Joint learning of Social Groups, Individuals Action and Sub-group Activities in Videos

Anonymous ECCV submission

Paper ID 761

1 Content

In the supplementary material, we provide more details for the evaluation of our social task in Section 5.2.

1.1 Evaluation Metrics

We evaluate our proposed framework for the social task by two metrics namely membership and social activity accuracy. Intuitively membership accuracy metric takes a social group assignment for each individual from our model and a ground truth assignment and then finds the best matching between them. The best matching can be efficiently computed by the Hungarian algorithm [1]. Figure 1 shows an example of calculating membership accuracy. Circles indicate individuals in the scene and the numbers inside the circles are social group IDs. By performing a Hungarian matching and substituting predicted social group IDs 1,2,3 by 3,1,2 respectively, membership accuracy can be calculated between the ground truth social group IDs and the predicted IDs which is 66% in this example.

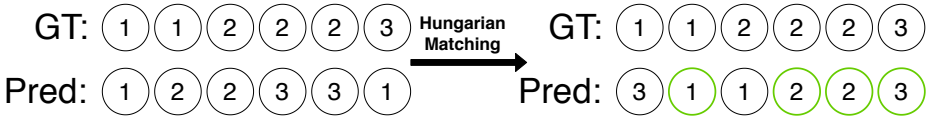

Fig. 1. An example of membership accuracy calculation. Circles indicate individuals in the scene and numbers in the circles indicate social group IDs.

1.2 Evaluation Settings

Since, we are the first to address the social task, we consider three set of different settings to evaluate the performance of our model. First two settings are the existing proposed methods for the group task which are evaluated in a social task setting. In the first setting indicated by [group] in Table 4 of the paper, similar to the group task, we consider that the predicted social group IDs are the same as shown in Figure 2 (considering all the individuals in the scene in

one social group) and calculate the membership and social activity accuracies accordingly. In the GT[group], membership accuracy and social activity accuracy are calculated based on the assumption that the predicted social activity is perfect for all the individuals. Thus, GT[group] is an upper bound performance in this setting.

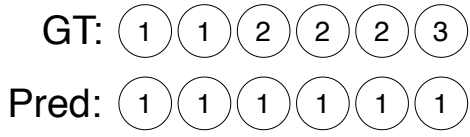

Fig. 2. Ground truth and predicted social group IDs in the [group] setting.

In the second setting indicated by [individual] in the table, we assume that all the predicted social group IDs are different as shown in Figure 3 (considering that each individual forms a social group) and calculate the membership and social activity accuracies accordingly. In the GT[individual], membership accuracy and social activity accuracy are calculated based on the assumption that the predicted social activity is perfect for all the individuals. Therefore, GT[individual] is an upper bound performance in this setting.

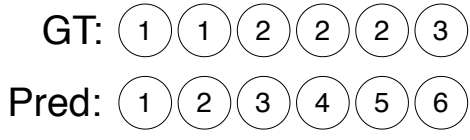

Fig. 3. Ground truth and predicted social group IDs in the [individual] setting.

Finally, in the last setting we give the model the flexibility to predict social groups. Ours[cluster] is the same as our proposed method for the group task in training. However, instead of predicting all the individuals in the same social group at test time, we perform a graph spectral clustering algorithm on the obtained matrix of connectivities between individuals from GAT and predict different social groups. In ours[learn2cluster], we learn the connectivities between individuals by adding a new term to our total loss function in training (graph partitioning loss) and perform graph spectral clustering algorithm on the obtained matrix of connectivities between individuals from GAT to predict social groups.

References

1. Kuhn, H.W.: The hungarian method for the assignment problem. Naval research logistics quarterly **2**(1-2), 83–97 (1955)