SCIENCE CHINA
Information Sciences

• NEWS & VIEWS •

May 2025, Vol. 68, Iss. 5, 157201:1–157201:2
https://doi.org/10.1007/s11432-024-4313-x

Open-source dataset for traffic light and countdown display detection in urban environment

Shuchun WAN[1,2], Haitao DING[1*], Jianwei ZHANG[1], Yunfeng HU[1,2] & Hong CHEN[3*]

[1]*State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 130022, China*
[2]*Department of Control Science and Engineering, Jilin University, Changchun 130022, China*
[3]*College of Electronic and Information Engineering, Tongji University, Shanghai 201804, China*

Received 27 May 2024/Revised 18 October 2024/Accepted 8 January 2025/Published online 3 April 2025

Citation Wan S C, Ding H T, Zhang J W, et al. Open-source dataset for traffic light and countdown display detection in urban environment. Sci China Inf Sci, 2025, 68(5): 157201, https://doi.org/10.1007/s11432-024-4313-x

Here we introduce an open-source dataset for traffic light and countdown display detection, which includes three subsets: a subset of traffic light data, a subset of traffic light and countdown display data, and a subset of non-motor vehicle and crosswalk signals data for academic and industrial research.

Traffic light and countdown display detection is very important for autonomous vehicle driving in urban environments. With significant breakthroughs in deep learning technology in the past decade, deep learning algorithms have shown better performance than traditional methods in object detection. The dataset is an important foundation for training, validating, and testing deep learning algorithms, and is significant for algorithm implementation. But to the best of our knowledge, there is no open-source standard dataset for traffic light and countdown display detection in China. The demo data of Apollo only contains 200 images [1]. The LaRA dataset was collected from streets in Paris, France, and includes 9168 images with a resolution of 640 × 480, the traffic light state is tagged as go, warning, stop, or ambuiguous[1)]. The Bosch Small Traffic Light Dataset contains 13427 images mainly collected in San Francisco and Palo Alto, California, USA, with a resolution of 1280 × 720 pixels, consisting of 24242 hand-labeled annotations [2]. The LISA dataset contains 43007 images collected in San Diego, USA, with a resolution of 1280 × 960 pixels, including samples during day and night time [3]. The open-source traffic light datasets mentioned above are excellent, but they cannot adapt well to domestic traffic scenes.

Firstly, we used smartphones to record the urban environment in Changchun, Wuhan at the resolution of 1920 × 1080 pixels from the summer of 2019 to the spring of 2025, the collection period was concentrated from 7:00 am to 24:00 pm, and the weather was covered with sunny, cloudy, rainy, snowy, foggy days. Then we extracted non-continuous frames with irregular time intervals and annotated the status of the traffic light and countdown display, as well as the number inside, according to GB14886-2016 Specification for road traffic signal setting and installation and GA/T 508-2014 Countdown display unit for traffic light to form the dataset. All three subsets were divided into three parts in an 8:1:1 ratio, respectively, for training, validation, and testing.

The traffic light subset contains 20295 images. The status of the traffic light is divided into red, amber, and green for motor vehicle signals, left turn, straight ahead, and right turn of direction signals, turn round traffic signals, and lamp off, a total of 16 statuses. We provide two annotation files: one named labels1 mainly contains the annotations of traffic lights at the first and second intersections, and the other named labels2 adds long-distance small traffic light annotations at the third, fourth, and fifth intersections based on the former with a resolution of 2 × 3 pixels for traffic lights at extreme distances. Based on label1, label2 mainly indicates the addition of 1879 motor vehicle signals red, 201 motor vehicle signals amber, 3459 motor vehicle signals green, 345 direction signals left turn red, 116 direction signals left turn green, and 687 light off states. The traffic light and countdown display subset contains 1000 images. The status of the numbers in the countdown display is divided into red 0–9, amber 0–3, green 0–9, and light off. The non-motor vehicle and crosswalk signals subset contains 479 images. The status of traffic lights is divided into red, amber, and green for non-motor vehicle signals, and the status of crosswalk signals is divided into red, green, and lamp off. We especially annotated the back view, left view, and right view of the traffic light and red, amber, and green in the left and right view. The label names, status, and quantity of traffic lights and countdown displays are detailed in Table 1.

Compared with previous datasets, this open-source dataset not only includes various weather conditions throughout the four seasons, but also includes non-motor vehicle signals, crosswalk signals and their multi-view, countdown displays, and remote small signal light annotations for intersections 2–5.

We have evaluated three subsets using the YOLOv8 al-

* Corresponding author (email: dinght@jlu.edu.cn, chenhong2019@tongji.edu.cn)

1) LaRA French traffic lights recognition public benchmarks. https://www.lara.prd.fr/benchmarks/trafficlightsrecognition.

© Science China Press 2025

Table 1 Dataset description. Subset A: traffic light; subset B: traffic light and countdown display; subset C: non-motor vehicle and crosswalk signals.

Subset	Label	Status	Amount	Subset	Label	Status	Amount
A	tlr	Motor vehicle signals red	12055	B	125	Countdown display 5 red	59
A	tly	Motor vehicle signals amber	1100	B	126	Countdown display 6 red	41
A	tlg	Motor vehicle signals green	18245	B	127	Countdown display 7 red	35
A	tlra	Direction signals left turn red	3422	B	128	Countdown display 8 red	38
A	tlya	Direction signals left turn amber	239	B	129	Countdown display 9 red	40
A	tlga	Direction signals left turn green	1709	B	130	Countdown display 0 red	389
A	tsra	Direction signals straight ahead red	58	B	131	Countdown display 1 green	26
A	tsya	Direction signals straight ahead amber	62	B	132	Countdown display 2 green	38
A	tsga	Direction signals straight ahead green	135	B	133	Countdown display 3 green	25
A	trra	Direction signals right turn red	156	B	134	Countdown display 4 green	45
A	trya	Direction signals right turn amber	22	B	135	Countdown display 5 green	37
A	trga	Direction signals right turn green	57	B	136	Countdown display 6 green	53
A	tura	Turn round traffic signals red	131	B	137	Countdown display 7 green	56
A	tuya	Turn round traffic signals amber	8	B	138	Countdown display 8 green	73
A	tuga	Turn round traffic signals green	125	B	139	Countdown display 9 green	60
A	tln	Lamp off	847	B	140	Countdown display 0 green	416
B	tlr	Motor vehicle signals red	746	B	141	Countdown display light off	1159
B	tly	Motor vehicle signals amber	84	C	tnr	Non-motor vehicle signals red	39
B	tlg	Motor vehicle signals green	290	C	tny	Non-motor vehicle signals amber	31
B	tlra	Direction signals left turn red	497	C	tng	Non-motor vehicle signals green	77
B	tlya	Direction signals left turn amber	159	C	tcr	Crosswalk signals red	209
B	tlga	Direction signals left turn green	188	C	tcg	Crosswalk signals amber	323
B	tsra	Direction signals straight ahead red	325	C	tcn	Crosswalk signals green	147
B	tsya	Direction signals straight ahead amber	109	C	tllv	Traffic light left view	115
B	tsga	Direction signals straight ahead green	89	C	tllvr	Traffic light left view red	23
B	trra	Direction signals right turn red	35	C	tllvy	Traffic light left view amber	48
B	trya	Direction signals right turn amber	233	C	tllvg	Traffic light left view green	12
B	tura	Turn round traffic signals red	91	C	tlrv	Traffic light right view	62
B	tuya	Turn round traffic signals amber	39	C	tlrvr	Traffic light right view red	16
B	tuga	Turn round traffic signals green	34	C	tlrvy	Traffic light right view amber	31
B	tln	Lamp off	188	C	tlrvg	Traffic light right view green	31
B	117	Countdown display 0 amber	51	C	tlbv	Traffic light back view	62
B	118	Countdown display 1 amber	96	C	tclv	Crosswalk signals left view	172
B	119	Countdown display 2 amber	92	C	tclvr	Crosswalk signals left view red	50
B	120	Countdown display 3 amber	66	C	tclvg	Crosswalk signals left view green	58
B	121	Countdown display 1 red	50	C	tcrv	Crosswalk signals right view	290
B	122	Countdown display 2 red	49	C	tcrvr	Crosswalk signals right view red	28
B	123	Countdown display 3 red	40	C	tcrvg	Crosswalk signals right view green	35
B	124	Countdown display 4 red	39	C	tcbv	Crosswalk signals back view	11

gorithm[2)]. For traffic light subset labels1, after 283 training epochs, the result is precision 0.922, recall 0.875, and map50 0.927. For traffic light subset labels2, after 126 training epochs, the result is precision 0.953, recall 0.812, map50 0.887. For the traffic light and countdown display subset, after 395 training epochs, the result is precision 0.734, recall 0.404, and map50 0.457. For the non-motor vehicle and crosswalk signals subset, after 357 training epochs, the result is precision 0.885, recall 0.864, and map50 0.913.

Usage. The open-source dataset is used for real-time traffic light and countdown display detection and status recognition. One may first build up an environment with PyTorch 1.13.0, Torchvision 0.14.0, CUDA 11.6.0, cuDNN 8.5.0, and Python 3.8.

Access methods. The dataset can be downloaded from https://pan.baidu.com/s/1jFGDr5eeWlVmJlYq55Xkpg?pwd=k2qp.

2) https://github.com/ultralytics/ultralytics.

References

1 Huang X Y, Cheng X J, Geng Q C, et al. The ApolloScape open dataset for autonomous driving. 2018. ArXiv:1803.06184

2 Behrendt K, Novak L, Botros R. A deep learning approach to traffic lights: detection, tracking, and classification. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2017. 1370–1377

3 Jensen M B, Philipsen M P, Mogelmose A, et al. Vision for looking at traffic lights: issues, survey, and perspectives. IEEE Trans Intell Transp Syst, 2016, 17: 1800–1815