

# Supplementary Material for

## A Generative Appearance Model for End-to-end Video Object Segmentation

Joakim Johnander<sup>1,3</sup>

Martin Danelljan<sup>1,2</sup>

Emil Brissman<sup>1,4</sup>

Fahad Shahbaz Khan<sup>1,5</sup>

Michael Felsberg<sup>1</sup>

<sup>1</sup> CVL, Linköping University, Sweden

<sup>2</sup> CVL, ETH Zürich, Switzerland

<sup>3</sup> Zenuity, Sweden

<sup>4</sup> Saab, Sweden

<sup>5</sup> IIAI, UAE

## Appendices

### A. Supplementary Video

Additional qualitative results are provided in a supplementary video (`proposed_and_rgmp_short.mp4`). We show several challenging sequences from the DAVIS2017 [1] and YouTube-VOS [3] validation sets, together with the corresponding predictions from both the proposed method (top), and RGMP [2] (bottom). The sequences contain scenarios where the target is occluded, moves rapidly, and where its appearance changes, for instance due to deformations or in-plane rotations. The proposed approach is able to model the target appearance sufficiently well to be able to deal with these situations, without resorting to extensive online fine-tuning.

### References

- [1] J. Pont-Tuset, F. Perazzi, S. Caelles, P. Arbeláez, A. Sorkine-Hornung, and L. Van Gool. The 2017 davis challenge on video object segmentation. *arXiv:1704.00675*, 2017. 1
- [2] S. Wug Oh, J.-Y. Lee, K. Sunkavalli, and S. Joo Kim. Fast video object segmentation by reference-guided mask propagation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7376–7385, 2018. 1
- [3] N. Xu, L. Yang, Y. Fan, J. Yang, D. Yue, Y. Liang, B. L. Price, S. Cohen, and T. S. Huang. Youtube-vos: Sequence-to-sequence video object segmentation. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part V*, pages 603–619, 2018. 1