Task Writeup

Himanchal Chandra
25/07/2021

Task:

This task involves creating a pipeline to test out inference from two different Human Activity recognition models (PyTorch and TensorFlow) which have been trained on the Kinetics dataset.

Pytorch Model:

1. Load the class labels which contain 400 labels and the input file which contains the video's name for inference.

2. Set the sample size to 112 and sample duration to 16.

3. Loading the pre-trained PyTorch model which utilizes Inflated 3D ResNet.

4. Iterating through the video list (input).

5. Create a temporary folder named tmp for storing the frames.

6. Extracting frames from video using FFmpeg.

7. Applying spatial and temporal transformations to the frames and passing them through the model.

8. Calculating the maximum occurring label across all frames and return that label.

9. Deleting the tmp folder.

10. Finally, printing the predicted label along with the video name.

Tensorflow Model:

1. Load the class labels which contain 400 labels and the input file which contains the video's name for inference.

2. Loading the pre-trained TensorFlow model which utilizes Inflated 3D Convnet.

3. Iterating through the video list (input).

4. Applying transformation to each video: Loading the video as frames using OpenCV, cropping each frame from the center (square), resizing the frame to 224 x 224, and finally normalizing each frame by dividing it by 255.

5. Passing the video through the TensorFlow model.

6. Getting the probabilities from the model using the softmax function on the logits.

7. Sorting the probabilities in decreasing order and getting the first label.

8. Finally, printing the predicted label along with the video name.

Results:

The inference was carried out using 20 short video clips which I downloaded from youtube.

1. PyTorch Model Result:

```
PRINTING THE RESULT USING PYTORCH MODEL:
video1.mp4 sticking tongue out
video2.mp4 sticking tongue out
video3.mp4 playing squash or racquetball
video4.mp4 riding unicycle
video5.mp4 snowboarding
video6.mp4 playing cricket
video7.mp4 ice climbing
video8.mp4 moving furniture
video9.mp4 petting animal (not cat)
video10.mp4 driving car
video11.mp4 pushing car
video12.mp4 breakdancing
video13.mp4 grooming dog
video14.mp4 making tea
video15.mp4 smoking
video16.mp4 driving car
video17.mp4 hurling (sport)
video18.mp4 training dog
video19.mp4 flipping pancake
video20.mp4 playing organ
```

2. TensorFlow Model Result:

```
PRINTING THE RESULT USING TENSORFLOW MODEL:
video1.mp4 checking tires
video2.mp4 crying
video3.mp4 walking the dog
video4.mp4 marching
video5.mp4 snowboarding
video6.mp4 flying kite
video7.mp4 making snowman
video8.mp4 cartwheeling
video9.mp4 petting animal (not cat)
video10.mp4 driving car
video11.mp4 faceplanting
video12.mp4 headbutting
video13.mp4 petting cat
video14.mp4 unboxing
video15.mp4 smoking
video16.mp4 driving car
video17.mp4 catching or throwing baseball
video18.mp4 walking the dog
video19.mp4 brushing teeth
video20.mp4 air drumming
```

Thank You!