Untapped

Using Computer Vision to Tap into On-premise Brand Intelligence for Beverage Companies

Nick Caldicott and Derek Wong

Supervisor: Ashish Pujari

June 6th, 2020

THE UNIVERSITY OF CHICAGO
GRAHAM SCHOOL
CONTINUING LIBERAL AND PROFESSIONAL STUDIES

Background


STELLA ARTOIS
BLUE MOON
BELGIAN WHITE
THE SPIRIT OF THE FINGER LAK
PALE ALE
LAGUNITAS
IPA
INDIA PALE ALE
LAGUNITAS
GOOSE
FOUR STAR
PILS
Heineken
SLY FOX
O'Reilly's
STOUT

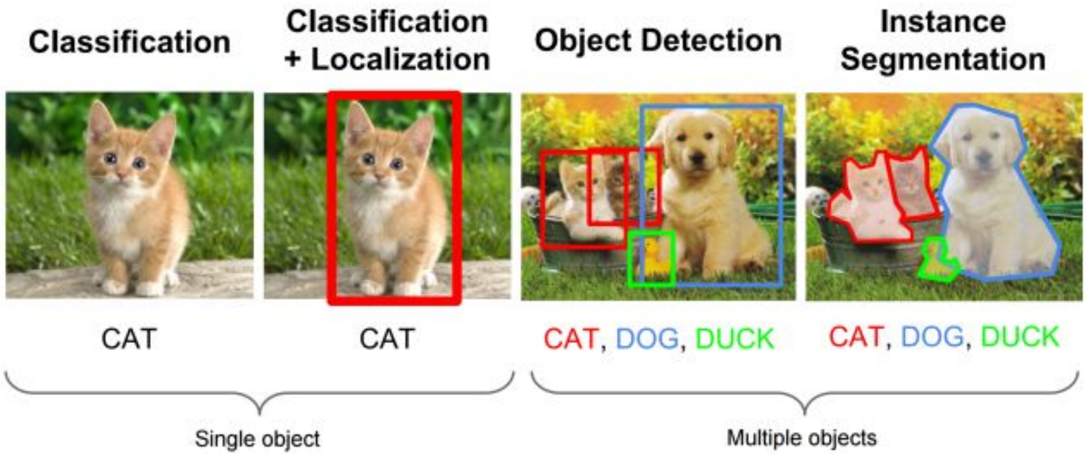
Computer vision 101

Computer Vision Tasks

Classification: “What single object is in this image?”

Object Detection: “What objects are in this image and where are they?”

Instance Segmentation: “What pixels belong to the objects in this image?”



Data


FIRESTONE
EASY JACK
IPA
UPSLOPE
BREWING COMPANY
LIMITED RELEASE
SAMUEL ADAMS
LAGUNITAS
PILS
OPHELIA
PYRAMID
THUNDER
BARREL

Defining the variables

Independent Variables

Red, Green, and Blue pixel values (0-255)

Original

Image Decomposition

R:	255	0	0
G:	0	255	0
B:	0	0	255

Dependent Variables

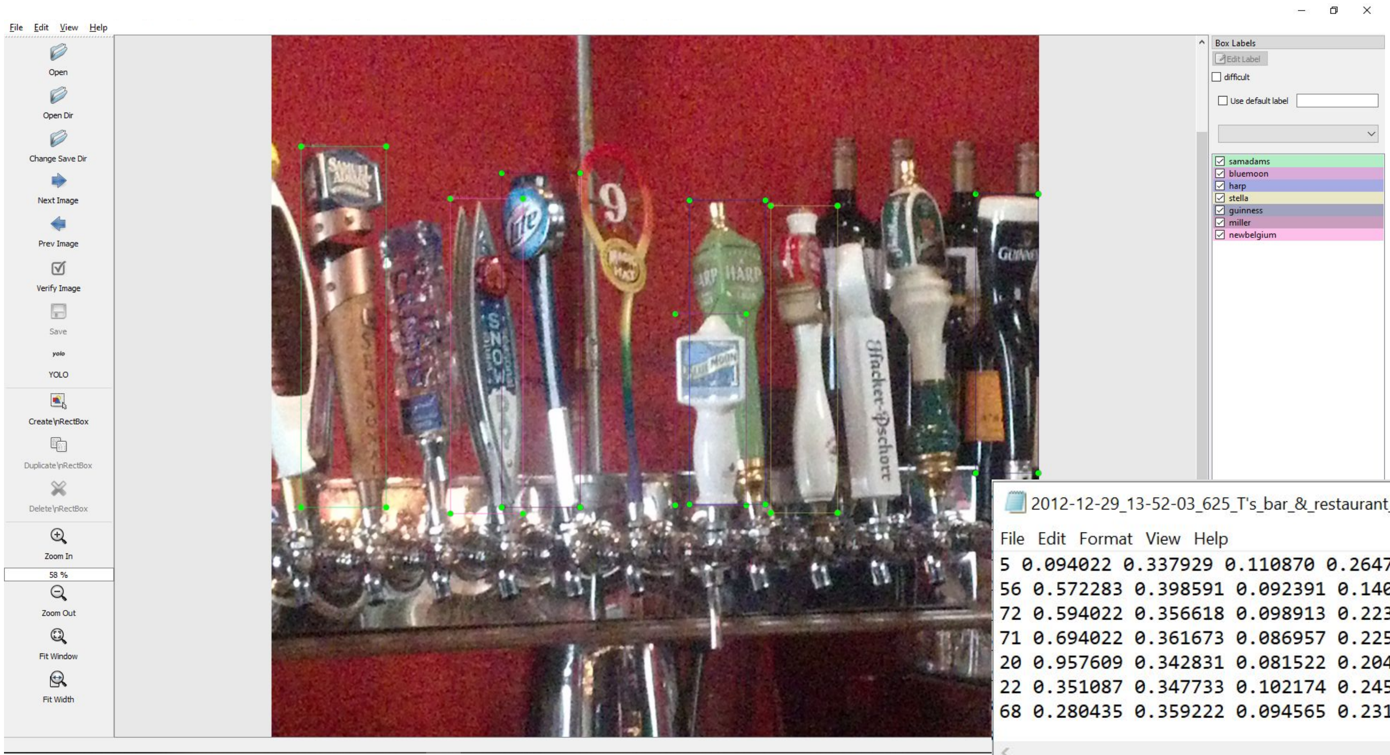
Bounding box and brand (class) of beer tap

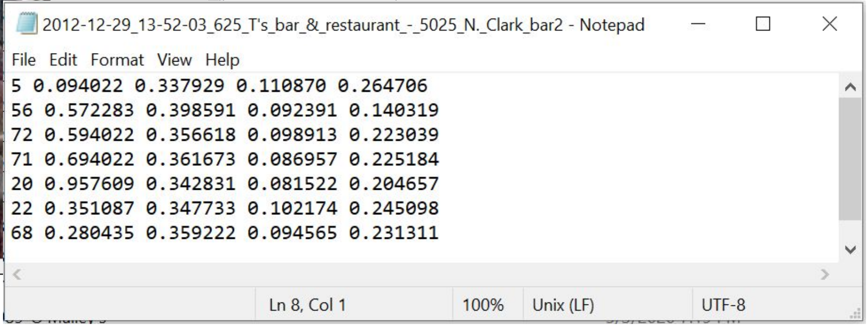
Image quality varies greatly

Good

Bad

Labeling images for consistent data capture





LabelImg application provides a quick and easy way to generate bounding box labels with corresponding underlying data.

Methodology

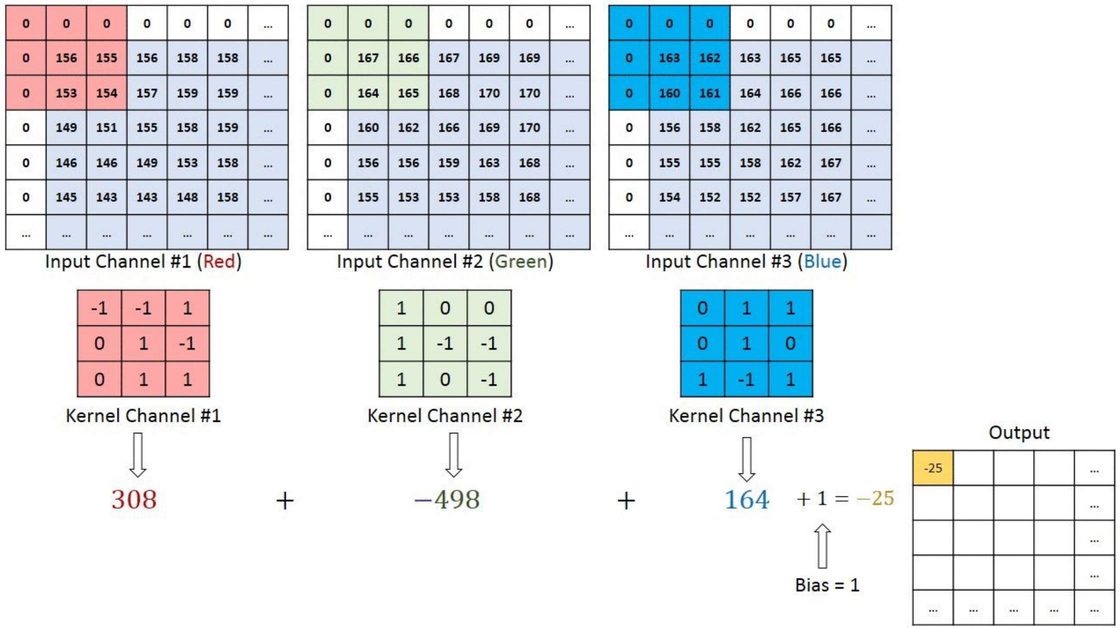
Feature extraction is critical for beer tap detection

Convolution

- Kernel weights slide left-to-right, top-to-bottom over an entire image
- Numeric pixel values are multiplied by the weights to derive output
- Operation performed iteratively to create transformations critical to detecting objects

Transfer Learning

- Weights from other models can be applied to our custom model to avoid relearning patterns common to most objects



“You Only Look Once” Model Overview


Create image grid


Generate class probability map


Plot bounding boxes
with confidence scores


Show final detections

Methodology

YOLO applies a single convolutional neural network to the entire image, divides the image into regions and predicts bounding boxes with corresponding probabilities for each class.

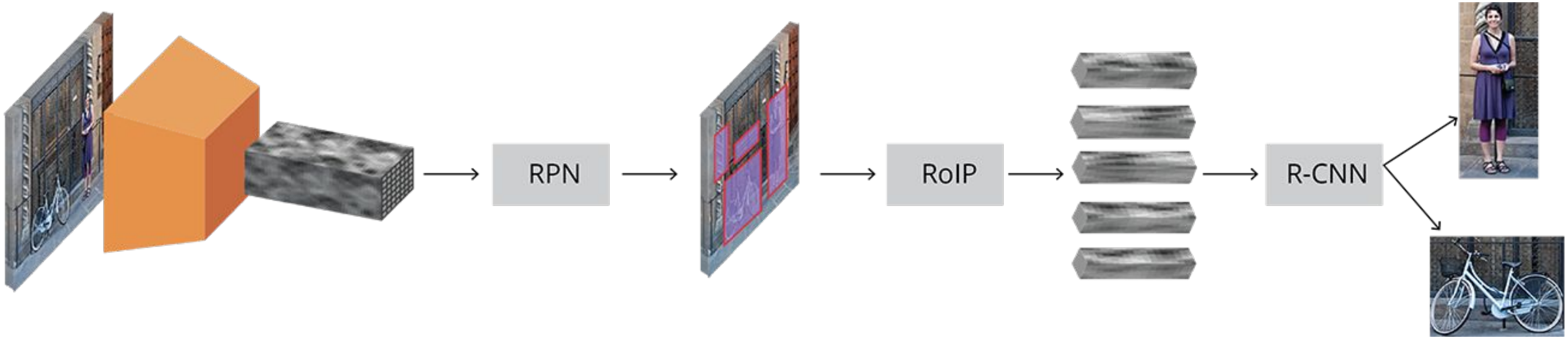
Benefits

- Higher precision, less incorrect predictions
- Supports real-time object detection

Limitations

- Tends to miss more detections
- Suffers with localization

Mask R-CNN Model Overview



Methodology

Mask R-CNN is a unified framework that first generates regions of interest and subsequently detects and predicts the probability of objects existing in these regions of interest.

Benefits

- High accuracy
- Great localization

Limitations

- Speed (compared to YOLO)
- Struggles with background noise

Evaluation Metrics

Recall: How many relevant taps are predicted?

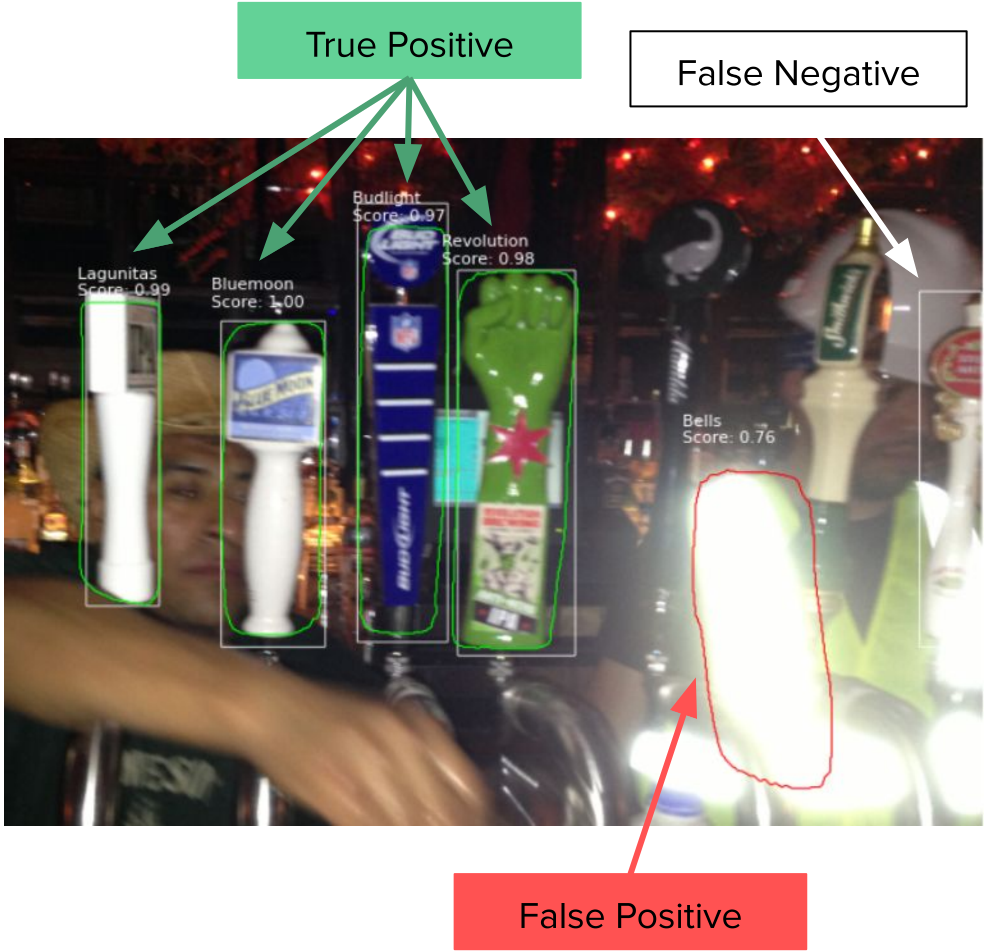
$$\frac{True\ Positive}{True\ Positive + False\ Negative}$$

Precision: How many predicted taps are relevant?

$$\frac{True\ Positive}{True\ Positive + False\ Positive}$$

Mean Average Precision (mAP)

The straight-average Precision at various levels of Recall



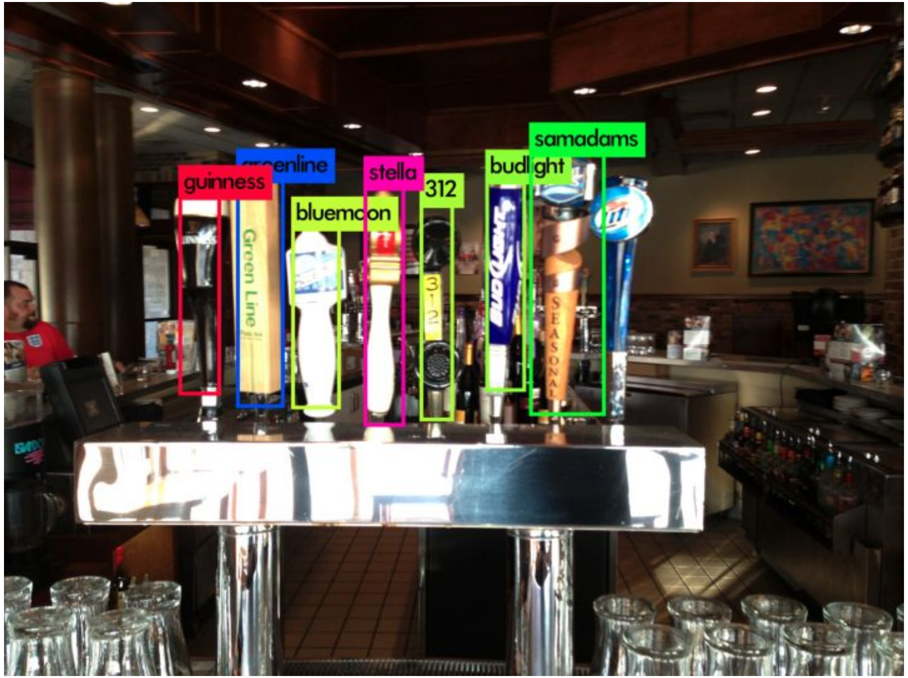
Findings


LAGUNITAS
IPA
INDIA PALE ALE
LAGUNITAS
THE
CENSORED
RICH COPPER ALE
Little Sumpin'
LIMITED RELEASE
LAGUNITAS

Head-to-head: Mask R-CNN vs YOLO

Mask R-CNN

YOLO



Head-to-head: Mask R-CNN vs YOLO

Mask R-CNN

YOLO

Mask R-CNN outperformed YOLO

Class	Mean Average Precision (on Test)		
	YOLO	Mask R-CNN	Difference
Stella	61%	94%	33%
Lagunitas	78%	92%	14%
Sam Adams	65%	91%	26%
Guinness	81%	92%	11%
Blue Moon	86%	93%	8%
Budlight	74%	91%	17%
312	81%	96%	15%
Green Line	85%	90%	5%
Goose Island	54%	85%	31%
Miller	47%	67%	20%
Founders	74%	82%	8%
Revolution	57%	93%	37%
Bells	60%	79%	19%
Half Acre	61%	78%	17%
3 Floyds	77%	78%	0%
Heineken	46%	54%	9%
Coors	53%	65%	12%
Budweiser	60%	78%	17%
Yuengling	73%	89%	16%
Allagash	80%	94%	14%
Total	**69.3%**	**86.5%**	**17.3%**
Total (top-5 test classes)	**73.7%**	**92.4%**	**18.7%**

Detection Quilt

Brand: Blue Moon

Recall: 86%

Precision: 90%

True Positive	False Negative	False Positive - No Ground Truth	False Positive - With Ground Truth

Detection Quilt

Brand: Green Line

Recall: 83%

Precision: 78%

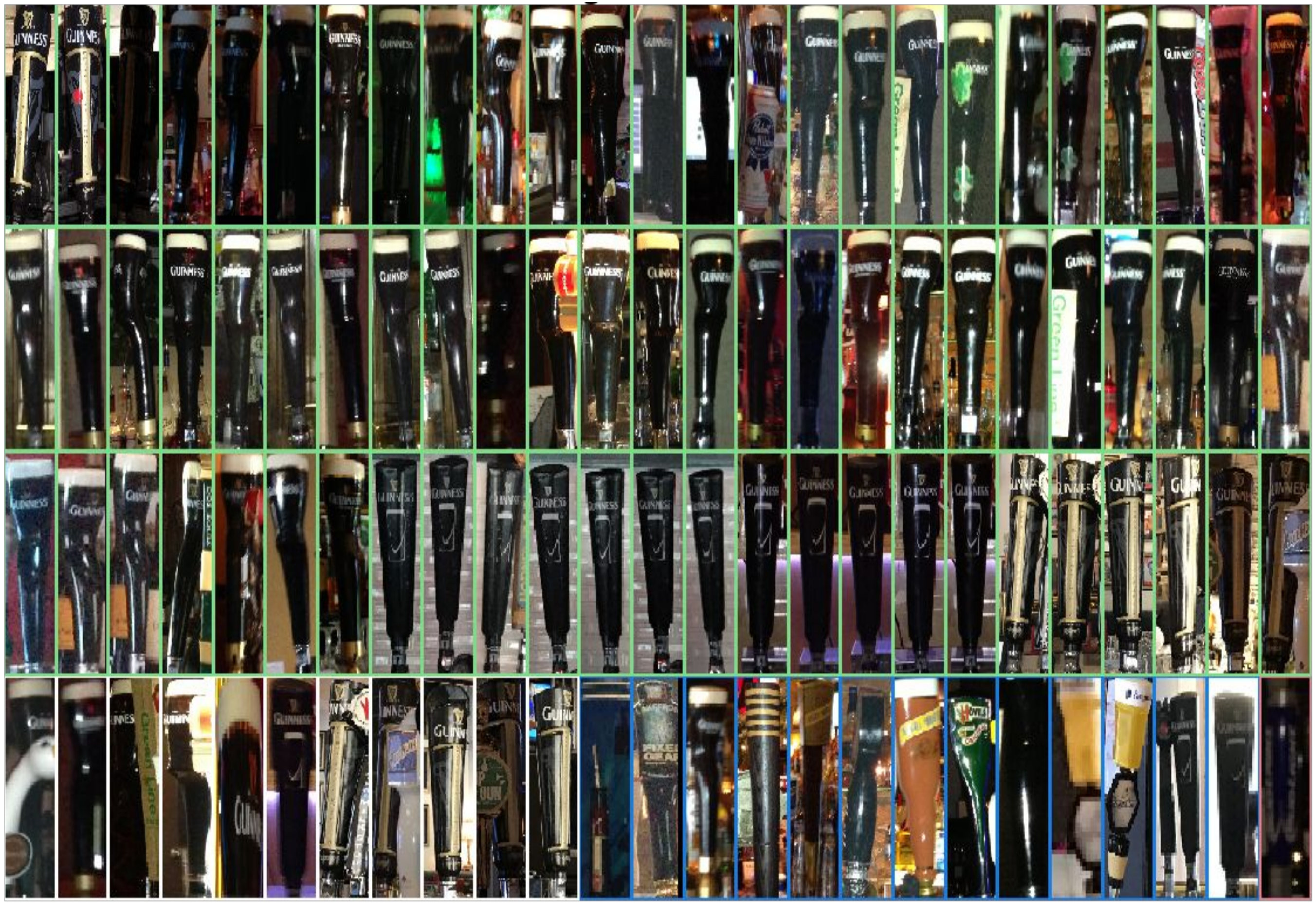
True Positive	False Negative	False Positive - No Ground Truth	False Positive - With Ground Truth

Detection Quilt

Brand: Guinness

Recall: 87%

Precision: 83%



True Positive	False Negative	False Positive - No Ground Truth	False Positive - With Ground Truth

Cheers!