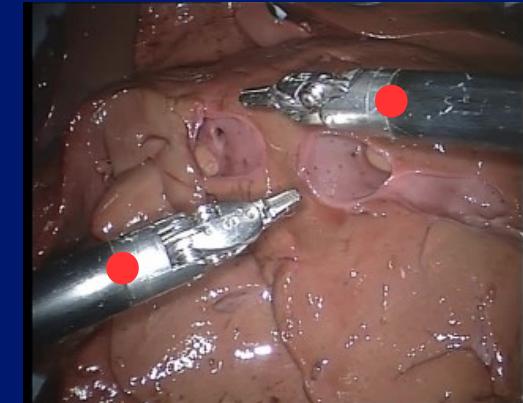


# Image-based segmentation and localization of surgical instruments using deep neural networks

Amelie Wagner  
bachelor thesis presentation

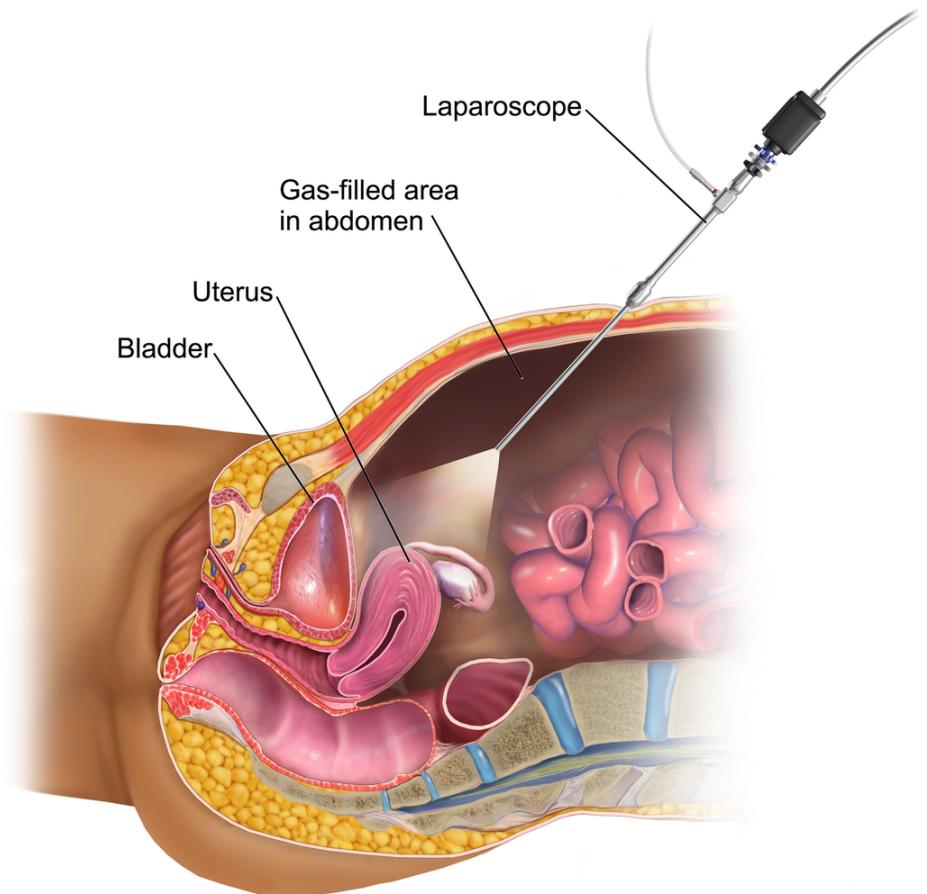


NATIONAL CENTER  
FOR TUMOR DISEASES  
PARTNER SITE DRESDEN  
UNIVERSITY CANCER CENTER UCC

Supported by:  
German Cancer Research Center  
University Hospital Carl Gustav Carus Dresden  
Carl Gustav Carus Faculty of Medicine, TU Dresden  
Helmholtz-Zentrum Dresden-Rossendorf

# Motivation

- Minimally Invasive Surgery (MIS)



[4]

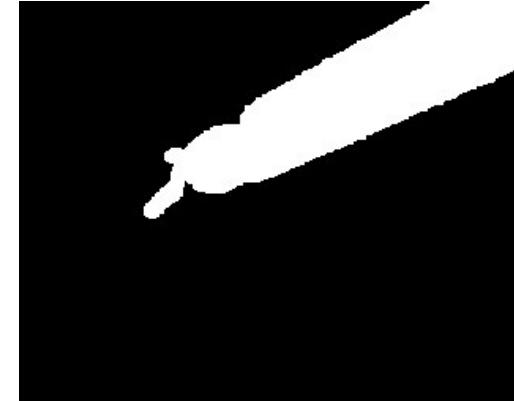
- Laparoscopic Instruments



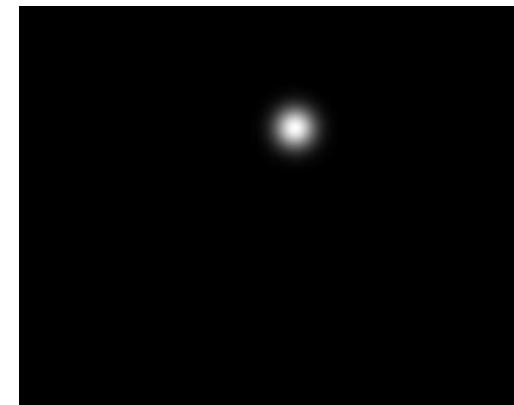
[5]

# Overview

Segmentation

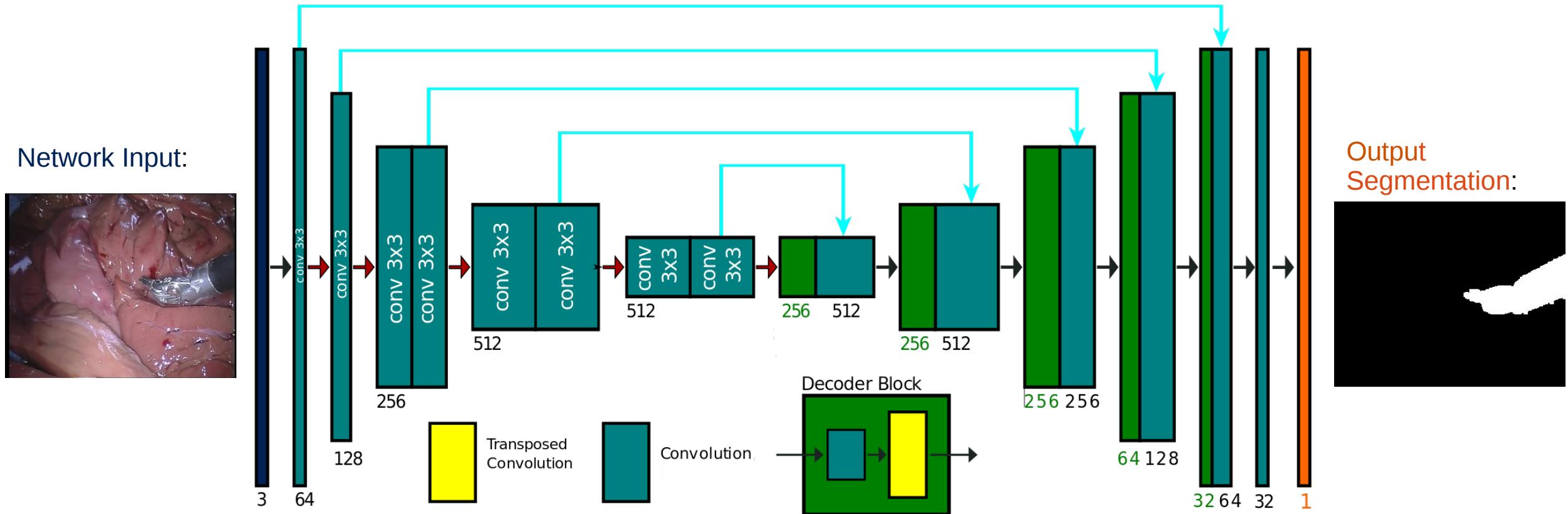


Localization

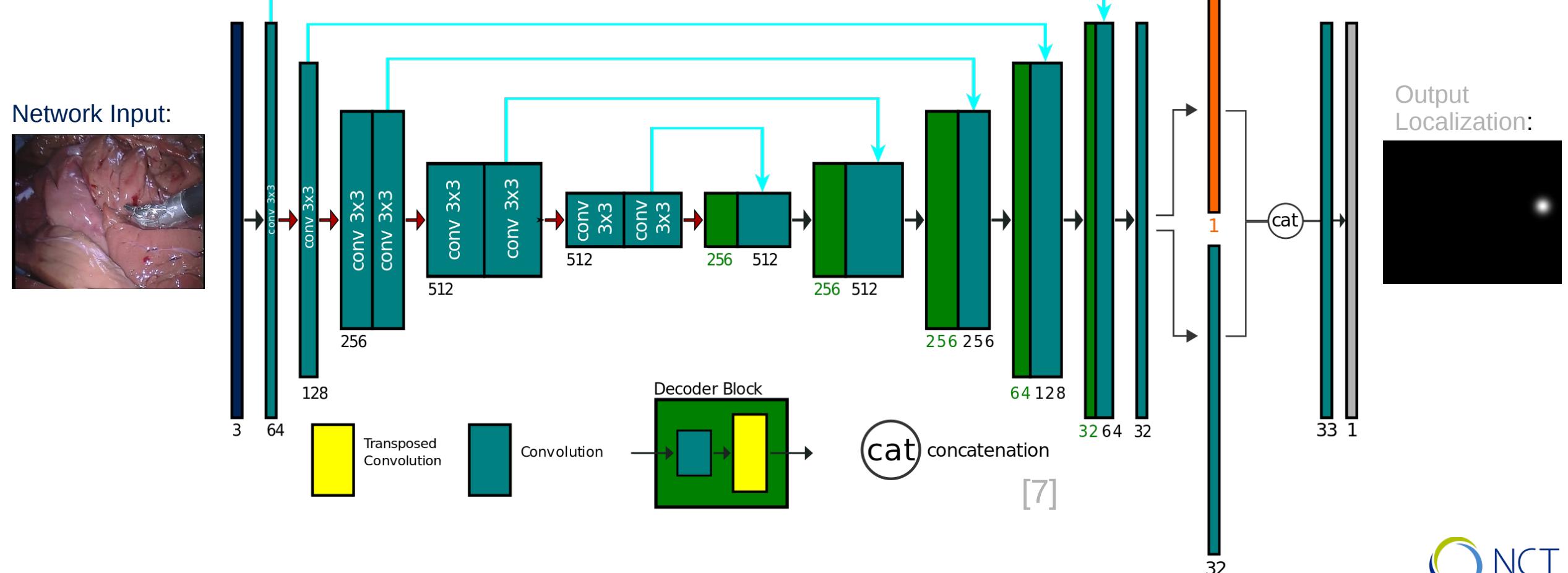


# Methods

# Method: Segmentation Network TernausNet-11 [1]



# Method: Localization Network [7]



# Concurrent Segmentation and Localization

Segmentation + Localization: Combined Loss

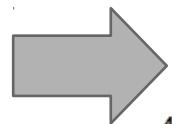
$$L(y_{seg}, y_{loc}, t_{seg}, t_{loc}) = (1 - \gamma) \cdot L_{seg}(y_{seg}, t_{seg}) + \gamma \cdot L_{loc}(y_{loc}, t_{loc})$$

$$L_{seg}(y_{seg}, t_{seg}) = BCE(y_{seg}, t_{seg}) - \log J_{binary}(y_{seg}, t_{seg})$$

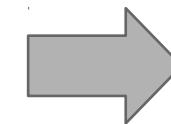
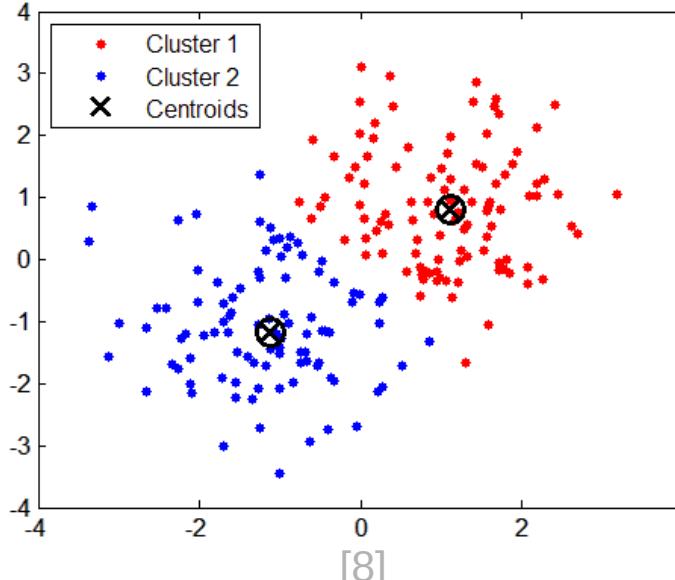
$$L_{loc}(y_{loc}, t_{loc}) = MSE(y_{loc}, t_{loc})$$

# Method: Localization Network

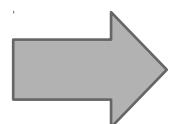
Output Localization:



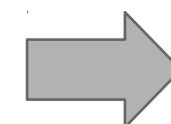
weighted k-means,  $k=1$



$$\begin{pmatrix} x \\ y \end{pmatrix}$$



weighted k-means,  $k=2$

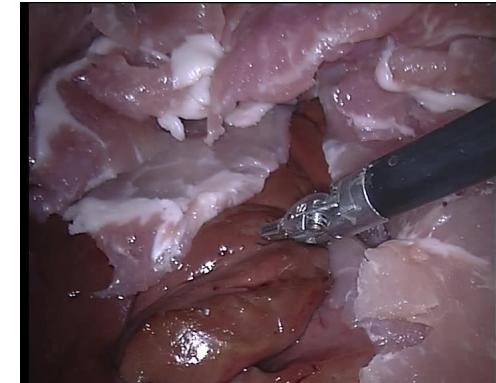
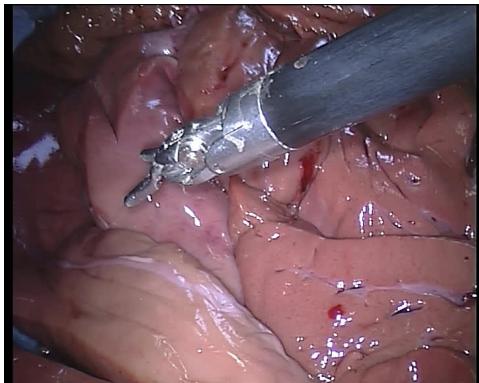


$$\begin{pmatrix} x \\ y \end{pmatrix} \quad \begin{pmatrix} x \\ y \end{pmatrix}$$

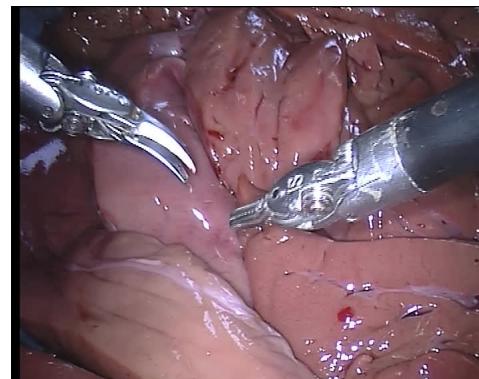
# Data

# Data: Endoscopic Vision Challenge 2015 [2]

4 Training Datasets

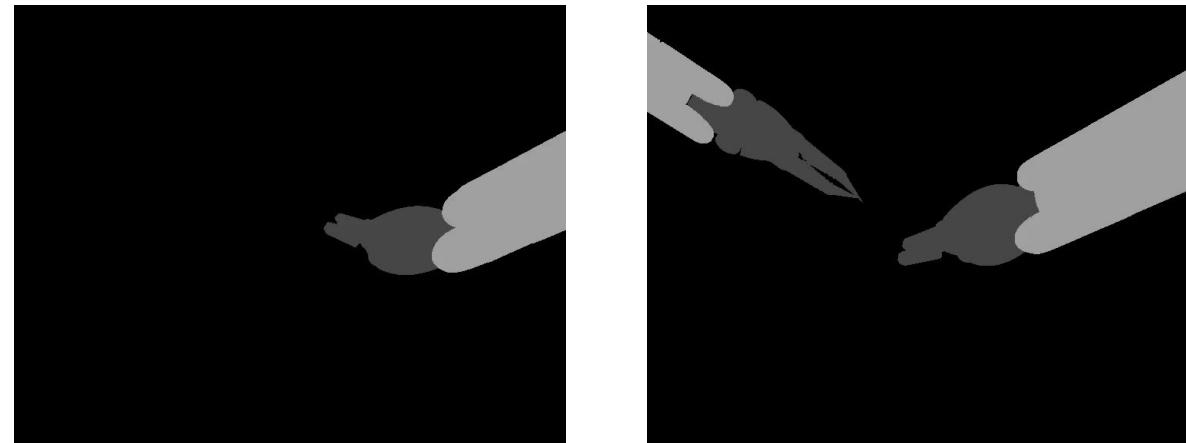


2 Testing Datasets



# Data: Endoscopic Vision Challenge 2015 [2]

Ground Truth Segmentation



Ground Truth Localization



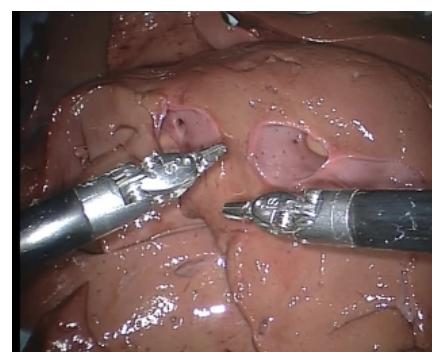
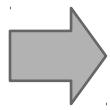
# Data Preparation

# Data Preparation: EndoVis15 [2]

Network Input

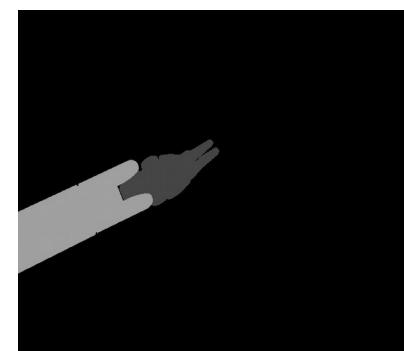
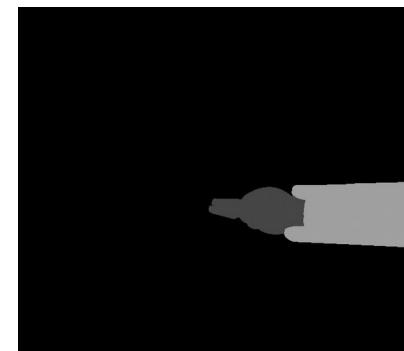


720×576 px

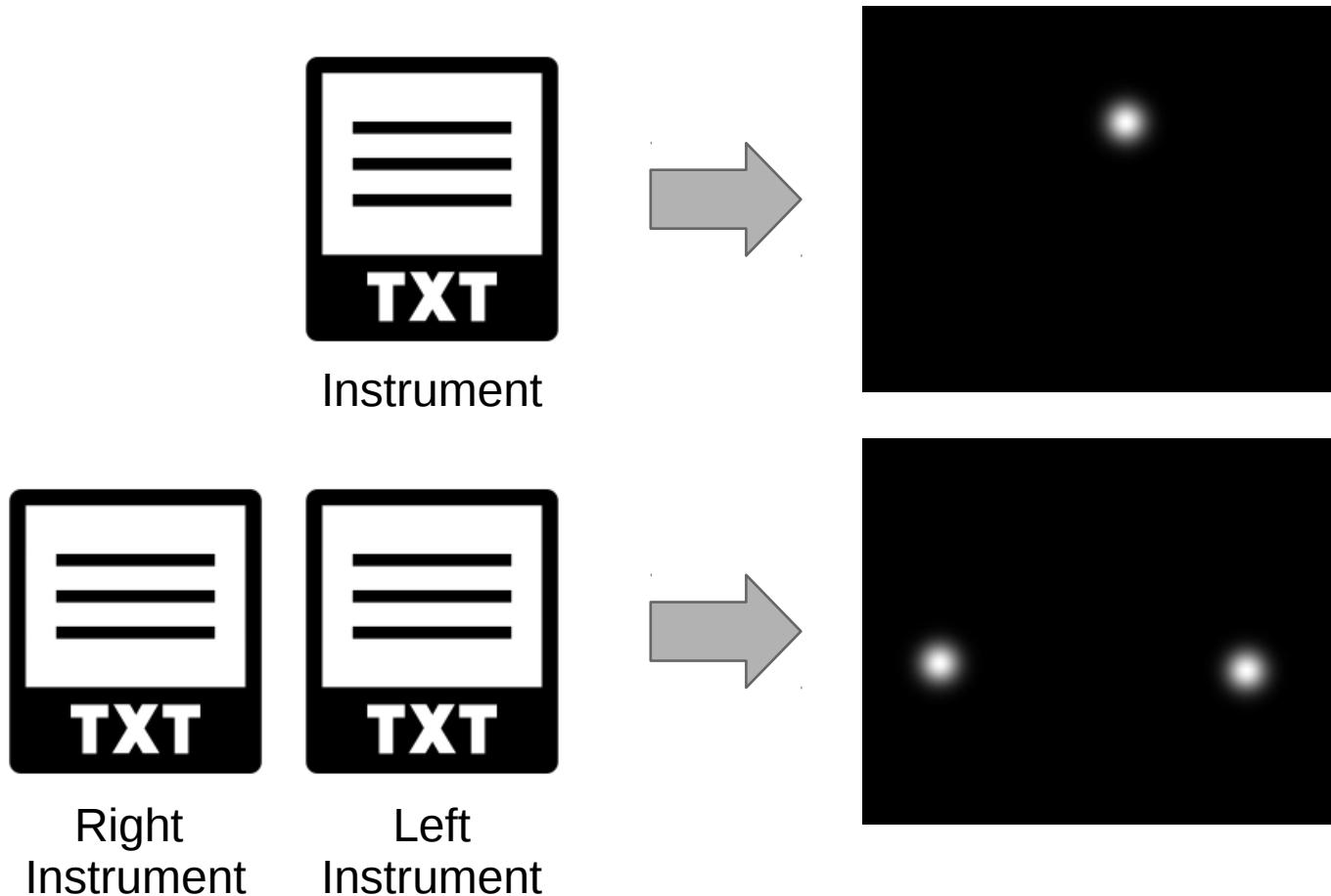


320×265 px

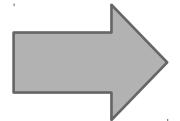
Ground Truth Segmentation



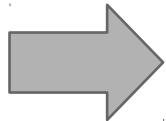
# Data Preparation: Localization EndoVis15 [2]



# Results



cross-validation



testsets

# Results Cross Validation

	Dataset1 left/right instr.	Dataset2	Dataset3	Dataset4	mean dist.
Laina et al.	39.0/30.8	9.7	10.9	13.0	20.7
LocNet	17.58/15.15	11.36	10.83	12.05	13.85

↓      ↓      ↓      ↓

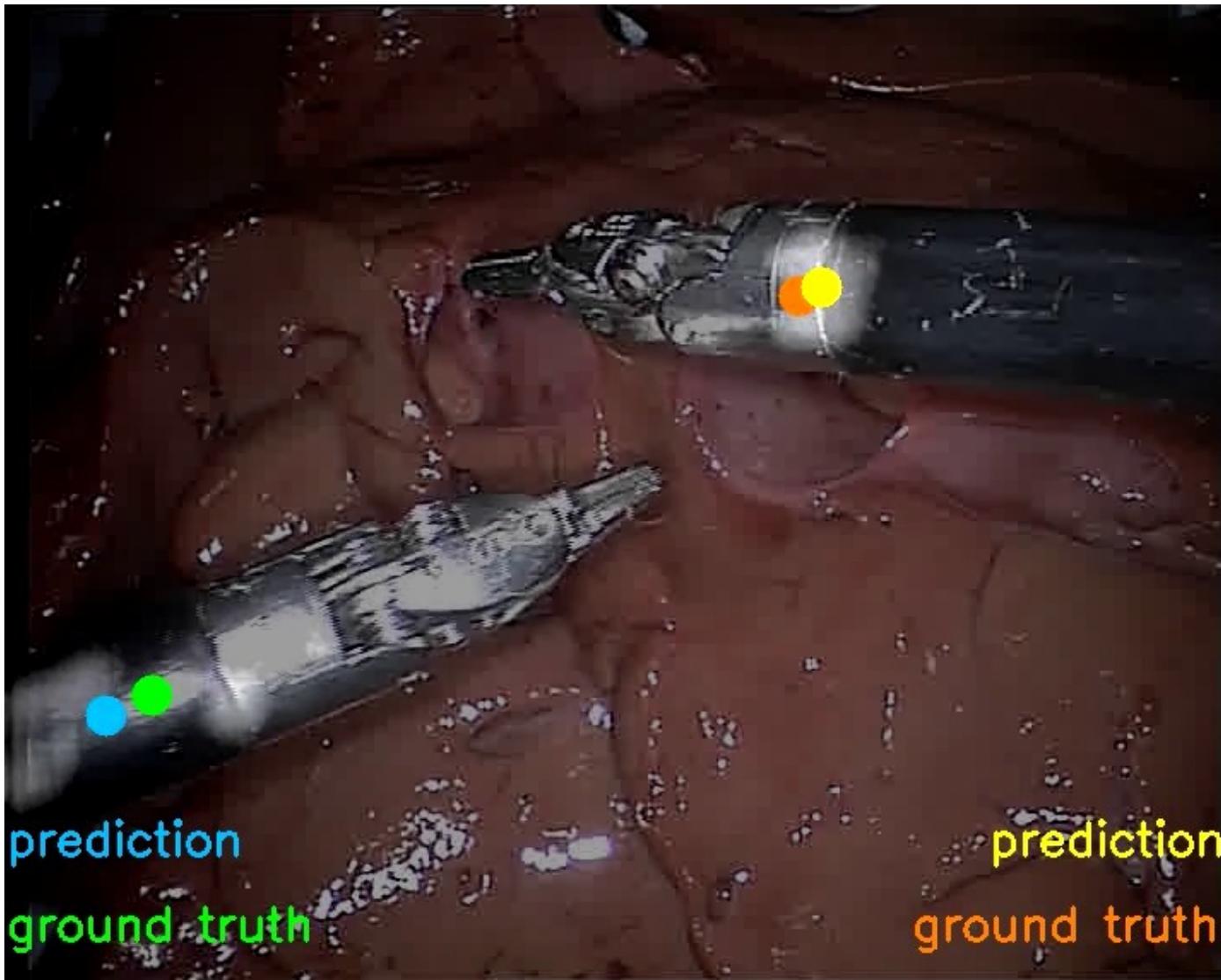


epochs      val.method       $\gamma$

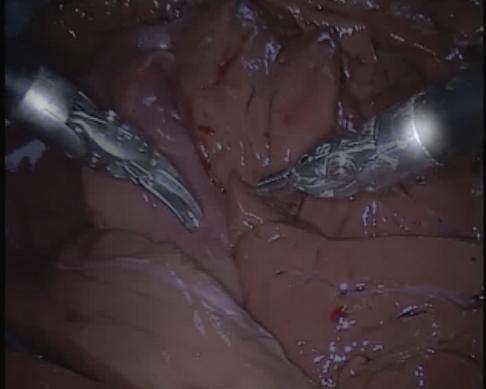
---

LocNet	50	crossval.	1/2
--------	----	-----------	-----

# Results Cross Validation



# Results Testsets

	Dataset5 left/right instr.	Dataset6 left/right instr.	mean dist.
Laina et al.	38.4/60.0	36.4/63.9	49.68
LocNet	89.80/117.91	88.17/120.02	106.4
	 ↓	 ↓	
	epochs	val.method	$\gamma$
LocNet	500	testset	1/2

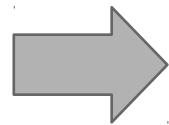
## Evaluation Conditions

	epochs	val.method	$\gamma$
LOC-01	600	crossval.	1/2
LOC-02	600	crossval.	2/3
LOC-03	50	crossval.	1/2
LOC-04	500	testset	1/2
LOC-05	100	testset	1/2
LOC-06	400	crossval.	1
LOC-07	100	testset	1
LOC-08	500	testset	1

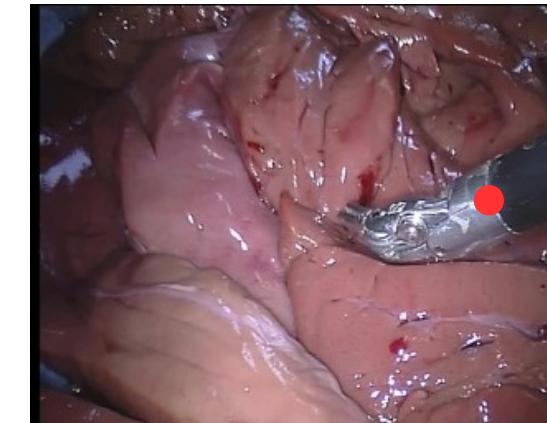
# Conclusion



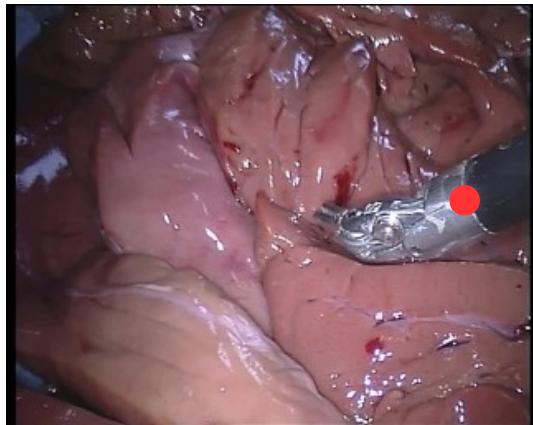
CNN



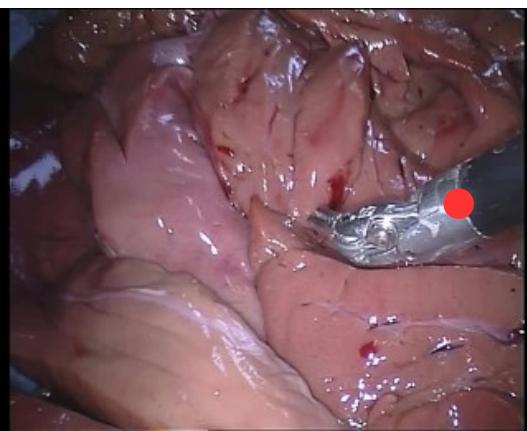
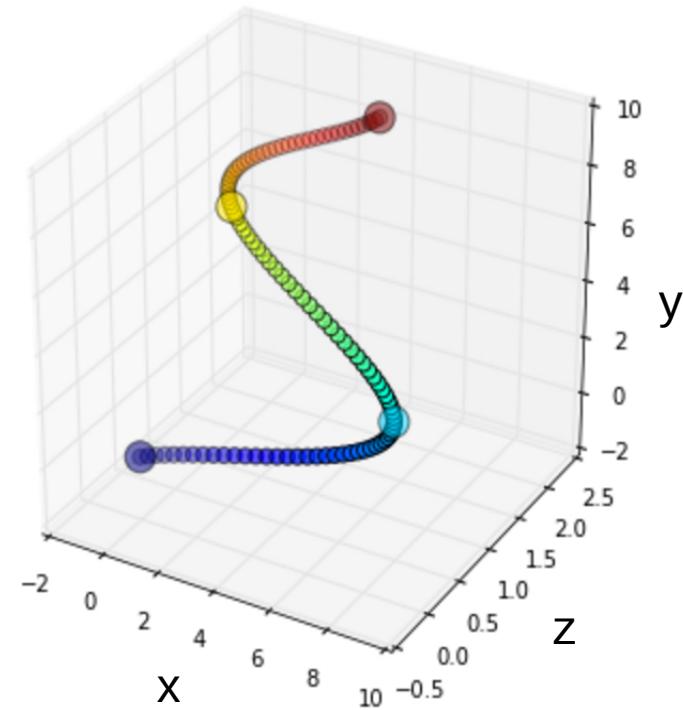
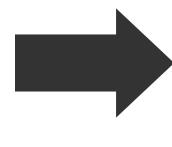
k-means



# Future Work



$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$



[9]

Thank you for your  
attention.



[6]

# Sources

- [1] A. Shvets, A. Rakhlin, A. A. Kalinin, and V. Iglovikov, "Automatic Instrument Segmentation in Robot-Assisted Surgery Using Deep Learning"
- [2] "MICCAI Endoscopic Vision Challenge: Subchallenge Instrument Segmentation and Tracking," <https://endovissub-instrument.grand-challenge.org/>, accessed: 2018-09-15
- [3] "MICCAI Endoscopic Vision Challenge: Subchallenge Robotic Instrument Segmentation," <https://endovissub2017-roboticinstrumentsegmentation.grand-challenge.org>, accessed: 2018-09-15
- [4] S. Blausen, "Medical gallery of Blausen Medical 2014," WikiJournal of Medicine, 2014
- [5] D. Bouget, M. Allan, D. Stoyanov, and P. Jannin, "Vision-based and marker-less surgical tool detection and tracking: a review of the literature," Medical Image Analysis, 2017
- [6] "#1 data scientist excuse slacking off", <https://me.me/i/the-1-data-scientist-excuse-for-legitimately-slacking-off-my-2aa76556da934f79b5818a02e7f23cf9>, accessed: 2018-11-20
- [7] I. Laina, N. Rieke, C. Rupprecht, J. P. Vizcaíno, A. Eslami, F. Tombari, and N. Navab, "Concurrent segmentation and localization for tracking of surgical instruments," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2017
- [8] "k-means example",  
[http://mines.humanoriented.com/classes/2010/fall/csci568/portfolio\\_exports/mvoget/cluster/kmeans\\_diagram.png](http://mines.humanoriented.com/classes/2010/fall/csci568/portfolio_exports/mvoget/cluster/kmeans_diagram.png), accessed: 2018-11-22
- [9]"3d trajectory" <https://mikeroberts3000.github.io/flashlight/>. accessed: 2018-11-23

# Method: Localization Network

$$L_{seg}(y_{seg}, t_{seg}) = BCE(y_{seg}, t_{seg}) - \log J_{binary}(y_{seg}, t_{seg})$$

$$BCE_{sigm}(y_{w,b}(x_j), t_j) = t_j \log(\sigma(y_{w,b}(x_j))) + (1 - t_j) \log(1 - \sigma(y_{w,b}(x_j)))$$

$$J_{binary}(y_{w,b}(x_j), t_j) = \frac{t_j \cdot y_{w,b}(x_j)}{t_j + y_{w,b}(x_j) - t_j \cdot y_{w,b}(x_j)}$$

$$L_{loc}(y_{loc}, t_{loc}) = MSE(y_{loc}, t_{loc}) = (t_{loc} - y_{loc})^2$$

$$MSE(y_{w,b}(x_j), t_j) = (t_j - y_{w,b}(x_j))^2$$

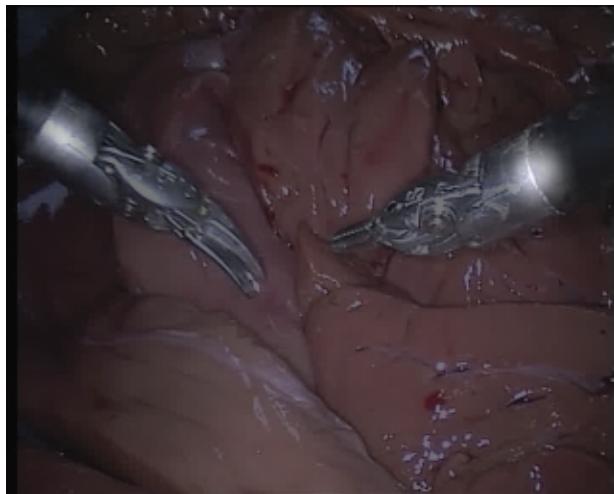
# Results Cross Validation

	Dataset1 left/right instr.	Dataset2	Dataset3	Dataset4	mean dist.
LOC-01	$21.24 \pm 12.5 / 17 \pm 10.4$	$13.45 \pm 5.9$	$12.73 \pm 5.7$	$14.22 \pm 23.5$	$16.27 \pm 19.6$
LOC-02	$22.65 \pm 31.1 / 15.5 \pm 22.5$	-	-	-	-
LOC-03	$17.58 \pm 12.1 / 15.15 \pm 9.4$	$11.36 \pm 6.4$	$10.83 \pm 6.3$	$12.05 \pm 8.9$	$13.85 \pm 10.0$
LOC-06	$372.24 \pm 91.7 / 299.62 \pm 251.1$	$27.66 \pm 109.9$	$20.47 \pm 82.13$	$21.48 \pm 80.2$	$152.5 \pm 207.9$



# Results Testsets

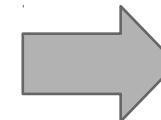
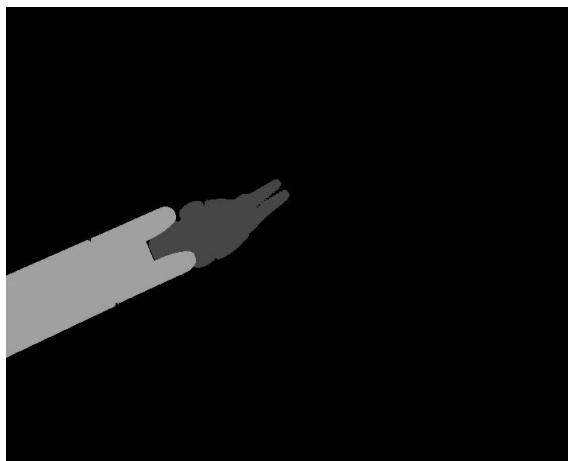
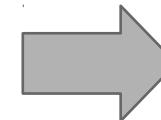
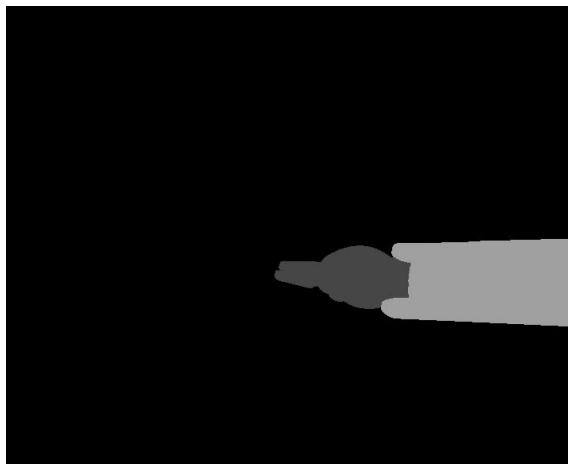
	Dataset5 left/right instr.	Dataset6 left/right instr.	mean dist.
LOC-04	$89.80 \pm 101.1 / 117.91 \pm 69.0$	$88.17 \pm 91.2 / 120.02 \pm 68.2$	$106.4 \pm 80.4$
LOC-05	$69.59 \pm 51.6 / 118.15 \pm 73.2$	$100.15 \pm 118.8 / 120.32 \pm 70.0$	$111.7 \pm 94.5$
LOC-07	$114.13 \pm 136.8 / 121.51 \pm 78.8$	$215.36 \pm 225 / 122.23 \pm 73.5$	$162.04 \pm 163.9$
LOC-08	$121.5 \pm 146.5 / 122.43 \pm 80.9$	$231.56 \pm 229.3 / 123.5 \pm 77.5$	$169.7 \pm 169.7$



# Data Preparation: Segmentation EndoVis15 [2]

## Ground Truth

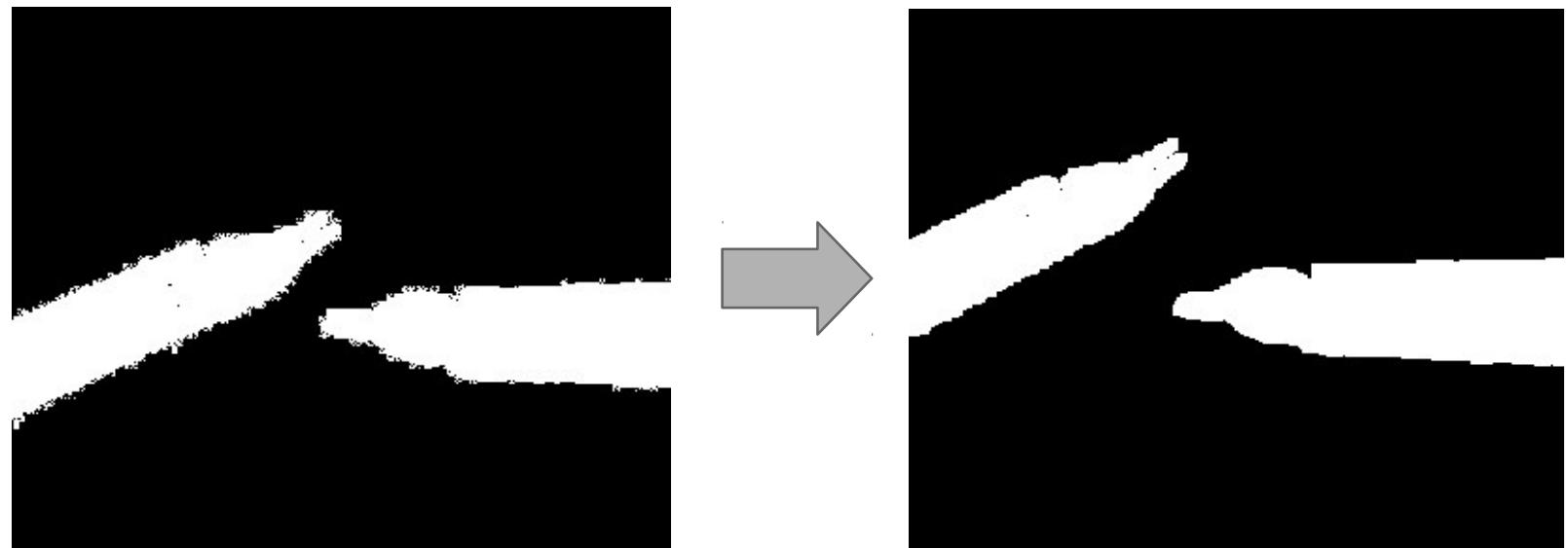
- Binary Mask



# Data Preparation: Segmentation EndoVis15 [2]

## Ground Truth

- Morphology Opening
- Downsampling

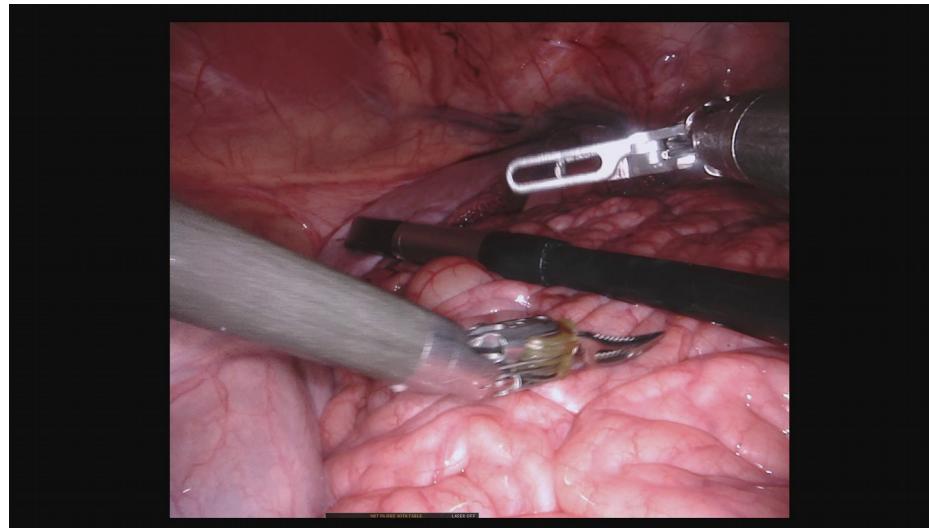


320×265 px

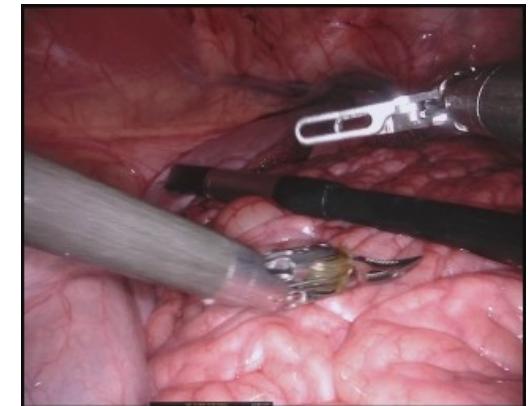
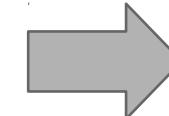
# Data Preparation: Segmentation EndoVis17 [3]

## Input Images

- Cropping
- Downsampling



1280×1024 px

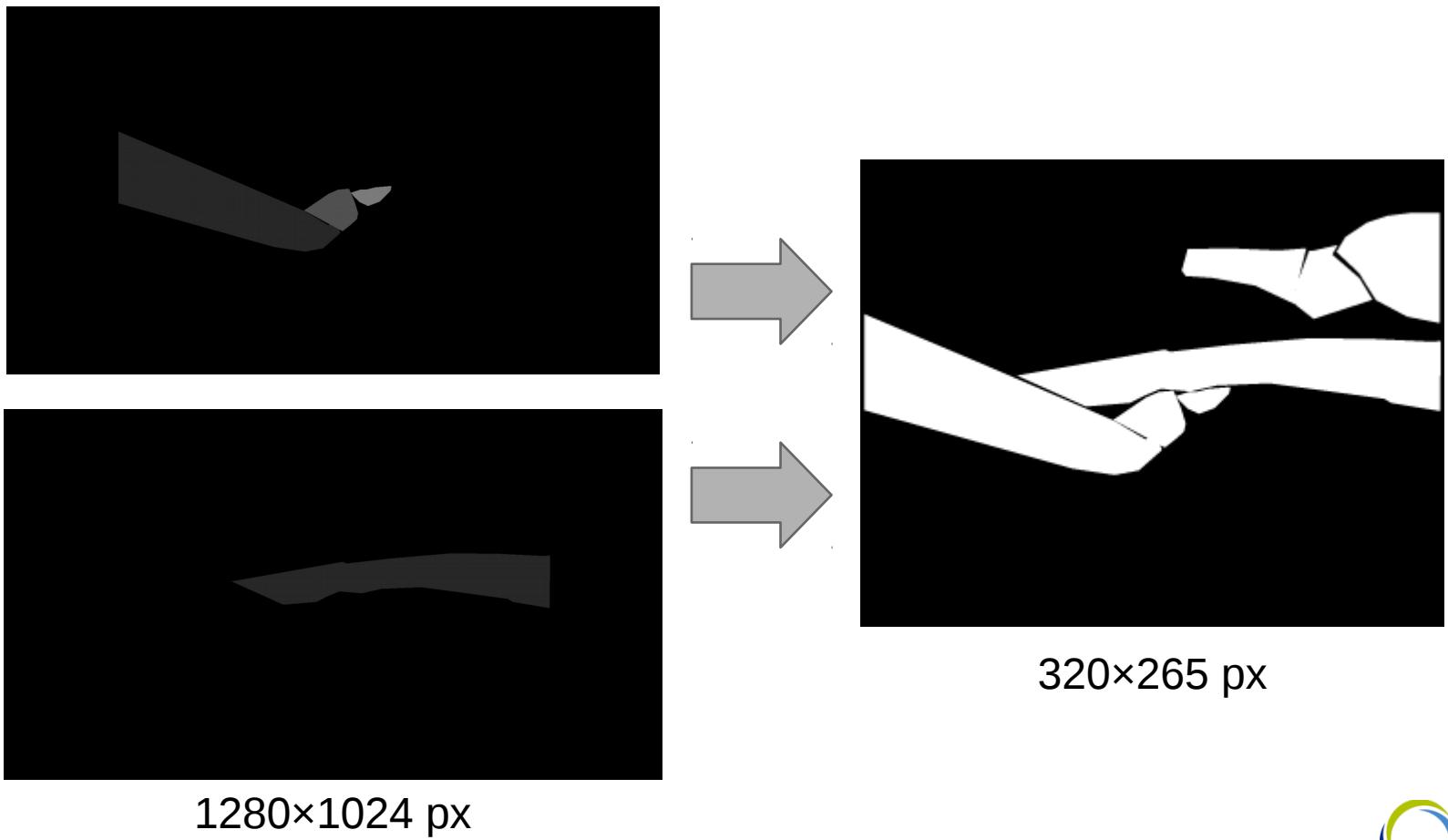


320×265 px

# Data Preparation: Segmentation EndoVis17 [3]

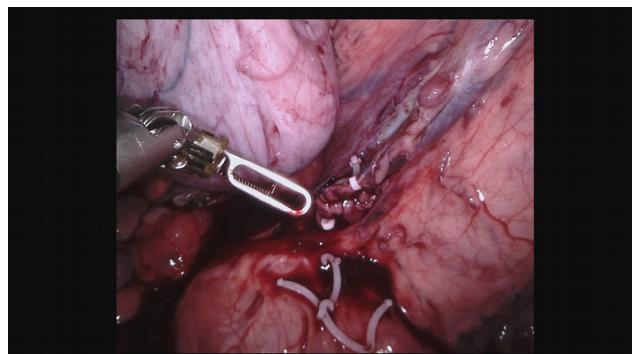
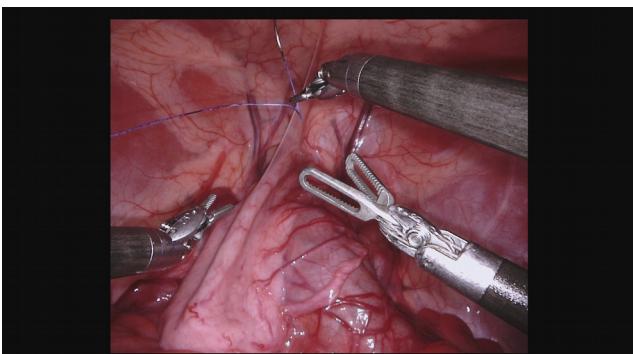
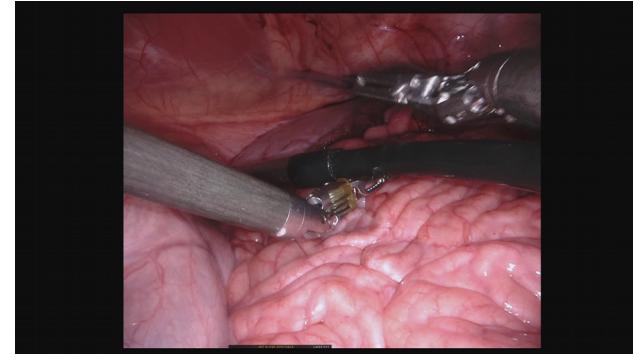
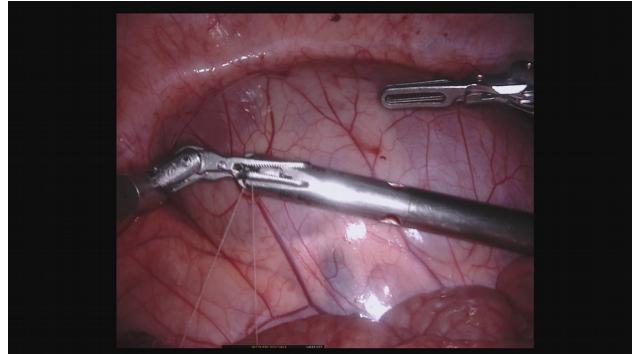
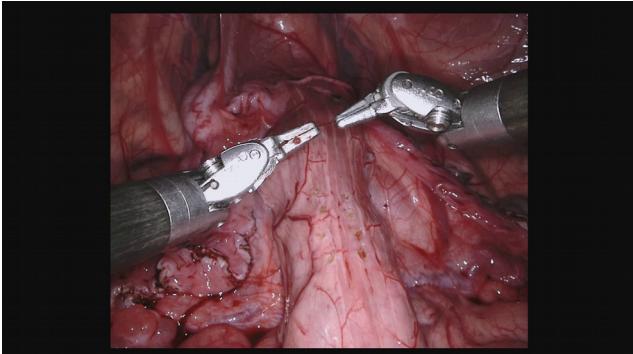
## Ground Truth

- Cropping
- Merging Instruments
- Binary Mask
- Downsampling



# Data: Endoscopic Vision Challenge 2017 [3]

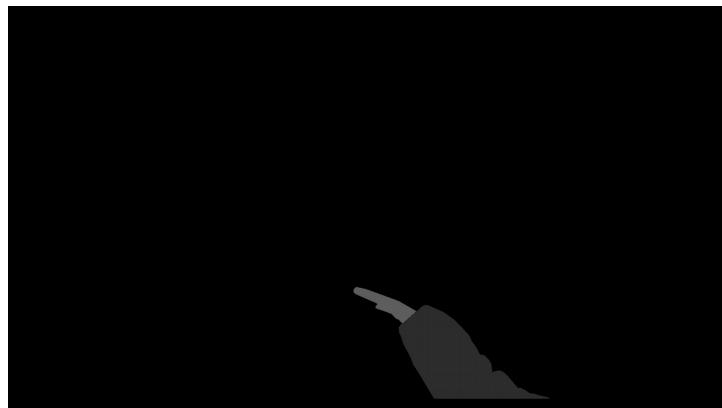
8 Training Datasets, 10 Testing Datasets



...

# Data: Endoscopic Vision Challenge 2017 [3]

Ground Truth Segmentation



...

Satzschrift 1, 20pt

*Auszeichnung*  
**Betonung**

Satzschrift 2, 16 pt

*Auszeichnung*  
**Betonung**

Satzschrift 3, 14 pt

*Auszeichnung*  
**Betonung**

- Satzschrift 1, 20pt

- *Auszeichnung*

- **Betonung**

- Satzschrift 2, 16 pt

- *Auszeichnung*

- **Betonung**

- Satzschrift 3, 14 pt

- *Auszeichnung*

- **Betonung**