Tracking by Segmentation

Jonáš Šerých, Jan Čech
Tracking by segmentation

- Correlation filter trackers
  - Usually axis-aligned bbox
  - Translation OK, Scale ok, rotation ?, Affine X

- Tracking by segmentation
  - Not limited to simple transformations
  - Segmentation captures reality better
Tracking by segmentation

Segmentation on first frame

Segmentation throughout the video
Tracking by segmentation

- DAVIS dataset – Densely Annotated Video Segmentation
  - 50 Full HD videos, 24 fps
  - 3455 annotated frames with pixel-level segmentation

https://youtu.be/8f9y17-OAwI

OSVOS architecture

1 – 2 x conv(3, 3, 64)
2 – maxpool, 3 x conv(3, 3, 128)
3 – maxpool, 3 x conv(3, 3, 256)
4 – maxpool, 3 x conv(3, 3, 512)
5 – maxpool, 3 x conv(3, 3, 512)
↑ – conv(3, 3, 16), bilinear_upsampling(H, W)
C – concat
L – conv(1, 1, 1), sigmoid


- Simple adaptation of VGG
- Video frames processed one by one independently
Training

1) Backbone CNN ImageNet pre-training
2) Segmentation pre-training (DAVIS, PASCAL VOC, …)
   - Strip fully connected layers
   - Convert into fully convolutional segmentation CNN
3) Fine-tuning on first frame

OnAVOS

- Newly appearing objects – issue for OSVOS

=> Online adaptation

Dilated (Atrous) convolutions

- Used in current state-of-the-art semantic segmentation DNNs
  - e.g. DeepLabv3+ (top on PASCAL VOC2012 leaderboard 20.4.2018)

- Increase receptive field
- Same number of parameters
- Same number of operations

MaskTrack

- Semantic segmentation CNN (DeepLabv2)
  - With dilated convolutions
- „Refine the segmentation from last frame“
- Mask as additional input channel - RGB+Mask

Input frame $t$

Mask estimate $t-1$

MaskTrack ConvNet

Refined mask $t$


https://youtu.be/Ze7dKwwAw8o
https://youtu.be/G8RbuKI_784
LucidTrack

- Like MaskTrack + uses optical flow
- **Complex data augmentation**
  - Very good results without segmentation training phase!
  - Very good even without ImageNet pretraining!

Original image:

Generated image:

[https://youtu.be/7aZI0BjmrA8](https://youtu.be/7aZI0BjmrA8)
[https://youtu.be/QrsR5w-HR14](https://youtu.be/QrsR5w-HR14)