Towards Self-Perception in Augmented Virtuality: Hand Segmentation with Fully Convolutional Networks

Ester Gonzalez-Sosa, Pablo Perez, Redouane Kachach, Jaime J. Ruiz, and Alvaro Villegas
Nokia Bell Labs

Abstract

In this work, we propose the use of deep learning techniques to segment items of interest from the local region **to increase self-presence** in Virtual Reality (VR) scenarios. Our goal is to segment hand images from the perspective of a user wearing a VR headset. We create the **VR Hand Dataset**, composed of more than 10.000 images, including variations of hand position, scenario, outfits, sleeve and people. We also describe the procedure followed to automatically generate groundtruth images and create synthetic images. Preliminary results look promising.

Self Perception

- Reduce isolation
- Ease interaction

Virtual Reality

Augmented Virtuality

Hand Segmentation

Local Reality

VR Hand Dataset




Chroma Key | Foreground | Background | Synthetic Image

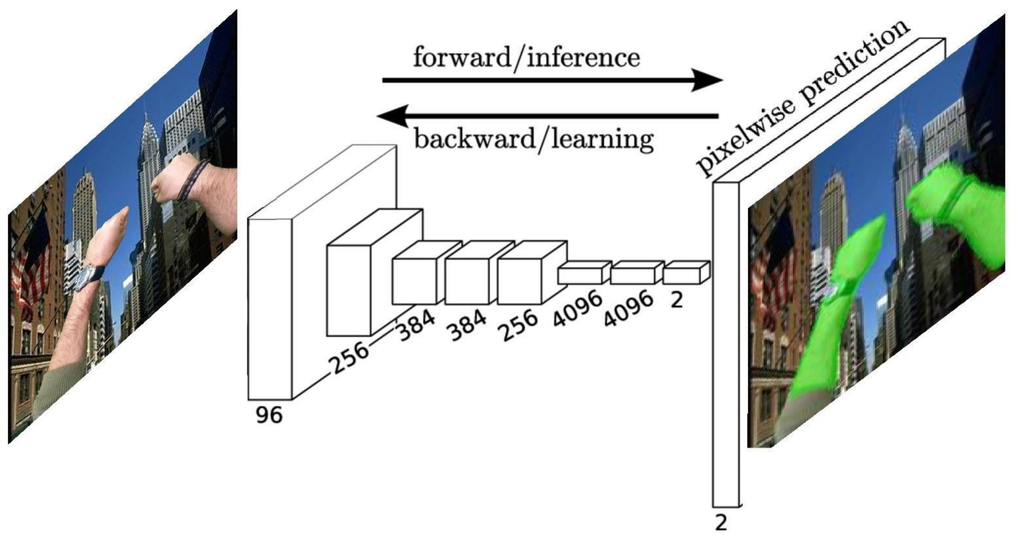
- Acquisition of **Chroma key**
- HSV Filtering to get **Foreground**

$$f(x) = \begin{cases} 1 & \text{if } H(x,y) \leq 0.22 \wedge H(x,y) \geq 0.45 \wedge S(x,y) \geq 0.20 \\ 0 & \text{otherwise} \end{cases}$$

- Frame Selection and Preprocessing
- Combination of Chroma key masked with Foreground with **Background** to get the **Synthetic Image**

Hand Segmentation

Fully Convolutional Networks



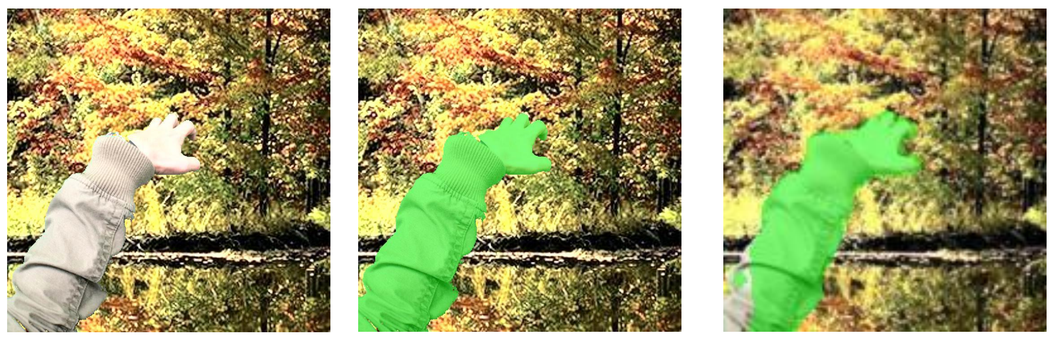
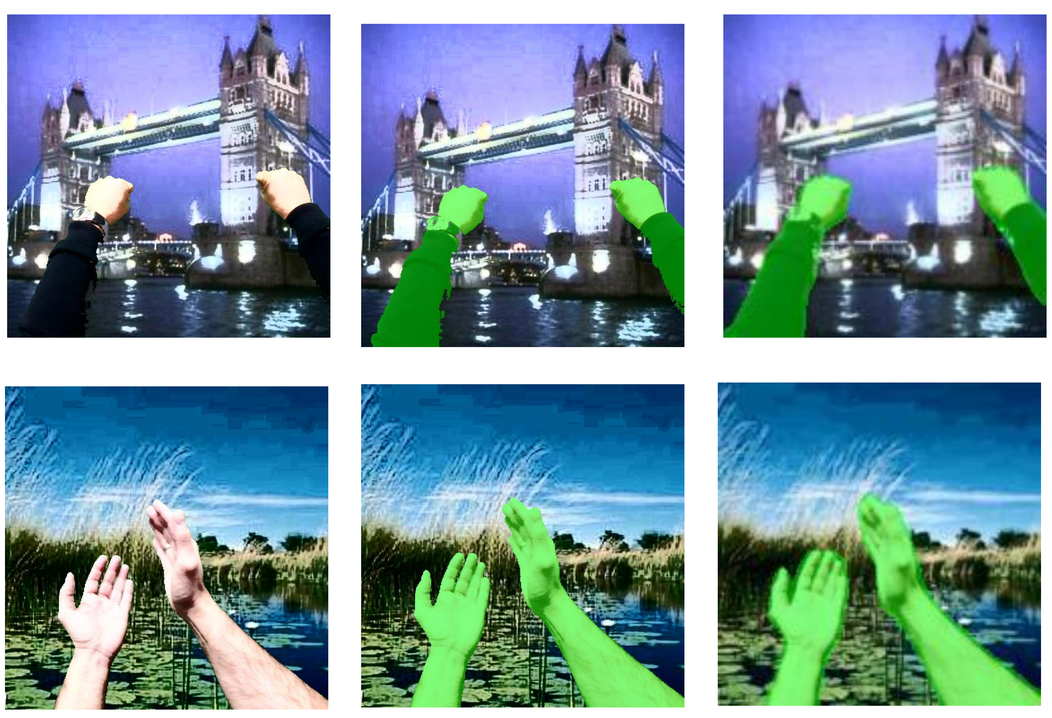
- Inspired by [1]
- 2 classes: hand and background
- Transfer learning from VGG-16 pre-trained model

Future Work

- Deployment on embedded devices
- Quantitative results in terms of IoU
- Test generalization capabilities
- Further exploration of Semantic Segmentation

Results

Original | Groundtruth | FCN

References

[1] LONG J., SHELHAMER E., DARRELL T.: Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015), pp. 3431–3440.

NOKIA Bell Labs