Video Object Segmentation using Deep Learning

Update Presentation, Week 6

Zack While

Advised by: Rui Hou, Dr. Chen Chen, and Dr. Mubarak Shah

June 23, 2017

Youngstown State University
1. Previous Work

2. Current Work

3. Upcoming Work
Previous Work
DAVIS 2017 Issues
■ Mask CNN by He et al. (arXiv, April 2017)
DAVIS Mask Change
DAVIS 2016 Initial Results
Current Work
Goals

- Fine-tuning and further working on the model for the DAVIS 2016 dataset.

- Reading related material as needed.
Goals

- Fine-tuning and further working on the model for the DAVIS 2016 dataset.

- Reading related material as needed.
Loss Curve

![Graph showing the loss curve and learning rate over iterations. The x-axis represents the number of iterations, ranging from 0 to 60000, with logarithmic scale for both axes. The loss curve shows a rapid decrease followed by a plateau, while the learning rate plot shows a sharp drop at iteration 40000.]
Figure 1: IoU: 0.3605
Figure 1: IoU: 0.3605
Results (2)

**Figure 2:** IoU: 0.2132

- **Actual Clip**
- **Ground Truth Mask**
- **Predicted Mask Heatmap**
- **Predicted Mask, th=0.65**
Figure 2: IoU: 0.2132
Figure 3: IoU: 0.3944
Results (3)

Figure 3: IoU: 0.3944

Actual Clip

Predicted Mask Heatmap

Ground Truth Mask

Predicted Mask, th=0.45
Quantitative Results

Table 1: IoU Comparison

<table>
<thead>
<tr>
<th>Method</th>
<th>Average IoU (DAVIS 2016)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCF-C3D</td>
<td>0.3479</td>
</tr>
<tr>
<td>FST</td>
<td>0.558</td>
</tr>
<tr>
<td>CUT</td>
<td>0.552</td>
</tr>
<tr>
<td>NLC</td>
<td>0.551</td>
</tr>
<tr>
<td>MSG</td>
<td>0.533</td>
</tr>
<tr>
<td>KEY</td>
<td>0.498</td>
</tr>
<tr>
<td>CVOS</td>
<td>0.482</td>
</tr>
</tbody>
</table>

Comparison data provided by davischallenge.org/soa_compare.html.
Quantitative Results

Table 1: IoU Comparison

<table>
<thead>
<tr>
<th>Method</th>
<th>Average IoU (DAVIS 2016)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCF-C3D</td>
<td>0.3479</td>
</tr>
<tr>
<td>FST</td>
<td>0.558</td>
</tr>
<tr>
<td>CUT</td>
<td>0.552</td>
</tr>
<tr>
<td>NLC</td>
<td>0.551</td>
</tr>
<tr>
<td>MSG</td>
<td>0.533</td>
</tr>
<tr>
<td>KEY</td>
<td>0.498</td>
</tr>
<tr>
<td>CVOS</td>
<td>0.482</td>
</tr>
</tbody>
</table>

Comparison data provided by davischallenge.org/soa_compare.html.
Observations

- Segmentation struggles with heavy background/camera movement, similar objects in background.

- Other models trained on different datasets, we trained only on DAVIS 2016 training subset (30 clips).

- In the ground truth images, background pixels greatly outnumber foreground pixels.

- May be decreasing learning rate too early.
Observations

- Segmentation struggles with heavy background/camera movement, similar objects in background.

- Other models trained on different datasets, we trained only on DAVIS 2016 training subset (30 clips).

- In the ground truth images, background pixels greatly outnumber foreground pixels.

- May be decreasing learning rate too early.
Observations

- Segmentation struggles with heavy background/camera movement, similar objects in background.

- Other models trained on different datasets, we trained only on DAVIS 2016 training subset (30 clips).

- In the ground truth images, background pixels greatly outnumber foreground pixels.

- May be decreasing learning rate too early.
Observations

- Segmentation struggles with heavy background/camera movement, similar objects in background.

- Other models trained on different datasets, we trained only on DAVIS 2016 training subset (30 clips).

- In the ground truth images, background pixels greatly outnumber foreground pixels.

- May be decreasing learning rate too early.
Upcoming Work
Plan for Next Week

- **Increasing weight values for foreground to balance loss function.**
  - One-Shot Video Object Segmentation by Caelles et al. (arXiv, April 2017)

- **Looking into other datasets to train the model on.**
  - SegTrack v2 (treated as binary classification)
  - DAVIS 2017 (treated as binary classification)

- **Generally improving the model.**
  - Decreasing the learning rate later at a later iteration during training.
Plan for Next Week

- Increasing weight values for foreground to balance loss function.
  - One-Shot Video Object Segmentation by Caelles et al. (arXiv, April 2017)

- Looking into other datasets to train the model on.
  - SegTrack v2 (treated as binary classification)
  - DAVIS 2017 (treated as binary classification)

- Generally improving the model.
  - Decreasing the learning rate later at a later iteration during training.
Plan for Next Week

- Increasing weight values for foreground to balance loss function.
  - One-Shot Video Object Segmentation by Caelles et al. (arXiv, April 2017)

- Looking into other datasets to train the model on.
  - SegTrack v2 (treated as binary classification)
  - DAVIS 2017 (treated as binary classification)

- Generally improving the model.
  - Decreasing the learning rate later at a later iteration during training.
Plan for Next Week

- Increasing weight values for foreground to balance loss function.
  - One-Shot Video Object Segmentation by Caelles et al. (arXiv, April 2017)

- Looking into other datasets to train the model on.
  - SegTrack v2 (treated as binary classification)
  - DAVIS 2017 (treated as binary classification)

- Generally improving the model.
  - Decreasing the learning rate later at a later iteration during training.
Plan for Next Week

- Increasing weight values for foreground to balance loss function.
  - One-Shot Video Object Segmentation by Caelles et al. (arXiv, April 2017)

- Looking into other datasets to train the model on.
  - SegTrack v2 (treated as binary classification)
  - DAVIS 2017 (treated as binary classification)

- Generally improving the model.
  - Decreasing the learning rate later at a later iteration during training.
Plan for Next Week

- Increasing weight values for foreground to balance loss function.
  - One-Shot Video Object Segmentation by Caelles et al. (arXiv, April 2017)

- Looking into other datasets to train the model on.
  - SegTrack v2 (treated as binary classification)
  - DAVIS 2017 (treated as binary classification)

- Generally improving the model.
  - Decreasing the learning rate later at a later iteration during training.
Plan for Next Week

- Increasing weight values for foreground to balance loss function.
  - One-Shot Video Object Segmentation by Caelles et al. (arXiv, April 2017)

- Looking into other datasets to train the model on.
  - SegTrack v2 (treated as binary classification)
  - DAVIS 2017 (treated as binary classification)

- Generally improving the model.
  - Decreasing the learning rate later at a later iteration during training.
Questions?