**BeaverDam: Video Annotation Tool for Computer Vision Training Labels**

by

Anting Shen

A thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science
in

Engineering - Electrical Engineering and Computer Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Kurt Keutzer, Chair
Professor Only Somewhat Important Guy

Fall 2016

BeaverDam: Video Annotation Tool for Computer Vision Training Labels

## Abstract

BeaverDam: Video Annotation Tool for Computer Vision Training Labels

by

Anting Shen

Master of Science in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Kurt Keutzer, Chair

Your abstract should be limited to 350 words. If it is longer, ProQuest will truncate and/or edit it so that it is less than 350 words. This is a very boring dissertation that nobody will probably ever read.

Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah.

Blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah

blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah.

And so it goes...

Professor Kurt Keutzer
Thesis Committee Chair

# Contents

## Acknowledgements

Thanks to Professor Keutzer and everyone in our group for the motivation and advice, Kosta for the great support, and Bichen Wu & Byung Gon Song for sharing in the struggles.

I'd also like to acknowledge Allen Wang, Sean Zhu, and Gabriel Arreola for contributing code to this project, coming up with awesome logos, and puns.

# Chapter 1

# Introduction

## 1.1 Introduction

Deep learning applications in recent years have come to require rapidly growing amounts of labeled training data. Often, accuracies can be boosted by adding data as much as by spending years on algorithmic development. For example, on the VOC07 benchmark, Faster-RCNN [1] with VGG-16 was able to eliminate 27.5% of errors in the much older R-CNN [2] backed by an equally old neural network architecture (mAP improved from 58.5 to 69.9). However, simply by including additional data from VOC12 and COCO, 29.5% of the remaining error was eliminated (mAP improved from 69.9 to 78.8). Therefore, for real-world application development, data can be cheaper and more effective than scientists. While many existing tools support image classification – it is even built into Amazon Mechanical Turk (MTurk) – and some tools support bounding box labeling in images, few tools exist for frame-by-frame labeling in videos. VATIC [3] stands out as being one of the best, as not only does it make high quality annotations one of its main goals, but also cost and scalability. Our work borrows and improves upon many concepts and results from VATIC's user studies, but we focus on two additional goals that we find extremely important in creating datasets for real applications in addition to research.

The first of these is researcher happiness. Although VATIC extensively tested its "User

Interfaces", we argue in chapter 2 that both the annotators and the experimenters are users, and the interfaces should be smooth for both when creating a tool.

## 1.2 Related Work

Vatic

Things vatic cite

# Chapter 2

# Experimenter Interface

Another great chapter!

Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah.

Figure 2.1. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah.

Blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah (see Figure 3.1) blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah

blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah.

say, “Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah.”

And so it goes...

# Chapter 3

# Annotator Interface

Another great chapter!

Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah.

Figure 3.1. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah.

Blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah (see Figure 3.1) blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah

blah blah blah blah blah blah blah blah blah blah blah blah. Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah.

say, “Blah blah blah blah blah blah blah blah blah blah blah blah blah blah blah.”

And so it goes...

# Chapter 4

# Conclusion

Ha! I'm done!

# References

[1] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” in *Advances in Neural Information Processing Systems (NIPS)*, 2015.

[2] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Computer Vision and Pattern Recognition*, 2014.

[3] C. Vondrick, D. Patterson, and D. Ramanan, “Efficiently scaling up crowdsourced video annotation,” *International Journal of Computer Vision*, pp. 1–21, 10.1007/s11263-012-0564-1. [Online]. Available: http://dx.doi.org/10.1007/s11263-012-0564-1

# Appendix A

# Allocated spectrum for communications

source: Comsearch [?]

| Category | Allocation |
|---|---|
| Microwave | 609 MHz |
| Broadcast | 423 MHz |
| Satellite | 188 MHz |
| Point-to-Multipoint | 203 MHz |
| PCS/Cellular | 193 MHz |
| ISM | 110 MHz |
| Other | 1274 MHz |

Table A.1. Spectrum under 3 GHz