# Supplementary for Simultaneous Detection and Removal of Dynamic Objects in Multi-view Images

Gagan Kanojia Shanmuganathan Raman
Indian Institute of Technology Gandhinagar, India
{gagan.kanojia, shanmuga}@iitgn.ac.in

## 1. Results and Comparisons

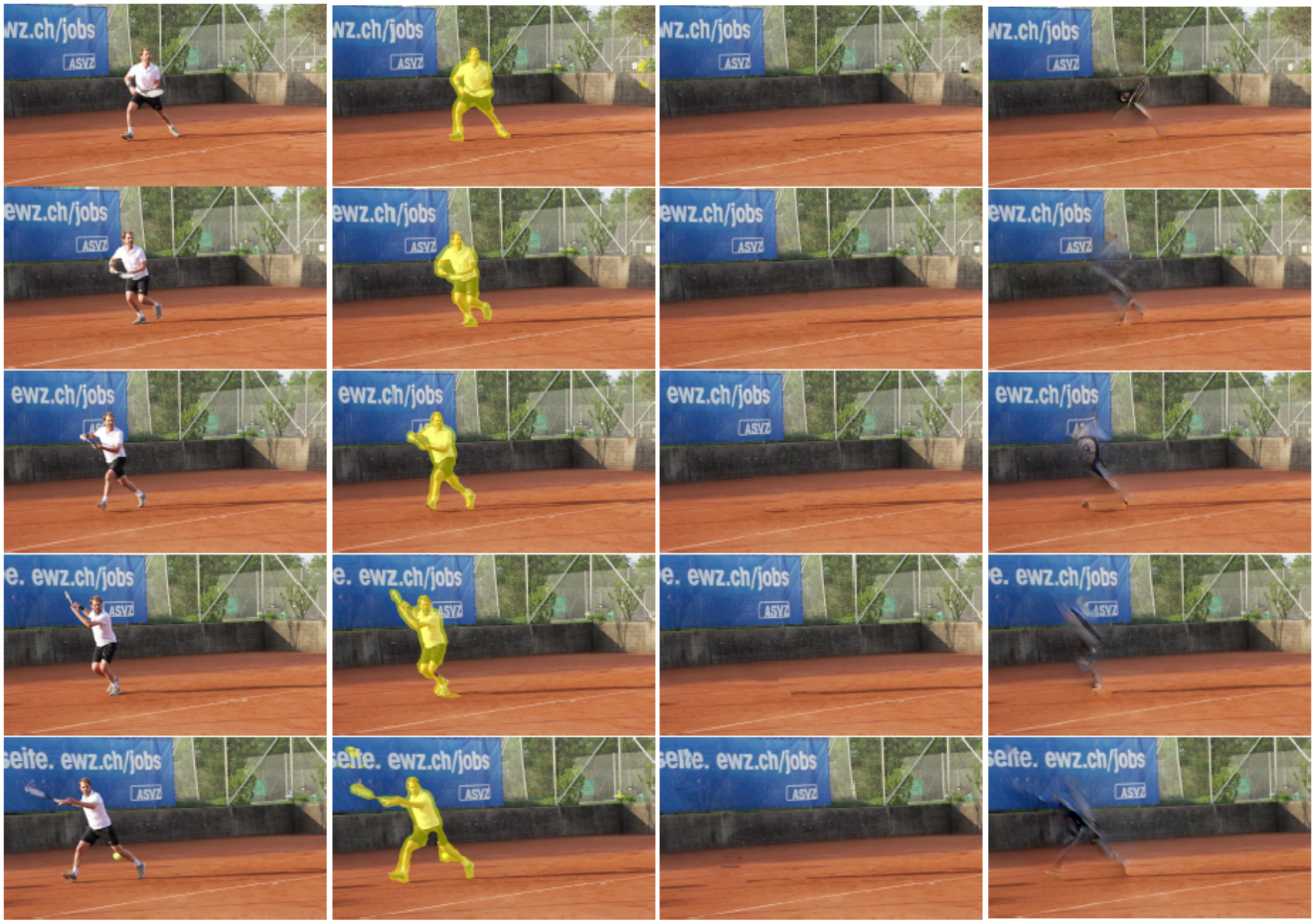
Table 1 shows the quantitative results of the dynamic object detection obtained using the proposed approach on the multi-view sets extracted from the DAVIS dataset. The first column shows the mean Jaccard index obtained on the input set. The second and third columns show the obtained true positives and false negatives. In our case, positive refers to the regions belonging to dynamic objects and negative refers to the static regions. It can be observed that we obtain high true positive rate and low false negative rate. Since, we not only want to detect the dynamic regions but also remove them, high true positives rate is very crucial for the task.

Figure 1, 2, 3, 4 show the dynamic object detection and removal results obtained on the image set extracted from the tennis, bmx-bumps, bmx-trees and paragliding scene of DAVIS dataset [5] using the proposed approach, respectively. They also show the comparison of the image completion obtained using Iizuka *et al.* [3] (when the dynamic map is provided to it) with the image completion obtained while simultaneously detecting the dynamic objects using the proposed approach. Figure 5, 6, 7 show the dynamic object detection and removal results obtained on the image set extracted from the cars6, people1, and rabbits04 scene of of Freiburg Berkeley Motion Segmentation Dataset [1] using the proposed approach, respectively. They also show the comparison of the image completion obtained using Iizuka *et al.* [3] (when the dynamic map is provided to it) with the image completion obtained while simultaneously detecting the dynamic objects using the proposed approach.

Figure 9, 10, 11, 12 show the comparison between the dynamic object detection results obtained on the skateboard image set [2] and image sets extracted from dog-gooses, paragliding and boxing-fisheye scene of the DAVIS dataset [5] using Kanojia et al. [4] and the results obtained using the proposed approach, respectively.

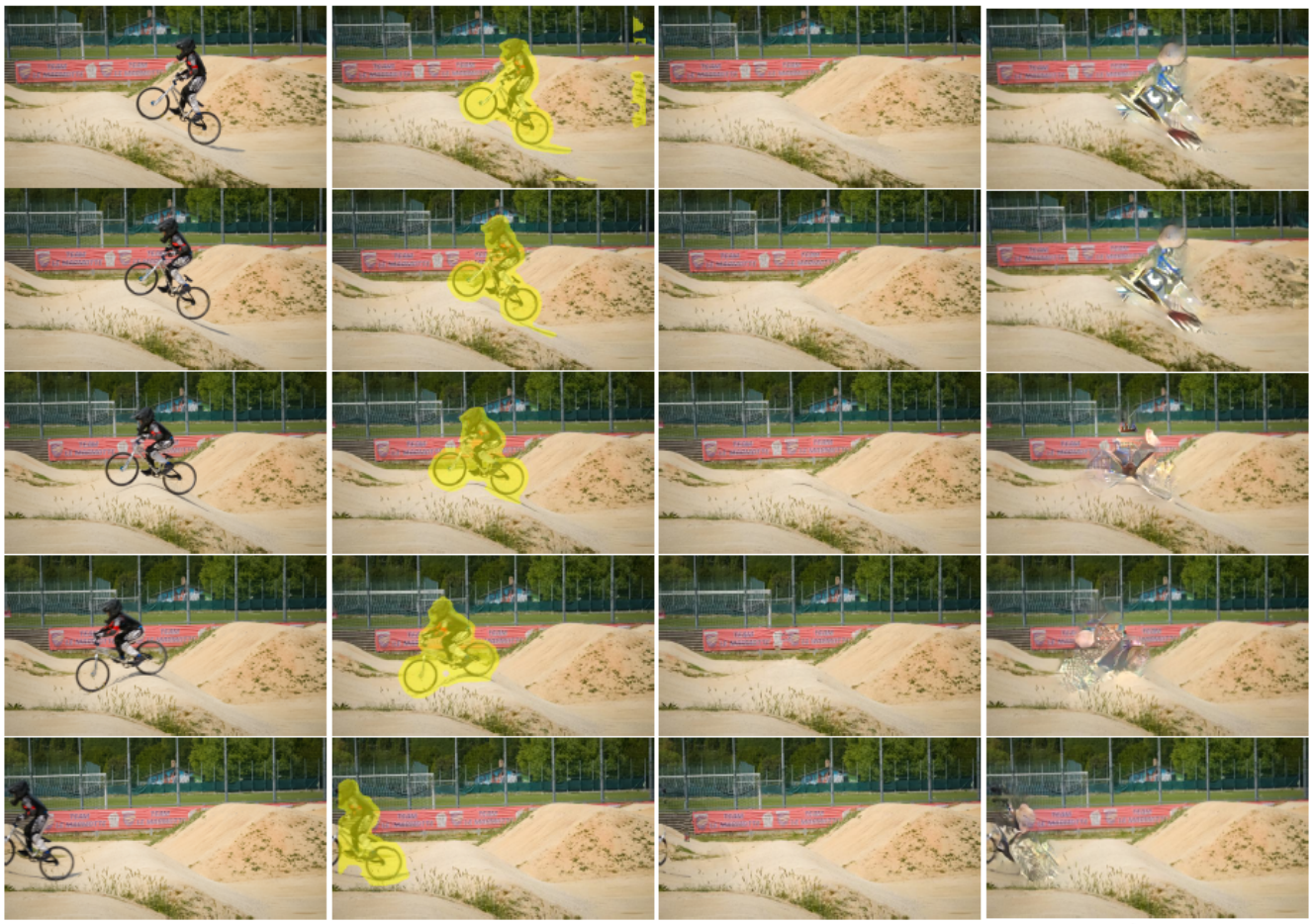
| Multi-view set | Jaccard Index | True positive | False Negatives |
|---|---|---|---|
| boxing-fisheye | 0.58 | 0.96 | 0.04 |
| rollerblade | 0.56 | 0.998 | 0.002 |
| tennis | 0.65 | 0.98 | 0.02 |
| dog-gooses | 0.43 | 0.91 | 0.09 |
| bmx-bumps | 0.37 | 0.99 | 0.01 |
| bmx-trees | 0.29 | 0.94 | 0.06 |

Table 1. The table shows the quantitative results of the dynamic object detection obtained using the proposed approach on the multi-view sets extracted from the DAVIS dataset [5].



(a) Input multi-view set (b) Ours (c) Ours (d) [3]

Figure 1. The figure shows the dynamic object detection and removal results obtained on the image set extracted from the tennis scene of DAVIS dataset [5] using the proposed approach. It also shows the comparison of the image completion obtained using Iizuka *et al.* [3] (when the dynamic map is provided to it) with the image completion obtained while simultaneously detecting the dynamic objects using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) and (c) shows the dynamic object detection and the removal results obtained using the proposed approach, respectively. (d) shows the image completion obtained using [3] (when the dynamic map is provided to it).



(a) Input multi-view set (b) Ours (c) Ours (d) [3]

Figure 2. The figure shows the dynamic object detection and removal results obtained on the image set extracted from the bmx-bumps scene of DAVIS dataset [5] using the proposed approach. It also shows the comparison of the image completion obtained using Iizuka *et al.* [3] (when the dynamic map is provided to it) with the image completion obtained while simultaneously detecting the dynamic objects using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) and (c) shows the dynamic object detection and the removal results obtained using the proposed approach, respectively. (d) shows the image completion obtained using [3] (when the dynamic map is provided to it).

(a) Input multi-view set (b) Ours (c) Ours (d) [3]

Figure 3. The figure shows the dynamic object detection and removal results obtained on the image set extracted from the bmx-trees scene of DAVIS dataset [5] using the proposed approach. It also shows the comparison of the image completion obtained using Iizuka *et al.* [3] (when the dynamic map is provided to it) with the image completion obtained while simultaneously detecting the dynamic objects using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) and (c) shows the dynamic object detection and the removal results obtained using the proposed approach, respectively. (d) shows the image completion obtained using [3] (when the dynamic map is provided to it).

(a) Input multi-view set (b) Ours (c) Ours (d) [3]

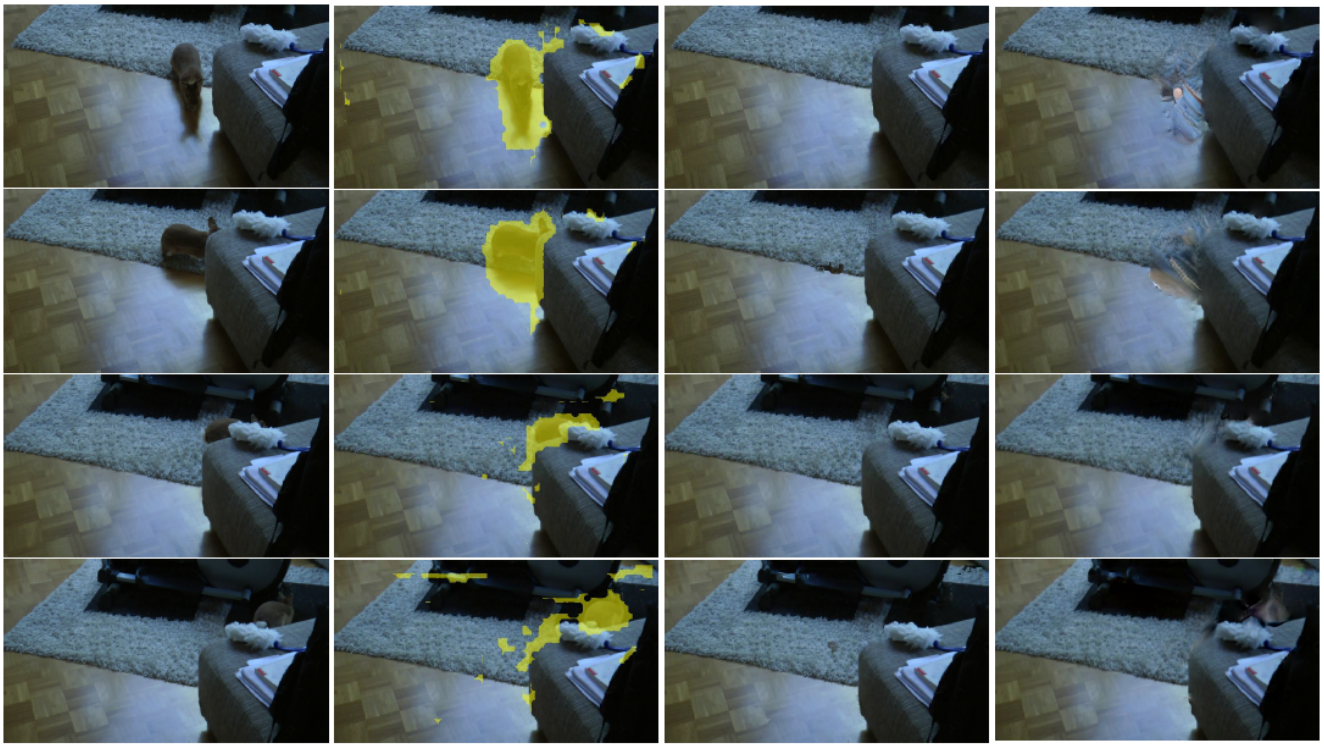
Figure 4. The figure shows the dynamic object detection and removal results obtained on the image set extracted from the paragliding scene of DAVIS dataset [5] using the proposed approach. It also shows the comparison of the image completion obtained using Iizuka *et al.* [3] (when the dynamic map is provided to it) with the image completion obtained while simultaneously detecting the dynamic objects using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) and (c) shows the dynamic object detection and the removal results obtained using the proposed approach, respectively. (d) shows the image completion obtained using [3] (when the dynamic map is provided to it).

(a) Input multi-view set (b) Ours (c) Ours (d) [3]

Figure 5. The figure shows the dynamic object detection and removal results obtained on the image set extracted from the cars6 scene of Freiburg Berkeley Motion Segmentation Dataset [1] using the proposed approach. It also shows the comparison of the image completion obtained using Iizuka *et al.* [3] (when the dynamic map is provided to it) with the image completion obtained while simultaneously detecting the dynamic objects using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) and (c) shows the dynamic object detection and the removal results obtained using the proposed approach, respectively. (d) shows the image completion obtained using [3] (when the dynamic map is provided to it).

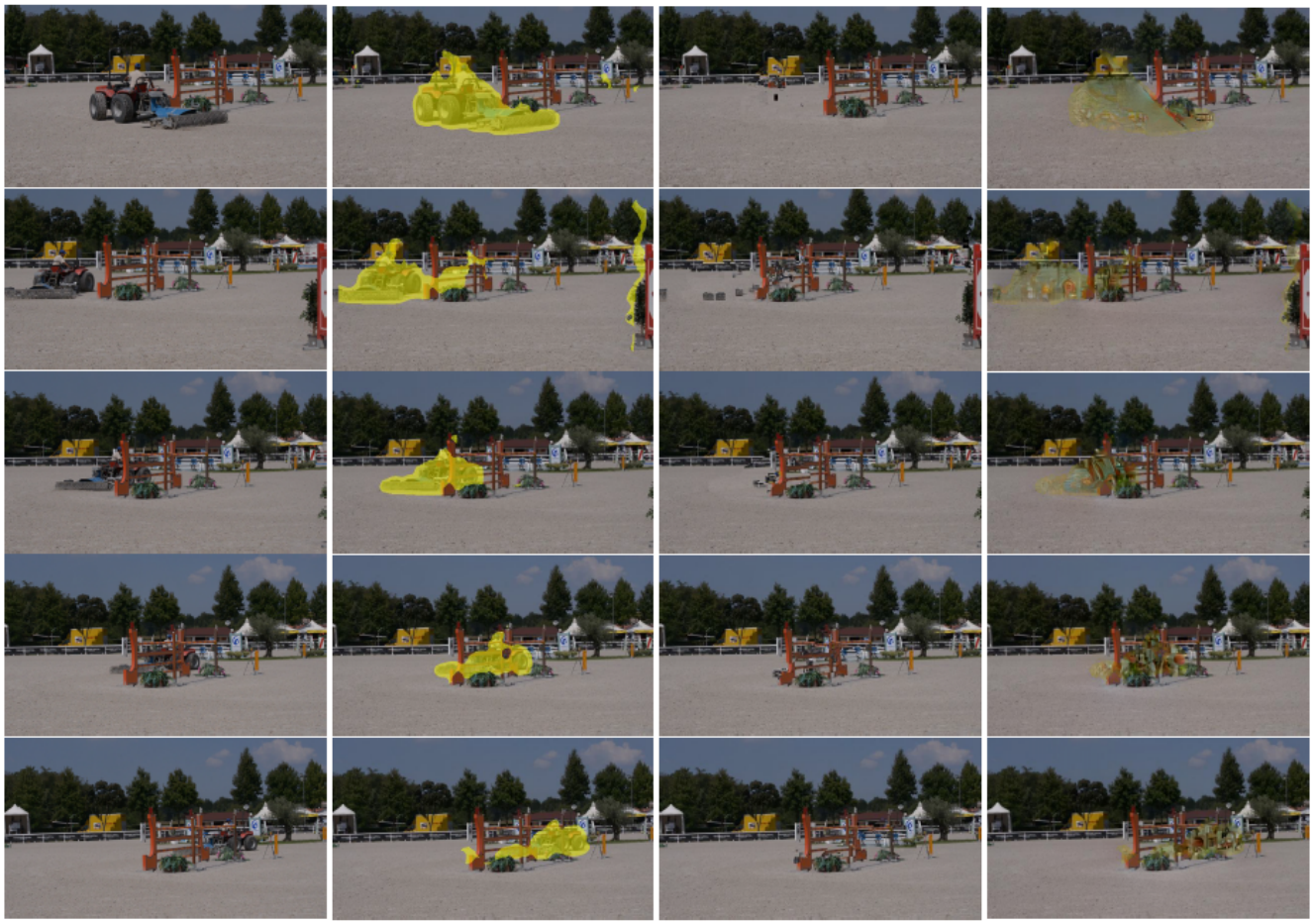
(a) Input multi-view set (b) Ours (c) Ours (d) [3]

Figure 6. The figure shows the dynamic object detection and removal results obtained on the image set extracted from the people1 scene of Freiburg Berkeley Motion Segmentation Dataset [1] using the proposed approach. It also shows the comparison of the image completion obtained using Iizuka *et al.* [3] (when the dynamic map is provided to it) with the image completion obtained while simultaneously detecting the dynamic objects using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) and (c) shows the dynamic object detection and the removal results obtained using the proposed approach, respectively. (d) shows the image completion obtained using [3] (when the dynamic map is provided to it).



(a) Input multi-view set (b) Ours (c) Ours (d) [3]

Figure 7. The figure shows the dynamic object detection and removal results obtained on the image set extracted from the rabbits04 scene of Freiburg Berkeley Motion Segmentation Dataset [1] using the proposed approach. It also shows the comparison of the image completion obtained using Iizuka *et al.* [3] (when the dynamic map is provided to it) with the image completion obtained while simultaneously detecting the dynamic objects using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) and (c) shows the dynamic object detection and the removal results obtained using the proposed approach, respectively. (d) shows the image completion obtained using [3] (when the dynamic map is provided to it).

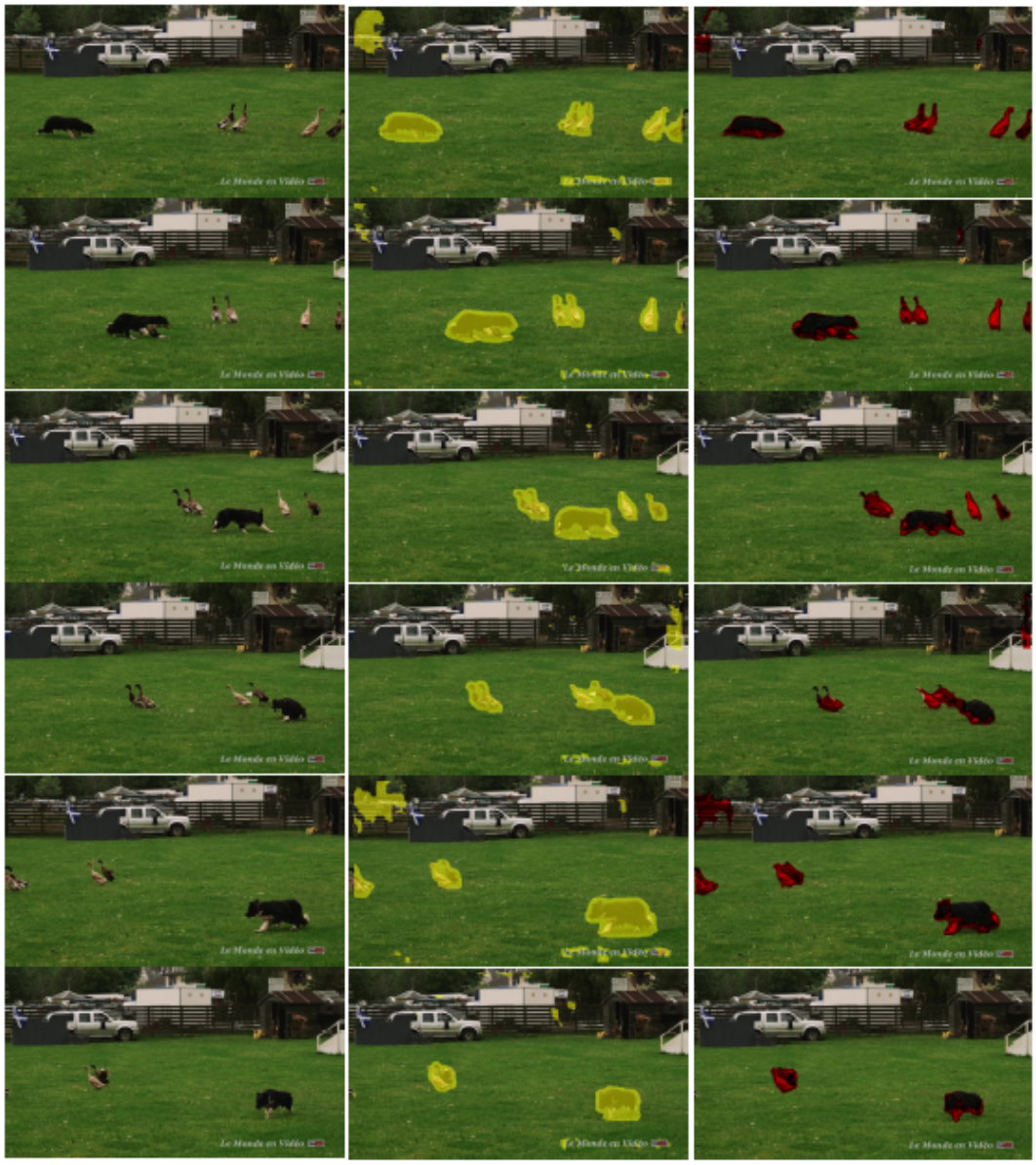


(a) Input multi-view set (b) Ours (c) Ours (d) [3]

Figure 8. The figure shows the dynamic object detection and removal results obtained on the image set extracted from the tractor-sand scene of Freiburg Berkeley Motion Segmentation Dataset [1] using the proposed approach. It also shows the comparison of the image completion obtained using Iizuka *et al.* [3] (when the dynamic map is provided to it) with the image completion obtained while simultaneously detecting the dynamic objects using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) and (c) shows the dynamic object detection and the removal results obtained using the proposed approach, respectively. (d) shows the image completion obtained using [3] (when the dynamic map is provided to it).

(a) Input multi-view set (b) Ours (c) [4]

Figure 9. The figure shows the comparison between the dynamic object detection results obtained on the skateboard set used in [4] using Kanojia et al. [4] and the results obtained using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) shows the dynamic object detection obtained using the proposed approach. (c) shows the dynamic object detection results obtained using the approach by Kanojia et al. [4].



(a) Input multi-view set (b) Ours (c) [4]

Figure 10. The figure shows the comparison between the dynamic object detection results obtained on the image set extracted from dog-gooses scene of the DAVIS dataset [5] using Kanojia et al. [4] and the results obtained using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) shows the dynamic object detection obtained using the proposed approach. (c) shows the dynamic object detection results obtained using the approach by Kanojia et al. [4].

(a) Input multi-view set (b) Ours (c) [4]

Figure 11. The figure shows the comparison between the dynamic object detection results obtained on the image set extracted from paragliding scene of the DAVIS dataset [5] using Kanojia et al. [4] and the results obtained using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) shows the dynamic object detection obtained using the proposed approach. (c) shows the dynamic object detection results obtained using the approach by Kanojia et al. [4].

(a) Input multi-view set (b) Ours (c) [4]

Figure 12. The figure shows the comparison between the dynamic object detection results obtained on the image set extracted from boxing-fisheye scene of the DAVIS dataset [5] using Kanojia et al. [4] and the results obtained using the proposed approach. (a) shows the reference images of the input multi-view image set with dynamic objects. (b) shows the dynamic object detection obtained using the proposed approach. (c) shows the dynamic object detection results obtained using the approach by Kanojia et al. [4].

## References


[1] T. Brox, J. Malik, and P. Ochs. Freiburg-berkeley motion segmentation dataset (fbms-59). In *European Conference on Computer Vision (ECCV)*, 2010.

[2] A. Dafni, Y. Moses, S. Avidan, and T. Dekel. Detecting moving regions in crowdcam images. *Computer Vision and Image Understanding*, 160:36–44, 2017.

[3] S. Iizuka, E. Simo-Serra, and H. Ishikawa. Globally and locally consistent image completion. *ACM Transactions on Graphics (TOG)*, 36(4):107, 2017.

[4] G. Kanojia and S. Raman. Patch-based detection of dynamic objects in crowdcam images. *The Visual Computer*, pages 1–14.

[5] F. Perazzi, J. Pont-Tuset, B. McWilliams, L. Van Gool, M. Gross, and A. Sorkine-Hornung. A benchmark dataset and evaluation methodology for video object segmentation. In *Computer Vision and Pattern Recognition*, 2016.