

Article

Real-Time Navigation Roads: Lightweight and Efficient Convolutional Neural Network (LE-CNN) for Arabic Traffic Sign Recognition in Intelligent Transportation Systems (ITS)

Alaa A. Khalifa ¹, Walaa M. Alayed ^{2,*} , Hesham M. Elbadawy ³  and Rowayda A. Sadek ¹

¹ Department of Information Technology, Faculty of Computers & Artificial Intelligence, Helwan University, Cairo 11795, Egypt; alaa.abdelaziz@fci.helwan.edu.eg (A.A.K.); rowayda_sadek@fci.helwan.edu.eg (R.A.S.)

² Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

³ Network Planning Department, National Telecommunication Institute (NTI), Cairo 12577, Egypt; heshamelbadawy@ieee.org

* Correspondence: wmalayed@pnu.edu.sa

Abstract: Smart cities are now embracing the new frontier of urban living, with advanced technology being used to enhance the quality of life for residents. Many of these cities have developed transportation systems that improve efficiency and sustainability, as well as quality. Integrating cutting-edge transportation technology and data-driven solutions improves safety, reduces environmental impact, optimizes traffic flow during peak hours, and reduces congestion. Intelligent transportation systems consist of many systems, one of which is traffic sign detection. This type of system utilizes many advanced techniques and technologies, such as machine learning and computer vision techniques. A variety of traffic signs, such as yield signs, stop signs, speed limits, and pedestrian crossings, are among those that the traffic sign detection system is trained to recognize and interpret. Ensuring accurate and robust traffic sign recognition is paramount for the safe deployment of self-driving cars in diverse and challenging environments like the Arab world. However, existing methods often face many challenges, such as variability in the appearance of signs, real-time processing, occlusions that can block signs, low-quality images, and others. This paper introduces an advanced Lightweight and Efficient Convolutional Neural Network (LE-CNN) architecture specifically designed for accurate and real-time Arabic traffic sign classification. The proposed LE-CNN architecture leverages the efficacy of depth-wise separable convolutions and channel pruning to achieve significant performance improvements in both speed and accuracy compared to existing models. An extensive evaluation of the LE-CNN on the Arabic traffic sign dataset that was carried out demonstrates an impressive accuracy of 96.5% while maintaining superior performance with a remarkably low inference time of 1.65 s, crucial for real-time applications in self-driving cars. It achieves high accuracy with low false positive and false negative rates, demonstrating its potential for real-world applications like autonomous driving and advanced driver-assistance systems.

Keywords: traffic signs; convolutional neural networks; ResNet



Citation: Khalifa, A.A.; Alayed, W.M.; Elbadawy, H.M.; Sadek, R.A. Real-Time Navigation Roads: Lightweight and Efficient Convolutional Neural Network (LE-CNN) for Arabic Traffic Sign Recognition in Intelligent Transportation Systems (ITS). *Appl. Sci.* **2024**, *14*, 3903. <https://doi.org/10.3390/app14093903>

Academic Editor: João M. F. Rodrigues

Received: 23 February 2024

Revised: 4 April 2024

Accepted: 28 April 2024

Published: 2 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, the idea of “smart cities” has gained a lot of attention. Smart cities can solve many of the problems that traditional cities face by combining various intelligent systems such as infrastructure, energy, transportation, and management [1].

Intelligent transportation systems (ITS) play an important and effective role in smart cities, as they have the ability to enhance the efficiency and safety of transportation in smart cities through the use of advanced technologies and data analysis techniques [2–4]. One of these techniques is traffic sign detection and classification, which enhance traffic management, decrease congestion, and increase safety [5].

Traffic sign detection and classification systems depend on computer vision techniques and machine learning algorithms, which give the ability to autonomous vehicles, connected cars, and intelligent transportation systems to have many advantages, one of which is increased road safety. By detecting and recognizing road signs, drivers can be alerted to speed limits, stop signs, pedestrian crossings, and other important traffic regulations. This approach can decrease the risk of accidents and make the road safer for drivers and pedestrians [6]. Also, through the analysis of the traffic signs, signal-control systems can dynamically adjust timing, leading to a decrease in travel time, reducing congestion, achieving efficient traffic flow, and improving transportation efficiency [7].

Hence, we introduce a Lightweight and Efficient Convolutional Neural Network (LE-CNN) architecture that makes use of computer vision and machine learning techniques to identify and detect traffic signs in real-time. By minimizing the time and effort needed to detect traffic signs, our system seeks to offer users in smart cities a seamless and effective driving experience.

The remaining contents of this paper have been arranged as follows: Section 2 provides a literature review of previous work on traffic sign detection and classification, while Section 3 covers the dataset's organization. In Section 4, the architecture of the proposed LE-CNN model is discussed. Section 5 discusses the experiment and its findings before moving on to Section 6, which provides the conclusion and recommendations for future work.

2. Literature Review

ITS have become essential to the development of effective, secure, and environmentally friendly mobility solutions. They are based on the integration of advanced sensor systems, communications networks, data analytics, and intelligent algorithms. These systems are essential for addressing the intricate problems posed by contemporary urbanization, providing creative ways to ease traffic, improve security, and lessen negative environmental effects [8]. ITS are leading the way in reshaping transport as traditional infrastructure buckles under mounting demand and urban populations rise. ITS have different models that serve different purposes, ranging from traffic prediction and congestion management to incident detection and route optimization.

Traffic sign detection can be considered a critical aspect of ITS, which has many benefits and contributions in terms of self-driving, increasing road safety, and developing traffic control systems. Traffic sign detection can have a substantial effect on road safety and traffic management. Accurate and timely detection has a critical role in informing drivers of speed limits and other regulatory requirements, which enhance overall road safety. Much research has been done on traffic sign detection using different approaches. The early research used color-based segmentation and template matching. For example, Horak et al. (2016) [9] proposed a method to detect European traffic signs using the color-based segmentation method to detect the image and then using feature-based recognition methods for categorizing the images into different classes. Their results showed the success of the model, with an accuracy of 93%. However, these models suffered from variation in lighting and complex backgrounds.

While some research has been done based on machine learning algorithms, Kuş et al. (2008) [10] proposed a method that used the Scale Invariant Feature Transform (SIFT) to extract traffic sign features from the image, then used the K-means for recognition, and finally used the support vector machine (SVM) for the classification process. This model was tested with images with different weather and light conditions, and the results showed that the model achieved an accuracy of 93%. Liu et al. (2022) [11] proposed a model called HOG_SVM that used features of the histogram of oriented gradient (HOG) and SVM networks with the German Traffic Sign Recognition Benchmark (GTSRB) dataset, and this model achieved an accuracy of 76.35%. However, these methods achieve efficient results with small datasets and require high computational costs.

Using CNN, researchers have the ability to learn features and relations between them automatically from raw images without the need to extract image features manually.

Different versions of CNN, such as AlexNet, ResNet, VGGNet, and LetNet, have been adapted to work on the problem of traffic sign detection. Zhou et al. (2018) [12] proposed an Improved VGG (IVGG) model that is a modified version of the Visual Geometry Group (VGG) model. This model takes advantage of data augmentation and transfer learning techniques, and it achieved an accuracy of 99% when trained on the GTSRB dataset; however, this model cannot classify blurry images or images with a dark background.

Simran et al. (2022) [13] proposed a CNN-based model that trained on a small dataset consisting of 25 samples, and the results showed the ability of the model to achieve an accuracy of 98.13% at the training time. Radu et al. (2020) [14] proposed a model based on the LeNet architecture that trained on the German traffic sign dataset GTSRB and achieved an accuracy of 97%. On the other side, Latif et al. (2023) [15] succeeded in the proposition of the ResNet-based model that has the ability to classify Arabic traffic signs using the Arabic Traffic Signs (ArTS) dataset with a high accuracy of 96.14%, but this accuracy was affected by the training time consumed, which reached 8425 s. After conducting this literature review, we aimed to develop a model that achieved high accuracy, taking into account reducing the time spent on the training and testing process.

3. Dataset

To our knowledge, there are two datasets for Arabic traffic signals: ArTS [16] and ATTICA [17]. A developed labeled dataset of images of Arabic traffic signs, known as ArTS, was created at Prince Mohammad Bin Fahd University in Al-Khobar, Kingdom of Saudi Arabia. Researchers can access it publicly. Applications pertaining to autonomous vehicles, traffic monitoring, and traffic safety can benefit from the dataset in various ways. Additionally, it can be helpful in creating applications for those who are blind or visually impaired and cannot read the signs. The photos were taken in Alkhobar and Dammam, two of the largest cities in Saudi Arabia's Eastern Province. The captured images were for the 24 most common Arabic traffic signs that are described in Figure 1, such as Slow, Stop, No Parking, and others.

This dataset is available in two versions. The first version contains a total of 2718 real images that were randomly divided into an 80% training set (2200 images) and a 20% test set (518 images). The second version contains a total of 57,078 images that were created by applying some augmentation techniques to the 2718 real captured images, such as resizing, flipping, scaling, zooming, cropping, and rotation, as shown in Figure 2. The augmented images were saved as RGB images in various sizes. These 57,078 images were randomly partitioned into an 80% training set (46,200 images) and a 20% testing set (10,878 images).

Sign name in English	Sign name in Arabic	Sign Image	No. of Images	Sign name in English	Sign name in Arabic	Sign Image	No. of Images
Front or Right	الاتجاه الى الامام او اليمين		144	Front or Left	الاتجاه الى الامام او اليسار		144
Slow	تمهل		107	Stop	قف		117
Left Turn	منحني الى اليسار		118	Right Turn	منحني الى اليمين		117
Narrow From Right	الطريق يضيق من اليمين		103	Narrow From Left	الطريق يضيق من اليسار		105
U Turn	الاتجاه الى الامام او الخلف		102	No U Turn	ممنوع الدوران للخلف		104
Parking	موقف		126	No Parking	ممنوع الوقوف		110
Rotor	الاتجاه المستديري		102	Right or Left	المرور يمينا او يسارا		104
Road Hump	مطب صناعي		124	No Overtaking	التجاوز محظور		112
No Horn	ممنوع التزمير		132	Pedestrian Crossing	عبور المشاه		111
Speed 30 km/hr	ممنوع تجاوز السرعة 30 كم/الساعة		132	Speed 40 km/hr	ممنوع تجاوز السرعة 40 كم/الساعة		114
Speed 50 km/hr	ممنوع تجاوز السرعة 50 كم/الساعة		110	Speed 60 km/hr	ممنوع تجاوز السرعة 60 كم/الساعة		109
Speed 80 km/hr	ممنوع تجاوز السرعة 80 كم/الساعة		101	Speed 100 km/hr	ممنوع تجاوز السرعة 100 كم/الساعة		80

Figure 1. Most common traffic signs along with the total number of photos gathered.

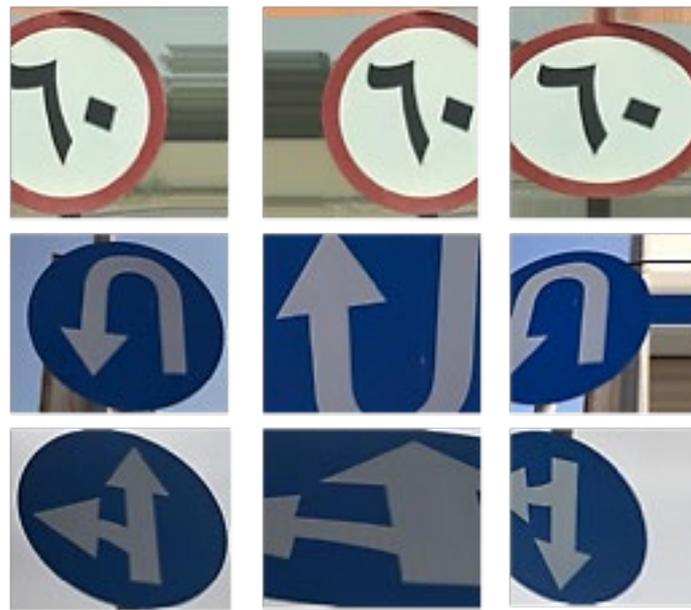


Figure 2. Samples of the traffic signs with different augmentation techniques.

4. Lightweight and Efficient Convolutional Neural Network (LE-CNN)

LE-CNN is a model that was built on the foundation of the CNN due to its ability to extract local and global features from the input images, such as edges, colors, shapes, and other features that are important to solving the target problem and that can happen through the convolutional layers [18]. Through the pooling layers, the different relationships between the image objects and the surrounding background can be learned and handled [19]. Also, the CNN model has the ability to deal with complex relationships between image features through the activation functions [20].

The LE-CNN was trained and tested using the ArTS dataset with Arabic and English traffic signs. The proposed model, as shown in Figure 3, consists of 9 layers, three of which are convolutional, three are dropout layers, two are pooling layers, and one is a flatten layer.

- **Input Layer:** Initial information entered into the model, where the pre-processed image is fed into the model. The input image is a three-channel image with input shape (32.32.3) that was captured by a camera on the road.
- **Convolutional Layers:** The core of the CNN model. It uses learnable filters such as (64 & 32) and small kernel sizes (such as 5×5 & 3×3) to extract features like edges, lines, or shapes. These features are the basics for detecting and identifying the shapes and the symbols of the traffic signs. The output of a convolutional layer Z can be calculated as follows:

$$Z = f\left(\sum_{i=1}^n (X * W_i) + b\right) \quad (1)$$

where $*$ refers to the convolution operation, n is the number of filters, f is the activation function used (ReLU), X is the input image, and W is the filter.

- **Max Pooling Layers:** Layers that are responsible for down-sampling the features maps by taking the maximum value for every small grid to reduce size and computational cost.
- **Drop Out Layer:** This layer is responsible for deactivating a certain percentage, such as 0.25, 0.25, and 0.5 (p), of neurons during training to prevent overfitting.
- **Flatten Layer:** This layer reshapes the output that comes from the previous layer from multi-dimensional into a one-dimensional vector to be then used with the dense layer.
- **Output Layer:** Typically, the ultimate layer is utilized to classify activity signs by anticipating the likelihood of them having a place in each conceivable activity sign lesson utilizing the softmax work.

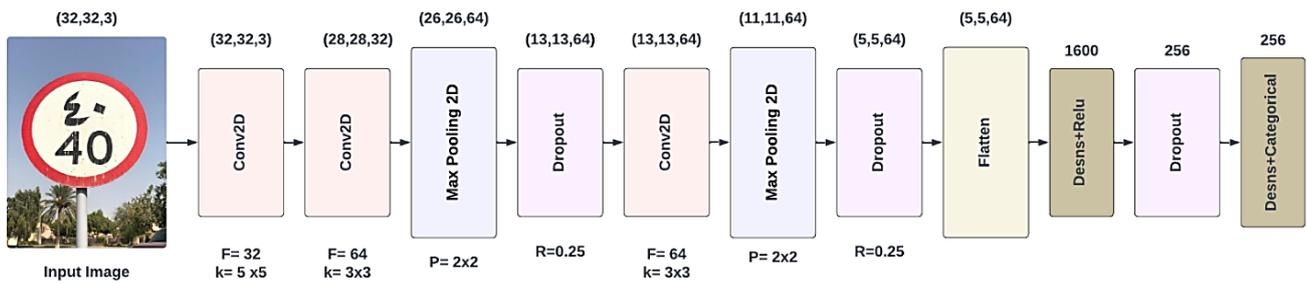


Figure 3. Proposed CNN model architecture.

5. Results

In our experiments, we use the second version of the ArTS dataset, and in order to validate and to generalize the abilities of the LE-CNN, we conducted two experiments. In the first experiment, we did not use the k-fold [21] cross validation technique, so the pre-partitioned files of the dataset were used, which are 80% training and 20% testing. In the second experiment, k-fold cross-validation with $k = 5$ was used, where all the 57,078 images were separated into five sub-groups at random, and the ratio of samples in each subset was the same for each category. Four subsets were taken as the training data in turn, and one subset as the testing data, from each training subset; 20% of the data was taken for the validation. In this section, we will present the experimental preprocessing, the setup and environment of the experiments, and the experimental results.

5.1. Experiment Preprocessing

The ArTS dataset contains images with different dimensions, and the input shape of the proposed model requires the images to be in the shape (32, 32, 3), so we needed to resize the input images to be in the shape (32, 32) and make sure they are three-dimensional images.

5.2. Experimental Setup & Environment

The experiments in this paper were conducted on CoLab [22] using Python 3, and the hyper-parameters that were used to train the model were:

- Learning rate is 0.001.
- Mini-batch size is 32.
- Optimizer is Adam [23].
- Training using one GPU.
- Epochs is 100.

5.3. Performance Evaluation Metrics of Model

The performance of the proposed model was evaluated using some metrics such as accuracy, recall, precision, F1 score, and the confusion matrix [24,25]. Accuracy measures the accuracy of recognizing traffic signs, while precision measures the proportion between the true positives to the total number of detected signs. Recall measures the ratio of true positives to the total number of ground-truthed signs, and the F1 score combines precision and recall into a single metric:

$$Accuracy = \frac{(TN + TP)}{(TP + TN + FP + FN)} \quad (2)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (3)$$

$$Sensitivity/Recall = \frac{TP}{(TP + FN)} \quad (4)$$

$$F1 - Score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{5}$$

where *TP* is the number of times the model correctly predicted the positive category, *TN* is the number of times the model correctly predicted the negative category, *FP* shows how many times the model was incorrect in its prediction of the positive category, and *FN* is the number of incorrect negative category predictions made by the model.

5.4. Experiment Results and Discussion

Figure 4 shows the confusion matrix of the LE-CNN, while Table 1 shows the different performance metrics for the LE-CNN model when trained without using the k-fold cross validation and when trained with the k-fold cross validation. The similarity of the results could be related to the small and homogeneous nature of the dataset.

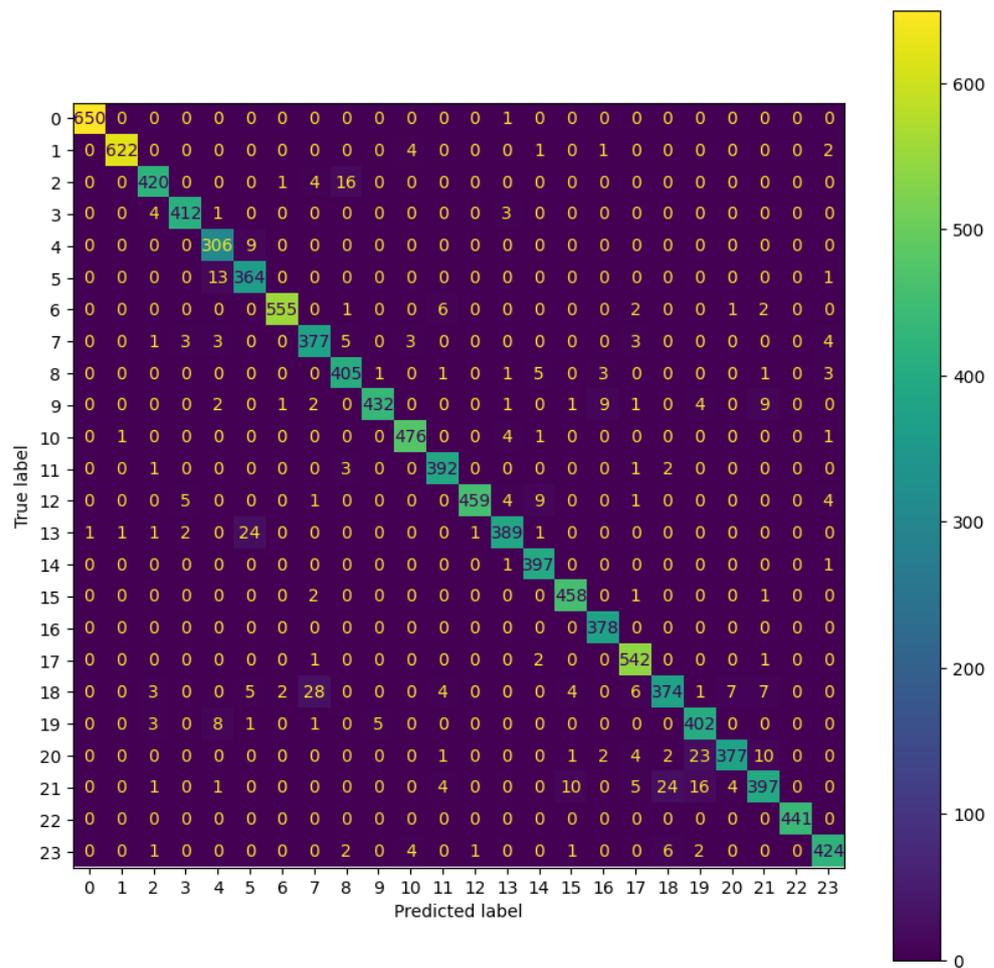


Figure 4. Confusion matrix of the LE-CNN model.

Table 1. Performance metrics for the LE-CNN model.

Metrics	Without Using K-Fold	Using K-Fold
Accuracy	96.5%	96.5%
Precision	96.66%	96.66%
Recall	96.83%	96.83%
F1-Score	96.74%	96.74%

A comparison of the experiment training and validation accuracies is presented in Figure 5. The fact that both values are nearly equal to each other suggests that the model fits the data well and exhibits neither overfitting nor underfitting.

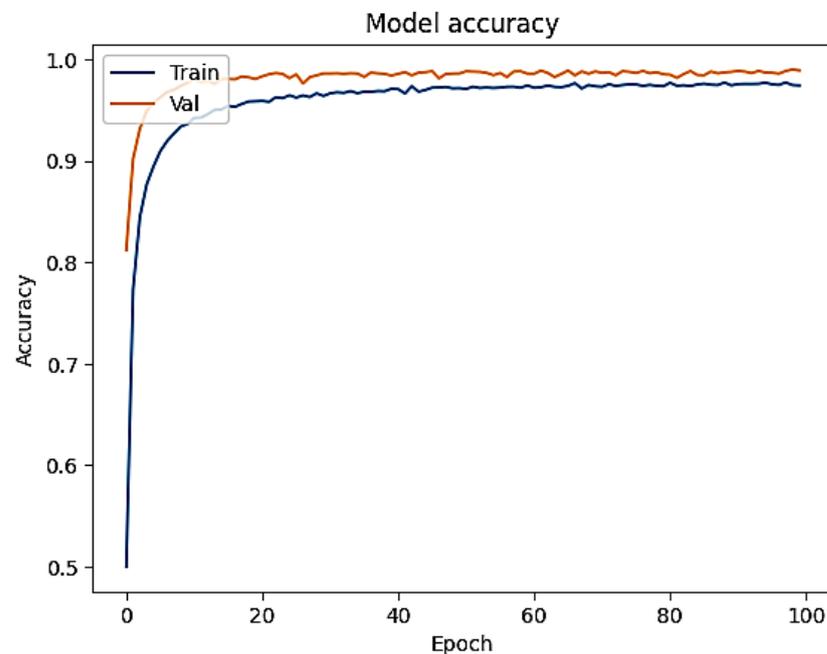


Figure 5. Comparison of the proposed LE-CNN model's training and validation accuracy.

After that, the training and validation losses for the proposed LE-CNN model with different epochs were computed and are displayed in Figure 6. The fact that both values are convergent towards zero suggests that there is no overfitting or underfitting in the model. This also shows how well the model can be used as a predictive tool in general.

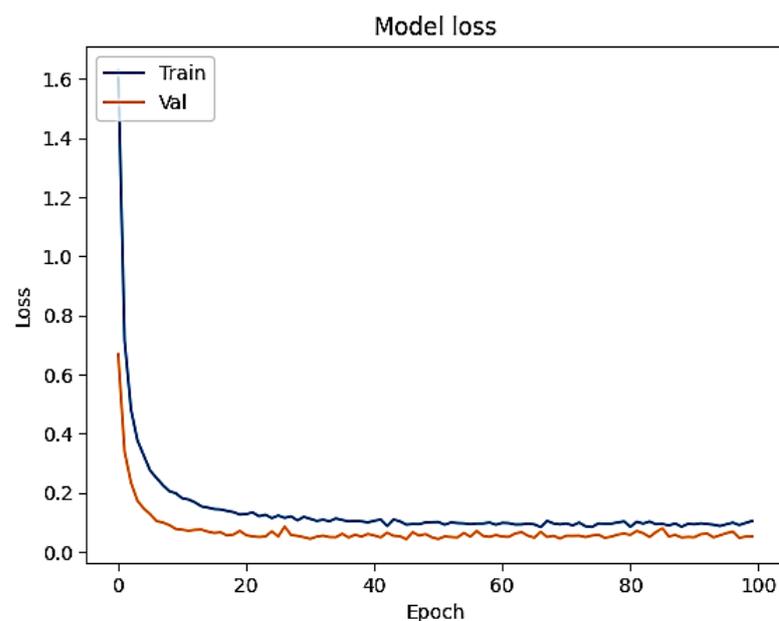


Figure 6. Comparison of the proposed LE-CNN model's training and validation loss.

Finally, a comparison was made between our proposed LE-CNN model and the ResNet, optimized ResNet, and ViT transformers [15]. As shown in Table 2, the LE-CNN model performed with an impressive accuracy of 96.5% while maintaining superior perfor-

mance with a remarkably low inference time of 1.65 s. Also, LE-CNN was able to obtain a verification time less than the time required for the ResNet and the optimized ResNet v2 model, which was 719 s when trained without the k-fold verification technique, while LE-CNN was able to obtain a verification time less than the time required for the optimized ResNet v2 model, which was 3595 s when trained with the k-fold verification technique. On the other hand, the VIT transformer succeeded in achieving faster validation and testing times than the LE-CNN model, but this was at the expense of transformer accuracy.

Table 2. Comparison between the proposed LE-CNN model and other recent models.

Model	Layers	Val. Acc.	Test Acc.	Val. Time without K-Fold	Val. Time with k-Fold	Testing Time
ResNet v1, 2023 [15]	8	95.73%	87.45%	1758 s	-	-
Optimized ResNet v1, 2023 [15]	56	99.18%	96.14%	8425 s	-	-
VIT Transformer [26]	12	79.19%	73.63%	271 s	-	0.04 s
LE-CNN	9	99.4%	96.5%	719 s	3595 s	1.65 s

6. Conclusions

This paper has shown the importance and advantages of using traffic sign detection and classification in ITS in smart cities through different techniques and technologies, especially computer vision and machine learning techniques. This paper presents a novel LE-CNN architecture that combines high accuracy with real-time processing capabilities. This efficient and robust model paves the way for reliable, secure, and efficient traffic sign recognition in self-driving cars, contributing to safer and more autonomous navigation across diverse regions like the Arabic world. Our findings demonstrated that the provided LE-CNN model was highly accurate at classifying the traffic signs while outperforming optimized ResNet v2 in this regard, scoring a test accuracy of 96.5% with a training time of 719 s as opposed to 96.14% with a training time of 8425 s.

In conclusion, in this paper, we have proposed a system that has the potential to alleviate the problems associated with traffic sign classification, which ensures the reduction in traffic congestion to enhance user satisfaction. Considering that latency in applications such as self-driving cars is highly important, more work is required to consider the process of capturing the sensory image, framing, and pre-processing. Also, collision avoidance timing and distance requirements in relation to vehicle speed on highways versus on crowded streets need to be considered.

Author Contributions: Conceptualization, A.A.K., W.M.A., H.M.E. and R.A.S.; methodology, A.A.K. and R.A.S.; software, A.A.K.; validation, W.M.A., H.M.E. and R.A.S.; writing—original draft preparation, A.A.K.; writing—review and editing, W.M.A., H.M.E. and R.A.S.; project administration, H.M.E. and R.A.S.; funding acquisition, W.M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Princess Nourah Bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R500), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are openly available at <https://data.mendeley.com/datasets/4tznkn45mx/1> (accessed on 29 April 2024).

Acknowledgments: The authors would like to acknowledge Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R500), Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia for supporting this project.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Moura, F.; De Abreu, E.; Silva, J. Smart Cities: Definitions, evolution of the concept and examples of initiatives. In *Industry, Innovation and Infrastructure*; Encyclopedia of the UN Sustainable Development Goals; Springer: Cham, Switzerland, 2019; pp. 1–9. [CrossRef]
2. Sadek, R.; Khalifa, A. Vision Transformer-Based Intelligent Parking System for Smart Cities. In Proceedings of the 20th International ACS/IEEE International Conference on Computer Systems and Applications AICSSA 2023, Giza, Egypt, 4–7 December 2023. Intelligent and Sustainable Vehicle Networking (ISVN) Workshop.
3. Radi, W.; El-Badawy, H.; Mudassir, A.; Kamel, H. Traffic Accident Management System for Intelligent and Sustainable Vehicle Networking. In Proceedings of the 20th International ACS/IEEE International Conference on Computer Systems and Applications AICSSA 2023, Giza, Egypt, 4–7 December 2023. Intelligent and Sustainable Vehicle Networking (ISVN) Workshop.
4. Radi, W.; El-Badawy, H. Enhanced Implementation of Intelligent Transportation Systems (ITS) based on Machine Learning Approaches. In Proceedings of the 20th International ACS/IEEE International Conference on Computer Systems and Applications AICSSA 2023, Giza, Egypt, 4–7 December 2023. Intelligent and Sustainable Vehicle Networking (ISVN) Workshop.
5. Iqbal, K.; Khan, M.A.; Abbas, S.; Hasan, M.Z.; Fatima, A. Intelligent Transportation System (ITS) for Smart-Cities using Mamdani Fuzzy Inference System. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*, 94. [CrossRef]
6. Abougarair, A.J.; Elmaryul, M.; Aburakhis, M. Real time traffic sign detection and recognition for autonomous vehicle. *Int. Robot. Autom. J.* **2022**, *8*, 82–87. [CrossRef]
7. Balu, V. Smart traffic light control system. *Int. J. Adv. Res. Eng. Technol.* **2020**, *11*, 20542059.
8. Pagano, P. *Intelligent Transportation Systems: From Good Practices to Standards*; CRC Press: Boca Raton, FL, USA, 2021.
9. Horák, K.; Číp, P.; Davídek, D. Automatic traffic sign detection and recognition using colour segmentation and shape identification. *MATEC Web Conf.* **2016**, *68*, 17002. [CrossRef]
10. Kuş, M.C.; Gokmen, M.; Etaner-Uyar, A.Ş. Traffic sign recognition using Scale Invariant Feature Transform and color classification. In Proceedings of the 2008 23rd International Symposium on Computer and Information Sciences, Istanbul, Turkey, 27–29 October 2008. [CrossRef]
11. Liu, Y.; Zhong, W. A novel SVM network using HOG feature for prohibition traffic sign recognition. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 6942940. [CrossRef]
12. Zhou, S.; Liang, W.; Li, J.; Kim, J. Improved VGG model for road traffic sign recognition. *Comput. Mater. Contin.* **2018**, *57*, 11–24. [CrossRef]
13. Simran; Sristi, T.; Shilpi, K.; Radhey, S. Detection of Traffic Sign Using CNN. *Recent Trends Parallel Comput.* **2022**, *9*, 14. [CrossRef]
14. Radu, M.; Costea, I.M.; Stan, V.A. Automatic Traffic Sign Recognition Artificial Intelligence—Deep Learning Algorithm. In Proceedings of the 2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Bucharest, Romania, 25–27 June 2020. [CrossRef]
15. Latif, G.; Alghmgham, D.A.; Maheswar, R.; Alghazo, J.; Sibai, F.N.; Aly, M.H. Deep learning in Transportation: Optimized driven deep residual networks for Arabic traffic sign recognition. *Alex. Eng. J.* **2023**, *80*, 134–143. [CrossRef]
16. Latif, G.; Alghazo, J.; Alghmgham, D.A.; Alzubaidi, L. ArTS: Arabic Traffic Sign Dataset. *Mendeley Data V1* **2020**. [CrossRef]
17. Boujemaa, K.S.; Akallouch, M.; Berrada, I.; Fardousse, K.; Bouhoute, A. ATTICA: A dataset for Arabic Text-Based traffic panels detection. *IEEE Access* **2021**, *9*, 93937–93947. [CrossRef]
18. Ghosh, A.; Sufian, A.; Sultana, F.; Chakrabarti, A.; De, D. Fundamental Concepts of Convolutional Neural Network. *Intell. Syst. Ref. Libr.* **2019**, *172*, 519–567. [CrossRef]
19. Bera, S.; Shrivastava, V.K. Effect of pooling strategy on convolutional neural network for classification of hyperspectral remote sensing images. *IET Image Process.* **2020**, *14*, 480–486. [CrossRef]
20. Onwujekweg, G.; Yoon, V. Analyzing the Impacts of Activation Functions on the Performance of Convolutional Neural Network Models. AI and Semantic Technologies for Intelligent Information Systems (SIGODIS). 2020. Available online: https://aisel.aisnet.org/amcis2020/ai_semantic_for_intelligent_info_systems/ai_semantic_for_intelligent_info_systems/4/ (accessed on 29 April 2024).
21. Anguita, D.; Ghelardoni, L.; Ghio, A.; Oneto, L.; Ridella, S. The ‘K’ in K-fold Cross Validation. *Eur. Symp. Artif. Neural Netw.* **2012**, *102*, 441–446.
22. Google Colaboratory. Available online: <https://colab.research.google.com/> (accessed on 21 March 2024).
23. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980. [CrossRef]
24. Sokolova, M.; Japkowicz, N.; Szpakowicz, S. Beyond accuracy, F-Score and ROC: A family of discriminant measures for performance evaluation. In Proceedings of the Australasian Joint Conference on Artificial Intelligence, Hobart, Australia, 4–8 December 2006; Lecture Notes in Computer Science. pp. 1015–1021. [CrossRef]

25. Vakili, M.; Ghamsari, M.; Rezaei, M.R. Performance analysis and comparison of machine and deep learning algorithms for IoT data classification. *arXiv* 2020, arXiv:2001.09636. Available online: <http://arxiv.org/pdf/2001.09636.pdf> (accessed on 29 April 2024).
26. Farzipour, A.; Manzari, O.N.; Shokouhi, S.B. Traffic Sign Recognition Using Local Vision Transformer. In Proceedings of the 2023 13th International Conference on Computer and Knowledge Engineering (ICCKE), Mashhad, Iran, 1–2 November 2023. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.