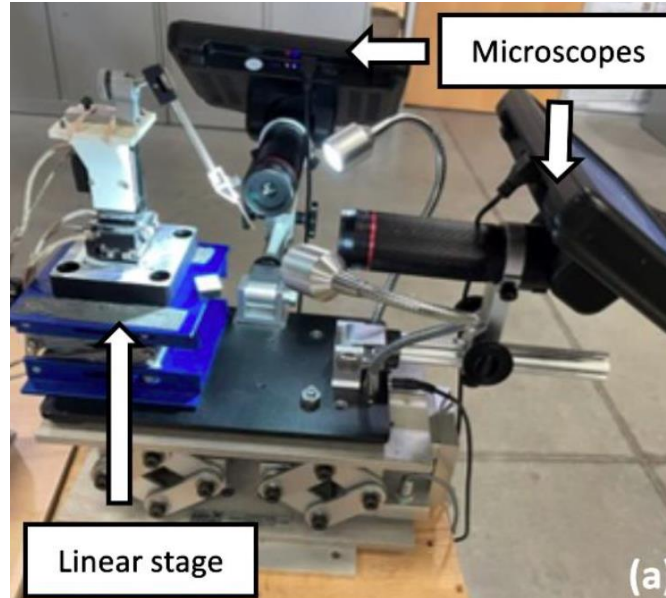


Background Reading Review

Team 16: Deep Learning Kinematics for Continuum Wire
Manipulator

Wenxuan Li, Zheyu Zhou, Zheyuan Zhang

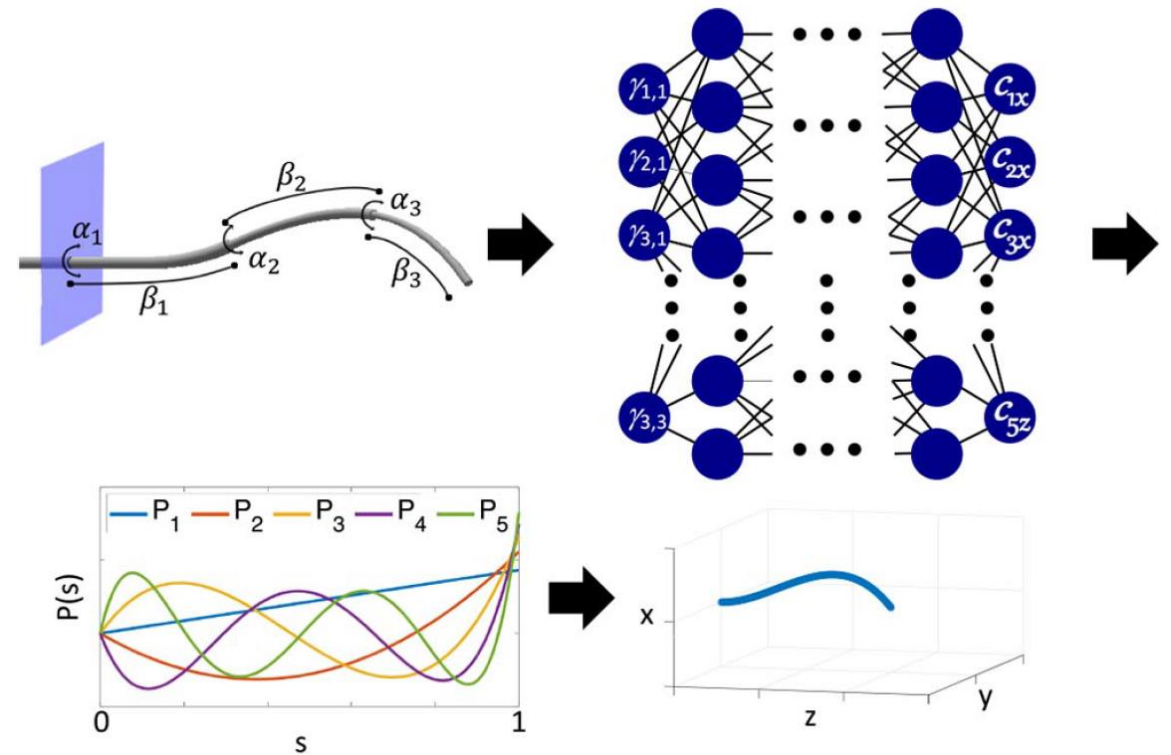
Project Overview



- The project goal is to develop a Deep Learning based kinematics model for the novel continuum wire manipulator for future surgery implementation.

Paper Selection 1

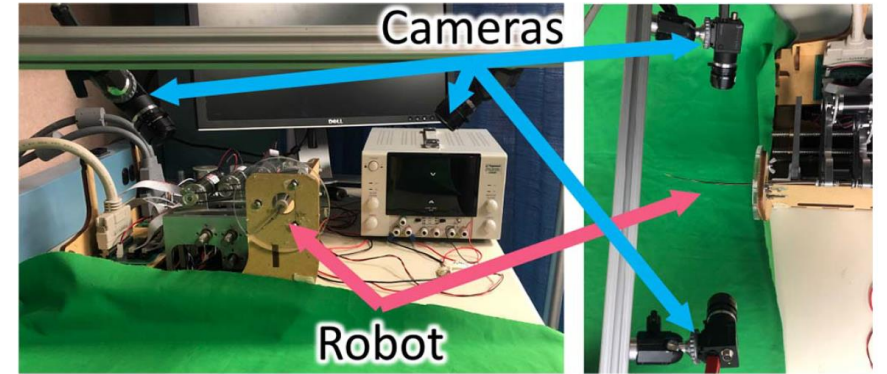
- Paper 1: Learning the Complete Shape of Concentric Tube Robots
- Author: Alan Kuntz , Armaan Sethi, Robert J. Webster, III, and Ron Alterovitz
- Conference: IEEE Transactions on Medical Robotics and Bionics (May, 2020)
- Relevance:
 - Introduced data-driven, deep-learning based method to estimate the kinematics of concentric tube soft robot.
 - Discussed Shape from Silhouette Computer vision technique to obtain ground truth data
 - Discussed Neural Network model and architecture that is optimal in accuracy



(Kuntz, Sethi, Webster III, & Alterovitz, 2020)

Technical Approach(1)

- Shape from Silhouette
 - Generate ground truth, real world 3D data from 2D images

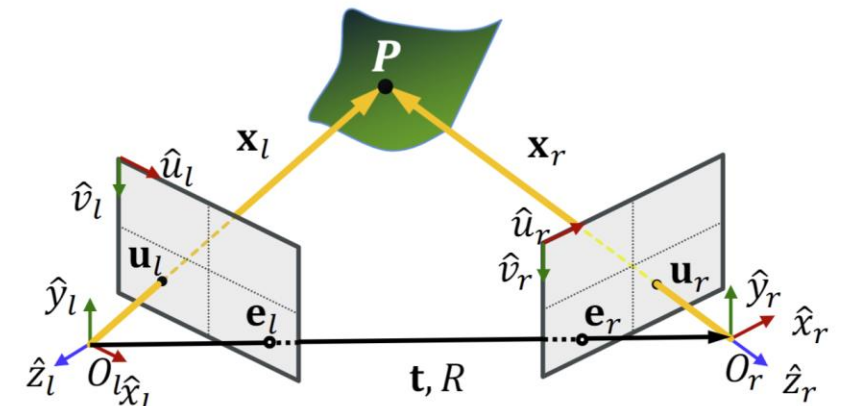


(Kuntz, Sethi, Webster III, & Alterovitz, 2020)

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X_o \\ Y_o \\ Z_o \\ 1 \end{bmatrix} = \mathbf{K} {}^c \mathbf{M}_o \begin{bmatrix} X_o \\ Y_o \\ Z_o \\ 1 \end{bmatrix}$$

$${}^{c_2} \mathbf{M}_{c_1} = {}^{c_2} \mathbf{M}_o \cdot {}^o \mathbf{M}_{c_1} = {}^{c_2} \mathbf{M}_o \cdot ({}^{c_1} \mathbf{M}_o)^{-1} = \begin{bmatrix} {}^{c_2} \mathbf{R}_o & {}^{c_2} \mathbf{t}_o \\ \mathbf{0}_{3 \times 1} & 1 \end{bmatrix} \cdot \begin{bmatrix} {}^{c_1} \mathbf{R}_o^T & -{}^{c_1} \mathbf{R}_o^T \cdot {}^{c_1} \mathbf{t}_o \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}$$

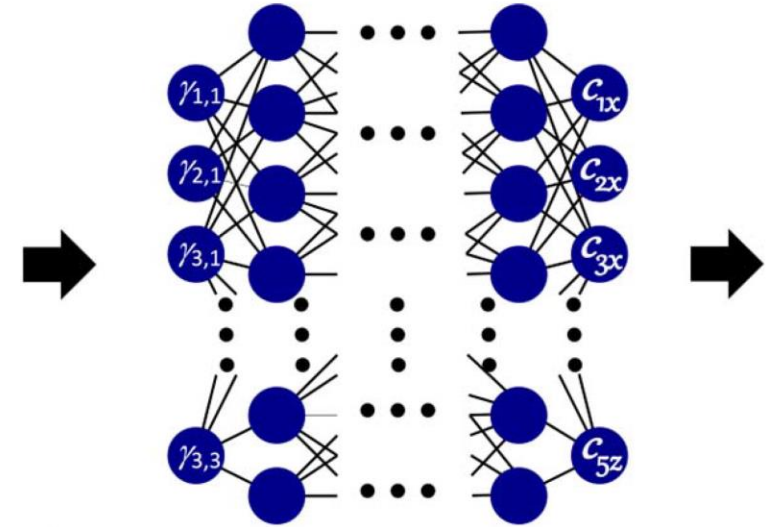
(OpenCV, (n.d.))



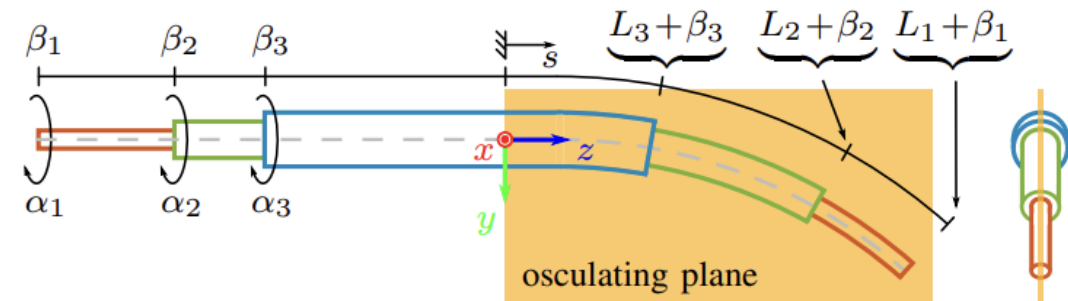
(Jones, 2023)

Technical Approach(2)

- Neural Network Model
 - Feed forward, Fully Connected
 - Parametric rectified linear unit activation function
 - Input
 - $\gamma_i := \{\gamma_{1,i}, \gamma_{2,i}, \gamma_{3,i}\} = \{\cos(\alpha_i), \sin(\alpha_i), \beta_i\}$
 - Output
 - 15 coefficients for 5 orthonormal polynomial basis functions in x, y, and z



(Kuntz, Sethi, Webster III, & Alterovitz, 2020)



(Grassmann, Modes, & Burgner-Kahrs, 2018)

Experiment

- Goal: Find optimal Neural Network Architecture
 - Numbers of hidden layers varying from 3 to 7
 - Numbers of nodes per hidden layer varying from 15 to 60
 - Networks trained with different dataset (Sim, Real, Sim+Real)

- Result:

- The optimal Neural Network Architecture is 3 layers, 30 nodes per layer, using 100,000 simulated data pretrained model, and transfer learning with 10,000 real-world data

AVERAGE LEARNED MODEL ACCURACY RESULTS (MEAN ± STD IN MM), MAXIMUM DEVIATION ALONG SHAFT

Data Type	Model Architecture (width x depth)								
	3x15	3x30	3x60	5x15	5x30	5x60	7x15	7x30	7x60
Sim	6.22 ± 3.95	6.28 ± 3.93	6.32 ± 3.95	6.28 ± 3.94	6.30 ± 3.95	6.31 ± 3.94	6.32 ± 3.98	6.31 ± 3.94	6.32 ± 3.96
Real	3.93 ± 2.50	3.66 ± 2.40	4.08 ± 2.41	4.38 ± 2.43	3.81 ± 2.35	4.11 ± 2.46	3.95 ± 2.34	3.65 ± 2.31	4.68 ± 2.58
Sim+Real	3.69 ± 2.45	3.35 ± 2.39	3.49 ± 2.42	4.68 ± 2.90	3.48 ± 2.49	3.58 ± 2.49	3.56 ± 2.50	3.53 ± 2.48	3.40 ± 2.41

AVERAGE LEARNED MODEL ACCURACY RESULTS (MEAN ± STD IN MM), MEAN SQUARED ERROR ALONG SHAFT

Data Type	Model Architecture (width x depth)								
	3x15	3x30	3x60	5x15	5x30	5x60	7x15	7x30	7x60
Sim	12.2 ± 16.4	12.4 ± 16.4	12.3 ± 16.5	12.1 ± 16.5	12.3 ± 16.6	12.3 ± 16.6	12.3 ± 16.7	12.3 ± 16.5	12.3 ± 16.7
Real	4.73 ± 6.29	3.50 ± 5.13	5.03 ± 6.61	4.44 ± 4.89	3.49 ± 4.72	4.06 ± 5.21	3.96 ± 4.81	3.54 ± 4.64	4.96 ± 5.08
Sim+Real	3.60 ± 5.44	3.03 ± 4.84	3.31 ± 5.27	5.94 ± 8.56	3.25 ± 5.53	3.47 ± 5.86	3.36 ± 5.63	3.36 ± 5.66	3.11 ± 5.24

AVERAGE LEARNED MODEL ACCURACY RESULTS (MEAN ± STD IN MM), SUM OF DEVIATIONS ALONG SHAFT

Data Type	Model Architecture (width x depth)								
	3x15	3x30	3x60	5x15	5x30	5x60	7x15	7x30	7x60
Sim	48.7 ± 21.8	48.9 ± 22.0	48.8 ± 21.8	48.6 ± 21.7	48.8 ± 22.0	48.8 ± 21.8	48.8 ± 22.0	48.7 ± 21.8	48.8 ± 22.0
Real	31.2 ± 13.7	25.8 ± 10.9	33.2 ± 12.0	30.1 ± 11.5	25.8 ± 10.6	28.4 ± 10.7	28.6 ± 11.3	26.8 ± 10.7	31.8 ± 11.5
Sim+Real	25.7 ± 11.4	23.2 ± 10.9	24.5 ± 11.2	33.5 ± 14.3	24.3 ± 11.0	25.1 ± 11.5	24.7 ± 10.9	24.8 ± 11.1	23.8 ± 10.8

(Kuntz, Sethi, Webster III, & Alterovitz, 2020)

Experiment

- Goal: Learned Model v.s. Physics-Based Model
 - Learned model: 3x30 Sim+Real
 - Physical-Based model: Cosserat Rod Equations
 - 1,000 test configurations compare with ground truth
- Result:
 - The Learned model outperformed the Physical-Based model

Maximum Error Deviation

	Physics-Based (mm)	Learned (mm)
Minimum	1.11	0.61
Maximum	46.68	30.18
Mean	6.32 ± 3.95	3.35 ± 2.39

Mean Squared Error

	Physics-Based (mm)	Learned (mm)
Minimum	0.30	0.05
Maximum	246.62	82.58
Mean	12.32 ± 16.60	3.03 ± 4.84

Sum of Error Deviation

	Physics-Based (mm)	Learned (mm)
Minimum	7.79	3.11
Maximum	181.63	85.89
Mean	48.79 ± 21.87	23.23 ± 10.90

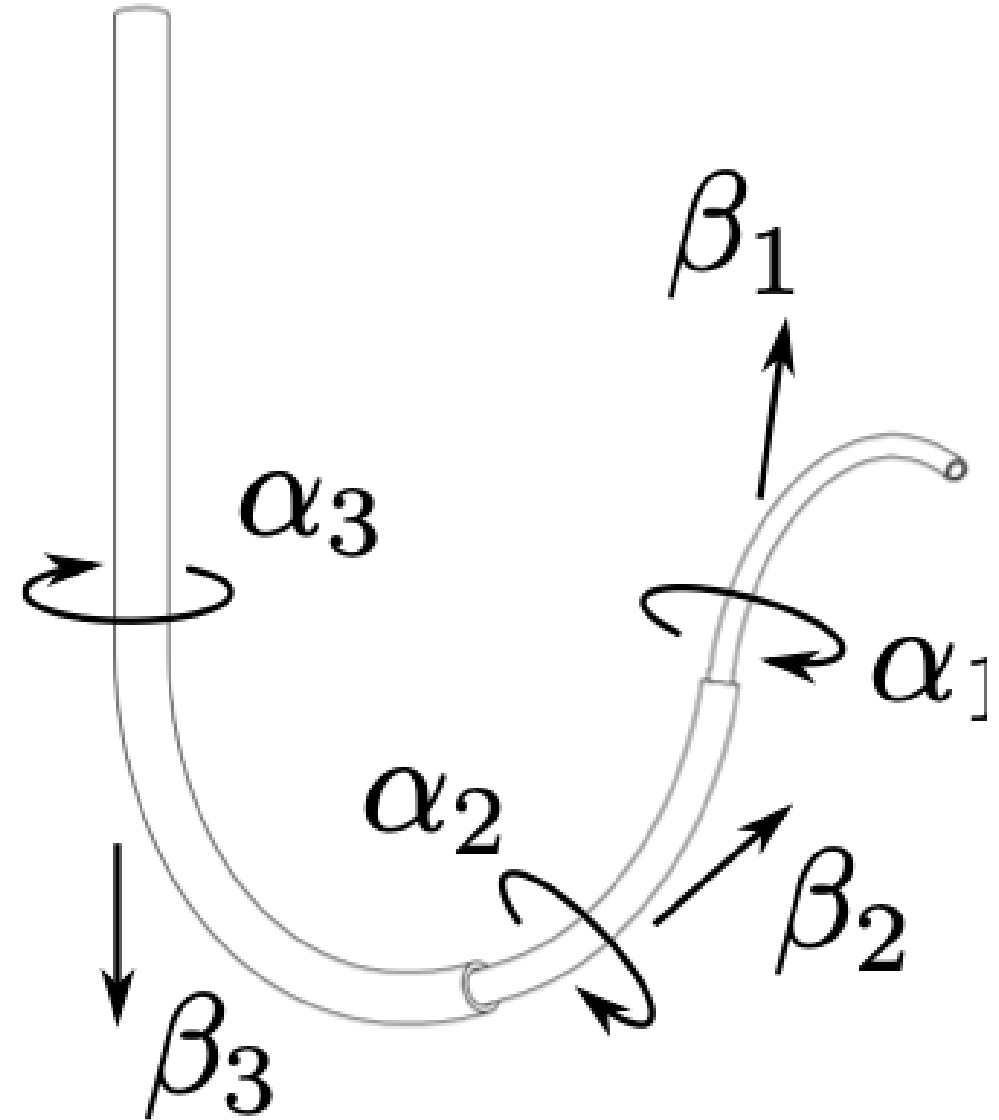
(Kuntz, Sethi, Webster III, & Alterovitz, 2020)

Critical Review

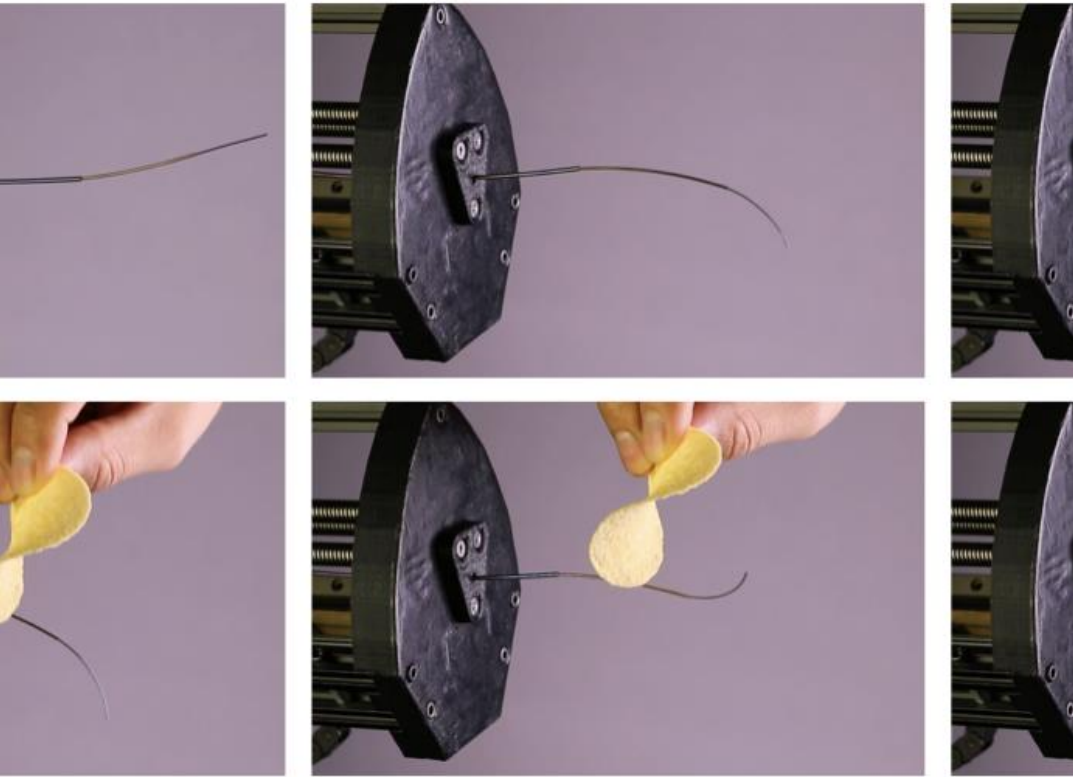
- Pros
 - Give an insight of a data-driven, deep-learning based method to model soft robot kinematics
 - Introduced a Computer Vision based technique to generate real-world soft robot 3D data
 - Explained the Neural Network input output model design
 - Showed the method of finding the optimal Neural Network architecture
- Cons
 - Other possible real-world 3D data generation method (ex. EM tracking)
 - Possible error in the CV generated ground truth 3D data
 - No comparison of Learned model on all-real data v.s. Physics-Based model

Paper Selection 2

- Paper 2:
On the Merits of Joint Space and Orientation Representations in Learning the Forward Kinematics in SE(3)
- Reinhard M. Grassmann, Jessica Burgner-Kahrs
- This paper investigates the influence of different joint space and orientation representations on the approximation of forward kinematics using shallow artificial neural networks.



Why do we choose this paper?



compliant reaction of a concentric tube continuum while interacting with a potato chip. Its highly non-linear characteristic of its for

(Grassmann & Burgner-Kahrs, 2019)

- Paper investigates different representations for approximating forward kinematics using artificial neural networks with ReLU.
- Our project involves development/testing of ROS-controlled robotic system requiring appropriate representations and neural networks for kinematics.
- Chosen as related work for guidance on appropriate neural networks for approximating forward kinematics in robotic system.

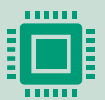
Background



Neural networks can be used to approximate the forward kinematics function.



Joint space and orientation representation can significantly impact the input-output characteristics of the feedforward networks.



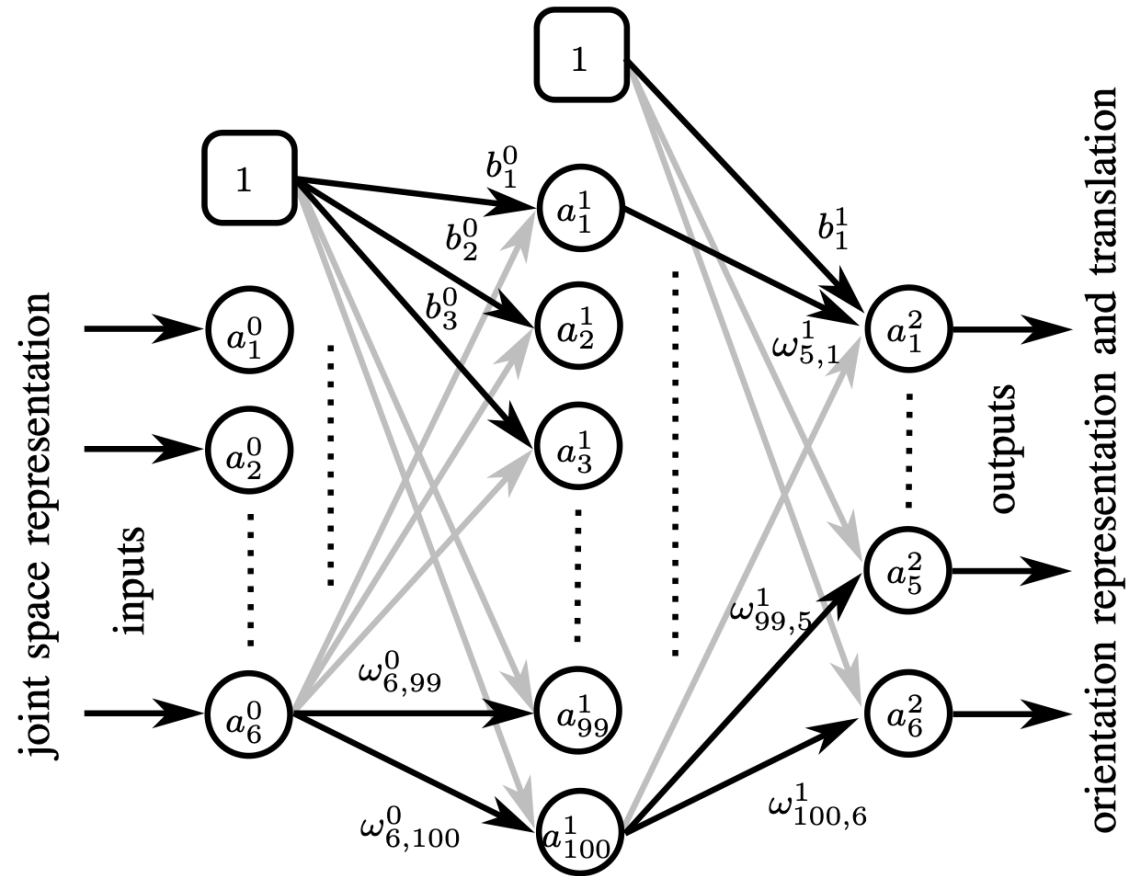
Suitable inputs and outputs can make feedforward networks more efficient in robotic applications.



Studies on forward kinematics approximation have mostly neglected orientations, which is a major drawback in real-world scenarios.

Relevant Technical Approach

- Feedforward networks are used to approximate the forward kinematics in robotic systems.
- The ReLU activation function is used in the hidden layer for faster training.
- Linear activation functions are used for the output layer.
- The weights of the ReLU activation functions are initialized with HE-initialization.
- The Adam optimizer is used with a mini-batch size of 128 to optimize the weights and biases of the feedforward network.
- Second-order optimizers are impractical in almost all useful cases.



(Grassmann & Burgner-Kahrs, 2019)

Significance

- ⑩ Feedforward networks are used to approximate the forward kinematics in robotic systems by approximating smooth functions in a compact set.
- ⑩ The architecture of the feedforward network uses ReLU activation functions in the hidden layer for computational efficiency and fast training.
- ⑩ Adam optimizer converges faster than a vanilla gradient-descent optimizer and that neither under-fitting nor over-fitting occurred.

Critical Review

Pros:

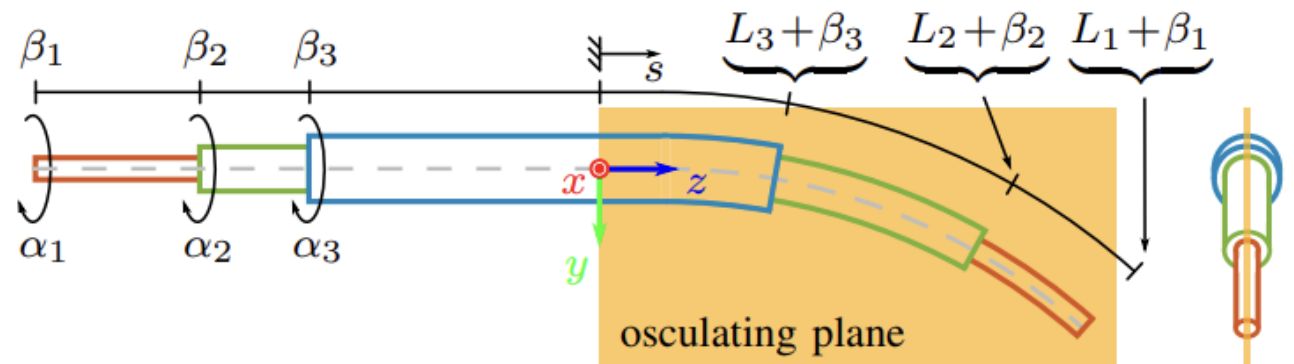
- Provides with a series of workable settings -> as our prior setting up
- Provides with a series of real-world concerns -> take into consideration

Cons:

- Lack of potential limitations or challenges associated with using feedforward networks
- Lack of GPU settings and conditions of conducting experiments
- Lack of environments that code built in
- Findings of other methods part are too specific to apply, which seem irrelevant to our project

Paper Selection 3

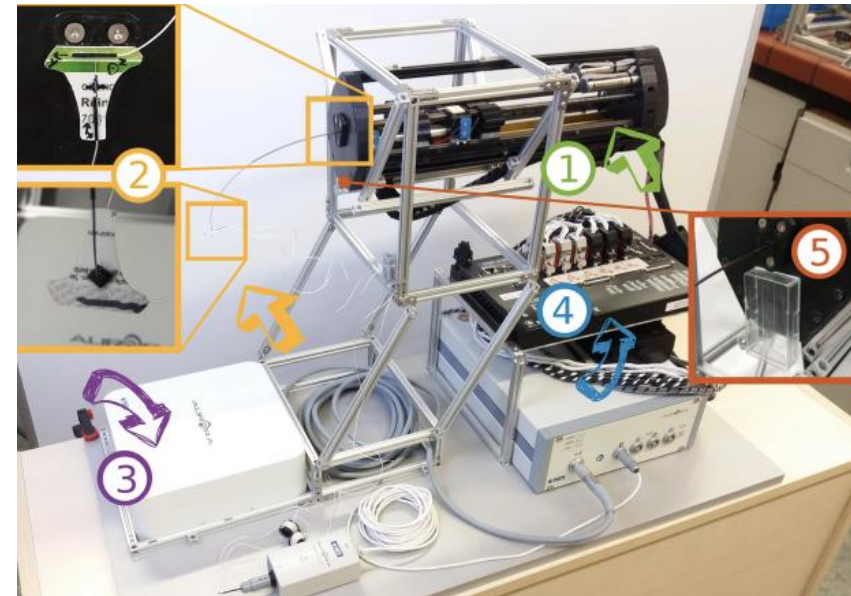
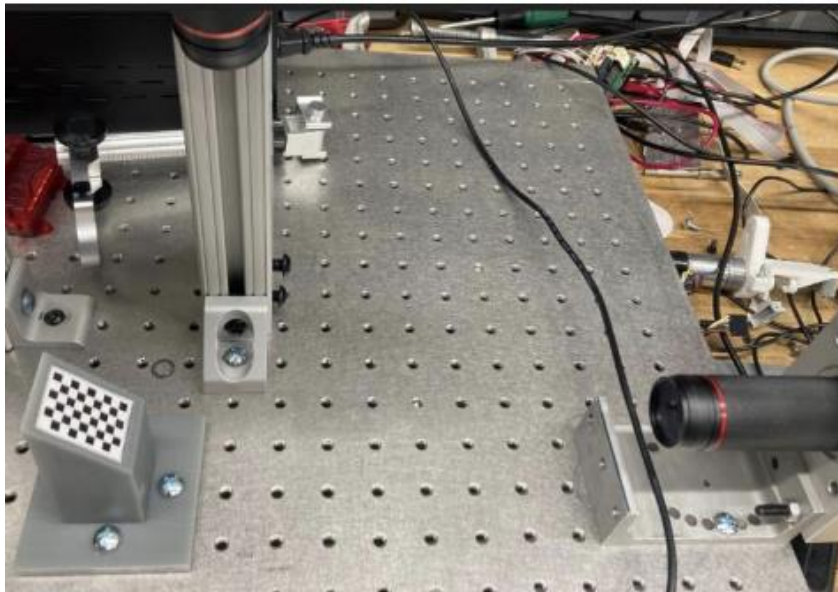
- Paper 3: Learning the Forward and Inverse Kinematics of a 6-DOF Concentric Tube Continuum Robot in SE(3)
- Reinhard Grassmann, Vincent Modes, and Jessica Burgner-Kahrs
- The paper presents a deep learning-based approach to estimate the kinematics of a concentric tube continuum robot, addressing the non-linearities in the elastic interactions between the pre-curved tubes that make modeling their kinematics challenging.



(Grassmann, Modes, & Burgner-Kahrs, 2018)

Why do we choose this paper?

- We chose this paper as a reference for studying the kinematics of a pre-shaped Nitinol wire manipulator for minimal invasive eye surgery, as both robots share similar challenges in accurately predicting their kinematics due to their deformable nature.
- The paper also provides insights into the construction of an experimental platform and the collection of data, which can serve as an inspiration for developing a similar experimental setup and collecting data for accurately predicting the kinematics of the Nitinol wire manipulator.



Background and Significance

- **Background**

- Traditional physics-based models are computationally demanding and require highly non-linear modeling efforts, making a data-driven approach an asset.
- Artificial neural networks using activation functions to approximate kinematics, using measurement data to train, validate and test the proposed approach proved to be effective for soft-bodied robots.

- **Significance**

- Effective platform building and data processing methods with data driving approach.
- Demonstrate the effectiveness of using feedforward neural networks with ReLU activation functions for kinematic approximation

Technical Approach (1)

$$\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \eta^{(1)} & \epsilon_1^{(1)} & \epsilon_2^{(1)} & \epsilon_3^{(1)} \\ \eta^{(2)} & \epsilon_1^{(2)} & \epsilon_2^{(2)} & \epsilon_3^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ \eta^{(i)} & \epsilon_1^{(i)} & \epsilon_2^{(i)} & \epsilon_3^{(i)} \\ \vdots & \vdots & \vdots & \vdots \\ \eta^{(n-1)} & \epsilon_1^{(n-1)} & \epsilon_2^{(n-1)} & \epsilon_3^{(n-1)} \\ \eta^{(n)} & \epsilon_1^{(n)} & \epsilon_2^{(n)} & \epsilon_3^{(n)} \end{bmatrix} \begin{bmatrix} \eta \\ \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix}$$

$$\mathbf{t} = \frac{1}{n} \sum_{i=1}^n \mathbf{t}^{(i)},$$

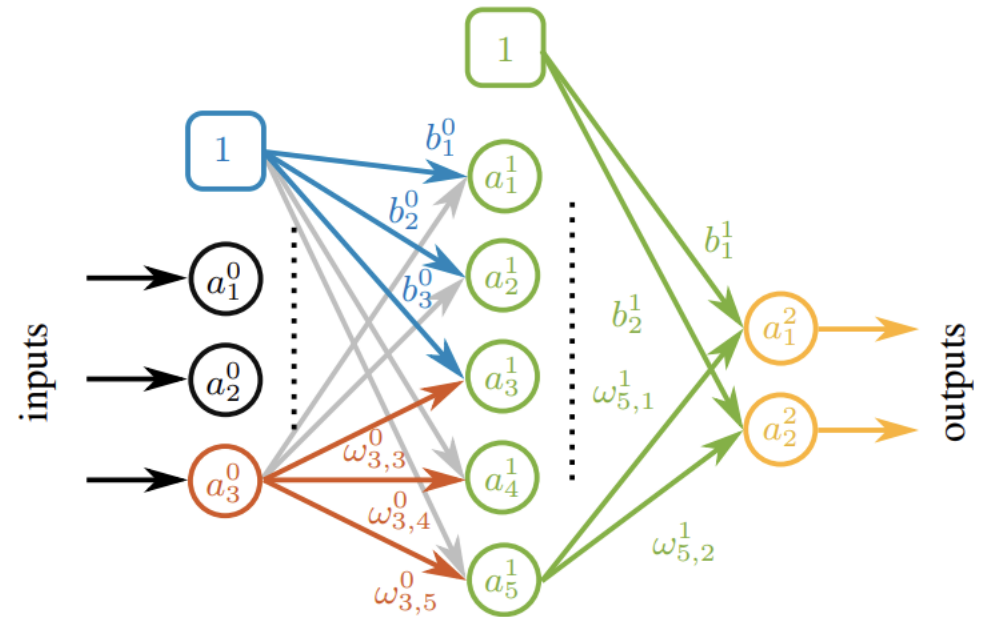
$$[\eta \quad \epsilon_1 \quad \epsilon_2 \quad \epsilon_3]^T = \left(\mathbf{Q}_m^T \mathbf{Q}_m \right) \mathbf{Q}_m^T \mathbf{1}$$

Random sampling method of joint space Q to generate data.
Multiple measurements of end-effector poses per sample to reduce error.

(Grassmann, Modes, & Burgner-Kahrs, 2018)

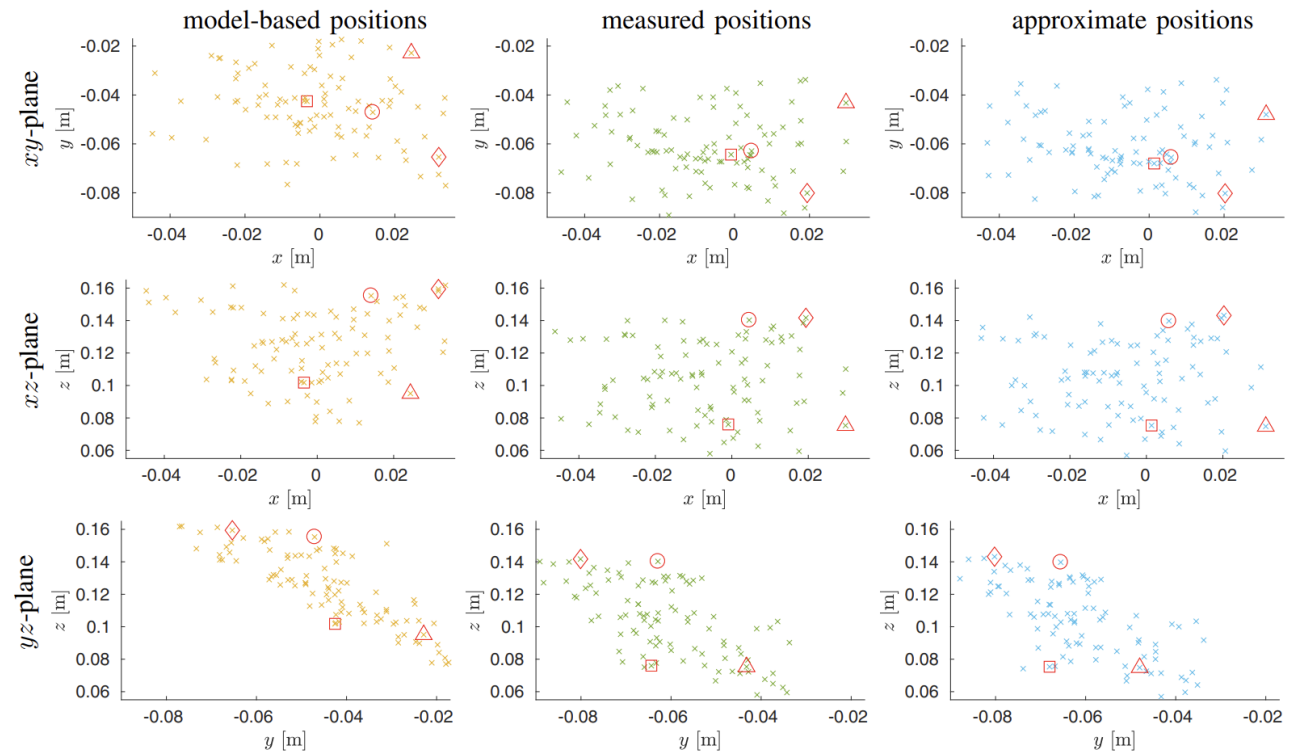
Technical Approach (2)

- Feedforward neural networks with ReLU activation functions used for kinematics approximation.
- Number of neurons and layers determined in preliminary examinations, with HE-initialization method for weight initialization and Adam optimizer for optimization.
- Results compared with test set and Kirchhoff rod theory, with observation that trigonometric functions improve accuracy and neural network approach is faster and more accurate than traditional models.



Results

- Figure shows 100 positions taken from the test set Stest. They are compared with positions calculated by a Kirchhoff based unloaded kinematic model.
- The model-based approach achieves an approximation error in position $e_x = 27.4 \pm 5.1$ mm and in orientation $e_\theta = 88.3 \pm 37.3$ °. The approximated pose of the feedforward network results in approximation errors of $e_x = 2.7 \pm 1.2$ mm and $e_\theta = 1.3 \pm 0.9$ °.



Critical Review

- **Pros:**
 - Accurate kinematics estimation of manipulator
 - Experimental platform for data collection
 - Feedforward neural network with ReLU activation function
 - Insights into transfer learning method
 - Evaluation of forward kinematics approximation
- **Cons:**
 - Lack of discussion on comparison of data-driven approach with the Physics-Based model
 - Insufficient discussion on robustness and reliability of the test system
- **Potential Next Steps:**
 - Conduct more experiments to evaluate robustness and reliability
 - Compare performance with other kinematic estimation methods
 - Explore combining data-driven and physics-based approaches

Conclusion

Generate Real-world Data

Learning the Complete Shape of Concentric Tube Robots

Learning the Forward and Inverse Kinematics of a 6-DOF Concentric Tube Continuum Robot in $SE(3)$

Design Neural Network Model


On the Merits of Joint Space and Orientation Representations in Learning the Forward Kinematics in $SE(3)$

Learning the Forward and Inverse Kinematics of a 6-DOF Concentric Tube Continuum Robot in $SE(3)$

Learning the Complete Shape of Concentric Tube Robots

Transfer Learning

Learning the Complete Shape of Concentric Tube Robots

- 
1. Kuntz, A., Sethi, A., Webster, R. J., 3rd, & Alterovitz, R. (2020). Learning the Complete Shape of Concentric Tube Robots. *IEEE transactions on medical robotics and bionics*, 2(2), 140–147. <https://doi.org/10.1109/tmrb.2020.2974523>
 2. Jones, C. (2023, March 9). *Uncalibrated Binocular Stereo* [Lecture]. Johns Hopkins. https://jhu.instructure.com/courses/35057/files/6581057?module_item_id=2693902
 3. Kong-man, C., Baker, S. and Kanade, T. (2005). Shape-from-silhouette across time part I: Theory and algorithms. *International Journal of Computer Vision*, vol. 62, no. 3, pp. 221–247.
 4. Grassmann, R.M., & Burgner-Kahrs, J. (2019). On the Merits of Joint Space and Orientation Representations in Learning the Forward Kinematics in SE(3). *Robotics: Science and Systems XV*.
 5. Grassmann, R., Modes, V., & Burgner-Kahrs, J. (2018). Learning the Forward and Inverse Kinematics of a 6-DOF Concentric Tube Continuum Robot in SE(3). In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 5125-5132).
 6. OpenCV (n.d.). *Basic concepts of the homography explained with code*. https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html