

Supplying Material for SSN3D

Latest Revision: September 14, 2020

Given their is no page limitation, we present our figures and tables without saving space.

1 Table with citation

As we can see, AAAI's ucite label including author name and year and publication. This makes a long row for our table with citation. Due to page limitation, we dont show the ucite label in our paper, and the full version is as follows:

Methods	top-1	top-5	top-10
LFDA [Pedagadi et al. 2013]	32.9	68.5	82.2
KISSME [Koestinger et al. 2012]	36.5	67.8	78.8
LADF [Li et al. 2013]	39.0	76.8	89.0
STF3D [Liu et al. 2015]	44.3	71.7	83.7
TDL [You et al. 2016]	56.3	87.6	95.6
MARS [Zheng et al. 2016]	53.0	81.4	-
SeeForest [Zhou et al. 2017]	55.2	86.5	91.0
CNN+RNN [McLaughlin et al. 2016]	58.0	84.0	91.0
Seq-Decision [Zhang et al. 2018]	60.2	84.7	91.7
ASTPN [Xu et al. 2017]	62.0	86.0	94.0
QAN [Liu et al. 2017]	68.0	86.8	95.4
RQEN [Song et al. 2018]	77.1	93.2	97.7
STAN [Li et al. 2018]	80.2	-	-
Snippet [Chen et al. 2018]	79.8	91.8	-
Snippet+OF [Chen et al. 2018]	85.4	96.7	98.8
VRSTC [Hou et al. 2019]	83.4	95.5	97.7
AP3D [Gu et al. 2020]	86.7	-	-
SSN3D	88.9	97.3	98.8

Table 1: Comparison with related methods on iLIDS-VID

Methods	top-1	top-5	top-10	mAP
Mars [Zheng et al. 2016]	68.3	82.6	89.4	49.3
SeeForest [Zhou et al. 2017]	70.6	90.0	97.6	50.7
Seq-Decision [Zhang et al. 2018]	71.2	85.7	91.8	-
Latent Parts [Li et al. 2017]	71.8	86.6	93.0	56.1
QAN [Liu et al. 2017]	73.7	84.9	91.6	51.7
K-reciprocal [Zhong et al. 2017]	73.9	-	-	68.5
RQEN [Song et al. 2018]	77.8	88.8	94.3	71.7
TriNet [Hermans et al. 2017]	79.8	91.3	-	67.7
EUG [Wu et al. 2018]	80.8	92.1	96.1	67.4
STAN [Li et al. 2018]	82.3	-	-	65.8
Snippet [Chen et al. 2018]	81.2	92.1	-	69.4
Snippet+OF [Chen et al. 2018]	86.3	94.7	98.2	76.1
VRSTC [Hou et al. 2019]	88.5	96.5	97.4	82.3
AP3D [Gu et al. 2020]	90.1	-	-	85.1
SSN3D	87.3	94.6	97.8	78.1

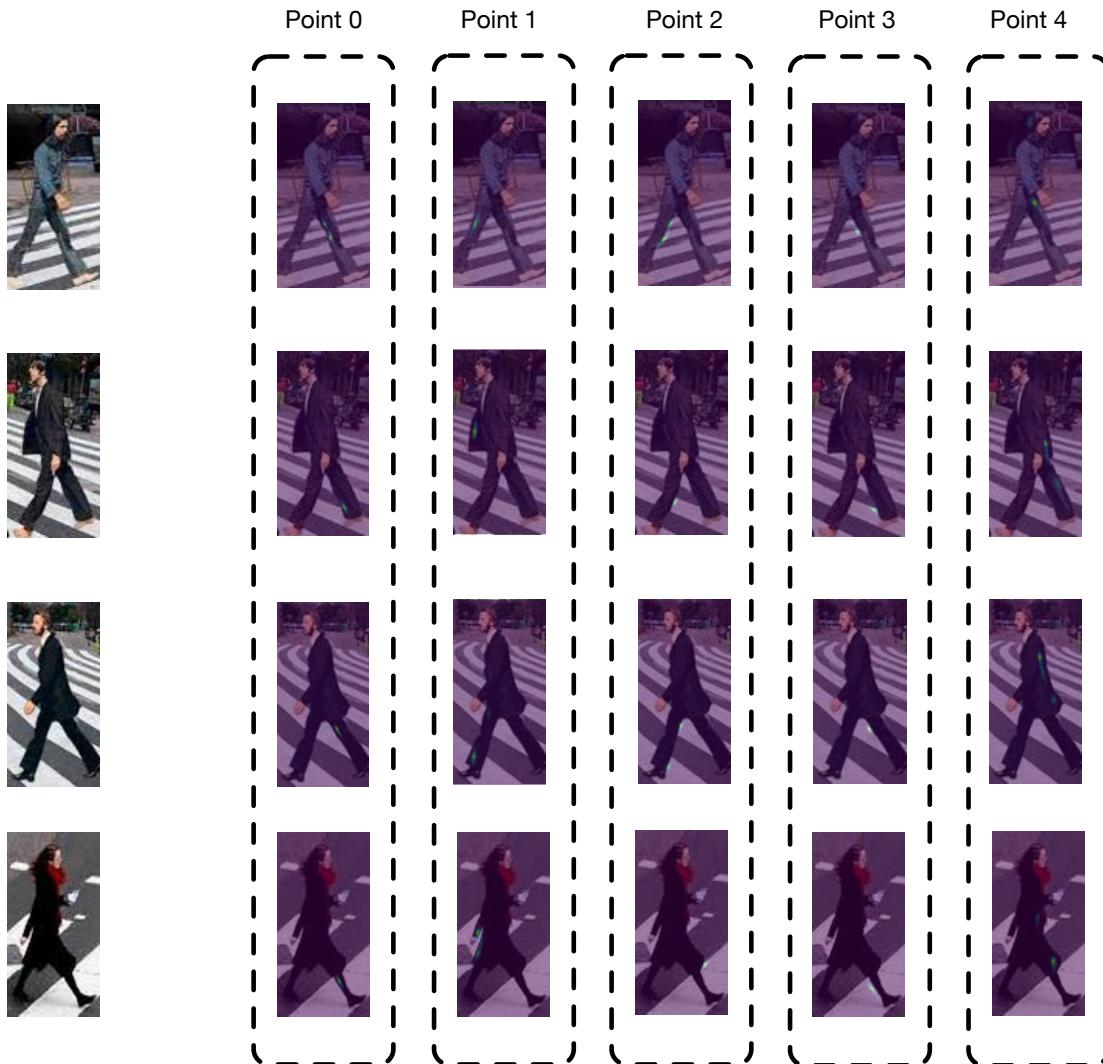
Table 2: Comparison with related methods on MARS

Methods	top-1	top-5	top-10	mAP
EUG [Wu et al. 2018]	83.6	94.6	97.6	78.3
VRSTC [Hou et al. 2019]	95.0	99.1	99.4	93.5
AP3D [Gu et al. 2020]	96.3	-	-	95.6
SSN3D	96.8	98.6	99.4	96.3

Table 3: Comparison with related methods on DukeMTMC

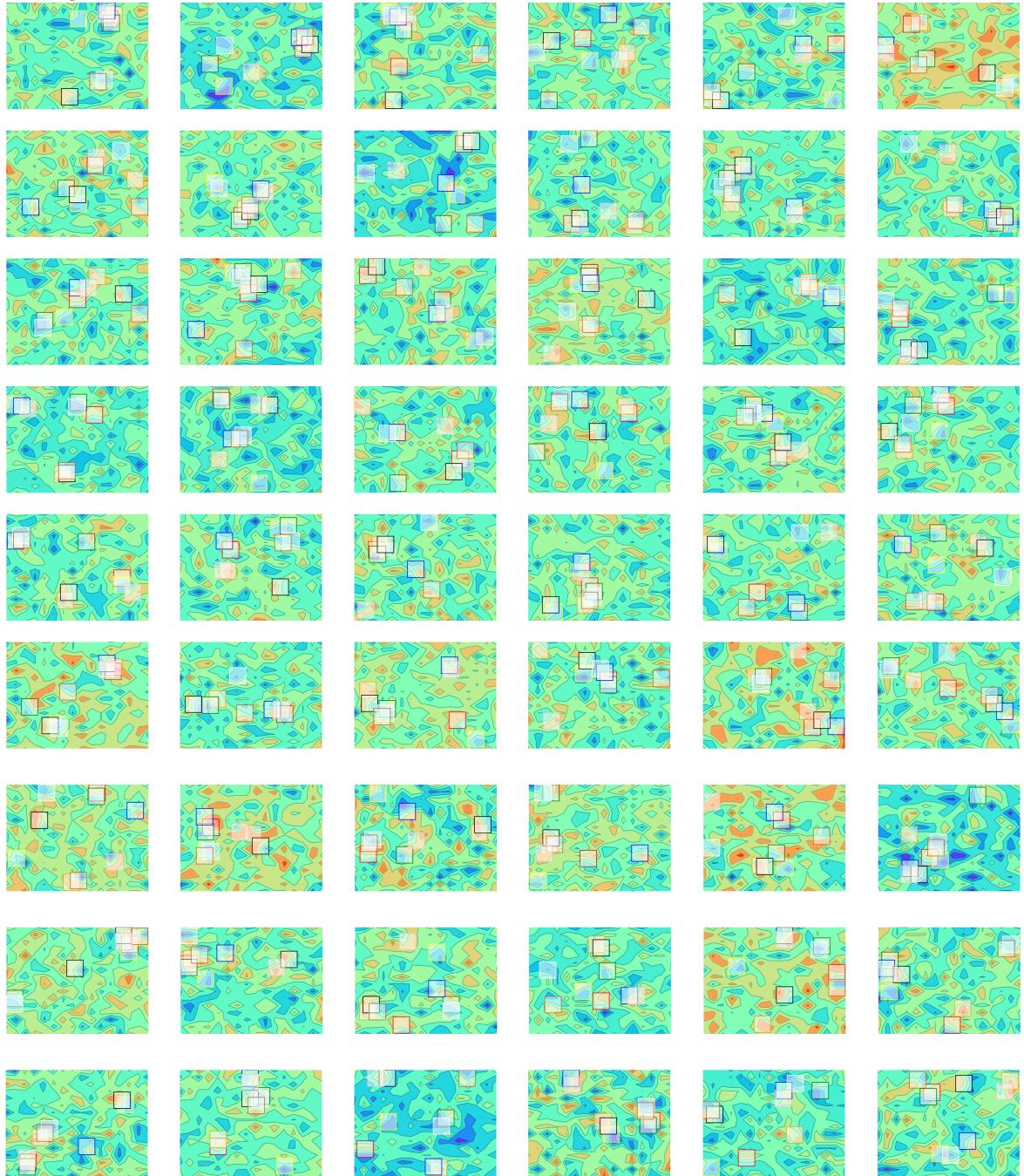
2 More visualized figure on Pedstrain

Figure 1: SSN trained in unsupervised strategy on Pedstrain128. Compared to SSN with semi-supervised learning, SSN selects the discriminative parts, but it is less accurate or stable. 1) Selected parts are discriminative as most of them are contours of person. 2) In some pictures, multiple parts with similar pixel value within a frame are selected, which shows the deficiency of unsupervised learning handling similar parts. 3) for some person, the selected part are not always stable, parts of legs and clothes are selected from multiple frames.



3 More visualized figure on Amber Abstract

Figure 2: More of Amber Abstract: SSN is so powerful to extract patterns. The selected parts are sharing some similarities.



References

- Dapeng Chen, Hongsheng Li, Tong Xiao, Shuai Yi, and Xiaogang Wang. Video person re-identification with competitive snippet-similarity aggregation and co-attentive snippet embedding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1169–1178, 2018.
- Xinqian Gu, Hong Chang, Bingpeng Ma, Hongkai Zhang, and Xilin Chen. Appearance-preserving 3d convolution for video-based person re-identification. In *ECCV*, 2020.
- Alexander Hermans, Lucas Beyer, and Bastian Leibe. In defense of the triplet loss for person re-identification. *arXiv preprint arXiv:1703.07737*, 2017.
- Ruibing Hou, Bingpeng Ma, Hong Chang, Xinqian Gu, Shiguang Shan, and Xilin Chen. Vrstc: Occlusion-free video person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7183–7192, 2019.
- Martin Koestinger, Martin Hirzer, Paul Wohlhart, Peter M Roth, and Horst Bischof. Large scale metric learning from equivalence constraints. In *2012 IEEE conference on computer vision and pattern recognition*, pages 2288–2295. IEEE, 2012.
- Dangwei Li, Xiaotang Chen, Zhang Zhang, and Kaiqi Huang. Learning deep context-aware features over body and latent parts for person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 384–393, 2017.
- Shuang Li, Slawomir Bak, Peter Carr, and Xiaogang Wang. Diversity regularized spatiotemporal attention for video-based person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 369–378, 2018.
- Zhen Li, Shiyu Chang, Feng Liang, Thomas S Huang, Liangliang Cao, and John R Smith. Learning locally-adaptive decision functions for person verification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3610–3617, 2013.
- Kan Liu, Bingpeng Ma, Wei Zhang, and Rui Huang. A spatio-temporal appearance representation for video-based pedestrian re-identification. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3810–3818, 2015.
- Yu Liu, Junjie Yan, and Wanli Ouyang. Quality aware network for set to set recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5790–5799, 2017.
- Niall McLaughlin, Jesus Martinez del Rincon, and Paul Miller. Recurrent convolutional network for video-based person re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1325–1334, 2016.
- Sateesh Pedagadi, James Orwell, Sergio Velastin, and Boghos Boghossian. Local fisher discriminant analysis for pedestrian re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3318–3325, 2013.
- Guanglu Song, Biao Leng, Yu Liu, Congrui Hetang, and Shaofan Cai. Region-based quality estimation network for large-scale person re-identification. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
- Yu Wu, Yutian Lin, Xuanyi Dong, Yan Yan, Wanli Ouyang, and Yi Yang. Exploit the unknown gradually: One-shot video-based person re-identification by stepwise learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5177–5186, 2018.
- Shuangjie Xu, Yu Cheng, Kang Gu, Yang Yang, Shiyu Chang, and Pan Zhou. Jointly attentive spatial-temporal pooling networks for video-based person re-identification. In *Proceedings of the IEEE international conference on computer vision*, pages 4733–4742, 2017.
- Jinjie You, Ancong Wu, Xiang Li, and Wei-Shi Zheng. Top-push video-based person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1345–1353, 2016.

Jianfu Zhang, Naiyan Wang, and Liqing Zhang. Multi-shot pedestrian re-identification via sequential decision making. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6781–6789, 2018.

Liang Zheng, Zhi Bie, Yifan Sun, Jingdong Wang, Chi Su, Shengjin Wang, and Qi Tian. Mars: A video benchmark for large-scale person re-identification. In *European Conference on Computer Vision*, pages 868–884. Springer, 2016.

Zhun Zhong, Liang Zheng, Donglin Cao, and Shaozi Li. Re-ranking person re-identification with k-reciprocal encoding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1318–1327, 2017.

Zhen Zhou, Yan Huang, Wei Wang, Liang Wang, and Tieniu Tan. See the forest for the trees: Joint spatial and temporal recurrent neural networks for video-based person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4747–4756, 2017.