DensePose Challenge Intro:

Joint COCO and Mapillary Recognition Workshop

September 9th, ECCV 2018 Sunday

Rıza Alp Güler, INRIA/CentraleSupélec

Video Source: https://www.youtube.com/watch?v=Dhkd_bAwwMc

DensePose Team

Rıza Alp Güler

Natalia Neverova

Vasil Khalidov

Iasonas Kokkinos

CentraleSupélec

Inria
INVENTEURS DU MONDE NUMÉRIQUE

Human-body analysis: from coarse to fine

Input Image

Image Classification

Is there a person in this image?

Yes? **No?**

Image Classification

Human-body analysis: from coarse to fine

Input Image

Person Detection

Localize persons in the image.

Image Classification

Person Detection

Human-body analysis: from coarse to fine

Input Image

Person Segmentation

Segment person instances

Image Classification

Person Detection

Person Segmentation

Human-body analysis: from coarse to fine

Input Image

Part Segmentation

Segment semantically meaningful body parts.

Image Classification

Person Detection

Person Segmentation

Part Segmentation

Human-body analysis: from coarse to fine

Input Image

Pose Estimation

Localize joints of the persons in the images.

Image Classification

Person Detection

Person Segmentation

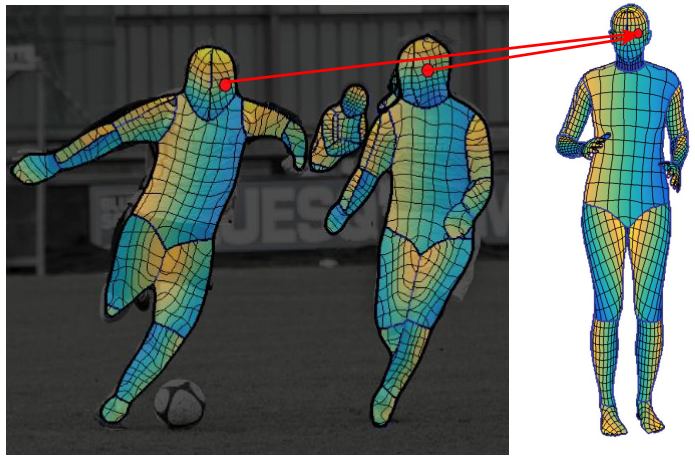
Part Segmentation

Pose Estimation

Human-body analysis: from coarse to fine

Input Image

Dense Pose Estimation



Find correspondence between all pixels and a 3D model.

Image Classification

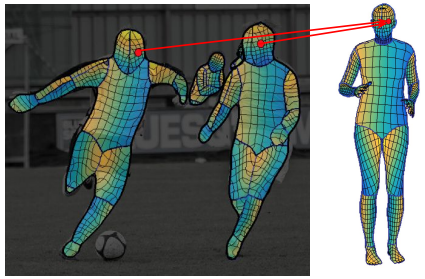
Person Detection

Person Segmentation

Part Segmentation

Pose Estimation

DensePose



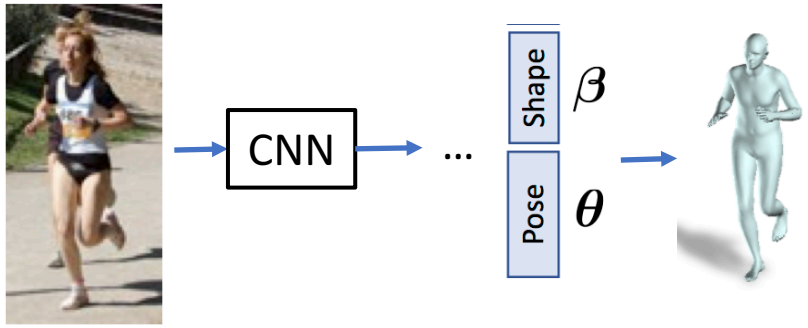
SMPL Model:

Loper et al. "SMPL: A skinned multi-person linear model." ACM Transactions on Graphics (TOG) 34.6 (2015): 248.

SMPLify:

Bogo et al. “Keep it SMPL: Automatic Estimation of 3D Human Pose and Shape from a Single Image” ECCV 2016

SMPL Parameter Regression:



Kanazawa et al. “End-to-end Recovery of Human Shape and Pose” CVPR 2018

Pavlakos et al. “Learning to Estimate 3D Human Pose and Shape from a Single Color Image” CVPR 2018

Dense Correspondences to SMPL model

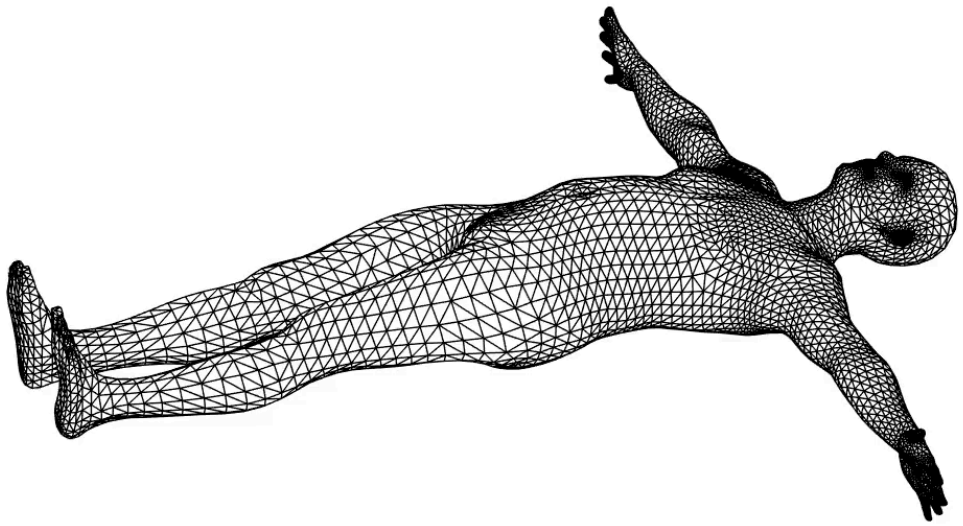
UP (model fitting)

SURREAL (synthetic)

DensePose (manual)

Unite the People Dataset (UP):
"Unite the people: Closing the loop between 3d and 2d human representations." Lassner, et al. (CVPR 2017)
SURREAL Dataset :
"Learning from synthetic humans" Varol, et al. (CVPR 2017)
DensePose-COCO Dataset :
"DensePose: Dense Human Pose Estimation" Guler, et al. (CVPR 2018)

Mesh Charting



simple parts

local UV coordinates

Video Source: https://www.youtube.com/watch?v=Dhkd_bAwwMc

Image-to-Surface correspondence

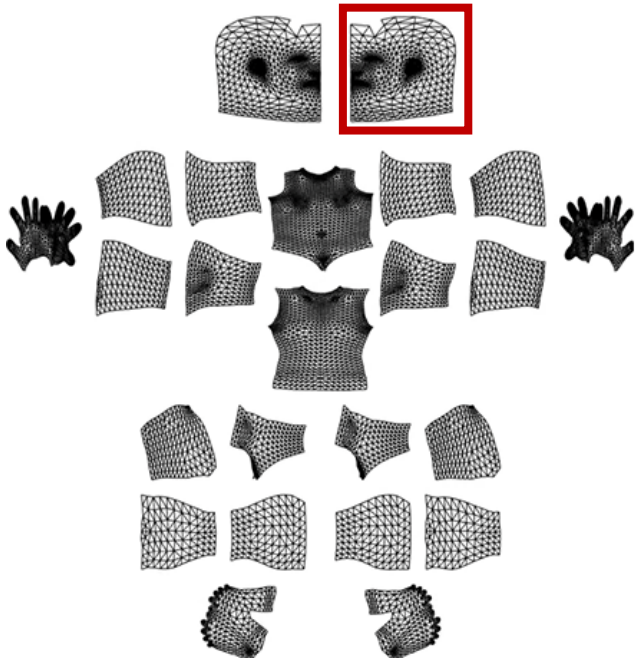
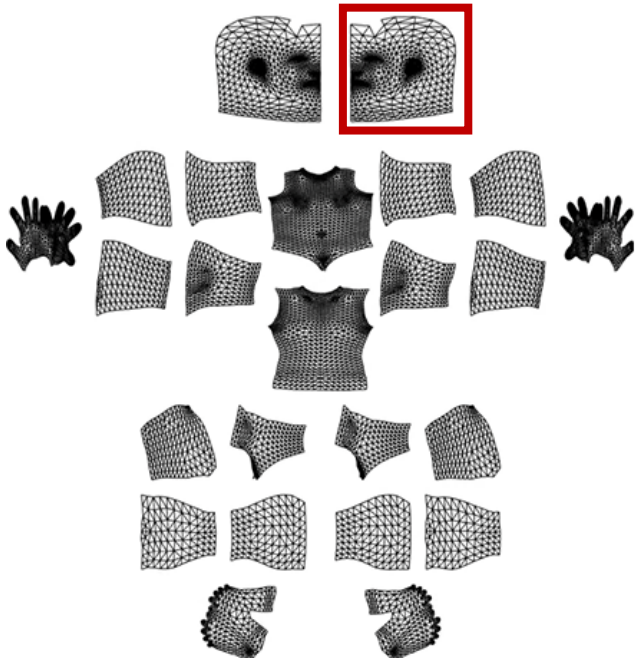


Image-to-Surface correspondence



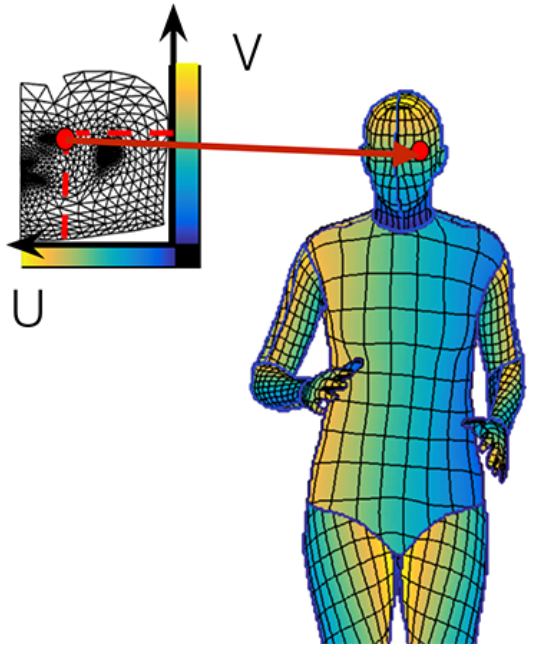
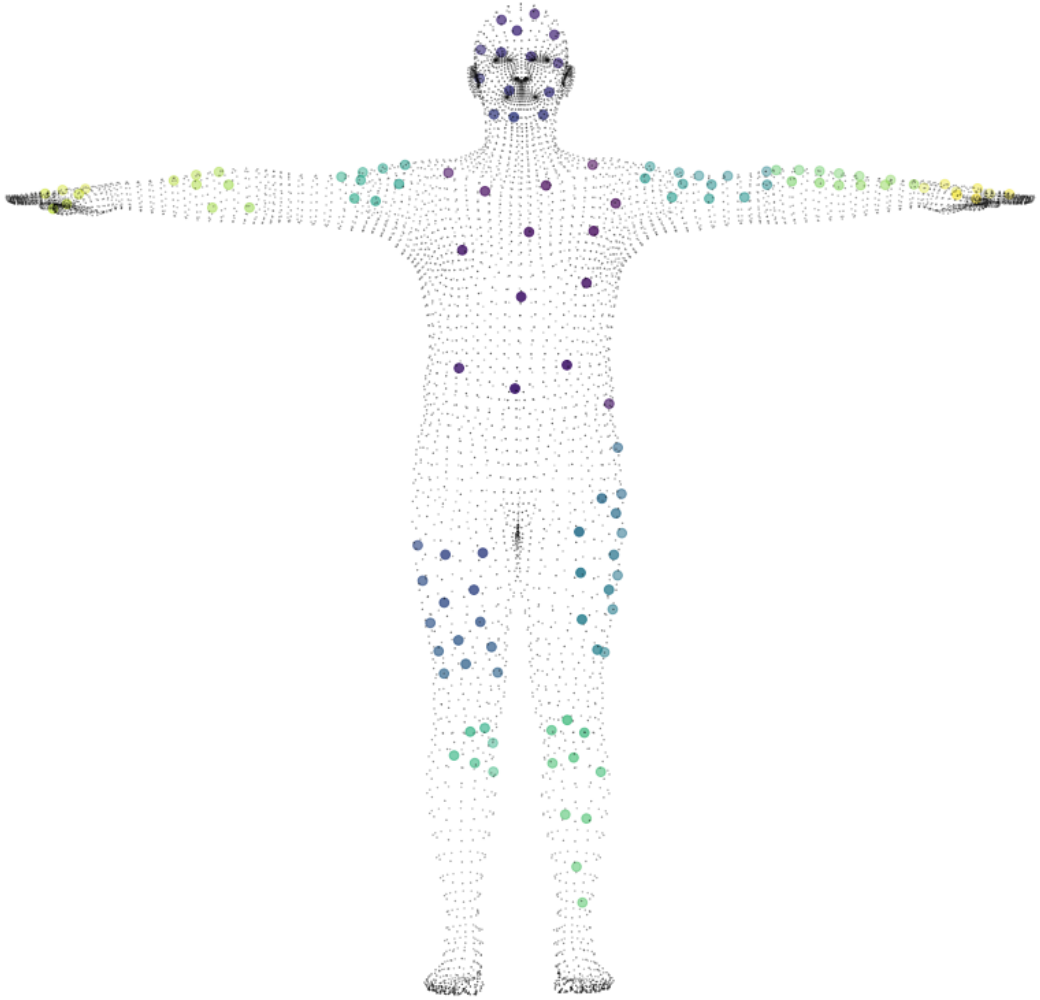
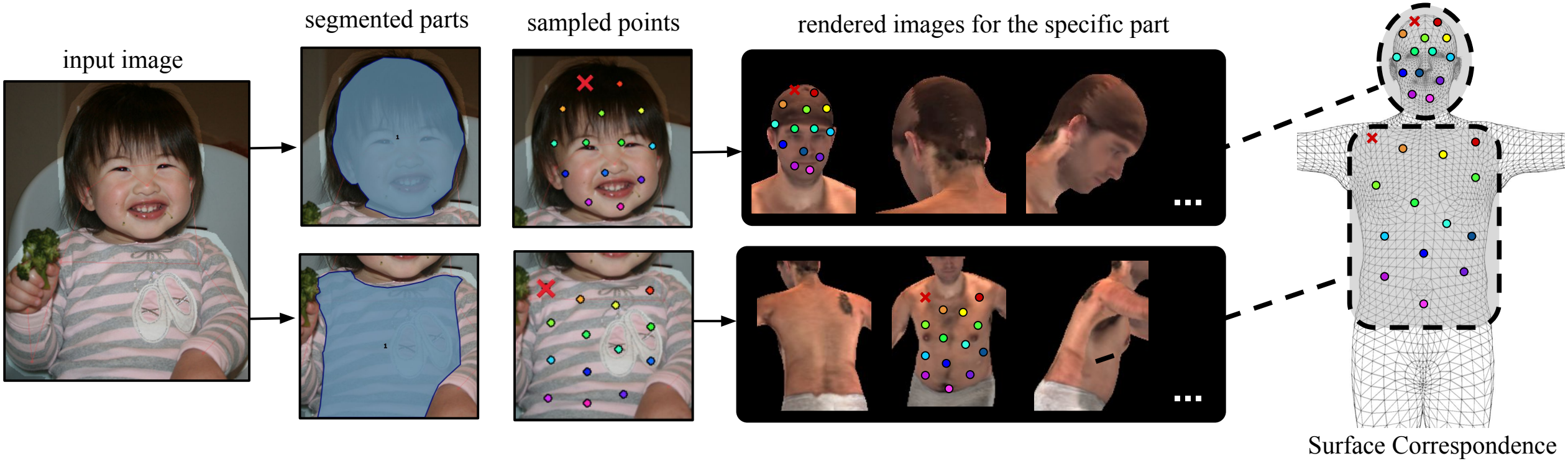


Image-to-Surface annotations



Annotation pipeline-I

Annotation pipeline-II



DensePose-COCO Dataset

densepose.org

Image U coordinates V coordinates

DensePose-COCO Dataset

densepose.org

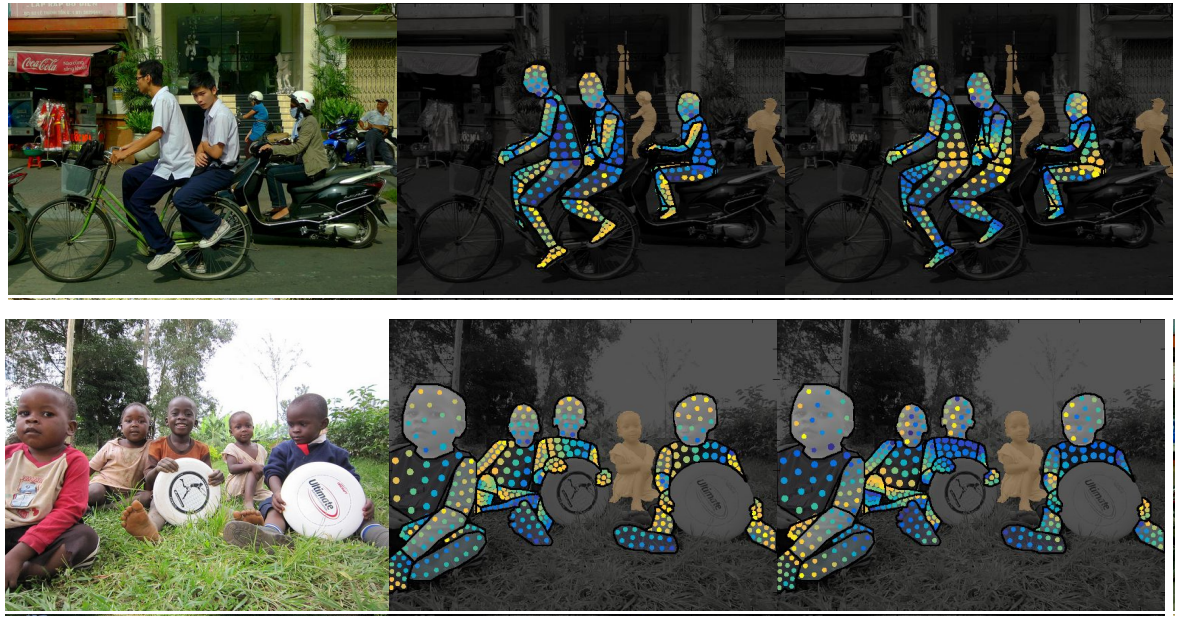
Image

U coordinates

V coordinates

densepose.org

DensePose-COCO dataset



DensePose-PoseTrack dataset

Posetrack Dataset:
Andriluka, Mykhaylo, et al. "Posetrack: A benchmark for human pose estimation and tracking." CVPR 2018.

https://github.com/facebookresearch/Densepose

Patch Indices

U coordinates

V coordinates

Points on the image

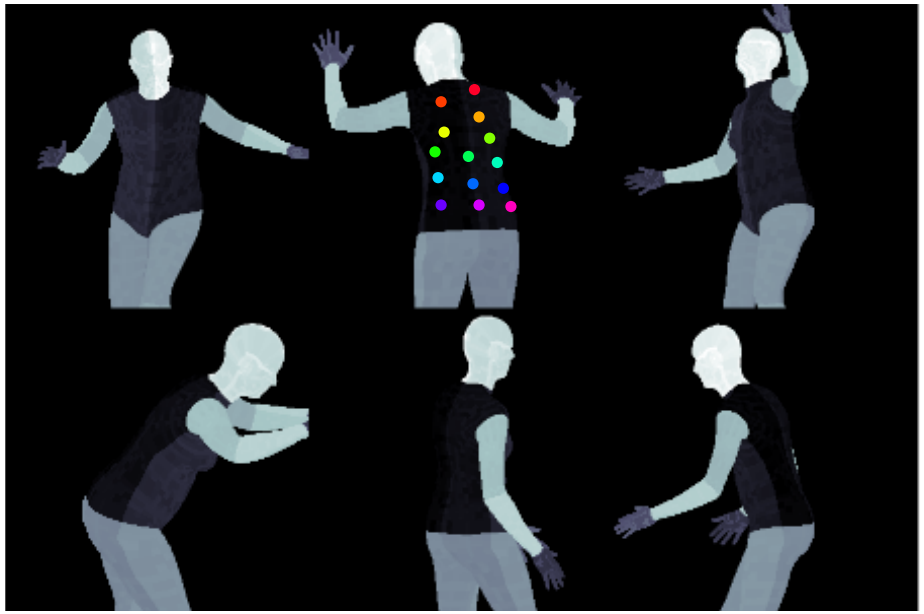
Points on the SMPL model

Points on the SMPL model

Annotator Performance

Rendered Image (SURREAL)

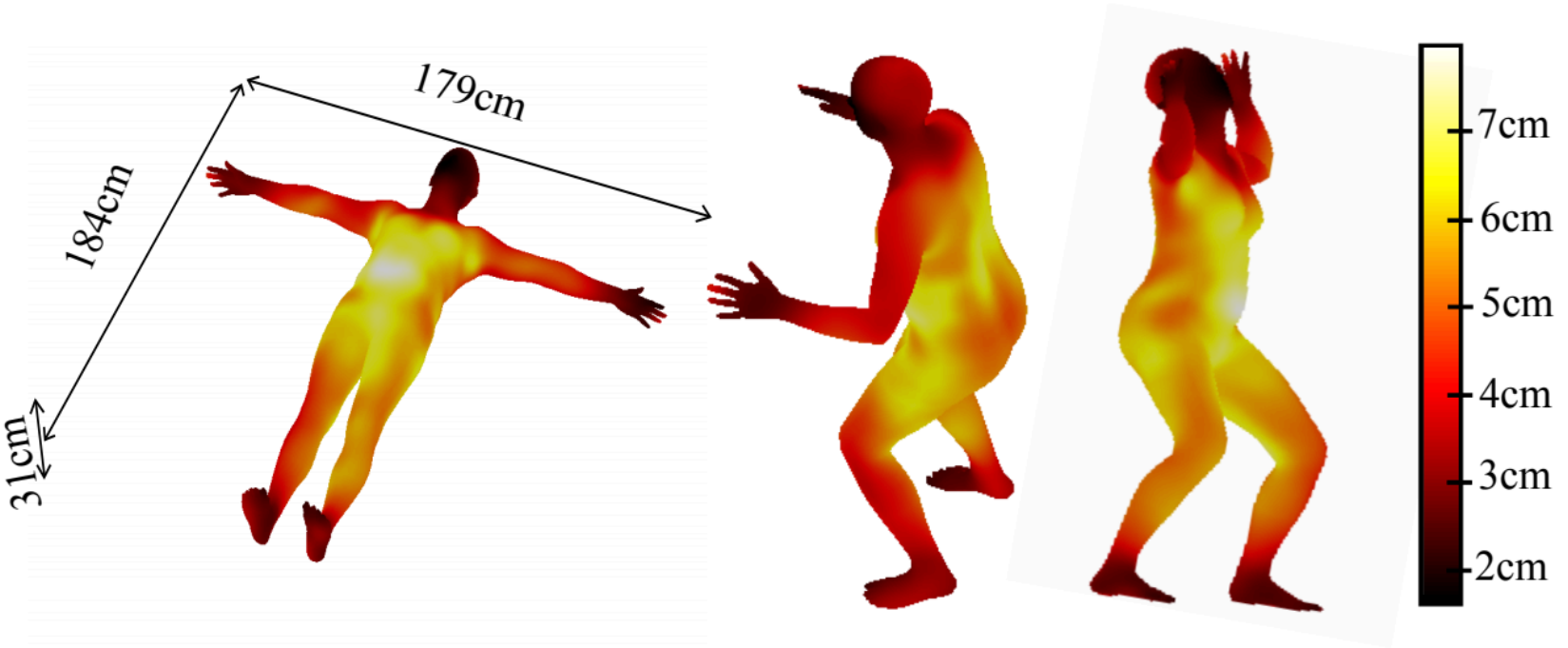
Sampled Points



Collected Points

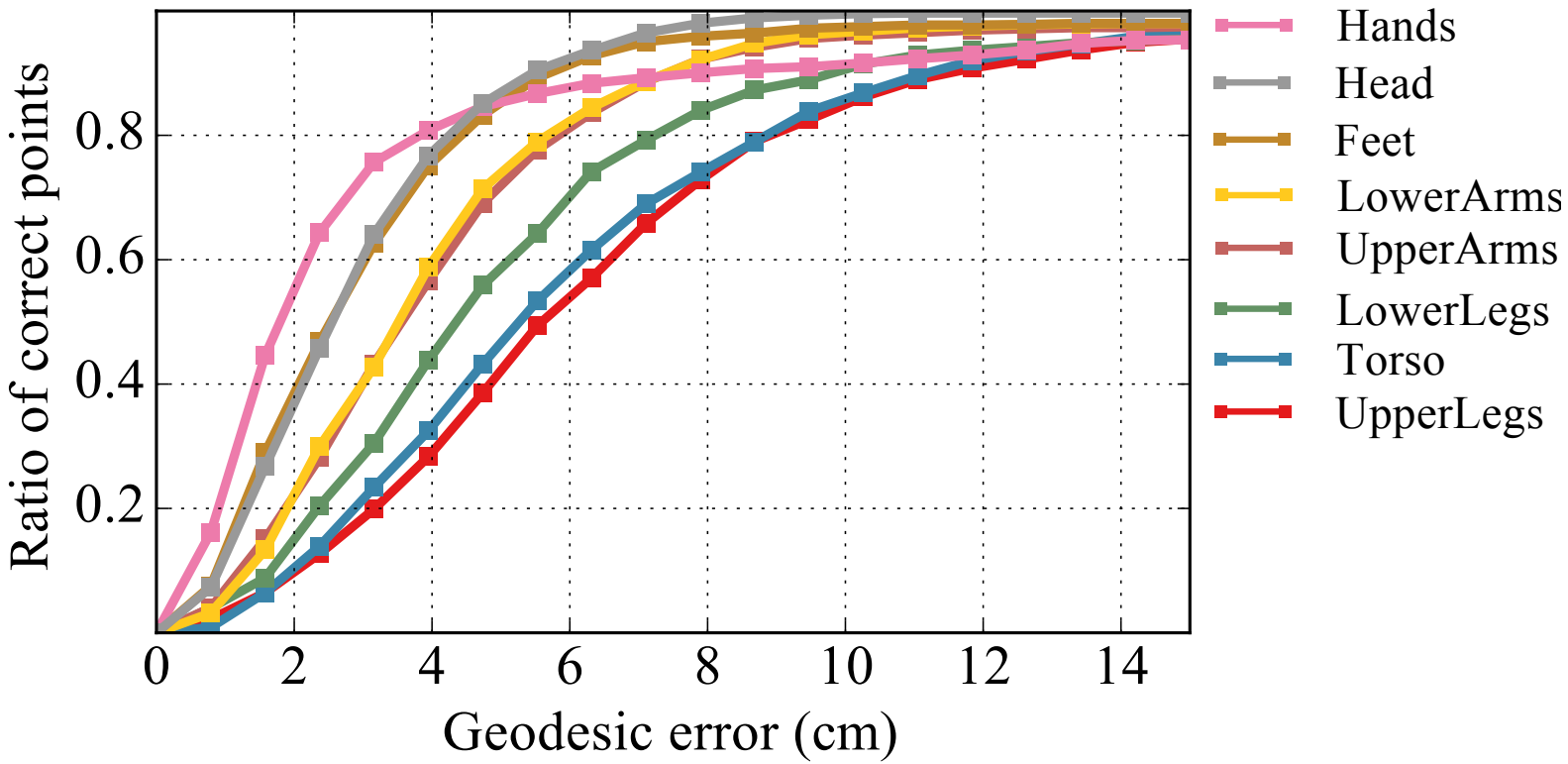
Geodesic distances

Annotator Performance



Average annotator error

Annotator Performance



Evaluation

Geodesic distances on the template

Geodesic Point Similarity (GPS) for instance based frameworks:

$$GPS = \frac{1}{|P|} \sum_{p_i \in P} \exp\left(\frac{-d(\hat{p}_i, p_i)^2}{2\kappa(p_i)^2} \right)$$

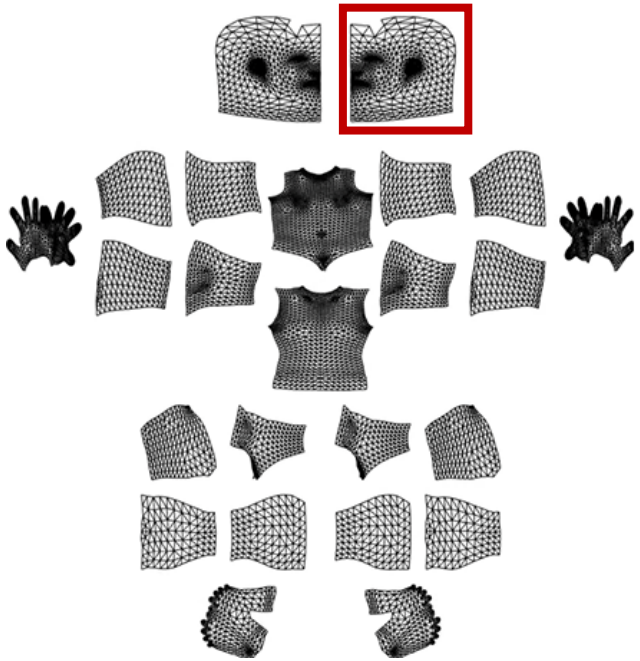
P → set of ground truth points annotated for a person

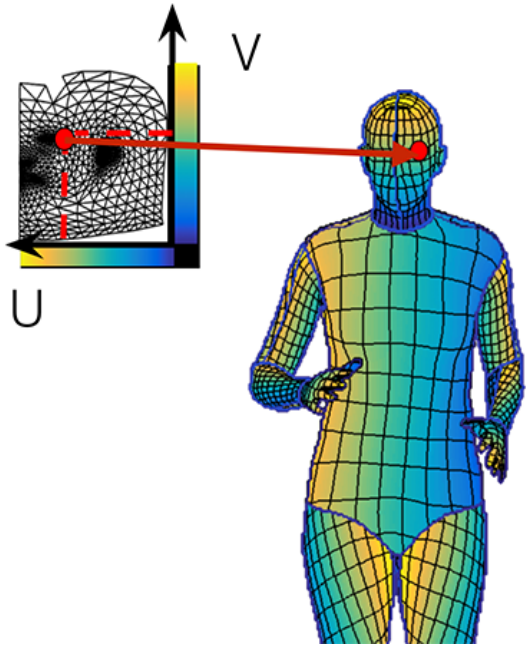
$d(\hat{p}_i, p_i)$ → Geodesic distance on the surface

$\kappa(p_i)$ → per-part normalization factor.

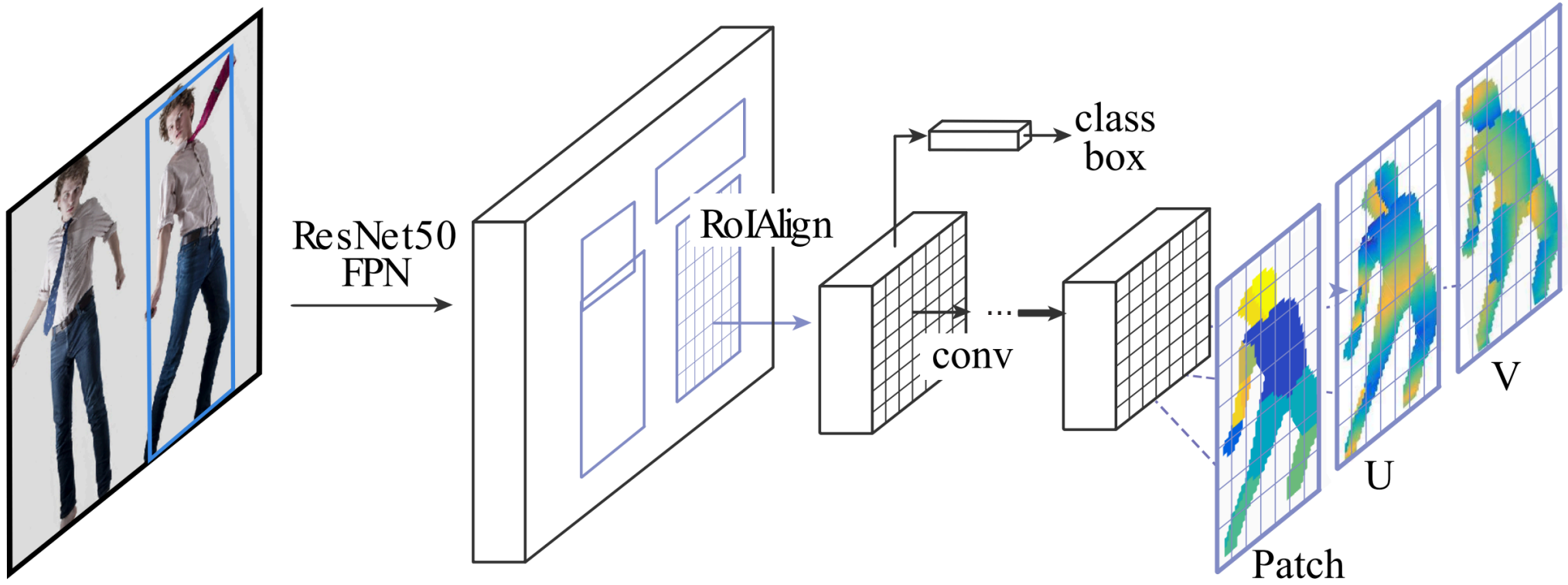
Measure AP between GPS = 0.5 - 0.95

Baseline: DensePose-RCNN

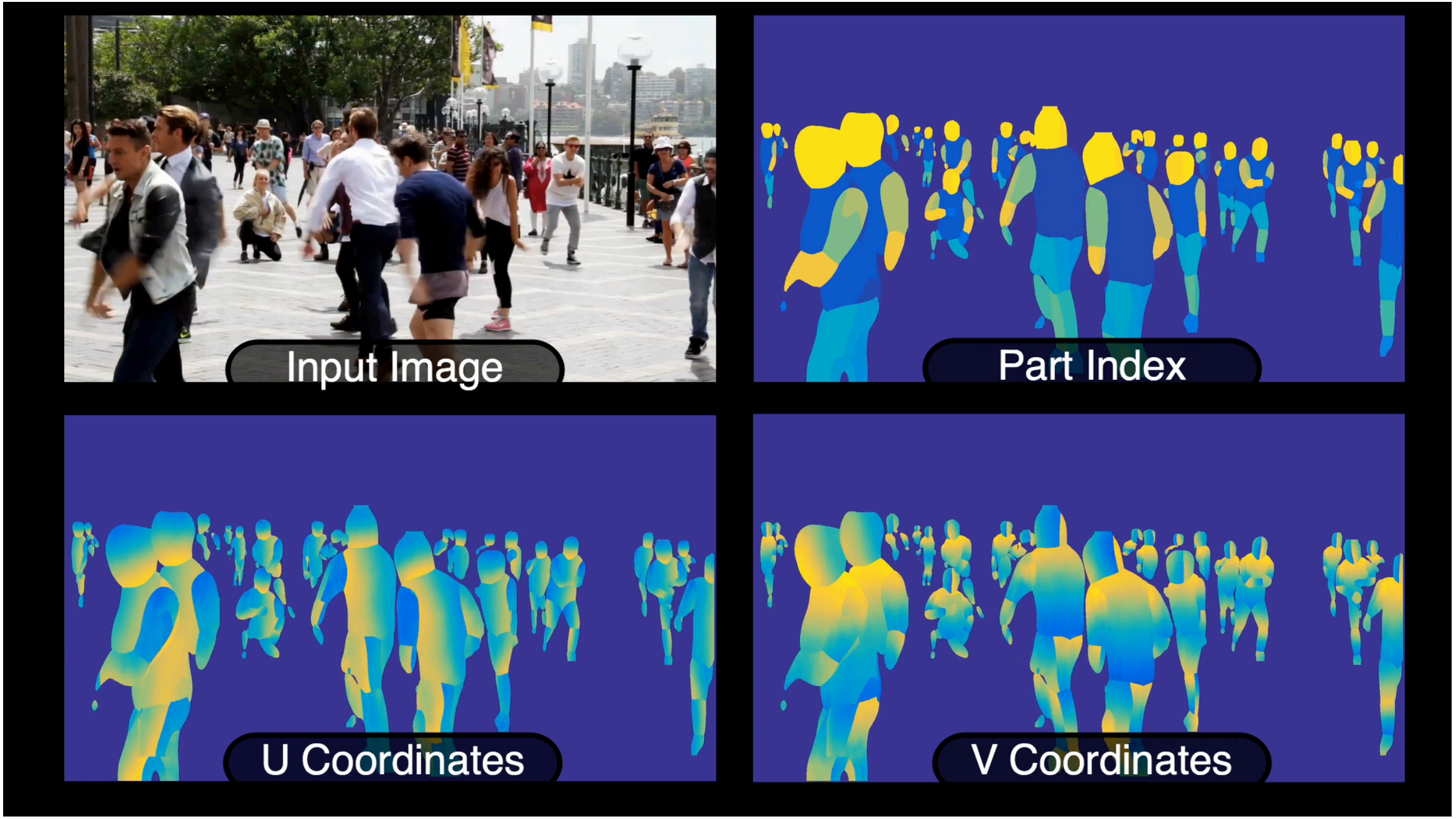




DensePose-RCNN Architecture



He, Kaiming, et al. "Mask R-CNN." *(ICCV 2017)*



Video Source: https://www.youtube.com/watch?v=Dhkd_bAwwMc

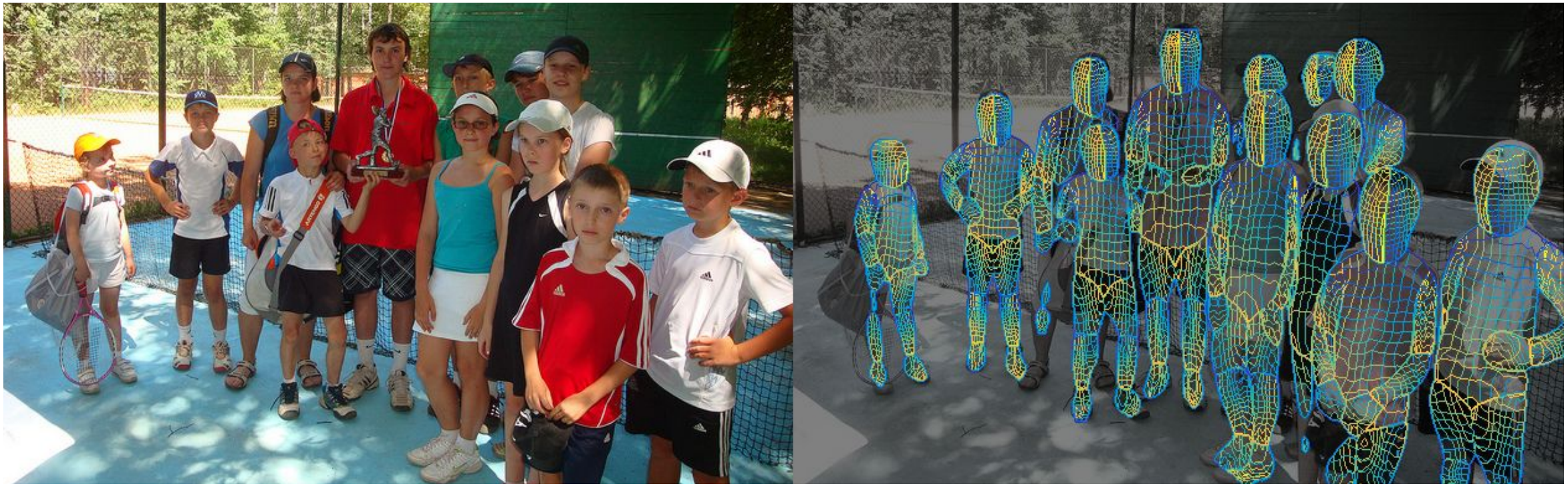
Baseline: DensePose-RCNN

DensePose-RCNN Model Zoo:

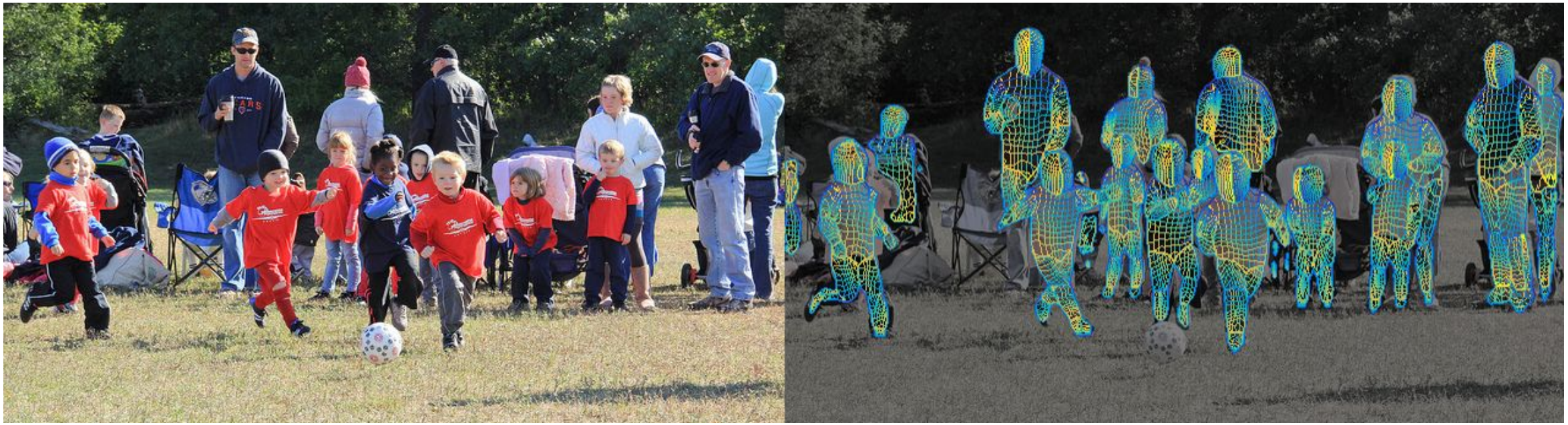
see: *github.com/facebookresearch/DensePose*

Model	AP	AP50	AP75	APm	APl
ResNet 50 + FPN	0.4892	0.8490	0.5078	0.4384	0.5059
ResNet 50 + FPN (mask, keypoints)	0.5075	0.8606	0.5373	0.4356	0.5265
ResNet101 + FPN	0.5147	0.8660	0.5601	0.4716	0.5291
ResNext101+ FPN	0.5554	0.8908	0.6080	0.5067	0.5676

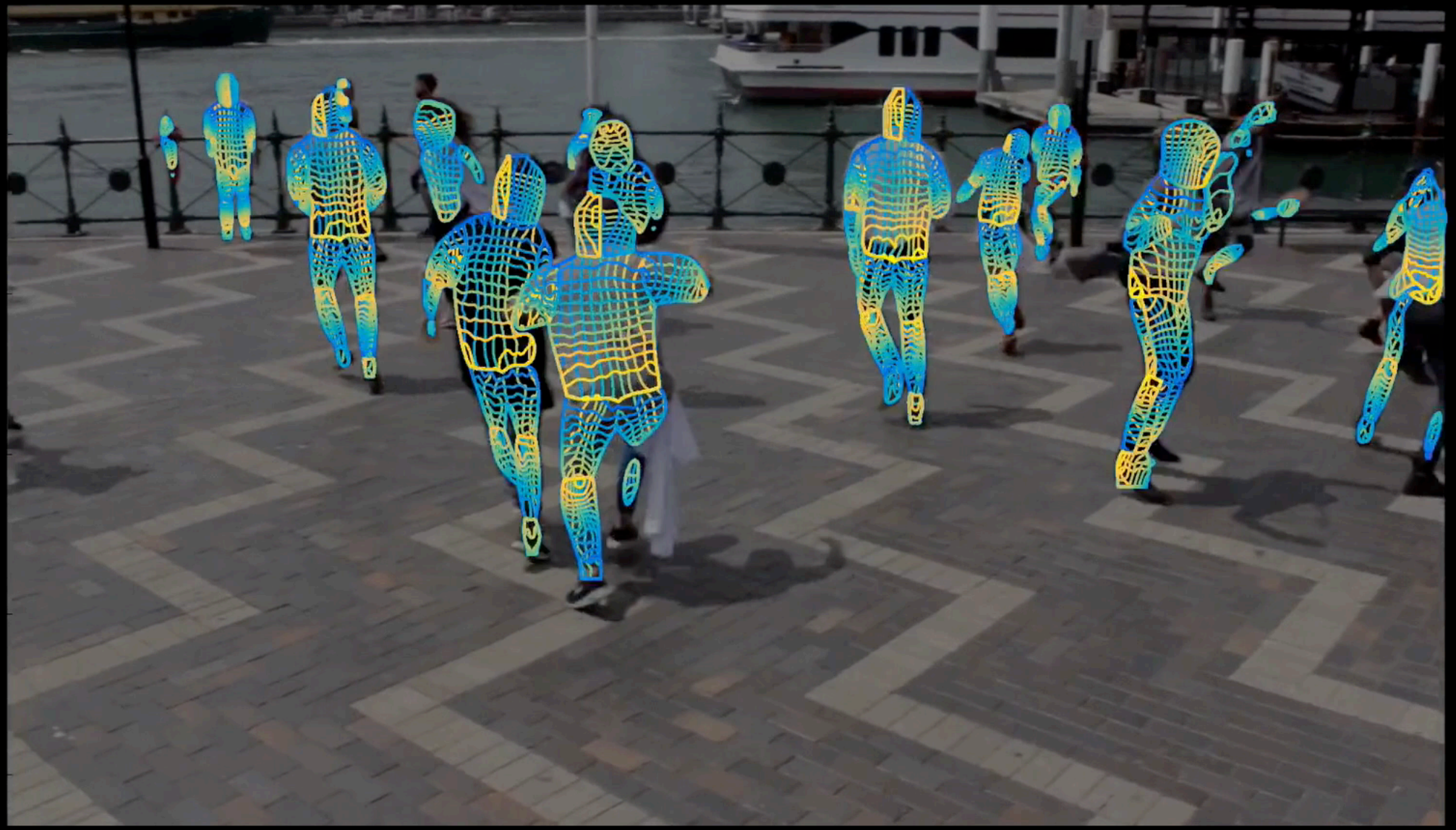
DensePose-RCNN Results



DensePose-RCNN Results



DensePose-RCNN Results



Video Source: https://www.youtube.com/watch?v=Dhkd_bAwwMc

Textures trom SURREAL dataset:
"Learning from synthetic humans”
Varol, Gül, et al. CVPR 2017

Winning entry

Parsing R-CNN

Team members:

Lu Yang (BUPT Priv Lab); Qing Song (BUPT Priv Lab); Zhihui Wang (BUPT Priv Lab)

Winner

Model	AP	AP50	AP75	APm	APl
ResNet 50 + FPN	0.4892	0.8490	0.5078	0.4384	0.5059
ResNet 50 + FPN (mask, keypoints)	0.5075	0.8606	0.5373	0.4356	0.5265
ResNet101 + FPN	0.5147	0.8660	0.5601	0.4716	0.5291
ResNext101+ FPN	0.5554	0.8908	0.6080	0.5067	0.5676
Sound of silent	0.57 (4)	0.87 (5)	0.66 (2)	0.48 (5)	0.61 (3)
ML_Lab	0.57 (3)	0.89 (3)	0.64 (4)	0.51 (3)	0.59 (4)
PlumSix	0.58 (2)	0.89 (2)	0.66 (3)	0.50 (4)	0.61 (2)
BUPT-PRIV	**0.64**	**0.92**	**0.75**	**0.57**	**0.67**

Baselines

Challenge entries

Winner!

DensePose Team

Rıza Alp Güler

Natalia Neverova

Vasil Khalidov

Iasonas Kokkinos

CentraleSupélec

Inria
INVENTEURS DU MONDE NUMÉRIQUE