# 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

# Table of Contents

1. Previous Work

2. Current Work

3. Upcoming Work

# Previous Work

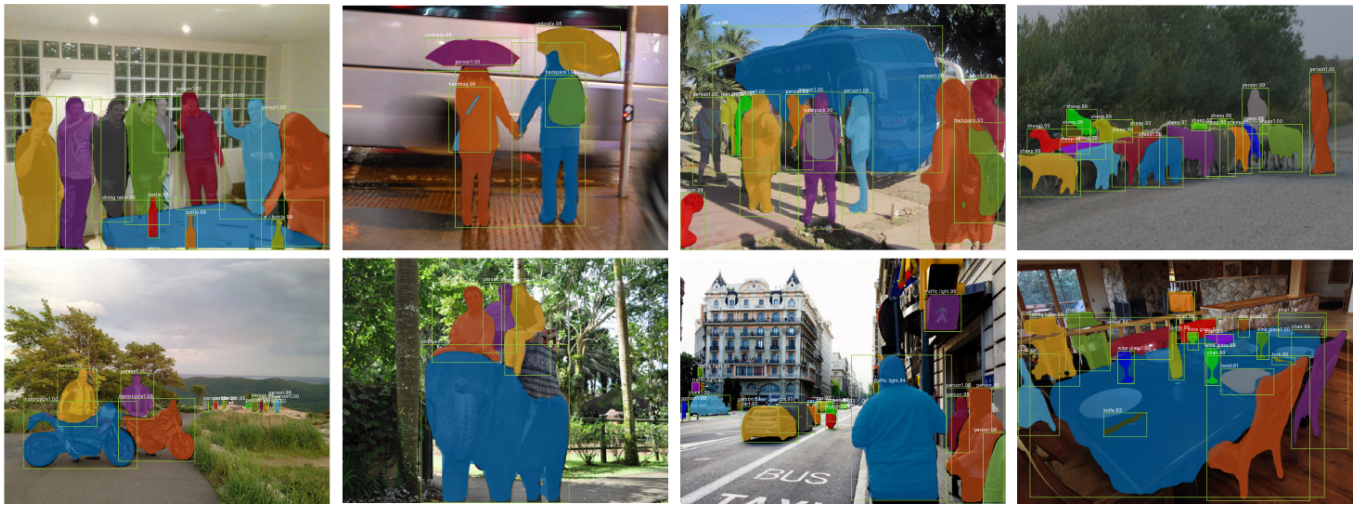
# DAVIS 2017 Issues





# Literature Review

- **Mask CNN** by He et al. (arXiv, April 2017)



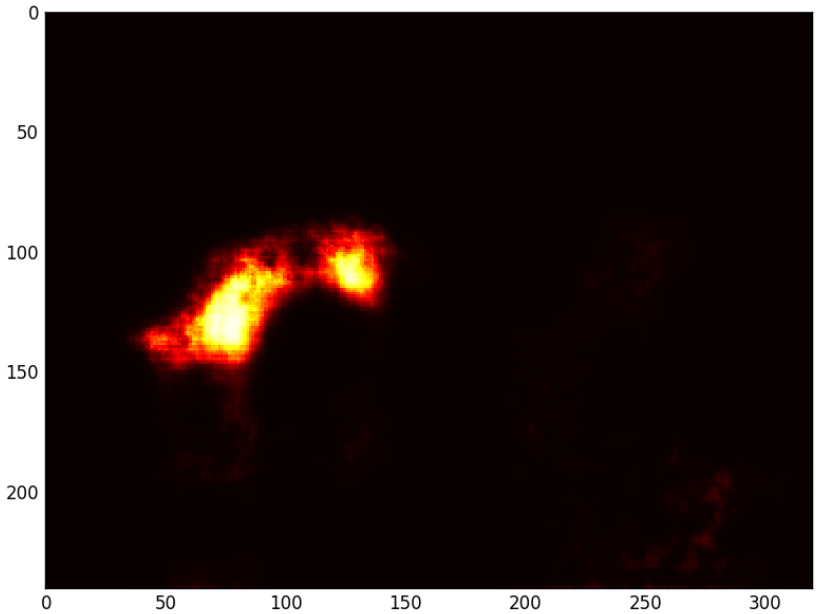
# DAVIS Mask Change





# DAVIS 2016 Initial Results





# Current Work

# Goals

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

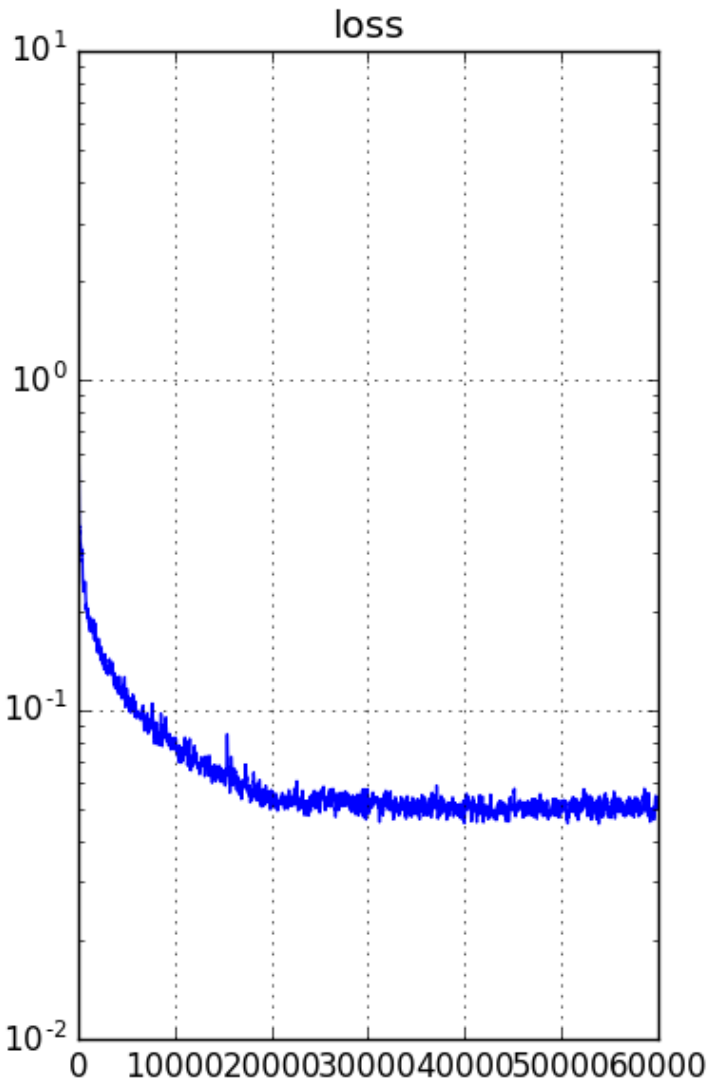
- Reading related material as needed.

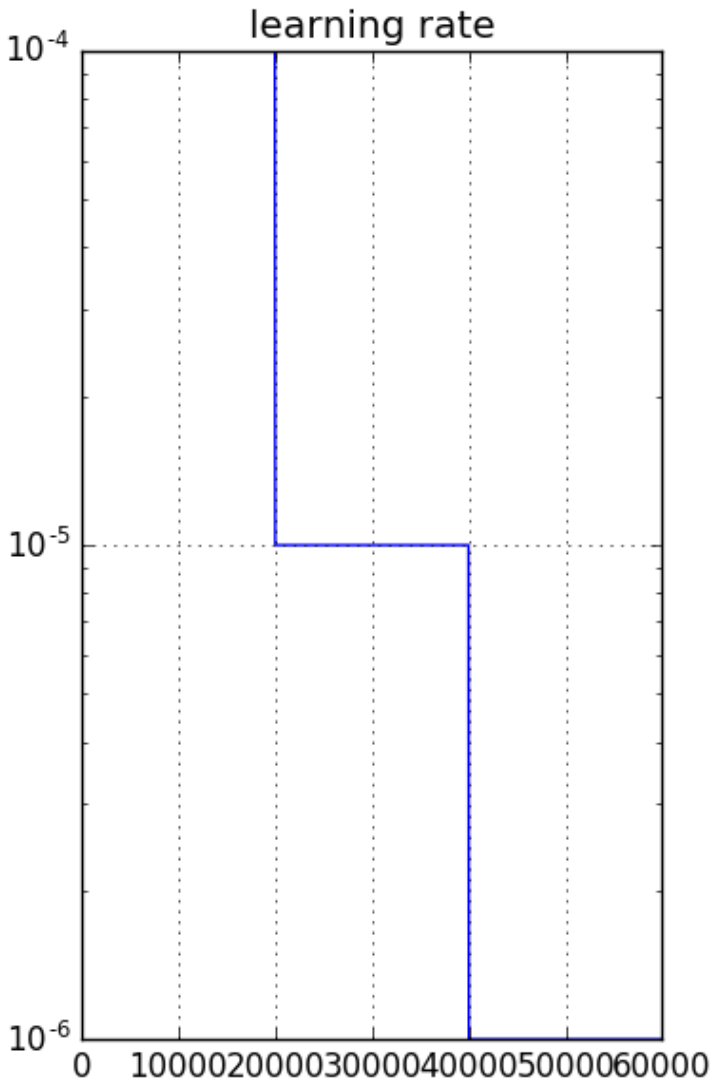
# Goals

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

- Reading related material as needed.

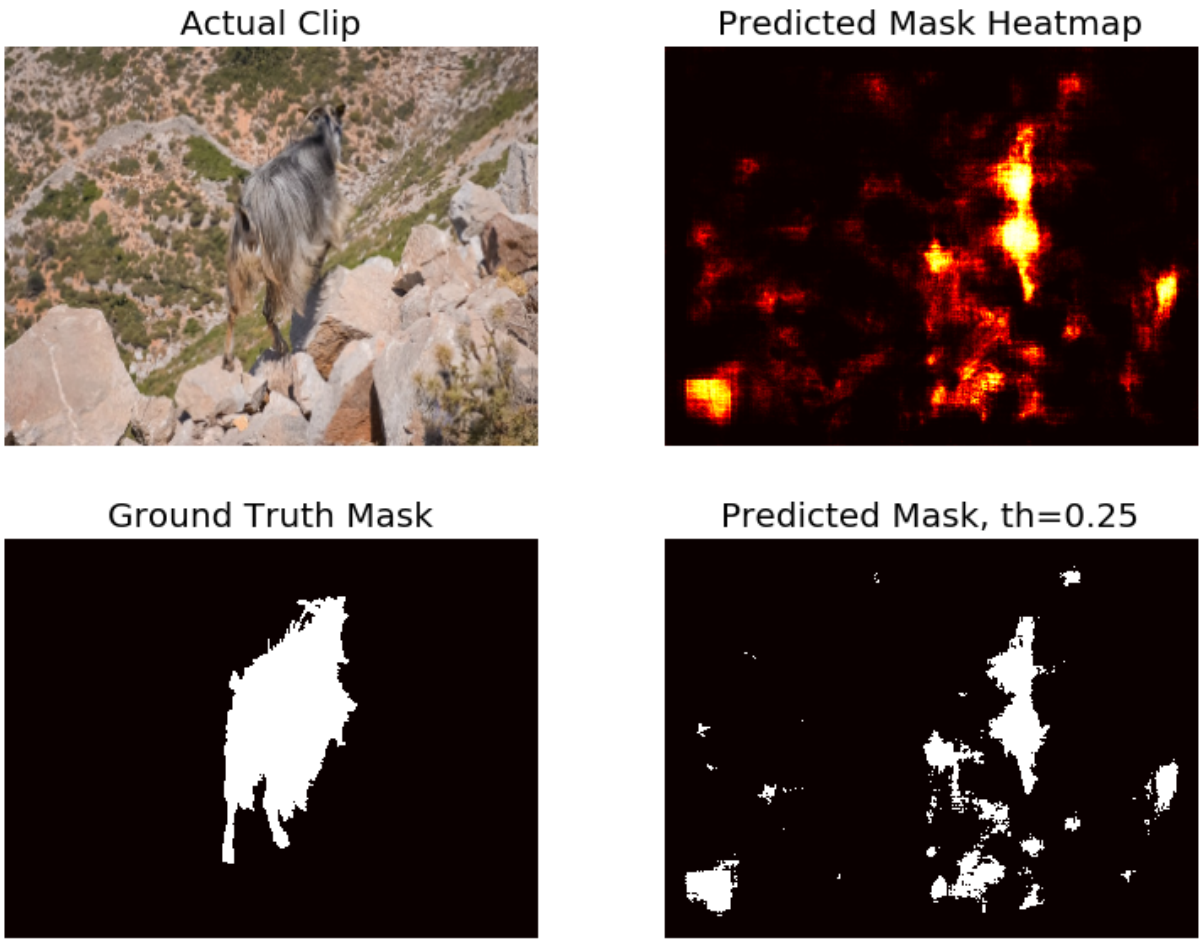
# Loss Curve


loss
$10^{1}$
$10^{0}$
$10^{-1}$
$10^{-2}$
0
10000
20000
30000
40000
50000
60000


learning rate
$10^{-4}$
$10^{-5}$
$10^{-6}$
0
10000
20000
30000
40000
50000
60000

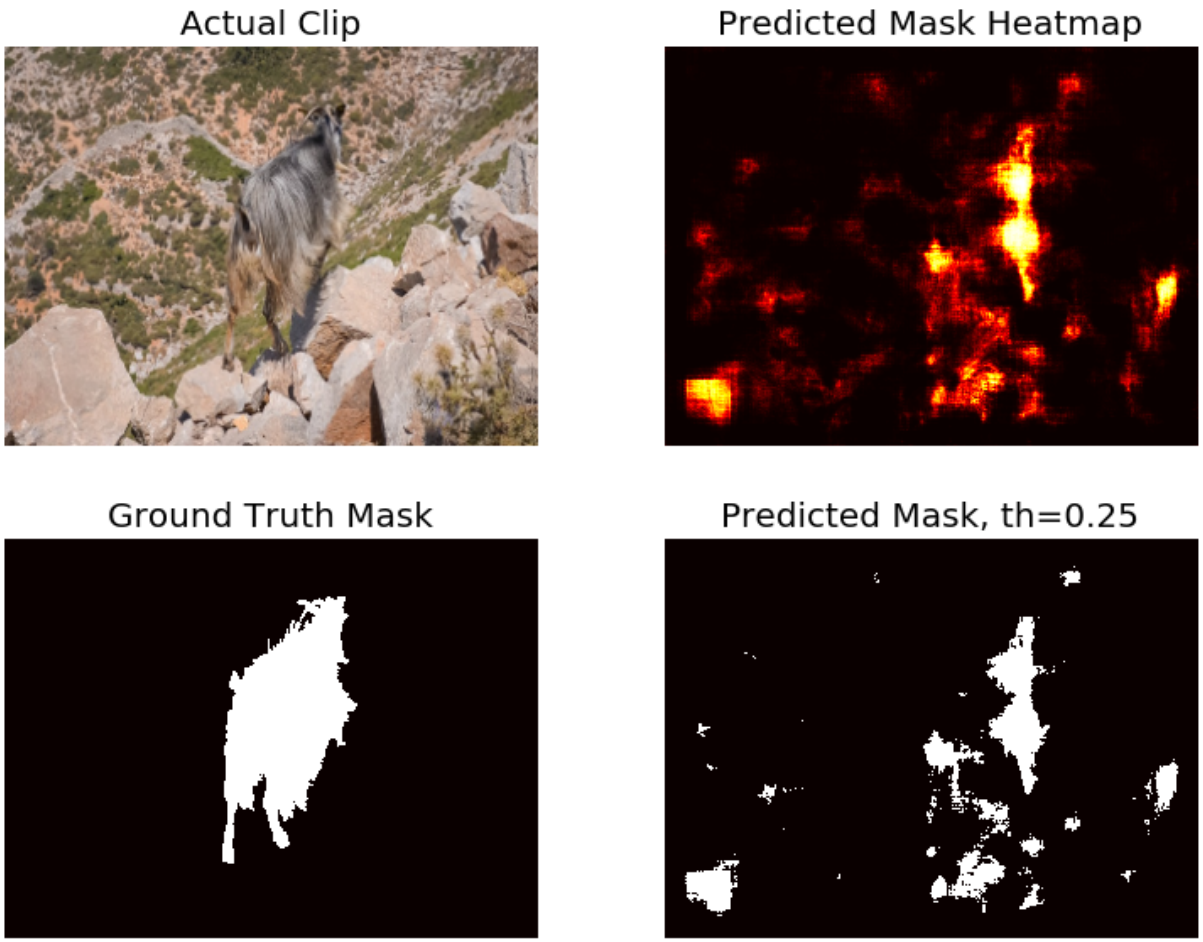
# Results (1)

**Figure 1:** IoU: 0.3605



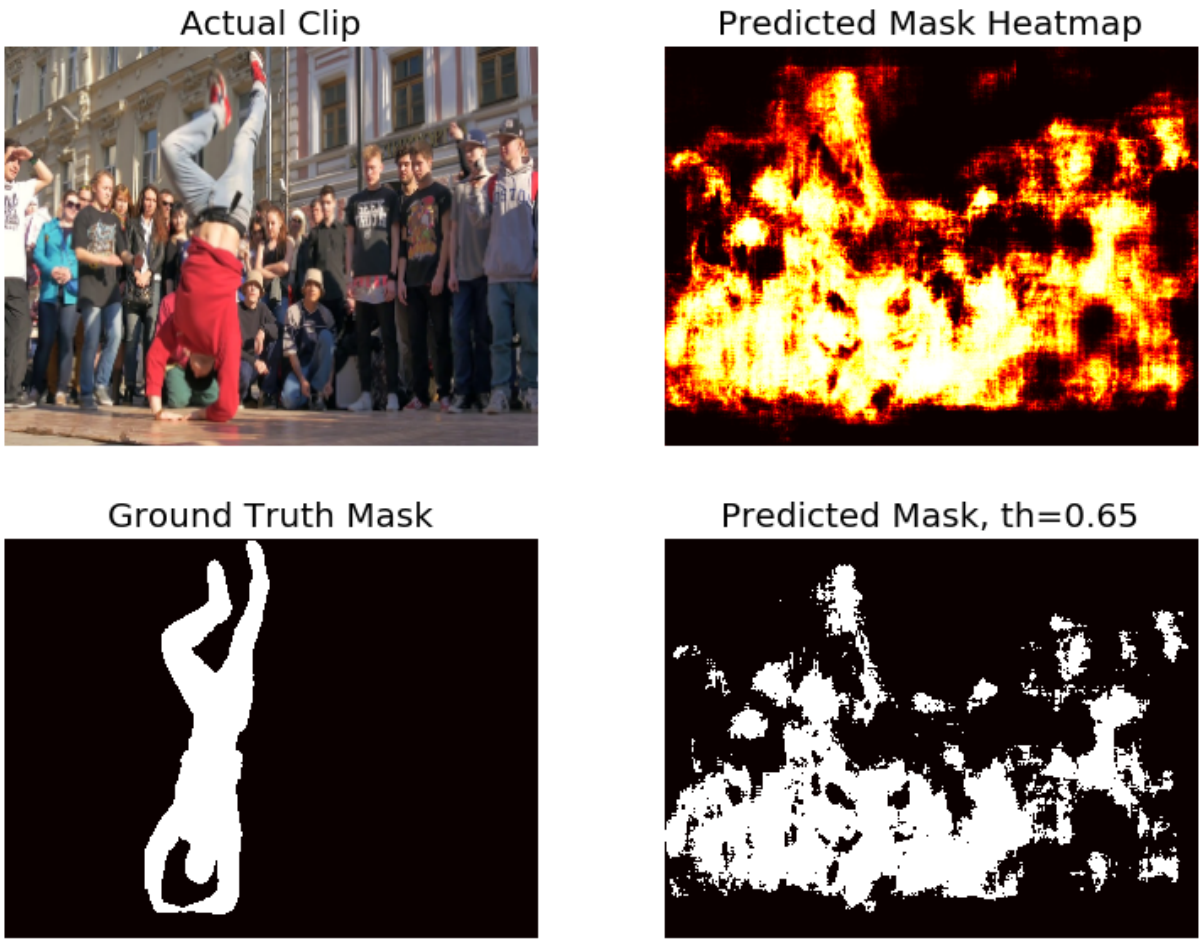
# Results (1)

**Figure 1:** IoU: 0.3605



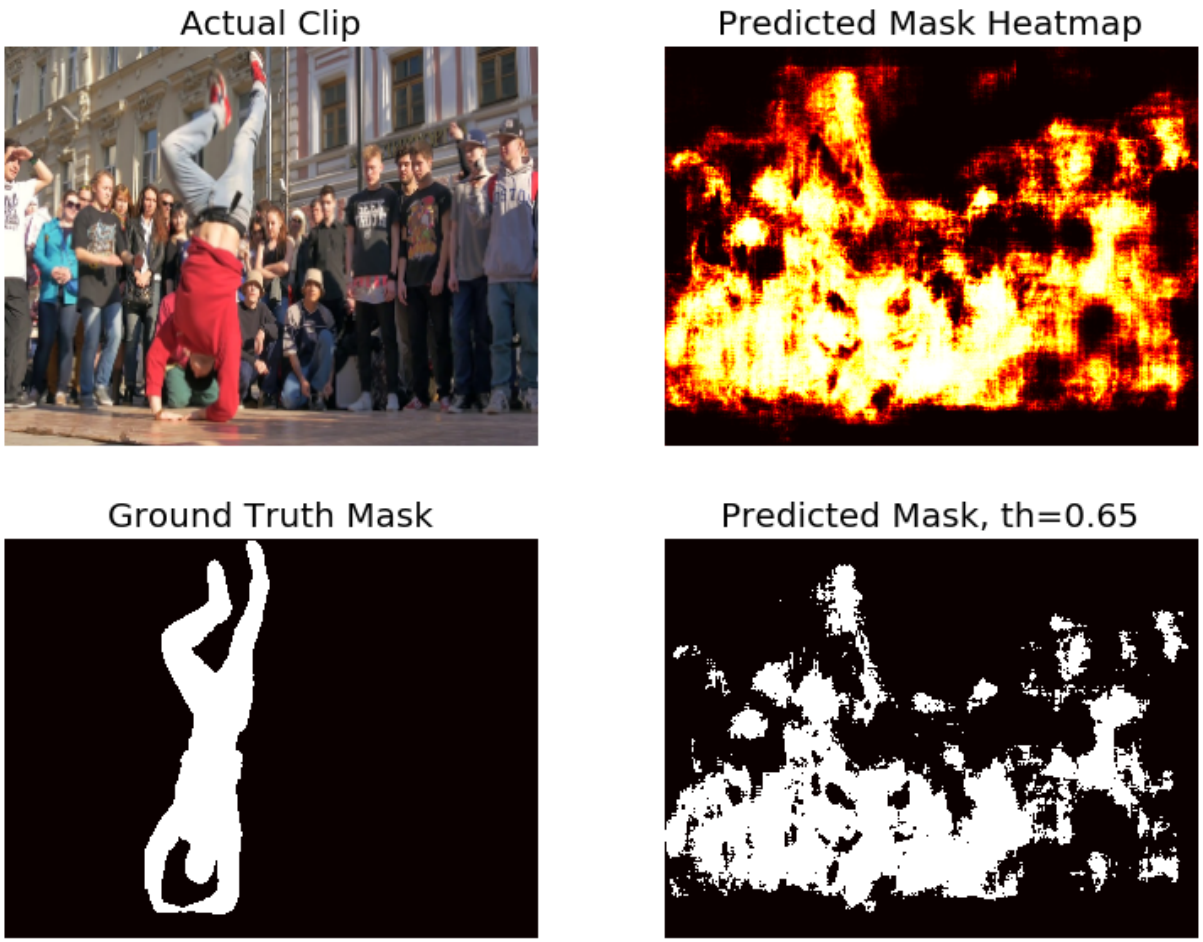
# Results (2)

**Figure 2:** IoU: 0.2132



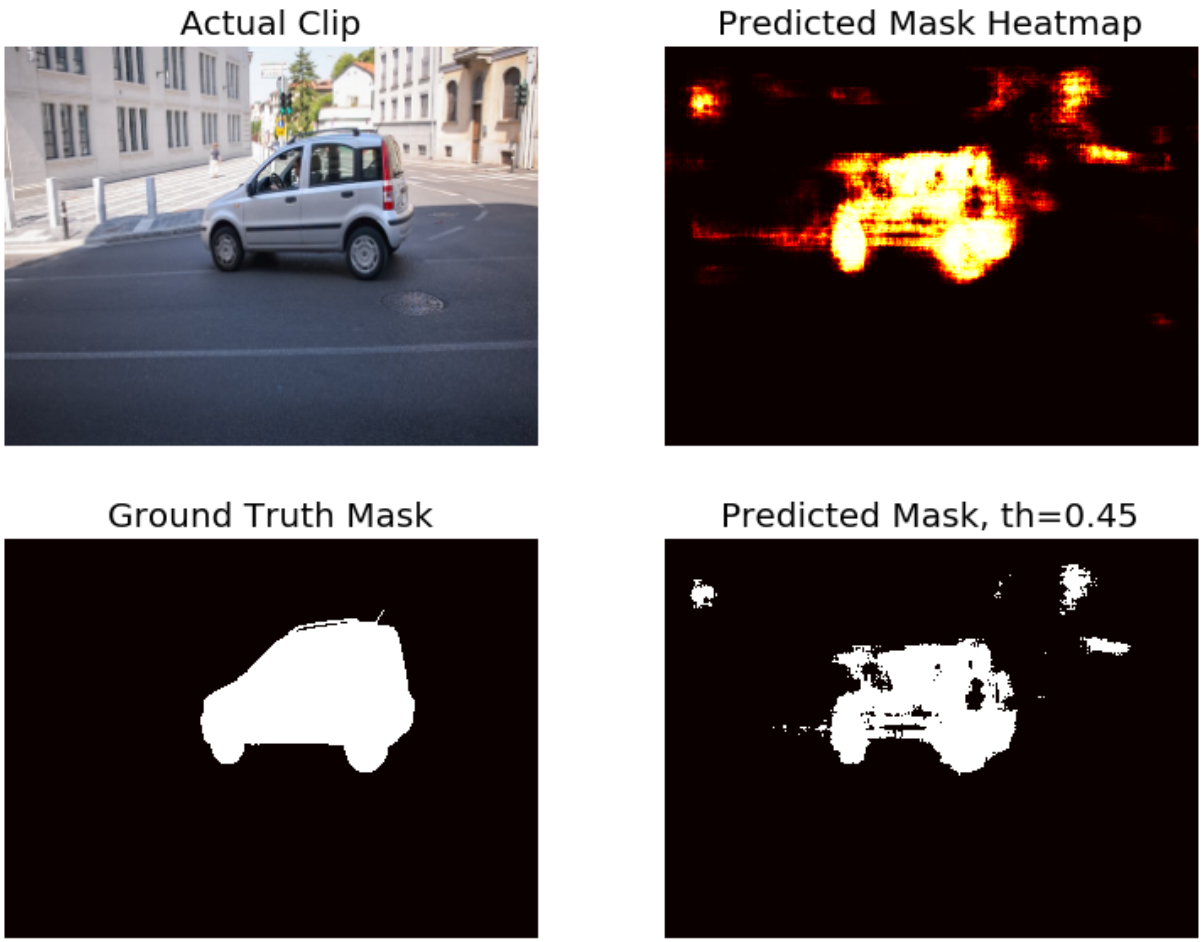
# Results (2)

**Figure 2:** IoU: 0.2132



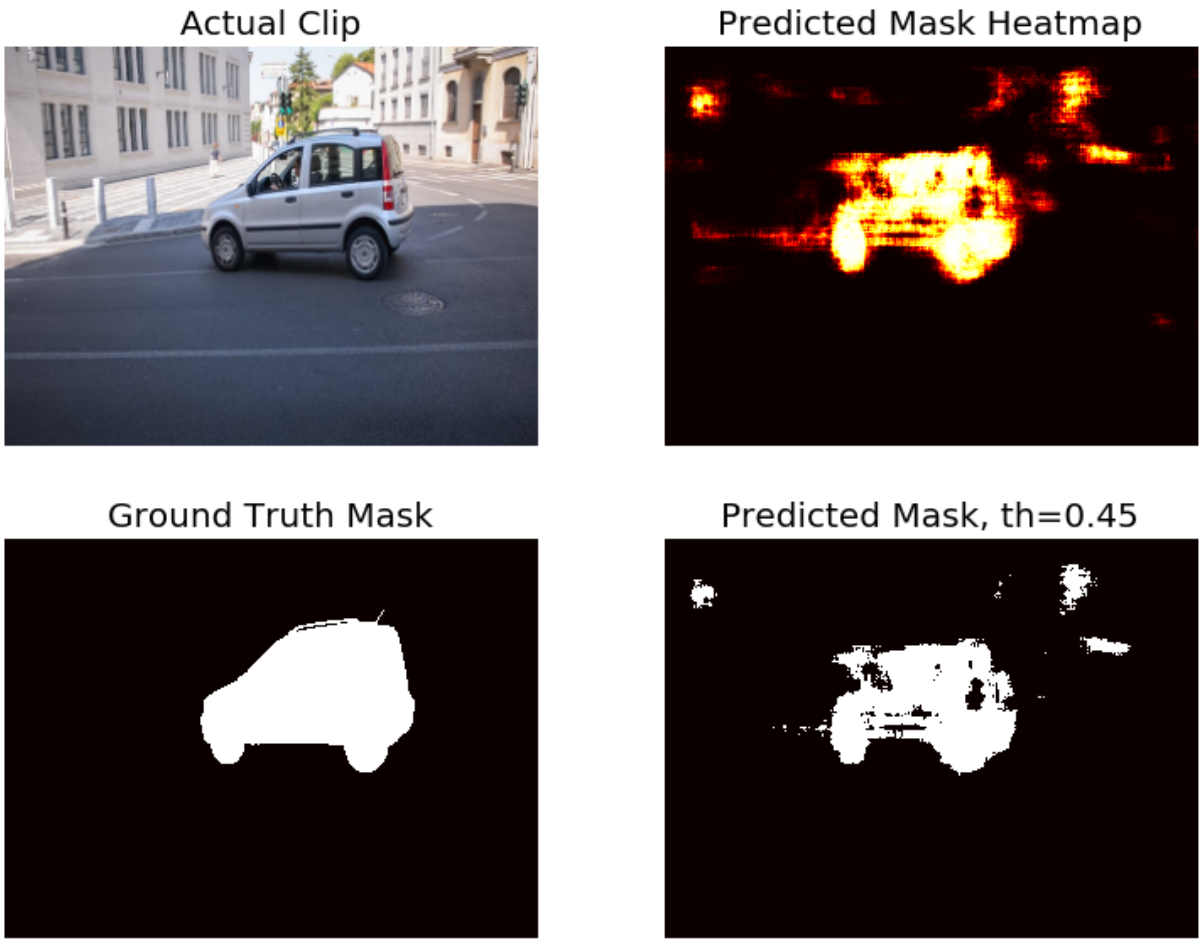
# Results (3)

**Figure 3:** IoU: 0.3944



# Results (3)

**Figure 3:** IoU: 0.3944



# Quantitative Results

**Table 1:** IoU Comparison

| Method | Average IoU (DAVIS 2016) |
|---|---|
| UCF-C3D | 0.3479 |
| FST | 0.558 |
| CUT | 0.552 |
| NLC | 0.551 |
| MSG | 0.533 |
| KEY | 0.498 |
| CVOS | 0.482 |

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

# Quantitative Results

**Table 1:** IoU Comparison

| Method | Average IoU (DAVIS 2016) |
|---|---|
| UCF-C3D | 0.3479 |
| FST | 0.558 |
| CUT | 0.552 |
| NLC | 0.551 |
| MSG | 0.533 |
| KEY | 0.498 |
| CVOS | 0.482 |

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?