

# 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.**

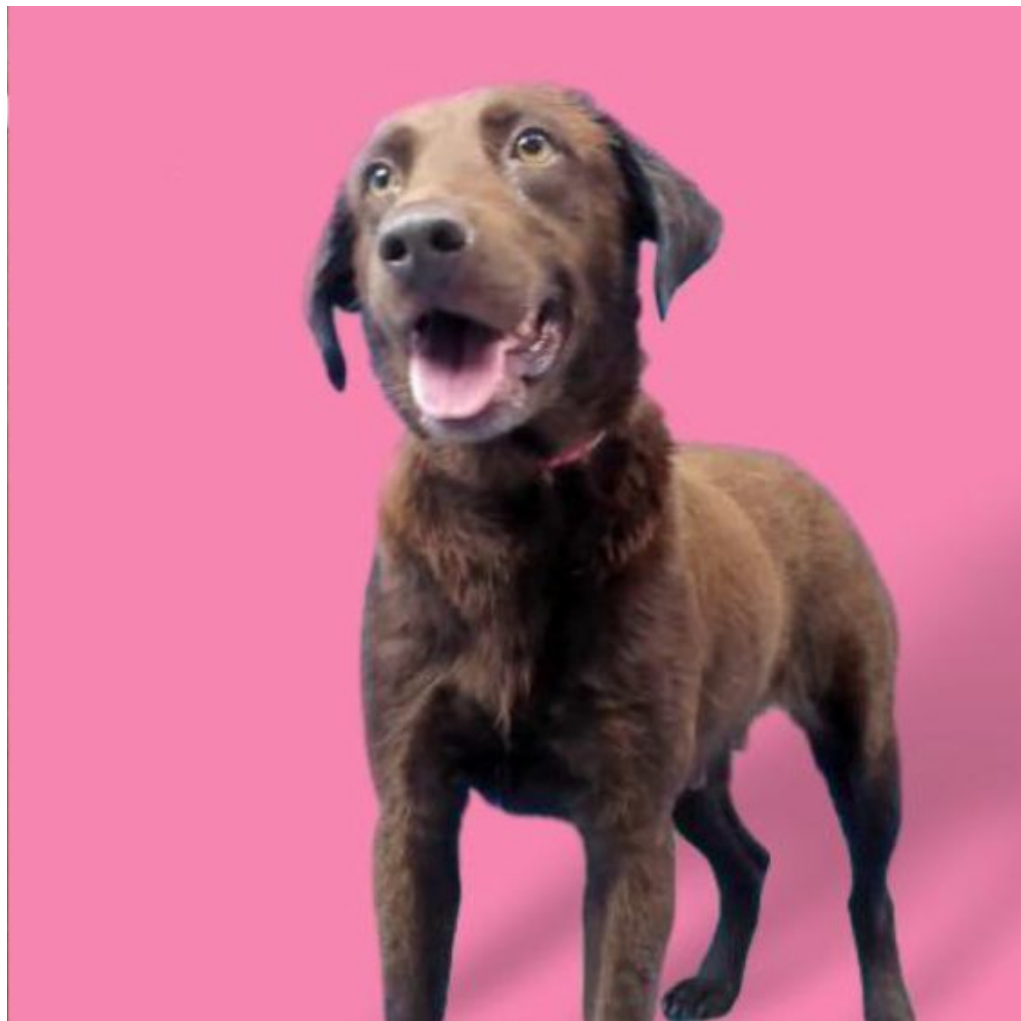


**ALL YOU NEED IS A  
CUTE CAT IMAGE OR VIDEO**



**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!**

**adoptimize®**



# 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



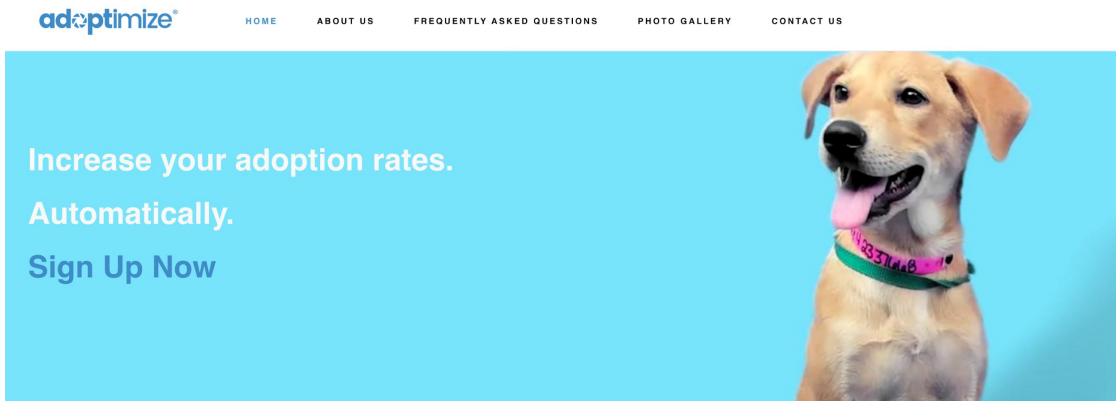
Saves Lives



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
<ul style="list-style-type: none"><li>○ Fur covering face</li><li>○ Not facing camera</li><li>○ Unwilling to sit still</li></ul>	<ul style="list-style-type: none"><li>○ Environment</li><li>○ Equipment and camera</li><li>○ Resources available</li></ul>

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



# CHALLENGES

Video Quality	Dataset
<ul style="list-style-type: none"><li>○ Filmed in kernels</li><li>○ Low light environment</li><li>○ Unstable camera</li><li>○ Camera quality (phone vs. laptop)</li></ul>	<ul style="list-style-type: none"><li>○ Small number of cat videos compared to dogs</li><li>○ No labeled data</li><li>○ Ambiguous: what makes a good photo?</li><li>○ Measuring success</li></ul>

The data initially poses some challenges...

Hello  
my name is

GOOD PHOTO?

# 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

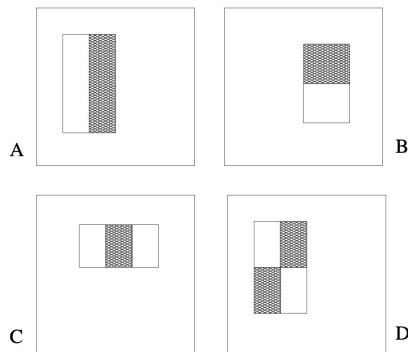
- Detection of Front-Facing Cat Head and Cat Eyes
- Variance of Laplacian
- Ratio of Cat Head Area to Frame Area



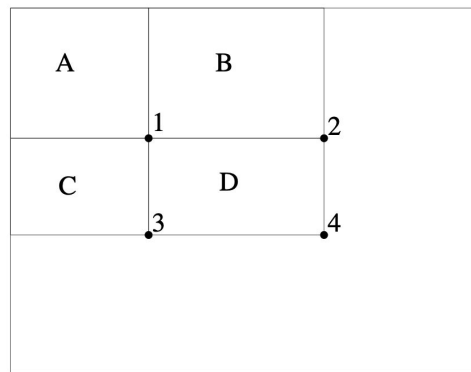


# LITERATURE REVIEW: VIOLA-JONES

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



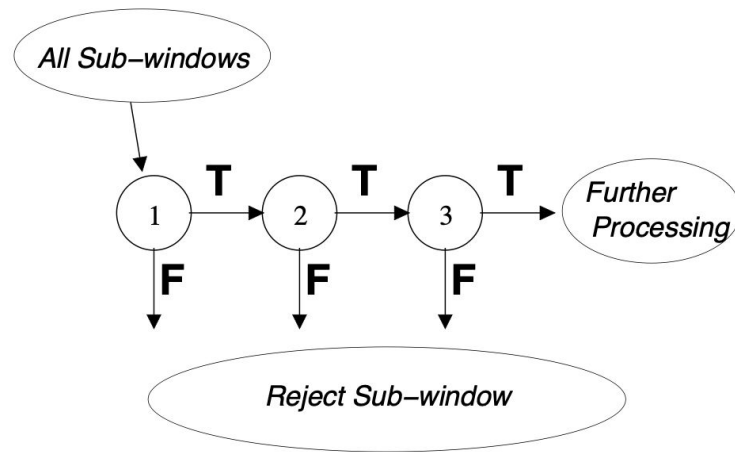
Haar-like Features



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

# LITERATURE REVIEW: VIOLA-JONES

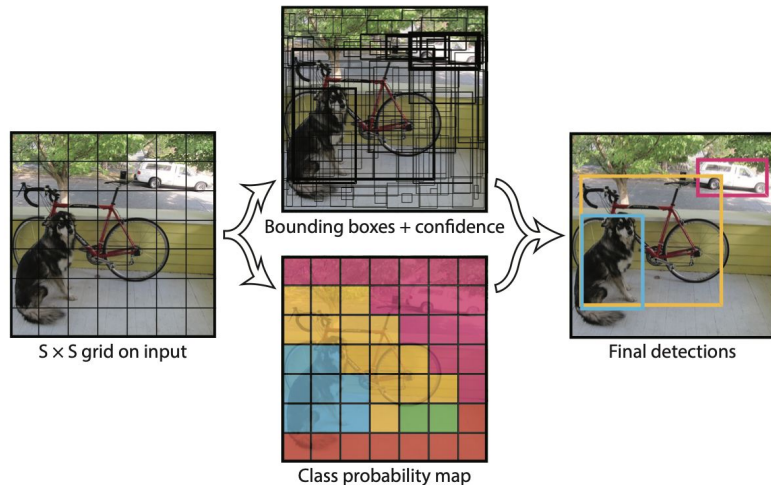
- *Rapid Object Detection using a Boosted Cascade of Simple Features*
- Haar-like Features
  - Pre-Compute Integral Image
- AdaBoost on Decision Stumps
- Cascade
- Sliding Windows



High-level view of cascade approach

# LITERATURE REVIEW: YOLO

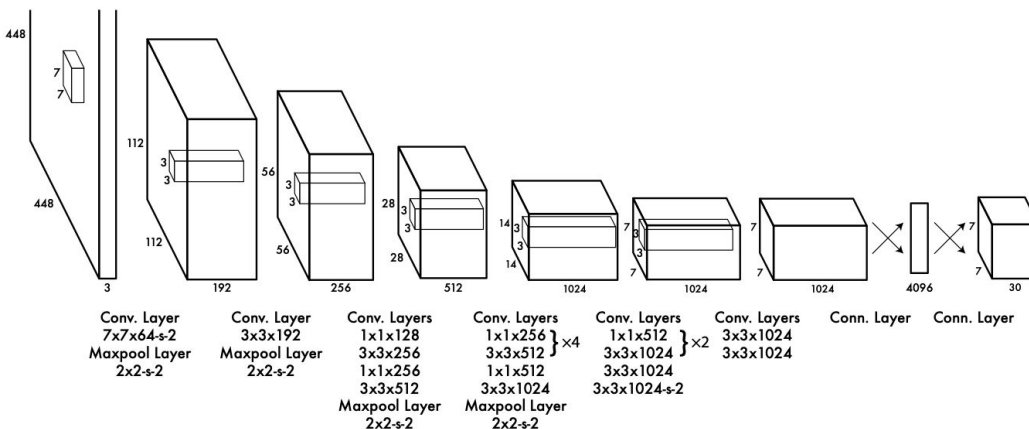
- *You Only Look Once: Unified, Real-Time Object Detection*
- Simultaneous Box and Class Proposal
- Simplicity: CNN
- Optimized for Speed



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

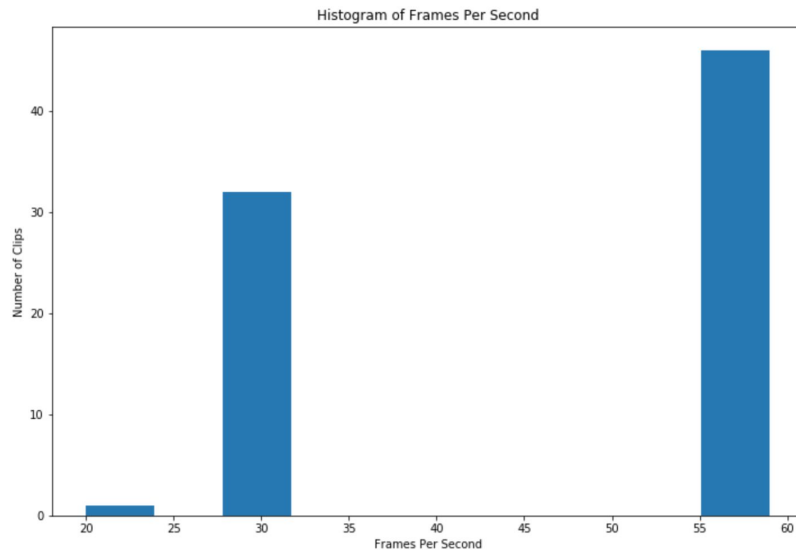
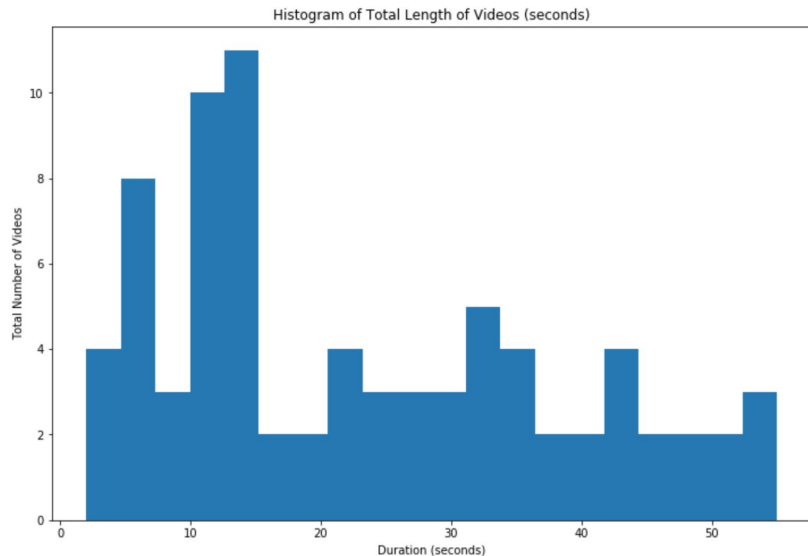
- *You Only Look Once: Unified, Real-Time Object Detection*
- Simultaneous Box and Class Proposal
- Simplicity: CNN
- Optimized for Speed



YOLO architecture; note only convolutional and fully connected layers

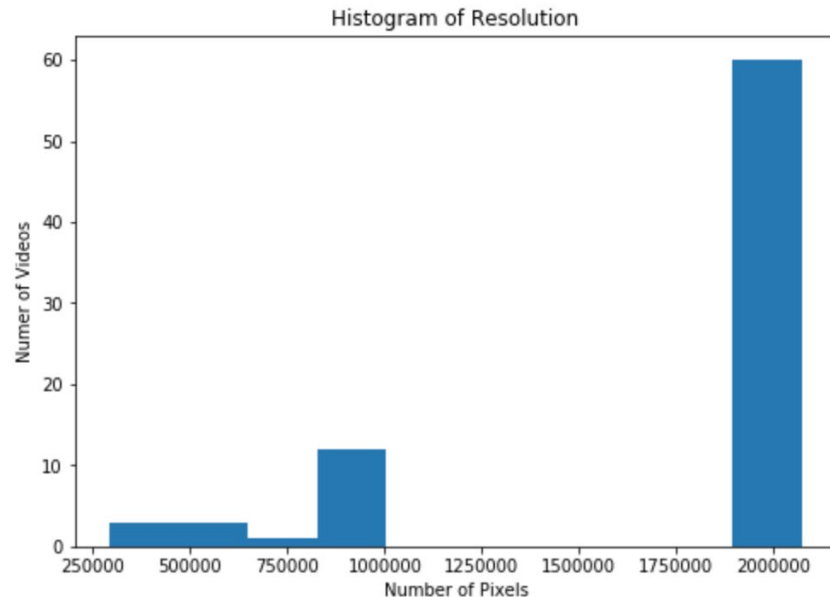
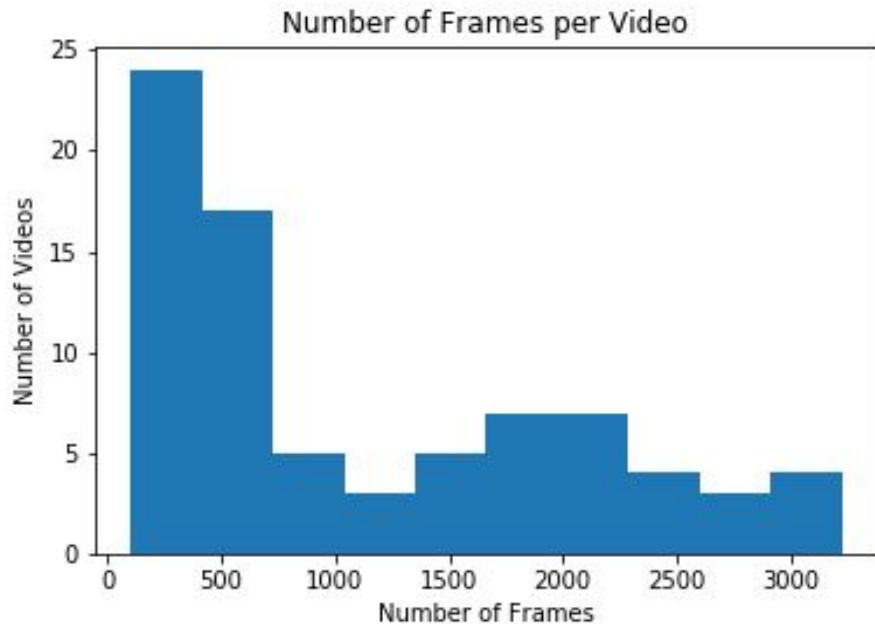
# EDA: GENERAL DATA SENSE

- 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



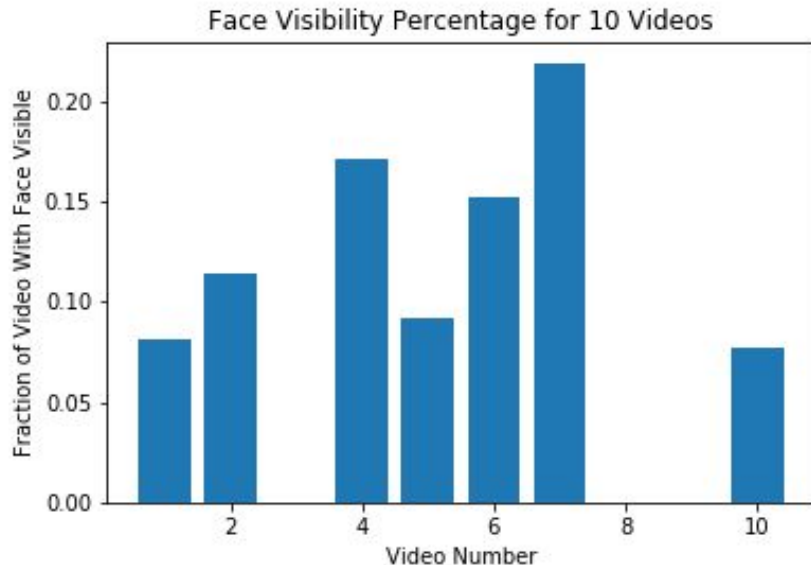
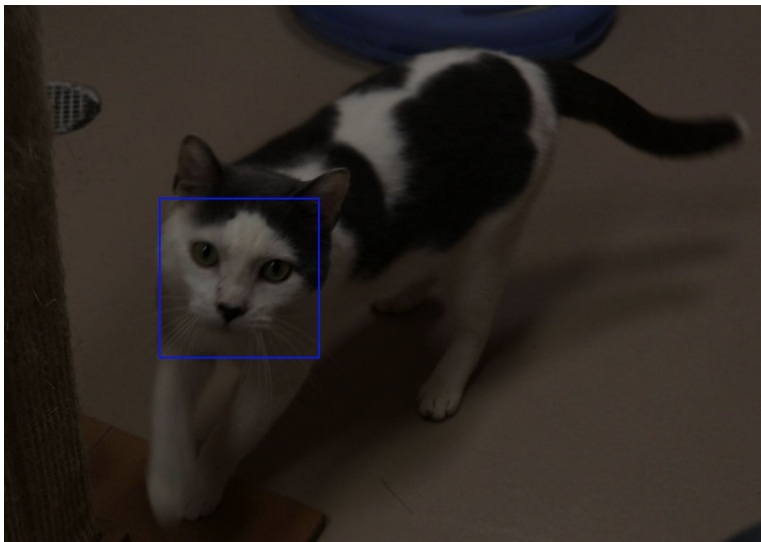


# EDA: GENERAL DATA SENSE

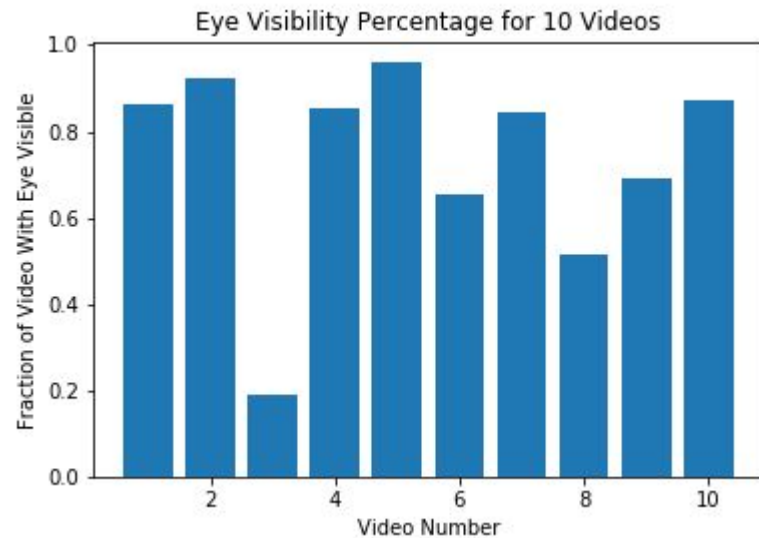
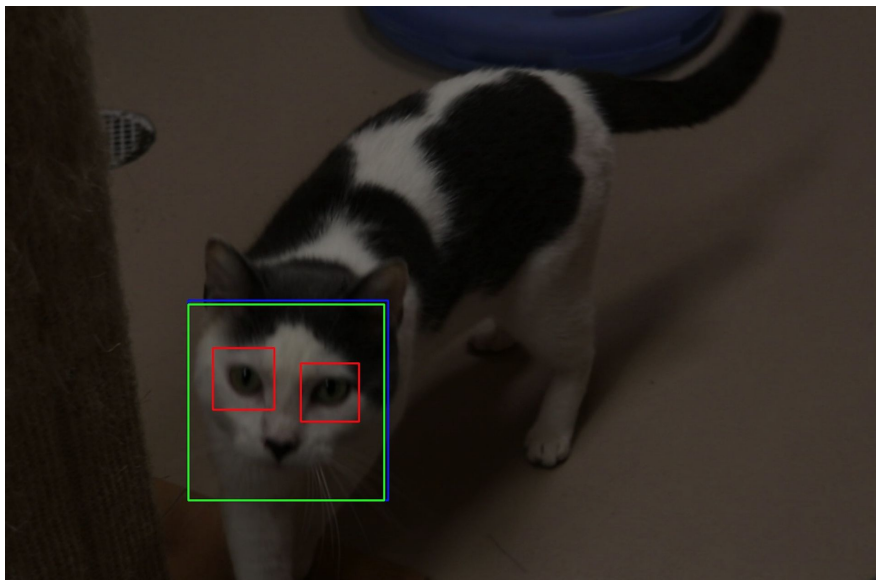


# EDA: CAT FACE DETECTION

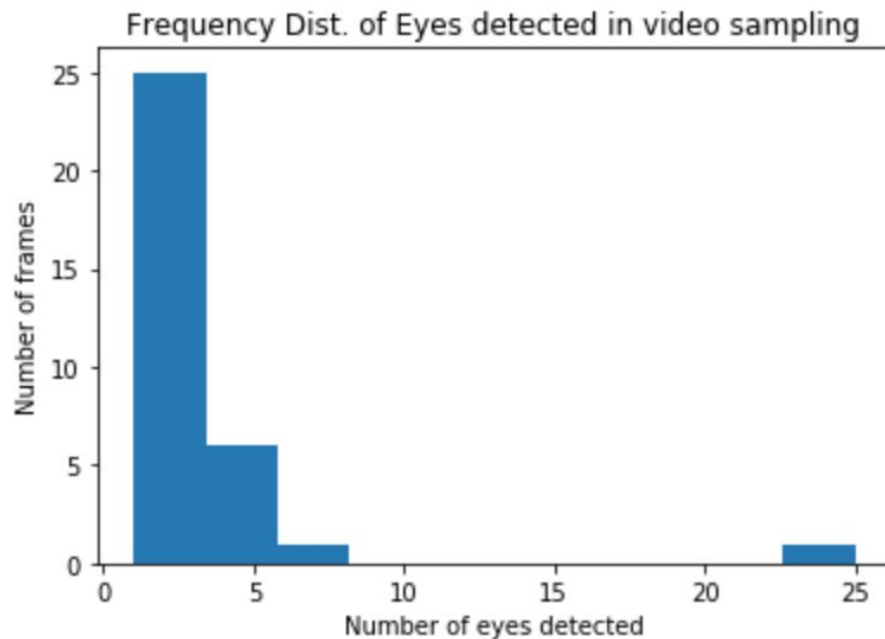
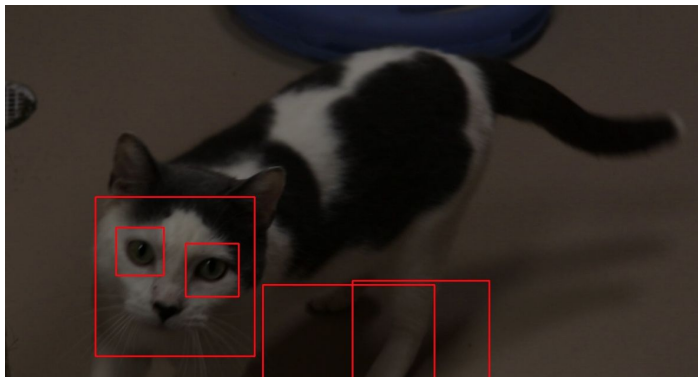
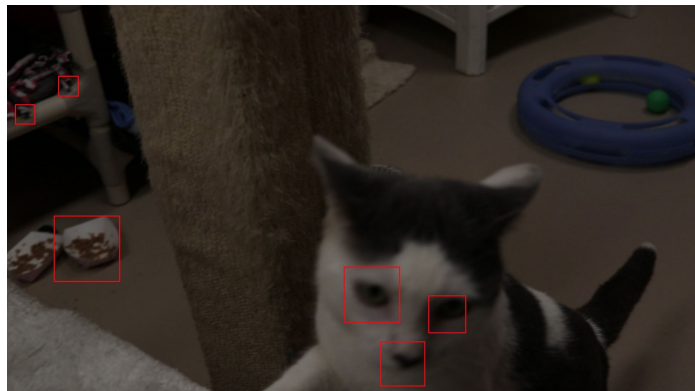
- Ran cat face detection using Haar Cascade
- Subsample of 10 videos
- Every 10th frame per video



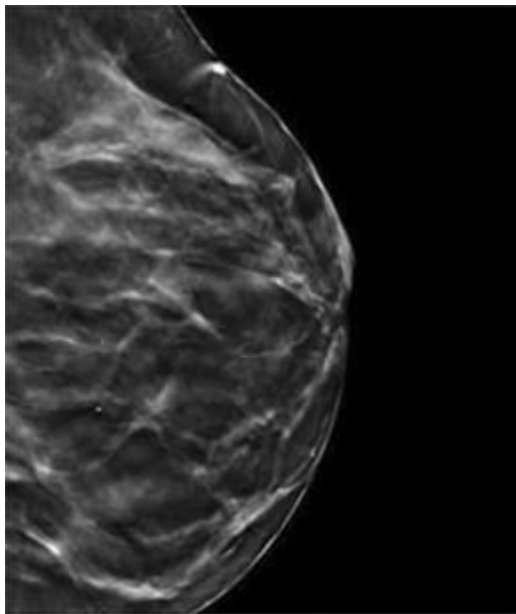
# EDA: EYE DETECTION



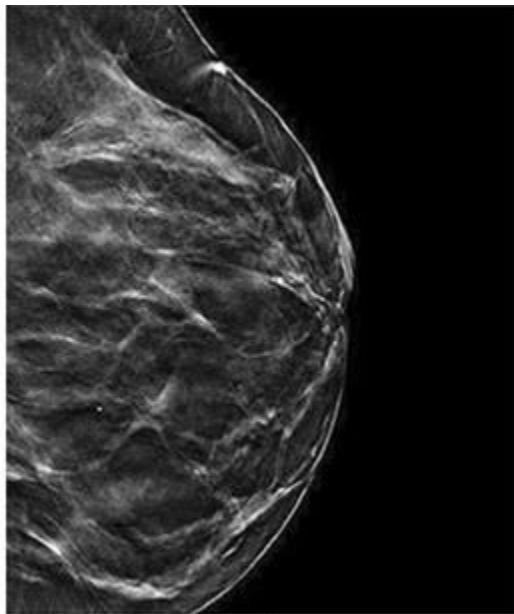
# EDA: EYE DETECTION



# EDA: MEASURES OF SHARPNESS



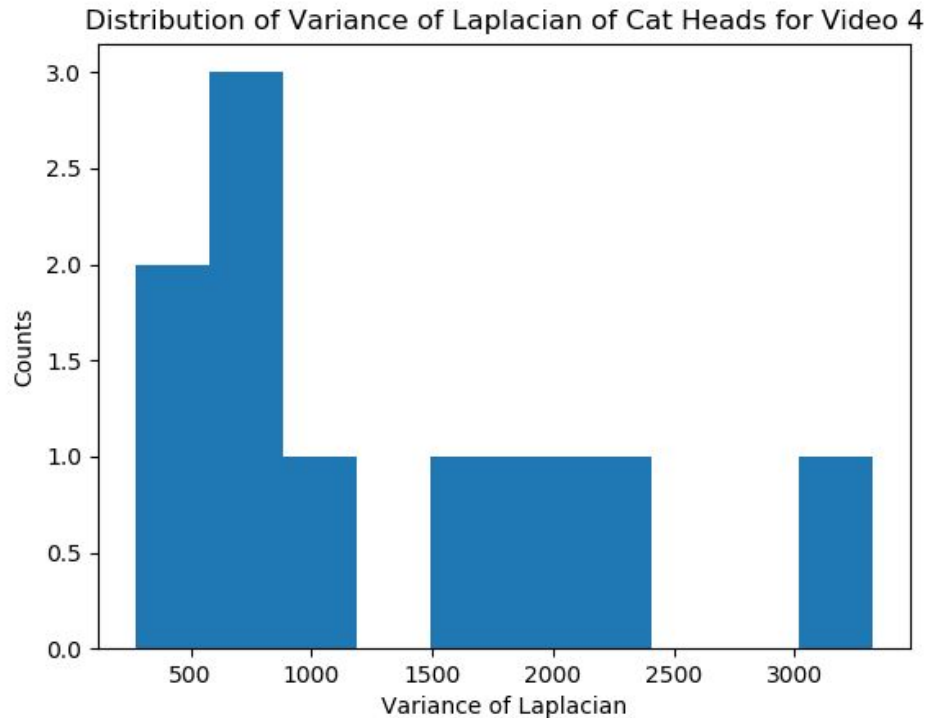
Less sharp



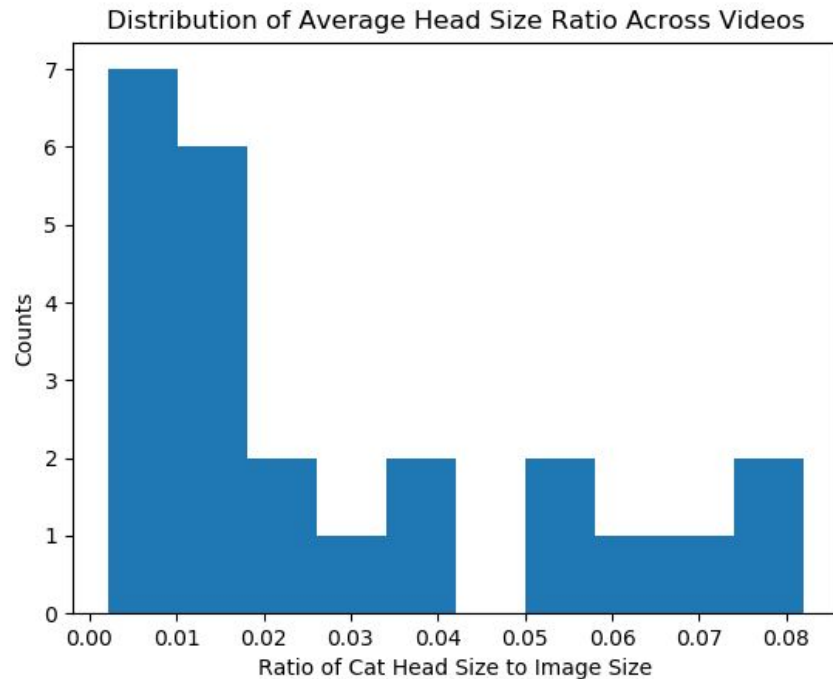
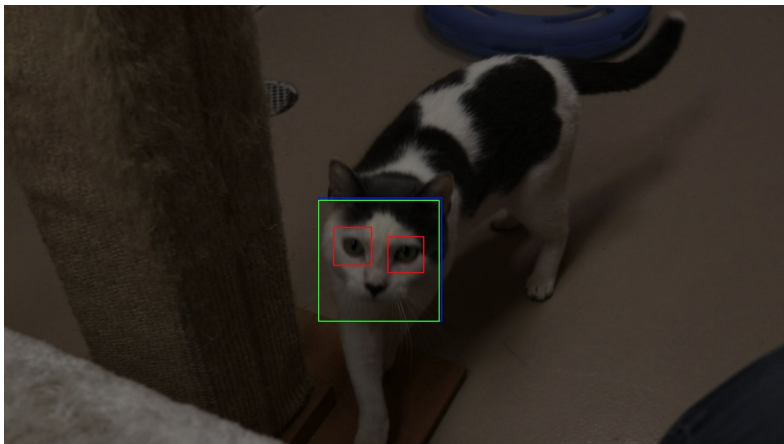
More sharp



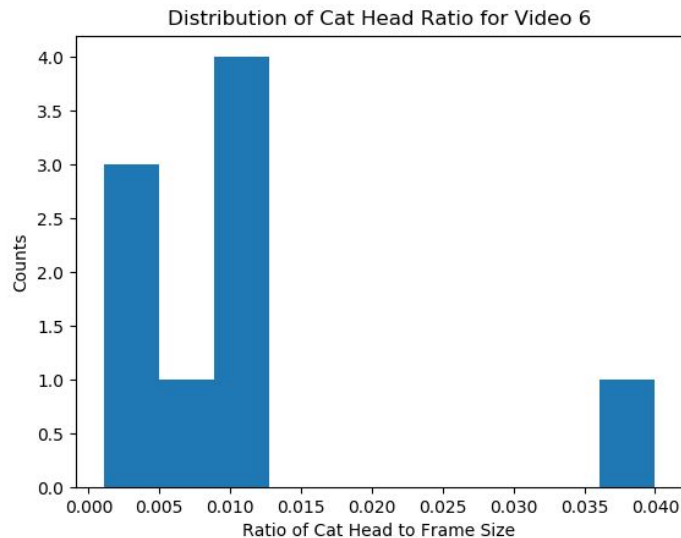
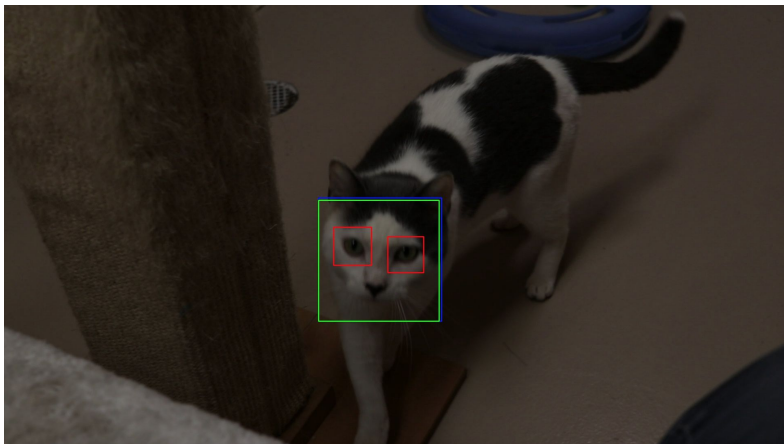
# EDA: MEASURES OF SHARPNESS



# EDA: HEAD SIZE RATIOS



# EDA: HEAD SIZE RATIOS



# NEXT STEPS

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



# THANK YOU

Questions?

