Problem

How can we fastly train video classification models?
Approach

- Development of video-handling condensation
- Changed data loader
  - Can get data from Newton cluster
- Changed model architecture
  - Architecture with video backbone
Baseline Model: Design

- Dataset: UCF-101
- Architecture: ResNet-18 3D
- Percentages of original data
  - Full Data - 100%
  - Random - 10%
  - Herding - 10%
- Epochs: 90
Baseline Model: Design

- UCF-101
  - 13,320 videos
  - 101 action categories:
    - Rope Climbing
    - Apply Eye Makeup
    - etc.
  - Aim: encourage research into action recognition
Baseline Model: Results

<table>
<thead>
<tr>
<th></th>
<th>Full Data</th>
<th>Random</th>
<th>Herding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 1 Acc</td>
<td>46.803</td>
<td>???</td>
<td>???</td>
</tr>
<tr>
<td>Top 5 Acc</td>
<td>70.332</td>
<td>???</td>
<td>???</td>
</tr>
</tbody>
</table>
Baseline Model: Road Blocks

- Problems with Newton Cluster
  - Unrecognized file or directory
    - Was not an issue for past jobs
    - Fixed by loading python module before jobs
  - Memory issue
    - Also was not an issue for past jobs
    - Fixed by allocating more GPU for jobs via more nodes
- Program takes long to run => need to improve program's time
- Challenge: obtaining synthesized videos
Next Steps

- Achieve testing accuracy results and resulting videos for coreset and distribution matching synthesized videos
- Get results for same the dataset with different sample percentages (e.g., 15%, 20%, 25%)
- Get results based on varying architectures and at least one other dataset
Thank You

Any Questions?