To Run

Step 1: Extract Features

Run “python main.py --feature_type i3d --device_ids 0 --video_paths ./sample_video_path”

“python main.py --feature_type vggish --device_ids 0 --video_paths ./sample_video_path”

This generates an output folder with name “output”

Step 2: Generate Captions

Run the following with correct file paths from source folder:

```
python test/test.py \
    --prop_generator_model_path best_prop_model.pt \
    --pretrained_cap_model_path best_cap_model.pt \
    --vggish_features_path test/”vgg_features_sample_path” \
    --rgb_features_path test/”i3d_rgb.npy” \
    --flow_features_path test/”i3d_flow.npy” \
    --duration_in_secs “duration” \
    --device_id 0 \
    --max_prop_per_vid 100 \
    --nms_tiou_thresh 0.4
```

This generates captions , sample output would look like this :

[{'start': 0.0, 'end': 4.9, 'sentence': 'We see the closing title screen'}, {'start': 2.7, 'end': 29.0, 'sentence': 'A woman is seen running down a track and down a track while others watch on the sides'}, {'start': 19.6, 'end': 33.3, 'sentence': 'The man runs down the track and jumps into a sand pit'}, {'start': 0.0, 'end': 13.0, 'sentence': 'A man is seen running down a track and leads into a large group of people running around a track'}]