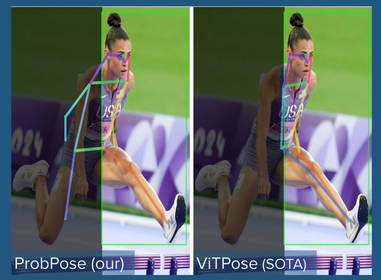
Do not forget out-of-image points when modeling people and evaluating pose

Try the online demo!



Takeaways:

- Standard evaluation protocols ignore “false positives”. We fix it.
- Models should predict whether a keypoint is out-of-image or not
- Probability maps are more versatile and interpretable than heatmaps while keeping SOTA localization

ProbPose: A Probabilistic Approach to 2D Human Pose Estimation

Miroslav Purkrabek and Jiri Matas
Visual Recognition Group, Czech Technical University in Prague

Out-of-image keypoints

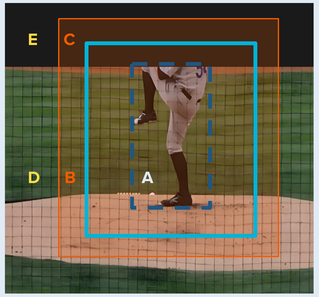
Out-of-image keypoints have been ignored by both models and metrics

➜ Models are not penalized for hallucinating points inside the image

	A	B	C	D	E
COCO train	96.2	3.5	0.0	0.2	0.0
COCO val	95.8	3.9	0.0	0.2	0.0
CropCOCO	68.8	2.2	23.5	0.1	5.3
OCHuman	99.2	0.8	0.0	0.0	0.0

The first dataset with out-of-image points

Areas where a keypoint could be located



bounding box

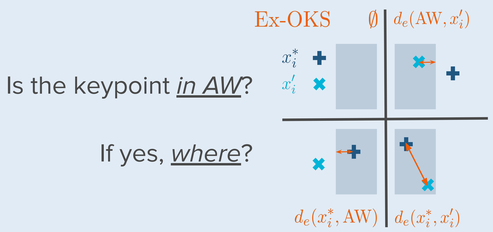
pose model input

activation window (AW)
= the range of joint locations
= the domain of the heatmap

Extended OKS

(= Object Keypoint Similarity)

Extension of OKS to a new out-of-image aware metric



Presence probability

Q: Is the point *in the AW?*

Trained with cross entropy on cropped images

Probability maps

(= probmaps, PMs)

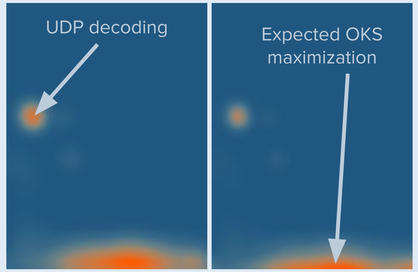
Q: If the point is in the AW, *where?*

Trained with a new pixel-wise OKS loss

$$R_{exp}(x_i) = \big(1 - OKS(x_i)\big) \cdot p_L(x_i)$$

$$\mathcal{L}_{OKS}(x_i) = (1-\alpha)R_{exp}(x_i) + \alpha g(x_i)$$

Location + Calibrated probabilistic estimate without Gaussian assumption



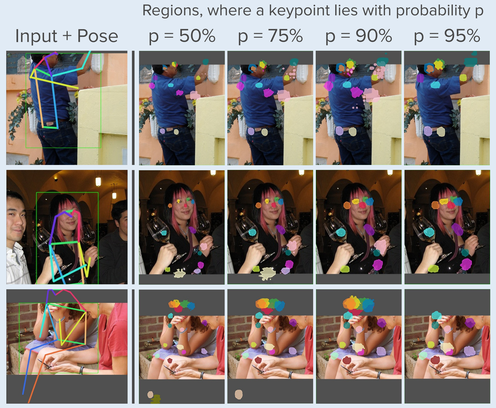
Decoding PMs maximizes expected OKS

$$\mathbb{E}_{OKS}(x_i) = \sum_{x_j \in PM} p_L(x_j) \cdot OKS(x_i, x_j)$$

Results

ProbPose predicts out-of-image keypoints
even better than GT!

ProbPose predicts calibrated, non-gaussian probability maps



In-image keypoints: SOTA on benchmarks
Out-of-image keypoints: superior to SOTA
Presence probability:
replaces confidence for in/out prediction

	COCO		CropCOCO	
	mAP	Ex-mAP	mAP	Ex-mAP
SWIN-t	73.5	72.9	71.3	65.0
ViTPose-s	75.9	75.3	72.7	66.5
ProbPose-s	76.6	76.4	81.7	73.9