IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields

David HURTADO
Zijie NING

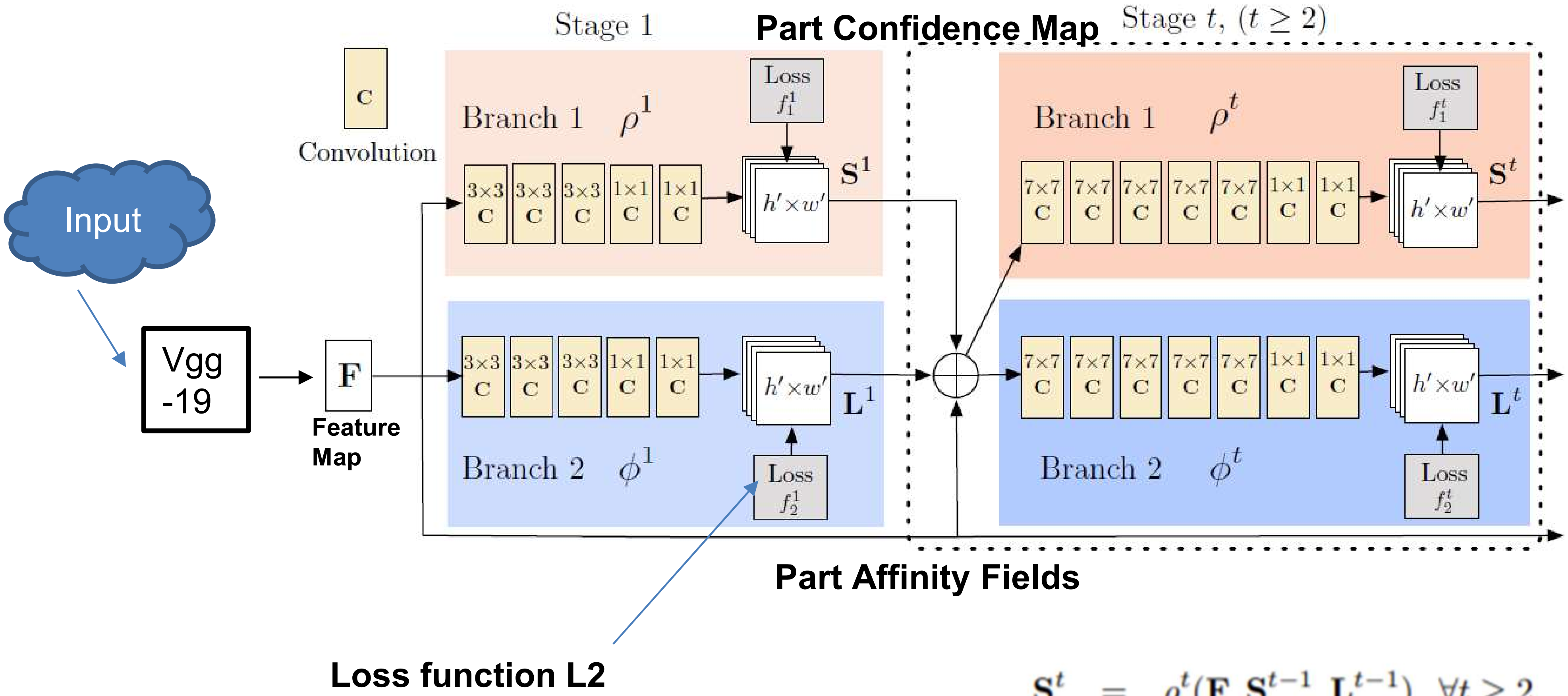
Interest

Inferring the pose of multiple people in images on real time.

Challenges:

- Unkown number of people (Position & Scale)
- People Interaction (Occlusion, spatial interference)
- Runtime Complexity

Architecture

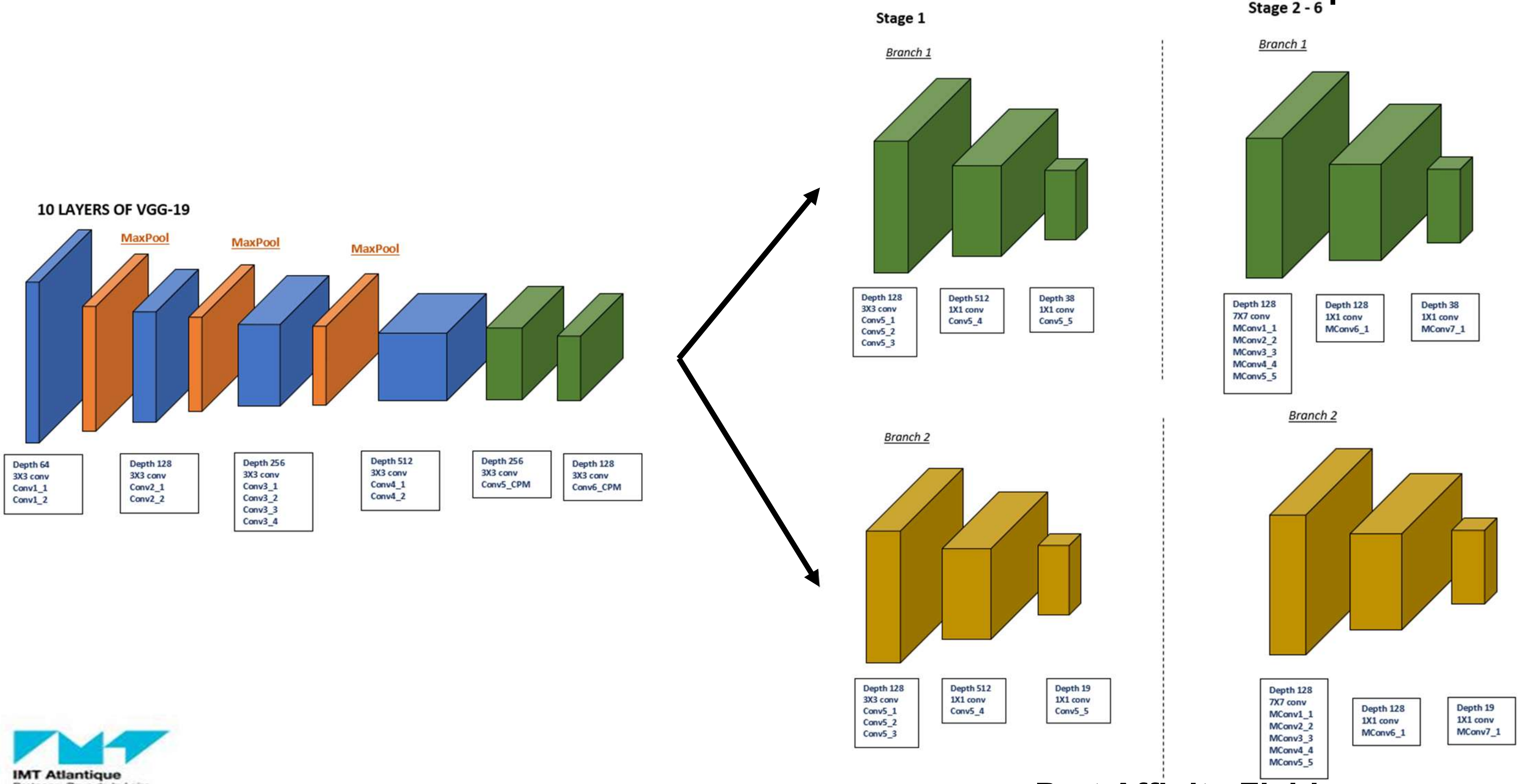


$$
\begin{aligned}
\mathbf{S}^t &= \rho^t(\mathbf{F}, \mathbf{S}^{t-1}, \mathbf{L}^{t-1}), \ \forall t \geq 2, \\
\mathbf{L}^t &= \phi^t(\mathbf{F}, \mathbf{S}^{t-1}, \mathbf{L}^{t-1}), \ \forall t \geq 2,
\end{aligned}
$$

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Architecture

Part Confidence Map



Part Affinity Fields

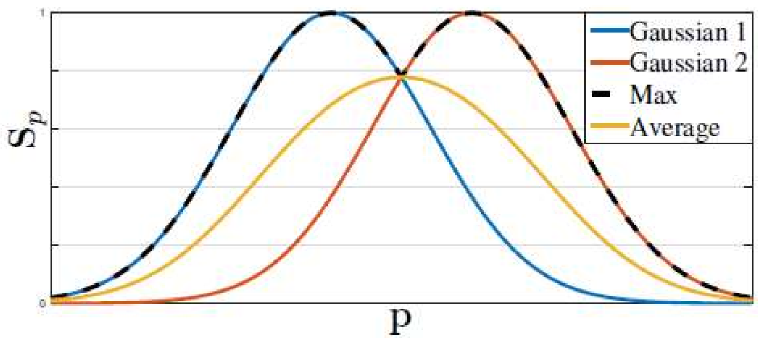
Part Confidence Map

$$\mathbf{S}^*_{j,k}(\mathbf{p}) = \exp\left(-\frac{||\mathbf{p} - \mathbf{x}_{j,k}||_2^2}{\sigma^2}\right)$$

A 2D representation of the belief that a particular body part occurs at each pixel location

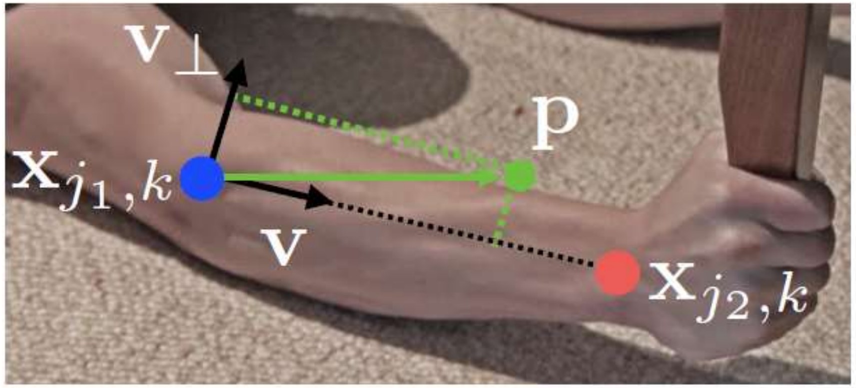
$$\mathbf{S}^*_j(\mathbf{p}) = \max_k \mathbf{S}^*_{j,k}(\mathbf{p}).$$

each visible part j for each person k



Part Affinity Fields

A field of 2D vectors encoding the direction that points from one joint of the limb to the other.



$$\mathbf{L}^*_{c,k}(\mathbf{p}) = \begin{cases} \mathbf{v} & \text{if } \mathbf{p} \text{ on limb } c,k \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbf{L}^*_c(\mathbf{p}) = \frac{1}{n_c(\mathbf{p})} \sum_k \mathbf{L}^*_{c,k}(\mathbf{p})$$

$$E = \int_{u=0}^{u=1} \mathbf{L}_c\left(\mathbf{p}(u)\right) \cdot \frac{\mathbf{d}_{j_2} - \mathbf{d}_{j_1}}{||\mathbf{d}_{j_2} - \mathbf{d}_{j_1}||_2} du.$$

two candidate part locations d_{j1} and d_{j2}

Identifying Limbs

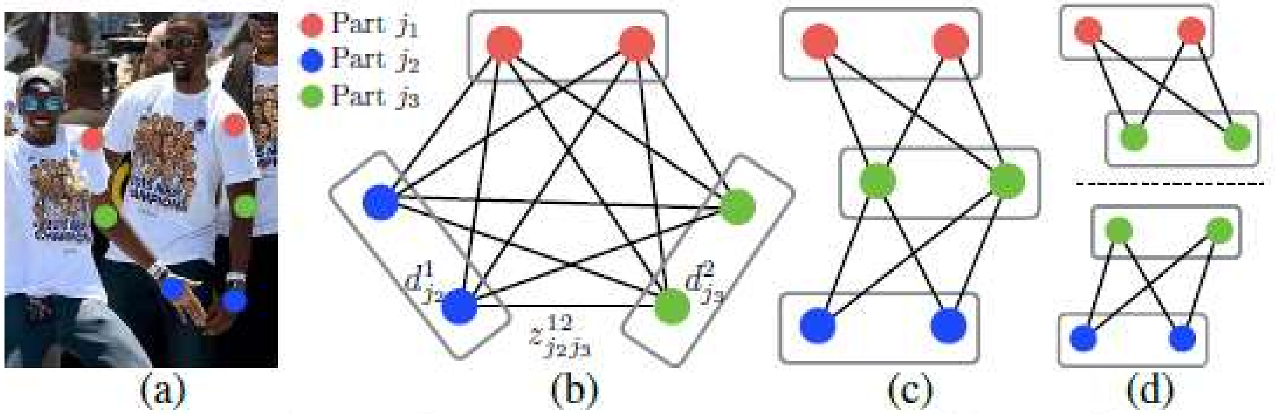


Figure 6. Graph matching. (a) Original image with part detections (b) K-partite graph (c) Tree structure (d) A set of bipartite graphs

b. The edges are all possible connections between pairs of detection candidates

c. No two limbs of the same type (e.g., left forearm) share a part.

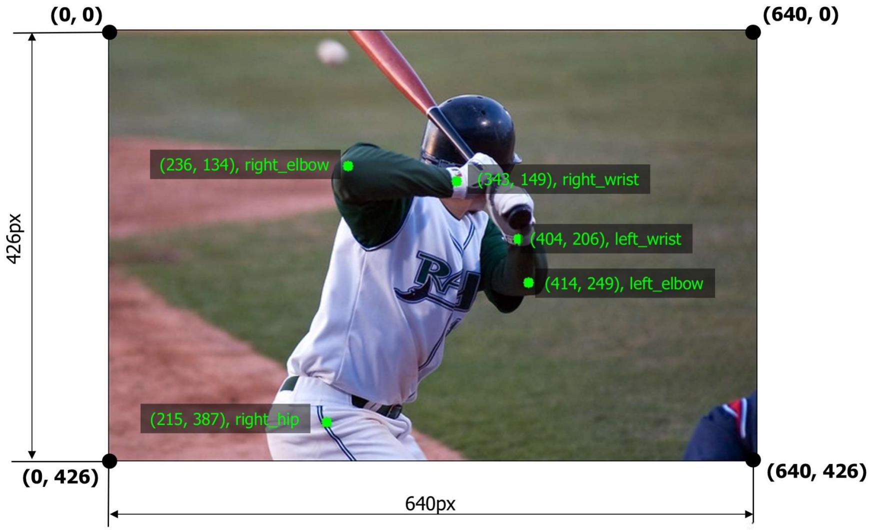
d. Decomposition

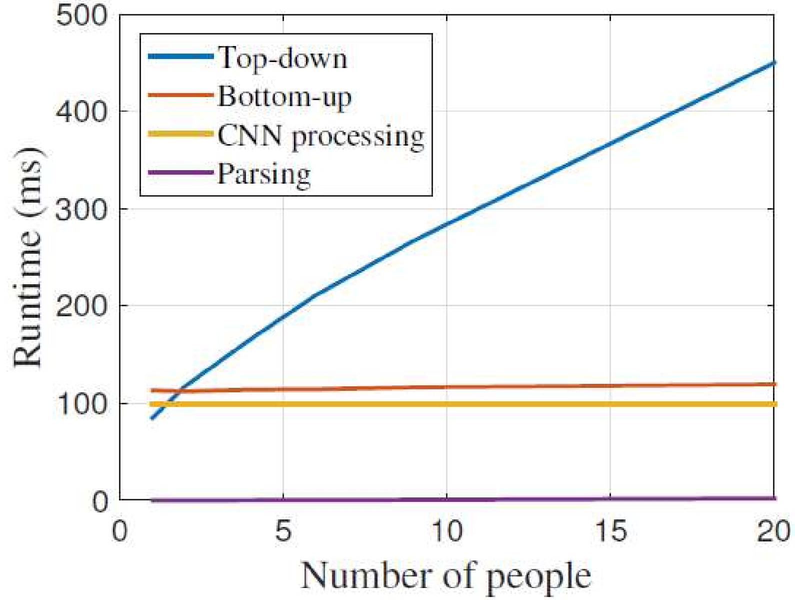
$$\max_{\mathcal{Z}} E = \sum_{c=1}^{C} \max_{\mathcal{Z}_c} E_c$$

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Test & Result

COCO Dataset & MPII MultiPerson Dataset



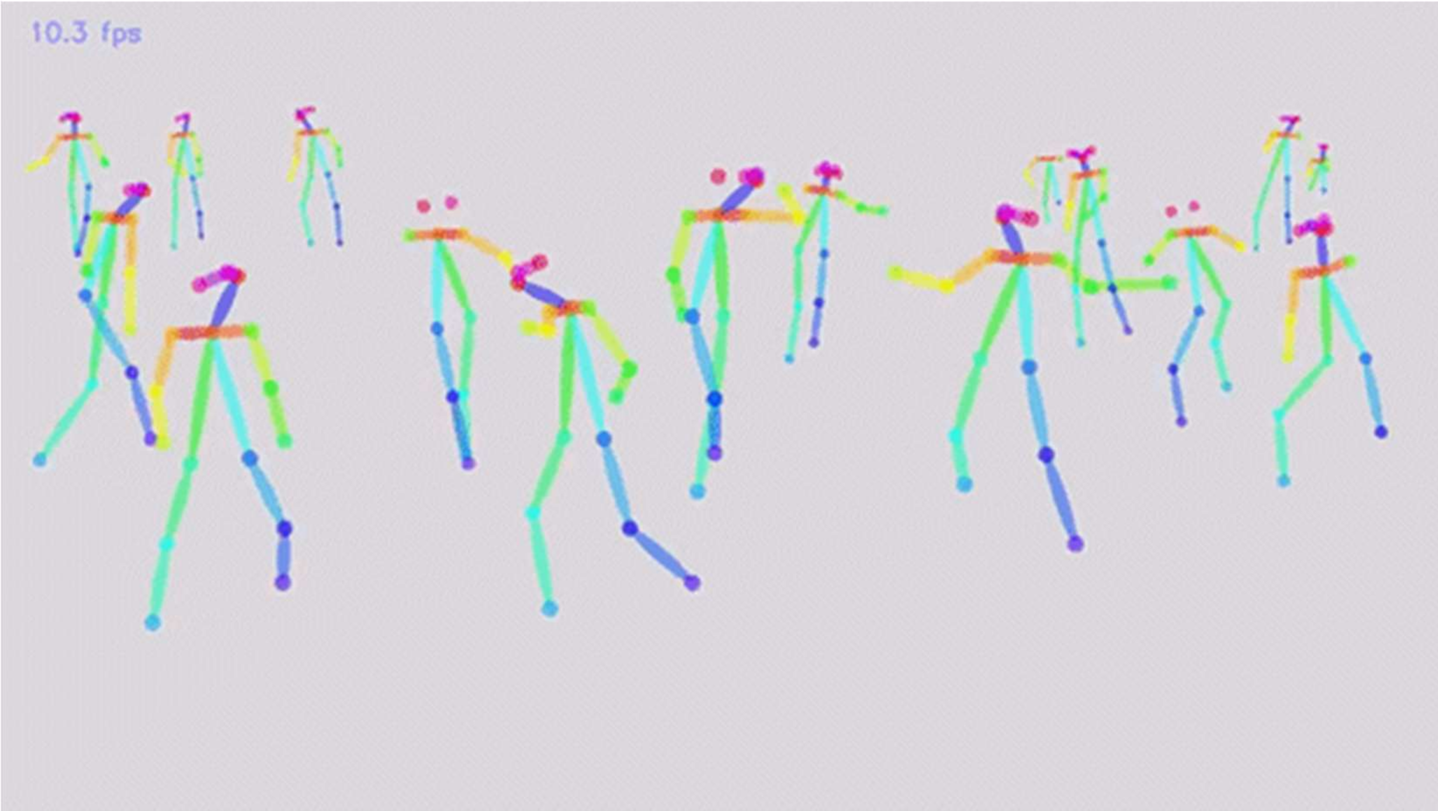


Team	AP	AP^{50}	AP^{75}	AP^{M}	AP^{L}
Test-challenge					
Ours	**60.5**	**83.4**	**66.4**	55.1	**68.1**
G-RMI [19]	59.8	81.0	65.1	**56.7**	66.7
DL-61	53.3	75.1	48.5	55.5	54.8
R4D	49.7	74.3	54.5	45.6	55.6
Test-dev					
Ours	**61.8**	**84.9**	**67.5**	57.1	**68.2**
G-RMI [19]	60.5	82.2	66.2	**57.6**	66.6
DL-61	54.4	75.3	50.9	58.3	54.3
R4D	51.4	75.0	55.9	47.4	56.7

Method	Hea	Sho	Elb	Wri	Hip	Kne	Ank	mAP	s/image
Subset of 288 images as in [22]									
Deepcut [22]	73.4	71.8	57.9	39.9	56.7	44.0	32.0	54.1	57995
Iqbal et al. [12]	70.0	65.2	56.4	46.1	52.7	47.9	44.5	54.7	10
DeeperCut [11]	87.9	84.0	71.9	63.9	68.8	63.8	58.1	71.2	230
Ours	**93.7**	**91.4**	**81.4**	**72.5**	**77.7**	**73.0**	**68.1**	**79.7**	**0.005**
Full testing set									
DeeperCut [11]	78.4	72.5	60.2	51.0	57.2	52.0	45.4	59.5	485
Iqbal et al. [12]	58.4	53.9	44.5	35.0	42.2	36.7	31.1	43.1	10
Ours (one scale)	89.0	84.9	74.9	64.2	71.0	65.6	58.1	72.5	0.005
Ours	**91.2**	**87.6**	**77.7**	**66.8**	**75.4**	**68.9**	**61.7**	**75.6**	**0.005**

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Test & Result

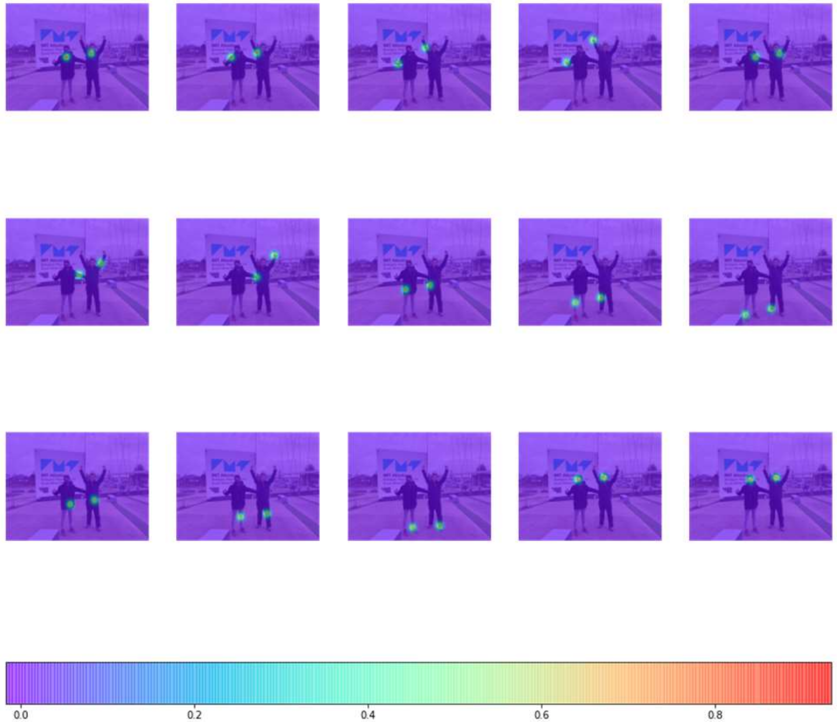


Our Work

1. Network Visualization

2. Webcam Integration

3. Pretrained Network Modification



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

https://github.com/mm0806son/2110D_DeepLearning

THANKS FOR YOUR ATTENTION

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom