



Large Cigar Detection using Computer Vision



Shuvam Keshari¹, Tanvi Anand¹, Juhan Lee², Rachel Ouellette², Dhiraj Murthy¹, Grace Kong²

¹Computational Media Lab, UT Austin, ²Department of Psychiatry, Yale School of Medicine

Introduction

- The presence of large cigars such as premium cigars and large manufactured cigars on social media is less studied
- The development of methods to automatically identify cigar content would facilitate enhanced surveillance of large cigar content on social media platforms
- We seek to use computer vision techniques to automatically detect large cigars using images from Reddit (specifically: relevant subreddits)
- Reddit has enabled the rapid spread of cigar-content. As of Dec, 2022, the 'r/cigars' subreddit has 161,000 members actively engaging in cigar-content, averaging 50 posts per day

Methods

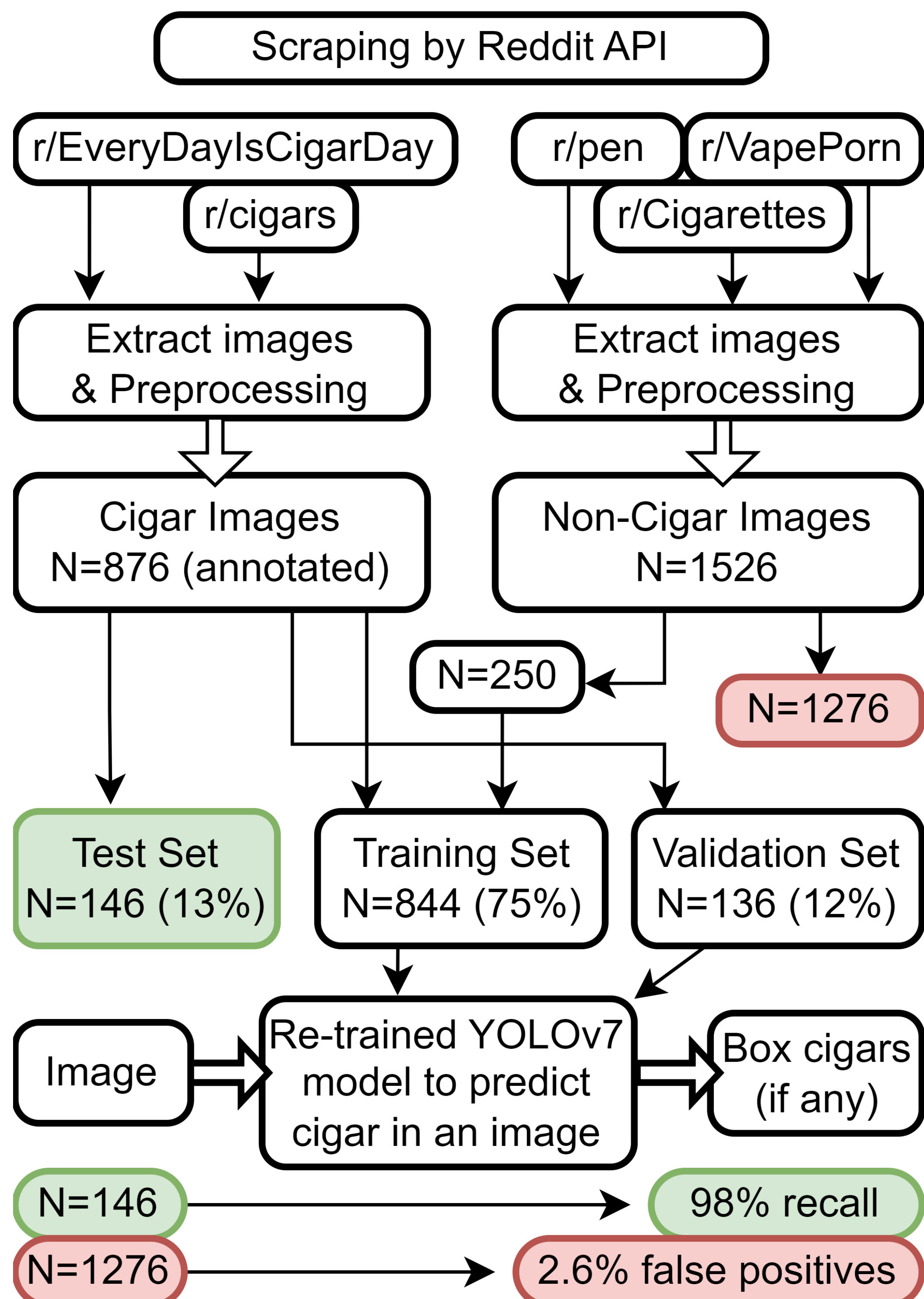
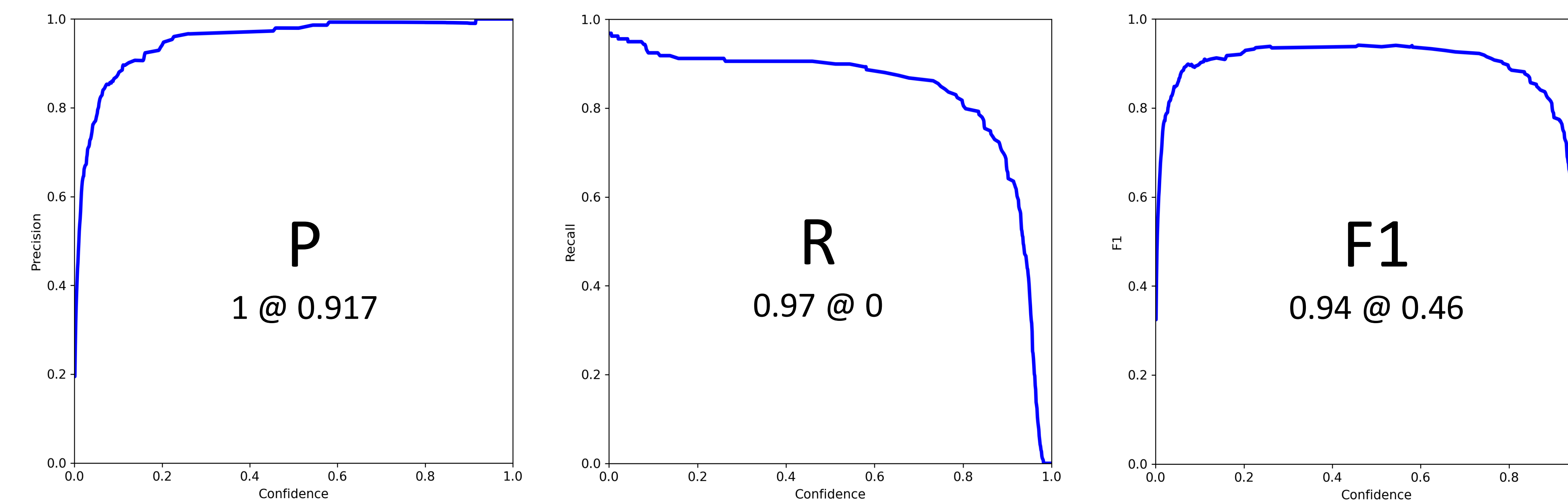


Image Categories and Model Results

Cigar and cigarette	Cigar among cigarettes	Cigar on a glass	Multiple cigars
Cigar detected (82% confidence)	Cigar detected (92% confidence)	Cigar detected (78% confidence)	Cigars detected (92%, 82% and 68% confidence)

Precision, Recall and F1 curves v/s confidence



Results

- Our model was able to correctly classify a large cigar with an F1 score of 0.983 and a recall value of 0.98 on the test set images that contain only large cigars (N=146).
- This means that it locates approximately 98 out of 100 images on average that contain cigar related content (Recall: Of all real positive cases, how many are predicted positive?)
- Our model was also able to distinguish cigars from similar objects (i.e. cigarettes, vapes, pens) with an accuracy of 97.4% (1276 non-cigar images) implying only 2.6% false positives
- We observed in our dataset that there were some social media posts/images that contain boxes of cigars or cigars placed consecutively together side by side which our model is not trained to identify. This can be pursued further to extend the research

Conclusion

- We developed automated, scalable computer vision models that successfully detected a wide range of large cigar images on Reddit and distinguished between large cigars and other similar looking products (e.g., cigarettes, pens and vapes)
- Our methods may facilitate quicker detection and monitoring of large cigar-related content and promotions on Reddit and other social media platforms
- Our method is especially useful in contributing automated and interpretable analyses of image/video data on a massive scale

Future Work

- Future work can extend our object detection model to detect other cigar related content on other social media platforms such as YouTube, Instagram, Snapchat, Facebook, etc.
- Additional work is needed to extend these frameworks to identify co-occurrence (e.g., if a cigar, hand, drink or cigar-cutter are in close proximity, it can be concluded that the person is smoking a cigar)
- Collaboration with regulatory agencies is needed to identify how these models can support regulation of tobacco content on social media

Acknowledgements

Funding: National Institutes of Health

Grant #: R01DA049878 (PI: Kong, G)

Disclosure: There is no conflict of interest in this study

Contact

Shuvam Keshari: skeshari@utexas.edu

Dhiraj Murthy: dhiraj.murthy@austin.utexas.edu