Volume 10, Issue 6, June – 2025 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jun1226

A Comparative Study of License Plate Recognition (LPR) Datasets and Benchmarks

Krishna Kumar Sahu[1]; Sudhanshu Shekhar Dadsena[2]; Komal Yadav[3]

[1]Student; [2,3]Assistant Professor

[1;2;3]Department of CSE Shri Rawatpura Sarkar University, Raipur

Publication Date: 2025/06/23

Abstract: License Plate Recognition (LPR) systems are vital components of modern intelligent transportation systems. Their performance heavily depends on the availability of high-quality datasets and reliable benchmarking techniques. This paper provides a comparative analysis of widely used LPR datasets and benchmarks, highlighting their unique characteristics, use cases, and limitations. The study aims to guide researchers in selecting appropriate datasets for training and evaluating LPR models.

Keywords: *License Plate Recognition, Dataset, Benchmark, Intelligent Transportation, OCR, Deep Learning.*

How to Cite: Krishna Kumar Sahu; Sudhanshu Shekhar Dadsena; Komal Yadav (2025) A Comparative Study of License Plate Recognition (LPR) Datasets and Benchmarks. *International Journal of Innovative Science and Research Technology*, 10(6), 1538-1539. https://doi.org/10.38124/ijisrt/25jun1226

I. INTRODUCTION

License Plate Recognition (LPR), also known as Automatic Number Plate Recognition (ANPR), is used for vehicle identification in applications like traffic monitoring, toll collection, and law enforcement. As LPR models transition from traditional to deep learning-based methods, the need for large, diverse, and annotated datasets becomes more significant.

II. OVERVIEW OF LPR SYSTEMS

➢ *An LPR System Typically Involves:*

- Image Acquisition: Capturing vehicle images using surveillance cameras.
- License Plate Detection: Locating the plate in the image.
- Character Segmentation: Splitting the plate into individual characters.
- Optical Character Recognition (OCR): Recognizing the alphanumeric text.

Performance varies based on image quality, lighting, angle, and regional plate formats.

III. PUBLICLY AVAILABLE LPR DATASETS

➢ *UFPR-ALPR (Brazil)*

- Contains over 4,500 images of real-world vehicles.
- Captured in diverse conditions (night/day, occlusions).
- Includes annotations for detection and recognition.

➢ *CCPD (China)*

- Over 250,000 images.
- Annotated with bounding boxes and plate text.
- Captures plates from multiple angles and lighting scenarios.

➢ *AOLP (Taiwan)*

- Designed for application-oriented scenarios.
- Three subsets: Access Control (AC), Traffic Law Enforcement (LE), Road Patrol (RP).
- Useful for evaluating LPR in varying environments.

➢ *Synth-LPR (Synthetic)*

- Artificially generated dataset using computer graphics.
- Useful for training models where real data is limited.
- Allows variation in font, size, lighting, and angle.

➢ *Caltech Cars (USA)*

- Primarily for vehicle detection.
- Can be adapted for LPR with manual annotation.

IV. BENCHMARKING AND EVALUATION METRICS

➢ *To Evaluate LPR Systems, Standard Metrics are used:*

- Detection Accuracy: Correct identification of plate regions.
- Recognition Accuracy: OCR performance on detected plates.

- Precision/Recall/F1-Score: To evaluate character-level or full plate recognition.
- Frame Per Second (FPS): For real-time applications.

V. COMPARATIVE ANALYSIS

Table 1 Comparative Analysis

Dataset	Region	Size	Annotations	Conditions Covered
UFPR-ALPR	Brazil	~4,500	Yes	Day/Night, Occlusion
CCPD	China	~250,000	Yes	Multi-angle, Lighting
AOLP	Taiwan	~2,000	Yes	Controlled Scenarios
Synth-LPR	NA	Custom	Yes	Fully Synthetic
Caltech	USA	~1,000	No (manual)	Side Views, Daylight

VI. CHALLENGES AND FUTURE DIRECTIONS

- Region-specific Variations: Plate formats vary by country.
- Data Privacy: Use of real-world images must respect legal and ethical standards.
- Synthetic vs. Real Data: Synthetic data helps but may not generalize well.
- Need for Multilingual OCR: Especially for countries with non-Latin scripts.

VII. CONCLUSION

This paper summarizes key LPR datasets and benchmarks essential for developing and evaluating modern LPR systems. The comparative insights can help researchers choose the most suitable dataset based on region, application, and model requirements. The field continues to evolve with the integration of synthetic data, deep learning, and multilingual capabilities.

REFERENCES

[1]. R. S. Laroca et al., "A Robust Real-Time Automatic License Plate Recognition Based on the YOLODetector," 2019.
[2]. X. Xu et al., "Towards End-to-End Car License Plates Detection and Recognition with Deep NeuralNetworks," 2018.
[3]. S.-L. Chen et al., "Application-Oriented License Plate Recognition," 2012.
[4]. T. Younes et al., "Synthetic Data for License Plate Recognition," arXiv, 2019.
[5]. Caltech Cars Dataset, http://www.vision.caltech.edu/Image_Datasets/CaltechCars/