

CLASSIFICATION OF SIGN LANGUAGE USING IMAGE SEGMENTATION AND HAND POSE ESTIMATION



(Source 1)

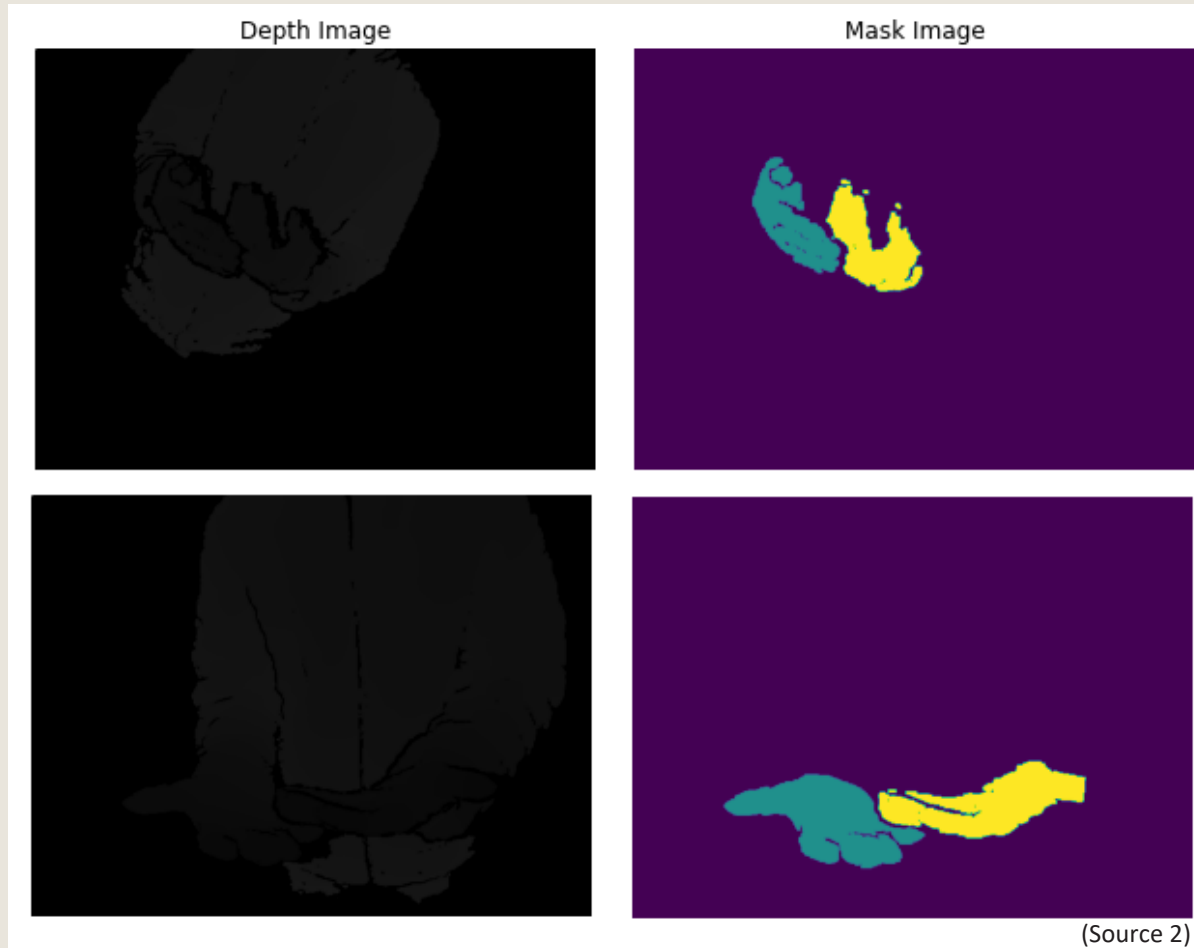
AGENDA

1. Motivation
2. Datasets
3. Methodology
 1. Hand Segmentation
 2. Hand Pose Estimation
 3. Sign Language Classification
4. Results
5. Live-Demo

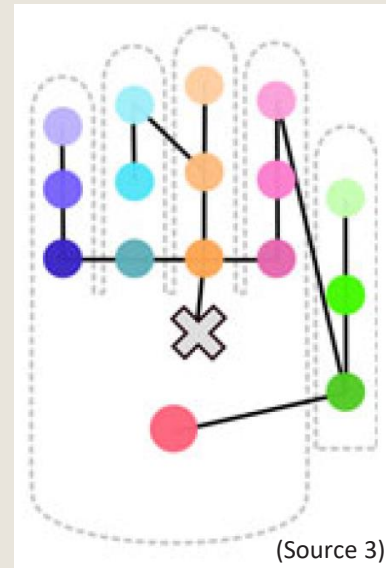
I. MOTIVATION



2. DATASET - HANDSEG



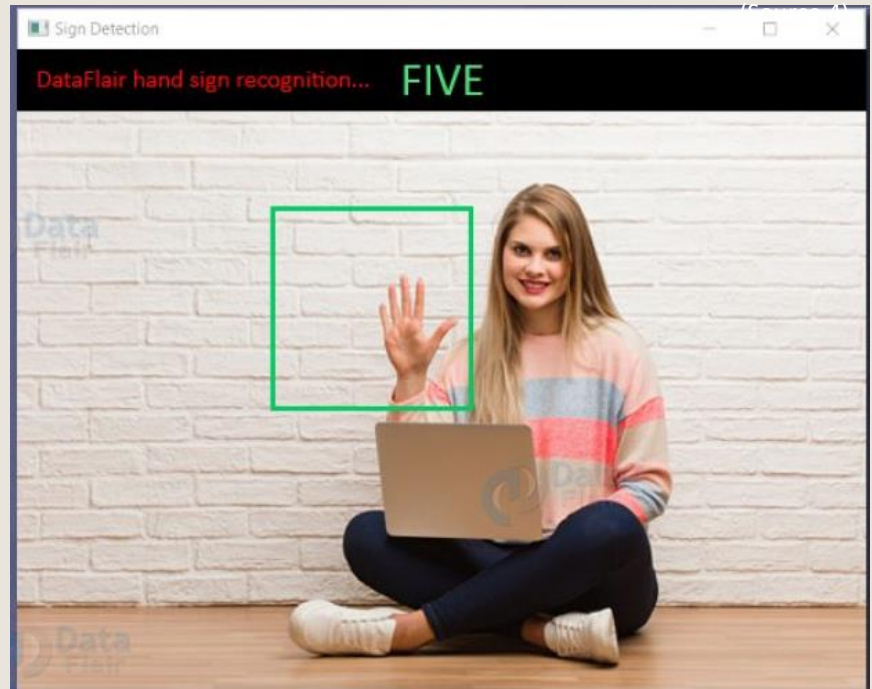
2. DATASET – ICVL HAND POSTURE



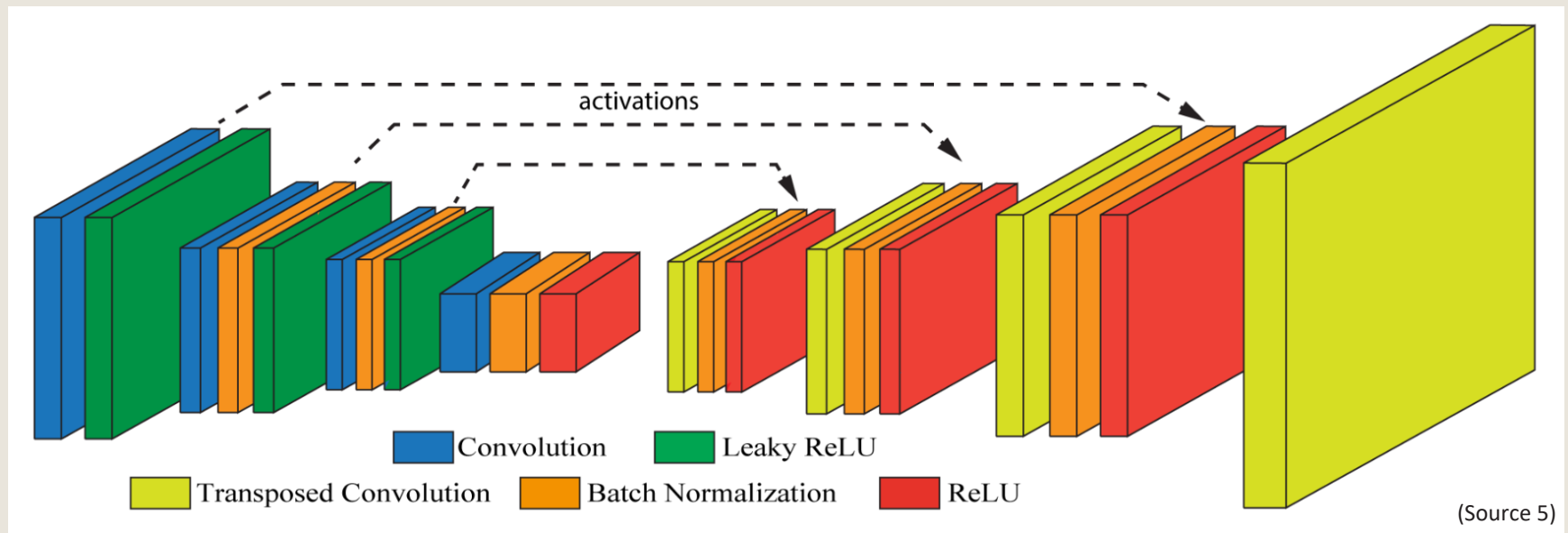
3. METHODOLOGY

Pipeline:

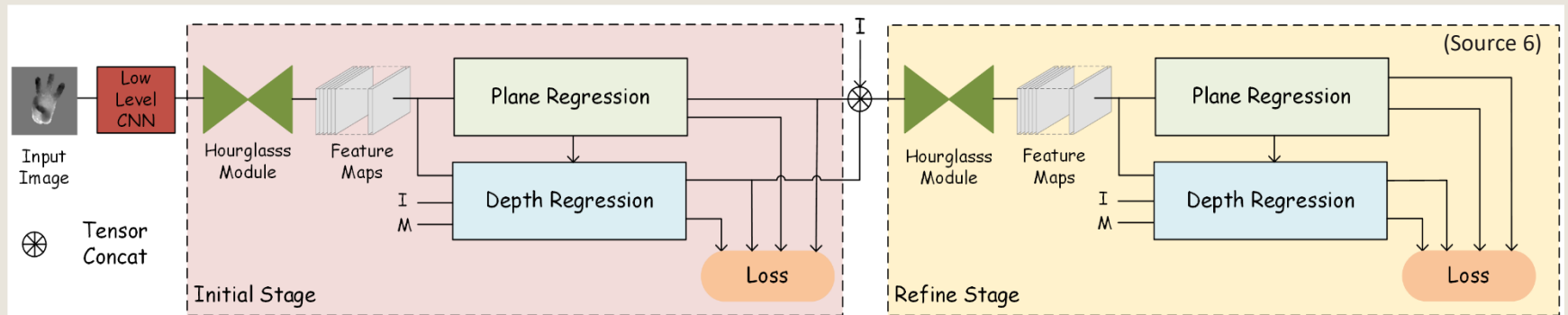
1. Preprocessing
2. Hand Segmentation
3. Bounding Box Prediction
4. Hand Pose Estimation
5. Classification of American Sign Language



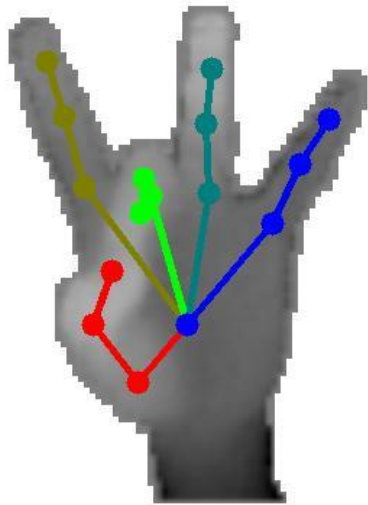
3.1 HAND SEGMENTATION



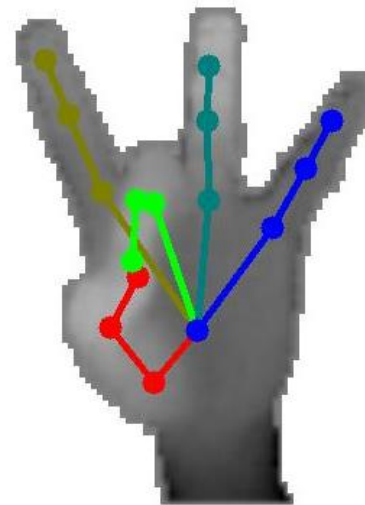
3.2 HAND POSE ESTIMATION



3.2 HAND POSE ESTIMATION

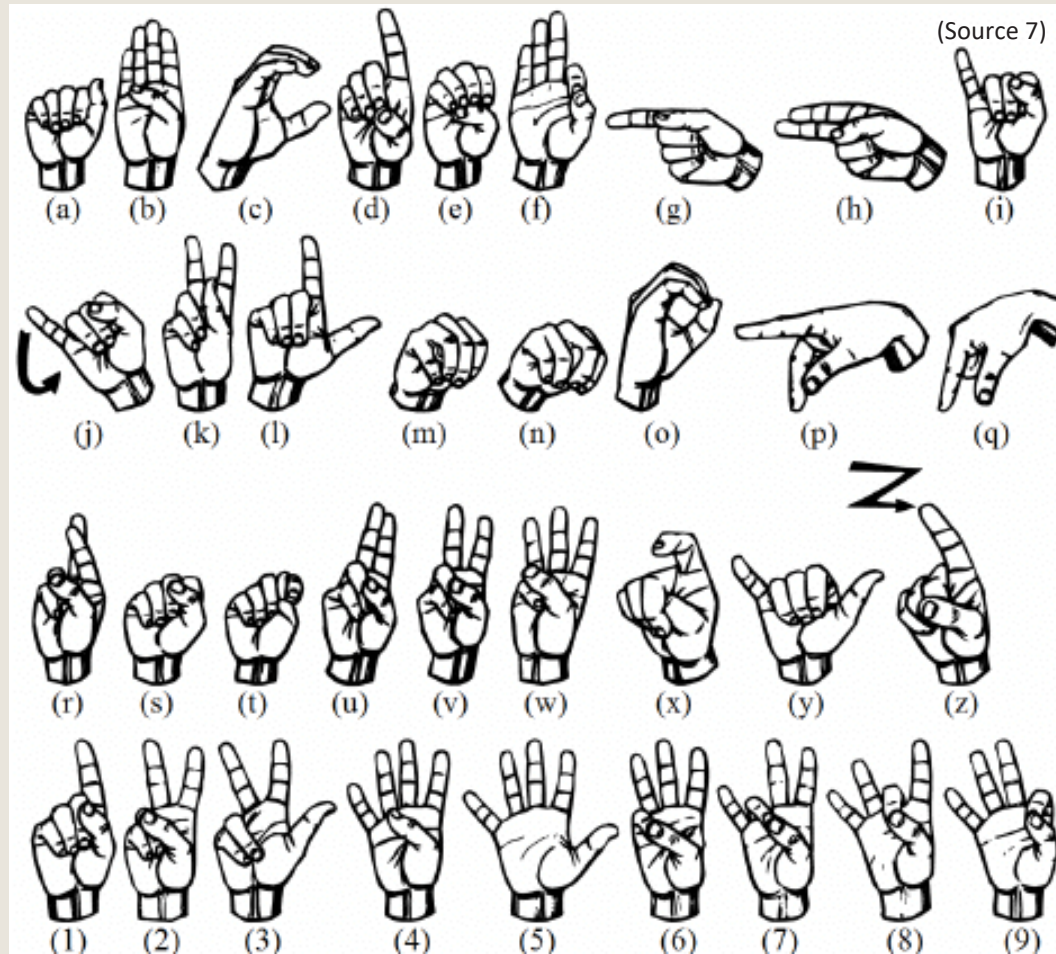


Predicted joints



Ground truth

3.3 SIGN LANGUAGE CLASSIFICATION



4. RESULTS

5.1 Hand Segmentation:

Dataset	Mean IoU
Own implementation	0.66
HandSeg	0.87
SegNet	0.89

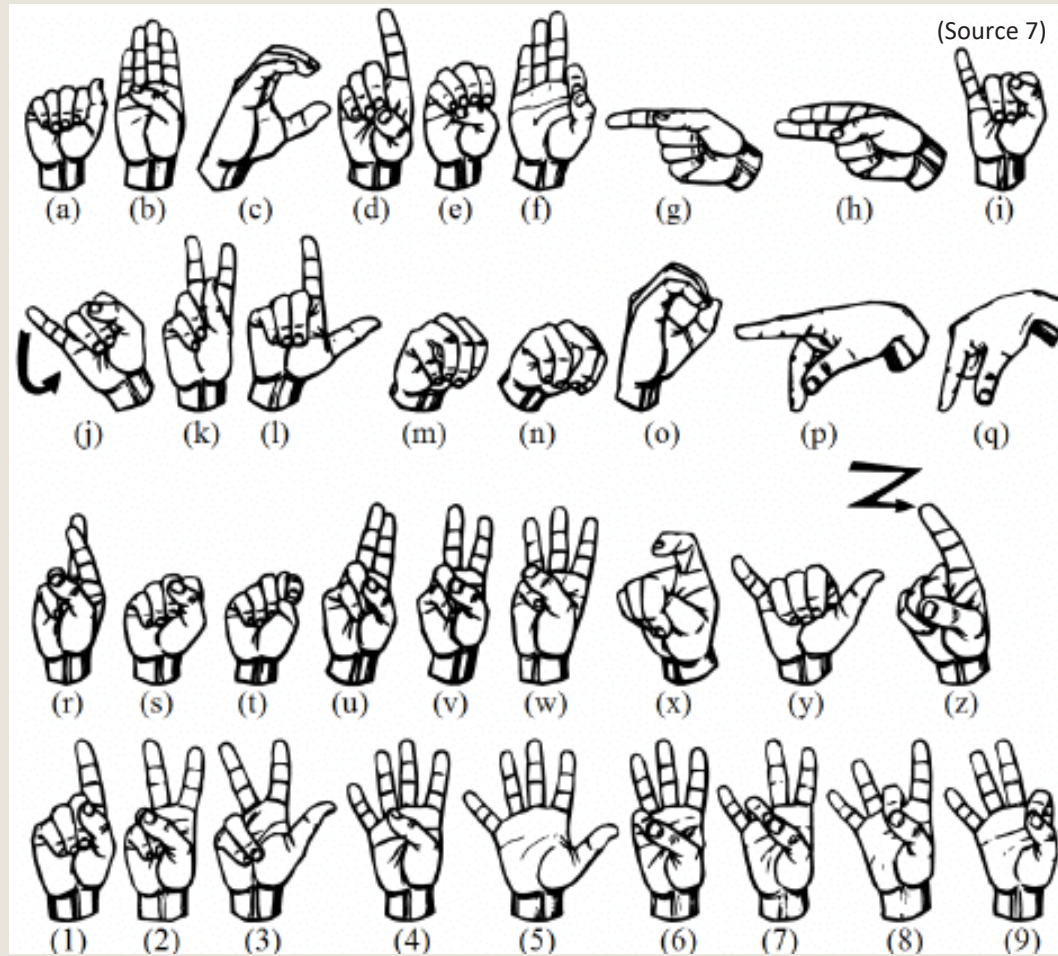
5.2 Bounding Box Prediction:

	Amount of boxes
Fully covered Hand	166
Missing Joints	30 (mean cut: 2.07)
Unusable	4

5.2 Hand Pose Estimation:

Mean 3D Error	3.46
---------------	------

5. LIVE-DEMO



ABBILDUNGSVERZEICHNIS

Source 1:

https://en.wikipedia.org/wiki/American_Sign_Language#/media/File:American_Sign_Language_ASL.svg

Source 2:

Bojja, Abhishake Kumar, et al. "Handseg: An automatically labeled dataset for hand segmentation from depth images." *2019 16th Conference on Computer and Robot Vision (CRV)*. IEEE, 2019.

Source 3:

Tang, Danhang, et al. "Latent regression forest: structured estimation of 3d hand poses." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.7 (2016): 1374-1387.

Source 4:

<https://data-flair.training/blogs/sign-language-recognition-python-ml-opencv/>

Source 5:

A. K. Bojja, F. Mueller, S. R. Malireddi, M. Oberweger, V. Lepetit, C. Theobalt, K. M. Yi, and A. Tagliasacchi, Conference on Computer and Robot Vision, "Handseg: An automatically labeled dataset for hand segmentation from depth images," in 16th, CRV 2019, IEEE, May 29-31, 2019, Kingston, ON, Canada

Source 6:

Zhang, Xingyuan, and Fuhai Zhang. "Pixel-wise regression: 3d hand pose estimation via spatial-form representation and differentiable decoder." *arXiv preprint arXiv:1905.02085* (2019)

Source 7:

B. Kang, S. Tripathi, and T. Q. Nguyen, in 3rd IAPR Asian Conference on Pattern Recognition, "Real-time sign language fingerspelling recognition using convolutional neural networks from depth map," ACPR 2015, , November 3-6, 2015, IEEE, Kuala Lumpur, Malaysia