Unsupervised and Semi-Supervised Domain Adaptation for Action Recognition from Drones

Supplementary Material

Jinwoo Choi[1,2] Gaurav Sharma[1] Manmohan Chandraker[1] Jia-Bin Huang[2]
[1]NEC Labs America [2]Virginia Tech

1. Ablation study on the losses

Table 1 gives the performances with the instance only adaptation for different configurations of the embedding-based framework in the different source and target label sets setting. If we use neither the adversarial loss nor the triplet loss and use only the cross entropy loss, the performance is 10.8%. If we add adversarial loss, the performance improves by 25.9%, relatively. If we use triplet loss only, we get 33.3% relative improvement over using cross entropy loss only. When both the triplet and adversarial losses are taken into account the performance is improved by 74.0% compared to cross entropy loss only. These results highlight two things. First, both the adversarial and triplet losses are important for good performance. Second, triplet loss performs better than cross entropy loss in the setting where the source and target label sets are different.

Table 1: Ablation study on the losses (*val* set) in the **different source and target label sets** setting. Our instance-based domain adaptation method is used for the ablation study.

Adversarial	triplet	Cross entropy	Acc (%)	Gain (%)
×	×	✓	10.8	0.0
✓	×	✓	13.6	25.9
×	✓	×	14.4	33.3
✓	✓	×	18.8	74.0

2. Implementation Details

Table 2: Correspondences between classes of the NEC-DRONE dataset and the Kinetics dataset for the same label set for source and target setting.

Drone dataset class	Kinetics dataset classes
walking	marching
running	jogging, running on treadmill
jumping	high jump, jumping into pool
drinking water from a bottle	drinking beer
throwing an object	throwing axe, throwing ball, throwing discuss, shot put, javelin throw
shaking hands	shaking hands
hugging	hugging

Same label set for source and target. We use the Kinetics dataset as source dataset and the proposed drone dataset as target dataset. Since the two datasets do not share exactly the same classes, we subsample the two datasets to obtain similar classes. We manually choose 13 classes from Kinetics dataset and 7 classes from our drone dataset which have similar or closely related actions to construct the source and target datasets. Table 2 gives the class correspondences between the classes of drone dataset and the Kinetics dataset.