

WiSe – Slide Segmentation in the Wild

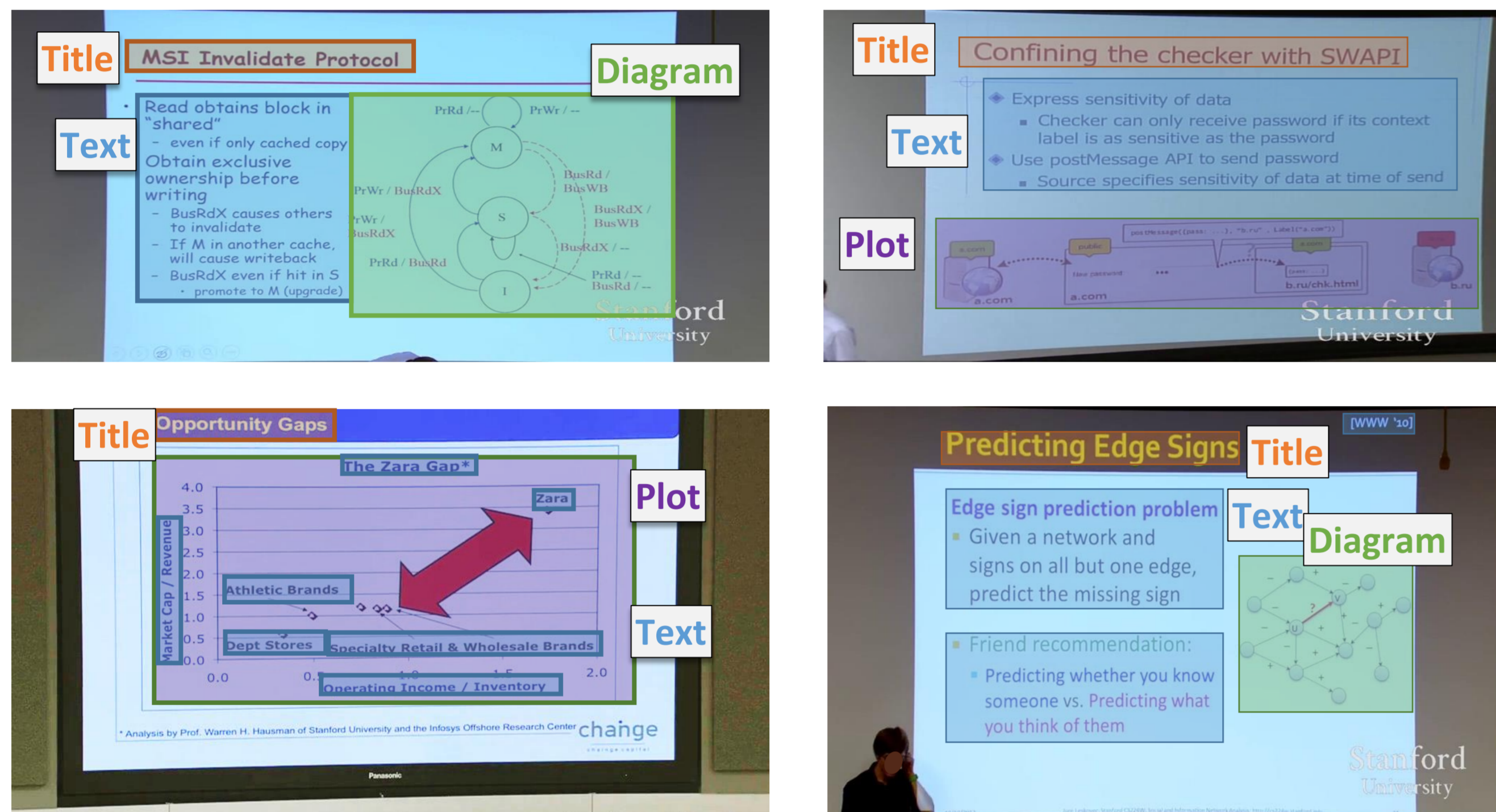
Monica Haurilet, Alina Roitberg, Manuel Martinez, Rainer Stiefelhagen

Motivation

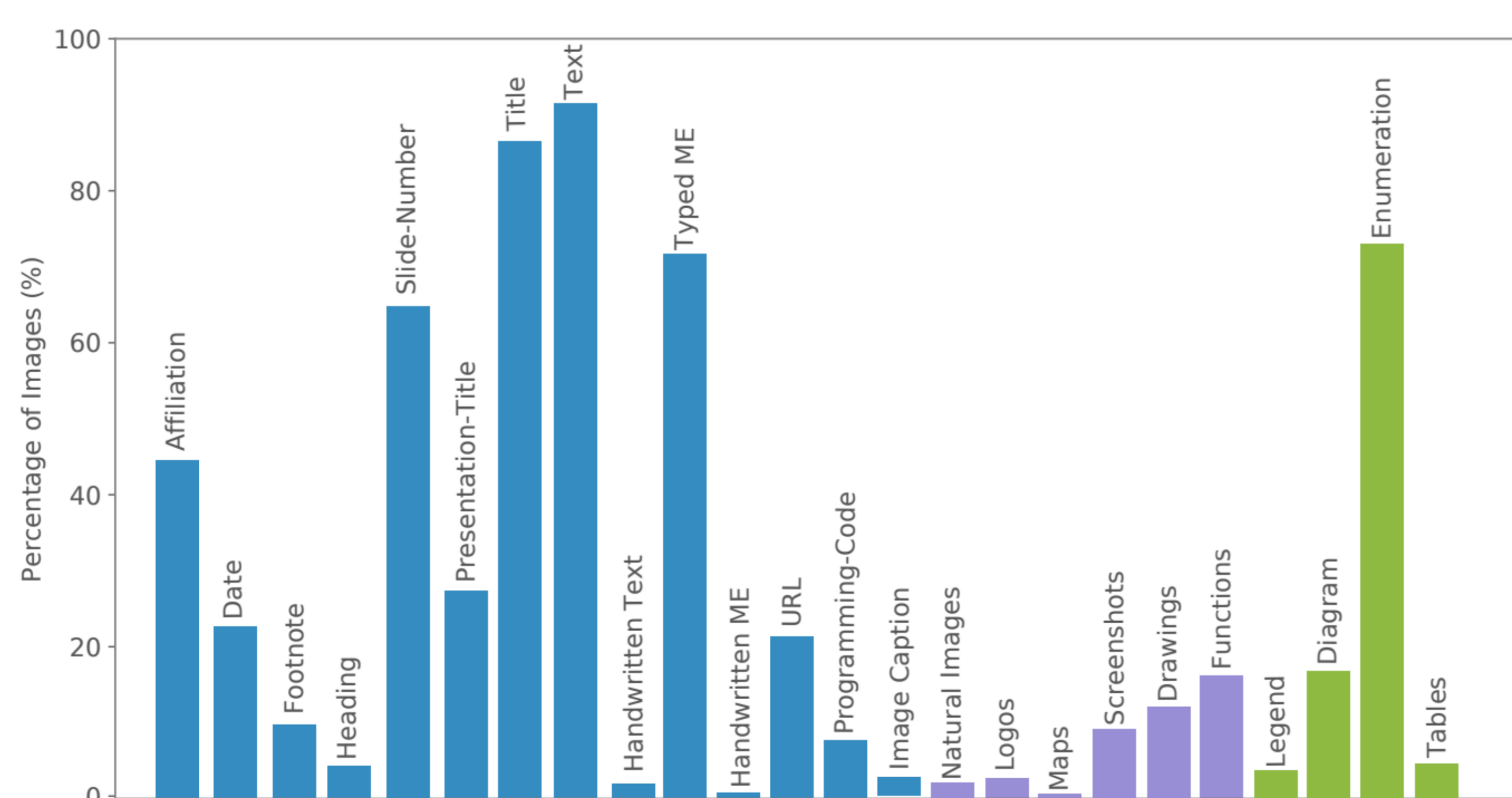
- Slides are a popular way for dissemination of information
- Only page analysis of presentation slides in digital format considered in the past
- Slides captured during lectures are more complex and helpful for visually impaired
- We introduce WiSe – a dataset with fine-grained labels of 1300 slides
- We assign each pixel a set of labels from 25 possible document classes
- The dataset is made publicly available to the community

Slides in the Wild

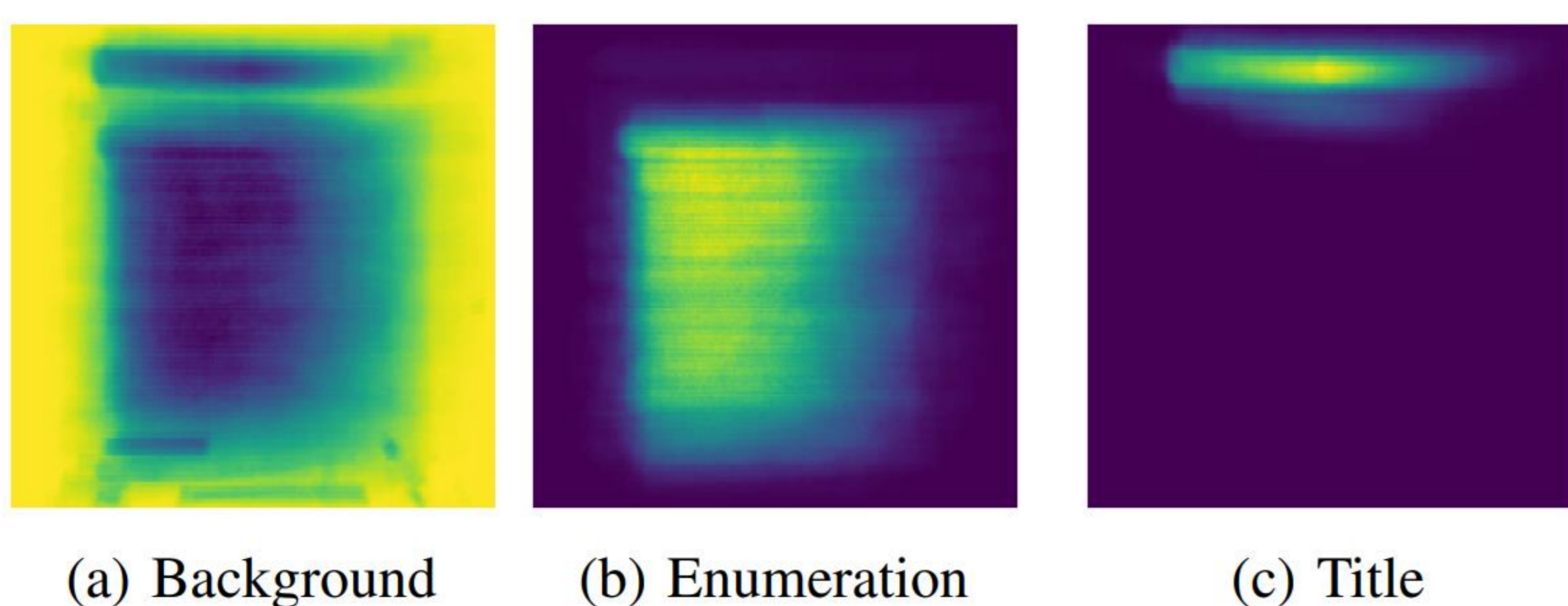
- Examples of presentation slides captured during lecture with assigned labels



Class Distribution



Location Variance



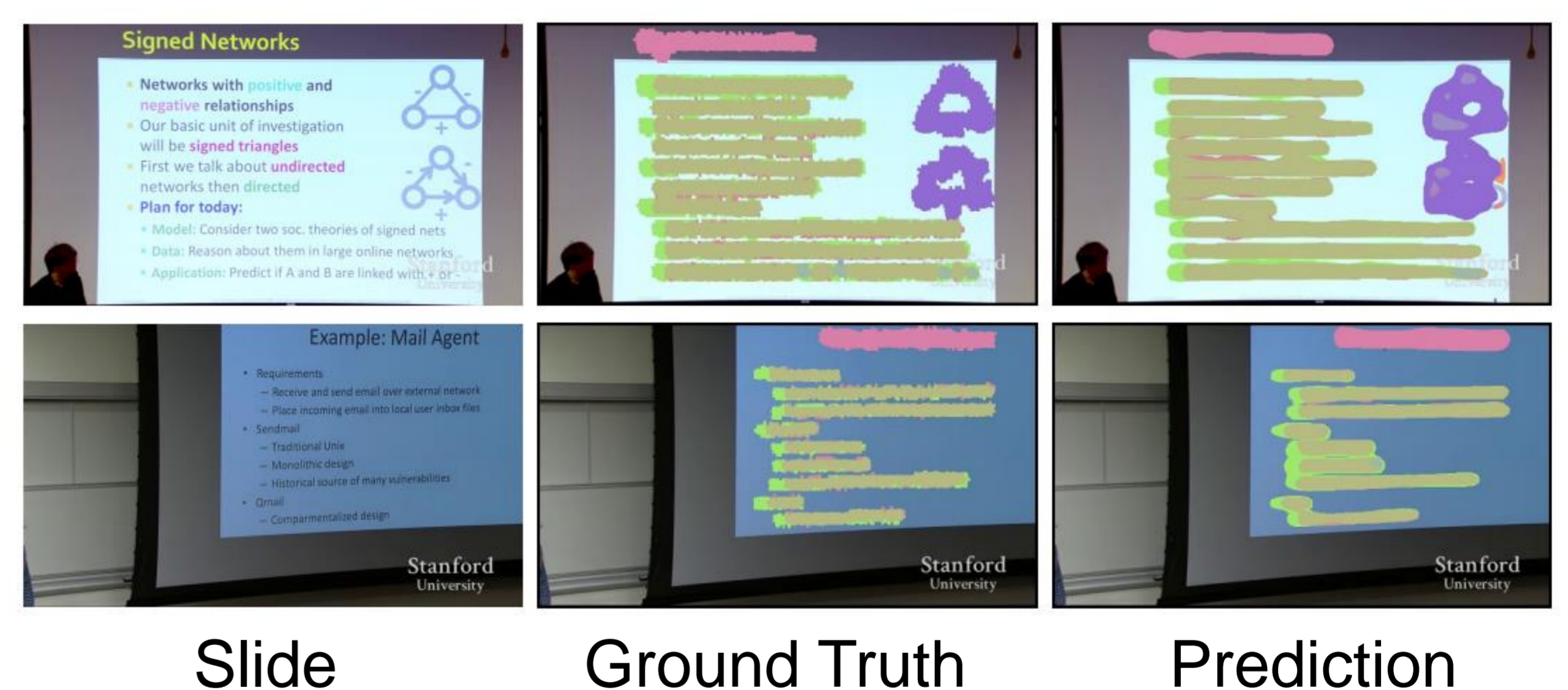
Datasets for Page Segmentation

Dataset	# Pages	# T-Cls.	# I+S-Cls.	Wild	Pxw.	Ovl.
Magazines						
RDCL17	77	10	2	X	✓	X
E-Books						
CM	244	12	3	X	✓	X
Papers						
CS-150	150	2	2	X	X	X
DSSE-200	200	2	3	X	X	✓
SectLabel	347	20	3	X	X	✓
CS-Large	3100	2	2	X	X	X
Street-View						
SVT	350	1	0	✓	X	✓
Presentation Slides						
SPaSe	2000	14	10	X	✓	✓
WiSe (ours)	1335	14	10	✓	✓	✓

Page Segmentation Results

Model	mIOU	pAcc	pIOU	mbAcc
Baseline Methods				
Uniform	0.2	0.0	0.0	50.0
Background	3.0	76.1	76.1	50.0
Neural Networks				
FCN-8s	18.3	81.7	59.8	59.8
DeepLab	35.8	88.3	90.4	72.2
DeepLab+L	37.2	88.5	90.4	72.8

Example Predictions



Contact

Email haurilet@kit.edu

Website <https://cvhci.anthropomatik.kit.edu/~mhaurile>

Download WiSe

<https://cvhci.anthropomatik.kit.edu/data/WiSe>

