Datasets ready-to-use

Mar 3, 2025 Ke, Wang, Wang, Gu, Sun

Charter Updated!

Gameplans...

PC/Laptop Identification using scene text detection and recognition (STDR) - Focus on tag extraction
- Two Stage Process

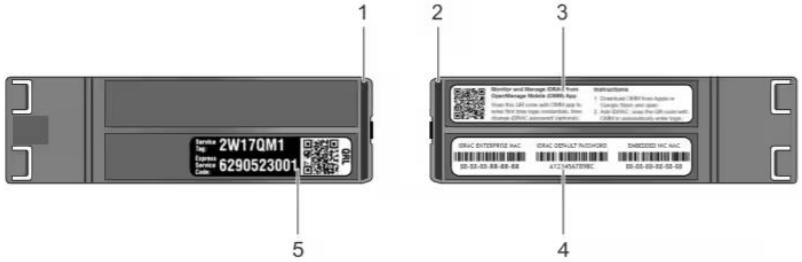
o Stage 1: High resolution camera that identifies text and label

Use YOLO (maybe V7) to create an ROI around the text (object detection)

Do some transforms to make text easy to read

o Stage 2:

Extract the text and read it
Provide the deciphered text



Jerome Barczykowski

Digitally signed by Jerome Barczykowski
DN: cn=Jerome Barczykowski, o=D3 Embedded, ou, email=jerome@d3embedded.com, c=US
Date: 2025.02.28 16:32:37 -05'00'

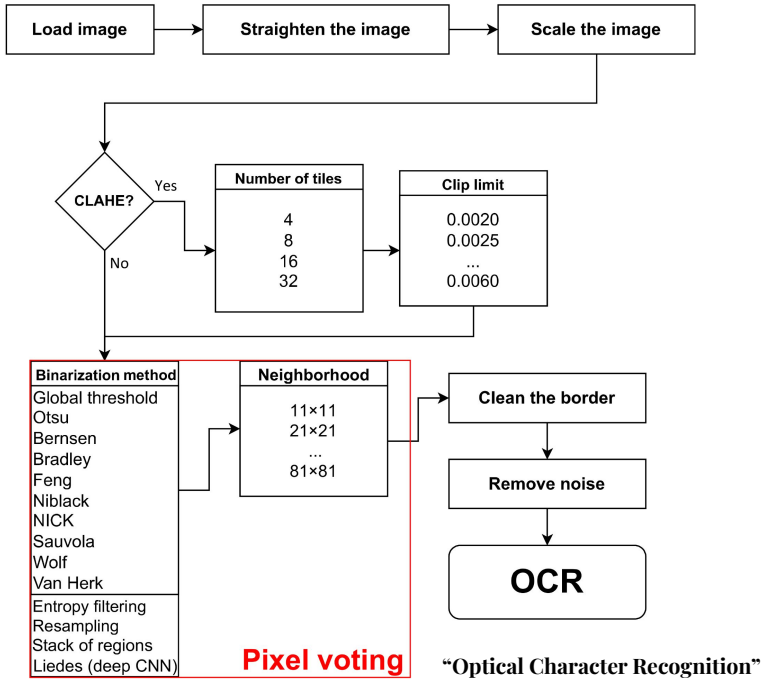
“Pixel Democracy” - Malinkski & Okarma (2023)

Scaling: Images are **upscaled** (e.g. via bicubic interpolation) to clarify small text or **downscaled** to reduce noise and computational Proper scaling ensures small IC markings are large enough for OCR without excessive noise.

Straightening (Deskew): Images are rotated to align text horizontally, preventing character skew. This is typically done by edge detection and a Hough transform to find the text Straightening reduces recognition errors caused by tilted characters.

Contrast Enhancement: Histogram equalization (especially CLAHE – contrast-limited adaptive histogram equalization) redistributes intensities to improve contrast This makes faint markings more distinguishable. CLAHE is used instead of basic equalization to avoid amplifying

Morphological Cleaning: Operations like **opening**, **closing**, and **boundary cleaning** are applied (usually after initial binarization) to remove small noise blobs and smooth character edges. This cleans up spurious pixels and gaps so that OCR reads continuous character shapes.



How to vote?

Binarization is the process of converting a grayscale image (which has pixel values ranging from 0 to 255) into a binary (black-and-white) image, where each pixel is either **black (0)** or **white (1)**, making OCR life easier.

Our candidates (at the same time the VOTERS!!!):
Otsu's Method (Global)
Bernsen's Method (Local)
Van Herk's Method (Morphological)
Bradley's Method (Integral Image)
Niblack's Method (Statistical Local)
Sauvola's Method (Adaptive Local)
Wolf's Method (Contrast-Normalized Local)
Feng's Method (Contrast-Enhanced Local)
NICK Method (Variance-Adjusted Local)
Entropy Filtering (Preprocessing-Based)
Stack of Regions Method (Multi-Layer Adaptive)
Resampling Method (Background Estimation)

Voting process: (each pixel of in an image is actually a ballot)

Apply Multiple Binarization Methods

Each method generates a **binary image** (where each pixel is either **black (0)** or **white (1)**).

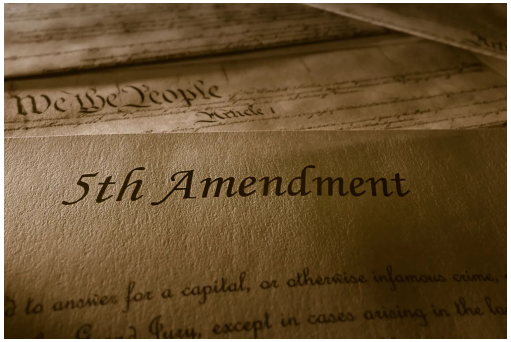
Pixel-Wise Majority Voting

For each pixel location, take the **binary output from all 11 methods**.

The final pixel value is set to the **majority vote** (i.e., if most methods classify it as black, it remains black; otherwise, it turns white).

Optimize the Voting Set (pruning the worst two candidates)

Result? They plead the fifth...



"Even for evenly illuminated images, the task of reading IC decals remains challenging for our current methods. This is expected since the decals are **difficult** to read even for the human eye. Therefore, **future research should focus** on **developing more advanced** OCR algorithms and image processing techniques to accurately recognize IC decals.

One of potential directions for further research is to c**ombine the strengths of different binarization methods** into a single approach. The best results may be achieved by using statistical or voting-based approaches [47] to choose the most appropriate binarization method for each image based on the specific characteristics of the image. The combination of five adaptive binarization algorithms based on pixel voting proposed in the paper leads to very promising results.
"

electronics

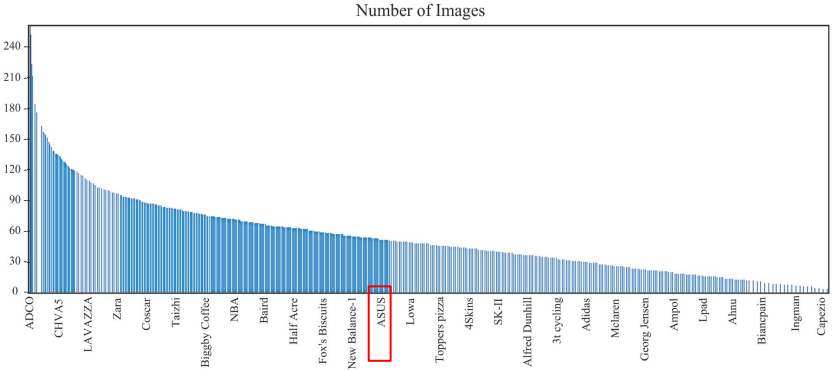
an Open Access Journal by MDPI

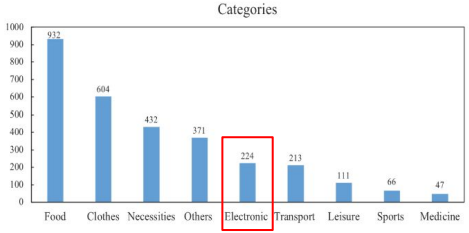
IF: 2.9

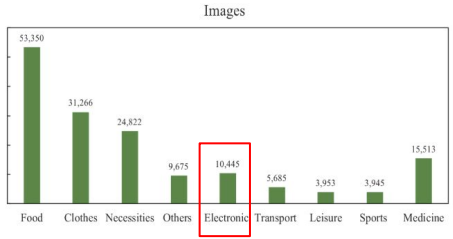
LogoDet - 3K

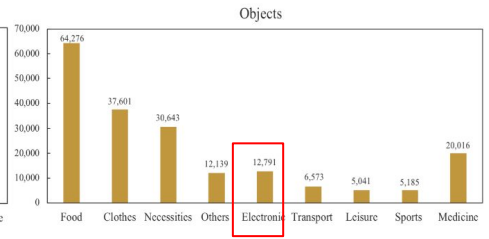
Overview

LogoDet-3K, the largest logo detection dataset with full annotation, which has **3,000 logo categories**, about **200,000 manually annotated** logo objects and **158,652 images**. LogoDet-3K creates a more challenging benchmark for logo detection, for its higher comprehensive coverage and wider variety in both logo categories and annotated objects compared with existing datasets.







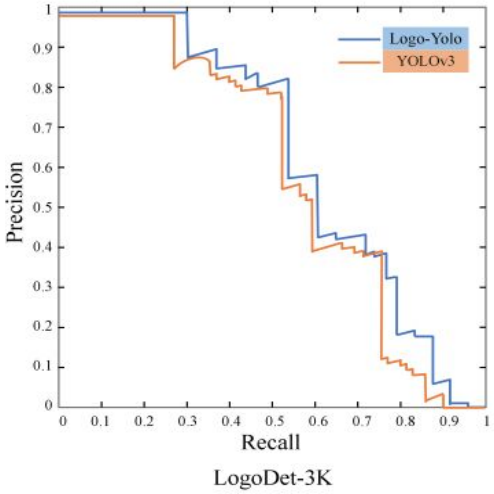


Pretrained Model: LogoDet-3K (Wang et al., 2020)

TABLE V: Comparison of baselines on different benchmarks (%).

Benchmarks	Methods	Backbones	mAP
LogoDet-3K-1000	Faster RCNN [13]	ResNet-101	45.16
	SSD [14]	VGGNet-16	43.32
	RetinaNet [22]	ResNet-101	52.10
	FPN [41]	ResNet-101	49.63
	Cascade R-CNN[35]	ResNet-101	48.14
	Distance-IoU [23]	DarkNet-53	53.06
	YOLOv3 [15]	DarkNet-53	55.21
	Logo-Yolo	**DarkNet-53**	**58.86**
LogoDet-3K-2000	Faster RCNN [13]	ResNet-101	41.86
	SSD [14]	VGGNet-16	38.97
	RetinaNet [22]	ResNet-101	49.00
	FPN [41]	ResNet-101	47.91
	Cascade R-CNN[35]	ResNet-101	46.32
	Distance-IoU [23]	DarkNet-53	51.69
	YOLOv3 [15]	DarkNet-53	52.32
	Logo-Yolo	**DarkNet-53**	**56.42**
LogoDet-3K	Faster RCNN [13]	ResNet-101	38.30
	SSD [14]	VGGNet-16	34.47
	RetinaNet [22]	ResNet-101	44.32
	FPN [41]	ResNet-101	42.84
	Cascade R-CNN[35]	ResNet-101	41.23
	Distance-IoU [23]	DarkNet-53	46.34
	YOLOv3 [15]	DarkNet-53	48.61
	Logo-Yolo	**DarkNet-53**	**52.28**

"Logo-Yolo, which incorporates Focal loss and CIoU loss into the state-of-the-art YOLOv3 framework for large-scale logo detection. Logo-Yolo can solve the problems of multi-scale objects, logo sample imbalance and inconsistent bounding-box regression. It obtains about 4% improvement on the average performance compared with YOLOv3, and greater improvements compared with reported several deep detection models on LogoDet-3K." (Wang et al., 2020)

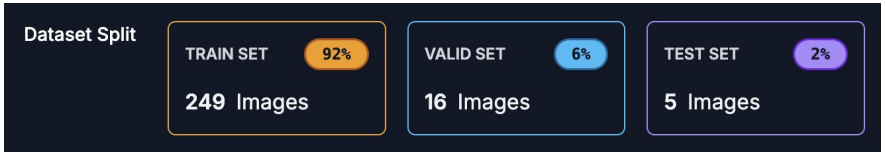


Licence:
Obtained through request: Google Drive link: https://drive.google.com/file/d/1GOYqRy7yoBVAt9Ih375NWtoZZTXN6OET/view?usp=sharing

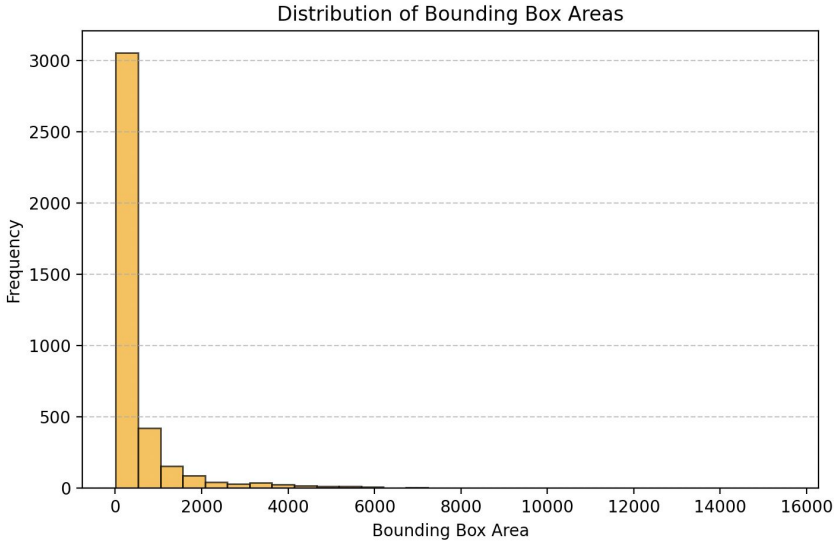
Non-Electronic Text Detection Dataset: Beverage

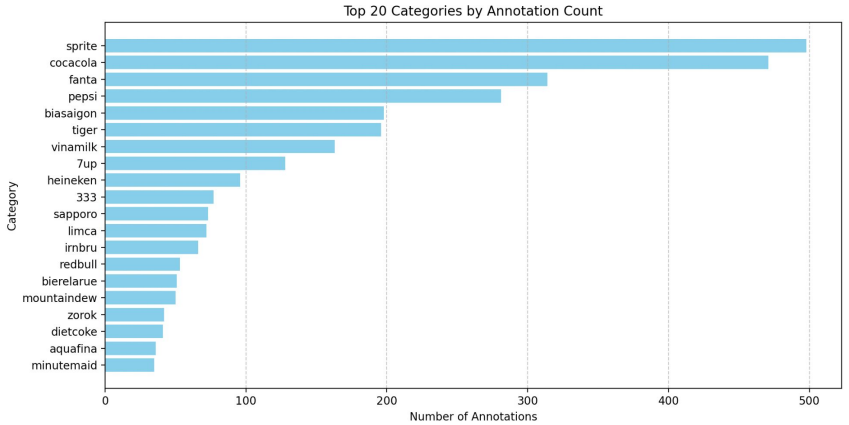
- 104 images and 124 beverage classes (like 7up, A&W, Coca-Cola, etc.)
- Available under a CC BY 4.0 license
- 1.3k views and 38 downloads .

Source: Roboflow
https://universe.roboflow.com/baovippro318-gmail-com/beverage-9ox8m



Non-Electronic Text Detection Dataset: Beverage



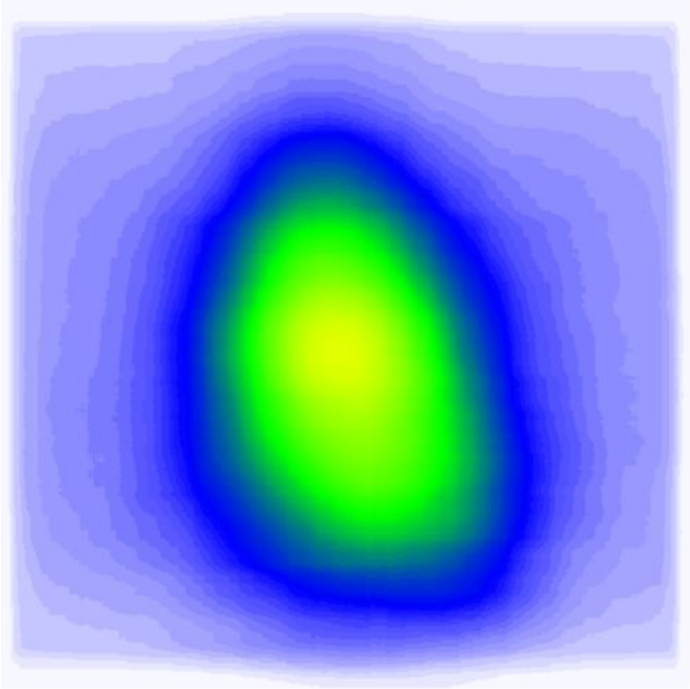


Non-Electronics Detection Dataset: License Plate Recognition

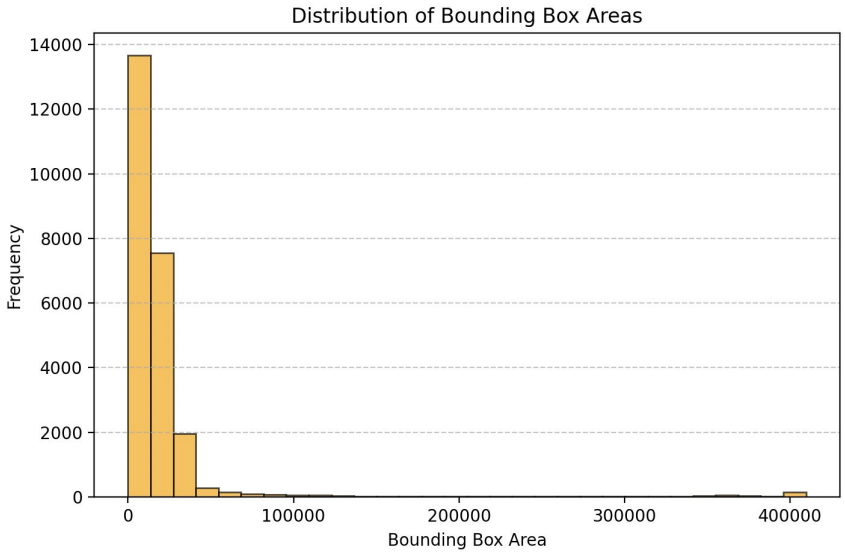
- Aimed at enabling automated license plate detection using object detection techniques
- 9811 images
- 10219 annotations
- 70% Training, 20% Validation, 10% Testing
- LICENSE CC BY 4.0

Source: Roboflow
https://universe.roboflow.com/test-vaxvp/license-plate-project-adaad

Non-Electronics Detection Dataset: License Plate Recognition



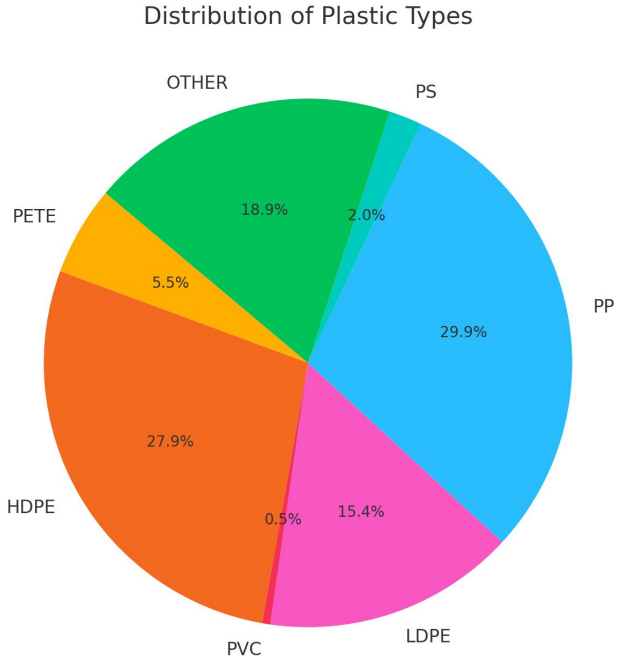
Annotation Heatmap



Recycle label dataset

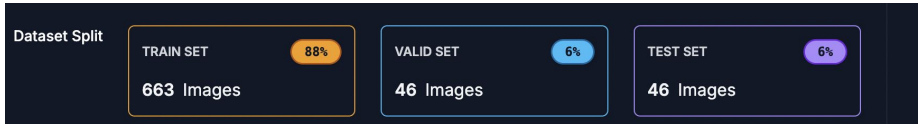
This dataset is instrumental for training models aimed at accurately identifying ROI and classifying various plastic types. It can enhancing recycling processes and contributing to effective waste management strategies. The dataset consists of the following classes of plastics:

1. **PETE** (Polyethylene Terephthalate)
2. **HDPE** (High-Density Polyethylene)
3. **PVC** (Polyvinyl Chloride)
4. **LDPE** (Low-Density Polyethylene)
5. **PP** (Polypropylene)
6. **PS** (Polystyrene)
7. **OTHER**



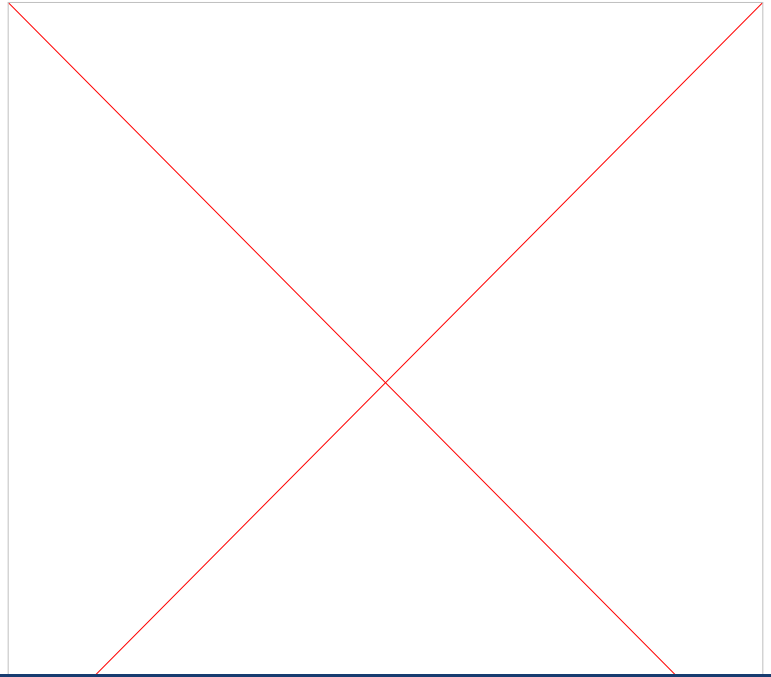
License:

Nutrition Content Label Dataset



This dataset is designed for **object detection** of nutrition tables using **YOLOv8**. It consists of **313 images** containing nutrition labels from various food packages. The dataset is useful for training computer vision models to **detect and extract nutritional information** from product packaging.

https://universe.roboflow.com/lizaza/nutrition-table/dataset/2



Car License Plate Dataset

Car License Plate Detection

576 Code

Data Card | Code (222) | Discussion (2) | Suggestions (0)

https://www.kaggle.com/datasets/andrewmvd/car-plate-detection

Demo: Label Detection + OCR