

CSC249 Final Project

Hecheng Sun

Task 1: Actor-Action Classification

Method description:

- For this task, I used the provided “a2d_dataset.py” program to preprocess my data. The input images are in the format of 224x224x3, and all has 44 valid classes. The provided annotations are processed into a one-hot encoding vector format, with 1 to represent classifying to that class and 0 to represent not classifying to that class.
- The network architecture I used for this task is ResNet152. I used PyTorch’s built-in ResNet class to build my “Network.py” and adjusting the final FC layer’s output dimension to 43 to match the class number of the dataset. I also initialize the network with the pretrained model.
- The loss function is cross entropy. The optimization method is Adam optimization, with a learning rate of 0.0001.
- All other hyperparameters should be the default values in the given skeleton code. This means an epoch of 10 and a batch size of 32.

Novelty of the method:

- There are a few novelties of this implementation.
- First, it uses the technique of transfer learning to initialize the network, making the convergence easier.
- Second, instead of using the sigmoid to map the output of the FC layer to an interval between 0 and 1, I changed the target label from “0 or 1” to “-1 or 1”. This seemingly simple change can tremendously benefit the training by allowing the loss function to backpropagate the loss with the correct sign. This increased the F1 score from around 15 to around 27.
- Third, after training, I used different cutoffs to check the resulted F1 score. After a few experiments, I found out that 0.1 is the best cutoff for the validation set. This change of cutoff increased the F1 score to 30+.

Performance on validation set:

- Precision: 21.4 Recall: 65.0 F1: 30.5