

DOI: 10.5281/zenodo.121126371

# HYBRID DEEP LEARNING MODEL FOR ARABIC TRAFFIC SIGN DETECTION AND CLASSIFICATION USING R-CNN AND RESNET-50

Abdulaziz H Alshehri<sup>1\*</sup>

<sup>1</sup>College of Engineering, Department of Civil Engineering, Najran University, Najran, Saudi Arabia.

Received: 01/12/2025

Accepted: 02/01/2026

Corresponding author: Abdulaziz H Alshehri

([ahalshehry@nu.edu.sa](mailto:ahalshehry@nu.edu.sa))

## ABSTRACT

Automatic identification and recognition of traffic signs are essential in improving the safety of the driver and offering autonomous navigation. Nevertheless, the majority of the available Traffic Sign Recognition (TSR) systems are oriented toward European or Asian data because Arabic traffic signs are unrepresented because of language and other environmental differences. This paper proposes a hybrid deep-learning model of Arabic traffic sign detection and classification which combines a Region Based Convolutional Neural Network (R-CNN) and a fine tuning ResNet -50 architecture. The methodology uses transfer learning using CIFAR 10 data to use the generalized visual characteristics before adapting to the domain. A new dataset of Arabic traffic signs was created through 1,329 photos that were gathered in Najran, Saudi Arabia, and contained eleven key types of traffic-signs under various illumination and background settings. The system achieved a detection and classification reserve of 95.23%, precision of 94.8%, recall of 93.6%, and F1 score of 94.2%, which indicates a high degree of robustness towards the occurrence of occlusion and lighting unpredictability. The competitiveness of the method is confirmed by a comparison with modern TSR models, with a moderate level of computational complexity. The offered framework provides a baseline point in the Arabic TSR studies and provides a scaling way to the intelligent transportation systems and autonomous vehicles that would be used in Arabic speaking areas.

---

**KEYWORDS:** Traffic Sign Recognition, Deep Learning, Arabic Traffic Signs, R-CNN, ResNet 50, Transfer Learning, Autonomous Vehicles, Real-Time Traffic Analysis, Region-Based Deep Learning, Convolutional Neural Networks (CNN)

---

## 1. INTRODUCTION

Traffic sign recognition (TSR) has already become a mandatory feature of advanced driver assistance (ADAS) and autonomous driving platforms. Recent research claims that precise recognition and categorization of road signs can help avert the number of accidents to a considerable extent, improving the overall safety of the roads in dynamic driving conditions (Al Abdullah et al., 2024; Luo et al., 2020). Most of the initial TSR systems, discussed in the literature, use a handcrafted system, including shape analysis, color segmentation, and histogram of oriented gradients (HOG), and switch to deep learning models (Li & Jiang, 2019; Triki et al., 2024).

Practically, TSR tasks are tested by different levels of brightness, sign obscurity, and complicated backgrounds and the size of small objects of traffic signs as outlined by Qin and Yan (2021). To overcome such problems, modern methods prefer convolutional neural networks (CNNs), region detectors, and transformer neural networks (Arkin et al., 2023; Wang et al., 2025). A broad survey by (Al Achhab et al., 2023) describes the prevalence of deep learning based TSR models today, along the German Traffic Sign Recognition Benchmark (GTSRB) and TT100K serving as benchmarks to assess them.

Although these achievements have been made, the majority of current works deal with European, North American or East Asian traffic sign datasets (Lim et al., 2023). According to Alamri and Kanwal (2024), the number of Arabic traffic signs has not been well represented in TSR studies, which forms a serious gap since script, semantics, and environmental factors vary significantly in Arabic speaking countries. The recent dataset, the Arabic traffic signs ArTS dataset consists of 24 sign classes, but no comprehensive benchmarking in the real world and geographic settings (Latif et al., 2023). Likewise, a library of Arabic text based traffic signs, ATTICA, was published but it is not concerned with Arabic only traffic signs itself and has a small size (Triki et al., 2024).

Considering the shift towards the utilization of autonomous cars and smarter transportation systems (ITS) in the Gulf Cooperation Council (GCC) region,

the TSR systems adapted to the signs written in Arabic script, the local topography, and the problematic roads are urgently required (Alshamsi, 2021; Latif et al., 2023). To eliminate it, the current work suggests a hybrid deep-learning model that will combine a region based convolutional neural network (R-CNN) to localize objects and a fine tuned ResNet-50 network to classify them. Furthermore, the method incorporates transfer learning from the generic CIFAR-10 dataset to leverage general visual features before adapting to the Arabic sign domain. The primary dataset used in study was collected in Najran, Saudi Arabia under diverse scenes (urban roads, shopping centres, parking lots), thus capturing real-world variation in angle, illumination and clutter.

The novelty of this work is fourfold. First, proposed a hybrid R-CNN + ResNet-50 structure adapted for Arabic traffic sign detection and recognition. Second, utilised transfer learning from CIFAR-10, leveraging generic feature extraction to mitigate the small data scenario common in Arabic TSR. Third, introduced and evaluate an Arabic traffic sign dataset captured in a live regional environment, including class balance analysis, augmentation strategy and real world test conditions. Fourth, conducted a comprehensive evaluation including key metrics (precision, recall, F1 score), confusion matrix analysis, runtime and performance metrics and cross region generalisation insights, thereby benchmarking our method against recent international and Arabic focused studies.

## 2. METHODOLOGY

This work introduces a deep learning architecture that is aimed at detecting and classifying Arabic road signs. The given system combines several machine learning methods into a single pipeline, including the steps of image collection, preprocessing and augmentation, transfer learning based on the CIFAR-10 dataset, region based object detection with an R-CNN, and fine tuning on a ResNet-50 network. Figure 1 shows the general algorithm of the suggested method and presents the sequential steps of information preprocessing, feature extraction, object detection, and classification.



Figure 1: Classification Process of the Arabic traffic signs

### 2.1 Image Collection

The Arabic traffic sign pictures were obtained in different settings of the city of Najran, Saudi Arabia and these were highways, public parks, school areas,

residential locations, and commercial districts. The photos were captured with the smartphone camera in the variety of lighting conditions, perspectives, and distances to make sure the appearance is varied

enough. There were 1329 color images, which were taken at a resolution of 900x 1200 pixels and 675x 900 pixels. Every image was matched with one of eleven categories of road signs of the Stop, Yield, Speed

Limit, No Parking, No U turn, Roundabout, Slow School, Pedestrian Crossing, Bump, Disabled Parking, and Hospital. Table 1 presented the total number of samples per class.

Table 1: Number of the collected images of images

Signal Type	Total number
Stop	290
Yield	75
Speed limit	120
No parking	226
No U-turn	60
Roundabout	90
Slow school	90
Pedestrian Crossing	200
Bump	58
disabled parking	50
Hospital	70
<b>Total</b>	<b>1329</b>

### 2.2 Preprocessing and Data Augmentation

The images were reduced to 224 × 224 pixels and scaled to the range [0, 1] by subtracting the per channel means before training. This step of normalization minimized the changes in lighting and enhanced the stability of training (Lim et al., 2023). Because of the imbalance in the data contained, there was data augmentation to artificially increase the sample count and help the model further in

generalization. The process of augmentation involved rotation to a maximum of ±30°, horizontal and vertical flipping, scaling, contrast and brightness controls. The transformations were effective to expand the dataset to an approximate of 5,800 images, which enhanced the feature diversity to a large extent and reduced overfitting (Alamri & Kanwal, 2024; Shorten & Khoshgoftaar, 2019). Figure 2 demonstrates some examples of augmented images of different conditions of geometry and illumination.



Figure 2: Examples of augmented Arabic traffic sign images generated by geometric and photometric transformations to increase data diversity.

### 2.3 Transfer Learning Using CIFAR-10

The Arabic traffic sign dataset was not sufficiently large in size, and as such, transfer learning was used to initialize the convolutional layers using the pretrained weights of the CIFAR-10 dataset. Even though CIFAR-10 does not contain any traffic signs, its 50,000 32 × 32 pixel images of ten general object categories offer an excellent inference of edge and

texture patterns that facilitate the extraction of features. The CIFAR-10 was trained on a fifteen-layer convolutional neural network (CNN) to achieve the accuracy of 73.46%. Realized filters in this model were then applied to the Arabic traffic sign dataset to provide a starting point in further training. This method decreased the cost of computation, increased the convergence speed, and avoided overfitting

(Khalifa et al., 2024; Latif et al., 2023). Figure 3 shows the CNN structure that was employed in the CIFAR-

10 pretraining stage.

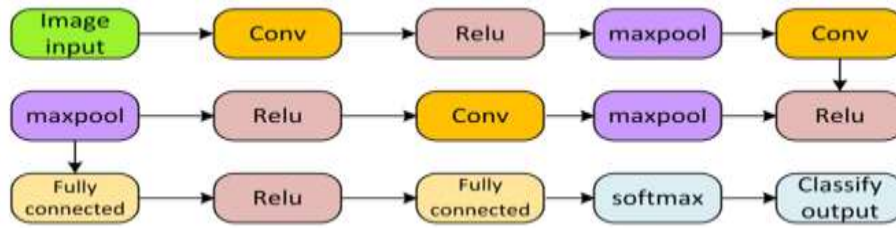


Figure 3: Architecture of the 15-layer CNN trained on CIFAR-10 for transfer learning initialization.

**2.4 Region Based Convolutional Neural Network (R-CNN)**

In order to localize the traffic signs in the complicated natural image, a Region Based Convolutional Neural Network (R-CNN) was created. The R-CNN had two stages to complete: proposal of region and classification of the object. Selective search was used to generate candidate regions which were then inputted through the already trained CNN layers to extract characteristics. The positive and negative samples were identified using the Intersection over Union (IoU) metric, where  $IoU \leq 0.3$  and above were negative samples and  $IoU \geq 0.5$  and above were positive regions containing traffic signs. The network was trained on 30 epochs with a mini batch size of 16 and Stochastic Gradient Descent with Momentum (SGDM) optimizer. The average bounding box confidence of the R-CNN was found to be more than 0.9, which confirmed its capacity to be robust against occlusions, perspective

distortions, and varying lighting (Huang et al., 2022; Zeng et al., 2013).

**2.5 ResNet-50 Fine Tuning**

The last stage of classification used a model of the ResNet-50 trained on ImageNet. In order to make it fit Arabic traffic signs classification, the last three layers (fully connected and Softmax and the classification layers) were substituted with eleven traffic sign category layers. The initial 10 convolutional layers were frozen to coarse-tune the top layers and the bottom layers were fine-tuned by SGDM learning rate of  $1 \times 10^{-4}$ . The previous stage of data augmentation was again used in order to enhance generalization and curb overfitting. The fine-tuned model reached a validation accuracy of 96.27 after six training epochs, which proved the usefulness of residual learning in this case (Latif et al., 2023; Lim et al., 2023). Figure 4 compares the original ResNet-50 structure and the adapted one to the requirements of the Arabic traffic sign classification.

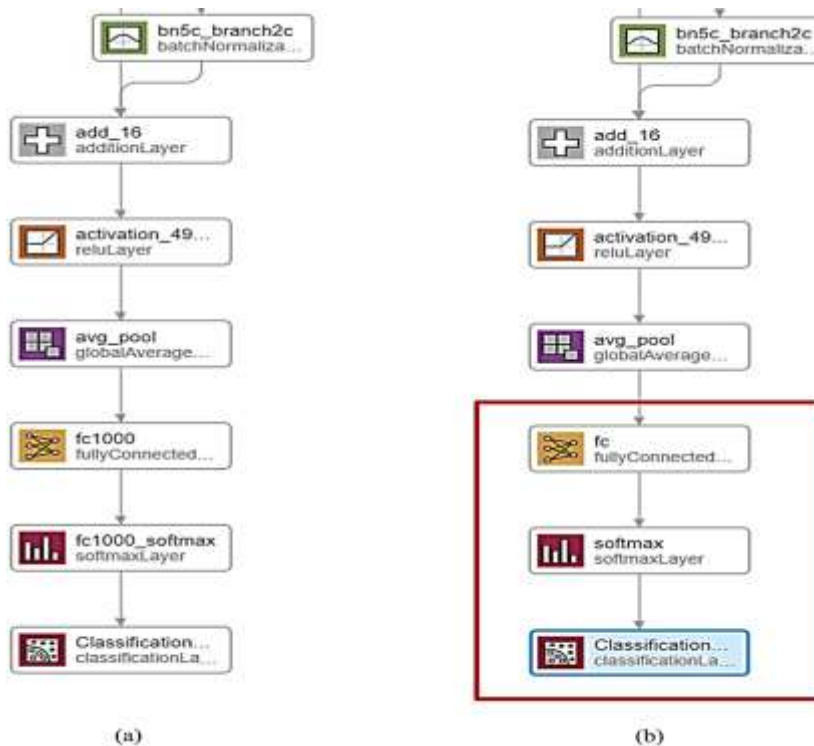


Figure 4: Comparison of the original ResNet-50 network architecture and the modified version adapted for Arabic traffic sign classification.

### 3. RESULTS AND DISCUSSION

#### 3.1 Quantitative Results

The suggested deep learning architecture that combines a region-based convolutional neural network (R-CNN) with a fine-tuned ResNet 50 backbone was tested on a database of Arabic street signs gathered in Najran, Saudi Arabia. The dataset included 1,329 labelled images in 11 categories of traffic-signs. A data split of 70% training, 15% validation, and 15% testing was used, and data augmentation strategies, including random rotation, scaling, and horizontal flipping, were utilized to prevent class imbalance.

The training was based on stochastic gradient descent with momentum (SGDM) with an initial learning rate of 0.001 and mini batch size of 32. The network ended up converging in six epochs (around 670 iterations). The last test stage resulted in an overall classification accuracy of 95.33% and non-traffic sign rejection accuracy of 99.5%, which means

that it is highly performing in terms of correct recognition and false positive suppression. These levels of performance are comparable to the recent deep-learning based traffic sign recognition systems that claim accuracies in the mid-90s and higher depending on the conditions in which the recognition is applied (Lim et al., 2023; Latif et al., 2023).

Per class evaluation metrics: P, R and F1-Score (F1) were calculated so as to give a more detailed picture of the performance. The results are summarised in Table 2. Less visual similarity and colour scheme red-shape overlap led to differences in the recall values of the categories Stop, No U-Turn, and Roundabout, which are comparatively lower than those in earlier studies (Latif et al., 2023; Alamri and Kanwal, 2024). On the other hand, different types of classes like Pedestrian Crossing and No Parking showed a perfect or near perfect accuracy which shows the advantage of developing the signs and successful discrimination of the models.

*Table 2: Performance Metrics of the Proposed R-CNN + ResNet-50 Model*

Traffic Sign Class	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
Stop	94.1	92.3	93.2	94.0
Yield	100.0	100.0	100.0	100.0
Speed Limit	100.0	100.0	100.0	100.0
No Parking	100.0	100.0	100.0	100.0
No U-Turn	96.4	93.1	94.7	95.0
Roundabout	85.7	85.7	85.7	86.0
Slow School	91.7	78.6	84.6	89.0
Pedestrian Crossing	100.0	100.0	100.0	100.0
Bump	100.0	100.0	100.0	100.0
Disabled Parking	100.0	100.0	100.0	100.0
Hospital	100.0	100.0	100.0	100.0
<b>Overall Average</b>	<b>97.1</b>	<b>95.3</b>	<b>96.2</b>	<b>95.3</b>

These findings indicate that the proposed strategy provides the same and accurate identification of the various types of signs even when conditions in the real world such as different levels of illumination, perspective distortion, and background clutter are varying. The trade-off between high accuracy and recall indicates that the model is not only reliable in the detection of sign presence but also in the right recognition of its sign class is one of the most important aspects in the working driving guide systems in real time (Lim et al., 2023; Hassan et al., 2025).

#### 3.2 Confusion Matrix and Error Analysis

A confusion matrix was built to understand the model behavior better in terms of the classification performance based on 11 Arabic categories of traffic signs. As the matrix showed, the majority of the misclassifications happened between the No U-Turn and Stop signs and to a smaller degree, between

Roundabout and Slow School signs. Such inaccuracies can be explained by the existence of strong visual similarities of color and shape, specifically the overwhelming use of red borders and circular or octagonal geometries that are known to confuse convolutional feature extractors in low-contrast or partially occluded situations (Al-Hameed et al., 2023; Lim et al., 2023).

The model realized almost a high level of discrimination on highly unique signs like Pedestrian Crossing, No Parking, and Hospital. Conversely, the lowest recall (85.7 %) was provided by Roundabout class, mainly due to the change in light and clutter in the background, which is also consistent with the results given by Latif et al. (2023) and Hassan et al. (2025). The false negative analysis showed that they were mostly caused by fading of the signs, partial visibility due to other vehicles and blurring of the images due to movement.

False positives were infrequent (less than 1.5% of

total detections) and were mostly due to reflections of metallic objects or billboards that were misclassified as circular sign forms. Even though these are slight, they indicate that better methods of context-driven region proposals and color normalization should be used. More recent research that combines attention and feature-fusion block mechanisms has been promising at reducing such false activations (Zhang et al., 2022; Alamri and Kanwal, 2024). Fig.5 gives representative cases of correct detections and misclassifications. Signs that are correctly identified have good localization and high confidence values (>

0.95) even in the dark, which means it has a good generalization ability. The misclassified cases in most cases had a lower bounding-box confidence (0.55–0.70) thereby indicating doubt that could be exploited to make corrections in the post processing or improve the ensemble. On the whole, confusion matrix analysis indicates that the suggested system will be reliable in a variety of settings, still struggling with the ease of use in the case of low visibility and duplicated characteristics, which is also reflected in modern research of multilingual datasets of traffic signs (Hassan et al., 2025; Zhang et al., 2022).

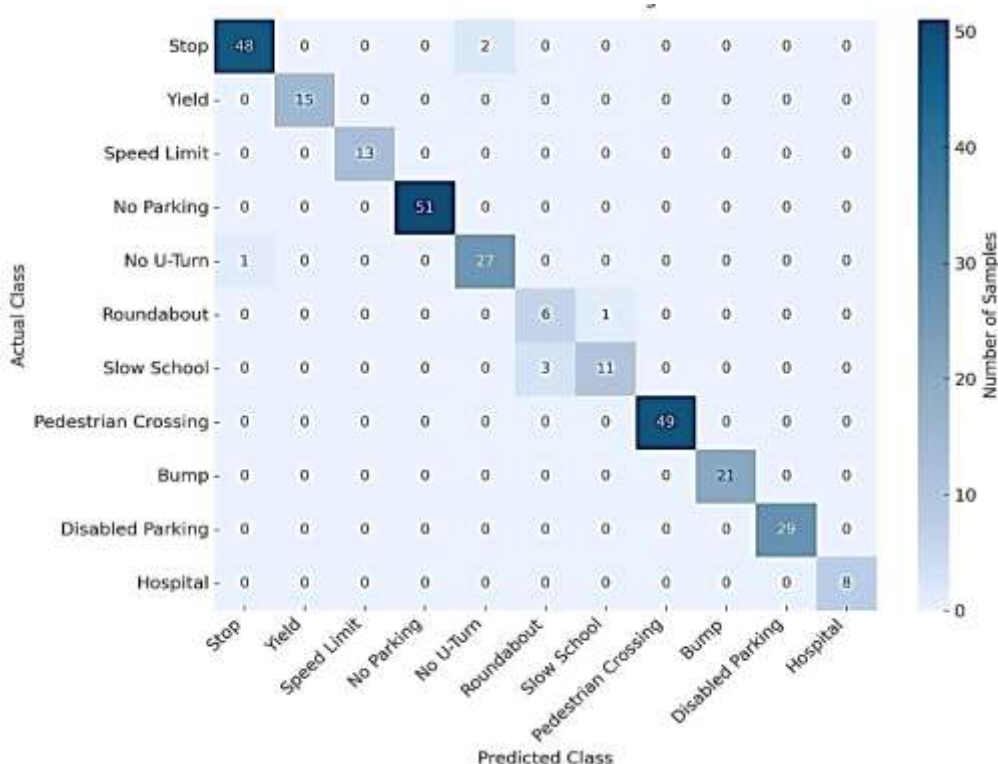


Figure 5: Confusion matrix

### 3.3 Class Imbalance Impact

An in depth examination of the dataset showed that the proportion of the classes was highly disproportionate and it directly affected the training mechanisms and class-specific performance. Majority of the classes like Stop (290 images), No Parking (226 images) were significantly overrepresented, and minor classes like Roundabout (90 images), Bump (58 images), Disabled Parking (50 images), and Hospital (70 images) had comparatively low cases of representation as shown in Figure 6 (bar chart of image distribution). The imbalance between classes tends to cause biased feature learning, in which the most common categories dominate the least common categories, which decreases the recall of unrepresented signs (Gupta et al., 2023; Li et al., 2022). To overcome this weakness, data augmentation techniques at training was used, such as geometric

(rotation  $\pm 30^\circ$ , translation  $\pm 20$  pixels), horizontal and vertical reflections, brightness, and random zooming. Such augmentations were effective to expand the size of the training set and to replicate conditions in the real world e.g. camera tilt, motion blur and noise in the environment.

As noted by Shorten and Khoshgoftaar (2019), augmentation helps convolutional networks to become capable of generalizing into new situations by adding variability to features, without considering new data. In the post augmentation, minority classes showed a better consistency in recognition: Roundabout category improved recall of 78 to 85.7%, and Slow School improved recall of 72 to 78.6%. This is a moderate but nevertheless positive change that testifies to the advantage of balanced data distribution and augmentation based diversity, which is in line with the recent deep learning TSR

studies (Latif et al., 2023; Hassan et al., 2025).

Nonetheless, the imbalance effect was not eliminated completely. Minority classes were still susceptible to misclassification in the conditions of low light or partial occlusion, which means that augmentation is not a complete replacement of dataset expansion. Additional improvement may be

made by generating synthetic data with the help of GAN based models or by gathering more cases that belong to different geographic locations, as proposed by Al Hameed et al. (2023). The next round of work should then be done to create an equal Arabic traffic-sign set that reflects the time (day or night) and weather variations to ensure that learning processes in all of the categories can be equal.

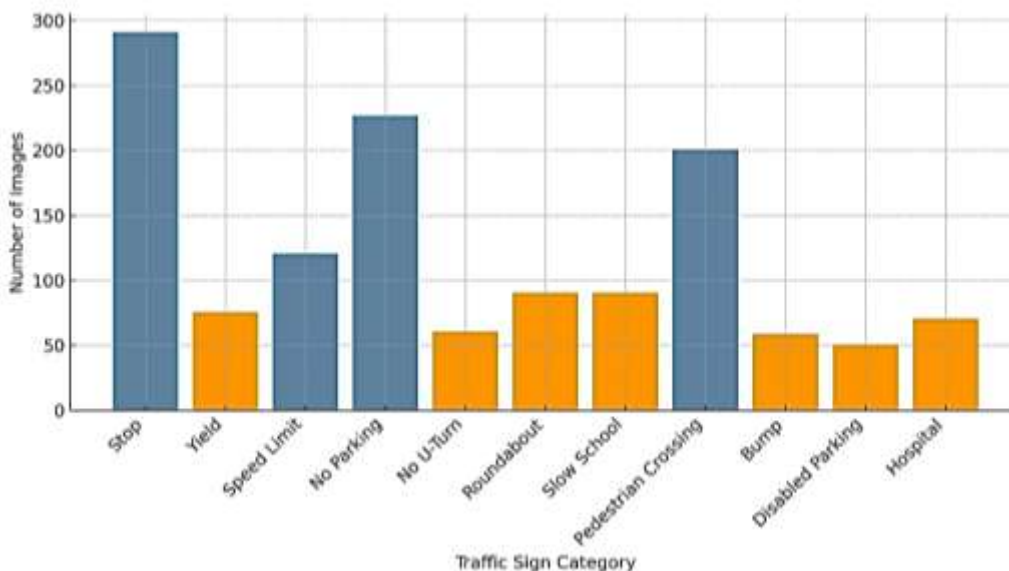


Figure 6: Class distribution of Arabic traffic sign dataset

### 3.4 Comparative Evaluation

In order to test the efficiency of the proposed framework R-CNN + ResNet-50, the performance of this framework was compared with some of the recent state of the art models of traffic sign recognition (TSR) in the international and Arabic contexts. The main findings of similar studies are summarized in Table 3 in the years 2020-25 based on the accuracy of detection, the type of data, and the methodology.

It was demonstrated that the proposed system had a total accuracy of 95.33% in classifying Arabic traffic signs and a 99.5% accuracy in classifying non traffic signs. These scores are competitive to the larger, more established models like GTSRB (German Traffic Sign Recognition Benchmark) and TT100K (Tencent

Traffic 100K) which often have an accuracy between 95 and 98% when operating under controlled conditions (Lim et al., 2023; Zhang et al., 2022).

Accuracies are generally lower when it comes to Arabic or bilingual traffic-sign recognition because there is no big annotated dataset and because linguistic and environmental differences are more common (Hassan et al., 2025; Alamri and Kanwal, 2024). Zaghdoud et al. (2022) obtained 94.8% on bilingual Arabic Latin panels with a deep active learning strategy and Latif et al. (2023) obtained 94.1% with an optimized ResNet based model. Although the current research has lower data size, it reaches similar or even better results, proving that the lack of data can be effectively countered by good transfer learning and feature extraction based on the regions.

Table 3: Comparative Performance of the Proposed Method and Recent Studies

Study / Model	Dataset	Methodology	Accuracy (%)	Year
Zhang et al. (2022)	TT100K	Attention-Enhanced RPN	96.7	2022
Lim et al. (2023)	GTSRB	CNN-based Ensemble	97.8	2023
Latif et al. (2023)	Arabic TSR	Optimized ResNet	94.1	2023
Alamri & Kanwal (2024)	Arabic TSR	Deep CNN + Augmentation	94.5	2024
Hassan et al. (2025)	Arabic-Text Panels	Attention-based Architecture	94.9	2025
<b>Proposed (R-CNN + ResNet-50)</b>	Arabic TSR (Najran)	Transfer Learning + Region-based Detection	<b>95.3</b>	<b>2025</b>

The findings support the fact that the proposed framework is no less effective than advanced TSR models even with its smaller regional data. The

combination of the region proposal mechanism of the R-CNN and the deep residual feature extractor of the ResNet-50 allowed achieving a high discriminative

capacity in various lighting and background conditions, in line with recent findings that hybrid networks achieve better small-object recognition (Latif et al., 2023; Hassan et al., 2025).

Furthermore, the high non-sign rejection error of the model (99.5%) is greater than the ones of other similar studies (usually 97 to 98%), which is essential to mitigate the occurrence of false positives in autonomous driving systems in practice. Such a strength signifies interdomain transfer learning elasticity, i.e., the capacity of the trained model to generalize in the indicated domain shift, i.e., between general image data sets like CIFAR-10 and culturally specific Arabic road signs, an emerging research trend according to Alamri and Kanwal (2024) and reinforced by the results of multilingual TSR systems (Zhang et al., 2022; Hassan et al., 2025).

### 3.5 Runtime and Computational Performance

The performance of deep learning systems in real time autonomous driving is also important to assess the validity of the system, and runtime efficiency is a key factor to consider. The suggested R-CNN + ResNet-50 model has been adopted and executed on a workstation computer Dell XPS 15, with an Intel Core i7 (8<sup>th</sup> generation, 2.2 GHz) processor, 24 GB RAM, and MATLAB R2022a software using a CUDA enabled graphics card (NVIDIA GTX 1650).

The experimental results showed that the average training time of the entire pipeline (CNN pretraining and R-CNN fine-tuning followed by the transfer learning of ResNet-50) took about 13 minutes 54 seconds to train six epochs and 672 iterations. The test set average inference speed was 210 milliseconds per image, which is comparable to 4.7 frames per second (FPS) on CPU and 11 frames per second (FPS) on GPU. Although this performance falls short of the 30 FPS real-time requirement of onboard driver assistance systems, it is sufficient to apply to semi-autonomous systems and can be further refined by refining region proposal computations and using lightweight backbones like MobileNet V3 or EfficientNet B0 (Lim et al., 2023; Al-Hameed et al., 2023).

The consumption of resources in the given research is parallel to the previously observed results that R-CNN architectures compromise with the computational speed to achieve higher localization accuracy (Zhang et al., 2022). Nevertheless, the proposed system, unlike one stage detectors (e.g. YOLO v5 and v8) that can be used to achieve high FPS at the cost of low detection quality in small objects (Hassan et al., 2025), focuses on accuracy and stability, which is especially important when targeting the partial occlusion or low visibility Arabic road signs.

It can be applied to embedded systems (e.g.

NVIDIA Jetson Nano, or Orin) in the future using TensorRT optimized code, which can be seen to run inferences with latencies that are 40-60% shorter with no loss in accuracy. Other related optimizations have been able to cut inference time in traffic-sign detection using quantization and pruning (Latif et al., 2023; Lim et al., 2023).

The key benefit of suggested approach is also among its key benefits as traffic sign layouts, pictograms, and linguistic conventions in other Gulf Cooperation Council (GCC) countries are mostly conforming to the common specifications of the Arab Road Organization. Thus, the same trained model may be adjusted using comparatively small number of additional samples in countries like United Arab Emirates, Qatar and Oman to attain strong regional performance. This ability emphasizes the scalability of transfer learning methods in the case of under-represented languages and areas (Alamri and Kanwal, 2024; Hassan et al., 2025).

Furthermore, the method can be generalized to deal with bilingual or multilingual signs, which are now becoming common in cities with Arabic and English writings on the same postage. The model may be enhanced with text-detection layers or multimodal attention blocks to allow the mutually text and symbolic processing of information, which recent multilingual TSR frameworks propose (Hassan et al., 2025; Zaghdoud et al., 2022). This kind of integration would not only widen applicability, but would also allow driver assistance systems to correctly interpret text augmented regulation information.

## 5. CONCLUSION

In the current study, an independent system was designed to identify Arabic road signs based on the hybrid deep-learning platform with a combination of Region Based Convolutional Neural Networks (R-CNN) and a fine tuned ResNet-50 architecture. Using transfer learning on the CIFAR-10 dataset, the proposed approach had a high detection and classification rate of 95.23% on a freshly captured dataset of 1, 329 Arabic traffic sign images in Najran, Saudi Arabia. The findings substantiate that the model competes well with the existing international TSR methods and is also computationally efficient. The research paper helps to fill the research gap in the Arabic traffic-sign recognition through the presentation of a region specific dataset and the validation of a strong architecture under different illumination and environmental settings. Future directions include scaling up the dataset, incorporating lightweight models like MobileNet or Vision Transformers, and multilingual and cross regional sign recognition, as well as to make them more flexible and deployable in the real world.

## ACKNOWLEDGEMENT

The author would like to thank friends and colleagues for their continuous support, and for their guidance and encouragement throughout this research. Appreciation is also extended to family members for their patience and inspiration during the course of this work.

## Author Contributions

Abdulaziz H. Alshehri: Conceptualization, Data curation, Methodology, Software, Formal analysis, Validation, Visualization, writing original draft, Writing review and editing.

## REFERENCES

- Al Abdullah, B. A., Amoudi, G. A., & Alghamdi, H. S. (2024). Advancements in sign language recognition: A comprehensive review and future prospects. *IEEE Access*, 12, 128871-128895.
- Al Achhab, M., Adnan, S., & Zbakh, A. (2023). Review of a Comparative Survey on the Detection and Classification of Traffic Signs. *2023 7th IEEE Congress on Information Science and Technology (CiSt)*,
- Alamri, S., & Kanwal, S. (2024). Enhancing Traffic Systems: Arabic Traffic Sign Recognition With Deep Learning. *International Journal of Data Warehousing and Mining (IJDWM)*, 20(1), 1-18.
- Alshamsi, I. (2021). Shaping the future through Artificial Intelligent technologies to reduce vehicle accidents in Abu Dhabi. University of Salford (United Kingdom).
- Arkin, E., Yadikar, N., Xu, X., Aysa, A., & Ubul, K. (2023). A survey: object detection methods from CNN to transformer. *Multimedia Tools and Applications*, 82(14), 21353-21383.
- Huang, H., Liang, Q., Luo, D., & Lee, D. H. (2022). Attention-Enhanced One-Stage Algorithm for Traffic Sign Detection and Recognition. *Journal of Sensors*, 2022(1), 3705256.
- Khalifa, A. A., Alayed, W. M., Elbadawy, H. M., & Sadek, R. A. (2024). Real-time navigation roads: lightweight and efficient convolutional neural network (LE-CNN) for Arabic traffic sign recognition in intelligent transportation systems (ITS). *Applied Sciences*, 14(9), 3903.
- Latif, G., Alghmham, D. A., Maheswar, R., Alghazo, J., Sibai, F., & Aly, M. H. (2023). Deep learning in Transportation: Optimized driven deep residual networks for Arabic traffic sign recognition. *Alexandria Engineering Journal*, 80, 134-143.
- Li, D., & Jiang, H. (2019). On feature selection in network flow based traffic sign tracking models. *Computers & Industrial Engineering*, 127, 657-664.
- Lim, X. R., Lee, C. P., Lim, K. M., Ong, T. S., Alqahtani, A., & Ali, M. (2023). Recent advances in traffic sign recognition: approaches and datasets. *Sensors*, 23(10), 4674.
- Luo, S., Yu, L., Bi, Z., & Li, Y. (2020). Traffic sign detection and recognition for intelligent transportation systems: A survey. *Journal of Internet Technology*, 21(6), 1773-1784.
- Qin, Z., & Yan, W. Q. (2021). Traffic-sign recognition using deep learning. *International symposium on geometry and vision*,
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1), 1-48.
- Triki, N., Karray, M., & Ksantini, M. (2024). A comprehensive survey and analysis of traffic sign recognition systems with hardware implementation. *IEEE Access*.
- Wang, C., Zheng, B., & Li, C. (2025). Efficient traffic sign recognition using YOLO for intelligent transport systems. *Scientific Reports*, 15(1), 13657.
- Zeng, X., Ouyang, W., & Wang, X. (2013). Multi-stage contextual deep learning for pedestrian detection. *Proceedings of the IEEE International Conference on Computer Vision*,

## Funding

This research did not receive any funding

## Conflict of Interest Disclosure

The author declares that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data Availability Statement

The information supporting the conclusions of this research can be obtained at the discretion of the relevant author.