



# Towards Unified Surgical Skill Assessment

Daochang Liu, Qiyue Li, Tingting Jiang, Yizhou Wang, Rulin Miao, Fei Shan, Ziyu Li

NELVT PKU, Peking University Cancer Hospital, Deepwise AI Lab, CFCS PKU, AIT PKU

{daochang, liqiyue, ttjiang}@pku.edu.cn



## Why automatic skill assessment

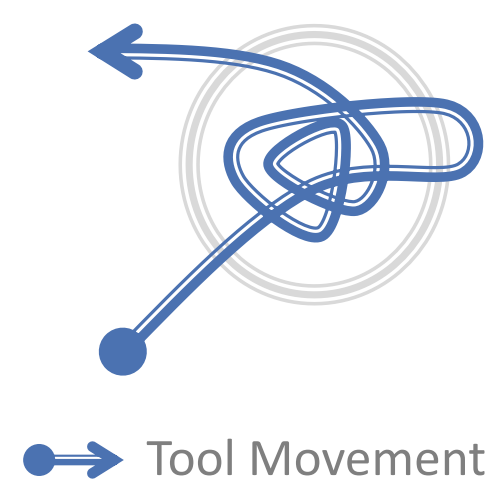
Crucial for patient safety  
 Laborious manual assessment  
 Enormous videos from ORs

## Applications

Efficient surgical training  
 Surgical quality control  
 Automatic surgical robot

## Tool Usage

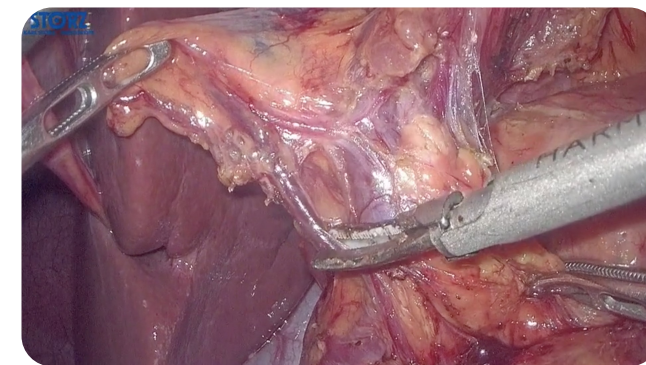
concentrated, smooth  
 vs.  
 dispersed, jerky



## Field Clearness

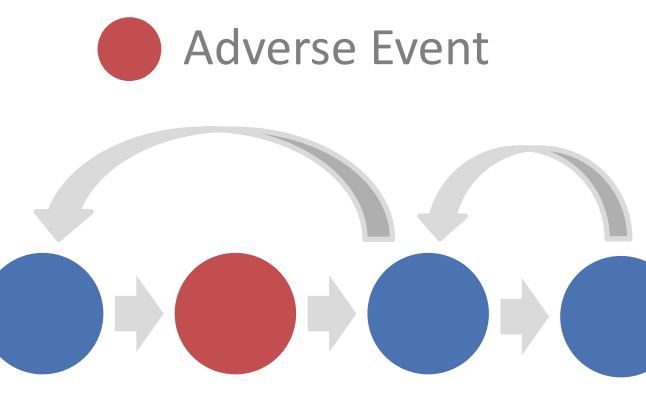
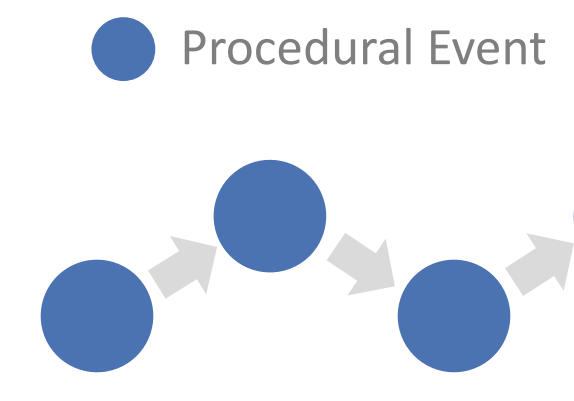
e.g., Skill Proxy

high visibility  
 vs.  
 limited visibility



## Event Pattern

linear, normal  
 vs.  
 loopy, adverse



**High Skill**

**Low Skill**

## Surgical skills are complex and have many aspects

Surgical tool usage -> Instrument proficiency and motion efficiency

Surgical skill proxy -> Indirect indicator correlated to surgical skills

Surgical event pattern -> Knowledge about a particular procedure

## Our Clinical Data

## Detailed Event Annotation (41 Classes)

20 gastrectomy videos

13 coarse-grained procedural events

Captured by laparoscopy

13 fine-grained procedural events

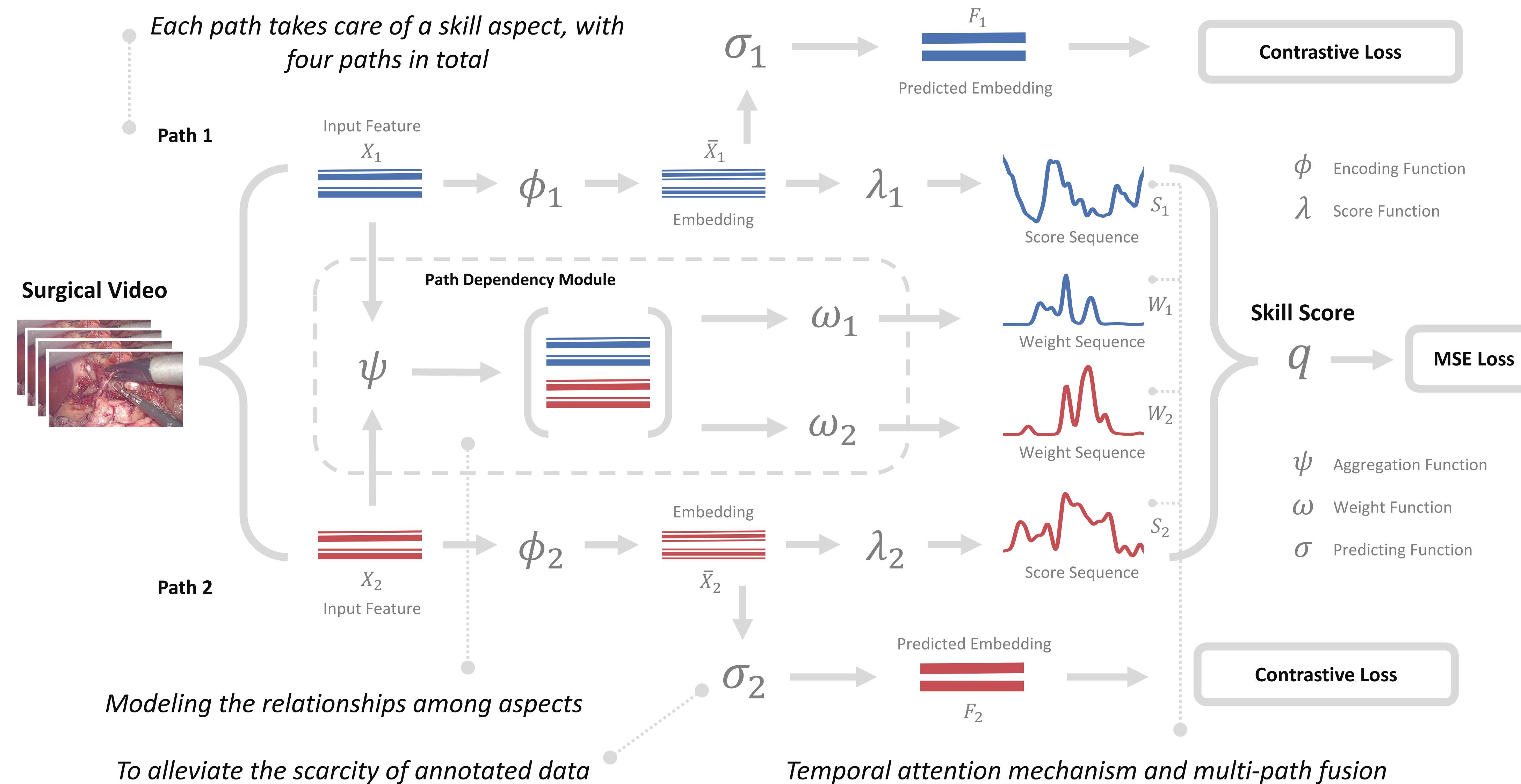
199 minutes per video

8 adverse events and 2 repair events

7 OSATS skill metrics

5 video recording events

## A unified multi-path framework



## Path Input Features

	Simulated Data	Clinical Data
<b>Visual</b>	Image-Net pretrained ResNet features	
<b>Tool</b>	Robotic kinematics	Histograms of tool masks
<b>Proxy</b>	Task completion time	Field clearness [Liu et al.]
<b>Event</b>	Gesture predictions from MS-TCN	Event predictions from MS-TCN

## Instantiation

Functions	
$\phi$	TCN
$\lambda$	MLP
$\psi$	Concat
$\omega$	MLP
$\sigma$	MLP

## Contrastive Loss

Positives: Embedding at the current frame

$$\mathcal{L}_{con} = - \sum_m \sum_{i=1}^L \log \frac{\exp(F_{m,i} \cdot \bar{X}_{m,i})}{\sum_{j \in \mathcal{N}_i} \exp(F_{m,i} \cdot \bar{X}_{m,j})}$$

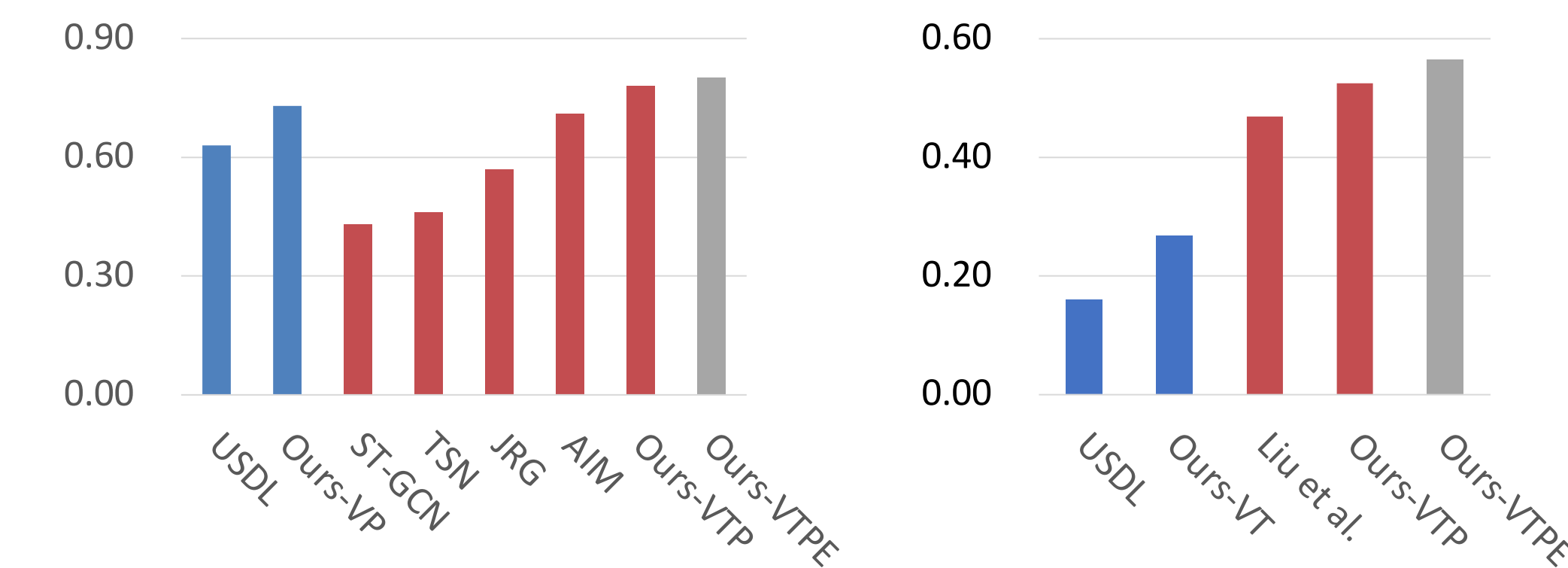
Predict the current using the past

$$F_{m,i} = \sigma_m(\bar{X}_{m,i-1})$$

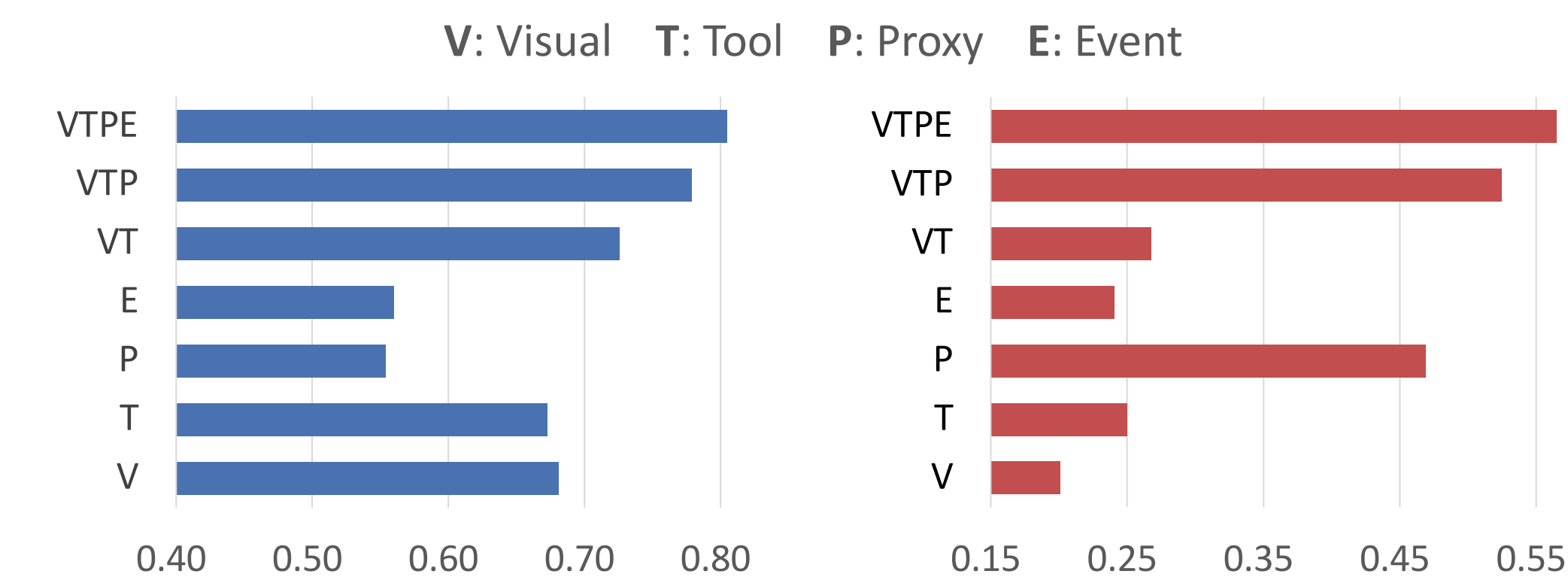
Negatives: Embedding at the others in a neighborhood

## Comparisons to State-of-the-Art

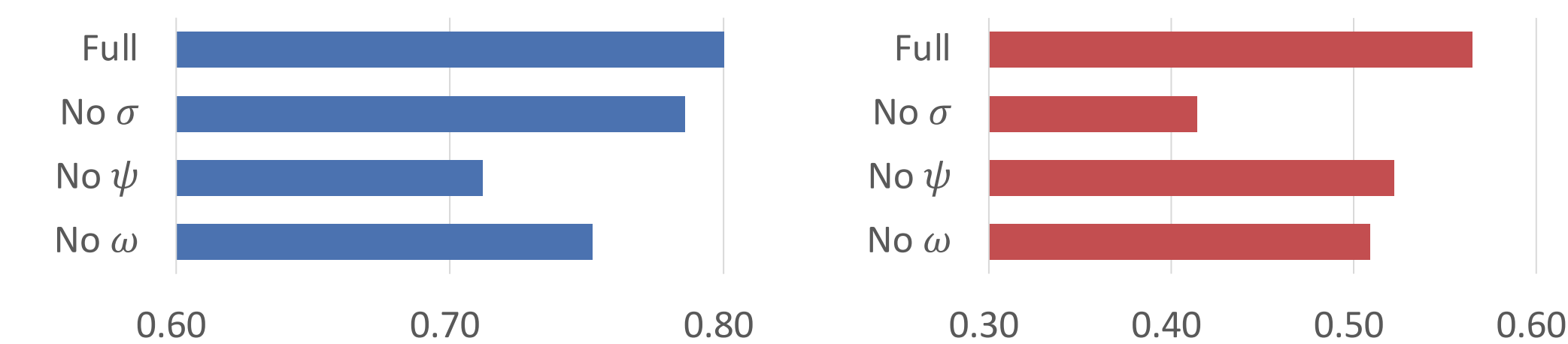
Bars with the same color can be directly compared in each plot



## Effects of Paths



## Effects of Components



## Simulated Data

## Clinical Data

## Conclusion

A flexible multi-path framework to assess surgical skills  
 Effective on both simulated and clinical surgery data

## Future Works

More paths and better functions  
 More clinical data from ORs

## Code Released

[github.com/Finspire13](https://github.com/Finspire13)