Improving Technical Proficiency in Robot-mediated Surgery Through Counterfactual Inquiry

Group 13: Hao, Ding

Mentor: Dr. Mathias Unberath

Motivation

- **Skill is among the strongest predictors for patient outcome**

  The higher surgeon skill, the better the outcome

- **Empower novice surgeons**

  "Translate" beginner-level commands to expert-level proficiency

- **Recent deep learning algorithm is powerful but vulnerable**

  Deep learning algorithms dominate large amount of benchmarks but suffer from generalization issues and low interpretability.

- **Causality brings more robustness and interpretability**

  With casual relations provided and causal inference mechanisms, we might be able to make the algorithm more robust and easier to interpret.

My Goal

- **Ultimate goal**

  Build a robust system to assist surgeons especially novice surgeons perform at expert-level by a simple query “Do something(task) somewhere(task context) as an expert”.

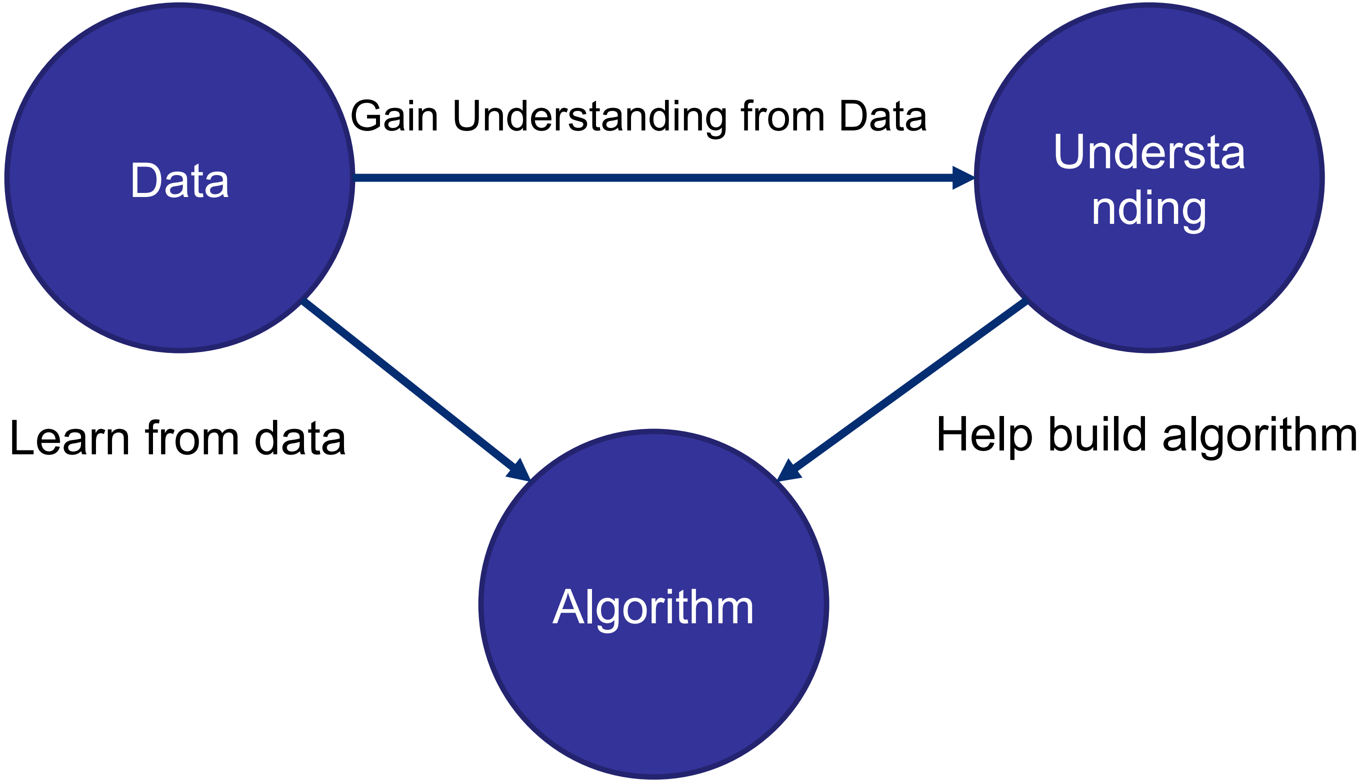
- **Practical subgoals**
  - Data

    Build a suitable dataset for causal inference study in surgery area.
  - Understandings

    Explore the difference between novice- and expert- level commands.
  - Algorithm

    Explore methods to incorporate causal inference mechanisms into deep learning algorithms of kinematic prediction for more robustness and interpretability.

Technical Approach - Overall



Technical Approach - Data

- Fundamentals: JHU-ISI Gesture and Skill Assessment Working Set（JIGSAW）

  The dataset was captured using the *da Vinci* Surgical System from eight surgeons with different levels of skill performing five repetitions of three elementary surgical tasks on a bench-top model

  The **JIGSAWS** dataset consists of three components:

  - kinematic data
  - video data
  - manual annotations: *gesture, skill level.*

Yixin Gao, S. Swaroop Vedula, Carol E. Reiley, Narges Ahmidi, Balakrishnan Varadarajan, Henry C. Lin, Lingling Tao, Luca Zappella, Benjam´ın B´ejar, David D. Yuh, Chi Chiung Grace Chen, Ren´e Vidal, Sanjeev Khudanpur and Gregory D. Hager, The JHU-ISI Gesture and Skill Assessment Working Set (JIGSAWS): A Surgical Activity Dataset for Human Motion Modeling, In Modeling and Monitoring of Computer Assisted Interventions (M2CAI) – MICCAI Workshop, 2014.

Technical Approach - Data

- Fundamentals: JIGSAW dataset
- Target: Temporally align JIGSAW samples for causal analysis
- Works to be done:
  - Manually segmenting and paring videos.
    Annotate the start and the end frame of each motion.
  - Making paired segments into same length
    Algorithm: dynamic time warping
    Distance function: l1 norm between normalized traveled distance

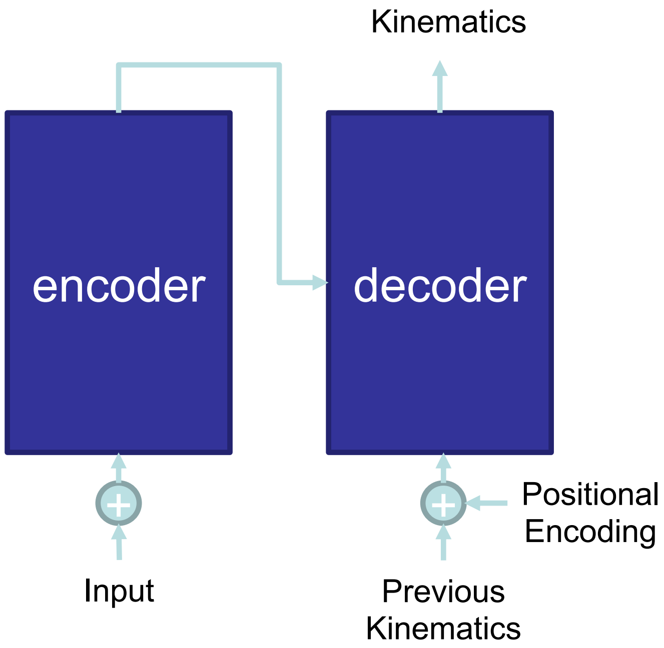
Technical Approach - Understandings

- Statistical analysis for some properties of the trajectory

    e.g. path length, time consumed, number of movements

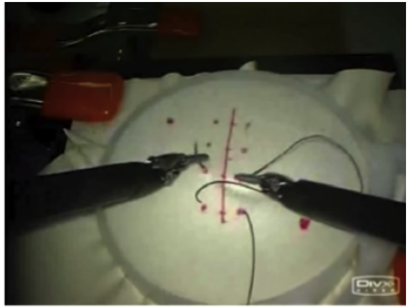
Technical Approach - Kinematic Prediction Network

- Architecture
  - Basic Transformer Network
- Input
  - Task information: task id and its context
  - Expected skill level: novice- or expert- level
  - Kinematics of previous frames
- Output
  - Kinematics of future frames

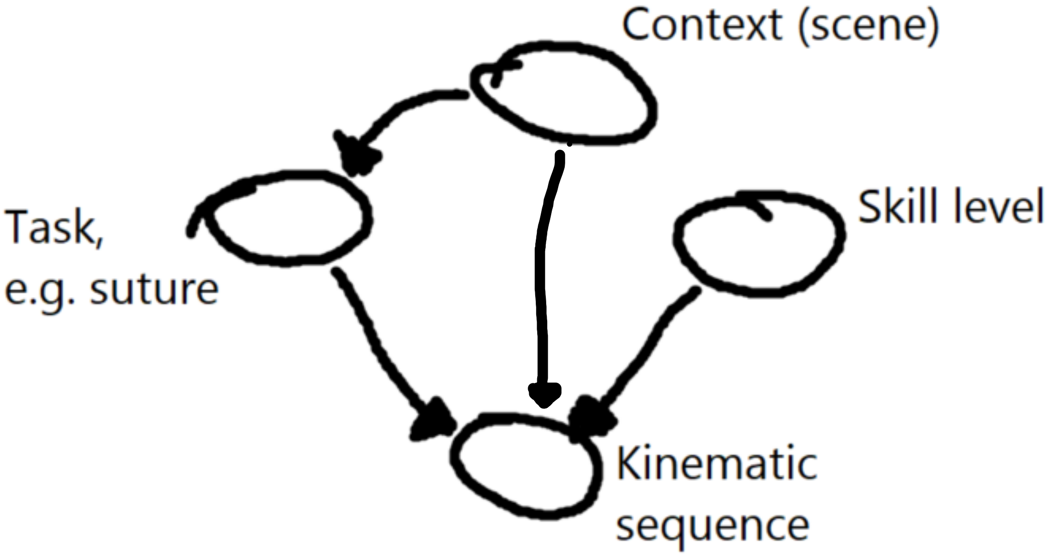


Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA, pages 5998–6008, 2017.

Technical Approach - Causal Inference



- **Treat the assistance as a counterfactual query**
  What would the robot commands have been if, contrary to fact, the surgeon were an expert?



Context: Phantom and its affordances

Technical Approach - Causal Inference

- Incorporate deep learning with the causal model

  There are several works that we can refer to.

  - Deep Structural Causal Models for Tractable Counterfactual Inference https://arxiv.org/abs/2006.06485
  - CausalGAN: Learning Causal Implicit Generative Models with Adversarial Training https://arxiv.org/pdf/1709.02023.pdf
  - Learning Functional Causal Models with Generative Neural Networks https://arxiv.org/abs/1709.05321
  - Counterfactual Generative Networks https://arxiv.org/pdf/2101.06046v1.pdf

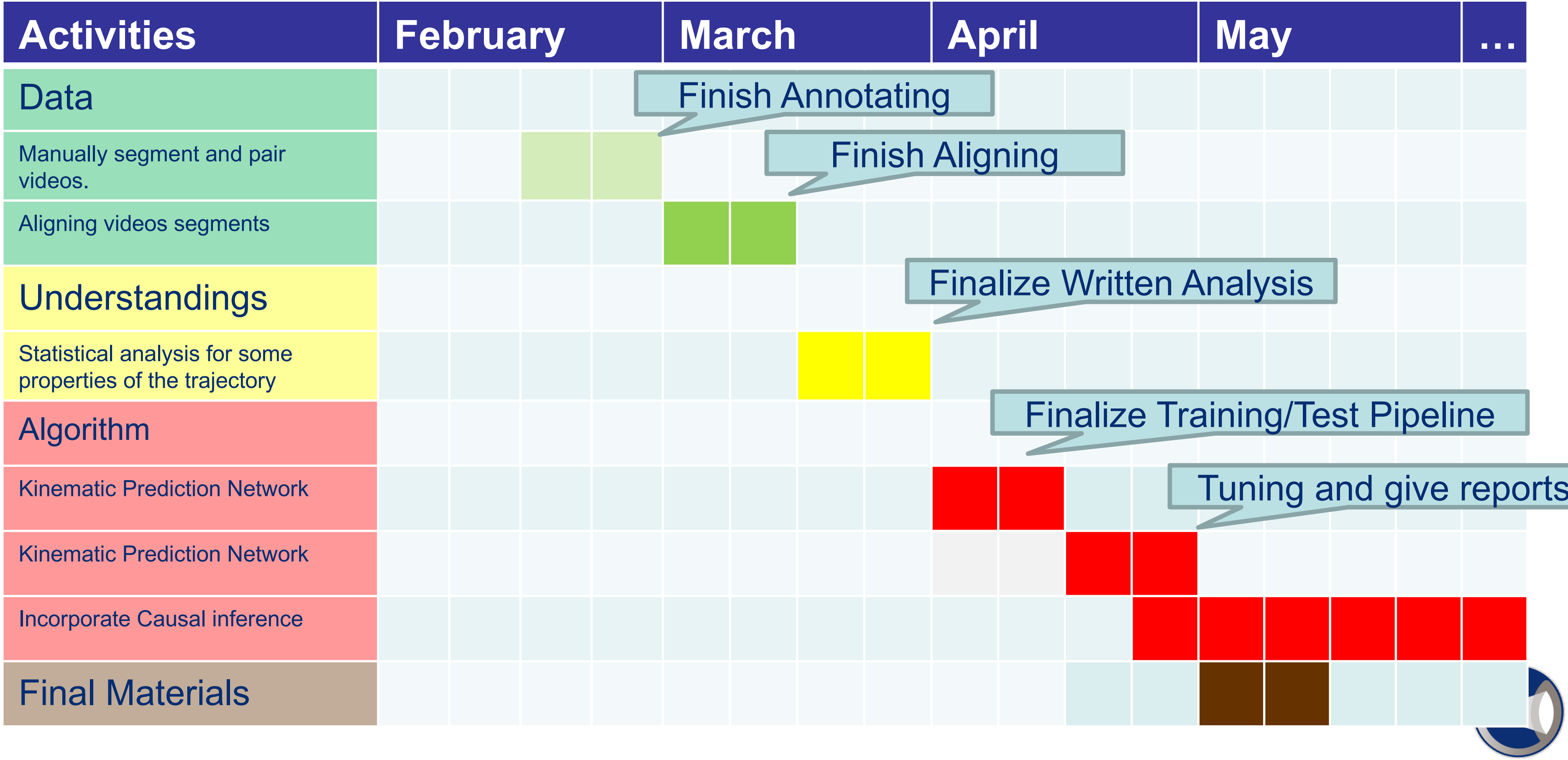
Dependencies

Dependency	Need	Status	Contingency Plan
JIGSAW dataset	Fundamental dataset	Acquired	N/A
Computational Resources	For Deep Learning experiments	Acquired	If crashed, ask Dr. Unberath for other computers, or acquire on some clouds resources (e.g. AWS).

Deliverables

Level	Activities	Deliverables
Minimal	Manually segment and pair videos. Implement DTW to align videos segments temporally	Task-aligned JIGSAW dataset for causal analysis
Expected	Statistical analysis for some properties of the trajectory.	Written analysis of understandings into novice- and expert-level robot command
Stretch	Implement the Kinematic Prediction Network Incorporate Causal inference mechanism into the DL methods	Development and evaluation of counterfactual model

Timeline


Activities
February
March
April
May
…
Data
Manually segment and pair videos.
Aligning videos segments
Understandings
Statistical analysis for some properties of the trajectory
Algorithm
Kinematic Prediction Network
Kinematic Prediction Network
Incorporate Causal inference
Final Materials
Finish Annotating
Finish Aligning
Finalize Written Analysis
Finalize Training/Test Pipeline
Tuning and give reports

Roles and Responsibilities

- The team:

  Hao Ding: Sole responsibility for all tasks

- Mentor:

  Dr. Mathias Unberath: Advising students for all tasks.

Management Plan

- Weekly meeting:

  Weekly meeting with Robotics group in ARCADE Lab including Dr. Mathias Unberath.

- Biweekly meeting:

  Biweekly personal meeting with Dr. Mathias Unberath.

Reading List

- Murat Kocaoglu, Christopher Snyder, Alexandros G. Dimakis, and Sriram Vishwanath. Causalgan: Learning causal implicit generative models with adversarial training. In 6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings. OpenReview.net, 2018

- hristos Louizos, Uri Shalit, Joris M. Mooij, David A. Sontag, Richard S. Zemel, and Max Welling. Causal effect inference with deep latent-variable models. In Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA, pages 6446–6456, 2017

- Nick Pawlowski, Daniel Coelho de Castro, and Ben Glocker. Deep structural causal models for tractable counterfactual inference. In Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual, 2020.

- Yu Kong and Yun Fu. Human action recognition and prediction: A survey. CoRR, abs/1806.11230, 2018

References

- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA, pages 5998–6008, 2017.

- Yixin Gao, S. Swaroop Vedula, Carol E. Reiley, Narges Ahmidi, Balakrishnan Varadarajan, Henry C. Lin, Lingling Tao, Luca Zappella, Benjam´ın B´ejar, David D. Yuh, Chi Chiung Grace Chen, Ren´e Vidal, Sanjeev Khudanpur and Gregory D. Hager, *The JHU-ISI Gesture and Skill Assessment Working Set (JIGSAWS): A Surgical Activity Dataset for Human Motion Modeling*, In Modeling and Monitoring of Computer Assisted Interventions (M2CAI) – MICCAI Workshop, 2014.

- Narges Ahmidi, Lingling Tao, Shahin Sefati, Yixin Gao, Colin Lea, Benjamın Bejar Haro, Luca Zappella, Sanjeev Khudanpur, Rene Vidal, Fellow, IEEE, Gregory D. Hager, *A Dataset and Benchmarks for Segmentation and Recognition of Gestures in Robotic Surgery,* Transaction of Biomedical Engineering, 2017.