Stroke Subject Dexterity Evaluation

Gavin Zhu

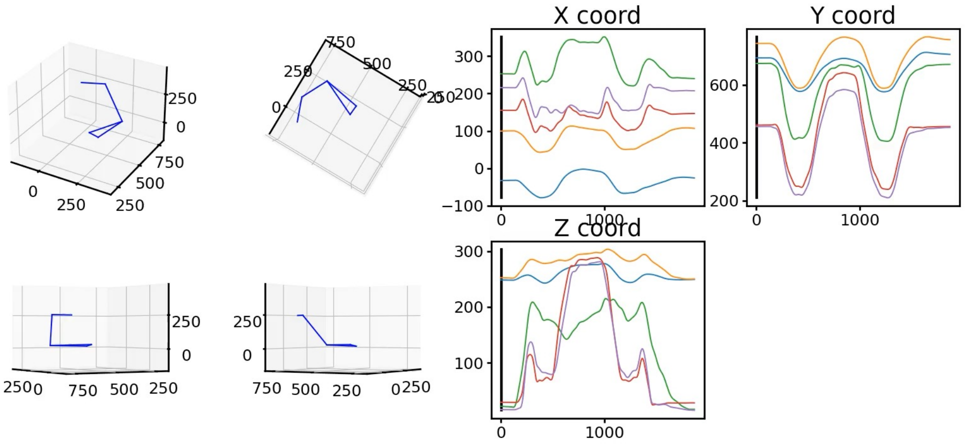
Lab Meeting

Sept 19th 2022

Background

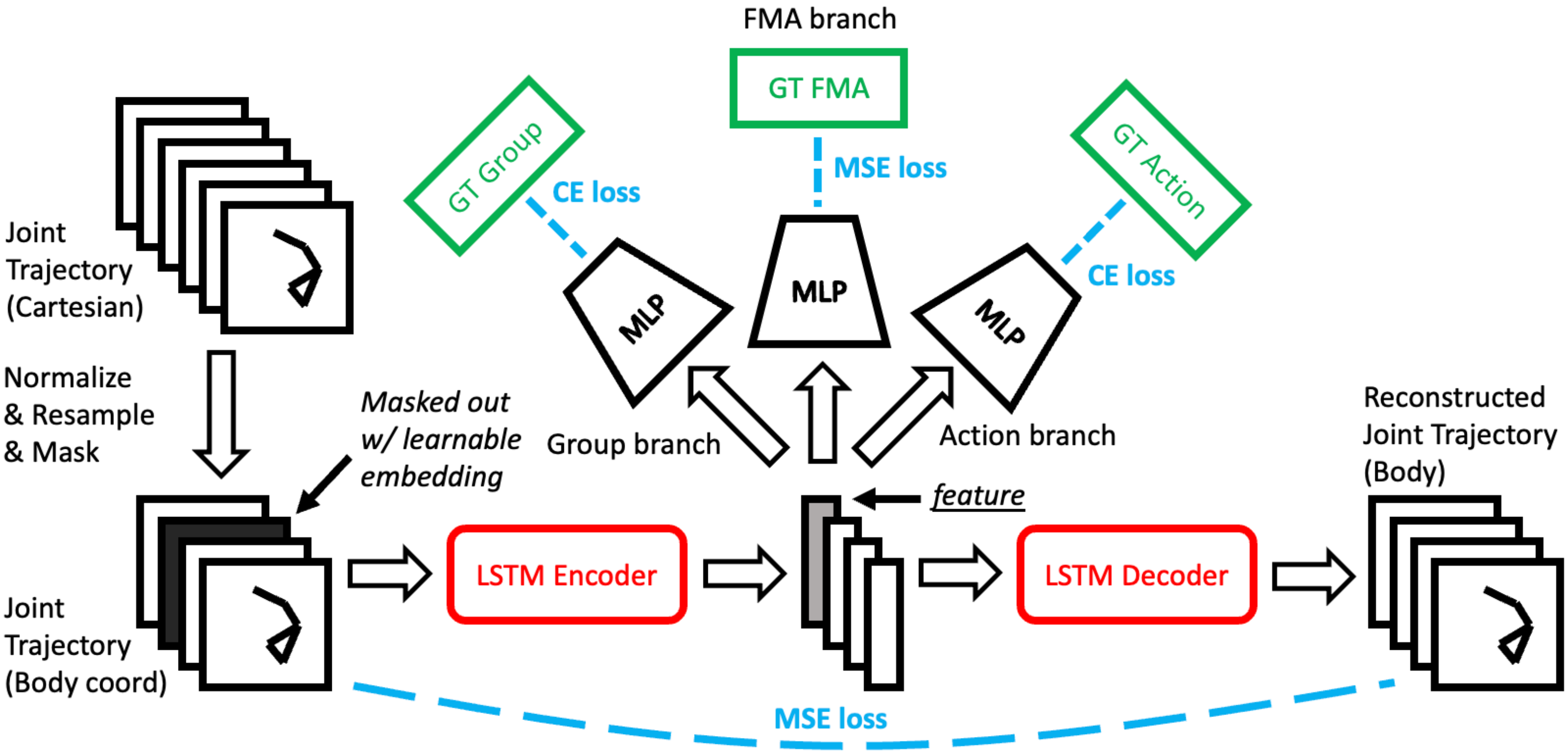
- Stroke causes great socio-economical burden
- Evaluation metric needed to guide personalized rehabilitation

Problem Definition: Building FMA Estimator for Motion Trajectories in The Wild



- **Fugl-Meyer Assessment:**
  - **Clinical dexterity evaluation base on a set of *script actions* on joints**
- **Project Goal:**
  - **Build FMA estimator for *unscripted actions* in daily activities (cooking)**
  - **(more generally) Construct feature descriptor for motion trajectories**

Current Approach: LSTM Based Feature Descriptor



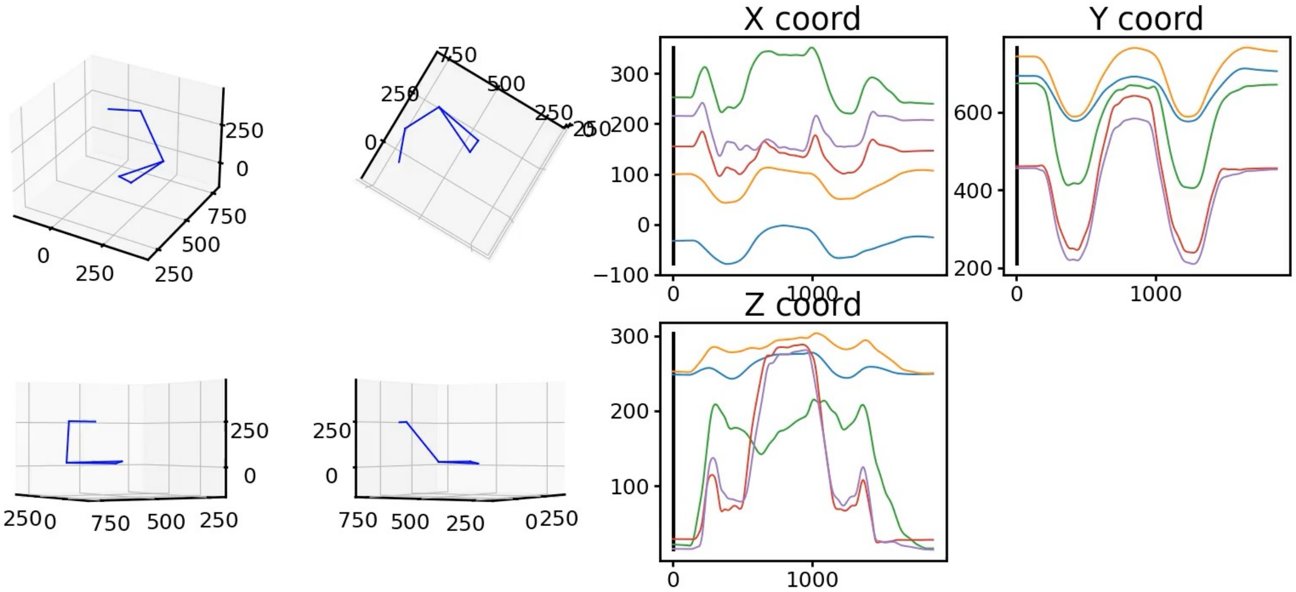
Preliminary Results on Downstream Tasks

	Action Classification (Accuracy ↑)		Group Classification (Accuracy ↑)		FMA Regression (MAE ↓)	
	Coordinates	FPCA	Coordinates	FPCA	Coordinates	FPCA
Adaboost	0.08	0.05	0.06	0.03	0.41	0.49
Random Forest	0.31	0.14	0.17	0.11	0.59	0.57
Linear SVM	0.20	0.15	0.13	0.09	0.50	0.51
MLP	0.04	0.12	0.03	0.16	0.51	0.52
Integrated	0.91		0.95		0.36	
Single Head	-		-		0.23	
Split by Subject	-		0.68		1.25	

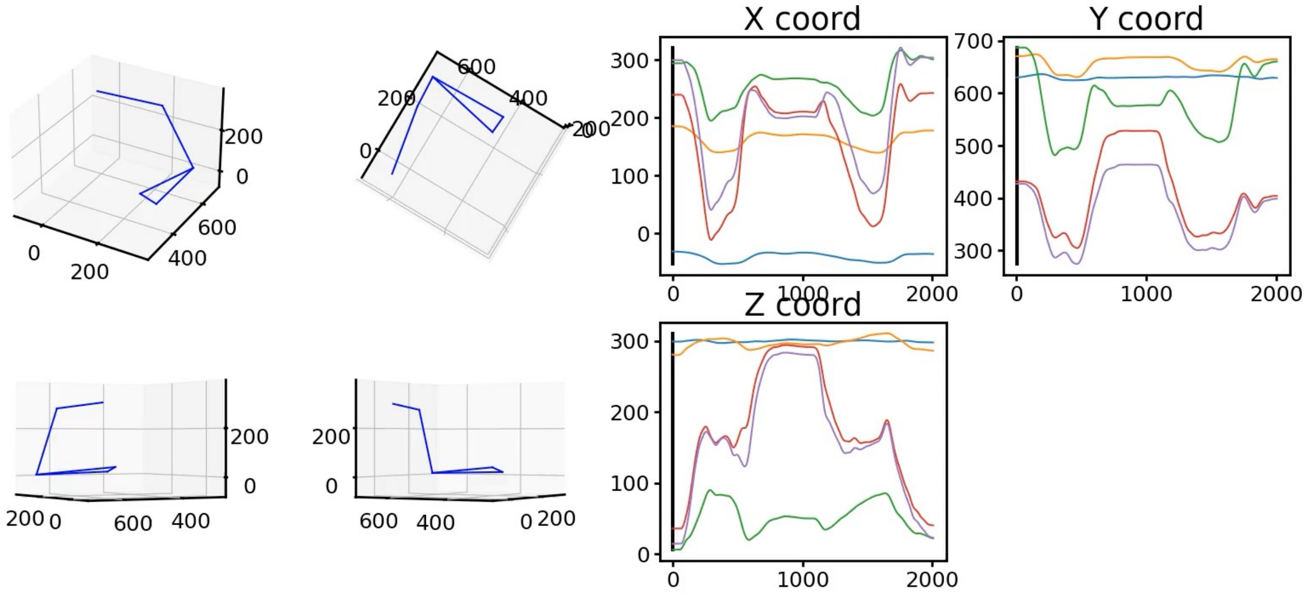
Problems & Potential Solutions

- Overfitting in by-subject split
  - Caused by limited data (only two with severe stroke)
  - More data augmentation
  - Augment action, original length, and/or other information at input
  - Local feature descriptor
- No large-scale unscripted action dataset for stroke patients
  - Learn healthy motion space from unscripted action of healthy subjects

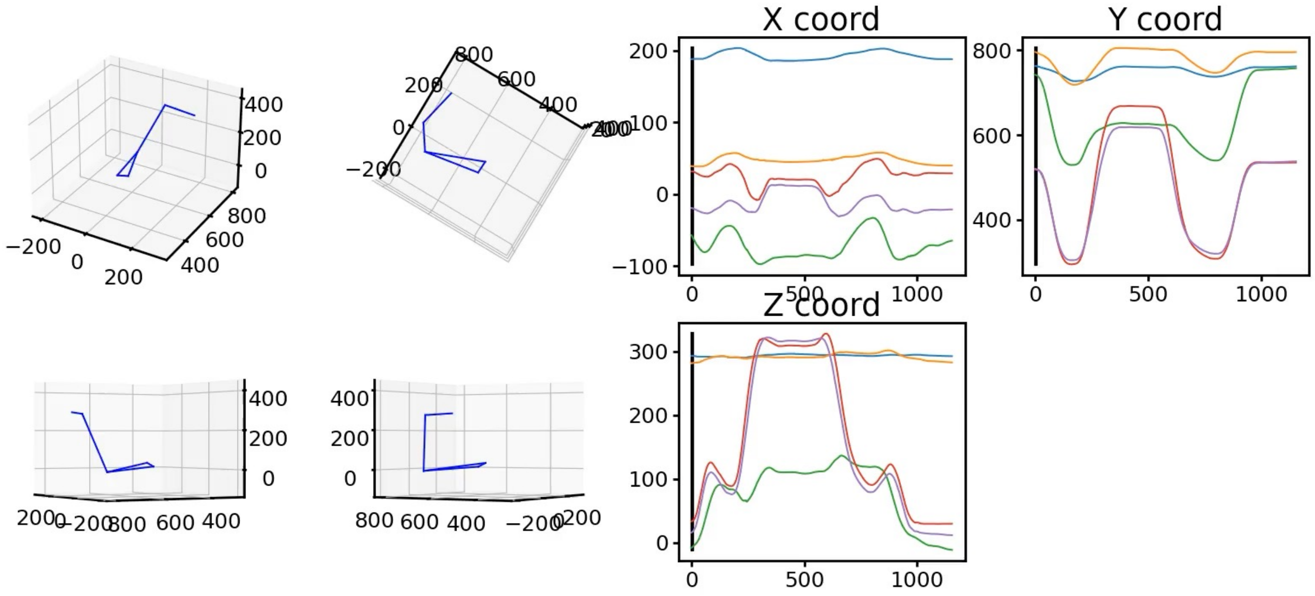
Supplementary: Qualitative Comparison Severe Stroke: FMA 14



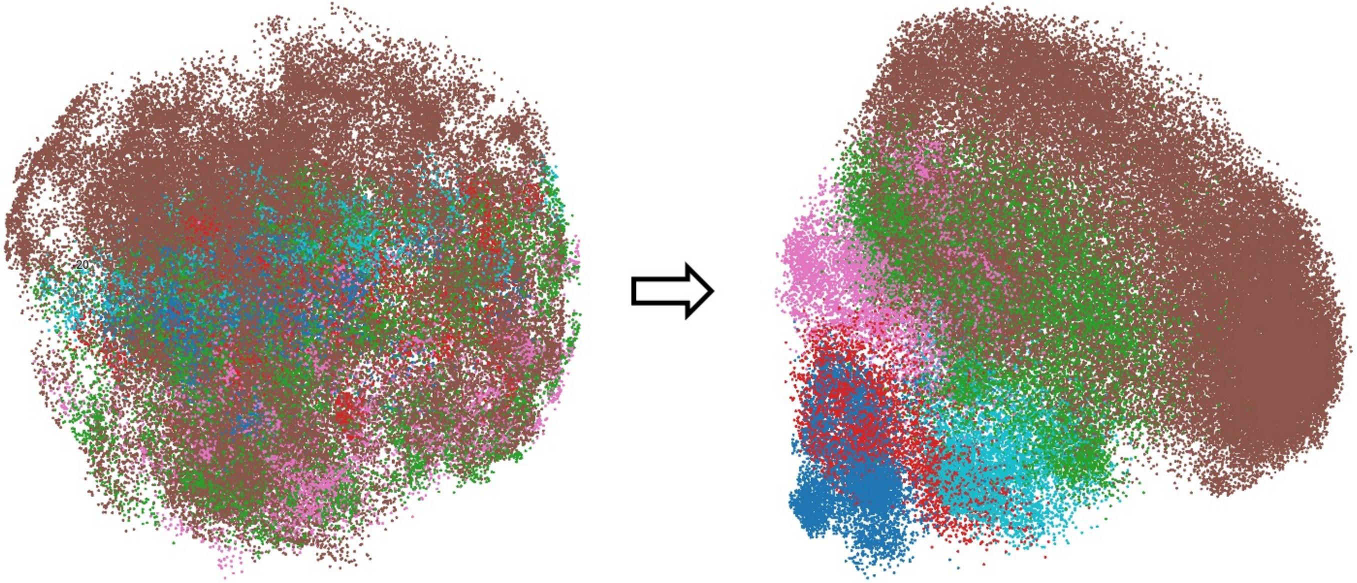
Supplementary: Qualitative Comparison
Moderate Stroke: FMA 17



Supplementary: Qualitative Comparison Healthy Subject: FMA 20



Supplementary: Feature Space After Training (colored by FMA score, PCA @ 27 variance)



EPFL

Stroke Analysis Discussion

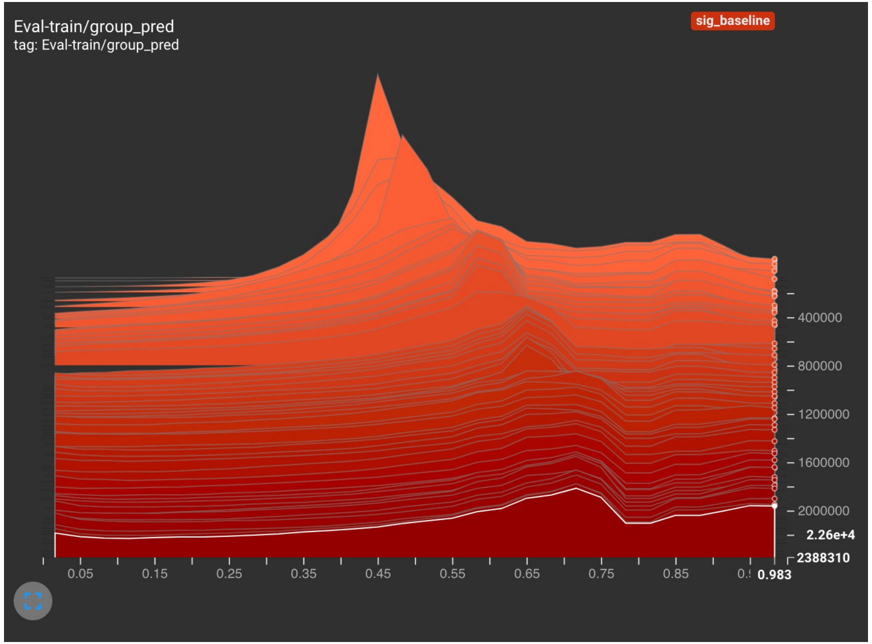
Gavin Zhu

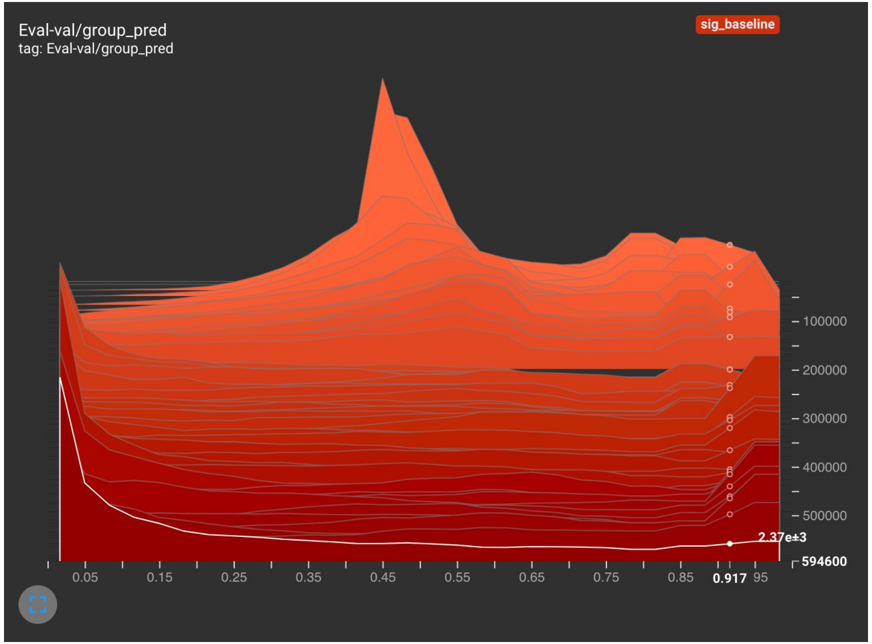
July 25th 2022

Overview

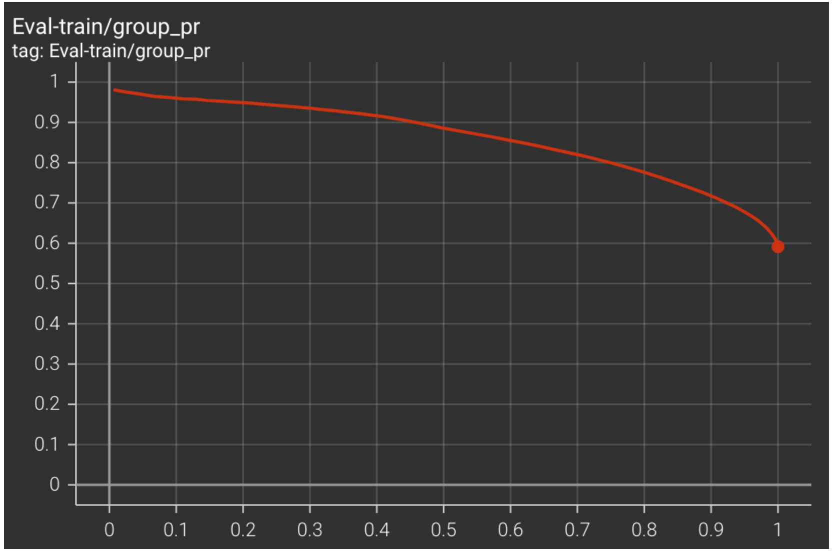
- Exploring local feature descriptor
- Action embedding integration

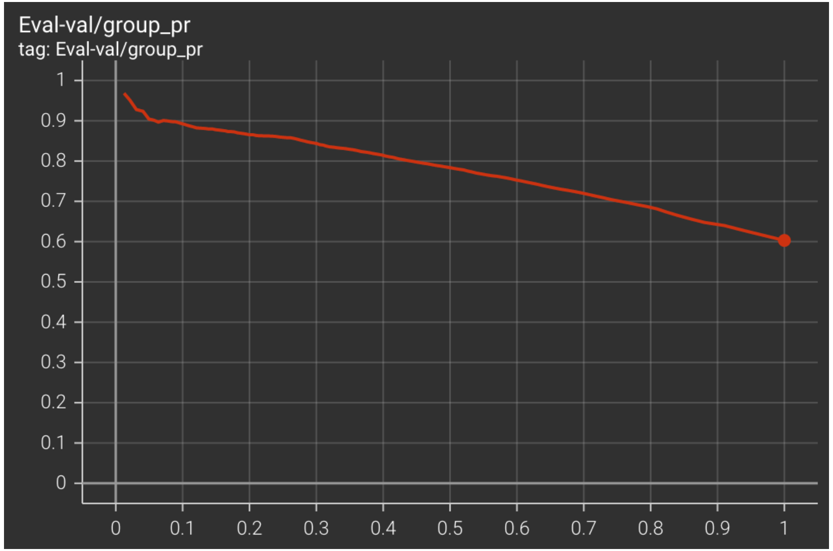
Exploring Local Feature Descriptor



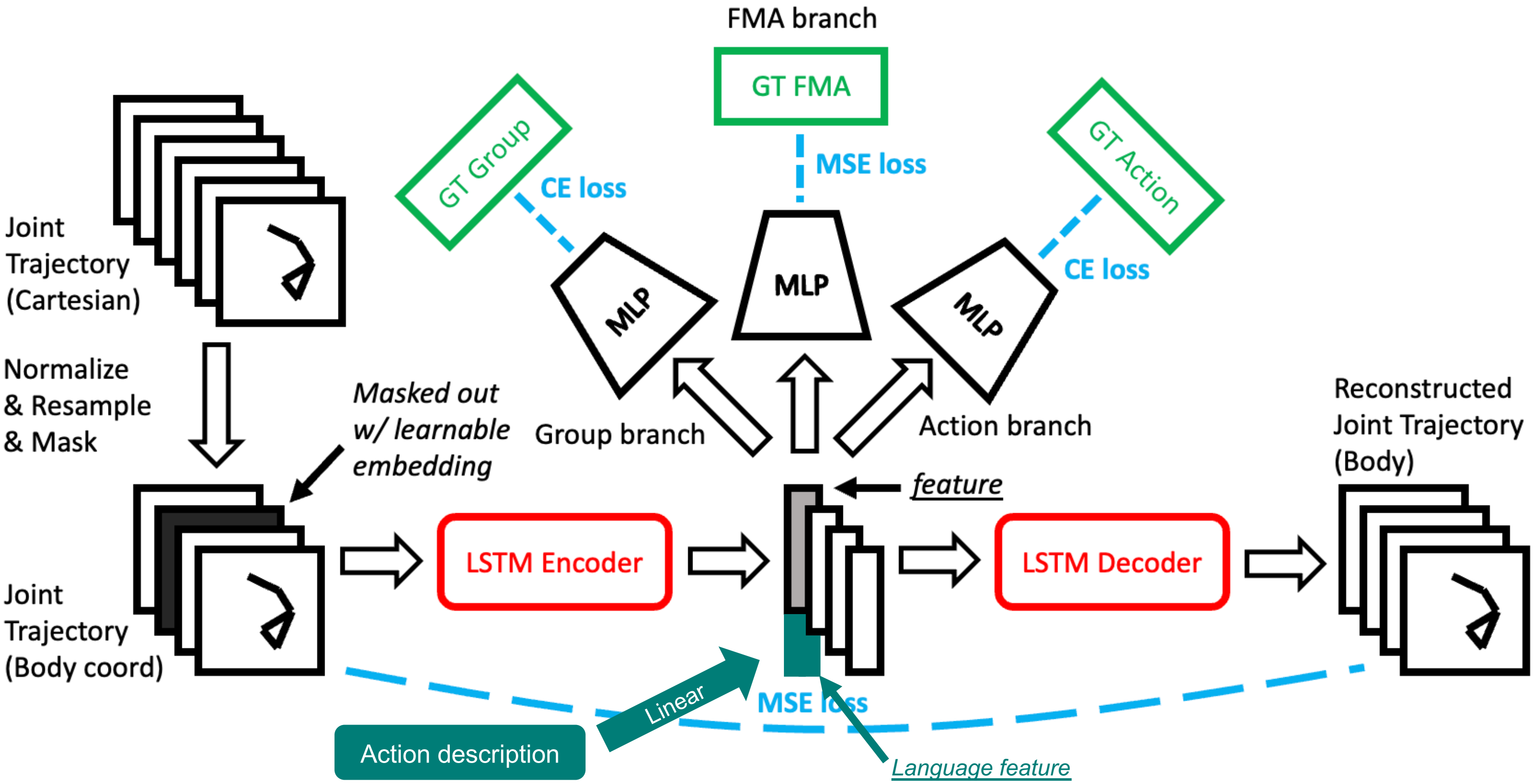


Exploring Local Feature Descriptor

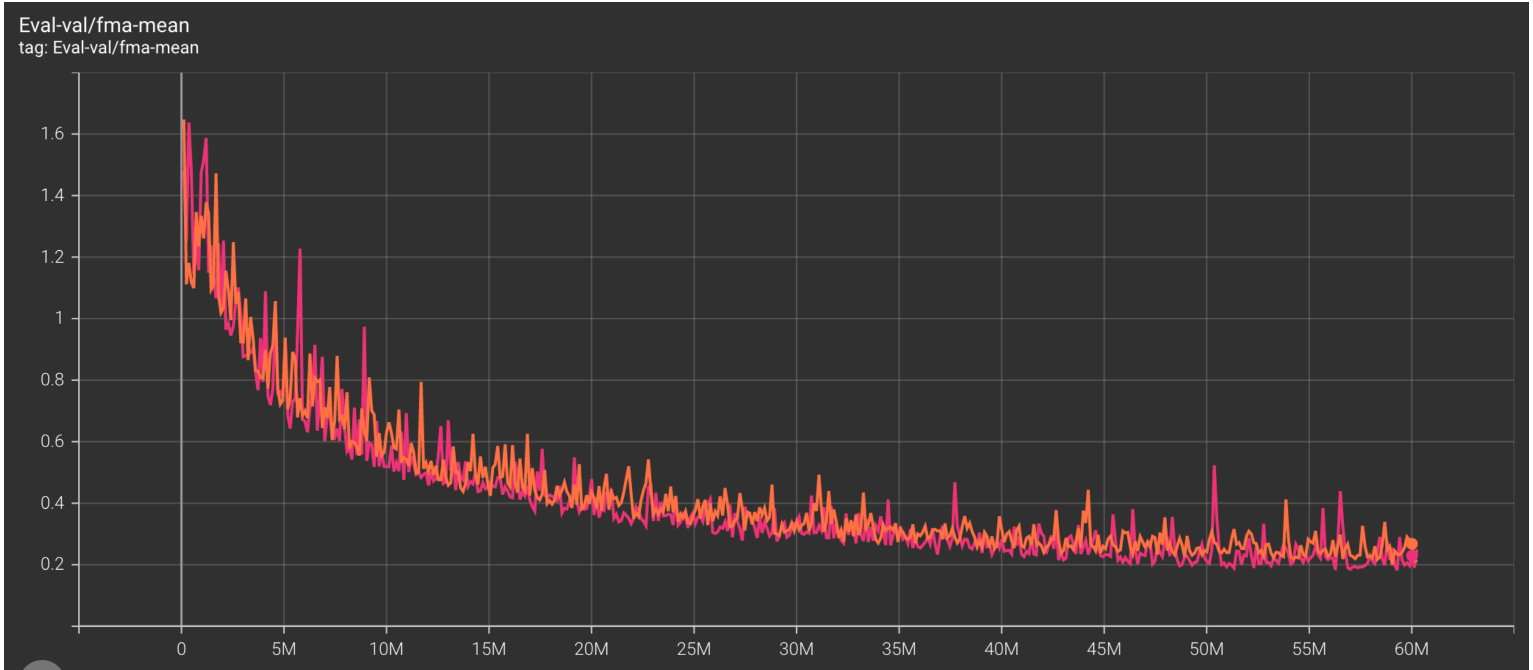




Action embedding integration

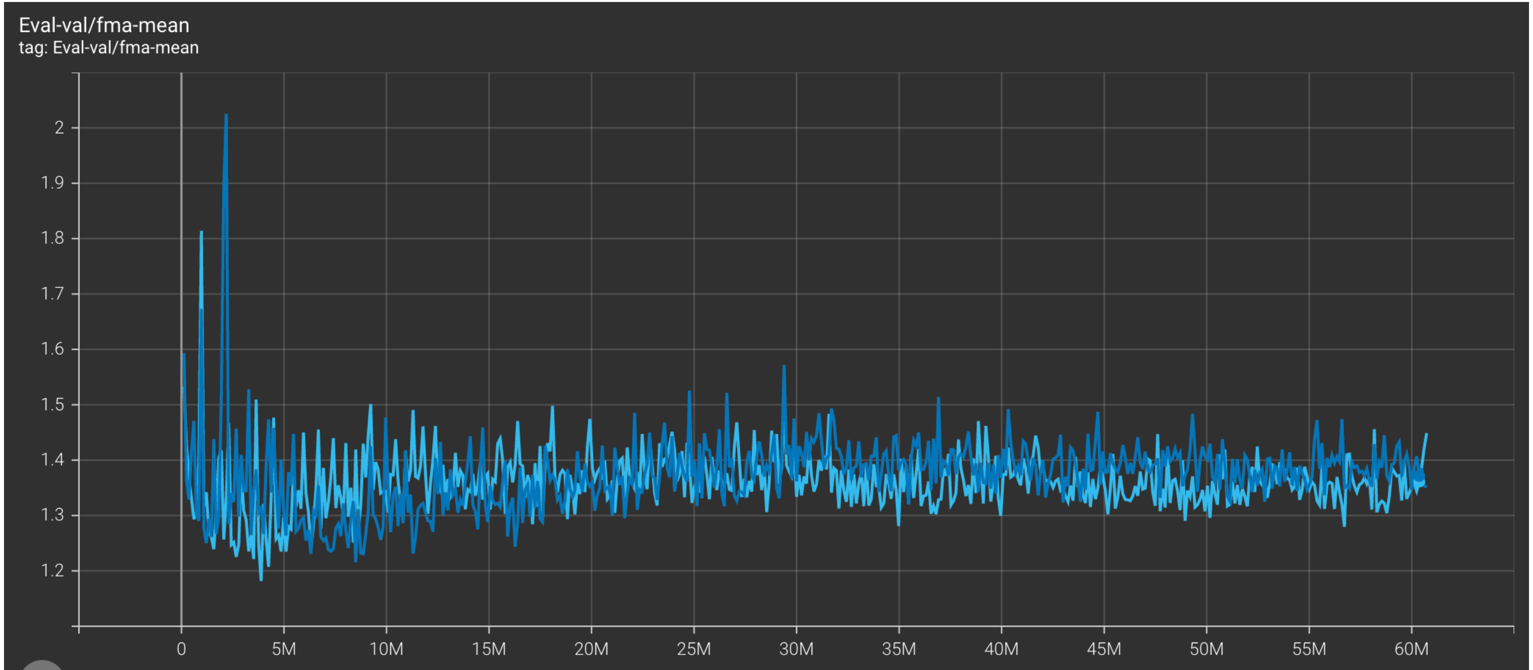


Action embedding integration



Name	Smoothed	Value	Step	Time	Relative
FMA_baseline	0.2679	0.2679	60.01M	Wed Jul 13, 17:05:15	1h 28m 24s
FMA_language	0.2293	0.2293	60.01M	Fri Jul 22, 19:12:38	59m 51s

Action embedding integration



Name	Smoothed	Value	Step	Time	Relative
FMA_LSTM1	1.371	1.371	60.36M	Tue Jul 19, 10:34:22	42m 5s
FMA_language_subject	1.361	1.361	60.36M	Fri Jul 22, 18:37:50	1h 3m 28s

Stroke Subject Dexterity Evaluation Progress Report

Gavin Zhu

Lab Meeting

Aug. 9th 2022

Previously…

Problems & Potential Solutions

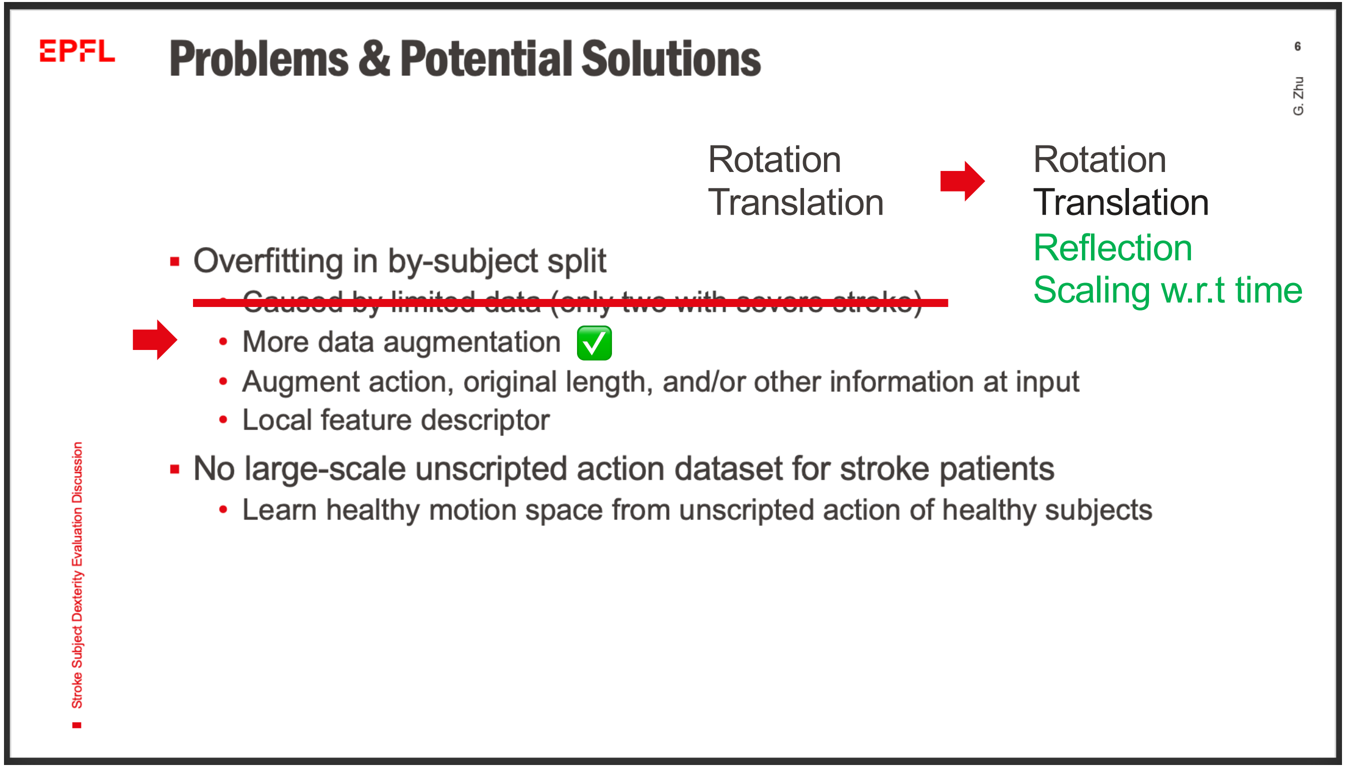
- Overfitting in by-subject split
  - Caused by limited data (only two with severe stroke)
  - More data augmentation
  - Augment action, original length, and/or other information at input
  - Local feature descriptor
- No large-scale unscripted action dataset for stroke patients
  - Learn healthy motion space from unscripted action of healthy subjects

Previously…

Problems & Potential Solutions

- Overfitting in by-subject split
  - ~~Caused by limited data (only two with severe stroke)~~ No new data
  - More data augmentation
  - Augment action, original length, and/or other information at input
  - Local feature descriptor
- No large-scale unscripted action dataset for stroke patients
  - Learn healthy motion space from unscripted action of healthy subjects

Previously…



Problems & Potential Solutions
6
G. Zhu
Rotation
Translation
Rotation
Translation
Reflection
Scaling w.r.t time
Overfitting in by-subject split
~~Caused by limited data (only two with severe stroke)~~
More data augmentation ✅
Augment action, original length, and/or other information at input
Local feature descriptor
No large-scale unscripted action dataset for stroke patients
Learn healthy motion space from unscripted action of healthy subjects
Stroke Subject Dexterity Evaluation Discussion

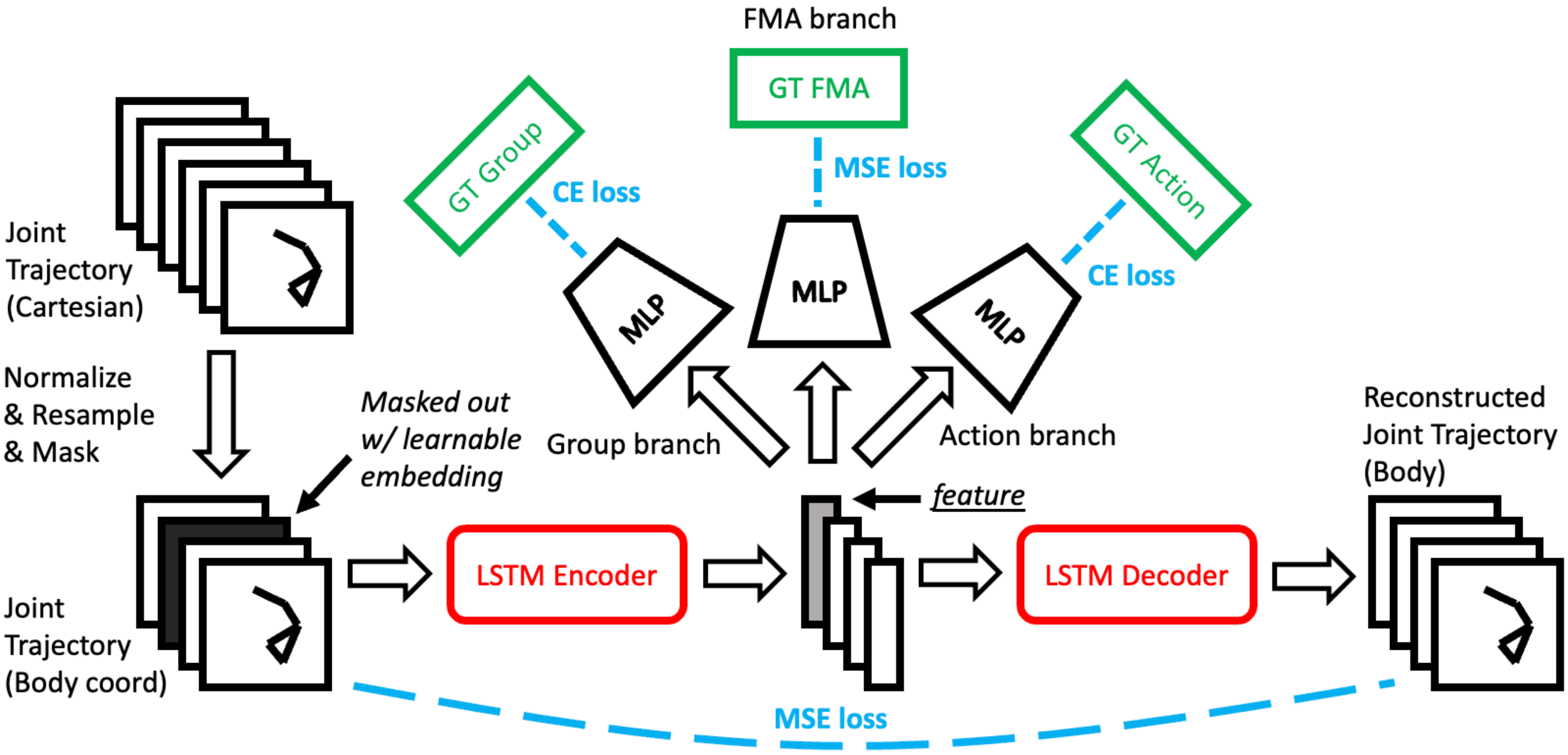
Previously…

G. Zhu

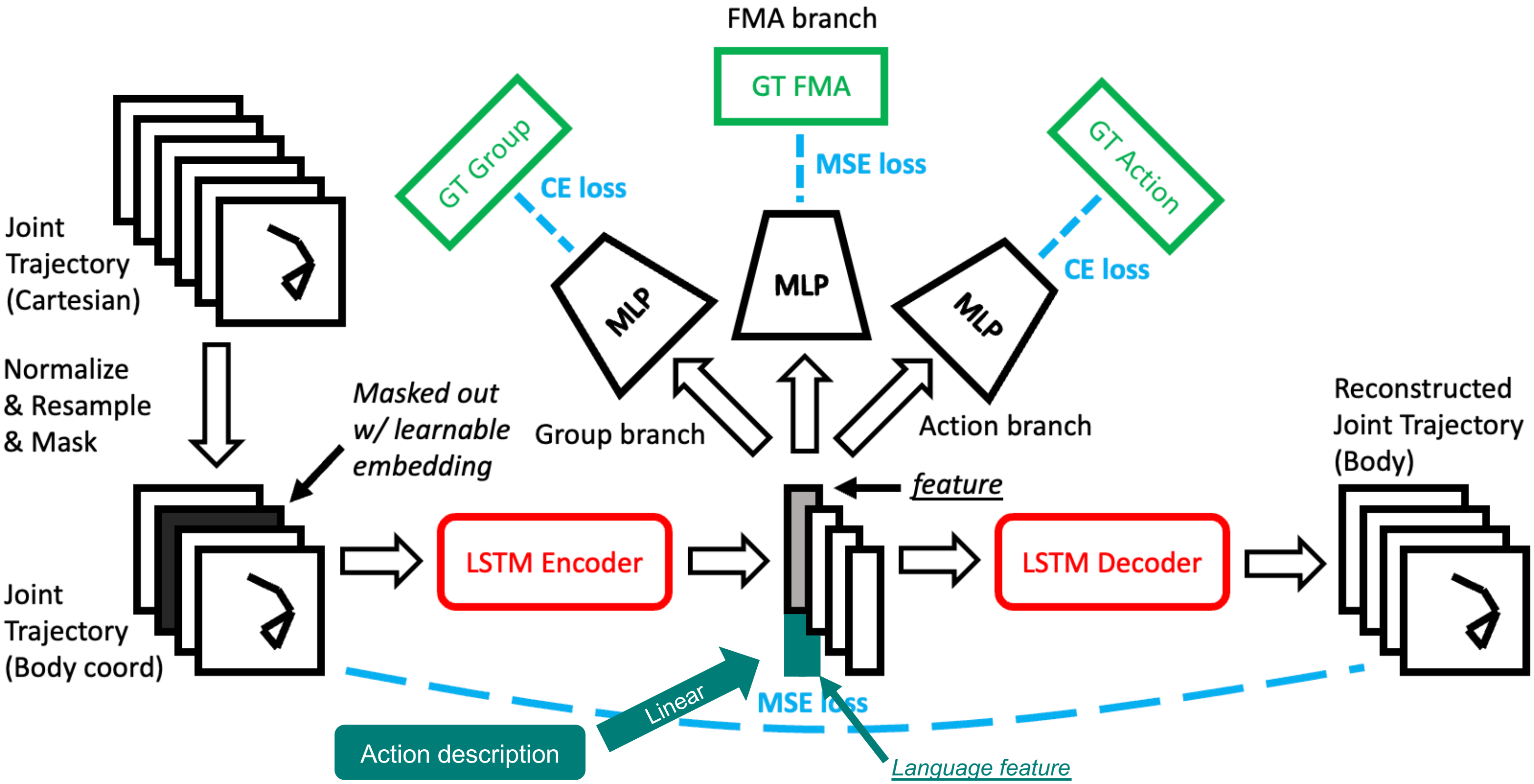
Problems & Potential Solutions

- Overfitting in by-subject split
  - ~~Caused by limited data (only two with severe stroke)~~
  - More data augmentation ✅
  - Augment action, original length, and/or other information at input
  - Local feature descriptor
- No large-scale unscripted action dataset for stroke patients
  - Learn healthy motion space from unscripted action of healthy subjects

Action embedding integration

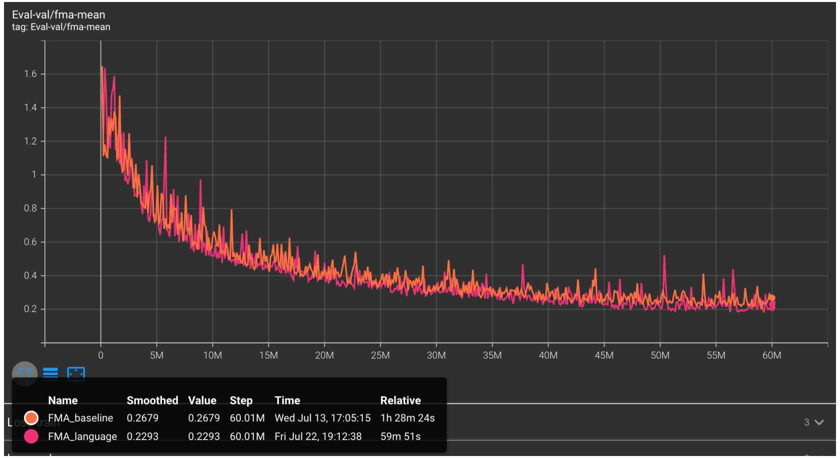


Action embedding integration

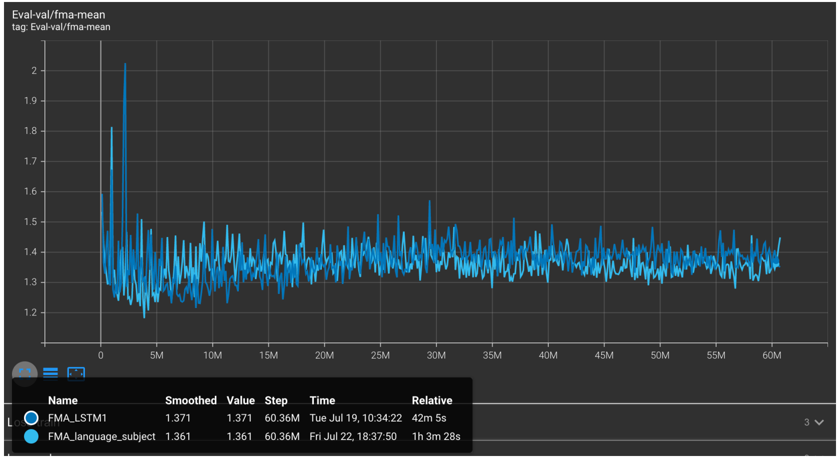


Action embedding integration

Split by trial



Split by subject



Takeaway: integrating action description *does not* improve performance

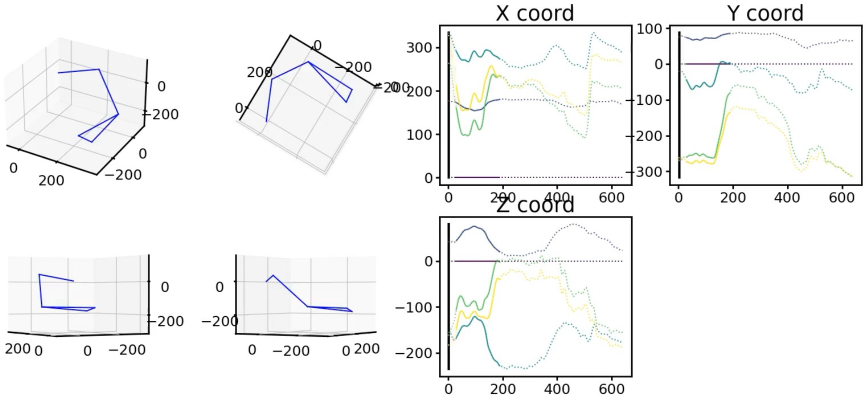
Previously…

Problems & Potential Solutions

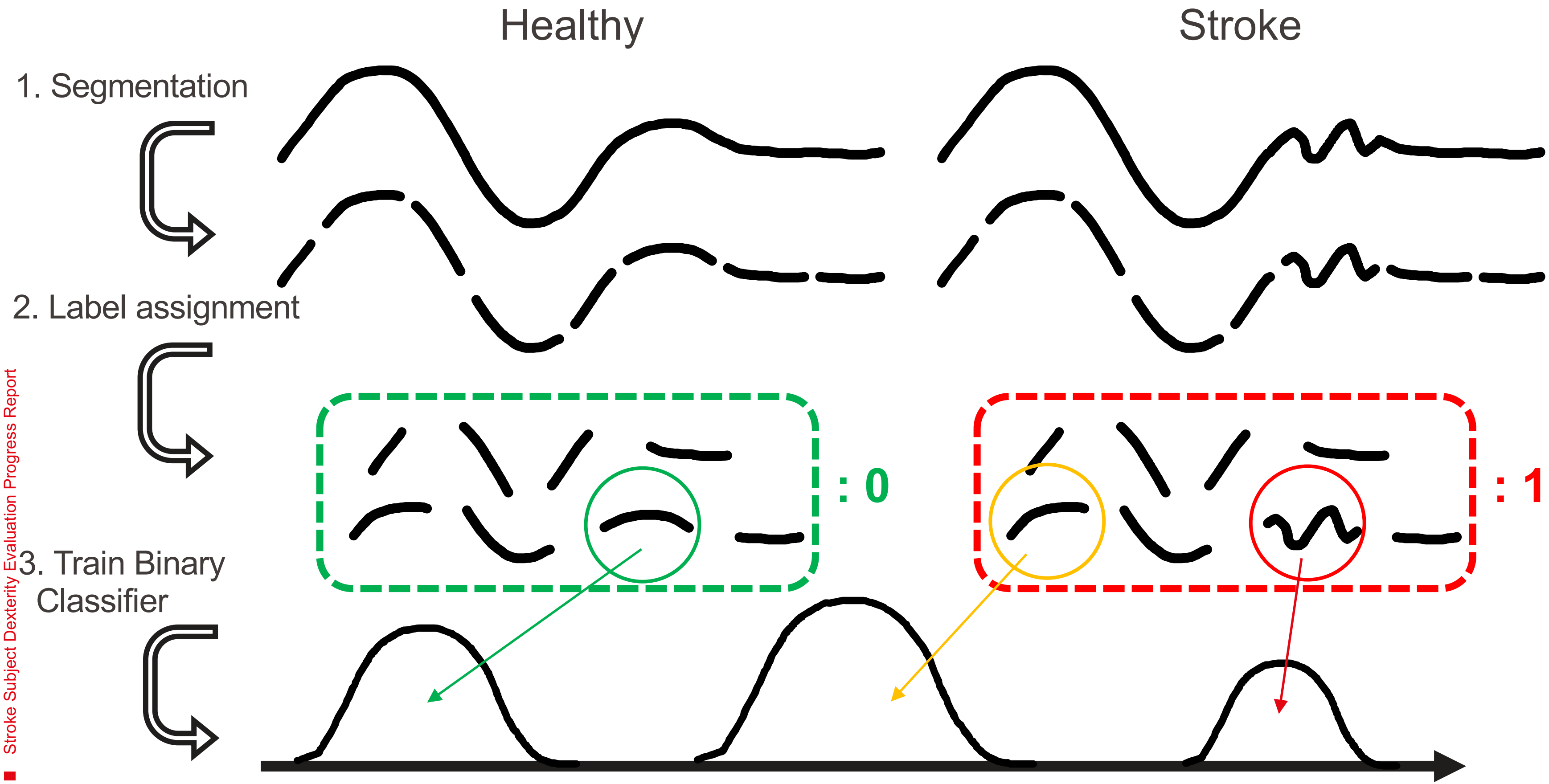
- Overfitting in by-subject split
  - ~~Caused by limited data (only two with severe stroke)~~
  - More data augmentation ✅
  - ~~Augment action, original length, and/or other information at input~~
  - Local feature descriptor
- No large-scale unscripted action dataset for stroke patients
  - Learn healthy motion space from unscripted action of healthy subjects

Local Feature Descriptor

- Problem Definition
  - Given a (potentially stroke/healthy) trajectory, try to locate it's *stroke-related segments* (with *weak supervision*)
- Motivation
  - Not the entire trajectory of stroke subject is stroke related
  - Fixed time-scale
  - Action agnostic



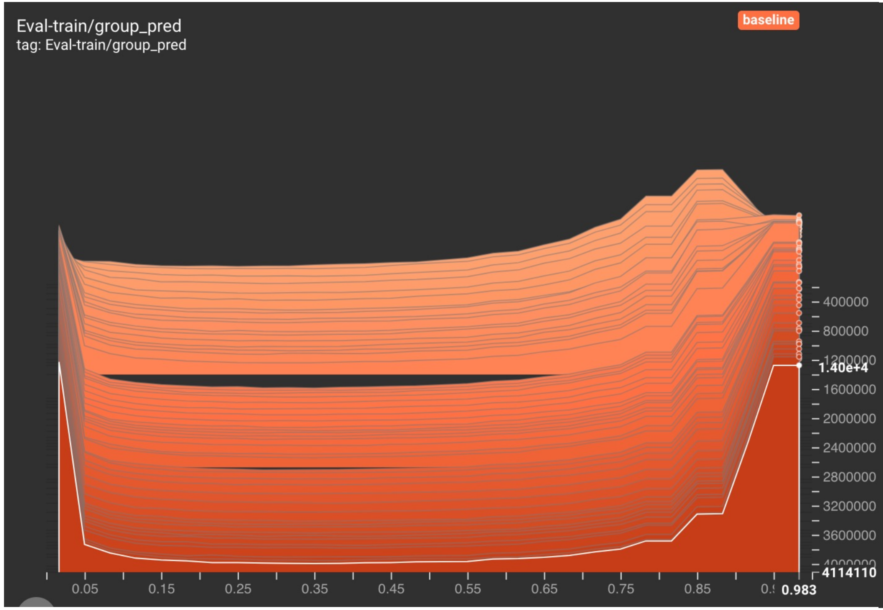
Naïve Training Schema (in Theory)

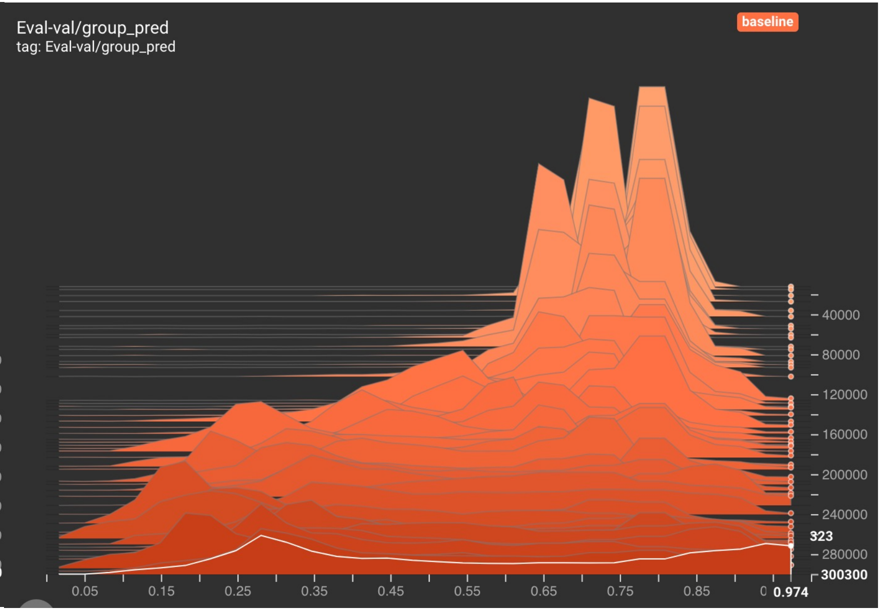


Naïve Training Schema (in Practice)

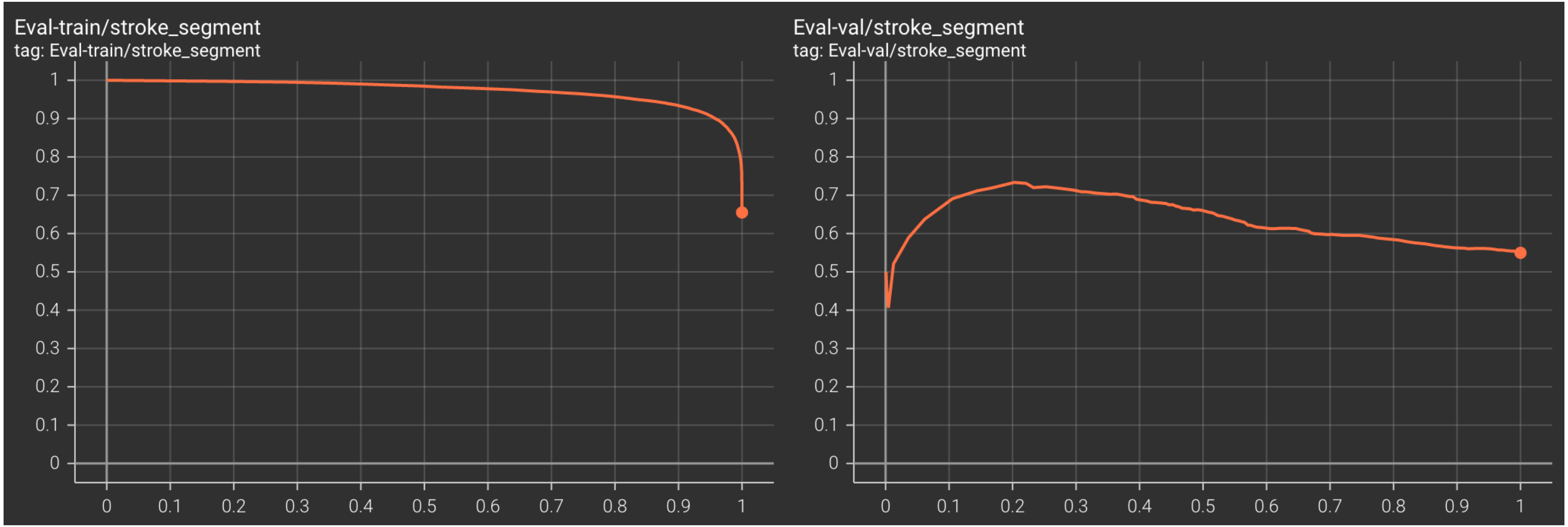
G. Zhu

Naïve Training Schema (in Practice)



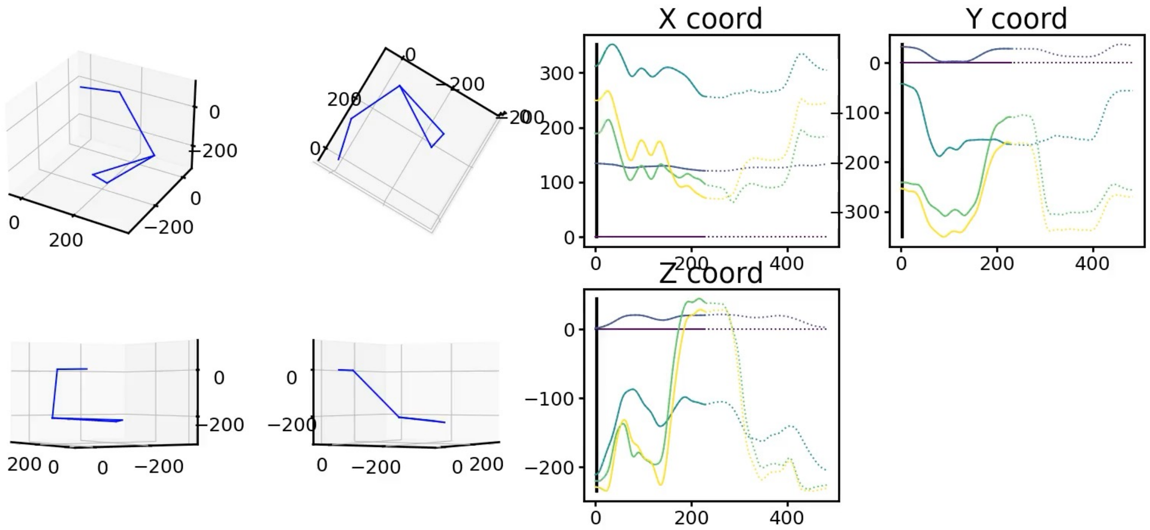


Naïve Training Schema (in Practice)

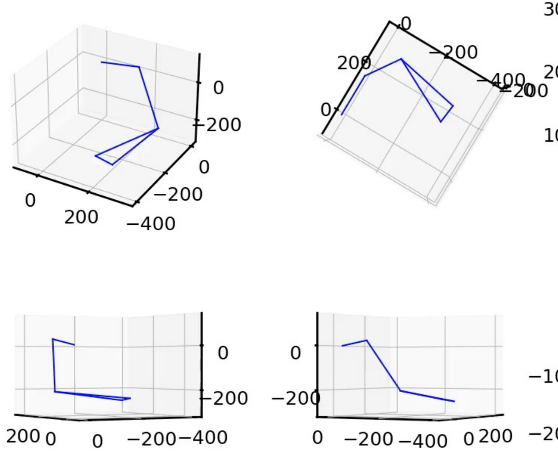


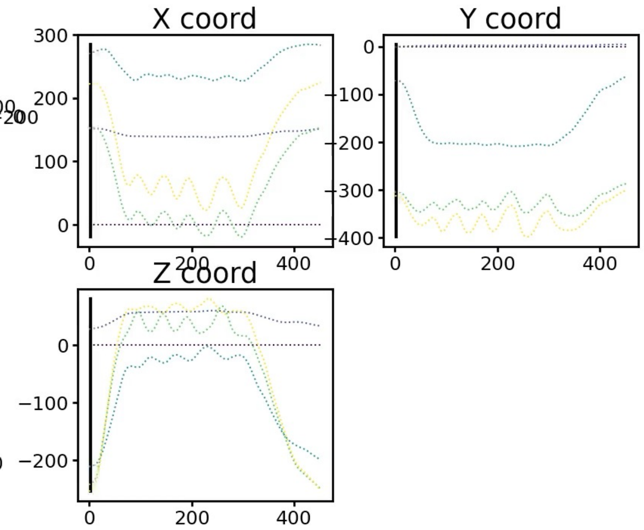
Seemingly overfitting but all kinds of hyper-parameter tuning does not help

Naïve Training Schema Case Study (Good Case: Reasonable Localization)

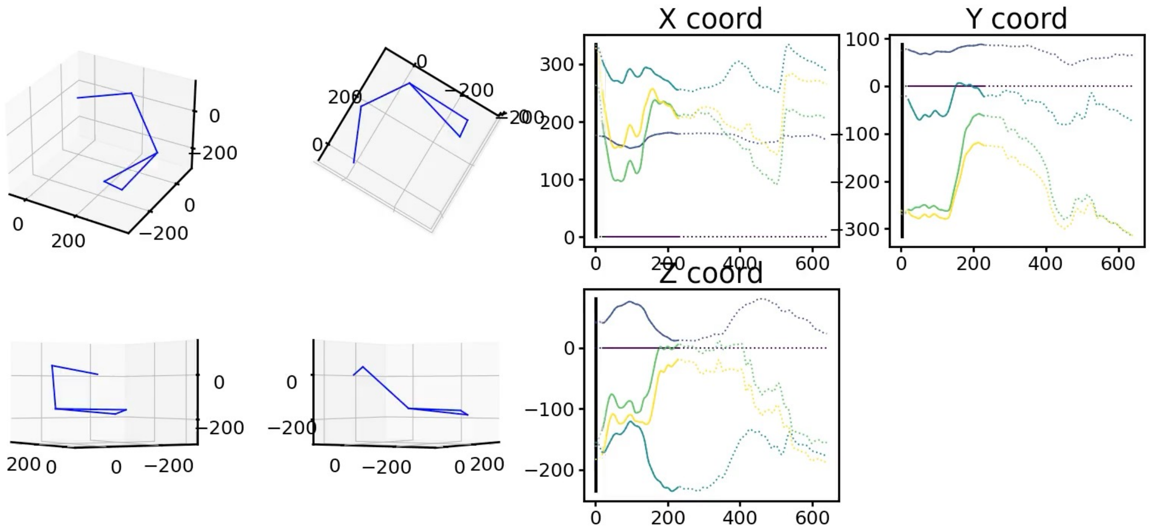


Naïve Training Schema Case Study (Good Case: Robust to Repeated Actions)

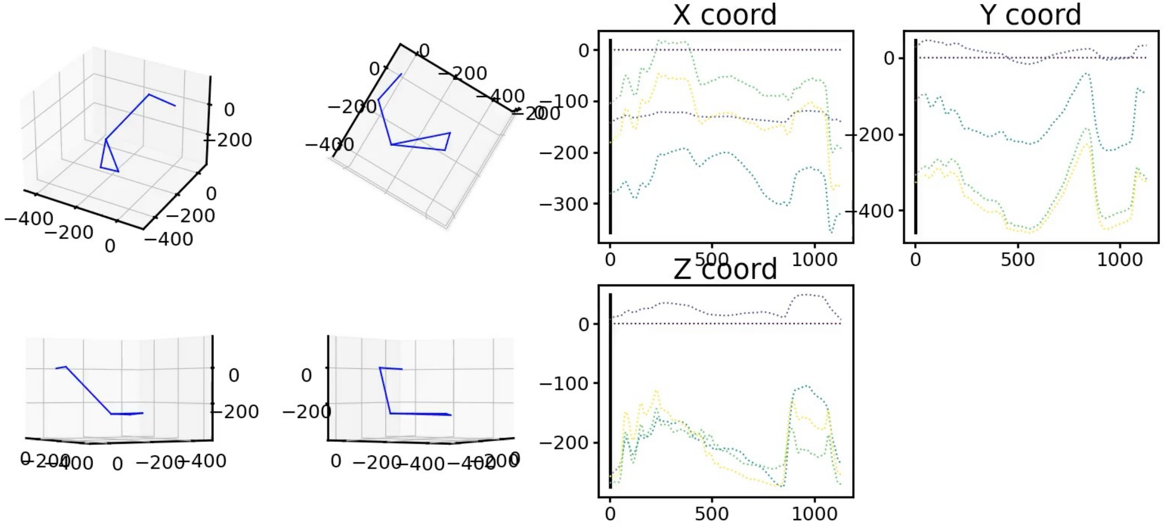




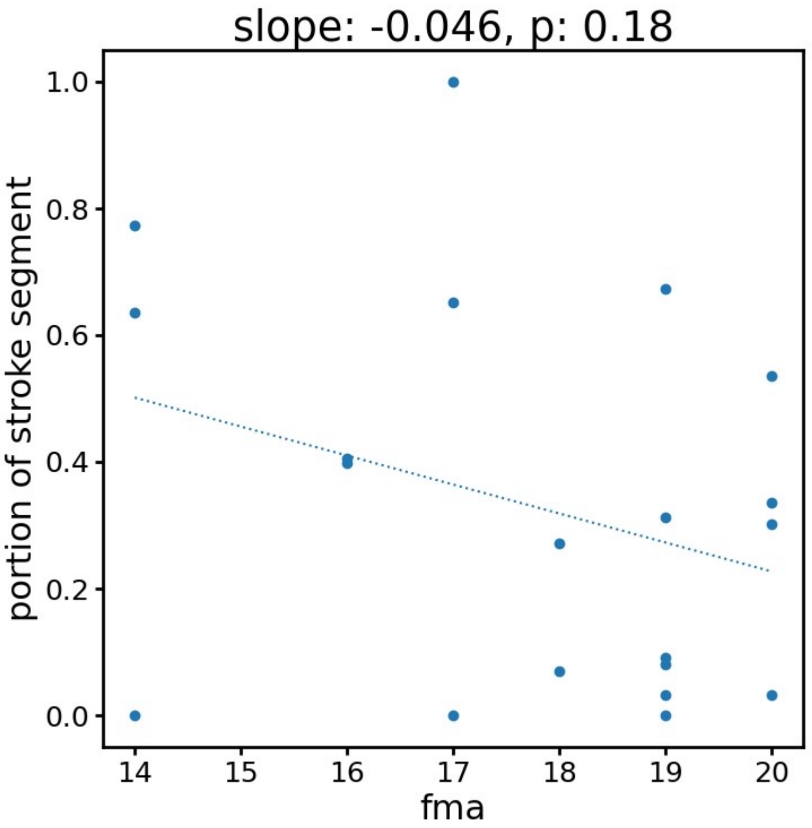
Naïve Training Schema Case Study (False Negative)



Naïve Training Schema Case Study (False Negative II)



Naïve Training Schema Best Correlation



Naïve Training Schema: Part II (with Handcrafted Features)

- Statistical feature of time series data
  - extreme value, variance, # of peaks, max change, skewness, etc.
- Used in anomaly detection and all kinds of other competitions
- (Comparatively) explainable & robust

Naïve Training Schema: Part II (Bootstrapping Handcrafted Features)

Algorithm 1: Boostrapping Relevant Features

Input: Potentially useful features $\mathcal{F}$ ($|\mathcal{F}| \sim 12k$)
Input: A set of segments $\mathcal{S}$ ($|\mathcal{S}| \sim 400k$) and their FMA $f : \mathcal{S} \to [0, 1]$
Output: A subset of $\mathcal{F}$ indicating relevant features

```
cnt ← 0
while cnt < 5 do
    relevant ← ∅
    visited ← ∅
    foreach batch 𝓑 ⊆ 𝓢 do
        F ← sample(𝓕)
        selected ← KendalTest(F, (𝓑, f(𝓑)))
        relevant ← relevant ∪ selected
        visited ← visited ∪F
    end
    if relevant ≠ visited then
        𝓕 ← 𝓕 \ visited ∪ relevant
        cnt ← 0
    else
        cnt ← cnt + 1
    end
end
return 𝓕
```

Previously…

EPFL

Problems & Potential Solutions

- Overfitting in by-subject split
  - ~~Caused by limited data (only two with severe stroke)~~
  - More data augmentation ✅
  - ~~Augment action, original length, and/or other information at input~~
  - Local feature descriptor
- No large-scale unscripted action dataset for stroke patients
  - Learn healthy motion space from unscripted action of healthy subjects

Weakly Supervised Stroke Evaluation

Gavin Zhu

Discussion

Aug. 22th 2022

Outline

- Current Work
  - Segment classification / stroke related segment localization
  - Regressing FMA from activities of daily living
- Potential Usage
  - Pinpoint impairment
  - Select representative evaluation action
- Future directions

Preview - Dataset

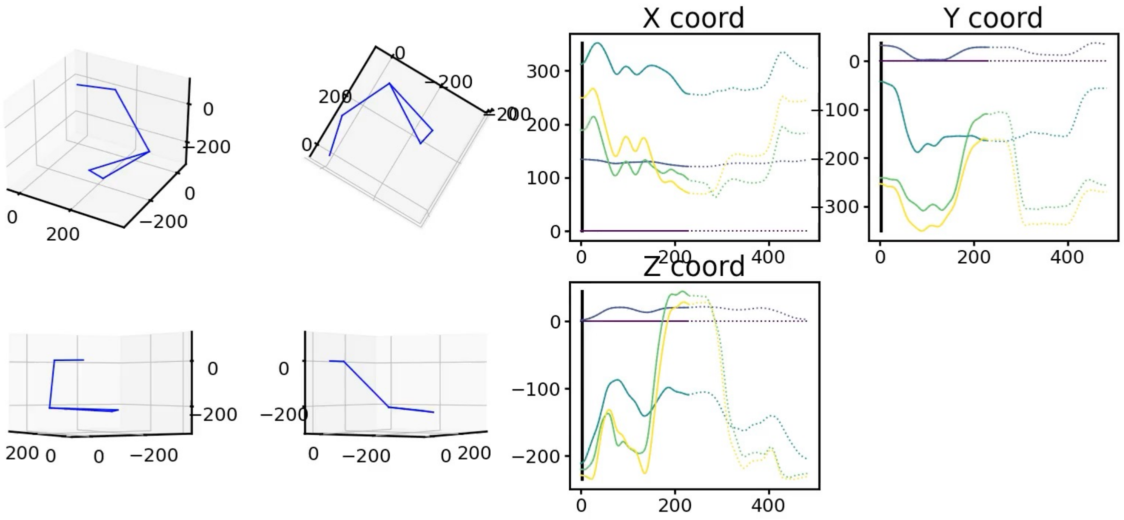
- 40 subjects (20 stroke)
  - FMA-UE (out of 20) count: 3@14, 2@16, 4@17, 2@18, 6@19, 4@20
- 30 activities of daily living in three categories
  - Simple: lifting hand from the table
  - Transitive: Reach and grasp a glass, drink for 3 seconds, and replace it in the initial position
  - Tool related: Reach and grasp a bottle, pour water into a glass, and replace the bottle in the initial position
- Each subject perform each trial 3 times

Segment Classification: Overview

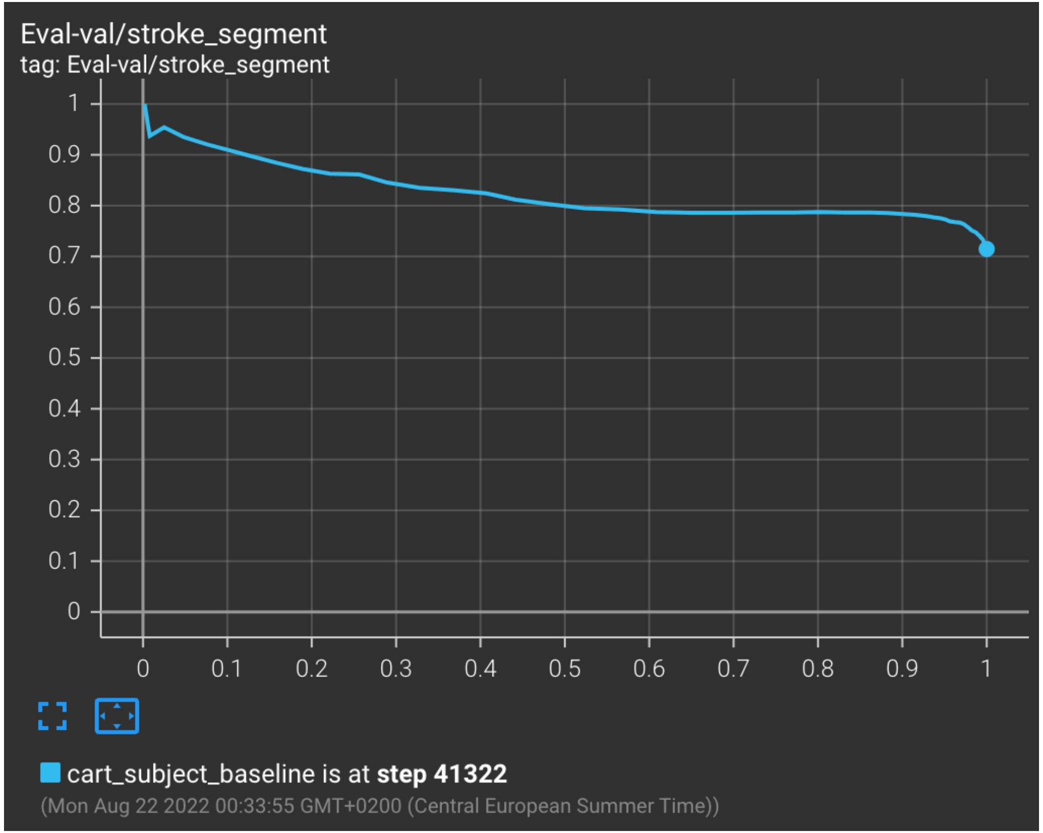
- Motivation
  - Many stroke behavior is “local” and *action agnostic*
  - Healthy subjects normally *do not* demonstrate stroke related segment
- Approach
  - Break a motion trajectory into fixed-length *segments*, design an estimator of the probability that a given segment is *only demonstrated by stroke subjects*
- Method
  - Classification formulation
    - CNN based feature extractor + MLP classifier
    - Handcrafted time series feature extractor + Boosting classifier
  - Anomaly detection formulation (TODO)
    - Train a model on healthy subject only, and inference on stroke for discrepancy

Segment Classification: Jerk Demo

“Reach and grasp an apple, mimic biting, and replace it in the initial position”

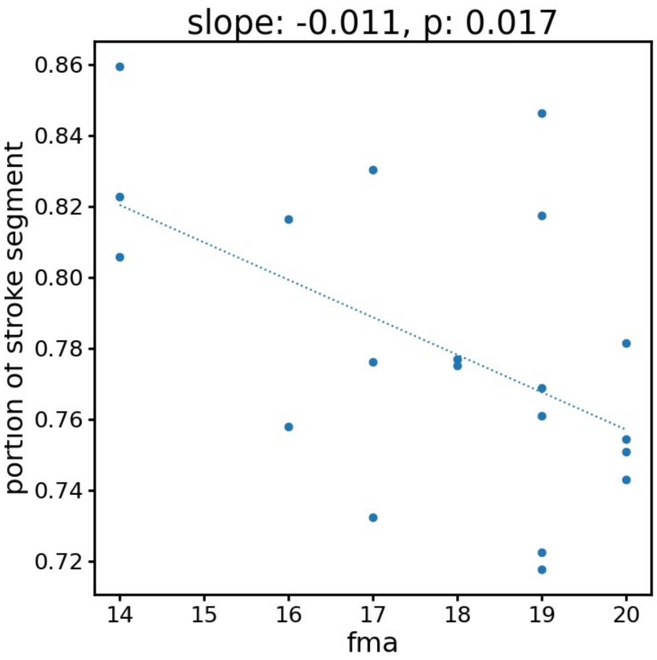


Segment Classification: PR Curve

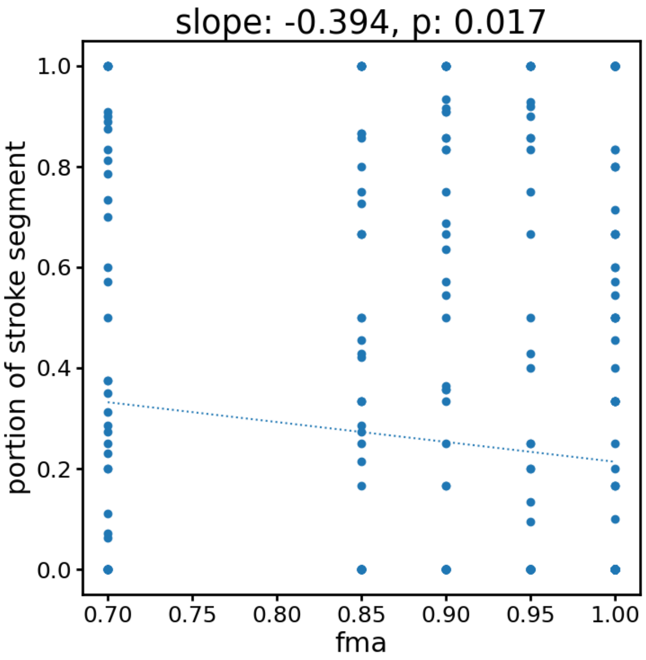


Segment Classification: Stroke Relevance

- CNN based per subject

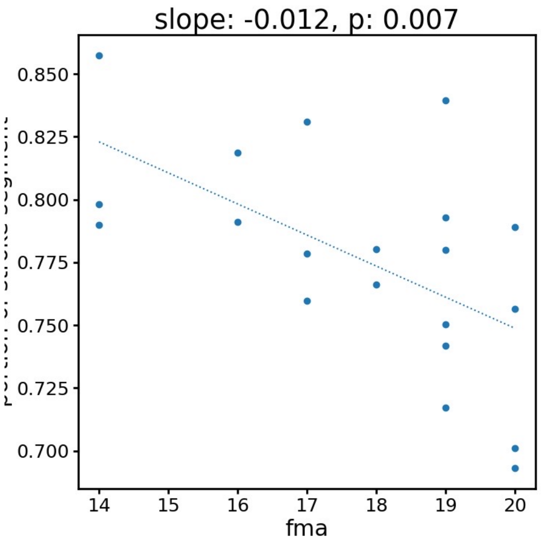


- Handcrafted feature per trial

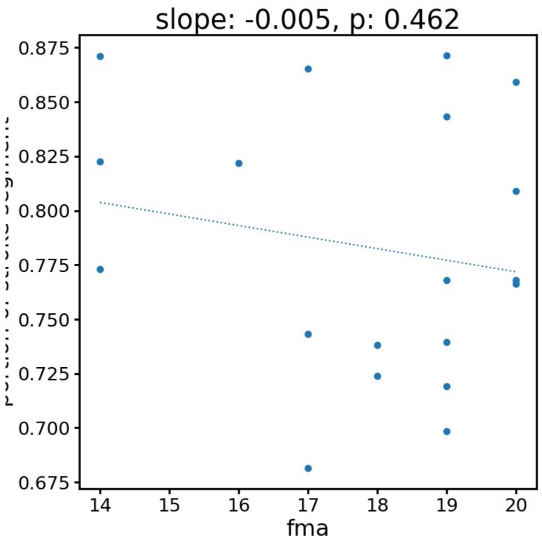


Potential Usage: Select Testing Actions

- Reach and grasp a phone receiver, carry it to own ear and hold for 3 seconds, and replace it in the initial position



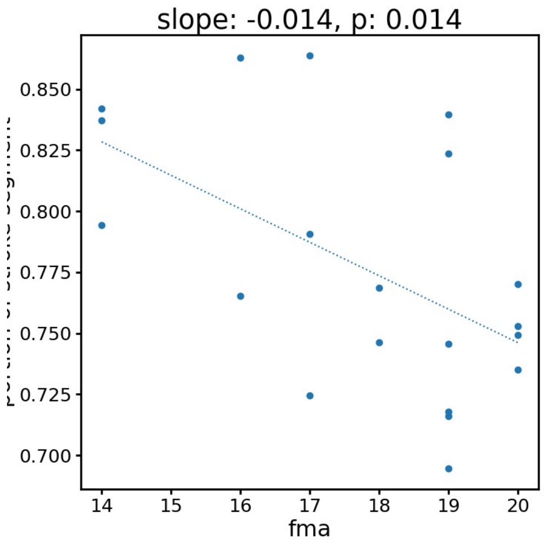
- Reach and grasp a book, put in on the table, and open it (from right side to left side)



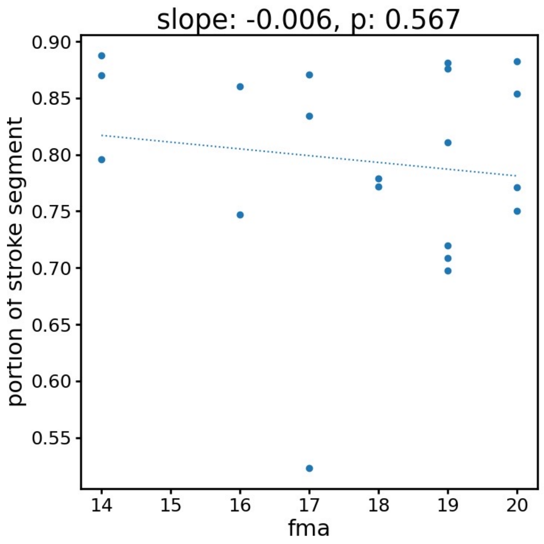
EPFL

Potential Usage: Select Testing Actions

- Reach and grasp a toothbrush, brush teeth, and put it inside a holder (on the right side of the table)

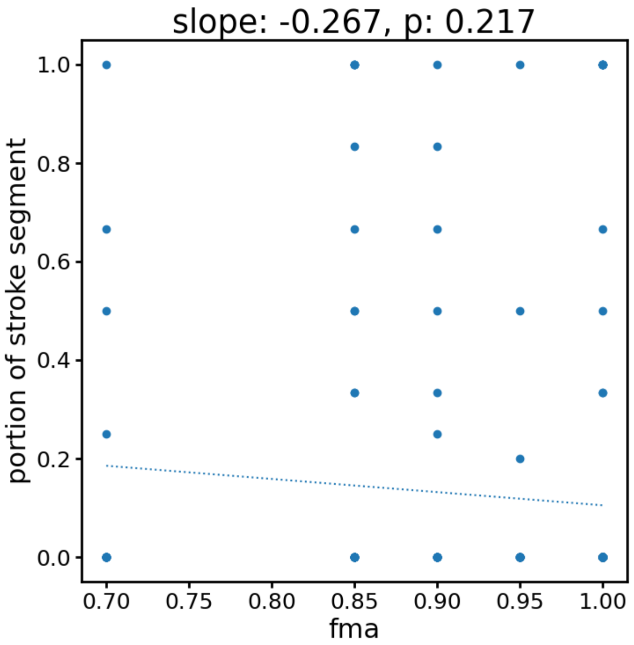


- Reach and grasp a laptop, then open it with one hand

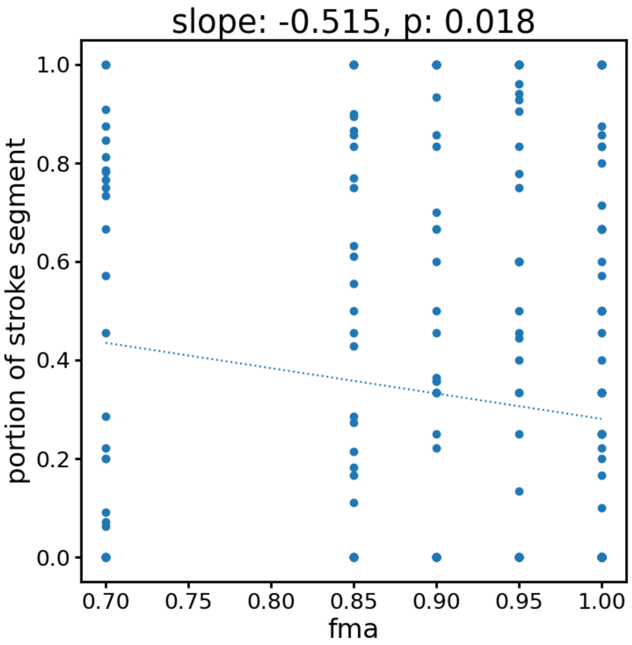


Potential Usage: Select Testing Actions

- Simple Actions



- Non-simple Actions

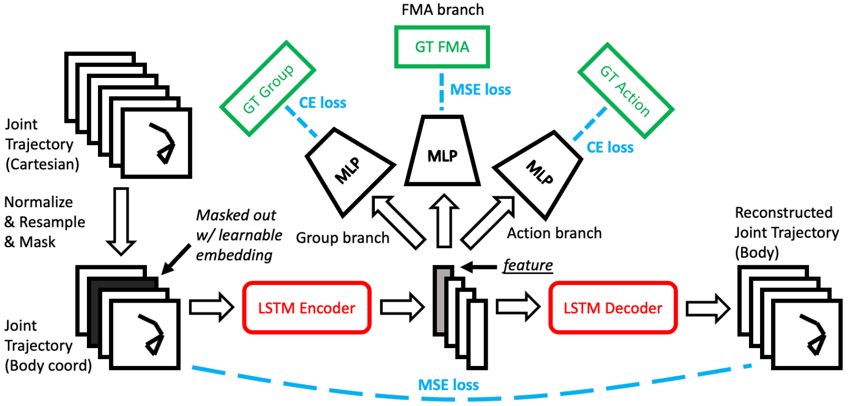


Future Directions

- Anomaly detection model for segment classification
- Segment / trial level annotation for better training / evaluation
- (Potentially) transfer the same method onto animal stroke model

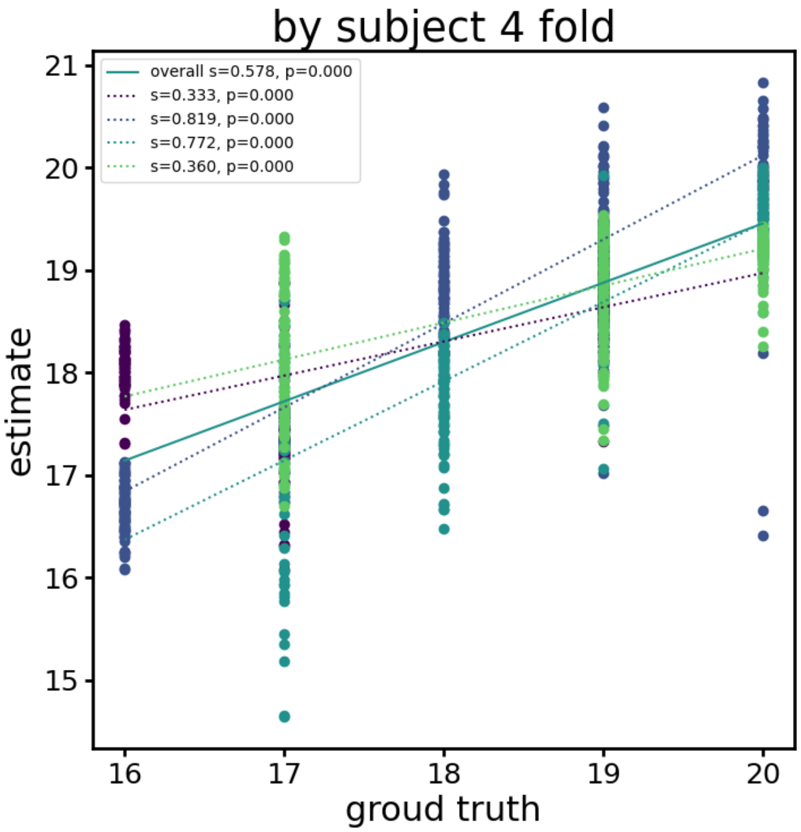
Direct FMA Estimator: Overview

- Motivation
  - Activities of daily living are more related to *functional* evaluation and provide more *home-based flexibility*
- Approach
  - Resample all actions into the same length, and perform regression
- Method



Direct FMA Estimator: Results

- Good by-trial train-test split
  - Mean Absolute Error: 0.23
  - Round to correct: 88%
- Poor by-subject train-test split
  - Mean Absolute Error: 0.73
  - Round to correct: 50%
  - Performance may be impacted by noisy label



Publication Formulation

- Central idea: local feature in stroke evaluation
- Useful
  - Pathological standpoint
  - Usage in stroke segment localization
  - Evaluation design (action ranking)
  - Direct FMA estimator (vs. global feature as baseline)
- Feasible
  - Baseline models that are significant
  - (Potentially) a dataset suitable for the task