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, AIIT 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 clearances e.g., Skill Tool Usage
- high visibility vs. limited visibility

Event Pattern
- linear, normal vs. loopy, adverse
- Surgical Video

Surgical tools 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

Surgical Video

Path 1
Input Feature \( \chi_1 \)
\( \phi_1 \)
Embedding
\( \lambda_1 \)
Precised Embedding
\( \sigma_1 \)
Contrastive Loss

Path 2
Input Feature \( \chi_2 \)
\( \phi_2 \)
Embedding
\( \lambda_2 \)
Precised Embedding
\( \sigma_2 \)
Contrastive Loss

Each path takes care of a skill aspect, with four paths in total

Path 2
Input Feature \( \chi_2 \)
\( \phi_2 \)
Embedding
\( \lambda_2 \)
Precised Embedding
\( \sigma_2 \)
Contrastive Loss

Path Dependency Module

Surgical Video

Path 1
Input Feature \( \chi_1 \)
\( \phi_1 \)
Embedding
\( \lambda_1 \)
Score Sequence
\( \omega_1 \)
Weight Sequence
\( q \)
MSE Loss

Path 2
Input Feature \( \chi_2 \)
\( \phi_2 \)
Embedding
\( \lambda_2 \)
Score Sequence
\( \omega_2 \)
Weight Sequence
\( q \)
MSE Loss

A unified multi-path framework

Path Input Features

Instantiation

Contrastive Loss

Path 1
Input Feature \( \chi_1 \)
\( \phi_1 \)
Embedding
\( \lambda_1 \)
Score Sequence
\( \omega_1 \)
Weight Sequence
\( q \)
MSE Loss

Path 2
Input Feature \( \chi_2 \)
\( \phi_2 \)
Embedding
\( \lambda_2 \)
Score Sequence
\( \omega_2 \)
Weight Sequence
\( q \)
MSE Loss

Contrastive Loss

Fusion

Path 1 predictions from MS-TCN
Path 2 predictions from MS-TCN

Predict the current using the past

Predicted Embedding

Negatives: Embedding at the others in a neighborhood

Temporal attention mechanism and multi-path fusion

Modeling the relationships among aspects

To alleviate the scarcity of annotated data

Comparisons to State-of-the-Art

Effects of Paths

Effects of Components

Our Clinical Data
- 20 gastrectomy videos
- Captured by laparoscopy
- 199 minutes per video
- 7 OSATS skill metrics

Detailed Event Annotation (41 Classes)
- 13 coarse-grained procedural events
- 13 fine-grained procedural events
- 8 adverse events and 2 repair events
- 5 video recording events

Future Works
- More paths and better functions
- More clinical data from ORs

Code Released

github.com/Finspire13