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

- Implement Youtube-BouncingBox in PyTorch.
    - Done.
        - videos->image of frames only.
            - subset cutting?
                - 23 classes->10 classes
                - 1000 segments per class(further action required)
- Implement related scripts.
    - Done.
        - Testing in progress.
- Perform the first training on TY-BB, gather performance data and evaluate further improvement space.
    - In progress(2-3 weeks).

Current challenges and blocks

- http error 429 too many requests
  - youtube blacklists fingerprints.
  - Possible solution: adding retry logic, download dataset from different clients.
  - Worst case: 1000 segments per class---should be sufficient.
- Yet to do:
  - Cropping images by bounding box.
  - Data comparison and conclusion.

My next steps

- Updated plan:
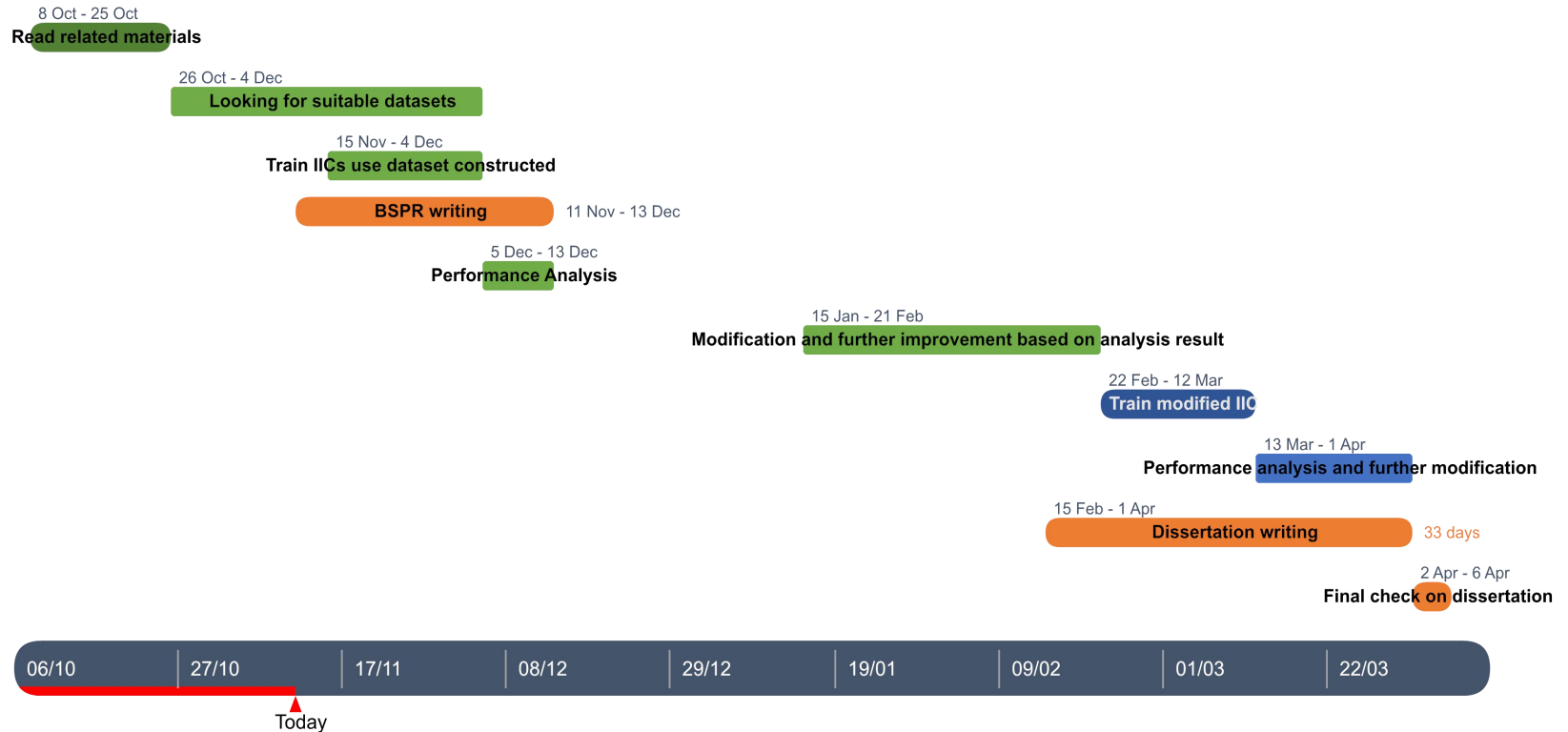
- Next steps:
  - Data gathering(in progress)
  - Dissertation writing.

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.*