

# $\text{\LaTeX}$ 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 activity 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.

Data	MSE Loss
BF train/test (Dense Trajectories)	0.114
BF train/test (ResNet embeddings)	-
BF train/test (Indices)	0.019
BF all (Dense Trajectories)	0.041
BF all (ResNet embeddings)	-
BF all (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 1. UTE

Data	MSE Loss
BF train/test (Dense Trajectories)	-
BF train/test (ResNet embeddings)	-
BF train/test (Indices)	-
BF all (Dense Trajectories)	-
BF all (RResNetSD embeddings)	-
BF all (Indices)	-
Cholec80 (i3d embeddings)	0.112
Cholec80 (ResNet embeddings)	0.019
Cholec80 (Indices)	0.018
UCF101-24 (i3d embeddings)	0.060
UCF101-24 (RSD embeddings)	-
UCF101-24 (Indices)	0.033

Table 2. RSDNet

## 4. Experiments

Why is this interesting!?

## 5. Conclusion

Data	MSE Loss
BF train/test (Dense Trajectories)	-
BF train/test (ResNet embeddings)	-
BF train/test (Indices)	-
BF all (Dense Trajectories)	-
BF all (RResNetSD embeddings)	-
BF all (Indices)	-
Cholec80 (i3d embeddings)	-
Cholec80 (ResNet embeddings)	-
Cholec80 (Indices)	-
UCF101-24 (i3d embeddings)	-
UCF101-24 (RSD embeddings)	-
UCF101-24 (Indices)	-
UCF101-24 (Frames & Boxes)	-

Table 3. ProgressNet

Data	MAE (minutes)
Cholec80 (i3d embeddings)	17.78
Cholec80 (RSD embeddings)	8.64
Cholec80 (Indices)	7.35

Table 4. RSD Predictions

without manual annotations. *IEEE Transactions on Medical Imaging*, 38(4):1069–1078, apr 2019. 1

## 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