ISSN: 2393-8048
International Advance Journal of Engineering, Science and Management (IAJESM)
Multidisciplinary, Multilingual, Indexed, Double Blind, Open Access, Peer-Reviewed, Refereed-International Journal.
SJIF Impact Factor =8.152, January-June 2025, Submitted in March 2025

A Review on Dataset Challenges and Model Generalization in Traffic Sign Recognition Using Artificial Intelligence

Priya Chaursia, Department of Engineering and Technology (M.Tech.), NIILM University, Kaithal (Haryana)
Dr. Mukesh Kumar Rana, Professor, Department of Engineering and Technology, NIILM University, Kaithal (Haryana)

Abstract

Traffic Sign Recognition (TSR) plays a pivotal role in intelligent transportation systems, driver assistance, and autonomous driving. While Artificial Intelligence (AI) and Deep Learning (DL) models have achieved impressive results in TSR, real-world performance often suffers due to dataset limitations and poor generalization capabilities. This review paper examines the core challenges related to dataset quality, diversity, annotation inconsistencies, and domain adaptability. It also discusses generalization issues of AI models across geographies, weather conditions, lighting variations, occlusions, and multilingual signs. We provide a comparative analysis of existing datasets (e.g., GTSRB, LISA, IDD, ITSD), review key AI models (YOLO, CNNs, MobileNet, etc.), and propose research directions to improve model robustness and generalization across diverse environments.

Keywords: Traffic Sign Recognition, Artificial Intelligence, Deep Learning.

I. Introduction

With the rapid advancement of artificial intelligence (AI) and its integration into the transportation sector, Traffic Sign Recognition (TSR) systems have become a critical component of modern intelligent transportation frameworks, particularly in Advanced Driver Assistance Systems (ADAS) and autonomous vehicles. These systems are designed to detect and interpret traffic signs in real time, enabling vehicles to make safe, informed decisions without human intervention. Accurate TSR can significantly reduce road accidents, support adherence to traffic laws, and provide contextual awareness in driver-assist technologies. According to a report by the World Health Organization (2023), nearly 1.19 million people die in road traffic accidents each year globally, with a substantial proportion attributed to human error and failure to observe road signage [1]. Thus, TSR technology has the potential to save lives and improve road safety when implemented effectively. However, despite remarkable progress in deep learning and computer vision over the past decade, TSR models often underperform in real-world scenarios, especially in developing countries like India. One of the primary reasons is the quality and diversity of training datasets. Most benchmark datasets, such as the German Traffic Sign Recognition Benchmark (GTSRB) and the LISA dataset from the U.S., are based on well-maintained, standardized road systems with consistent signage [2]. These datasets, though useful in structured environments, do not capture the diversity of sign styles, occlusion, damage, lighting conditions, multilingual content, or chaotic traffic environments commonly found in countries with heterogeneous infrastructure. In India, where over 1.53 lakh road accident deaths were recorded in 2022 (Ministry of Road Transport and Highways, Govt. of India), the role of TSR is even more critical [3]. Yet, local datasets such as the Indian Driving Dataset (IDD) and the Indian Traffic Sign Dataset (ITSD), while valuable, remain limited in scale, diversity, and consistent annotation. For example, the IDD includes over 10,000 traffic signs but suffers from class imbalance—certain rare signs have less than 50 labeled instances, impairing the model's ability to learn [4]. Moreover, annotations in regional datasets often lack standardization, with inconsistent bounding box dimensions, incorrect class labels, or missing multilingual text interpretation. This significantly hampers generalization, i.e., the model's ability to perform well on unseen data from different geographic, climatic, or cultural settings.

Another major hurdle is domain adaptation. A model trained on European traffic signs may perform well in its native context but fail drastically when applied to Indian or Southeast Asian roads due to variation in sign shape, font, language, color, and placement. A survey conducted by the Indian Institute of Science (2022) found that TSR models trained on GTSRB achieved over 96% accuracy on the German test set but dropped below 65% accuracy when tested on Indian road images [5], highlighting the critical issue of domain-specific generalization. Additionally, environmental variability such as poor lighting (nighttime driving), weather conditions (fog, rain, dust), and occlusions (trees, poles, vehicles) further degrade TSR performance. Many real-world images contain faded, vandalized, or partially visible signs, which are rarely represented in training datasets. According to a 2021 study by IIT Delhi, nearly 38% of signs in semi-urban areas in North India were either defaced or obscured, leading to misclassification or non-detection by conventional AI models [6]. Finally, multilingual traffic signs pose a unique challenge in countries like India, where signs may contain text in English, Hindi, and the local state language. Most current TSR models do not incorporate Optical Character Recognition (OCR) or Natural Language Processing (NLP) capabilities to interpret such text, limiting their effectiveness in diverse linguistic contexts [7].

Accurate detection and recognition of traffic signs is becoming increasingly important in intelligent transportation systems. Proper interpretation of traffic signs is essential for safety and compliance since they carry vital road information. Strong vision systems that can understand complicated driving surroundings in real time are in high demand due to the rising deployment of semi- and fully-autonomous vehicles. In the past, features including color histograms, edge detectors, and geometric form filters were used extensively for traffic sign detection. Although these methods could recognize some objects, their lack of generalizability and adaptability made them unstable in complex or changing settings [8]. Despite advancements in machine learning techniques like Random Forests and Support Vector Machines, feature engineering limitations continued to hinder their ability to fully enhance classification accuracy. Convolutional Neural Networks (CNNs) and other forms of deep learning have enabled tremendous advancements in image categorization and object detection. Feature extraction by hand is unnecessary when using CNNs because they can learn hierarchical feature representations automatically from raw image pixels. Many fields have begun to use them, including autonomous driving, face recognition, and medical imaging [9]. A CNN-based traffic sign identification system trained on the GTSRB dataset is presented in this research. The implementation goes a step further by utilizing a webcam and OpenCV in real time, proving that such systems may be deployed in practical applications. The suggested model is well-suited for incorporation into driver-assistance modules or autonomous navigation frameworks because of its high speed and minimal latency, which is achieved by utilizing data augmentation techniques, GPU acceleration, and PyTorch [10].

Therefore, this paper aims to provide a comprehensive review of the limitations of existing traffic sign datasets, explore challenges in achieving robust model generalization, and recommend potential strategies—such as synthetic data augmentation, active learning, transfer learning, multilingual OCR integration, and cross-domain validation—to enhance the applicability and accuracy of TSR systems in real-world settings. Through this analysis, we intend to bridge the gap between laboratory success and field deployment, particularly in the context of developing nations.

Pundir and Mahajan (2021)[11] developed the Indian Traffic Sign Dataset (ITSD), specifically curated to reflect the diversity and complexity of Indian roads, including occluded and degraded signs. Their dataset includes over 12,000 labeled signs spanning 45 classes. When tested with CNN models, the dataset exposed serious limitations in class balance, especially with rare categories having fewer than 50 samples. The authors concluded that without addressing such imbalances, generalization would remain poor. Grounded in postcolonial technological theory, they argued for datasets that reflect indigenous conditions rather than replicating Western benchmarks Varma et al. (2019)[12] introduced the Indian Driving Dataset (IDD), designed to showcase unstructured Indian road conditions such as chaotic signage,

ISSN: 2393-8048
International Advance Journal of Engineering, Science and Management (IAJESM)
Multidisciplinary, Multilingual, Indexed, Double Blind, Open Access, Peer-Reviewed, Refereed-International Journal.
SJIF Impact Factor =8.152, January-June 2025, Submitted in March 2025

unpredictable occlusions, and poor lighting. Their findings revealed a performance drop of over 30% in models trained on the German Traffic Sign Recognition Benchmark (GTSRB) when applied to IDD, highlighting the failure of Western-trained AI to adapt to Indian scenarios. The research emphasized situated AI theory, calling for AI systems developed within the environments where they are expected to operate . Meena and Sharma (2021)[13] In their study, they evaluated YOLOv3 and Faster R-CNN for detecting traffic signs under real-world Indian driving conditions. YOLOv3 offered faster inference times but was less accurate for partially occluded signs, while Faster R-CNN showed higher precision but needed more computation. By applying behavioral AI theory, they emphasized that TSR systems must be capable of interpreting incomplete or ambiguous cues—just as human drivers do in complex environments Bansal and Kumar (2021)[14] applied transfer learning using ResNet and MobileNet architectures on Indian traffic sign datasets. While MobileNet yielded better results on low-resource hardware, both models were limited in identifying weathered and misaligned signs. Their work echoed ecological systems theory, advocating that AI must adapt to the local context and environmental variables influencing visibility and interpretability of road signs. Sinha and Kapoor (2022)[15] proposed a hybrid model combining CNN and Support Vector Machines (SVM) to enhance detection of obscured and multilingual signs. While the hybrid model improved detection of low-quality signs, SVM was found sensitive to background clutter. Using multimodal learning theory, they argued that effective TSR must go beyond visual cues by integrating OCR and semantic understanding—particularly in regions with multilingual signage.

Raj and Verma (2020)[16] conducted a quality audit on Indian TSR datasets and identified significant annotation errors such as misaligned bounding boxes and mislabeled signs, affecting nearly 18% of samples. They concluded that annotation errors introduce structural biases, limiting a model's generalization capacity. Drawing from critical data studies, they argued that flawed annotations are not just technical problems but reflections of systemic oversight in data governance. Sharma et al. (2022)[17] developed a TSR system that integrates CNN with OCR to interpret multilingual traffic signs (English, Hindi, Tamil). This multimodal system improved detection of instruction-based signs but underperformed in conditions with motion blur or glare. The study employed linguistic relativity theory to argue that understanding traffic context requires models capable of processing regional languages alongside visual features. Joshi and Rathi (2021)[18] examined how data augmentation techniques—like Gaussian noise, rotation, and mirroring—improved CNN performance on traffic sign datasets. While classification accuracy increased by 12%, the authors observed that synthetic noise could not fully replicate real-world imperfections like dust, rust, or stickers. Based on simulation theory, they noted that artificial augmentations cannot substitute for real field data in training robust models. Upadhyay and Sneha (2022)[19] deployed YOLOv5-Tiny on a Raspberry Pi for real-time TSR and achieved 84% accuracy at minimal latency. However, they reported issues in detecting small and distant signs. Their work was rooted in embodied AI, which posits that intelligent behavior must emerge within the physical and computational constraints of the system. They concluded that AI for smart mobility must balance accuracy with efficiency. Saxena and Bhatnagar (2023)[20] studied domain adaptation methods like fine-tuning, adversarial training, and style transfer to improve the adaptability of models trained on GTSRB to Indian roads. Their domain-adapted model improved recognition accuracy by 18%, especially on occluded and faded signs. Guided by transfer learning theory, they asserted that cross-domain retraining is essential for generalizing across geographic contexts.

II. Importance of Traffic Sign Recognition

Traffic signs convey crucial information to drivers and automated systems. Recognizing signs accurately, regardless of lighting, weather, or visual obstructions, is critical for safety and

ISSN: 2393-8048
International Advance Journal of Engineering, Science and Management (IAJESM)
Multidisciplinary, Multilingual, Indexed, Double Blind, Open Access, Peer-Reviewed, Refereed-International Journal.
SJIF Impact Factor =8.152, January-June 2025, Submitted in March 2025

compliance. AI-based TSR systems must adapt to real-time challenges such as:

Blurred or Faded Signs – A Challenge of Visual Degradation

In real-world driving environments—especially in developing countries like India—traffic signs are often subjected to harsh environmental conditions that lead to significant visual degradation over time. Prolonged exposure to ultraviolet (UV) radiation from sunlight, repeated monsoonal rainfalls, air pollution, and dust accumulation collectively deteriorate the surface quality of traffic signboards. These environmental stressors gradually fade the reflective paint, rust metallic components, and blur the graphical or textual content on the signs. The problem is further intensified on highways and urban roads, where vehicles typically move at high speeds, resulting in motion blur in real-time image captures by dashcams or onboard vehicular cameras. The combined effect of environmental wear and motion-induced distortion poses a serious challenge for traditional Traffic Sign Recognition (TSR) systems that rely heavily on hard-coded features such as fixed color thresholds, shape-based filters, or edge detection algorithms. These classical methods require clean, sharp, and well-defined input to accurately detect and classify signs.

To overcome these limitations, modern AI-based TSR systems have shifted toward deep learning architectures—particularly Convolutional Neural Networks (CNNs), YOLO (You Only Look Once), RetinaNet, and EfficientDet—which are more resilient to image quality distortions. These models are designed to extract abstract, hierarchical features that are less sensitive to surface-level imperfections. Crucially, during model training, datasets are augmented with synthetic visual impairments such as Gaussian blur, motion blur, salt-and-pepper noise, reduced contrast, low brightness, and partial occlusions. This form of data augmentation enhances the model's robustness and its ability to generalize across degraded visual conditions. The Indian Driving Dataset (IDD), for example, is a critical resource in this regard—it includes real-life samples of faded, damaged, and poorly lit traffic signs from Indian roads, helping models learn to detect degraded signs in authentic environments. In addition to data augmentation, several preprocessing and architectural strategies have been introduced to enhance performance under degraded conditions. Image enhancement techniques, such as histogram equalization, deblurring filters, and contrast stretching, are often applied during preprocessing to improve visual clarity before feeding data into the model. Moreover, advanced models incorporate blur-tolerant loss functions that assign greater weight to uncertain or low-confidence predictions, encouraging the network to pay closer attention to difficult samples. These innovations have collectively enabled modern AI-based TSR systems to detect and classify traffic signs with significantly higher accuracy, even when the signs are partially faded, occluded, or blurred. The adoption of these techniques is crucial for achieving reliable sign recognition in real-time, safety-critical applications like Advanced Driver Assistance Systems (ADAS) and autonomous driving, especially in visually challenging settings prevalent across the Indian subcontinent.

Varying Regional Designs and Symbols – A Challenge of Non-Standardization

A significant obstacle in the deployment of Traffic Sign Recognition (TSR) systems in India lies in the non-standardized nature of traffic signage across different states and regions. Unlike countries that follow international standards such as the Vienna Convention on Road Signs or the Manual on Uniform Traffic Control Devices (MUTCD), Indian traffic signage suffers from a lack of centralized regulatory enforcement. As a result, variations in color schemes, font styles, symbol placements, and even sign shapes are commonly observed. In rural and semi-urban areas, signs are often hand-painted or fabricated by local contractors, leading to the use of unofficial pictograms and inconsistent dimensions. Such heterogeneity poses a serious challenge for traditional TSR models, which are designed to detect uniform and well-defined sign patterns. To overcome this challenge, modern AI-based TSR systems leverage transfer learning and domain adaptation techniques. Pretrained models on standardized datasets like the

German Traffic Sign Recognition Benchmark (GTSRB) are fine-tuned using region-specific datasets such as the Indian Driving Dataset (IDD) and the Bihar Traffic Sign Corpus (BTSC). This process enhances the model's adaptability to diverse real-world sign styles. Furthermore, researchers have introduced Generative Adversarial Networks (GANs) to synthetically generate a wide range of traffic sign variants that mimic regional styles, languages, and layouts. These synthetic images enrich training datasets without requiring labor-intensive manual data collection. Such approaches significantly broaden a model's exposure to variable signage patterns, enabling AI-driven TSR systems to perform reliably even when the appearance of traffic signs deviates from conventional norms.

Occlusion from Vehicles or Vegetation – A Challenge of Partial Visibility

Another persistent challenge in real-world TSR deployment, especially in Indian traffic environments, is the issue of occlusion. In congested urban areas and on highways, traffic signs are frequently obstructed by large vehicles such as trucks, buses, or vans that block line-of-sight for both human drivers and vehicle-mounted cameras. In addition, natural obstructions like overgrown tree branches, vines, or bushes—especially during the monsoon season—can partially or fully obscure signage. The presence of visual noise in the form of political banners, commercial advertisements, or unauthorized stickers further exacerbates the issue, making it difficult for traditional TSR systems, which depend on full visibility and distinct contours, to function accurately. To tackle partial visibility, AI-powered TSR systems incorporate temporal and spatial inference techniques. Models based on Mask R-CNN and other instance segmentation frameworks can isolate specific regions of interest, enabling them to recognize and interpret traffic signs even when only a portion is visible. Additionally, temporal convolutional networks and attention-based recurrent neural networks (RNNs) leverage sequential frame analysis to infer missing parts of signs using contextual information from preceding or following video frames. These models maintain a memory of past visual cues, allowing them to reconstruct or predict obscured signs with reasonable confidence. Some advanced systems also utilize context-aware reasoning, where the model factors in environmental cues such as the curvature of the road, proximity to intersections, vehicle GPS data, and typical sign placement zones. This allows the system to logically infer which sign might be present in a partially visible or occluded location. Collectively, these techniques significantly improve the robustness and accuracy of TSR models in complex, cluttered, and dynamically changing environments typical of Indian roads.

Multi-language Content (Especially in India) – A Challenge of Linguistic Diversity

One of the most intricate challenges faced by Traffic Sign Recognition (TSR) systems in India is the high degree of linguistic diversity reflected in roadside signage. Traffic signs across the country often contain text in three or more languages—commonly English, Hindi, and a third regional language such as Tamil, Bengali, Kannada, or Telugu. In many rural and remote regions, signage may omit English or Hindi altogether, presenting information solely in the local script. The co-existence of various writing systems such as Devanagari, Tamil, Malayalam, Gurmukhi, and others introduces significant complexities for traditional Optical Character Recognition (OCR) engines, which are typically optimized for Latin alphabets and struggle with multiple scripts, mixed fonts, and context-based language identification. The situation is further complicated by the condition of the signs themselves—text may be faded, distorted, skewed, partially occluded, or rendered in low contrast, which hampers accurate detection and classification. To address these challenges, modern TSR systems increasingly integrate advanced multilingual OCR solutions, such as Google's Tesseract 4.0—which supports over 100 languages—and more recent deep-learning-based OCR frameworks like PaddleOCR and EasyOCR, which are capable of script detection and robust multilingual interpretation. These systems are often paired with Vision Transformers (ViTs) and CNN-OCR hybrids, enabling simultaneous visual-symbol and text recognition. Some implementations also use language classification layers to automatically detect the script and invoke the correct OCR pipeline. In more advanced applications, neural machine translation (NMT) modules are embedded into the recognition system to translate the sign's message into the user's preferred language in real time—supporting dashboard alerts, voice assistants, or mobile driving apps. This is particularly crucial in autonomous vehicles or Advanced Driver Assistance Systems (ADAS), where accurate and instantaneous understanding of multilingual content can affect navigation, regulatory compliance, and overall road safety.

Real-Time Constraints in Dynamic Environments – A Challenge of Speed and Adaptability

Beyond the challenge of accuracy, one of the most critical technical barriers for TSR deployment is the need for real-time responsiveness in dynamic and fast-moving driving environments. Vehicles on highways often travel at speeds exceeding 100 km/h, leaving mere milliseconds for a TSR system to detect, classify, and react to a traffic sign. These systems must operate effectively under highly variable conditions, including intense glare, nighttime low-light visibility, heavy rain, fog, dust storms, and rapidly changing urban landscapes. Traditional computer vision algorithms are generally too computationally intensive or brittle to maintain consistent performance under such conditions, and their inference time often exceeds safe thresholds for real-time decision-making. To overcome these operational limitations, contemporary TSR systems utilize lightweight deep learning architectures such as YOLOv4-Tiny, MobileNet, EfficientNet-Lite, and Tiny RetinaNet. These models strike a balance between speed and accuracy, allowing for high frame-per-second (FPS) processing and sub-60 millisecond latency, making them viable for deployment on resource-constrained edge devices such as Raspberry Pi, NVIDIA Jetson Nano, Qualcomm Snapdragon, and other automotive-grade CPUs and GPUs. During model training, extensive data augmentation is performed to simulate adverse environmental conditions—e.g., rain streaks, motion blur, glare, and low-light scenarios—to ensure robustness in real-world driving contexts. Evaluation metrics now go beyond accuracy to include latency, throughput (FPS), and energy efficiency, all of which are critical for ensuring that TSR systems can reliably support driver alerts, adaptive cruise control, emergency braking, or lane-keeping assistance within stringent temporal constraints. Ultimately, real-time performance is a non-negotiable feature of TSR systems, essential for ensuring safety and functional integration with intelligent transportation systems in high-speed, dynamic environments.

III. Existing Datasets and Their Limitations

Effective Traffic Sign Recognition (TSR) systems depend fundamentally on access to high-quality, diverse, and representative datasets that mirror the intricacies of real-world driving environments. Over the years, several benchmark datasets have been developed to train and evaluate TSR models, each offering unique strengths but also presenting notable limitations—especially when applied to diverse and unstructured road settings like those in India.

The **German Traffic Sign Recognition Benchmark (GTSRB)** is among the most extensively used datasets in TSR research. It comprises accurately labeled images from 43 categories of German traffic signs, captured under controlled environmental and infrastructural conditions. While GTSRB excels in clarity, consistency, and annotation quality, it is inherently limited by its geographic and linguistic scope. As it only includes standardized German signage, it lacks the complexity of multilingual, occluded, or regionally varied traffic signs, making it insufficient for training models intended for global or non-uniform environments.

Similarly, the **LISA Traffic Sign Dataset** from the United States offers a broader variety in terms of weather and lighting conditions. This contributes to improved model resilience in dynamic real-world scenarios. However, like GTSRB, the LISA dataset is focused on North American traffic signs, primarily annotated in English. It lacks support for multilingual or multi-script traffic signage, which is essential in countries like India where signs often appear in multiple regional languages. Additionally, the relatively modest size of the LISA dataset limits its usefulness for training modern deep learning architectures that require large-scale, balanced datasets to avoid over-fitting. Recognizing the limitations of such datasets in the Indian context, the **Indian Driving Dataset (IDD)** was developed to reflect the real challenges of driving in India. IDD includes images of unstructured roadways, erratic traffic patterns, cluttered urban scenes, and non-standard sign designs. It effectively captures India's unique traffic scenarios, including occluded or partially damaged signs. While it marks a significant step forward for localized model development, IDD remains a work in progress. It suffers from class imbalance, where rare sign types are underrepresented, and inconsistencies in labeling accuracy, which can compromise model performance and generalization if used without further augmentation or validation.

To address the linguistic and visual diversity of Indian traffic systems more directly, the **Indian Traffic Sign Dataset (ITSD)** has been curated with an emphasis on multilingual and regionally diverse traffic signs. It includes images with challenges such as blur, shadow, occlusion, and contextual noise, making it especially valuable for training TSR models that must operate in the unpredictable real-world conditions typical of Indian roads. Nevertheless, despite its relevance and specificity, ITSD is still limited in scale and suffers from uneven class distribution. This restricts its effectiveness for comprehensive model training and highlights the need for further dataset expansion, standardized annotation protocols, and incorporation of rare or emerging traffic signs.

IV. Generalization Challenges in AI-based TSR Models

Despite notable advances in AI-powered TSR systems, several generalization challenges persist when deploying these models in real-world, diverse driving environments. These limitations are outlined below:

Domain Shift: It refers to a significant mismatch between the data distribution used during training and the conditions encountered during real-world deployment. Most AI-based Traffic Sign Recognition (TSR) models are developed using structured and standardized datasets like the German Traffic Sign Recognition Benchmark (GTSRB). These datasets assume uniform signage, consistent placement, and well-maintained road infrastructure. However, Indian roads present a vastly different scenario, characterized by irregular sign shapes and sizes, non-standard sign placements (such as on electric poles or tree trunks), and signs that are faded, rusted, or partially broken. Moreover, heavy visual occlusion from urban clutter, congested traffic, and informal structures like street vendor stalls further complicates detection. As a result, deep learning models trained on structured data struggle with accurate feature extraction, leading to high false negative rates and poor classification performance when deployed in heterogeneous environments.

Environmental Variability: Environmental factors significantly impact the robustness and generalizability of TSR systems. Variations in lighting, weather, and shadowing conditions alter the visual representation of traffic signs and degrade model performance. For instance, low illumination in night-time settings, dimly lit rural roads, or underpasses often leads to underexposed or blurry images where signs become indistinguishable. Adverse weather conditions such as rain, fog, dust storms, or smog reduce image clarity and introduce noise, which distorts crucial visual features. Shadows cast by trees, buildings, or moving vehicles may partially or completely obscure the signs, making it difficult for convolutional neural networks (CNNs) and object detectors to identify them correctly. These distortions negatively impact recall and classification accuracy, ultimately compromising the reliability of TSR systems in real-world scenarios.

Multilingual and Regional Sign Variants: Traffic signs in multicultural countries like India are often multilingual and vary widely across regions. Unlike standardized Western datasets, Indian signs may include text in Hindi, Tamil, Telugu, Bengali, and other regional scripts, either alone or in combination with English. However, most benchmark datasets like GTSRB or the LISA dataset are monolingual, primarily using Latin alphabets, which limits the model's ability to generalize to other languages. Furthermore, regional scripts such as Devanagari, Dravidian, and Bengali-Assamese have unique shapes and font styles that differ significantly from the alphabets used in training. In many cases, traffic signs incorporate mixed-language instructions or local idiomatic phrases, making the task even more complex. AI models trained without multilingual Optical Character Recognition (OCR) or script-specific feature extraction often fail to interpret these signs correctly, leading to misclassification or complete omission—especially in critical warning or regulation signs like "School Ahead" or "No Parking."

Annotation Inconsistencies: Accurate and consistent annotations are essential for training high-performing supervised learning models. However, many public traffic sign datasets suffer from annotation inconsistencies that undermine model robustness. Labeling errors, such as misclassified signs or incorrectly annotated bounding boxes, introduce noise into the learning process, leading to biased training and inaccurate predictions. Additionally, annotation format mismatches—where one dataset uses rectangular bounding boxes while another uses rotated or polygonal annotations—make it difficult to transfer or fine-tune models across datasets. Another common issue is class imbalance: rare or region-specific signs are underrepresented in most datasets, which causes models to be skewed toward more common classes. These inconsistencies collectively hinder the reliability of transfer learning and domain adaptation, resulting in overfitting on specific formats or environments and limiting the broader applicability of TSR models.

V. Review of AI Techniques Used

Technique	Advantages	Limitations	Analysis
CNN (Convolutional Neural Networks)	- Excellent at extracting spatial hierarchies from images - Capable of learning robust visual patterns	- Sensitive to image distortions - Performance drops with occluded or faded signs	CNNs form the foundation of most TSR models. They effectively learn shape and color patterns critical for classification but require large, well-annotated datasets. In challenging scenarios (e.g., dirt-covered or tilted signs), their feature recognition may degrade.
YOLO (You Only Look Once)	- Real-time object detection with high speed - Single-stage detector: detects and classifies simultaneously	- Often misses small objects or partially visible signs - Less robust in cluttered scenes	YOLO (especially versions v3–v5) is favored for real-time TSR due to its speed. However, in densely populated or visually complex environments (like Indian roads), small or partially obscured signs may be missed due to downsampling in initial layers.
SSD (Single Shot Multibox Detector)	- Balanced in terms of speed and accuracy - Performs well in medium-complexity environments	- Inferior to YOLO in low-light or highly dynamic conditions	SSD performs slightly slower than YOLO but with more accuracy on mid-sized signs. Its performance declines in extreme lighting variations (e.g., dusk/dawn), which limits its standalone use in unstructured road conditions.

ISSN: 2393-8048
International Advance Journal of Engineering, Science and Management (IAJESM)
Multidisciplinary, Multilingual, Indexed, Double Blind, Open Access, Peer-Reviewed, Refereed-International Journal.
SJIF Impact Factor =8.152, January-June 2025, Submitted in March 2025

MobileNet	- Lightweight architecture - Optimized for mobile and edge computing - Lower memory and computational requirements	- Compromised accuracy due to shallow depth - May underperform on complex sign features	MobileNet is widely used for deploying TSR on embedded systems in ADAS. Its speed and efficiency come at the cost of lower feature granularity, making it less effective in differentiating visually similar traffic signs (e.g., 30 km/h vs. 80 km/h).
Transfer Learning (e.g., ResNet, VGG)	- Pre-trained on large-scale datasets like ImageNet - Highly effective when annotated TSR data is scarce	- Domain shift issues when pre-trained on non-traffic datasets - Fine-tuning required for region-specific signs	Transfer learning enables leveraging deep features learned from large, generic datasets. While effective in accelerating training, mismatches between source (ImageNet) and target domain (multilingual road signs) can hinder generalization unless properly adapted through fine-tuning.

VI. Proposed Solutions and Future Directions

1. Train models on diverse datasets from various countries, lighting, weather, and signage formats to improve generalization across regions.
2. Use Generative Adversarial Networks (GANs) and simulators to generate rare or complex sign images, addressing class imbalance and edge cases.
3. Combine TSR with Optical Character Recognition (OCR) to interpret multilingual text-based signs, especially for countries like India.
4. Adopt AI-assisted labeling tools and active learning to reduce manual effort, improve consistency, and enhance dataset quality
5. Create standardized global evaluation protocols to ensure fair comparison and performance assessment across diverse TSR models and datasets.
6. Use techniques like few-shot learning, self-supervision, and contrastive learning to adapt models to new environments with minimal labeled data.
7. Deploy lightweight models (e.g., MobileNet, YOLOv8-Nano) using quantization and pruning for real-time performance on embedded systems.
8. Enable models to learn new signs over time without forgetting previous knowledge, using curriculum or lifelong learning methods.
9. Integrate explainability tools (e.g., Grad-CAM, SHAP) to visualize model decisions and improve trust in real-world applications.

VII. Conclusion

Artificial Intelligence (AI) has revolutionized Traffic Sign Recognition (TSR) systems by enabling them to process visual data at scale with impressive accuracy, thereby enhancing driver assistance systems and paving the way for autonomous vehicles. Techniques such as Convolutional Neural Networks (CNNs), Single Shot Detectors (SSDs), YOLO (You Only Look Once), and Transformer-based models have demonstrated state-of-the-art performance in recognizing and classifying traffic signs in benchmark datasets. However, these successes are largely confined to controlled experimental conditions or structured driving environments, such as those found in developed nations. A fundamental limitation still hinders widespread, reliable deployment: the absence of diverse, large-scale, well-annotated, and contextually rich datasets that reflect real-world complexities. Most existing datasets, such as GTSRB (German Traffic Sign Recognition Benchmark) and LISA (Laboratory for Intelligent and Safe

ISSN: 2393-8048
International Advance Journal of Engineering, Science and Management (IAJESM)
Multidisciplinary, Multilingual, Indexed, Double Blind, Open Access, Peer-Reviewed, Refereed-International Journal.
SJIF Impact Factor =8.152, January-June 2025, Submitted in March 2025

Automobiles), are heavily region-specific, recorded under ideal conditions with standardized traffic signs, clean backgrounds, consistent lighting, and clear visibility. These datasets rarely account for the chaotic and variable nature of real-world traffic scenes, especially in countries like India where signage can be weathered, occluded, tilted, vandalized, or even partially broken. Moreover, multilingual signs (e.g., English, Hindi, Tamil) and handwritten or regionally customized boards present substantial challenges that are not addressed in traditional datasets. As a result, AI models trained on such narrow data sources often exhibit poor generalization, leading to degraded performance in unstructured, heterogeneous environments. The problem is compounded by the underrepresentation of rare or edge-case traffic signs, inconsistent annotation standards, and the lack of multimodal data (such as textual overlays, speech cues, or environmental metadata). For instance, faded signs under low-light or rainy conditions are frequently misclassified or missed altogether. These limitations are particularly critical for autonomous systems that require high-confidence decision-making in real time. Without access to training data that mirrors the target deployment domain, even the most advanced AI architectures become unreliable. To overcome these challenges, there is an urgent need for the creation of multi-environment, multilingual, and multimodal datasets. Such datasets should span different countries, urban and rural areas, varied weather and lighting conditions, multiple languages, and diverse sign formats. Data augmentation using Generative Adversarial Networks (GANs), simulation environments (e.g., CARLA, SYNTHIA), and synthetic data generation techniques can play a crucial role in supplementing real-world data, especially for rare classes or adverse scenarios. Additionally, active learning frameworks that integrate human-in-the-loop annotations can improve labeling accuracy and reduce the burden of manual dataset curation. Moreover, AI-based TSR systems must evolve from static recognition tasks to context-aware, adaptive systems that can reason about occluded or ambiguous signs using temporal cues, vehicle telemetry, and map data. Incorporating multilingual OCR (Optical Character Recognition) can further enable interpretation of text-heavy signage, which is prevalent in developing regions. Evaluation protocols must also shift towards cross-domain benchmarking, where models are tested on unseen, domain-shifted data to truly assess their robustness and transferability.

References

1. World Health Organization, Global Status Report on Road Safety, Geneva, Switzerland, 2023.
2. J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, “The German Traffic Sign Recognition Benchmark: A Multi-Class Classification Competition,” in Proc. IEEE Int. Joint Conf. Neural Networks (IJCNN), 2011, pp. 1453–1460.
3. Ministry of Road Transport and Highways, Government of India, Road Accidents in India, New Delhi, India, 2022.
4. S. Varma, R. Subramanian, and A. Singh, “IDD: A Dataset for Exploring Problems of Autonomous Navigation in Unconstrained Environments,” in Proc. IEEE Winter Conf. Applications of Computer Vision (WACV), 2019, pp. 1743–1751.
5. Indian Institute of Science, Evaluation Report on Cross-Domain Generalization of Traffic Sign Recognition Models, Bengaluru, India, 2022.
6. IIT Delhi, Study on Traffic Sign Visibility in Semi-Urban India, New Delhi, India, 2021.
7. P. Pundir and A. Mahajan, “Indian Traffic Sign Dataset (ITSD),” International Journal of Computer Vision and Signal Processing, vol. 8, no. 2, pp. 59–67, 2021.
8. P. Dahiya, A. Mehta, and R. Aggarwal, “Traditional Techniques for Traffic Sign Detection: A Review,” International Journal of Signal Processing Techniques, vol. 5, no. 1, pp. 12–18, 2018.
9. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in Proc. IEEE Conf. Computer Vision and Pattern

Recognition (CVPR), 2016, pp. 779–788.
10. R. Gupta and K. Sharma, "Real-Time Traffic Sign Detection Using PyTorch and OpenCV," Journal of AI in Transport, vol. 4, no. 1, pp. 21–29, 2022.
11. P. Pundir and A. Mahajan, "Indian Traffic Sign Dataset (ITSD): A Step Towards Contextual AI in Traffic Sign Recognition," Journal of Artificial Intelligence and Data Science, vol. 9, no. 3, pp. 122–131, 2021.
12. S. Varma, R. Subramanian, and A. Singh, "IDD: A Dataset for Exploring Problems of Autonomous Navigation in Unconstrained Environments," in Proc. IEEE Winter Conf. Applications of Computer Vision (WACV), 2019, pp. 1743–1751.
13. N. Meena and R. Sharma, "Performance Evaluation of YOLOv3 and Faster R-CNN for Indian Traffic Sign Detection," International Journal of Computer Vision and Intelligent Systems, vol. 10, no. 2, pp. 109–117, 2021.
14. A. Bansal and P. Kumar, "Transfer Learning-Based Traffic Sign Recognition for Indian Conditions Using ResNet and MobileNet," International Journal of Deep Learning and AI Applications, vol. 8, no. 4, pp. 45–52, 2021.
15. K. Sinha and R. Kapoor, "A Hybrid CNN-SVM Model for Multilingual and Obscured Traffic Sign Recognition," International Journal of Artificial Intelligence in Transportation, vol. 6, no. 3, pp. 134–141, 2022.
16. D. Raj and S. Verma, "Impact of Annotation Errors on Traffic Sign Recognition Models: A Study on Indian Datasets," Journal of Intelligent Data Systems, vol. 7, no. 1, pp. 77–84, 2020.
17. R. Sharma, A. Iyer, and M. Pillai, "Multilingual Traffic Sign Detection Using CNN and OCR Fusion: A Case Study from India," IEEE Access, vol. 10, pp. 43712–43720, 2022.
18. S. Joshi and M. Rathi, "Improving Traffic Sign Recognition with Data Augmentation Techniques: An Indian Context," Pattern Recognition and Intelligent Systems Journal, vol. 9, no. 2, pp. 95–103, 2021.
19. A. Upadhyay and S. Raj, "Real-Time Traffic Sign Detection on Embedded Hardware Using YOLOv5-Tiny," Journal of Embedded AI Systems, vol. 11, no. 1, pp. 63–70, 2022.
20. K. Saxena and R. Bhatnagar, "Cross-Domain Adaptation for Traffic Sign Recognition in Indian Road Scenarios," International Journal of Computer Vision and Smart Mobility, vol. 8, no. 2, pp. 81–89, 2023.