Automatic Skin Labeling of Egocentric Images for Virtual Reality Applications

Heiko Raible
h_da Hochschule Darmstadt

Structure

Structure

- Introduction

Structure

- Introduction
- Research

Structure

- Introduction
- Research
- Implementation

Structure

- Introduction
- Research
- Implementation
- Evaluation

Structure

- Introduction
- Research
- Implementation
- Evaluation
- Conclusion

Introduction

Previous Work

My Contribution

Research

GrabCut

GrabCut

- rough foreground and background mask

GrabCut

- rough foreground and background mask
- iteratively optimizes masks
  - Gaussian Mixture Models
  - spatial information
  - graph cuts

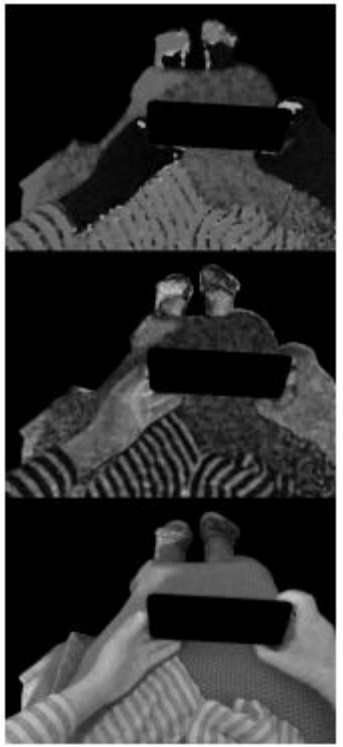
Color Spaces

Color Spaces

- RGB (red, green, blue)

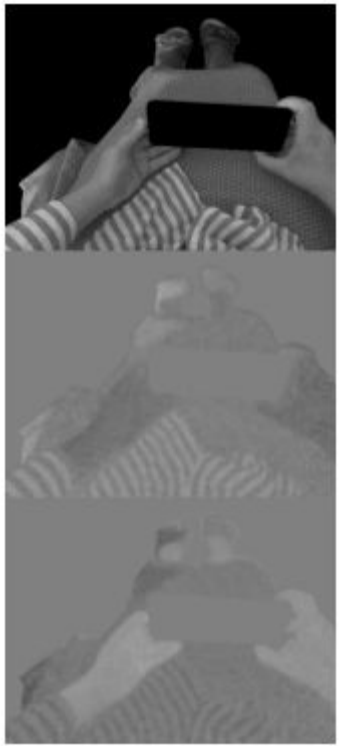
Color Spaces

- RGB (red, green, blue)
- HSV (hue, saturation, value)



Color Spaces

- RGB (red, green, blue)
- HSV (hue, saturation, value)
- YCrCb (luma, red-difference chroma, blue-difference chroma)



Color Spaces

- RGB (red, green, blue)
- HSV (hue, saturation, value)
- YCrCb (luma, red-difference chroma, blue-difference chroma)
- LAB (lightness, green-red, blue-yellow)

Otsu’s Method

Otsu’s Method

- operates on greyscale

Otsu’s Method

- operates on greyscale
- calculates variance of intensity per threshold

Otsu’s Method

- operates on greyscale
- calculates variance of intensity per threshold
- maximizes between-class variance

MediaPipe Hands

MediaPipe Hands

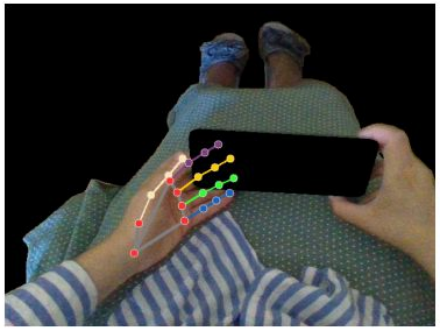
- lightweight Convolutional Neural Network (CNN)

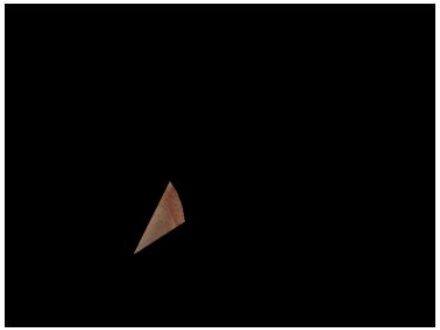
MediaPipe Hands

- lightweight Convolutional Neural Network (CNN)
- based on MediaPipe Framework

MediaPipe Hands

- lightweight Convolutional Neural Network (CNN)
- based on MediaPipe Framework
- detects hands
  ⇒ skin patch Region of Interest (ROI)





Mask R-CNN

Mask R-CNN

- ResNet (Residual Network)
  - addresses vanishing gradient problem (skip connections)

Mask R-CNN

- ResNet (Residual Network)
  - addresses vanishing gradient problem (skip connections)
- Faster R-CNN
  - region proposal network for object detection
  - region of interest pooling layer for classification and refinement

Mask R-CNN

- ResNet (Residual Network)
  - addresses vanishing gradient problem (skip connections)
- Faster R-CNN
  - region proposal network for object detection
  - region of interest pooling layer for classification and refinement
- Mask R-CNN
  - mask prediction branch to generate segmentation mask per object

Implementation

Filtering Methods

Filtering Methods

- color range inside ROI
  - RGB
  - HSV (no value)
  - YCrCb (no luminance)

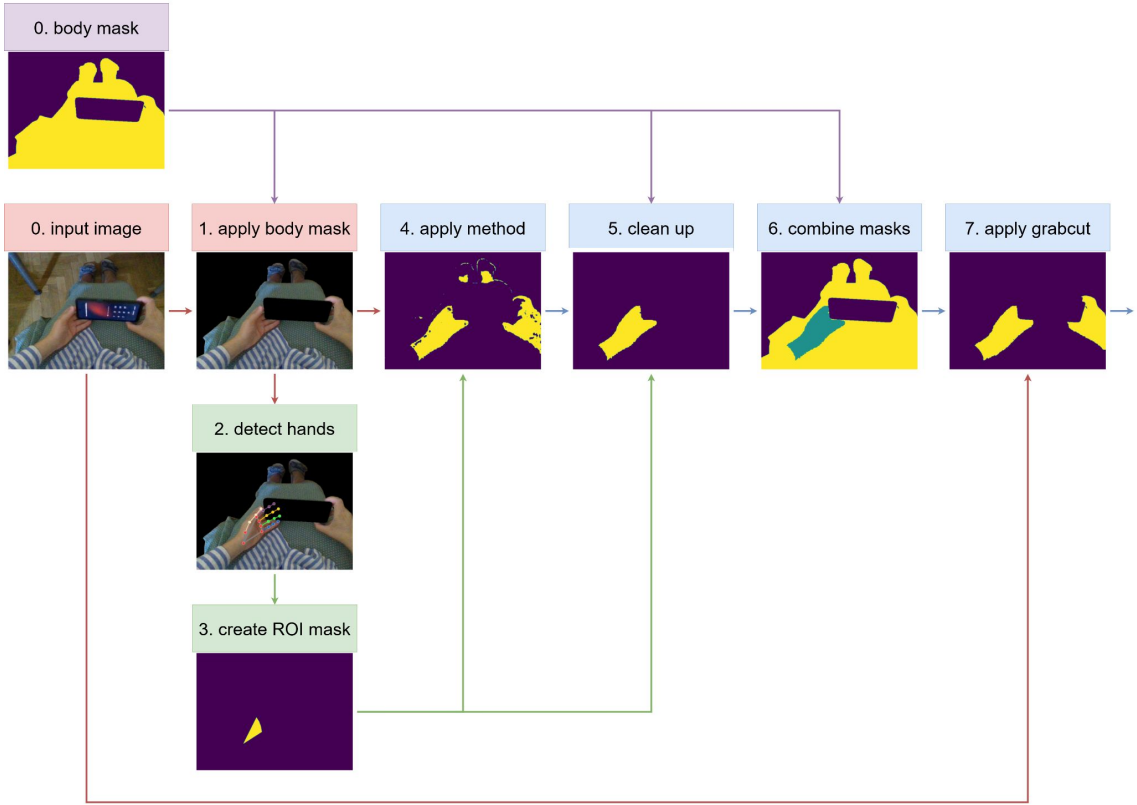
Filtering Methods

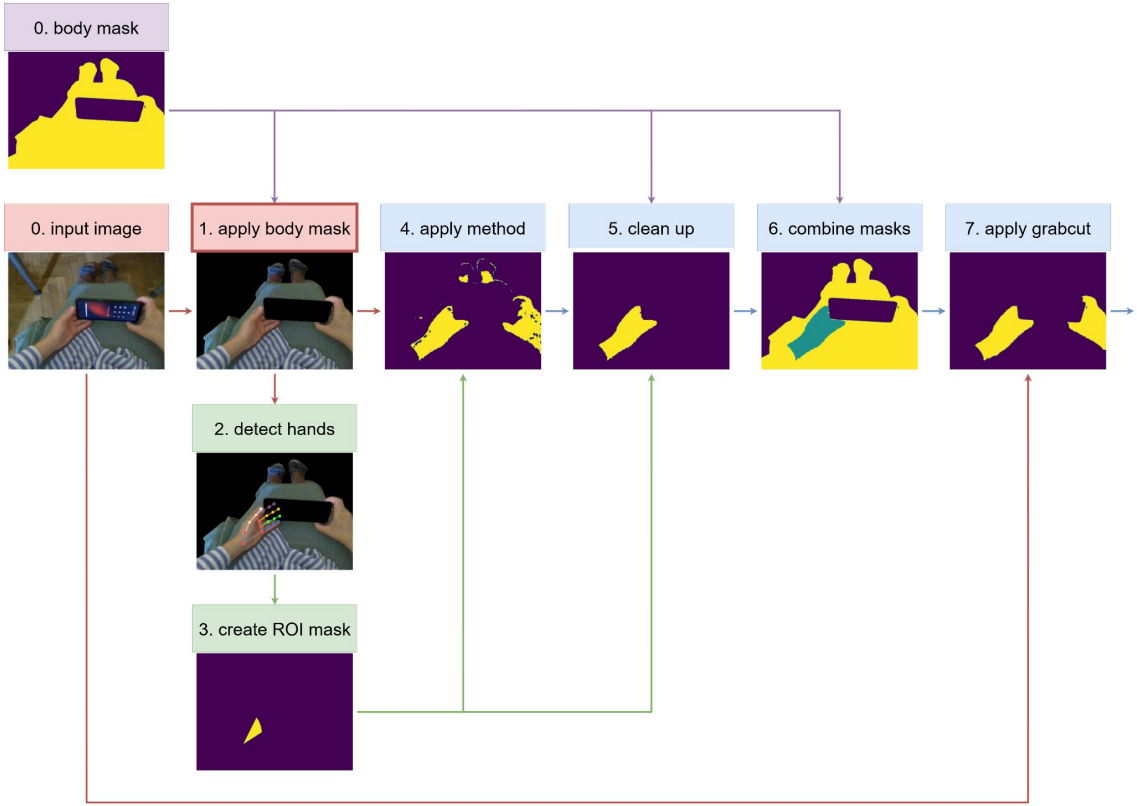
- color range inside ROI
  - RGB
  - HSV (no value)
  - YCrCb (no luminance)
- Otsu’s method

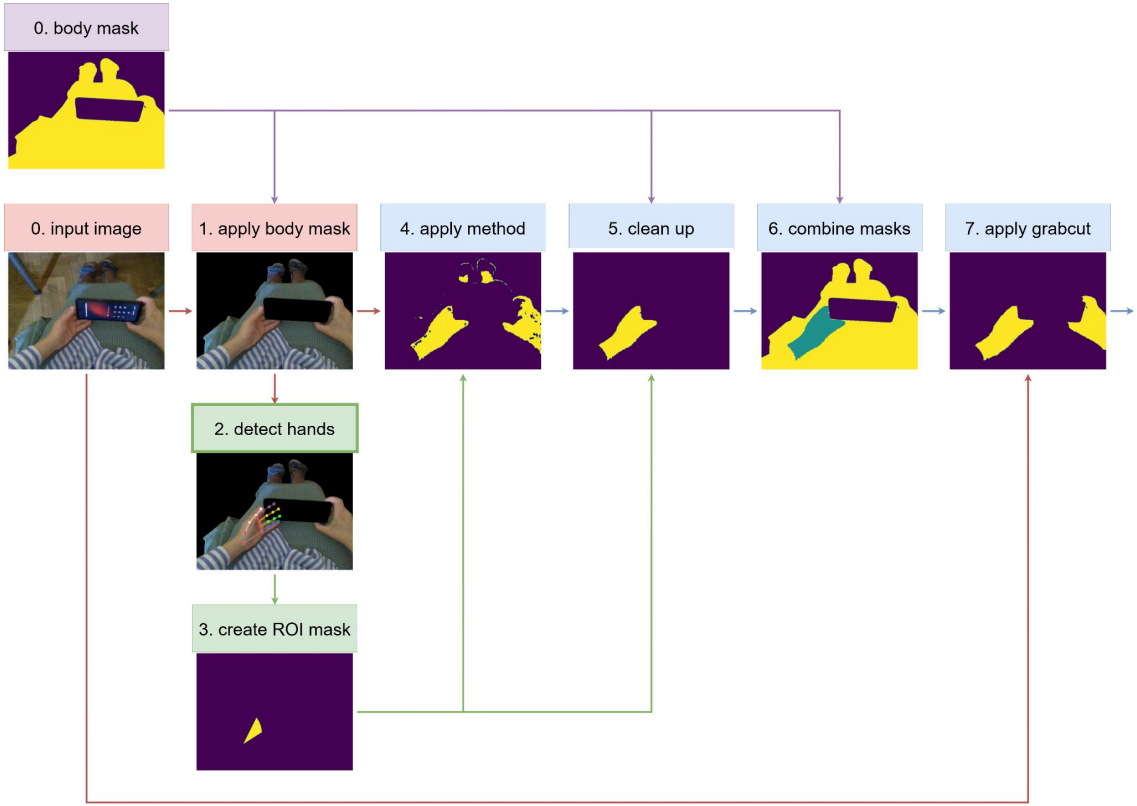
Filtering Methods

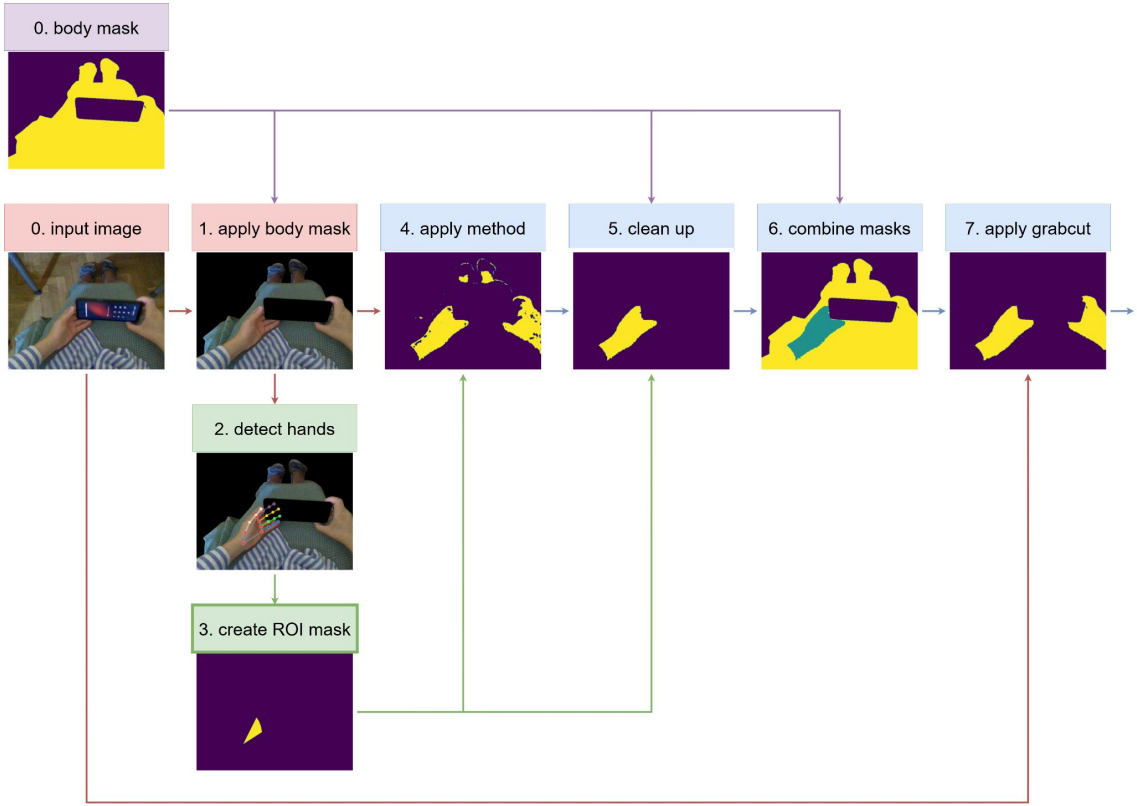
- color range inside ROI
  - RGB
  - HSV (no value)
  - YCrCb (no luminance)
- Otsu’s method
- histogram based inside ROI
  - HSV
  - LAB
  - YCrCb
  - HCrCb

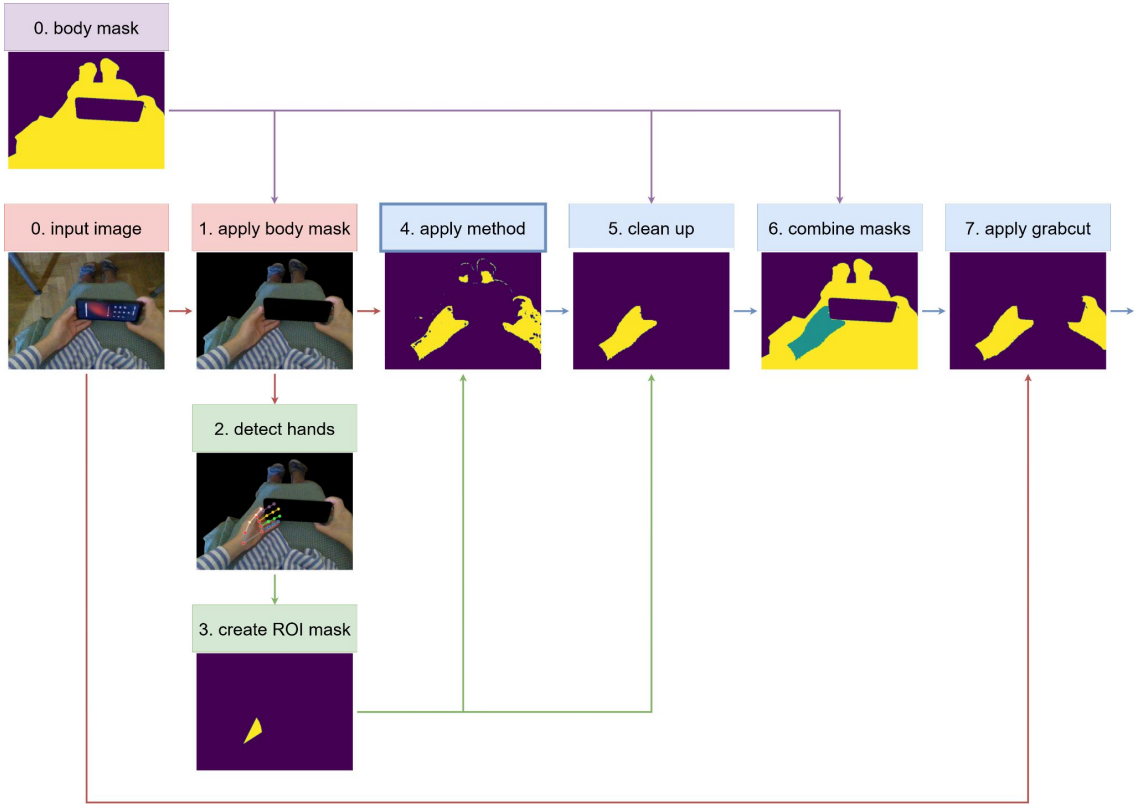
End-to-end System

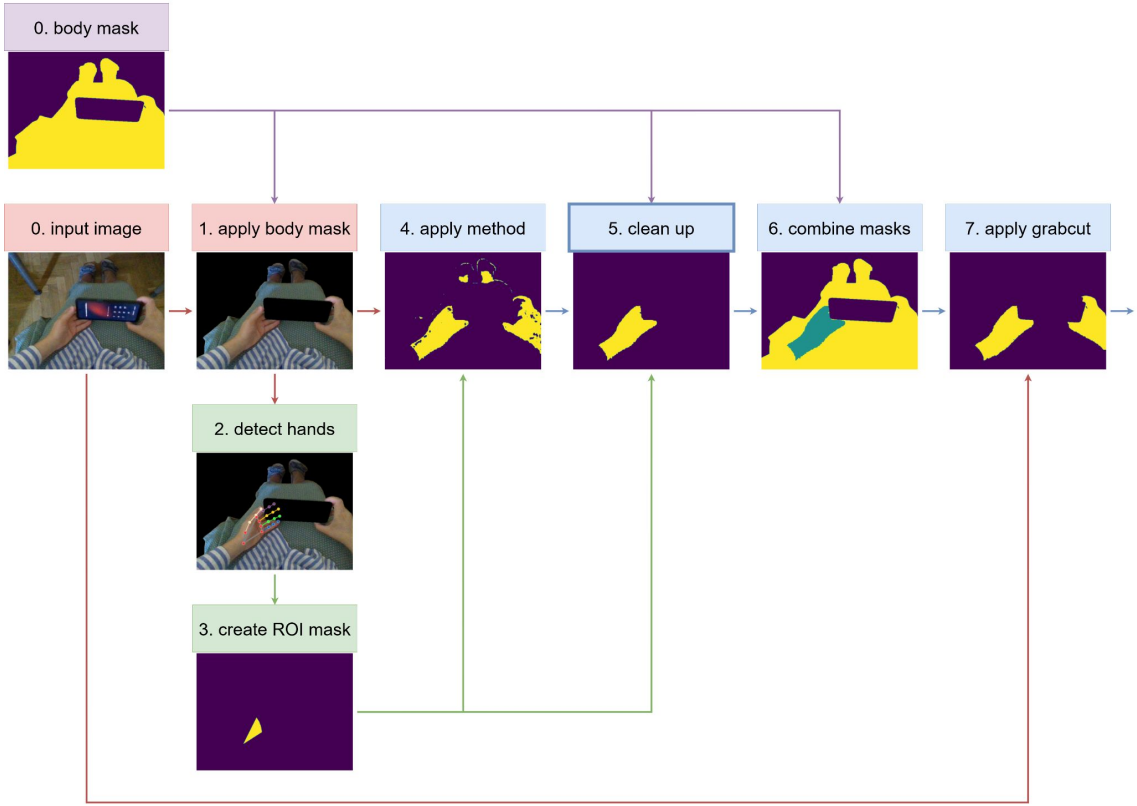

0. body mask
0. input image
1. apply body mask
4. apply method
5. clean up
6. combine masks
7. apply grabcut
2. detect hands
3. create ROI mask

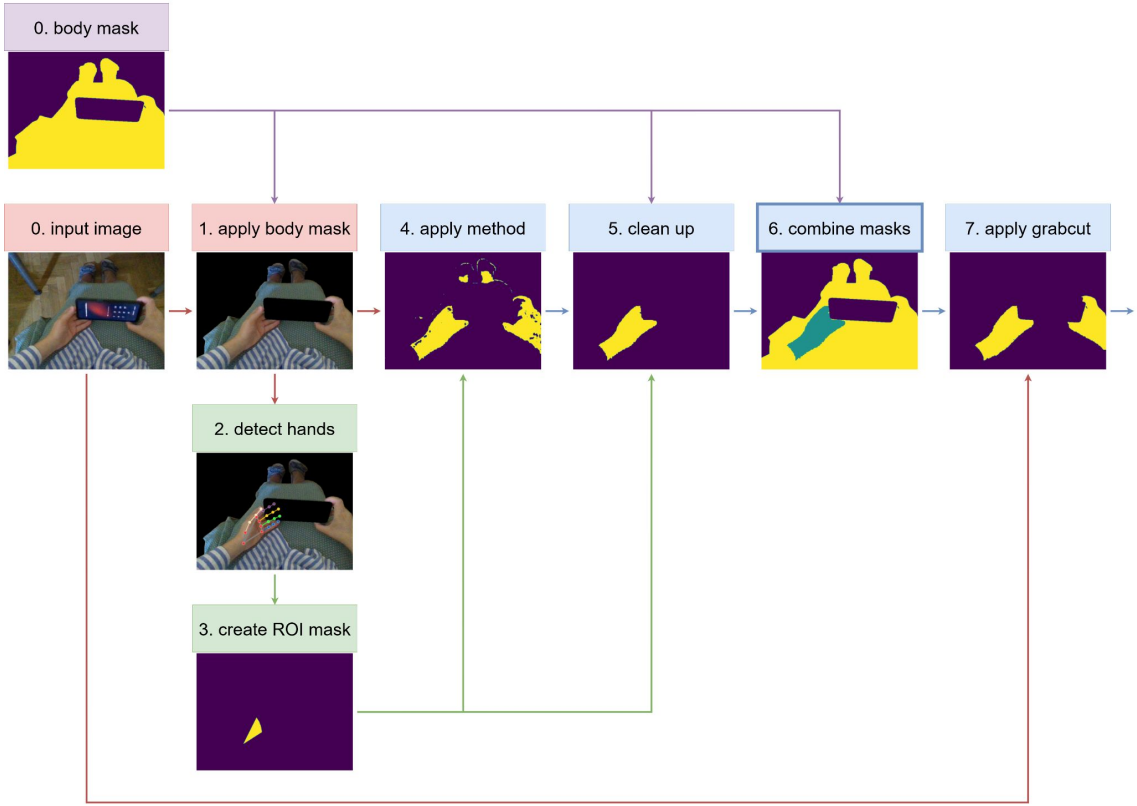

0. body mask
0. input image
1. apply body mask
4. apply method
5. clean up
6. combine masks
7. apply grabcut
2. detect hands
3. create ROI mask

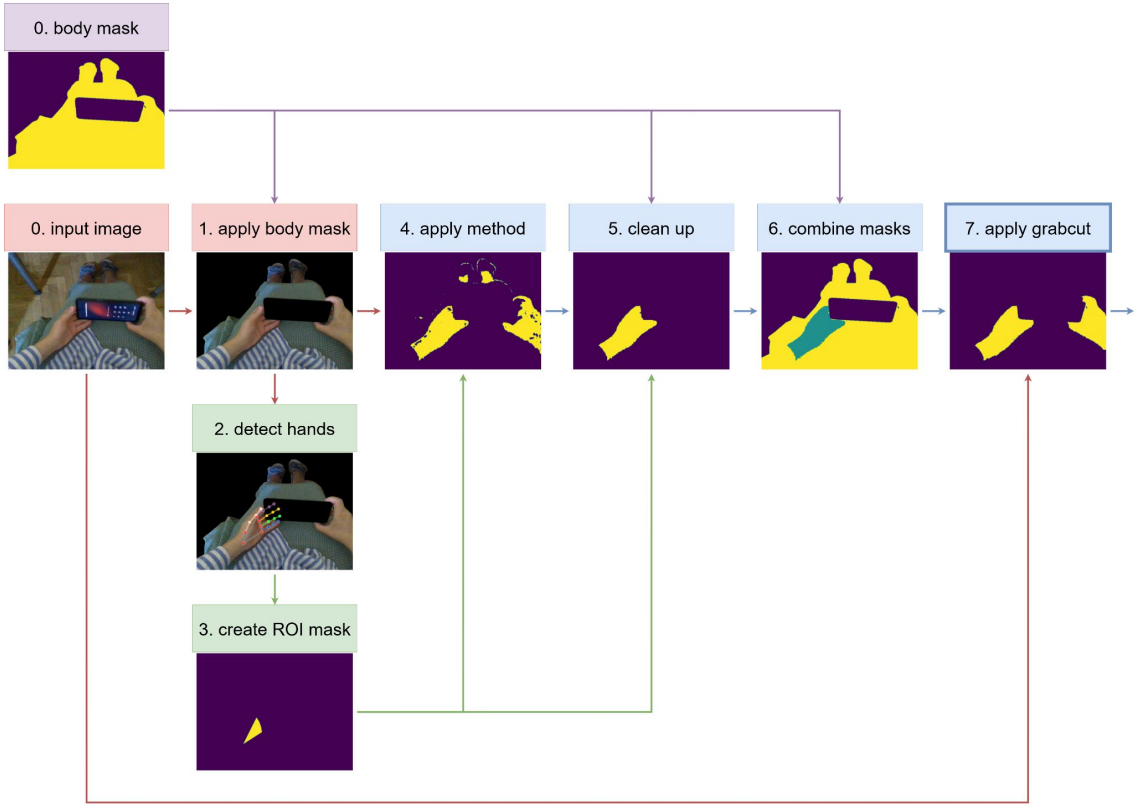

0. body mask
0. input image
1. apply body mask
4. apply method
5. clean up
6. combine masks
7. apply grabcut
2. detect hands
3. create ROI mask


0. body mask
0. input image
1. apply body mask
4. apply method
5. clean up
6. combine masks
7. apply grabcut
2. detect hands
3. create ROI mask


0. body mask
0. input image
1. apply body mask
4. apply method
5. clean up
6. combine masks
7. apply grabcut
2. detect hands
3. create ROI mask


0. body mask
0. input image
1. apply body mask
2. detect hands
3. create ROI mask
4. apply method
5. clean up
6. combine masks
7. apply grabcut


0. body mask
0. input image
1. apply body mask
2. detect hands
3. create ROI mask
4. apply method
5. clean up
6. combine masks
7. apply grabcut


0. body mask
0. input image
1. apply body mask
4. apply method
5. clean up
6. combine masks
7. apply grabcut
2. detect hands
3. create ROI mask

Mask R-CNN

Mask R-CNN

- Training
  - pretrained on COCO
  - fine-tuned for 20 epochs

Evaluation

Datasets

Datasets

- THU-READ
  - 158 images
  - multiple of same scene
  - many where skin = body
    ⇒ 35 body images

Datasets

- THU-READ
  - 158 images
  - multiple of same scene
  - many where skin = body
    ⇒ 35 body images
- Egocentric Bodies
  - 103 images
  - fewer of same scene
  - skin ≠ body

THU_READ_RGBD

joint-ep-of-thu-ego

Datasets

- THU-READ
- Egocentric Bodies
- Amazon Mechanical Turk

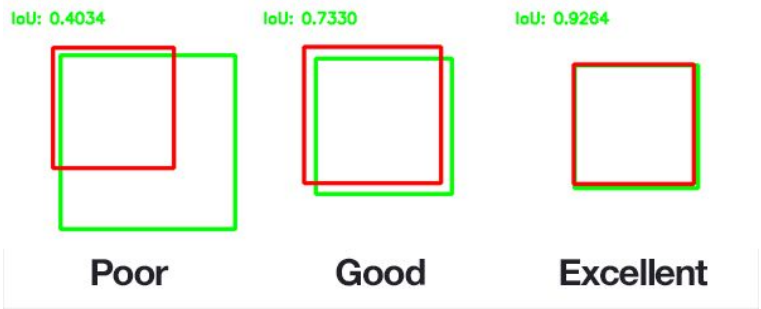
Datasets

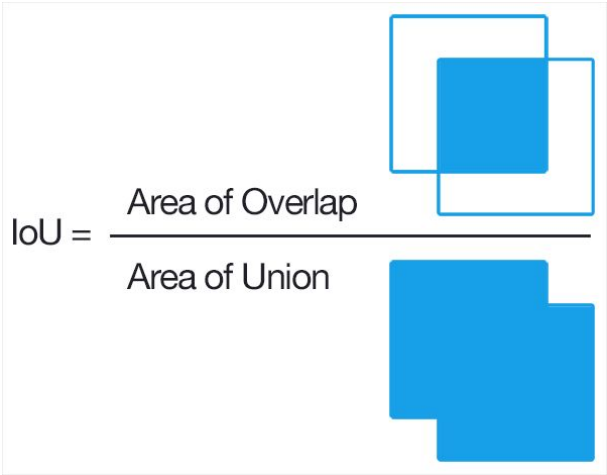
- THU-READ
- Egocentric Bodies
- Amazon Mechanical Turk
- end-to-end system

Datasets

- THU-READ
- Egocentric Bodies
- Amazon Mechanical Turk
- end-to-end system
- Mask R-CNN
  - data augmentation
    ⇒ 259 → 519 images
  - 200 random test dataset
    319 remaining training dataset

(mean) Intersection over Union (IoU)





Results

Results

- end-to-end system

Results

- end-to-end system
- GrabCut iterations

Results

- end-to-end system
- GrabCut iterations
- Mask R-CNN epochs

Results

- end-to-end system
- GrabCut iterations
- Mask R-CNN epochs
- illustrative examples

thresholding method	**THU_READ_RGBD** (158 images)			**THU_READ_RGBD_BODIES** (35 images)			**joint-ep-thu-ego** (103 images)		
	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut
baseline		81.7			49.0			40.3	
ROI mask	07.4	07.4	79.4	06.3	06.3	77.2	09.1	09.1	61.4
RGB	07.4	74.5	75.9	08.3	45.5	45.8	11.0	31.4	32.8
HS	06.7	71.0	76.7	07.0	48.9	49.7	12.7	48.1	46.4
CrCb	32.1	74.8	79.0	37.2	66.9	65.4	29.8	44.5	48.1
Otsu's	71.8	67.8	79.4	71.2	70.3	77.9	47.8	45.0	50.5
RGB · HS	06.5	69.7	76.8	06.9	48.8	50.1	11.8	44.6	46.5
RGB · CrCb	31.2	73.6	79.4	36.8	66.4	65.5	28.0	42.0	48.2
HS · CrCb	28.7	68.8	80.0	32.7	60.1	69.0	28.6	44.3	51.6
RGB · Otsu's	69.8	66.5	79.7	70.6	70.3	79.0	47.9	44.6	53.2
HS · Otsu's	66.7	63.1	79.0	63.9	64.4	77.3	54.5	48.1	60.0
CrCb · Otsu's	67.6	65.4	79.1	67.7	68.3	77.3	53.6	48.5	59.4
RGB · HS · CrCb	27.8	67.5	80.0	32.3	59.5	69.0	26.7	41.4	51.8
RGB · HS · CrCb · Otsu's	61.9	60.0	78.7	61.3	61.9	77.5	49.1	44.5	61.9
hist HSV	12.9	11.6	69.4	11.6	08.6	73.4	23.7	20.2	61.0
hist LAB	43.1	43.7	71.1	38.3	37.5	71.3	47.0	43.3	58.8
hist YCrCb	41.6	41.7	68.3	33.1	32.0	63.2	46.5	42.9	59.7
hist HCrCb	59.6	72.1	80.4	60.7	67.9	73.2	42.5	49.3	53.1
hist HCrCb · RGB	58.0	70.7	80.4	59.8	67.3	73.2	39.6	46.7	53.7
hist HCrCb · HS	54.6	65.9	80.0	53.5	59.9	73.2	39.7	46.8	55.7
hist HCrCb · CrCb	57.4	69.6	79.7	58.9	66.3	73.7	40.7	47.2	55.1
hist HCrCb · Otsu's	65.9	63.1	78.7	65.1	66.1	77.5	58.7	52.1	63.4

thresholding method	**THU_READ_RGBD** (158 images)			**THU_READ_RGBD_BODIES** (35 images)			**joint-ep-thu-ego** (103 images)		
	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut
baseline		81.7			49.0			40.3	
ROI mask	07.4	07.4	79.4	06.3	06.3	77.2	09.1	09.1	61.4
RGB	07.4	74.5	75.9	08.3	45.5	45.8	11.0	31.4	32.8
HS	06.7	71.0	76.7	07.0	48.9	49.7	12.7	48.1	46.4
CrCb	32.1	74.8	79.0	37.2	66.9	65.4	29.8	44.5	48.1
Otsu's	71.8	67.8	79.4	71.2	70.3	77.9	47.8	45.0	50.5
RGB · HS	06.5	69.7	76.8	06.9	48.8	50.1	11.8	44.6	46.5
RGB · CrCb	31.2	73.6	79.4	36.8	66.4	65.5	28.0	42.0	48.2
HS · CrCb	28.7	68.8	80.0	32.7	60.1	69.0	28.6	44.3	51.6
RGB · Otsu's	69.8	66.5	79.7	70.6	70.3	79.0	47.9	44.6	53.2
HS · Otsu's	66.7	63.1	79.0	63.9	64.4	77.3	54.5	48.1	60.0
CrCb · Otsu's	67.6	65.4	79.1	67.7	68.3	77.3	53.6	48.5	59.4
RGB · HS · CrCb	27.8	67.5	80.0	32.3	59.5	69.0	26.7	41.4	51.8
RGB · HS · CrCb · Otsu's	61.9	60.0	78.7	61.3	61.9	77.5	49.1	44.5	61.9
hist HSV	12.9	11.6	69.4	11.6	08.6	73.4	23.7	20.2	61.0
hist LAB	43.1	43.7	71.1	38.3	37.5	71.3	47.0	43.3	58.8
hist YCrCb	41.6	41.7	68.3	33.1	32.0	63.2	46.5	42.9	59.7
hist HCrCb	59.6	72.1	80.4	60.7	67.9	73.2	42.5	49.3	53.1
hist HCrCb · RGB	58.0	70.7	80.4	59.8	67.3	73.2	39.6	46.7	53.7
hist HCrCb · HS	54.6	65.9	80.0	53.5	59.9	73.2	39.7	46.8	55.7
hist HCrCb · CrCb	57.4	69.6	79.7	58.9	66.3	73.7	40.7	47.2	55.1
hist HCrCb · Otsu's	65.9	63.1	78.7	65.1	66.1	77.5	58.7	52.1	63.4

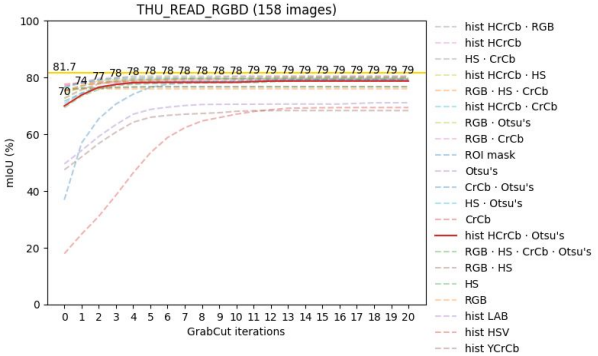
thresholding method	**THU_READ_RGBD** (158 images)			**THU_READ_RGBD_BODIES** (35 images)			**joint-ep-thu-ego** (103 images)		
	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut
baseline		81.7			49.0			40.3	
ROI mask	07.4	07.4	79.4	06.3	06.3	77.2	09.1	09.1	61.4
RGB	07.4	74.5	75.9	08.3	45.5	45.8	11.0	31.4	32.8
HS	06.7	71.0	76.7	07.0	48.9	49.7	12.7	48.1	46.4
CrCb	32.1	74.8	79.0	37.2	66.9	65.4	29.8	44.5	48.1
Otsu's	71.8	67.8	79.4	71.2	70.3	77.9	47.8	45.0	50.5
RGB · HS	06.5	69.7	76.8	06.9	48.8	50.1	11.8	44.6	46.5
RGB · CrCb	31.2	73.6	79.4	36.8	66.4	65.5	28.0	42.0	48.2
HS · CrCb	28.7	68.8	80.0	32.7	60.1	69.0	28.6	44.3	51.6
RGB · Otsu's	69.8	66.5	79.7	70.6	70.3	79.0	47.9	44.6	53.2
HS · Otsu's	66.7	63.1	79.0	63.9	64.4	77.3	54.5	48.1	60.0
CrCb · Otsu's	67.6	65.4	79.1	67.7	68.3	77.3	53.6	48.5	59.4
RGB · HS · CrCb	27.8	67.5	80.0	32.3	59.5	69.0	26.7	41.4	51.8
RGB · HS · CrCb · Otsu's	61.9	60.0	78.7	61.3	61.9	77.5	49.1	44.5	61.9
hist HSV	12.9	11.6	69.4	11.6	08.6	73.4	23.7	20.2	61.0
hist LAB	43.1	43.7	71.1	38.3	37.5	71.3	47.0	43.3	58.8
hist YCrCb	41.6	41.7	68.3	33.1	32.0	63.2	46.5	42.9	59.7
hist HCrCb	59.6	72.1	80.4	60.7	67.9	73.2	42.5	49.3	53.1
hist HCrCb · RGB	58.0	70.7	80.4	59.8	67.3	73.2	39.6	46.7	53.7
hist HCrCb · HS	54.6	65.9	80.0	53.5	59.9	73.2	39.7	46.8	55.7
hist HCrCb · CrCb	57.4	69.6	79.7	58.9	66.3	73.7	40.7	47.2	55.1
hist HCrCb · Otsu's	65.9	63.1	78.7	65.1	66.1	77.5	58.7	52.1	63.4

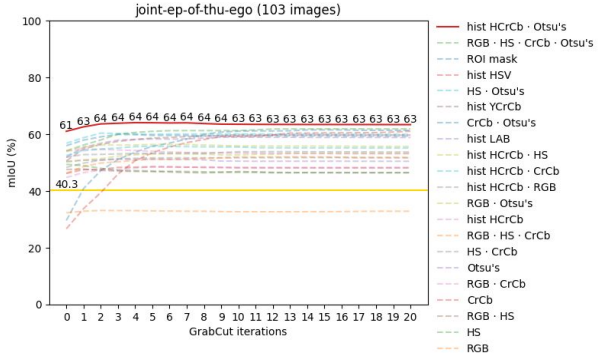
thresholding method	**THU_READ_RGBD** (158 images)			**THU_READ_RGBD_BODIES** (35 images)			**joint-ep-thu-ego** (103 images)		
	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut
baseline		81.7			49.0			40.3	
ROI mask	07.4	07.4	79.4	06.3	06.3	77.2	09.1	09.1	61.4
RGB	07.4	74.5	75.9	08.3	45.5	45.8	11.0	31.4	32.8
HS	06.7	71.0	76.7	07.0	48.9	49.7	12.7	48.1	46.4
CrCb	32.1	74.8	79.0	37.2	66.9	65.4	29.8	44.5	48.1
Otsu's	71.8	67.8	79.4	71.2	70.3	77.9	47.8	45.0	50.5
RGB · HS	06.5	69.7	76.8	06.9	48.8	50.1	11.8	44.6	46.5
RGB · CrCb	31.2	73.6	79.4	36.8	66.4	65.5	28.0	42.0	48.2
HS · CrCb	28.7	68.8	80.0	32.7	60.1	69.0	28.6	44.3	51.6
RGB · Otsu's	69.8	66.5	79.7	70.6	70.3	79.0	47.9	44.6	53.2
HS · Otsu's	66.7	63.1	79.0	63.9	64.4	77.3	54.5	48.1	60.0
CrCb · Otsu's	67.6	65.4	79.1	67.7	68.3	77.3	53.6	48.5	59.4
RGB · HS · CrCb	27.8	67.5	80.0	32.3	59.5	69.0	26.7	41.4	51.8
RGB · HS · CrCb · Otsu's	61.9	60.0	78.7	61.3	61.9	77.5	49.1	44.5	61.9
hist HSV	12.9	11.6	69.4	11.6	08.6	73.4	23.7	20.2	61.0
hist LAB	43.1	43.7	71.1	38.3	37.5	71.3	47.0	43.3	58.8
hist YCrCb	41.6	41.7	68.3	33.1	32.0	63.2	46.5	42.9	59.7
hist HCrCb	59.6	72.1	80.4	60.7	67.9	73.2	42.5	49.3	53.1
hist HCrCb · RGB	58.0	70.7	80.4	59.8	67.3	73.2	39.6	46.7	53.7
hist HCrCb · HS	54.6	65.9	80.0	53.5	59.9	73.2	39.7	46.8	55.7
hist HCrCb · CrCb	57.4	69.6	79.7	58.9	66.3	73.7	40.7	47.2	55.1
hist HCrCb · Otsu's	65.9	63.1	78.7	65.1	66.1	77.5	58.7	52.1	63.4

thresholding method	**THU_READ_RGBD** (158 images)			**THU_READ_RGBD_BODIES** (35 images)			**joint-ep-thu-ego** (103 images)		
	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut
baseline		81.7			49.0			40.3	
ROI mask	07.4	07.4	79.4	06.3	06.3	77.2	09.1	09.1	61.4
RGB	07.4	74.5	75.9	08.3	45.5	45.8	11.0	31.4	32.8
HS	06.7	71.0	76.7	07.0	48.9	49.7	12.7	48.1	46.4
CrCb	32.1	74.8	79.0	37.2	66.9	65.4	29.8	44.5	48.1
Otsu's	71.8	67.8	79.4	71.2	70.3	77.9	47.8	45.0	50.5
RGB · HS	06.5	69.7	76.8	06.9	48.8	50.1	11.8	44.6	46.5
RGB · CrCb	31.2	73.6	79.4	36.8	66.4	65.5	28.0	42.0	48.2
HS · CrCb	28.7	68.8	80.0	32.7	60.1	69.0	28.6	44.3	51.6
RGB · Otsu's	69.8	66.5	79.7	70.6	70.3	79.0	47.9	44.6	53.2
HS · Otsu's	66.7	63.1	79.0	63.9	64.4	77.3	54.5	48.1	60.0
CrCb · Otsu's	67.6	65.4	79.1	67.7	68.3	77.3	53.6	48.5	59.4
RGB · HS · CrCb	27.8	67.5	80.0	32.3	59.5	69.0	26.7	41.4	51.8
RGB · HS · CrCb · Otsu's	61.9	60.0	78.7	61.3	61.9	77.5	49.1	44.5	61.9
hist HSV	12.9	11.6	69.4	11.6	08.6	73.4	23.7	20.2	61.0
hist LAB	43.1	43.7	71.1	38.3	37.5	71.3	47.0	43.3	58.8
hist YCrCb	41.6	41.7	68.3	33.1	32.0	63.2	46.5	42.9	59.7
hist HCrCb	59.6	72.1	80.4	60.7	67.9	73.2	42.5	49.3	53.1
hist HCrCb · RGB	58.0	70.7	80.4	59.8	67.3	73.2	39.6	46.7	53.7
hist HCrCb · HS	54.6	65.9	80.0	53.5	59.9	73.2	39.7	46.8	55.7
hist HCrCb · CrCb	57.4	69.6	79.7	58.9	66.3	73.7	40.7	47.2	55.1
hist HCrCb · Otsu's	65.9	63.1	78.7	65.1	66.1	77.5	58.7	52.1	63.4

thresholding method	**THU_READ_RGBD** (158 images)			**THU_READ_RGBD_BODIES** (35 images)			**joint-ep-thu-ego** (103 images)		
	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut	mIoU (%) @ method	mIoU (%) @ clean up	mIoU (%) @ grabcut
baseline		81.7			49.0			40.3	
ROI mask	07.4	07.4	79.4	06.3	06.3	77.2	09.1	09.1	61.4
RGB	07.4	74.5	75.9	08.3	45.5	45.8	11.0	31.4	32.8
HS	06.7	71.0	76.7	07.0	48.9	49.7	12.7	48.1	46.4
CrCb	32.1	74.8	79.0	37.2	66.9	65.4	29.8	44.5	48.1
Otsu's	71.8	67.8	79.4	71.2	70.3	77.9	47.8	45.0	50.5
RGB · HS	06.5	69.7	76.8	06.9	48.8	50.1	11.8	44.6	46.5
RGB · CrCb	31.2	73.6	79.4	36.8	66.4	65.5	28.0	42.0	48.2
HS · CrCb	28.7	68.8	80.0	32.7	60.1	69.0	28.6	44.3	51.6
RGB · Otsu's	69.8	66.5	79.7	70.6	70.3	79.0	47.9	44.6	53.2
HS · Otsu's	66.7	63.1	79.0	63.9	64.4	77.3	54.5	48.1	60.0
CrCb · Otsu's	67.6	65.4	79.1	67.7	68.3	77.3	53.6	48.5	59.4
RGB · HS · CrCb	27.8	67.5	80.0	32.3	59.5	69.0	26.7	41.4	51.8
RGB · HS · CrCb · Otsu's	61.9	60.0	78.7	61.3	61.9	77.5	49.1	44.5	61.9
hist HSV	12.9	11.6	69.4	11.6	08.6	73.4	23.7	20.2	61.0
hist LAB	43.1	43.7	71.1	38.3	37.5	71.3	47.0	43.3	58.8
hist YCrCb	41.6	41.7	68.3	33.1	32.0	63.2	46.5	42.9	59.7
hist HCrCb	59.6	72.1	80.4	60.7	67.9	73.2	42.5	49.3	53.1
hist HCrCb · RGB	58.0	70.7	80.4	59.8	67.3	73.2	39.6	46.7	53.7
hist HCrCb · HS	54.6	65.9	80.0	53.5	59.9	73.2	39.7	46.8	55.7
hist HCrCb · CrCb	57.4	69.6	79.7	58.9	66.3	73.7	40.7	47.2	55.1
hist HCrCb · Otsu's	65.9	63.1	78.7	65.1	66.1	77.5	58.7	52.1	63.4

GrabCut iterations





Mask R-CNN epochs

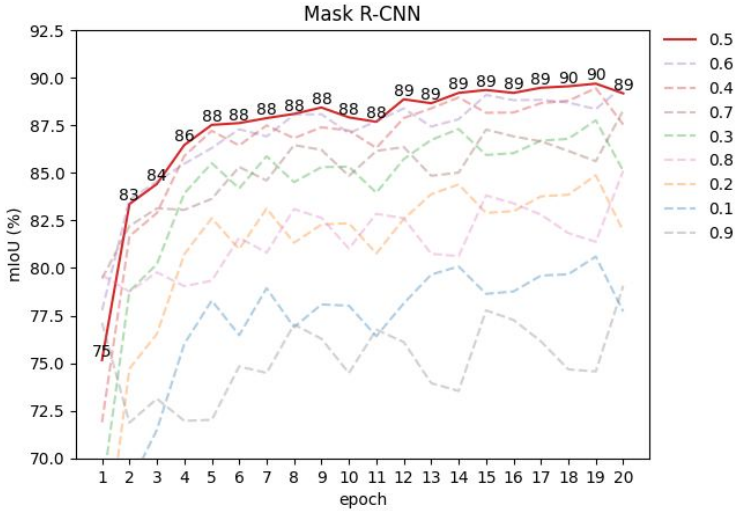
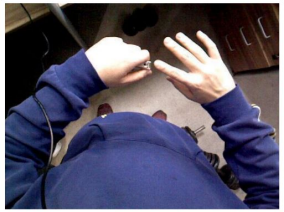
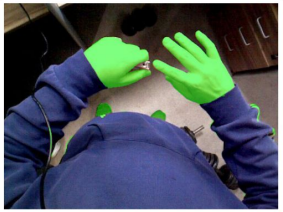
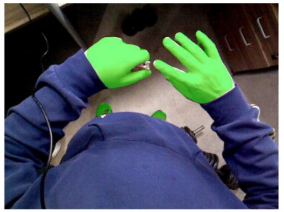
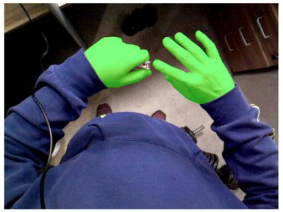
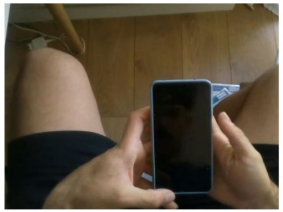
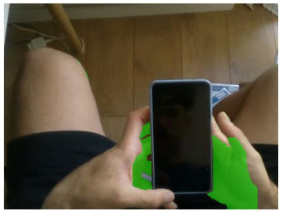
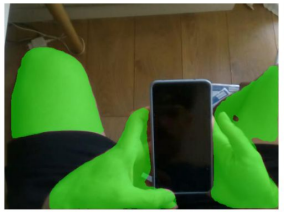
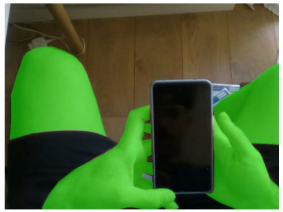


Image	hist HCrCb · Otsu's	Mask R-CNN	Label
			
			
			
			

Conclusion

Limitations

Limitations

- THU_READ_RGBD
  - 80% skin = body

Limitations

- THU_READ_RGBD
  - 80% skin = body
- in general
  - mistakes in labels
  - multiple images per scene

Limitations

- THU_READ_RGBD
  - 80% skin = body
- in general
  - mistakes in labels
  - multiple images per scene
- MediaPipe Hands

Limitations

- THU_READ_RGBD
  - 80% skin = body
- in general
  - mistakes in labels
  - multiple images per scene
- MediaPipe Hands
- Mask R-CNN results put into question

Conclusion

Conclusion

- skin segmentation for VR applications

Conclusion

- skin segmentation for VR applications
- color-based end-to-end system

Conclusion

- skin segmentation for VR applications
- color-based end-to-end system
- deep learning model

Conclusion

- skin segmentation for VR applications
- color-based end-to-end system
- deep learning model
- complimentary approach

Future Work

Future Work

- consider t-1 in video data

Future Work

- consider t-1 in video data
- less restrictive between mask and border

Future Work

- consider t-1 in video data
- less restrictive between mask and border
- if skin doesn’t touch border, k-means (k=2)

Future Work

- consider t-1 in video data
- less restrictive between mask and border
- if skin doesn’t touch border, k-means (k=2)
- additional, different clean up step after GrabCut

Future Work

- consider t-1 in video data
- less restrictive between mask and border
- if skin doesn’t touch border, k-means (k=2)
- additional, different clean up step after GrabCut
- investigate using semantic segmentation models

Future Work

- consider t-1 in video data
- less restrictive between mask and border
- if skin doesn’t touch border, k-means (k=2)
- additional, different clean up step after GrabCut
- investigate using semantic segmentation models
- more / higher quality datasets

Future Work

- consider t-1 in video data
- less restrictive between mask and border
- if skin doesn’t touch border, k-means (k=2)
- additional, different clean up step after GrabCut
- investigate using semantic segmentation models
- more / higher quality datasets
- generate initial background mask through other means (no labels)

Thank you for listening!