Convolutional Neural Networks (CNNs): Improving Performance

Further Explore the Potential of Invariant Information Clustering Using Video Datasets

Cai Shengqu

k1763920

Progress since the last meeting

- Data gathering, pair = (x, x'):
  - Use only the first frame as x.(Initial experiemens to show this is worth experiment on)
    - train and test on the full dataset.
      - 3 sample repeats, pair with (0,0,0)(baseline),(1,2,3),(2,4,6),(3,6,9)
    - train on 6000 video train partition, test on 3000 video test partition.
      - 5 sample repeats, pair with single frame:0(baseline), 1, 2, 3, 4. (5, 6, 7, 8, 9)*
  - Use all 10 frames as x.
    - 3 sample repeats.
      - (use single frame)*
      - use multiple frames.
        - train and test on 6000 train partition.
          - interval 0(baseline), 1*, 2, 3*
        - train on train, test on test
          - interval 0(baseline), 1, 2, 3*
        - (train and test on full dataset)*
  - Remove the random crop option, train on train test on test.(To show the effect of frames clearer)
    - 3 sample repeats
      - use single frame, pair with 0(baseline), 1, 2, 3, 4, (5,6,7,8,9)*.
      - use multiple frames, interval 1, 2*, 3*
    - 5 sample repeats, increment 1*
    - 10 sample repeats, increment 1*

- *: experiments planning but not yet started.

Current challenges and blocks

- Report.

- Design the best approach to represent data.
  - single frame vs different single frame.
  - single frame vs multi-frame.
  - multi-frame vs multi-frame with different interval.
  - train and test on the same data vs train and test on seperate partitions.
  - results without random cropping.
  - for each of which, average vs best

My next steps

- Updated plan:
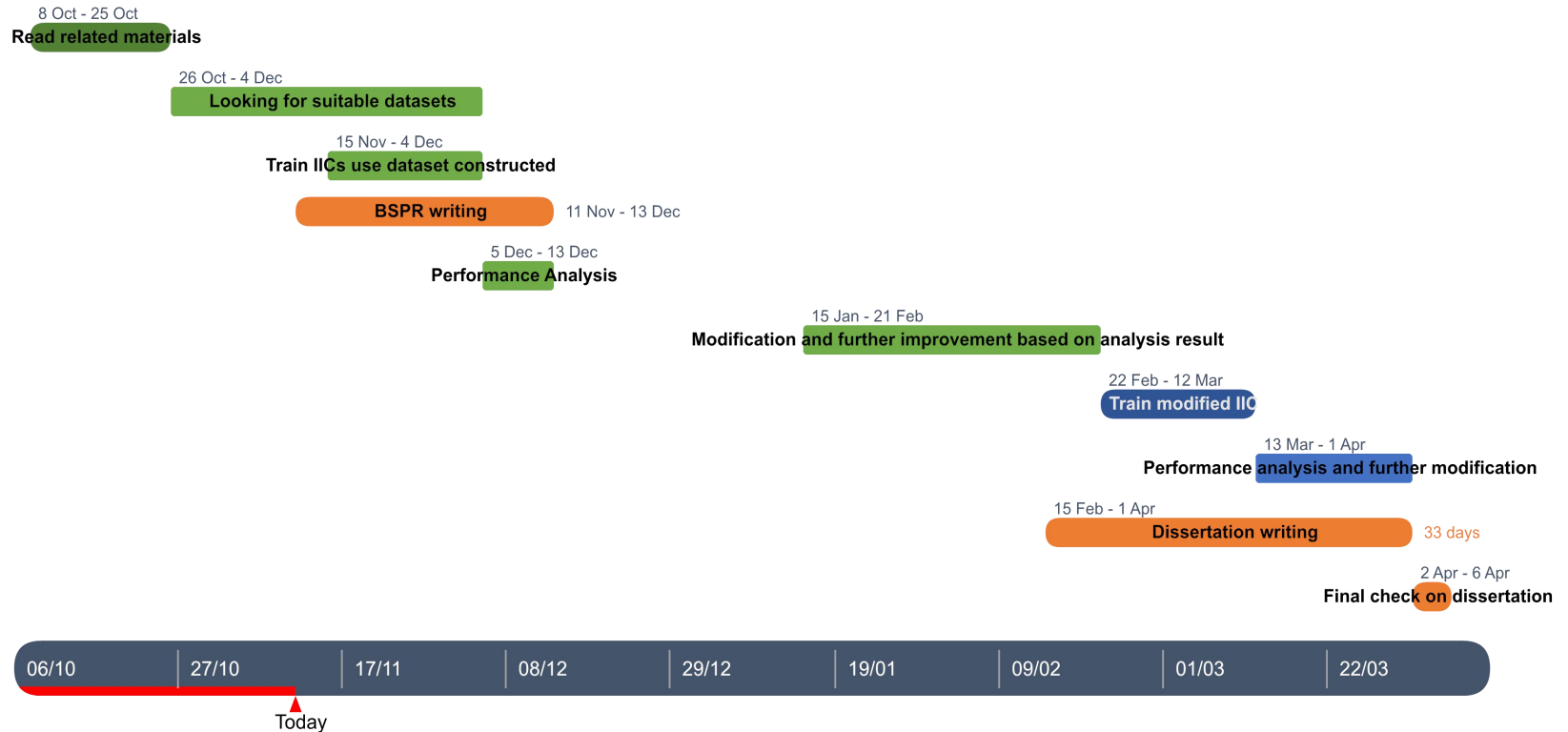
- Next steps:
  - Report.

Reference list

Xu Ji, Joao F Henriques, and Andrea Vedaldi. Invariant information distillation for unsupervised image segmentation and clustering. *arXiv preprint arXiv:1807.06653, 2018.*

E. Real, J. Shlens, S. Mazzocchi, X. Pan, and V. Vanhoucke. Youtube-boundingboxes: A large high-precision humanannotated data set for object detection in video. *arXiv preprint arXiv:1702.00824, 2017.*