An Indian Roads Dataset for Supported and Suspended Traffic Lights Detection

Sarita Gautam1(*),
Research Scholar, Panjab University, Chandigarh, India.
Email: Gautamsarry@gmail.com

Anuj Kumar2,
Assistant Professor, Panjab University, Chandigarh, India.
Email: Anuj_gupta@gmail.com

Abstract: Autonomous vehicles are growing rapidly, in well-developed nations like America, Europe, and China. Tech giants like Google, Tesla, Audi, BMW, and Mercedes are building highly efficient self-driving vehicles. However, the technology is still not mainstream for developing nations like India, Thailand, Africa, etc., In this paper, we present a thorough comparison of the existing datasets based on well-developed nations as well as Indian roads. We then developed a new dataset “Indian Roads Dataset” (IRD) having more than 8000 annotations extracted from 3000+ images shot using a 64 (megapixel) camera. All the annotations are manually labelled adhering to the strict rules of annotations. Real-time video sequences have been captured from two different cities in India namely New Delhi and Chandigarh during the day and night-light conditions. Our dataset exceeds previous Indian traffic light datasets in size, annotations, and variance. We prove the amelioration of our dataset by providing an extensive comparison with existing Indian datasets. Various dataset criteria like size, capturing device, a number of cities, and variations of traffic light orientations are considered. The dataset can be downloaded from here https://sites.google.com/view/ird-dataset/home

Keywords: Indian roads dataset, Self-driving, computer-vision, Developing- nations

1. Introduction:

Vision-based systems are trending these days. They are gaining popularity for self-driving cars and Driver Assistant systems. In well-developed nations, self-driving cars are becoming popular due to their safety, efficiency, and convenience. On the contrary in a highly-populated country like India, where technical problems (non-working traffic signals, missing traffic signals) and non-technical problems (highly-congested roads, multiple intersections, trees, obstructions) still prevail on the roads, self-driving technology has a long way to go. It’s challenging for self-driving technology to run efficiently on highly congested and non-maintained roads of underdeveloped nations. Talking about the challenges Traffic Light perception and detection is one of the major problems in accident-proof self-driving in India because the traffic lights are not well maintained and there might also be non-working traffic lights. Additionally, the challenges like recognition of moving objects such as cars, pedestrians, cyclists, motor-cyclist, and static objects like traffic lights and poles also hinder the efficiency of autonomous vehicles. These kinds of problems are being well addressed by researchers for well-developed nations. However, very few researchers have worked on Indian road data and a negligible amount of Indian data is available. Here we are proposing a new dataset “Indian Roads Dataset” (IRD) based on Indian roads captured from the roads of New Delhi and Chandigarh in day & night conditions. This dataset contains images from two different cities and presents traffic lights in two different orientations namely supported and suspended. Section II represents the state-of-art datasets based on well-developed and under-developed nations, section III shows the limitations and drawbacks of the existing datasets, section IV provides a detailed description of our IRD dataset, and section V represents the annotation tools that we have used for labelling and annotations. Section VI presents a technical evaluation of IRD dataset.

2. Review of Existing Datasets based on well-developed nations:

In Table-1 we present an overview of existing traffic light datasets based on well-developed nations. Table-1 clearly depicts that the datasets based on well-devolved nations are shot using high-definition cameras and have large number of high-quality images in them. The authors have classified various objects on the road by labelling different classes like car, pedestrians, pole, traffic lights, motorcycle, bicycle, truck, traffic cone, road pile, fence,
traffic light, pole, traffic sign, building, bridge, tunnel, overpass and many more. A Tabular comparison of datasets based on well-developed nations is given in Table-1 and Table-2.

| Reference | Dataset name | Camera | Size      | Location  | Resolution | Conditions                | Classes                              | Tested on |
|-----------|--------------|--------|-----------|-----------|------------|---------------------------|--------------------------------------|----------|
| [1]       | CDTD         | Sekonix AR0231 | 155,529 | Tokyo, Japan | 1280 X720 | Sunny, night, cloudy, rainy | Car, person, traffic-sign            | NA       |
| [2]       | Kitti vision Benchmark suit | Stereo camera | 12,919   | Karlsruhe | 1280 X720 | day                       | Car, Pedestrian                      | NA       |
| [3]       | Cityscapes   | Stereo camera | 5000 fine annotations, 20000 coarse annotations | Various cities | 1024 X 2049 | Day light | Car, truck, bus, bridge, tunnel, Pole, traffic light, traffic sign, motorcycle, bicycle etc. | NA       |
| [4]       | SIM10k       | GTA V engine | 10000 images of cars | NA | 1280 X720 | Clear, Overcast, Rainy, Night, Light snow | car | NA |
| [5]       | nuScenes     | 8 cameras have been used to capture images. | 93,000 2d annotated images | Boston, Pittsburgh, Las Vegas and Singapore | 1600x1200 | Clear, Overcast, Rainy, Night, Light snow, heavy snow | Car, Person, animal, stroller, wheelchair, air barriers, debris, motorcycle, trailer, truck, police, ambulance, bendy, construction_worker etc. | NVIDIA TITAN GPU |
| [6]       | BDD100k      | NA | 120,000,000 | New York, Berkeley, San Francisco | 1280 X 720 | Clear, rainy, snowy, overcast, cloudy, foggy, average | Person, car, sign, light, truck, bus, bike, rider, motor, train | NA       |
Table 1: Extensive overview of Datasets based on Well-developed nations

| Dataset                  | Camera Type                        | FPS | Depth Resolution | Feature                  | Annotation               | Graphics       |
|--------------------------|------------------------------------|-----|------------------|--------------------------|--------------------------|-----------------|
| Apollo Scape             | Stereo camera, HD cameras          | 10  | 1382 x 512       | Instance-level, pixel-level, Depth images | Vehicles, pedestrians, riders, car, truck, traffic cone, road pile, fence, traffic light, pole, traffic sign, building, bridge, tunnel, overpass | NA             |
| SJTU small traffic light dataset | NA                                  | 10  | 1080 X 1920      | Flickering/fluctuating traffic lights, illumination/exposure changes, multiple visible lights, false positives | Red, yellow, green, off, wait | NA             |
| DriveU traffic light dataset | Stereo camera                      | 232,039 | 640 X 480      | Daylight                 | Circle, arrow straight, arrow left, arrow straight left, arrow right, pedestrian, cyclist | NA             |
| YBY                      | AVT Make G-125c camera with 12-36mm Lens | 30,886 | 1292 X 964      | Daytime, Nightime        | Go, stop, go, arrow, left go, right go, left stop | 4 NVIDIA Jetson TX1 and two FPGs |
| DriveU traffic light dataset | NA                                  | 71,770 | 1280 X 720       | Day, night, city, residential | Green, Red, Yellow, green(left-turn), red(left-turn), yellow(left-turn) | 3 NVIDIA Titan X GPUs |

All the existing datasets based on well-developed nations are mentioned in section-II. To provide a comparative analysis, we compared the characteristic value of Kitti, Nuscenes, BDD100k, Apollo Scape, and DriveU datasets. Their infographics are given in Figures 1-6.

1. **Kitti**: Kitti vision benchmark dataset provides real road images for autonomous driving. The dataset contains more than 12000 images having 1382 x 512 resolution shot on 10FPS and divided in 11 different classes namely buildings, sky, tree, cars, sign, road, fence, pole, sidewalks, pedestrians, bicyclist etc. The infographics of Kitti dataset is provided in Figure 1.

2. **NuScenes**: A large-scale dataset for autonomous driving captured in Boston and Singapore using 6 cameras, 32-beam LiDAR and radars with 360-degree view cameras having about 10 classes such as cars, trucks, buses, trailers, pedestrians, construction vehicles, motorcycles, bicycles, barriers, and traffic cones. This dataset contains 1,400,000 images leading it to be one of the largest datasets publicly available. The infographics of NuScenes dataset is provided in the Figure 2.

3. **BDD100k**: This is a diverse dataset for heterogenous learning for multiple tasks. BDD100k was shot in multiple cities such as Berkeley, New York, San Francisco and Bay Area. The dataset contains more than 100,000 images and have classes like bus, light, sign, person, bike, truck, motor, car, train and rider. The infographics of BDD100k dataset is provided in the Figure 3.

4. **Apollo Scape**: An autonomous driving dataset shot in multiple cities and sites having high variance. The dataset contains more than 100,000 images captured from 10 different cities. The infographic for Apollo Scape has been shown in Figure 4.

5. **DriveU**: This dataset has been generated to be one of the largest and high variance datasets, this dataset contains more than 23000 annotations divided into 344 classes and captured from 11 different cities. The infographic for DTLD has been shown in Figure 5.
| Dataset     | Resolution | Frame Rate | Annotations/Images | Cities | Classes |
|-------------|------------|------------|--------------------|--------|---------|
| Kitti       | 1382 X 512 | 10 Fps     | 12919              | 1      | 11      |
| nuScenes    | 1600x1200  | 20 Fps     | 1,400,000          | 2      | 10      |
| BDD100k     | 1280 x 720 | 30 Fps     | 100k               | 4      | 10      |
| Apollo Scape| 3384 x 2710| 30 Fps     | 140,000            | 10     | 30      |
| DriveU      | 2048 x 1024| 15 Fps     | 232,039            | 11     | 344     |

Table 2: Characteristics comparison of different traffic light datasets.

Figure (1-5): Frequency Comparison between five datasets based on well-developed nations. Each of the infographic here presents number of annotations/images, number of cities the dataset have been shot and the number of classes in every dataset.

Figure 6: shows all the datasets based on well-developed nations in a comparable form.


## 2. Overview of Existing Datasets based on Indian roads:

Table-3 presents an overview of existing traffic light datasets based on Indian roads. After analysing Table 1 and Table-3, it can clearly be defined that the datasets based on Indian roads are not that mature. These datasets are captured using poor camera devices and have limited number of images in them. On top of that none of the datasets are made publicly available. Here we are providing our very own dataset captured from two Indian cities and making it publicly available to use. Further details are mentioned in section 5.

| Reference | Dataset name | Camera | Size/Annotations | Location | Resolution | Conditions | Classes | Tested on |
|-----------|--------------|--------|------------------|----------|------------|------------|---------|-----------|
| Ours      | IRD dataset  | Redmi note-9 pro max, having 64 megapixel camera. 3,334 (images + thumbnails) /8000+ annotations | New Delhi & Chandigarh, India | 3840 X 2160 | Daylight, night | Car, traffic-light, bus, person, Auto-rickshaw,bicycle, truck, motobike, green traffic light, red-light, timer, yellow light, green-light | NVIDIA GEFORCE GTX 1650 |
| [12]      | Private dataset | 16 megapixel | Pune, Maharastra | 4616 X 3464 for images and 1920 X 1080 | daylight | Straight, right, left, yellow,green | NVIDIA GEFORCE 940M GPU |
| [13]      | Private dataset | Normal video camera, Live feed videos | Indian roads | NA | daylight | Car, motorcycle | 6th Generation i5 clocked at 3.0GHz with Catalina MacOS and 8 Gigabytes RAM |
| [14]      | Private dataset | Normal video camera | Chandigarh, from PGI to Madhya Marg | NA | Day, night | Car, motocycle | Webster’s method |
| [15]      | Used Tesla’s autopilot features, BMW’s LiDar | Forward looking side camera, rearward looking side camera, rear view cameras, radar sensor’s data, | NA | NA | NA | NA | Arduino uno, IR sensor |
| [16]      | NA | Video footage of 50 frames per second | Tamil nadu | 336x 596 | daylight | Cars | Matlab , Arduino circuit |
| [17]      | Private dataset | Video frames and still images | Bengluru | NA | daylight | Lane, potholes, road signs images | NA |
| [18]      | Private dataset | Still images of road sign, roads with potholes, road images with lane markings | Indian roads | NA | daylight | Lane, potholes, road signs images, human shapes, bounding boxes | MatLab |
| [19]      | Used publicly available BSTLD | Images are acquired from a publicly available dataset | Not on Indian roads | 1280x720 | Busy streets, overcast, light rain, flickering light, illuminations, Traffic light | NA |
Table 3: Extensive overview of Datasets based on Indian roads

All the existing datasets based on Indian roads are mentioned in section 3. To provide a comparative analysis, we compared the characteristic value of our dataset with the other dataset. As analysed from table 3 not much information is provided about the Indian datasets. So, we present an infographic of these datasets in comparison to our IRD dataset. The infographic of Indian roads datasets is given in Figure 7. Since the Indian datasets are private so they do not have any name, here we are presenting these datasets by the author’s name.

| Dataset       | Resolution          | Frame Rate | Annotations/Image | Cities | Classes |
|---------------|---------------------|------------|-------------------|--------|---------|
| IRD           | 3840 X 2160         | 3FPS       | 8000+             | 2      | 14      |
| Kulkarni’s    | 4616 X 3464 for images and 1920 X 1080 | 20FPS      | 1237              | 1      | 5       |
| Sharma’s      | Not provided        | NA         | NA                | 1      | 2       |
| Sharma’s      | Not provided        | NA         | NA                | 1      | 2       |
| Ritwik’s      | Not provided        | NA         | NA                | 1      | 1       |
| Kumar’s       | Not provided        | NA         | NA                | 1      | 3       |
| Danti’s       | Not provided        | NA         | NA                | 1      | 5       |
| Swati’s       | Not provided        | NA         | NA                | 1      | 2       |

Table 4: Characteristics comparison of different Indian traffic light datasets.

Figure 7: Infographic of dataset based on Indian roads.

4. Limitations in Existing datasets: As analysed from Table-1 and Table-2, we can see that the dataset shot in well-developed nations/cities has a large number of images in them, captured using high-definition cameras or multiple cameras. Their capturing vehicles are well equipped with high-definition cameras, LiDar, radar sensors, etc. These datasets are highly useable as compared to Indian datasets which have been shot using minimal quality devices. The Indian datasets are not well organized and also a limited number of objects have been considered in them. So here we propose a new dataset having 3000+ images (combined with thumbnails) captured from the roads of New Delhi and Chandigarh shot in a day and night-light conditions using a 64 mega-pixel camera having a total of 8,520 annotations of different objects. Our dataset provides a better insight into Indian roads based on two different cities namely New Delhi and Chandigarh. We have considered 13 classes for different objects in our
dataset as mentioned Car, traffic light, bus, person, Auto-rickshaw, bicycle, truck, motorbike, green traffic light, red light, timer, yellow light, and green-light.

5. **Our Contribution (IRD):** We collected video sequences from the roads of New Delhi and Chandigarh cities of India. We used wagonR car and mounted a camera inside the car using a camera-holder. We used a mobile phone with a camera of 64 Megapixels for capturing the video sequences from two cities New Delhi and Chandigarh. In the city, Chandigarh video sequences from Housing board chowk to Panjab University were captured. It’s a 10 k.m. long route as shown in map Figure-8. For the New-Delhi video sequence from Govind-puri to Gurgaon, it’s a 10 k.m were captured. long route as shown in map Image-2. After capturing the videos, we extracted frames from them using the VLC media player with a frame of 20 fps. We have a total of 3050 images combined from both New Delhi and Chandigarh roads.

6. **Video data:** In the process of creating our dataset, we first captured video sequences from the roads of New Delhi and Chandigarh. The main purpose of creating this dataset was to create a hybrid dataset having traffic lights from different cities having different orientations. Like in Chandigarh we have supported traffic lights and in New Delhi suspended traffic lights are there. After capturing video sequences, we extracted frames from the videos using the VLC media player. We followed the following steps for image extraction:

1. Open the VLC player and pause the video.
2. Then under the tools tab, we clicked on preferences.
3. After that we landed on the Interface tab, where we need to select “All” preferences under the “show settings” bar.
4. This will lead us to “Advanced Preferences”, here under the video submenu we clicked on “filter” and select “scene filter” from the drop-down menu.
5. A new small window pane will open on the right side of the same dialogue box.
6. Here we can specify the image format, Height, width, filename prefix directory path, and the recording ratio.

7. The Recording ratio denotes the ratio of images to be recorded. Here recording ratio of 10 means one image out of ten images are stored. We kept recording ratio as 3.

8. Now after changing all the settings, we played the video. As soon as the video is done playing, go to the specified directory and here you can see images extracted from the video. Steps for frame extraction are given in the images from 1 to 4.

Figure 9: steps for image extraction from video sequences.

7. Real-time images of the dataset in dayLight: Images extracted from video frames in dayLight
8. **Real-time images of the dataset in nightLight:** Images extracted from video frames in nightLight

9. **Image Annotation:**

For the purpose of labelling and annotation we used a freely available image annotation tool “Super Annotate”. In the labelling process, we first defined 14 classes with different colour codes namely car, traffic light, green traffic light, red-light, green-arrow-light, pedestrian green light, pedestrian-red-light, yellow light, timer, person, auto-rickshaw, bicycle, motor-bike, truck, bus, board as shown in Figure 3. We created squared shape bounding boxes and some free-hand boundaries for labelling the objects in the images. Our dataset contains a total of 3050
images with 3840 X 2160 resolution. From figure 5 to 7 various windows and dialogue boxes of Super-annotate is shown.

**Figure 12:** Super Annotate tool: for labelling images.

**Figure 13:** Classes with different colour codes for labelling images.
10. **Evaluation:** Transfer learning is a process of reusing pre-trained knowledge and applying it to process similar kinds of new data. During the evaluation process, we tried to detect traffic lights and some other objects like people, cars, and other vehicles. We used the single shot detector model for object detection and trained our Inception_v3 model on it. We created a traffic light extraction model to create cropped images of traffic lights and then detected traffic lights and other objects from the cropped images as shown in Figure 12. Here are some of the cropped images as shown in Table 5. We evaluated our dataset on an NVIDIA GeForce GTX 1650 ti having GB RAM. Our system achieved 97.23% test accuracy. The line graph is shown in Figure 13. The evaluation results are shown in Table 5 along with an infographics of extracted features in

![Figure 14: The labelled objects and bounding boxes](image)
Figure 15: Cropped images of traffic lights using Single Shot Detector

Figure 16: Line graph of Detection accuracy using Inception_v3 model.

| S.no | Name of the folder | Number of images in each folder | Features extracted (Traffic light and similar looking objects) |
|------|---------------------|----------------------------------|---------------------------------------------------------------|
| 1.   | One                 | 252                              | 1496                                                          |
| 2.   | Two                 | 210                              | 1331                                                          |
| 3.   | Three               | 392                              | 2187                                                          |
| 4.   | Four                | 318                              | 1053                                                          |
| 5.   | Five                | 336                              | 1343                                                          |
| 6.   | Six                 | 1139                             | 780                                                           |
| 7.   | Seven               | 353                              | 1256                                                          |
| 8.   | Eight               | 350                              | 655                                                           |
| 9.   | Nine                | 350                              | 2507                                                          |
| 10.  | Ten                 | 274                              | 160                                                           |

Table 5: Shows number of features extracted from each folder.
12. Conclusion and Future scope: A manually annotated labelled dataset IRD is presented in this paper which is based on two different cities having two different types of traffic light orientations. This dataset can be used to detect traffic lights and other objects like persons, cars, other vehicles, etc. This dataset provides insight into Indian roads. Which can help motor industrialists to gain better insight into Indian roads and manufacture autonomous cars accordingly. Here we applied transfer learning and Single shot detector methods to identify traffic lights in an image. The limitation of our dataset lies in the fact that there is a limited number of images in the night light and in the future we will extend our dataset to night light, foggy and rainy weather and automatically remove blurry images captured due to camera motion. We will try to test our dataset in real-time using a real vehicle.

References:
[1] http://zhiqiangshen.com/projects/INIT/index.html
[2] http://www.cvlibs.net/datasets/kitti/
[3] https://www.cityscapes-dataset.com/
[4] https://paperswithcode.com/dataset/sim10k
[5] https://www.nuscenes.org/
[6] https://www.bdd100k.com/
[7] http://apolloscape.auto/
[8] https://github.com/Thinklab-SJTU/S2TLD
[9] https://www.uni-ulm.de/in/lui-drive-u/projekte/driveu-traffic-light-dataset/
[10] Ouyang, Z., Niu, J., Liu, Y., & Guizani, M. (2019). Deep CNN-based real-time traffic light detector for self-driving vehicles. IEEE transactions on Mobile Computing, 19(2), 300-313.
[11] Kim, J., Cho, H., Hwangbo, M., Choi, J., Canny, J., & Kwon, Y. P. (2018, November). Deep traffic light detection for self-driving cars from a large-scale dataset. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 280-285). IEEE.

[12] Kulkarni, R., Dhavalikar, S., & Bangar, S. (2018, August). Traffic light detection and recognition for self driving cars using deep learning. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (pp. 1-4). IEEE.

[13] Sharma, M., Bansal, A., Kashyap, V., Goyal, P., & Sheikh, T. H. (2021). Intelligent Traffic Light Control System Based On Traffic Environment Using Deep Learning. In IOP Conference Series: Materials Science and Engineering (Vol. 1022, No. 1, p. 012122). IOP Publishing.

[14] Sharma, I., & Pardeep, K. G. (2015). Study of automatic traffic signal system for chandigarh.

[15] Rao, S. S., & Desai, S. R. (2021). Machine Learning based Traffic Light Detection and IR Sensor based Proximity Sensing for Autonomous Cars. Available at SSRN 3883931.

[16] Ritwik, M. G., Krishna, D., Shreyas, T. R., & Phamila, Y. A. V. (2018, September). Road Congestion based Traffic Management System with Dynamic Time Quantum. In 2018 International Conference on Recent Trends in Advance Computing (ICRTAC) (pp. 1-6). IEEE.

[17] Kumar, T. N. R. (2015). A real time approach for indian road analysis using image processing and computer vision. IOSR J. Comput. Eng.(IOSR-JCE), 17(4), 2278-0661.

[18] Danti, A., Kulkarni, J. Y., & Hiremath, P. S. (2012). An image processing approach to detect lanes, pot holes and recognize road signs in Indian roads. International Journal of Modeling and Optimization, 2(6), 658.

[19] Mahesh .G, Satish Kumar .T, (2019, August). Real Time Traffic Light Detection by Autonomous Vehicles using Artificial Neural Network Techniques (UIITEE) Volume-8 Issue-10,( 2278-3075)

[20] Miss. Swati N. Zavar, Mrs. Yogita H. Sonawane, Mrs. Ashwini P. Bawdhankar, 2020, Image Analysis Driven Intelligent Traffic Light Control for Smart Cities, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) ICSITS – 2020 (Volume 8 – Issue 05),