FINDING THE (AD)OPTIMAL CAT PHOTO

MILESTONE 1 PRESENTATION

Kyra Ballard
Andrew Fu
Shravan Nageswaran
Emily Xie

Harvard University
IACS Capstone, Spring 2020

PROBLEM

- About 6.5M dogs and cats each year enter animal shelters, according to the ASPCA
- But approximately 1.5M of these are put down

That's 1 in 4 pets.


ALL YOU NEED IS A
CUTE CAT IMAGE OR VIDEO
AND THE INTERNET
WILL GO CRAZY


42258969
Female
1 year
35lbs.
Intake/vax. 07/19
Timid, but approaches.
Owner/Guardian Surrender

THE QUALITY OF THE PHOTO OF A PET IS CORRELATED TO ITS LIKELIHOOD OF BEING ADOPTED

GOALS

increase adoption rates

decrease euthanasia for shelter cats

SPONSORS

AUSTIN PETS ALIVE!

- Animal shelter based in Texas
- No-kill shelter
- Saved over **70,000 animal lives** since 2008

Austin Pets Alive! is not your average animal shelter. We pioneer innovative lifesaving programs designed to save the animals most at risk of euthanasia.

Adopt

Looking for a furry friend to add to the family? We have thousands of animals that would love to be part of your home.

Foster

Open your heart and home to a pet in need, and be the bridge to a dog or cat's forever home.

Volunteer

Our volunteers make lifesaving possible – become a volunteer today!

ADOPTIMIZE

- Software company
- Primary goals
  - Increase adoption
  - Decrease euthanasia
  - Increase shelter engagement
- Algorithm that optimizes image taking
  - For the best chance of adoption

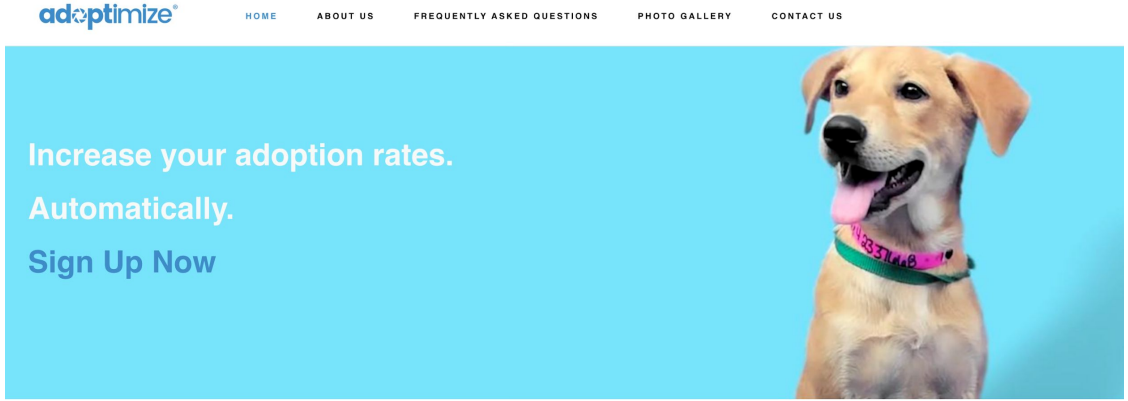
Increases Adoption Rates

Saves Lives

Increases Online Engagement

ADOPTIMIZE

- Dog model process
  - Takes in a **video** of the animal
  - Selects the **optimal shot**
  - Automatically **edits** the image
  - Outputs **enhanced optimal** image



IMPACT

124% increase in adoption
41% reduction in euthanasia

27% increase in adoption
56% reduction in euthanasia

COLLABORATION INFRASTRUCTURE

- Meetings with sponsors every 1-2 weeks
- Weekly meetings with team members
- **Shared Github repository** for project
  - Contains code files, data, reports
- Personal and shared copies
  - Separate exploration and functional product
  - Combine working parts

SCOPE OF WORK

in scope

- Model taking cat videos and outputting best frame
  - Length: <60s
  - Unobstructed view of a single cat
- Simple web app for mobile devices

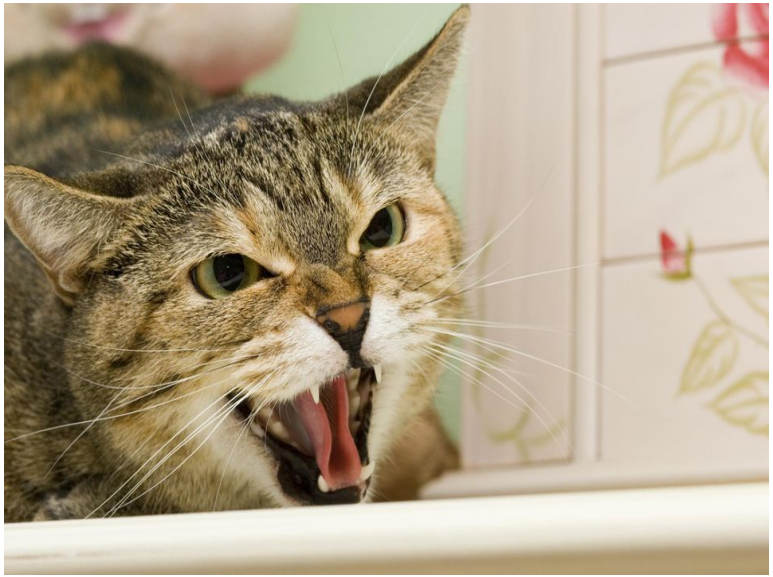
out of scope

- Front-end
- Measuring Adoption Rates

CHALLENGES

Behavior	Limitations
○ Fur covering face ○ Not facing camera ○ Unwilling to sit still	○ Environment ○ Equipment and camera ○ Resources available

Working with cats is **tough...**



CHALLENGES

Video Quality	Dataset
○ Filmed in kernels ○ Low light environment ○ Unstable camera ○ Camera quality (phone vs. laptop)	○ Small number of cat videos compared to dogs ○ No labeled data ○ Ambiguous: what makes a good photo? ○ Measuring success

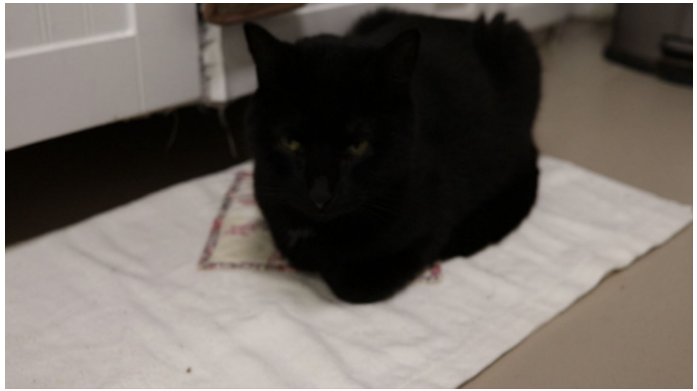
The data initially poses some **challenges...**

THE DATA

The Good

- Full body visible
- Looking directly at camera
- Clear, high quality image
- Good lighting

The Bad



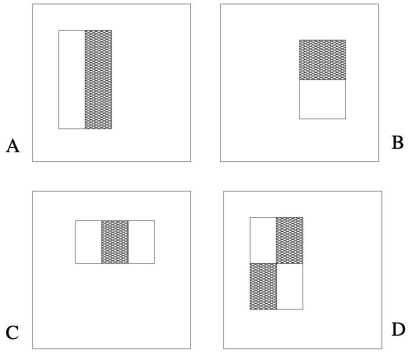
- Full body not visible
- Can’t distinguish facial features
- Looking away from camera
- Blurry image
- Darker area

HEURISTICS

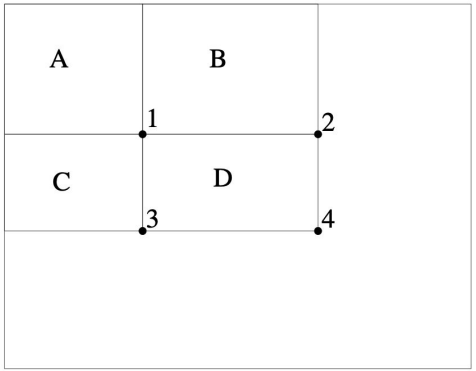
- Detection of Front-Facing Cat Head and Cat Eyes
- Variance of Laplacian
- Ratio of Cat Head Area to Frame Area

LITERATURE REVIEW: VIOLA-JONES

- *Rapid Object Detection using a Boosted Cascade of Simple Features*
- Haar-like Features
  - Pre-Compute Integral Image
- AdaBoost on Decision Stumps
- Cascade
- Sliding Windows



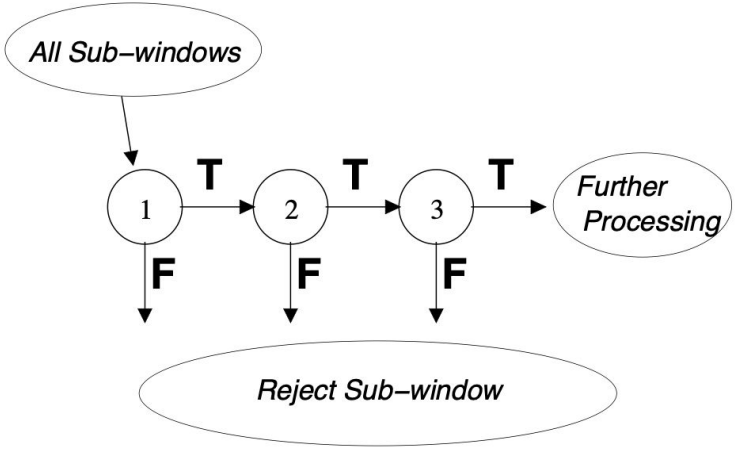
Haar-like Features



After post-processing, rectangle features can be calculated with array lookups as opposed to sums

LITERATURE REVIEW: VIOLA-JONES

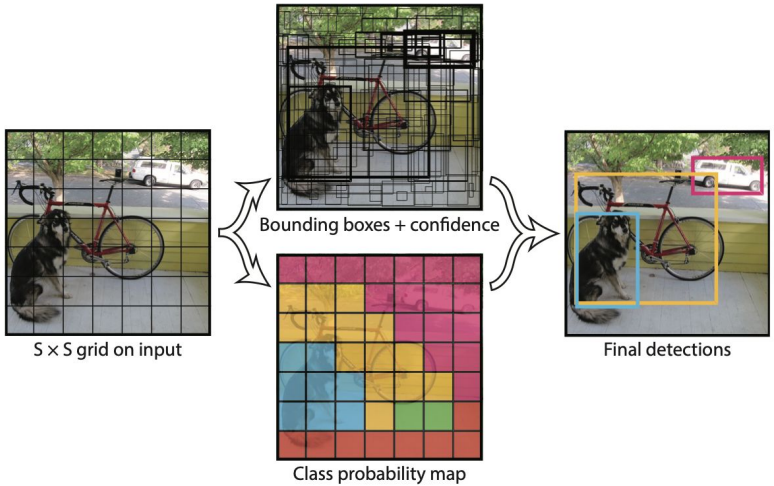
- *Rapid Object Detection using a Boosted Cascade of Simple Features*
- Haar-like Features
  - Pre-Compute Integral Image
- AdaBoost on Decision Stumps
- Cascade
- Sliding Windows



High-level view of cascade approach

LITERATURE REVIEW: YOLO

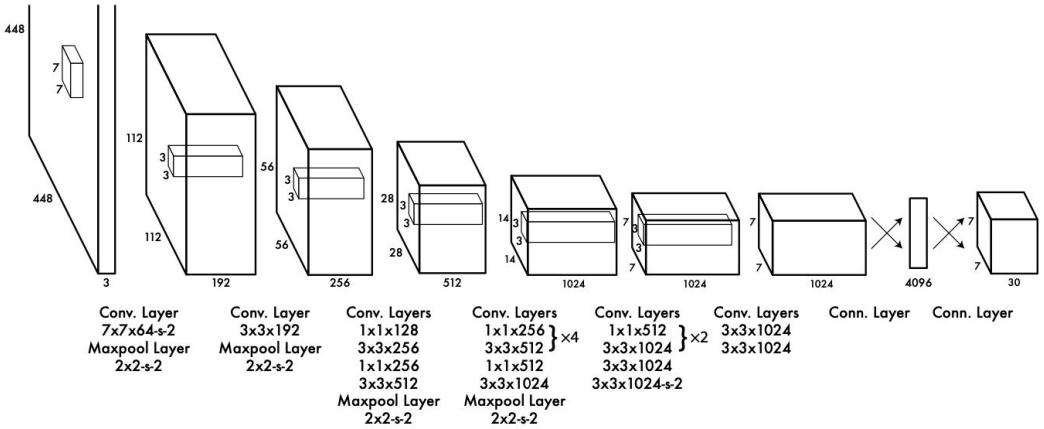
- *You Only Look Once: Unified, Real-Time Object Detection*
- Simultaneous Box and Class Proposal
- Simplicity: CNN
- Optimized for Speed



Each grid cell is responsible for producing exactly *B=2* bounding boxes representing existence of any object with center in the cell

LITERATURE REVIEW: YOLO

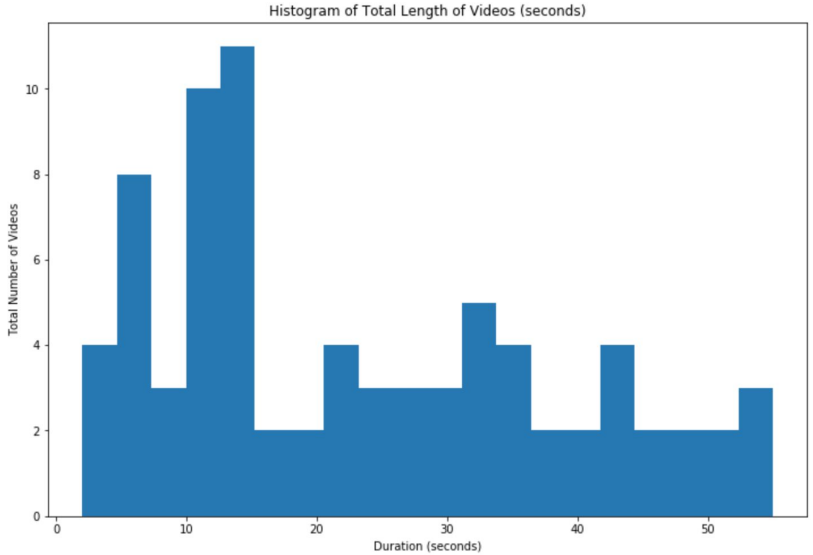
- *You Only Look Once: Unified, Real-Time Object Detection*
- Simultaneous Box and Class Proposal
- Simplicity: CNN
- Optimized for Speed

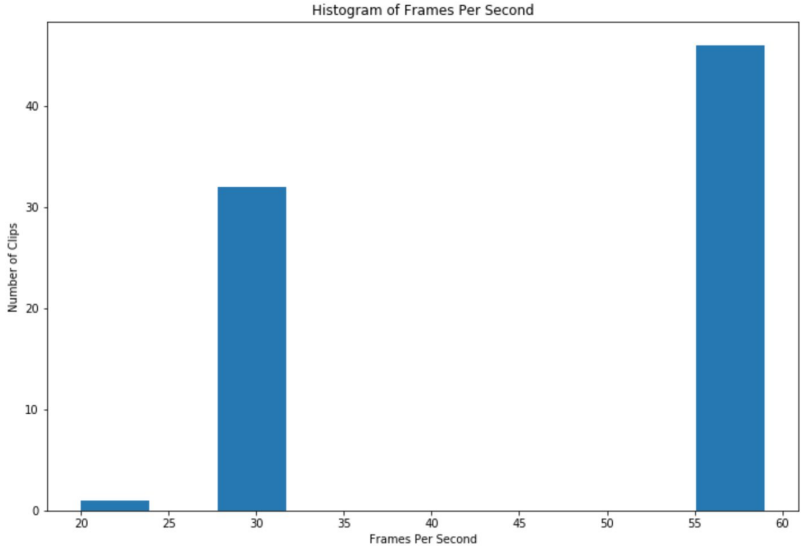


YOLO architecture; note only convolutional and fully connected layers

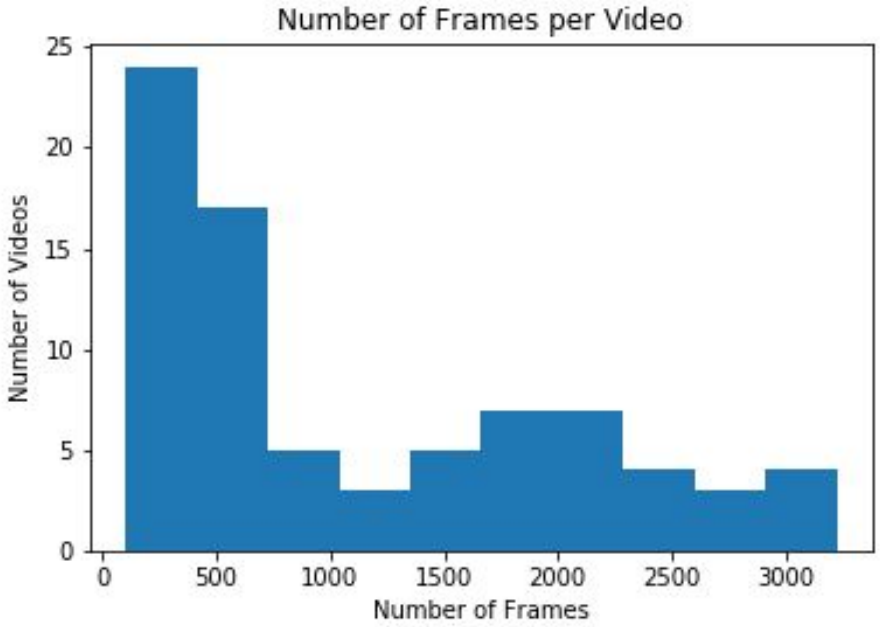
EDA: GENERAL DATA SENSE

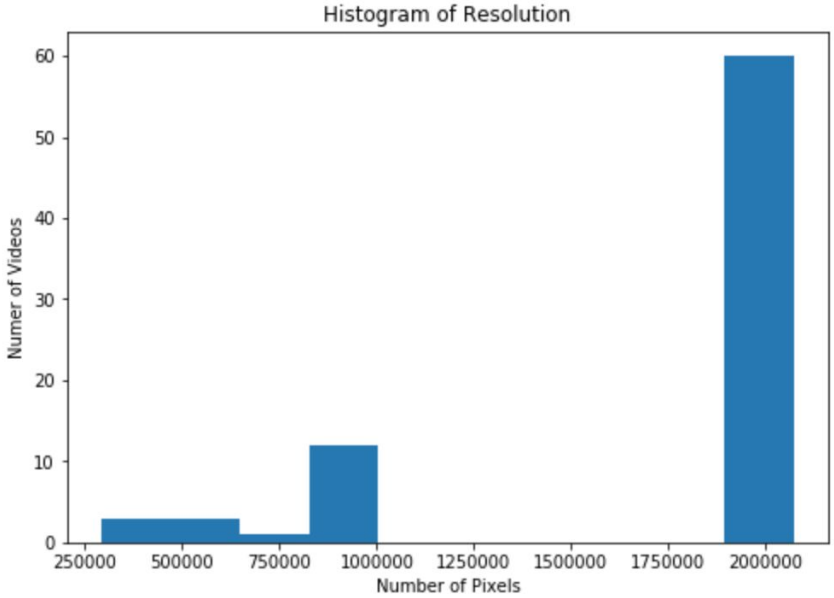
- Initial data cleaning yielded 80 videos
- Duration: Avg: 23 seconds. Min 2 seconds. Max 55 seconds
- FPS: ~60/40 breakdown on 60 fps vs 30





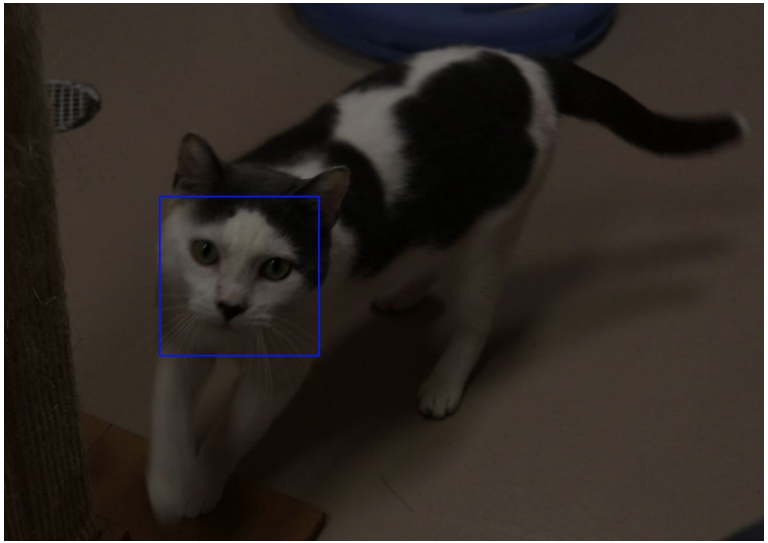
EDA: GENERAL DATA SENSE

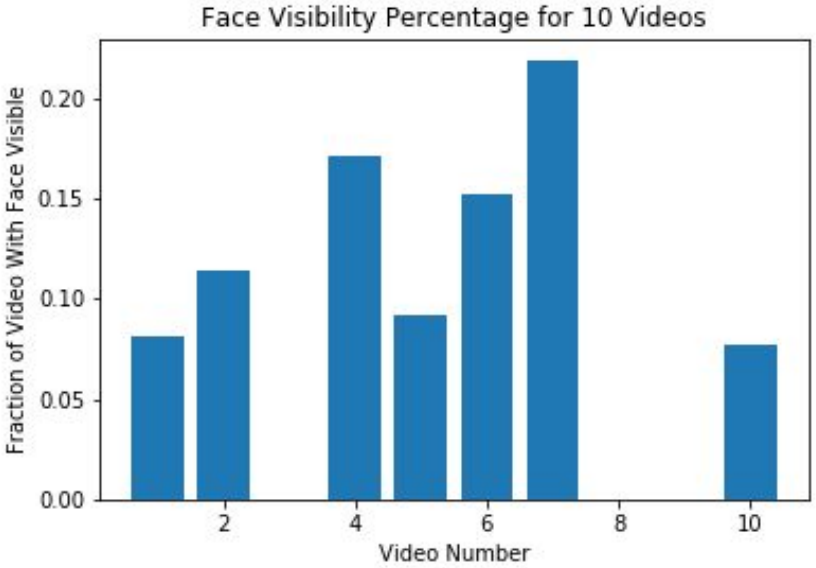




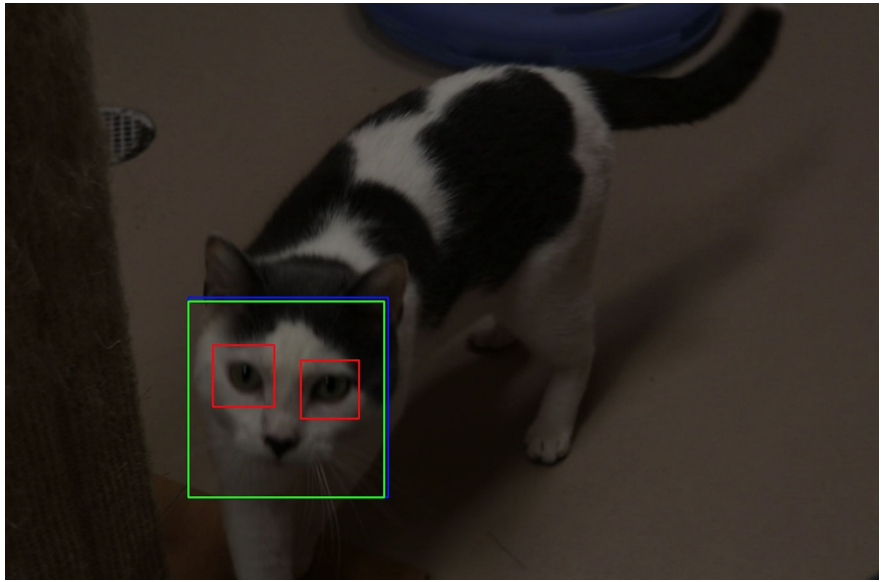
EDA: CAT FACE DETECTION

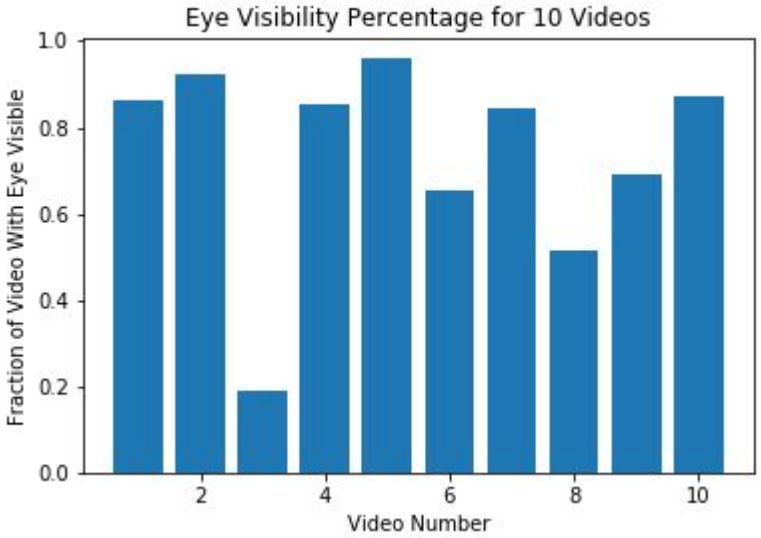
- Ran cat face detection using Haar Cascade
- Subsample of 10 videos
- Every 10th frame per video



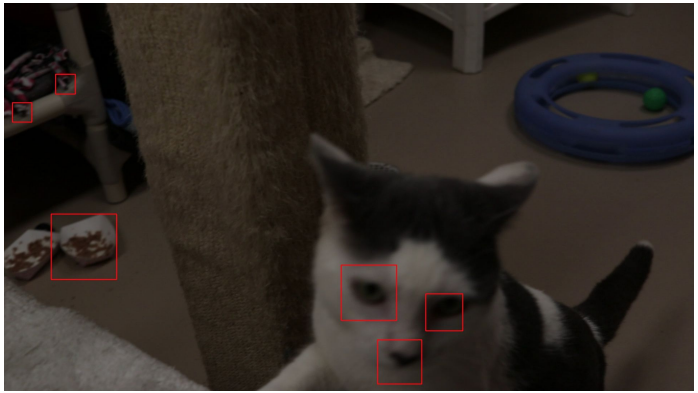


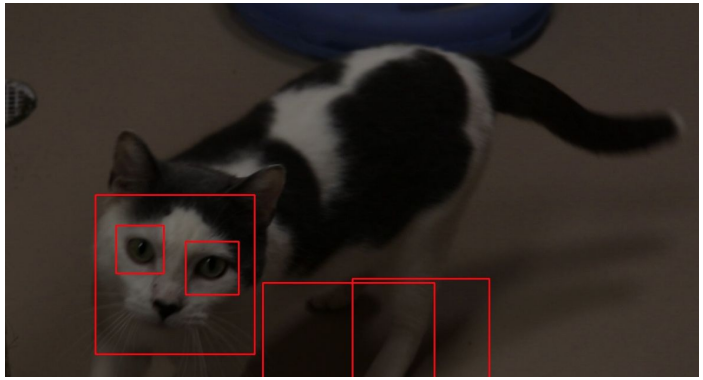
EDA: EYE DETECTION

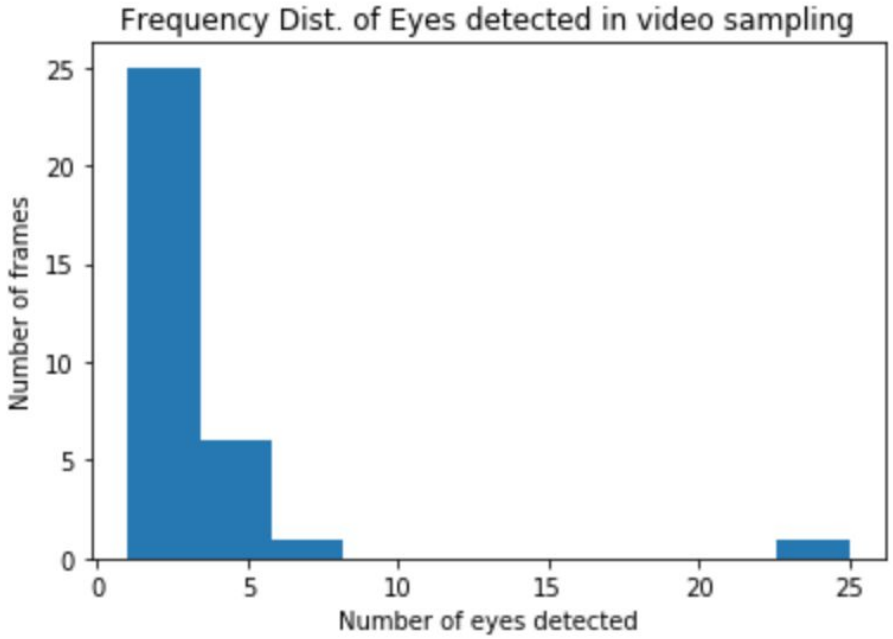




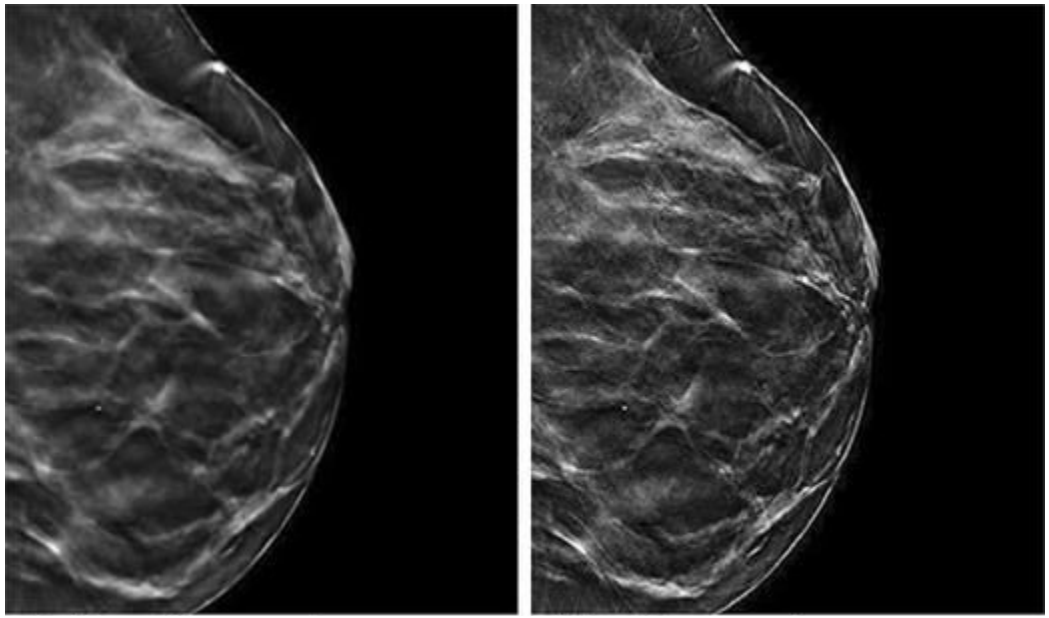
EDA: EYE DETECTION







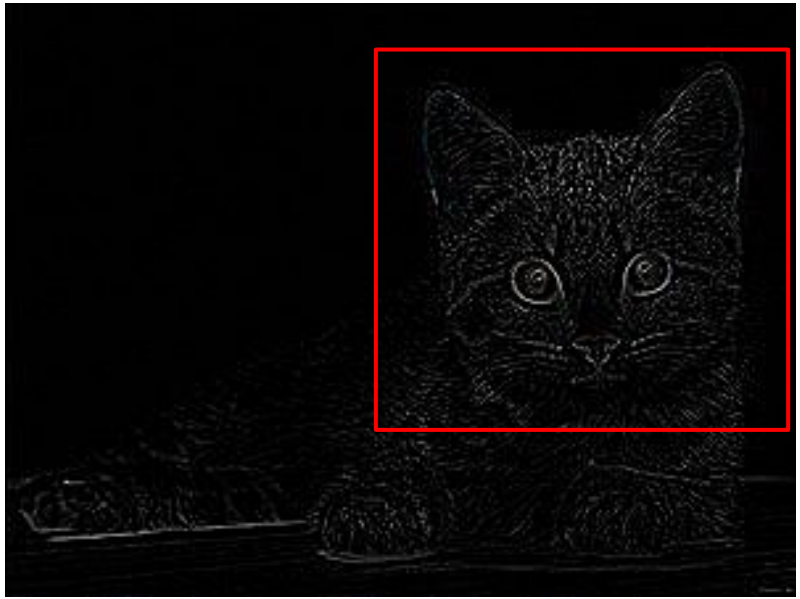
EDA: MEASURES OF SHARPNESS

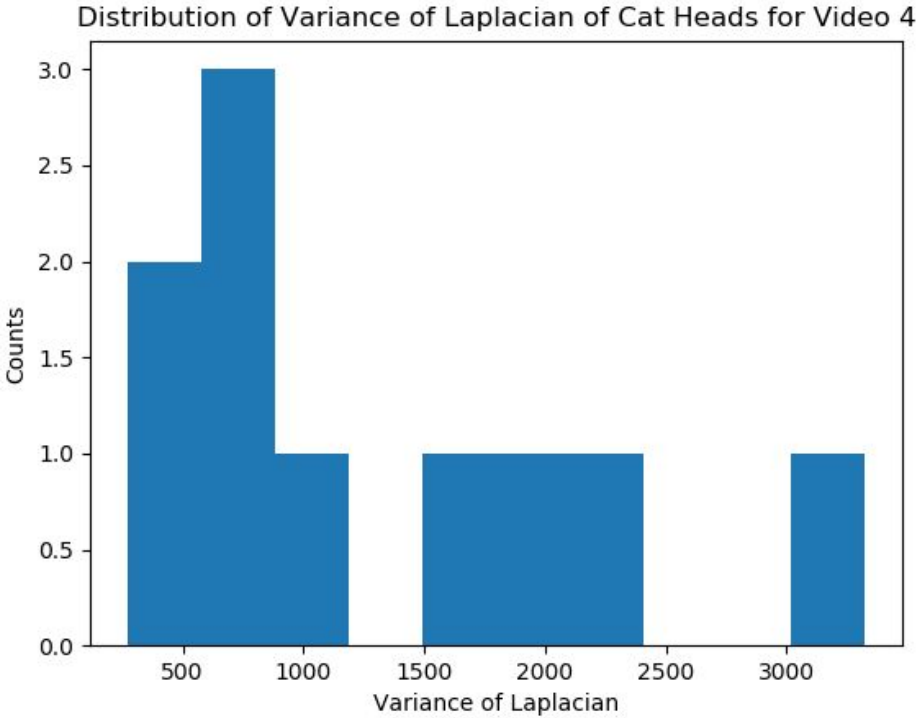


Less sharp

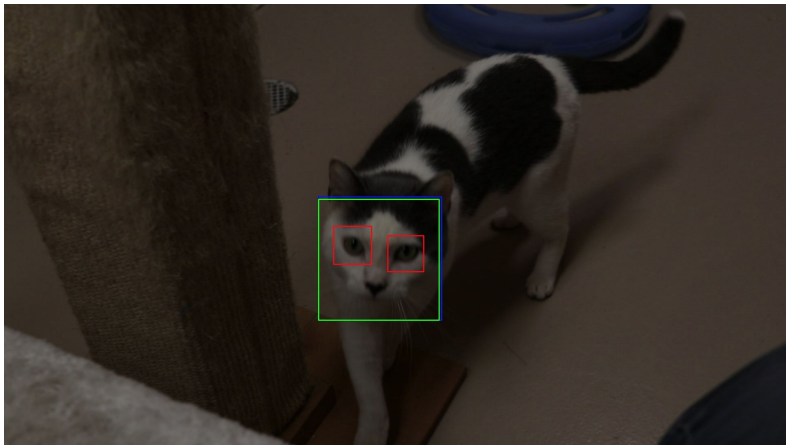
More sharp

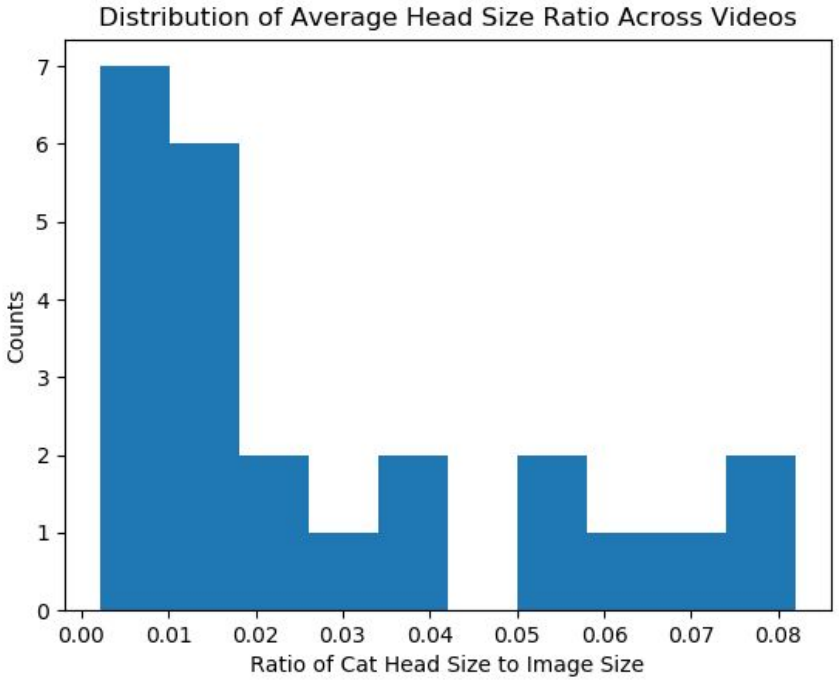
EDA: MEASURES OF SHARPNESS



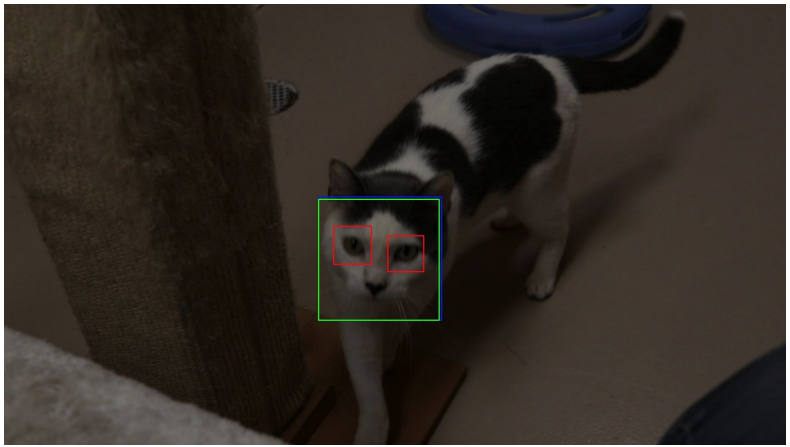


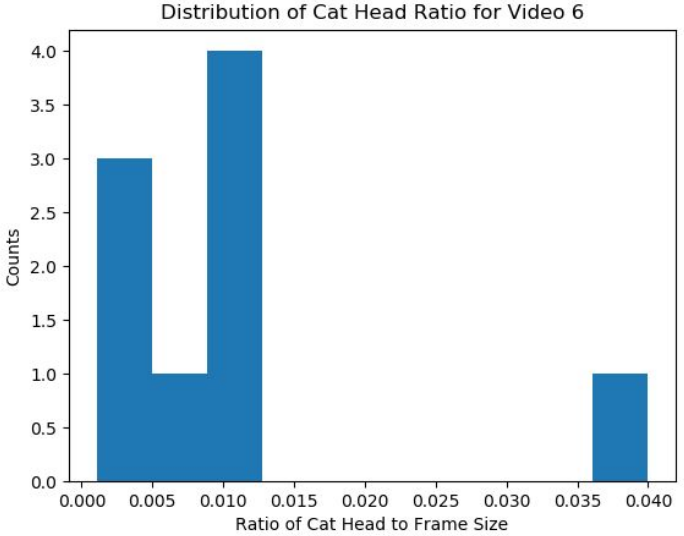
EDA: HEAD SIZE RATIOS





EDA: HEAD SIZE RATIOS





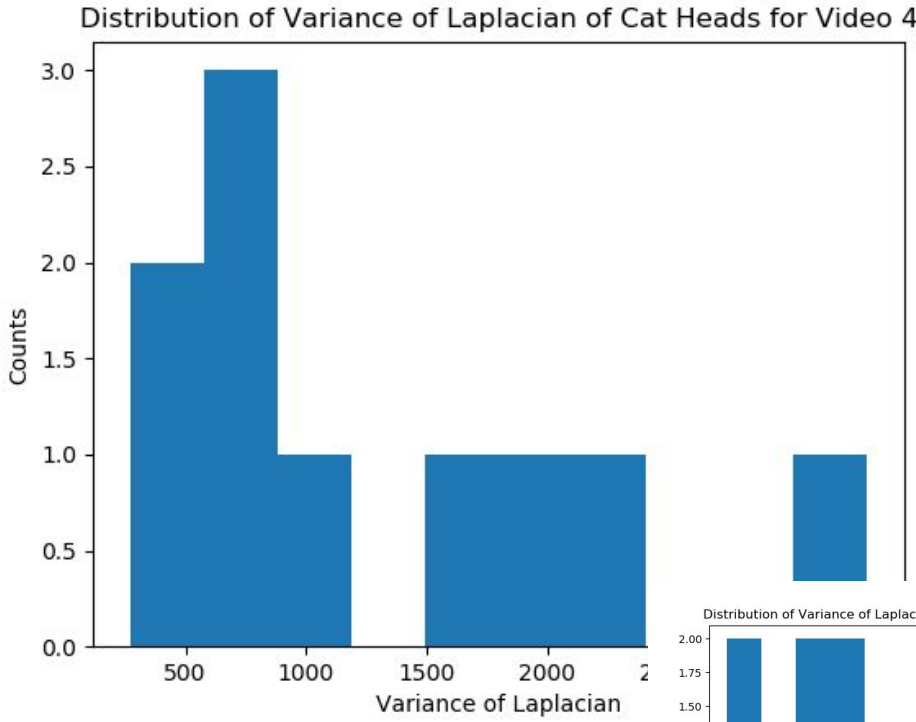
NEXT STEPS

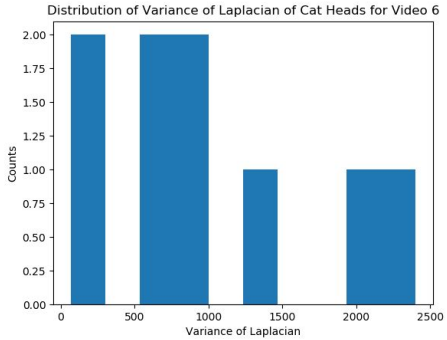
- Alternative Heuristics
- YOLO model
- Performance testing (double-blind experiments)

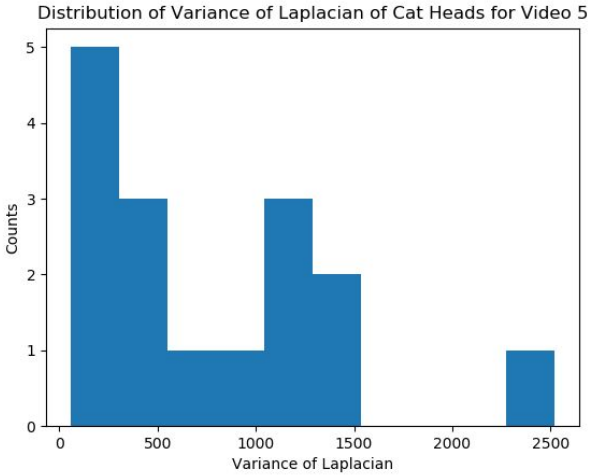
THANK YOU

Questions?

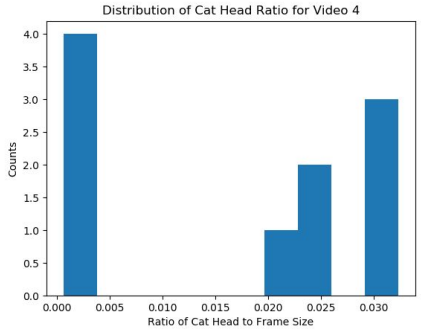
- A basic **problem statement,** including motivation and background info (4 minutes) => Shravan & Kyra
  - Statistics on euthanasia, etc.
  - Intro to sponsors
  - Basic problem statement (this and three points above are from lightning talk) -> State project
  - A summary of your **team** and **collaboration infrastructure =>** Kyra
    - Dog model
    - **Sponsors -> weekly-meetings, hands on, etc.**
    - **Github, Separate Jupyter notebook files for EDA, one canonical python file for finalized ideas that worked**
- Your **Scope of Work** (what are you focusing on; what are you **not** focusing on?) (1 minutes) => Emily
  - Our initial focus is on getting a good model / MVP
    - Prioritize: model building: a piece of software that given an image returns the best frame
      - This means we need to define "best"
    - Not priority: building out the web app itself.
      - While we want to build out an end-to-end pipeline, making it look pretty is not a priority
- **Learning goals and Domain understanding** – what did/do you need to learn in order to attack the problem (3 minutes)
  - Challenges => Kyra
    - Challenges with cats in the first place
    - Data was not at first what we envisioned... selectively choose it and also access some on the internet
  - Heuristics of a good cat photo => Andrew
  - What are some of the existing models present (see below)
- A brief description of **relevant knowledge** potentially including a **literature review**. (3.5 minutes) => Andrew
  - Haarcascade paper
    - This is where we discuss how it actually works
  - YOLO paper
    - (Something that we will attempt to look at for future model refinement)
- **Project ideas / Initial EDA** (minutes) => Emily & Shravan
  - Given the literature review, here is what we've decided to do at a high level: => Initial Jupyter notebook
    - Haarcascade detection using cats by this Joseph Howse guy
    - Laplacian variance for sharpness

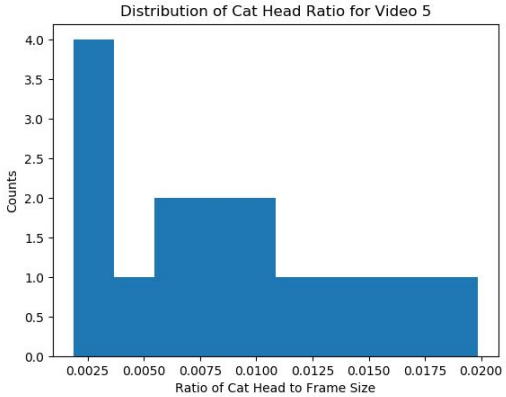

Distribution of Variance of Laplacian of Cat Heads for Video 4
Counts
Variance of Laplacian
0.0
0.5
1.0
1.5
2.0
2.5
3.0
500
1000
1500
2000

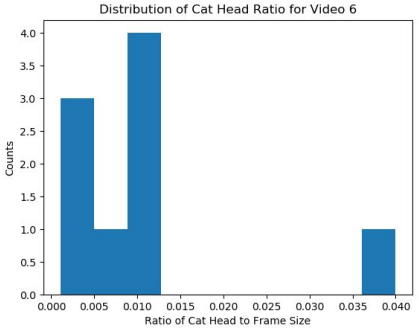

Distribution of Variance of Laplacian of Cat Heads for Video 6
Counts
Variance of Laplacian
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
0
500
1000
1500
2000
2500


Distribution of Variance of Laplacian of Cat Heads for Video 5
Counts
Variance of Laplacian
0
1
2
3
4
5
0
500
1000
1500
2000
2500

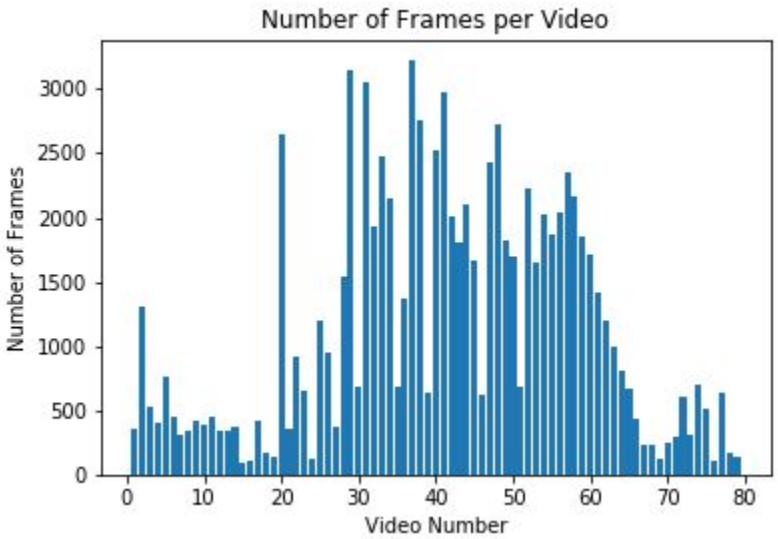
EDA

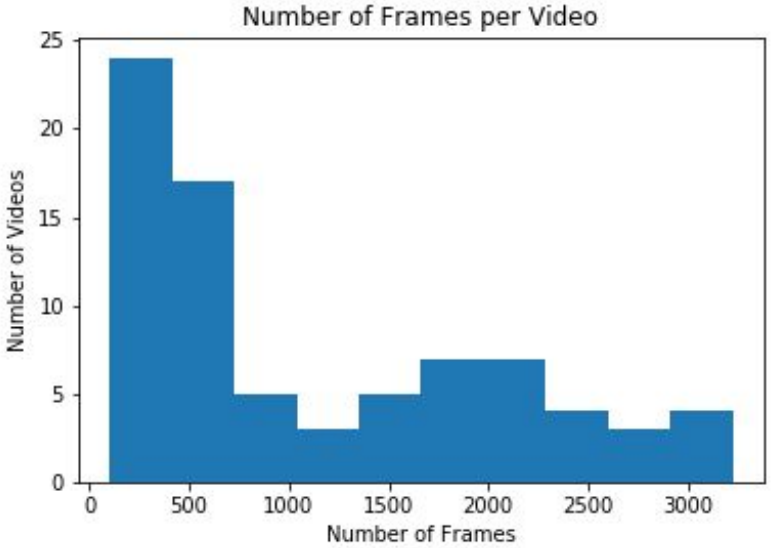

Distribution of Cat Head Ratio for Video 4
Counts
Ratio of Cat Head to Frame Size
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0


Distribution of Cat Head Ratio for Video 5
Counts
Ratio of Cat Head to Frame Size
0.0025
0.0050
0.0075
0.0100
0.0125
0.0150
0.0175
0.0200
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0


Distribution of Cat Head Ratio for Video 6
Counts
Ratio of Cat Head to Frame Size
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035
0.040
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0

EDA


Number of Frames per Video
Number of Frames
Video Number
0
500
1000
1500
2000
2500
3000
0
10
20
30
40
50
60
70
80


Number of Frames per Video
Number of Videos
Number of Frames
0
5
10
15
20
25
0
500
1000
1500
2000
2500
3000