LaTeX Author Guidelines for CVPR Proceedings

Anonymous CVPR submission

Paper ID *****

Abstract

The ABSTRACT is to be in fully justified italicized text, at the top of the left-hand column, below the author and affiliation information. Use the word "Abstract" as the title, in 12-point Times, boldface type, centered relative to the column, initially capitalized. The abstract is to be in 10-point, single-spaced type. Leave two blank lines after the Abstract, then begin the main text. Look at previous CVPR abstracts to get a feel for style and length.

1. Introduction

2. Related Work

Progress Prediction. Activity progress, how complete an acitivity is, was first introduced by [1]. They introduce ProgressNet, an LSTM based network, to predict activity progress on the ucf24 dataset.

Phase Prediction. The current phase an activity is in is often correlated to the progress of the activity, especially in linear activities.

Remaining Surgery Duration. The remaining time left in a surgery, or any other activity, is directly correlated to the progress of the activity. [7] use this to jointly predict surgery progress and remaining surgery duration.

3. Method

3.1. Definition of Progress

3.2. Networks

We implemented 3 different networks. ProgressNet [1], RSDNet [7], and UTE [4].

3.3. Data

We use the following datasets: Breakfast (BF) [2, 3], UCF101-24 [6], and cholec80 [5]. For BF we have dense trajectories, RSDNet embeddings, and ProgressNet embeddings. For UCF101-24 and Cholec80 we have i3d embeddings, RSDNet embeddings, and ProgressNet embeddings.

4. Experiments

Why is this interesting!?

5. Conclusion

References

[1] Federico Becattini, Tiberio Uricchio, Lorenzo Seidenari, Lamberto Ballan, and Alberto Del Bimbo. Am i done? predicting action progress in videos, 2017. 1

[2] H. Kuehne, A. B. Arslan, and T. Serre. The language of actions: Recovering the syntax and semantics of goal-directed human activities. In *Proceedings of Computer Vision and Pattern Recognition Conference (CVPR)*, 2014. 1

[3] Hilde Kuehne, Juergen Gall, and Thomas Serre. An end-to-end generative framework for video segmentation and recognition. In *Proc. IEEE Winter Applications of Computer Vision Conference (WACV 16)*, Lake Placid, Mar 2016. 1

[4] Anna Kukleva, Hilde Kuehne, Fadime Sener, and Juergen Gall. Unsupervised learning of action classes with continuous temporal embedding, 2019. 1

[5] Chinedu Innocent Nwoye, Tong Yu, Cristians Gonzalez, Barbara Seeliger, Pietro Mascagni, Didier Mutter, Jacques Marescaux, and Nicolas Padoy. Rendezvous: Attention mechanisms for the recognition of surgical action triplets in endoscopic videos. *Medical Image Analysis*, 78:102433, may 2022. 1

[6] Khurram Soomro, Amir Roshan Zamir, and Mubarak Shah. Ucf101: A dataset of 101 human actions classes from videos in the wild, 2012. 1

[7] Andru Putra Twinanda, Gaurav Yengera, Didier Mutter, Jacques Marescaux, and Nicolas Padoy. RSDNet: Learning to predict remaining surgery duration from laparoscopic videos without manual annotations. *IEEE Transactions on Medical Imaging*, 38(4):1069–1078, apr 2019. 1

	breakfast		cholec80		ucf101-24	
	Normal	Indices	Normal	Indices	Normal	Indices
UTE	0.114	0.019	0.050	0.024	0.103	0.034
RSDNet	0.067	0.058	0.024	0.024[1]	0.195	0.097[2]
ProgressNet	0.070	0.054	0.078	0.031	0.103	0.054[3]
Average		0.019		0.025		0.034
0.5		0.083		0.083		0.083
Random		0.166		0.166		0.166

Table 1. Normal Data vs. Indices (MSE Loss)

Data	MSE Loss
BF train/test (Dense Trajectories)	0.114
BF train/test (ResNet embeddings)	-
BF train/test (Indices)	0.019
BF train/train (Dense Trajectories)	0.041
BF train/train (ResNet embeddings)	-
BF train/train (Indices)	0.017
Cholec80 (i3d embeddings)	0.050
Cholec80 (ResNet embeddings)	-
Cholec80 (Indices)	0.024
UCF101-24 (i3d embeddings)	0.103
UCF101-24 (ResNet embeddings)	-
UCF101-24 (Indices)	0.034

Table 2. UTE

Data	MSE Loss
BF train/test (Dense Trajectories)	0.067
BF train/test (ResNet embeddings)	-
BF train/test (Indices)	0.058
Cholec80 (i3d embeddings)	0.031
Cholec80 (ResNet embeddings)	0.024
Cholec80 (Indices)	0.024
Cholec80 sampled (i3d embeddings)	0.116
Cholec80 sampled (ResNet embeddings)	0.05
Cholec80 sampled (Indices)	0.070
UCF101-24 (i3d embeddings)	0.195*
UCF101-24 (RSD embeddings)	-
UCF101-24 (Indices)	0.097*

Table 3. RSDNet

Data	MSE Loss
BF train/test (Dense Trajectories)	0.070
BF train/test (ResNet embeddings)	-
BF train/test (Indices)	0.054
Cholec80 (i3d embeddings)	0.078
Cholec80 (ResNet embeddings)	-
Cholec80 (Indices)	0.031
UCF101-24 (i3d embeddings)	0.103
UCF101-24 (RSD embeddings)	-
UCF101-24 (Indices)	0.054
UCF101-24 (Frames & Boxes)	-

Table 4. ProgressNet

Data	MAE (minutes)
Cholec80 (i3d embeddings)	12.58
Cholec80 (RSD embeddings)	11.06
Cholec80 (Indices)	11.21
Cholec80 sampled (i3d embeddings)	16.30
Cholec80 sampled (RSD embeddings)	12.95
Cholec80 sampled (Indices)	18.13

Table 5. RSD Predictions

Dataset	MoF	iou	Mean F1
Dense Trajectories	45.1	12.7	26.5
Indices	24.3	12.8	24.0
Raw Indices	25.2	12.4	24.7

Table 6. Unsupervised Action Classes