TSN Task Report

Training:

1. RGB model:
    a. We use a pretrained ResNet18 model, modified to have 25 output classes as in our dataset.
    b. No. of segments = 4
2. Optical Flow model:
    a. We also initialize with the pretrained version of ResNet18, because it was shown in the literature that RGB values can correlate with Optical Flow values
    b. While we are not 100% able to confirm this due to time constraints, we believe that initializing the first layer with the averages learned weights of the RGB model leads to more efficient learning
    c. No. of segments = 4
3. Notes on training:
    a. As would be expected, the preinitialization from ImageNet benefits the rgb model much more than the optical flow model, especially at the beginning.

Results:
- We faced problems managing the large memory needs of the different models, especially when training all of them and trying to keep them together in memory for comparison and fusion.
- RGB model reaches 75% accuracy with just one epoch, optical flow (6 epochs) without rgb init 61%, with it 63%. (In the best runs)
- RGB model stays at 75% best accuracy, more epochs lead to overfitting. Loss drops near 0 on the training set, but error gets higher on the test set. We would need more time for testing more parameters and transformations on the data.
- RGB model gets higher results because background information helps a lot in the case of our simple classes

Suggestions for improvement:
- Use more data or apply more transforms and data augmentation to avoid quick overfitting (Training keeps improving on training set, but stagnates on test set)
- Using more segments: We use 4 segments per video, however, in the literature they can use up to 64, so trying more segments would be a good idea
- Trying other baseline models, or deeper versions of ResNet

LOGS from training on one epoch for RGB, 6 for optical Flow in both of its modes:

========STARTING RGB TRAINING==========
starting epoch for RGB training: 0
/usr/local/lib/python3.7/dist-packages/torchvision/io/video.py:159: UserWarning: The pts_unit 'pts' gives wrong results and will be removed in a follow-up version. Please use pts_unit 'sec'.
  "The pts_unit 'pts' gives wrong results and will be removed in a "
  Batch 0: loss = 3.49269318580062744
  Batch 10: loss = 1.5362038612365723
  Batch 20: loss = 0.930392861366272
  Batch 30: loss = 0.5449123382568359
  Batch 40: loss = 0.32985010743141174
  Batch 50: loss = 0.18782034516334534
  Batch 60: loss = 0.16709424555301666
  Batch 70: loss = 0.19210827350616455
average loss on epoch 0: 0.6672223179539045
evaluating on validation set
current accuracy: 0.7219827771186829
========STARTING FLOW TRAINING==========
512
starting epoch for optical flow training: 0
  Batch 0: loss = 3.298482894897461
  Batch 10: loss = 2.4944210052490234
  Batch 20: loss = 2.273702383041382
  Batch 30: loss = 2.2339725494384766
  Batch 40: loss = 1.8463265895843506

```
    Batch 50: loss = 1.7042677402496338
    Batch 60: loss = 1.4837234020233154
    Batch 70: loss = 1.7284736633300781
average loss on epoch 0: 1.9706972726186116
evaluating on validation set
current accuracy: 0.45743533968925476
starting epoch for optical flow training: 1
    Batch 0: loss = 1.058258295059204
    Batch 10: loss = 1.0314205884933472
    Batch 20: loss = 1.1468636989593506
    Batch 30: loss = 0.9844346046447754
    Batch 40: loss = 0.7871050834655762
    Batch 50: loss = 1.2888381481170654
    Batch 60: loss = 1.0843921899795532
    Batch 70: loss = 0.9662673473358154
average loss on epoch 1: 1.1353025285402933
evaluating on validation set
current accuracy: 0.5444504022598267
starting epoch for optical flow training: 2
    Batch 0: loss = 0.6169149279594421
    Batch 10: loss = 0.9653003215789795
    Batch 20: loss = 0.6240341663360596
    Batch 30: loss = 0.6107838749885559
    Batch 40: loss = 0.7504445910453796
    Batch 50: loss = 0.6765245795249939
    Batch 60: loss = 0.629477322101593
    Batch 70: loss = 0.5399748086929321
average loss on epoch 2: 0.797319540977478
evaluating on validation set
current accuracy: 0.595097005367279
starting epoch for optical flow training: 3
    Batch 0: loss = 0.6231139898300171
    Batch 10: loss = 0.40583735704421997
    Batch 20: loss = 0.6060651540756226
    Batch 30: loss = 0.5236485600471497
    Batch 40: loss = 0.45296353101730347
    Batch 50: loss = 0.36324042081832886
    Batch 60: loss = 0.6118367314338684
    Batch 70: loss = 0.602269172668457
average loss on epoch 3: 0.5489975770314535
evaluating on validation set
current accuracy: 0.6096444129943848
starting epoch for optical flow training: 4
    Batch 0: loss = 0.34588488936424255
```

Batch 10: loss = 0.37583544850349426
Batch 20: loss = 0.3317928910255432
Batch 30: loss = 0.47672486305236816
Batch 40: loss = 0.31999877095222473
Batch 50: loss = 0.40093663334846497
Batch 60: loss = 0.22433197498321533
Batch 70: loss = 0.38348138332366943
average loss on epoch 4: 0.3584156254927317
evaluating on validation set
current accuracy: 0.615840494632721
starting epoch for optical flow training: 5
Batch 0: loss = 0.27572011947631836
Batch 10: loss = 0.2961585819721222
Batch 20: loss = 0.17818039655685425
Batch 30: loss = 0.250670850276947
Batch 40: loss = 0.12327300012111664
Batch 50: loss = 0.3387334644794464
Batch 60: loss = 0.26593053340911865
Batch 70: loss = 0.1790829598903656
average loss on epoch 5: 0.23902002463738123
evaluating on validation set
current accuracy: 0.6371228694915771
========STARTING FLOW TRAINING with RGB init=========
512
starting epoch for optical flow training: 0
    Batch 0: loss = 3.4189157485961914
    Batch 10: loss = 2.7579710483551025
    Batch 20: loss = 2.1881024837493896
    Batch 30: loss = 1.597906470298767
    Batch 40: loss = 1.785117745399475
    Batch 50: loss = 1.8426909446716309
    Batch 60: loss = 1.6441879272460938
    Batch 70: loss = 1.714063286781311
average loss on epoch 0: 2.042806951204936
evaluating on validation set
current accuracy: 0.46713361144065857
starting epoch for optical flow training: 1
    Batch 0: loss = 1.2274694442749023
    Batch 10: loss = 1.2197353839874268
    Batch 20: loss = 0.8888126015663147
    Batch 30: loss = 1.1054989099502563
    Batch 40: loss = 1.3322105407714844
    Batch 50: loss = 0.9872933030128479
    Batch 60: loss = 0.8750084042549133
    Batch 70: loss = 1.1038854122161865
average loss on epoch 1: 1.1333243981997172

```
evaluating on validation set
current accuracy: 0.5463362336158752
starting epoch for optical flow training: 2
    Batch 0: loss = 0.8103703856468201
    Batch 10: loss = 0.9812977910041809
    Batch 20: loss = 0.6072425246238708
    Batch 30: loss = 0.8291133046150208
    Batch 40: loss = 0.8448845148086548
    Batch 50: loss = 0.690477728843689
    Batch 60: loss = 0.7568544745445251
    Batch 70: loss = 0.5800045728683472
average loss on epoch 2: 0.7958146977424622
evaluating on validation set
current accuracy: 0.5522629022598267
starting epoch for optical flow training: 3
    Batch 0: loss = 0.5998767018318176
    Batch 10: loss = 0.48972102999687195
    Batch 20: loss = 0.4167863726615906
    Batch 30: loss = 0.47728854417800903
    Batch 40: loss = 0.33490923047065735
    Batch 50: loss = 0.4067733883857727
    Batch 60: loss = 0.5109941959381104
    Batch 70: loss = 0.6749112606048584
average loss on epoch 3: 0.5360585661729177
evaluating on validation set
current accuracy: 0.5918642282485962
starting epoch for optical flow training: 4
    Batch 0: loss = 0.552790105342865
    Batch 10: loss = 0.3796640634536743
    Batch 20: loss = 0.4437817335128784
    Batch 30: loss = 0.44314926862716675
    Batch 40: loss = 0.7322526574134827
    Batch 50: loss = 0.5619193911552429
    Batch 60: loss = 0.5024798512458801
    Batch 70: loss = 0.2594705820083618
average loss on epoch 4: 0.39306026140848793
evaluating on validation set
current accuracy: 0.6179956793785095
starting epoch for optical flow training: 5
    Batch 0: loss = 0.19502922892570496
    Batch 10: loss = 0.20212356746196747
    Batch 20: loss = 0.2918771803379059
    Batch 30: loss = 0.20447732508182526
    Batch 40: loss = 0.1685357689857483
    Batch 50: loss = 0.1713617891073227
    Batch 60: loss = 0.10694024711847305
    Batch 70: loss = 0.1959858387708664
```

```
average loss on epoch 5: 0.23295244336128235
evaluating on validation set
current accuracy: 0.6169180870056152
========TESTING FUSION==========

RuntimeError: CUDA out of memory. Tried to allocate 76.00 MiB (GPU 0;
14.76 GiB total capacity; 13.46 GiB already allocated; 69.75 MiB free;
13.66 GiB reserved in total by PyTorch)
```