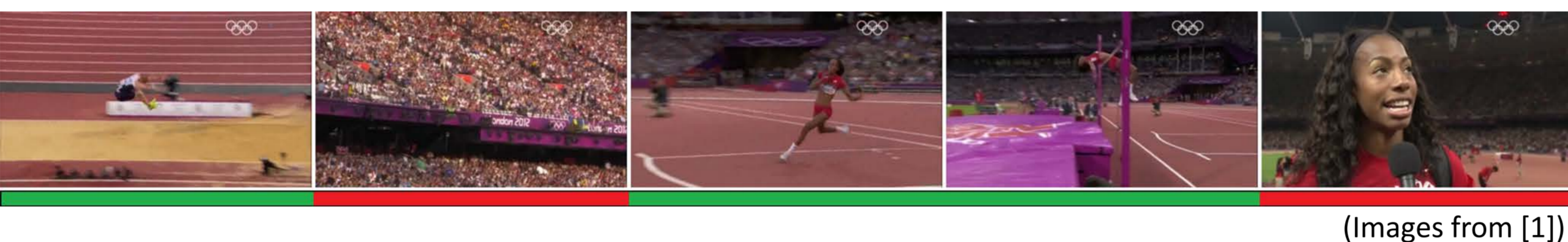


Investigation on Combining 3D Convolution of Image Data and Optical Flow to Generate Temporal Action Proposals

Patrick Schlosser, David Münch, Michael Arens

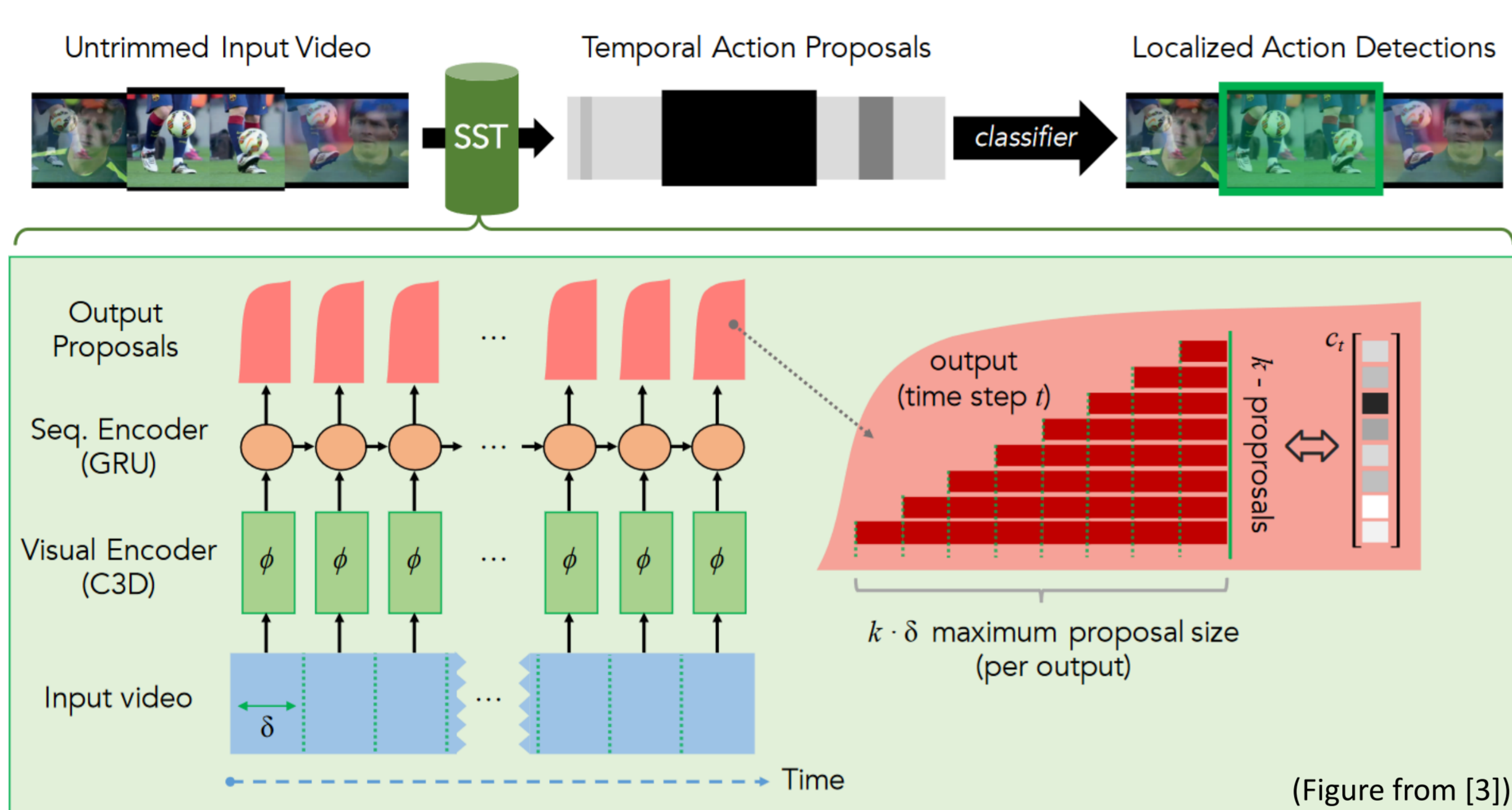
Motivation



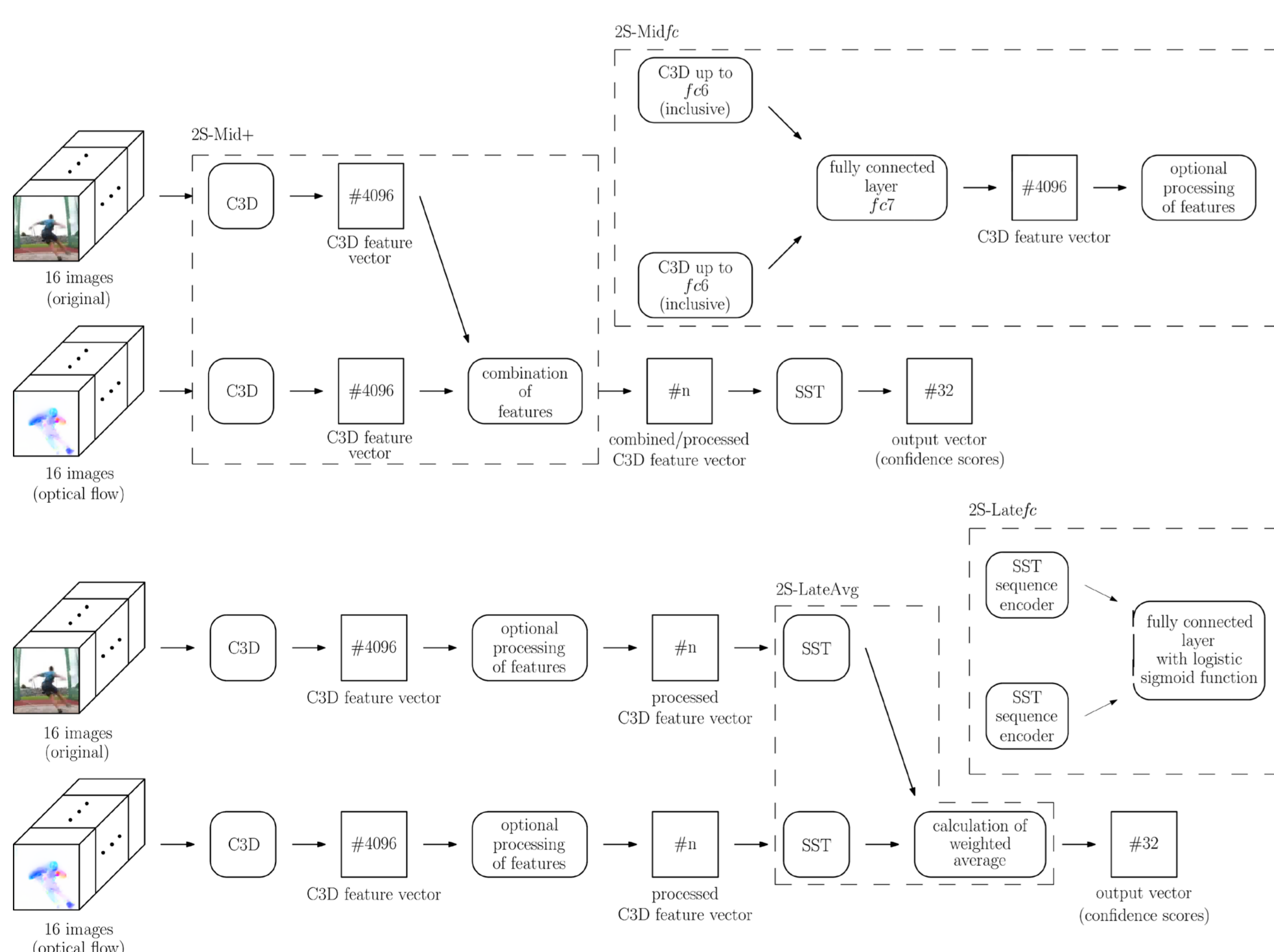
- Detection of time segments with 'actions of interest', e.g. all time segments with sporting activities from an untrimmed video.
- Preprocessing step for later classification of actions inside the time segments.

Approach

- C3D [2] and SST [3] network used as a basis for all model architectures.

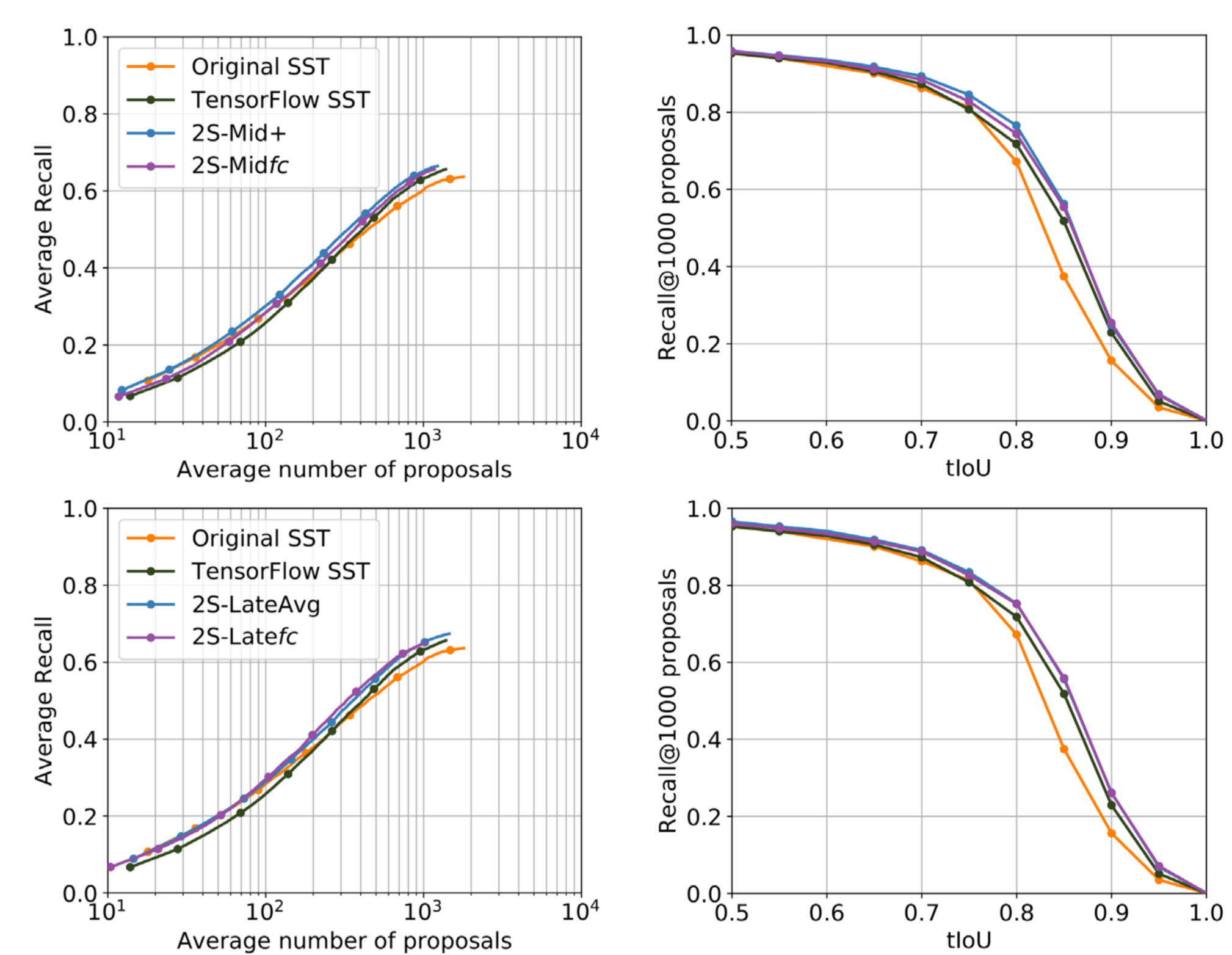


- Development of four different two-stream models with 3D convolutions on both streams (images + optical flow):



Evaluation & Results

- Evaluation on THUMOS'14 [4] dataset.
- All two-stream models achieve improvements.
- Improvements still achieved when optical flow is calculated by FlowNet2 [5] instead of Brox [6].



Network	Score
Original SST network	0.6025
TensorFlow SST network (images)	0.6295
TensorFlow SST network (optical flow)	0.6320
2S-Mid+	0.6497
2S-Midfc	0.6438
2S-LateAvg	0.6495
2S-Latefc	0.6466
2S-LateAvgFN (optical flow by FlowNet2)	0.6436

(Score: average recall at average 1000 proposals)

Conclusion

- Two-stream approach with 3D convolutions useful for temporal action proposals: all models achieved improvements.
- Improvements not bound to a single method of calculating optical flow.



References

- [1] A. Gorban, H. Idrees, Y.-G. Jiang, A. R. Zamir, I. Laptev, M. Shah, and R. Sukthankar. THUMOS challenge: Action recognition with a large number of classes. <http://www.thumos.info>, 2015.
- [2] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri. Learning spatiotemporal features with 3d convolutional networks. ICCV, 2015.
- [3] S. Buch, V. Escorcia, C. Shen, B. Ghanem, and J. C. Niebles. SST: Single-stream temporal action proposals. CVPR, 2017.
- [4] Y.-G. Jiang, J. Liu, A. R. Zamir, G. Toderici, I. Laptev, M. Shah, and R. Sukthankar. THUMOS challenge: Action recognition with a large number of classes. <http://crv.ucf.edu/THUMOS14/>, 2014.
- [5] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox. FlowNet 2.0: Evolution of optical flow estimation with deep networks. CVPR, 2017.
- [6] T. Brox, A. Bruhn, N. Papenberger, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. ECCV. Springer, 2004.

