and David Silver. Fictitious self-play in extensive-form games. In *International Conference on Machine Learning*, pages 805–813, 2015.

Radhini Chelvarajah, Brigid Leighton, Linda Martin, Wayne Smith, and Rachael Beldham-Collins. Cranial immobilisation—is there a better way? *Radiographer*, 51(1):29–33, 2004.

Andrew Belcher. *Patient Motion Management with 6-DOF Robotics for Frameless and Maskless Stereotactic Radiosurgery*. PhD thesis, The University of Chicago, 2017.

Olalekan Ogunmolu and Rodney Wiersma. A Real-Time Patient Head Motion Correction Mechanism for MRI-Linac Systems. In *Joint AAPM | COMP Meeting, Virtual*, 2020.

Isuru S Godage, Gustavo A Medrano-Cerda, David T Branson, Emanuele Guglielmino, and Darwin G Caldwell. Dynamics for variable length multisection continuum arms. *The International Journal of Robotics Research*, 35(6):695–722, 2016.

Federico Renda, Michele Giorelli, Marcello Calisti, Matteo Cianchetti, and Cecilia Laschi. Dynamic model of a multibending soft robot arm driven by cables. *IEEE Transactions on Robotics*, 30(5):1109–1122, 2014.

Rongjie Kang, David T Branson, Emanuele Guglielmino, and Darwin G Caldwell. Dynamic modeling and control of an octopus inspired multiple continuum arm robot. *Computers & Mathematics with Applications*, 64(5):1004–1016, 2012.