Available online at www.dergipark.org.tr/en

INTERNATIONAL ADVANCED RESEARCHES and ENGINEERING JOURNAL

Journal homepage: www.dergipark.org.tr/en/pub/iarej

Internation Open Access

Volume 10 Issue 02

August, 2026

e-ISSN: 2618-575X

Research Article

Middle-East Traffic Sign (METS) recognition: databases and models

Elena Battini Sonmez [a,*], ***Hamza Arnaout*** [a] ***and Albano Younes*** [a]

[a] *Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Istanbul Bilgi UniversityIstanbul, 34060, Turkey*

ARTICLE INFO

Article history:
Received 13 Dec 2025
Accepted 16 August 2026
Published 25 August 2026

Keywords:
Arabic traffic sign recognition
Computer vision,
Convolutional neural networks
Deep learning

ABSTRACT

Automatic Traffic Sign Recognition is a crucial sub-task for autonomous driving, playing a key role in reducing accidents and enabling safer, more intelligent navigation on the roads. This challenge remains largely unresolved in the Arabic world, unlike other regions. This paper reviews the current state-of-the-art in automatic Arabic road-signs recognition; it proposes a new model, which improves the best performance of existing databases; it merges the Arabic Traffic Signs collection with the Jordanian dataset, and it uses several data augmentation techniques to create the largest publicly available database in the Arabic world to date. Finally, it gives a benchmark algorithm for the newly introduced collection. The proposed model is a variation of the lightweight Residual Net architecture, which has been improved by a significant reduction in the number of parameters, minimization of down-sampling to maintain as much spatial detail as possible, and the incorporation of both Dropout and Batch Normalization to ensure better regularization and overfitting prevention. The performance of the introduced model on the Arabic Traffic Sign database surpasses the current state of the art in both accuracy and training time. The same architecture is also very successful when tested on the newly introduced picture collection.

1. Introduction

Accurate detection and classification of traffic signs is crucial for the realization of driverless vehicles, which can be used in the service sector, as well as the tourism and health sectors. That is, the driverless option can benefit several moving vehicles, ranging from taxis, tourist buses, golf buggies, to wheelchairs.

Automatic Traffic Sign Recognition (TSR) algorithms enable vehicles to read basic road signs, such as speed limits, when there is a threat to drivers or a change in direction, making driving and the overall road environment safer. A TSR system must be able to segment the original picture and correctly classify it. Deep learning algorithms are used to develop a successful system capable of handling multiple disturbances, including variable lighting, occlusion, rotation, translation, degradation, and noise.

Although several studies have examined European traffic signs, few have examined traffic signs in Arabic-speaking regions. Moreover, to date, there are only two publicly available datasets in this domain, the Jordanian [1] and the Saudi Arabian [2] traffic sign databases, which are neither big nor challenging enough. This shortage underscores the importance of building a larger database and an effective Traffic Sign Recognition system for the Arabic world.

Arabic TSR is an essential field of study because it addresses the development of embedded systems that can interpret traffic signs commonly used in Arabic regions, while accounting for disturbances such as simulated sandstorms, a common issue in desert environments. That is, as more people become accustomed to using electric vehicles equipped with autonomous technology, Arabic TSR algorithms have become an essential component of intelligent transport systems (ITS) and advanced driver assistance systems (ADAS). The main difference between Arabic and European TS is the presence of Arabic language in some signs; also, some of the road signs themselves are different in both design and meaning. For example, in Saudi Arabia and Jordan, warning signs usually contain Arabic text for (30km/h speed limit) and (speed bump), both shown in Fig. 1, which are not used in European countries. In addition, certain signs specific to the region are used only in the desert environment, such as the crossing of camels (Fig. 1), which are not found anywhere in Europe. These differences underscore the

* Corresponding author. Tel.: *+90 212 3115434*
E-mail addresses: elena.sonmez@bilgi.edu.tr (*E.B.Sonmez*), hamzamurad92@gmail.com *(H.Arnaout)*, albano.younes1234@gmail.com *(A.Younes)*
ORCID: *0000-0003-0090-984X* (*E.B.Sonmez*)*,, 0009-0002-7350-902X (H.Arnaout), 0009-0009-9113-6985 (A.Younes)*
DOI: 10.35860/iarej.1841529
© 2026, The Author(s). This article is licensed under the CC BY-NC 4.0 International License (https://creativecommons.org/licenses/by-nc/4.0/).

need for local TSR systems that can support specific sign shapes, symbols, and languages associated with different regions.

It is also important to underline that different Arabic countries use different TS. For example, Saudi Arabia's bilingual signage system is reflected in the "STOP" sign, which typically takes the form of a red octagon with Arabic and English text. On the other hand, Jordan frequently uses the same color and shape but only Arabic text, which may cause difficulties for drivers of other nationalities and for models trained on multilingual datasets. Moreover, warning signs in Egypt often employ distinctive icons or stylized fonts that are not widely used in Gulf countries. These differences demonstrate the importance of developing local TSR systems rather than relying on generic datasets such as the German Traffic Sign Recognition Benchmark (GTSRB), which primarily represent signage in Europe. In addition, these areas often experience sandstorms, dust haze, and intense sunlight, all of which can significantly impair visual clarity and affect recognition accuracy.

This work presents a robust, lightweight system specifically designed for Arabic traffic sign recognition. It also introduces the research community to the largest database of 131,604 Arabic signs, along with its baseline performance. The introduced collection contains several disturbance elements, including light (bright, dark, white, and yellow), translation, rotation, degradation, fog, and sandstorm images. The database is already divided into five folders, which randomly group images into subdirectories (train, valid, test) to allow reproducibility of our experiments. Databases and the code can be requested by authors.

The rest of the paper is organized as follows. Section 2 discusses related work on this subject and highlights the lack of Arabic TSR systems and databases. Section 3 describes the datasets currently available in the Arabic world and introduces a new collection built ad hoc to address the issue of TSR in the Arabic region. Section 4 details the data pre-processing and augmentation methods used and introduces the successful, robust Arabic TSR system proposed in this research. Section 5 describes the experiments conducted and presents their performance and running times. Finally, Section 6 summarizes our findings and suggests future work.

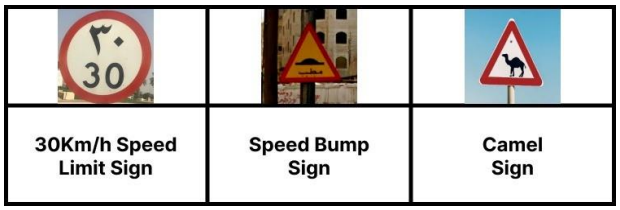


Figure 1. Samples of traffic signs specific to the Arabic region

2. Literature Review

Several classification algorithms and future extraction techniques have been used in the field of traffic sign recognition.

In 2020, Vincent et al. [3] proposed a Convolutional Neural Network (CNN) with four convolutional layers and four fully connected layers and achieved 98.44% in the German Traffic Sign Recognition Benchmark (GT- SRB) [4] database.

In 2021, Jonah and Orike [5] challenged the GTSRB database and compared the performance of ResNet50, VGG16, with a new 8-layer CNN. Their best performance reached an accuracy of 96.0%. Kerim and Efe [6] used several features, including color, Local Binary Patterns (LBP), and Histograms of Oriented Gradients (HOG), and a hybrid Artificial Neural Network (ANN) to reach the accuracy of 95% in the GTSRB database. Zaibi et al. [7] reached a performance of 99.84% on the GTSRB collection and 98.37% the Belgian Traffic Sign Data Set (BTSD) [8] using LeNet-5 model [9]. It is interesting to point out that LeNet-5 architecture was introduced by Lecun et al. in 1988, it is a lightweight model suitable for embedded applications such as ITS and ADAS.

In 2022, Abudhagir and Ashok [10] worked on the GTSRB database using LeNet-5 CNN architecture and obtained a performance of 98.50%. Bhatt et al. [11] created a hybrid database, that merges the GTSRB collection with a self-made Indian road sign database. They used data augmentation to improve the original performance of the model, which reached the top accuracy of 95.45% on the hybrid dataset, and 99.85% on the GTSRB alone. Ay et al. [12] compared the performance of several models, ranging from AlexNET, DarkNET-53, to EfficientNET-b0, in the GTSRB database. The best accuracy of 98.64% was achieved by EfficientNet. Zheng et al. [13] tested the performance of vision transformers for TSR. They worked with the German, Indian, and Chinese TS collections and proposed suggestions to improve the original transformer performance, which was not as good as that of CNNs.

In 2023, Lim et al. [14] made a comprehensive overview of the algorithms and databases for TSR. This paper compares the architecture, performance, data requirements, training time, and interpretability of machine learning (ML) and deep learning (DL) algorithms. It includes an extensive literature review that describes the features used by ML algorithms. It also provides a comprehensive survey of existing databases, including the German (GTSRB) collection, with 51,839 images belonging to 43 classes, and the Belgian (BTSD) databases, with 7,095 road signs divided into 62 classes, which are the largest and most widely used. Qu et al. [15] addressed the challenge of TSR under adverse weather conditions. This study improved the original YOLOv5 architecture by embedding traffic sign location information into the

attention channel and adding a prediction head for fine-grained detection. The same architecture, YOLOv5, was also improved by Wang et al. [16], who addressed the challenge of scaling. The proposed YOLO architecture is integrated with an improved feature pyramid model to handle this issue.

In 2024, Lu et al. [17] incorporated a multi-level squeeze feature perception module into the original YOLOv5 architecture to improve YOLO's performance for the detection of small traffic signs. In the same year, several papers have also proposed variations of the YOLO architecture, to mention a few [18] and [19].

In 2025, Tran et al. [20] introduced a novel architecture that improved traffic sign detection performance while maintaining real-time feasibility. The proposed model was evaluated on a custom dataset and compared with several baseline models, including ResNet, VGG, and MobileNet. Guo et al. [21] investigated on the performance of a lightweight adaptive robust traffic sign detection transformer.

The rest of this section focuses on papers related to the Arabic TSR.

In 2023, Latif et al. [22] compared the performance of several deep learning models on the Saudi Arabic Traffic Sign (ArTS) database. In more specific terms, the authors postulated two versions of the ResNet architecture, which were, in turn, optimized and featured as V1 (Standard), V2 (Standard), V1 (Optimized), and V2 (Optimized). Their work can be considered the gold standard in our study; we largely replicated their approach to directly compare results. Finally, we created a robust TSR system that improved their performance while running faster.

In 2024, Alqahtani et al. [23] challenged the Saudi Arabic TS database with a new architecture called MobileViTv2-MLP. Since they made a different division of (train, valid, test) sets, their performance is not comparable with the benchmark paper. However, this work is interesting because MobileViTv2-MLP is a good compromise between performance and computational complexity. This study was carried out in [24].

3. Datasets

A comprehensive overview of existing databases for TS is given in [14]. This paper focuses on the collections available for the Arabic region.

The **Jordanian Traffic Sign dataset** [1] contains traffic signs used in Jordan. Available on Kaggle, it allows researchers to become familiar with actual Middle Eastern Road signs, some of which contain symbols and texts in two languages. This collection is useful for investigations into traffic sign recognition, especially Arabic-script signs, and for developing methods applicable to TSR in the Arabic world. Figure 2 gives examples of images present in the Jordanian database. More in detail, this collection contains 9,246 images. The dataset has 7 classes, which are: No Entry, Hump, Stop, Pedestrian Cross, No Stop, Give Way, Pass Either. Table 1 displays the distribution of the images among each class.

The **Saudi Arabian** database [2] stores 2,718 field-captured images; the images belong to 24 classes, which are the most widely used road signs. The number of images has been increased using several data augmentation techniques, resulting in an augmented database of 57,078 images. The emphasis of this dataset is on attributes unique to Saudi traffic signs, including their shapes, colors, and the texts used, which mostly use Arabic. Figure 3 shows examples of traffic signs in the Saudi Arabian database. The 24 classes are the most common traffic signs used in the Arabic world; the Jordanian collection includes 3 additional classes not present in the Saudi Arabian database: No Entry, No Stop, and Give Way.

This paper introduces a new database, which has been created by merging the Jordanian and Saudi Arabic databases; we call this new collection the **Middle-East Traffic Sign (METS) database**. In more detail, the original 2,718 images of the Saudi Arabic dataset were joined with the Jordanian one, which contains 9,246 images for a total of 2,718 + 9,246 = 11,964 signs; the total number of classes of the METS database is 24+3 = 27. To establish consistency, only the original images (not previously augmented) from both datasets were used, preserving the original train, test, and validation splits. Afterward, data augmentation techniques were applied to enlarge the dataset, making it more diverse and useful for training models for ITS and ADAS systems in the Middle East. As a result, the METS dataset comprises exactly 131,604 images, making it the largest Arabic traffic sign dataset available to date. Table 2 details the distribution of the classes.



Figure 2. Jordanian database

Table 1. Distribution of images of the Jordanian Traffic Sign dataset

Class Name	Image Count	Percentage (%)
Hump	1,524	16.48%
Stop	1,443	15.61%
Give Way	1,314	14.21%
No Stop	1,305	14.11%
No Entry	1,242	13.43%
Pass Either	1,215	13.14%
Pedestrian Cross	1,203	13.01%
Total	9,246	100.00%



Figure 3. Saudi Arabian database

Table 2. Distribution of images among traffic sign classes in the METS dataset

Class Name	Image Count	Percentage (%)
front or right	1,584	1.20%
front or left	1,584	1.20%
road hump	18,128	13.77%
left turn	1,298	0.99%
narrow from right	1,133	0.86%
no u turn	1,089	0.83%
parking	1,452	1.10%
no horn	1,210	0.92%
right turn	1,144	0.87%
narrow from left	1,232	0.94%
u turn	1,386	1.05%
rotor	14,410	10.95%
slow	14,509	11.02%
stop	1,287	0.98%
no overtaking	1,122	0.85%
right or left	1,177	0.89%
pedestrian crossing	880	0.67%
no parking	1,452	1.10%
speed 30	1,254	0.95%
speed 40	1,210	0.92%
speed 50	1,199	0.91%
speed 60	1,111	0.84%
speed 80	17,160	13.04%
speed 100	1,122	0.85%
give way	13,662	10.38%
no entry	14,355	10.91%
no stop	14,454	10.98%
Total	131,604	100.00%

4. The Proposed System

The architecture of the model proposed in this research paper is a custom variation of the lightweight Residual Network (ResNet) [25]. However, the original ResNet architecture was significantly optimized for efficiency and applicability to real-world settings. It has about 750,000 parameters, which is about 15 times fewer than regular ResNets, yet it achieves competitive or higher accuracy. The proposed model reduces excessive down-sampling to maintain spatial details, which is particularly important when dealing with small or low-resolution traffic signs. Moreover, it improves regularization with Batch Normalization and Dropout, unlike most variants of ResNet, which are usually reliant on Batch Normalization alone. This custom architecture is implemented with TensorFlow and Keras and trained on Intel's 13th-gen i7-13620H CPU using different augmented and merged datasets. The result is an efficient, lightweight, and robust model that accurately identifies traffic signs under noisy or degraded visual conditions.

4.1 Data Pre-processing

To ensure consistent processing across datasets, several preprocessing steps were applied. First, all images were resized to 48×48 pixels to match the exact input shape required by the proposed architecture. In the case of the Jordanian dataset (which includes images of varying sizes and aspect ratios), central cropping was applied to remove unnecessary background and retain only the traffic sign. These pre-processing steps helped standardize the input data, reduce irrelevant visual information, and improve model performance across different experiments.

4.2 Data Augmentation

To increase dataset diversity and improve model robustness under real-world conditions, large-scale data augmentation methods were applied. These augmentations perfectly simulate the conditions prevalent in Arabic regions due to environmental disturbances, including sandstorms, light, fog, night driving, and degradation.

Lighting Variations: Four different lighting conditions were simulated:

•Brighter: Increased overall brightness to mimic harsh sunlight.

•Night: Reduced brightness with a subtle purple tint to simulate nighttime visibility of traffic signs.

•Yellow: Applied a yellow tint to represent illumination from the car headlights.

•White: Added a white overlay to simulate overexposure and glare.

Fog Simulation: Foggy conditions were simulated by adding random noise to the image, creating a hazy appearance that degrades contrast and mimics the appearance of fog.

Sandstorm Simulation: To mimic the effect of a sandstorm, a combination of Gaussian blur, slight random noise, and an orange-tinted overlay was applied. This created a visual obstruction consistent with desert driving environments.

Image Degradation: To simulate physical wear and tear on traffic signs, the following effects were introduced:

•Scratches: Brown colored scratches were drawn across the traffic sign to resemble rust or aging.

•Holes: Randomly shaped and sized regions were masked using colors sampled from the corners of the image (usually sky or background) to simulate missing or corroded parts of the sign. A light Gaussian blur was also applied to blend these effects.

4.3 Model Summary

The proposed architecture is a variation of the ResNet model, which is the best architecture used by the benchmark paper [25]. Since they did not provide the code, we implemented it from scratch; that is, we did not use a pre-trained ready model. More in detail, the proposed architecture has the following layers:

• An input layer accepting an image of size 48 × 48 × 3

• A convolutional layer with 32 filters of size 3 × 3, applied with stride 1 and 'same' padding, followed by Batch Normalization, a ReLU activation, and a MaxPooling layer with a 3 × 3 window, stride 2, and 'same' padding to reduce spatial dimensions.

•Residual Block Group 1 (32 filters) represented in Figure 4 and followed by a RELU layer

•Residual Block Group 2 (64 filters) represented in Figure 5 and followed by a RELU layer

•Residual Block Group 3 (128 filters) represented in Figure 6 and followed by a RELU layer

•Final Classification Head where a Global Average Pooling layer is used to collapse spatial dimensions, followed by a Dense layer with 256 units and ReLU activation. To improve generalization, a Dropout layer with a rate of 0.5 is applied, followed by Batch Normalization. The final output layer is a Dense layer with (number of classes) units and a softmax activation, producing the class probabilities for traffic sign classification.

This architecture includes approximately 750,000 trainable parameters, making it about 15 times smaller than usual ResNets while still achieving competitive or better performance. The model design minimizes down-sampling to preserve spatial detail and incorporates both Dropout and Batch Normalization to improve generalization and reduce overfitting.

5. Experiments and Results

The performance of the proposed model has been tested in several experiments.

5.1 Experiment 1 on the Saudi Arabic dataset

It is the same experiment conducted in the benchmark paper, Latif et al. [22], using the original Saudi dataset consisting of 2,718 images. Like Latif et al., the dataset was pre-split into 80% training and 20% testing; we then split the 80% into 75% and 25%, using the latter as the final split of 60% training, 20% testing, and 20% validation. Like in the benchmark paper, we added each original image's respective augmented versions, totaling 57,078 images. Our model achieved an accuracy of 96.68% in 1072 seconds of training time. This outperformed the former benchmark model, achieving faster and more accurate results, as detailed in Table 3.

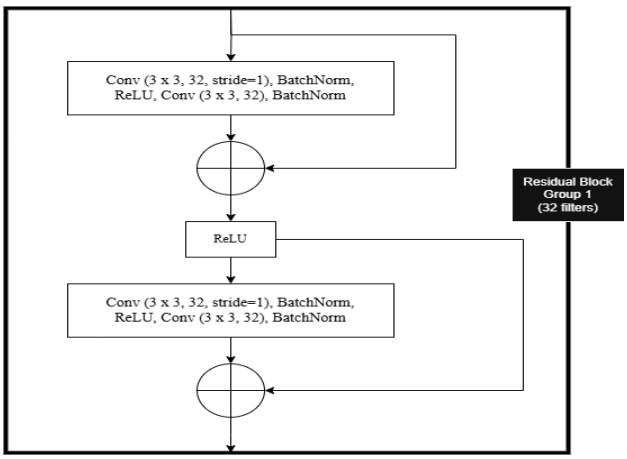


Figure 4: Residual Block Group 1

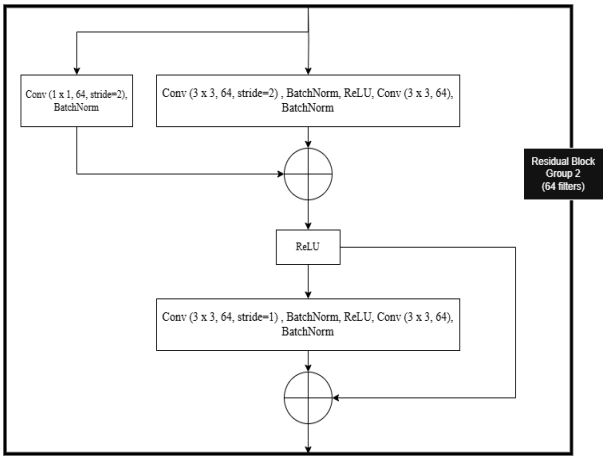


Figure 5: Residual Block Group 2

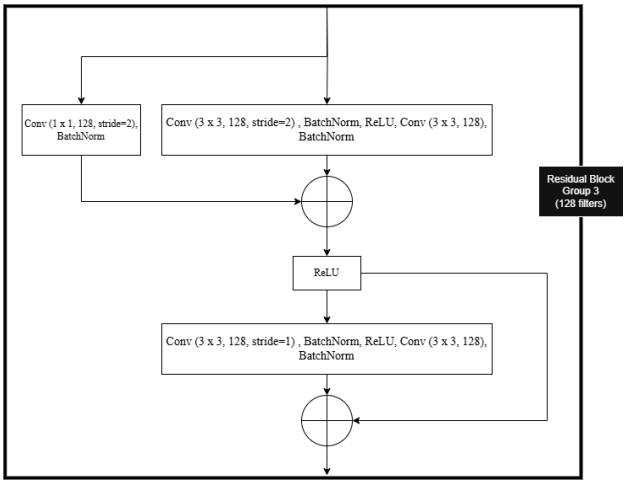


Figure 6: Residual Block Group 3

5.2 Experiment 2 on the Custom Augmented Saudi dataset

This experiment addresses the future work proposed by Latif et al. [22]. That is, the original Saudi dataset, consisting of 2,718 images, was augmented by extra transformations, including sandstorm simulation, degradation, fog, and light (Bright, Night, Yellow, White). The augmented images were returned to their respective sets with the same 60/20/20 split as in experiment 1, resulting in a dataset of exactly 76,104 images. Our architecture achieved 96.41% accuracy in 1135 seconds, demonstrating its robustness under poor and adverse visual conditions. The given metrics for the benchmark models V1 (Standard), V2 (Standard), V1 (Optimized), and V2 (Optimized) are the results of our code, which carefully reproduces the models proposed by Latif et al. All results are summarized in Table 4.

Table 3. Performance comparison between benchmark models and the proposed architecture, on the original Saudi Arabian dataset.

Version	Layers	Test Acc. (%)	Test Loss	Train Time (sec)
V1 (Standard) [22]	8	87.45	0.3162	1758
V2 (Standard) [22]	11	86.29	0.4750	1210
V1 (Optimized) [22]	56	96.14	0.2765	8425
V2 (Optimized) [22]	29	94.40	0.2757	5023
Ours	49	96.68	0.1659	1072

Table 4. Performance comparison between benchmark models and the proposed architecture, on the Custom Augmented Saudi dataset (L = number of layers)

Version	L	Test Acc. (%)	Test Loss	Train Time (sec)
V1 (Standard)	8	87.45	0.3162	1758
V2 (Standard)	11	86.29	0.4750	1210
V1 (Optimized)	56	96.14	0.2765	8425
V2 (Optimized)	29	94.40	0.2757	5023
Ours	49	96.41	0.1891	1135

5.3 Experiment 3 on the Middle East Traffic Sign (METS) dataset

This experiment merges the original Saudi collection of 2,718 images (split into 60/20/20) with the Jordanian Traffic Sign Dataset, which consists of 9,246 images and follows its own train/test/validation structure, totaling 11,964 images. Pictures in the Jordanian dataset were already augmented. The merged database was used to test generalization in different regional sign formats. The precision of our model was 94.99%, and the training time was 355 seconds. The given metrics for the benchmark models V1 (Standard), V2 (Standard), V1 (Optimized), and V2 (Optimized) are the results of our code, which carefully reproduces the models proposed by Latif et al. All results are detailed in Table 5.

5.4 Experiment 4 on the Augmented METS dataset

Our final and most comprehensive experiment involves splitting the original Saudi dataset (2,718 images) into 5 random splits and merging each split with the Jordanian dataset, yielding 11,964 images; we then applied data augmentation to each split, resulting in a total of 131,604 images. In this experiment, full data augmentation was used, including fog, sandstorm, degradation, light, rotation (left and right), and translation. Figure 7 shows the different types of data augmentations applied. That is, every original image is now augmented 10 times, adding several disturbance elements.

Table 5. Performance comparison between benchmark models and our proposed architecture on the METS database.

Version	Layers	Test Acc. (%)	Test Loss	Train Time (sec)
V1 (Standard)	8	87.45%	0.3162	1758
V2 (Standard)	11	86.29%	0.4750	1210
V1 (Optimized)	56	96.14%	0.2765	8425
V2 (Optimized)	29	94.40%	0.2757	5023
Ours	49	94.99%	0.1219	355

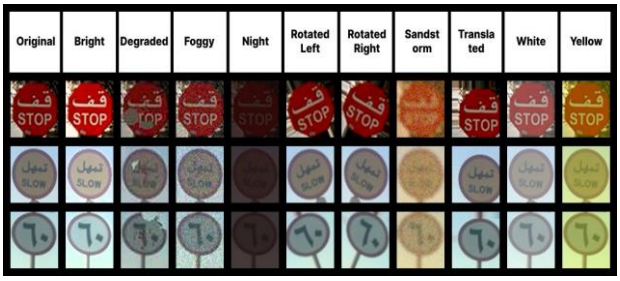


Figure 7. Custom Data Augmentation

Table 6. Performance comparison between benchmark models and our proposed architecture on the Augmented METS dataset.

Version	Layers	Test Acc. (%)	Test Loss	Train Time
V1 (Standard)	8	87.45	0.3162	1758 s
V2 (Standard)	11	86.29	0.4750	1210 s
V1 (Optimized)	56	96.14	0.2765	8425 s
V2 (Optimized)	29	94.40	0.2757	5023 s
Our Average Over 5 splits	49	96.84	0.1405	3940 s
Best Split perf	49	97.20	0.1507	4379s

Our model achieved the highest accuracy (average) of 96.84% in 3940 seconds in all experiments, while the highest achieved accuracy (Split 4 out of the 5) was 97.20% in 4379 seconds. The given metrics for the benchmark models V1 (Standard), V2 (Standard), V1 (Optimized), and V2 (Optimized) are the results of our code, which carefully reproduces the models proposed by Latif et al. All results are summarized in Table 6.

6. Conclusions

This paper introduces an effective and innovative deep learning approach to Arabic TSR, addressing the growing need for region-specific intelligent transport systems in Arabic-speaking countries. The proposed architecture is a custom variation of a lightweight Residual Network, optimized for speed, reduced parameter count, and enhanced generalization through a combination of Dropout and Batch Normalization.

In addition to developing a high-performing model, this work contributes to the largest publicly available Arabic traffic sign dataset, created by merging the Saudi Arabian and Jordanian datasets and applying extensive data augmentation. The proposed model was also validated across various configurations in four distinct experiments, achieving a top accuracy of 97.2% on a challenging dataset of exactly 131,604 images. Our results demonstrate not only high accuracy but also strong computational efficiency, outperforming standard lightweight models such as ResNet-18 across both performance and training time. These findings confirm the model's applicability for real-world TSR applications in Arabic regions and contribute to the ongoing advancement of deep learning solutions for autonomous vehicles and intelligent transport systems. As future work, we plan ablation studies to isolate the impact of Dropout, Batch Normalization, and reduced down-sampling; also, we consider addressing the class imbalance issue, and we aim to deploy the model in real time on devices such as the Raspberry Pi, enabling Arabic traffic sign recognition in embedded autonomous systems.

To ensure the systems' full applicability across the Middle East region, additional traffic sign classes specific to different countries must be incorporated, which forms an essential part of our future work. This includes, for example, signs such as camel crossing, desert road hazard warnings, and other culturally and regionally unique traffic signs not currently present in our merged Saudi–Jordanian dataset. Expanding the dataset to include these classes will enhance the model's robustness, adaptability, and accuracy in recognizing a broader range of signs encountered in real-world driving scenarios across the Arabic-speaking world.

Declaration

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author(s) also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

The code and the databases can be requested by Corresponding Author.

Author Contributions

Hamza Arnaout and Albano Younes developed the methodology and performed the analysis. Elena Battini Sonmez supervised and improved the study. All authors wrote the manuscript together.

References

1. Hwei, K., Jordanian (Arabic) traffic signs in yolo format. https://www.kaggle.com/datasets/khaledhweij/Jordanian-traffic-signs. (Last access: 15.08.2026).
2. Latif, G., J. Alghazo, D.A. Alghmgham and L. Alzubaidi, ArTS: Arabic traffic sign dataset. Mendeley Data. 2020. (Last access: 15.08.2026).
3. Vincent, M. A., K.R. Vidya and P.M. Santhosh, *Traffic Sign Classification using Deep Neural Network*. IEEE Recent Advances in Intelligent Computational Systems (RAICS). 2020. India: p. 13–17.
4. Stallkamp, J., M. Schlipsing, J. Salmen and C. Igel, *The German Traffic Sign Recognition Benchmark: A Multi-Class Classification Competition*. International Joint Conference on Neural Networks. 2021. CA, USA: p. 1453–1460.
5. Jonah, S. and S. Orike, *Traffic Sign Classification Comparison Between Various Convolution Neural Network Models*. International Journal of Scientific and Engineering Research, 2021. **12**: p. 165–171.
6. Kerim, A. and M. Efe, *Recognition of Traffic Signs with Artificial Neural Networks: A Novel Dataset and Algorithm.* International Conference on Artificial Intelligence in Information and Communication (ICAIIC). 2021. Korea (South): p. 171–176.
7. Zaibi, A., A. Ladgham and A. Sakly, *Lightweight Model for Traffic Sign Classification Based on Enhanced LeNet-5*

Network. Journal of Sensors, 2021. **1**: p. 8870529.

8. Timofte, R., K. Zimmermann and L. van Gool, *Multi-View Traffic Sign Detection, Recognition, and 3D Localization.* Ninth IEEE Computer Society Workshop on Application of Computer Vision. 2009. Utah, USA: p. 1–8.
9. Lecun, Y., L. Bottou, Y. Bengio and P. Haffner, *Gradient-based learning applied to document recognition*. Proceedings of the IEEE, 1998. **86** (11): p. 2278–2324.
10. Syed Abudhagir, U. and N. Ashok, *Highly Sensitive Deep Learning Model for Road Traffic Sign Identification*. Mathematical Statistician and Engineering Applications, 2022. **71** (4).
11. Bhatt, N., P. Laldas and V.B. Lobo, *A Real-Time Traffic Sign Detection and Recognition System on Hybrid Dataset Using CNN*. 7th International Conference on Communication and Electronics Systems (ICCES). 2022. India: p. 1354–1358.
12. Ay, G., A. Durdu and B. Nesimioglu, *Accuracy Comparison of CNN Networks on GTSRB Dataset*. Journal of Artificial Intelligence and Data Science, 2022. **2**: p. 63–68.
13. Zheng, Y. and W. Jiang, *Evaluation of Vision Transformers for Traffic Sign Classification*. Wireless Communications and Mobile Computing, 2022. **1**: p. 3041117.
14. Lim, X., C. Lee, K. Lim, T. Ong, A. Alqahtani and A. Ali, *Recent Advances in Traffic Sign Recognition: Approaches and Datasets*. Sensors (Basel) 2023. **23**(10): p. 46–74.
15. Qu, S., X. Yang, H. Zhou and Y. Xie, *Improved Yolov5-Based for Small Traffic Sign Detection Under Complex Weather*. Scientific Reports 2023. **13**: p. 16219.
16. Wang, J., Y. Chen, Z. Dong and M.Gao, *Improved YoloV5 Network for Real- Time Multi-Scale Traffic Sign Detection*. Neural Computing and Applications, 2023. **35**(10): p. 7853–7865.
17. Lu, Z., Z. Zhu, G. Li and J. Chen, *Enhancing Small Target Traffic Sign Detection with ML_SAP in YoloV5s*. Scientific Reports, 2024. **14**(1): p. 25904.
18. Luo, Y., Y. Ci, S. Jiang and X. Wei, *A Novel Lightweight Real-Time Traffic Sign Detection Method Based on an Embedded Device and YoloV8*. Real-Time Image Proc, 2024. **21**(2).
19. Han, Y., F. Wang, W. Wang, X. Zhang and X. Li, *End-Yolo: Multi-Scale Traffic Sign Detection Method in Complex Scenes*. Digital Signal Processing, 2024. **153**: p. 104615.
20. Tran, H., N. Nguyen, N. Le, T. Le and A.D. Nguyen, *Grounding Dino and Distillation-Enhanced Model for Advanced Traffic Sign Detection and Classification.* International Journal of Autonomous Vehicles, Engineering Science and Technology, 2025. **64**: p. 102028.
21. Guo, Z., Y. Wang and N. Chen, *LAR-TSDETR: a lightweight adaptive robust traffic sign detection transformer with multi-scale fusion for real-time recognition under challenging conditions.* Measurement Science and Technology, 2025. 36: p.096129.
22. Latif, G., D.A. Alghmgham, R. Maheswar, J. Alghazo, F. Sibai and M.H. Aly, *Deep Learning in Transportation: Optimized Driven Deep Residual Networks for Arabic Traffic Sign Recognition*. Alexandria Engineering Journal, 2023. **80**: p. 134–143.
23. Alqahtani, A., E. Alqaysi, W. Alsarhani, C. Lee, A. Alsharafi, M. Alzahrani and A. Almosa, *MobileVITv2-MLP: Arabic Traffic Sign Recognition with Enhanced Lightweight Vision Transformers*. In Proc. SPIE 13517, 17th International Conference on Machine Vision (ICMV). 2024. Edinburg, UK.
24. Alqahtani, A. and M. Naseer, *Enhancing Arabic traffic sign recognition through synthetic data generation using diffusion models.* In Proc. SPIE 14114, 18th International Conference on Machine Vision (ICMV). 2026. Paris, FR.
25. He K., X. Zhang, S. Ren and J. Sun, *Deep Residual Learning for Image Recognition*. In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016.