Alexandria Engineering Journal 80 (2023) 134–143

Contents lists available at ScienceDirect

Alexandria Engineering Journal

journal homepage: www.elsevier.com/locate/aej

Original Article

Deep learning in Transportation: Optimized driven deep residual networks for Arabic traffic sign recognition

Ghazanfar Latif [a,*], Danyah Adel Alghmgham [a], R. Maheswar [b], Jaafar Alghazo [c], Fadi Sibai [d], Moustafa H. Aly [e,*]

[a] *Computer Science Department, Prince Mohammad bin Fahd University, Khobar, Saudi Arabia*
[b] *Department of Applied Cybernetics, Faculty of Science, University of Hradec Kralove, Czech Republic*
[c] *Artificial Intelligence Research Initiative, College of Engineering and Mines, University of North Dakota, Grand Forks, ND 58202, USA*
[d] *College of Engineering and Architecture, Gulf University for Science and Technology, Mubarak Al-Abdullah, 32093, Kuwait*
[e] *Electronics and Communication Engineering Department, College of Engineering and Technology, Arab Academy for Science, Technology and Maritime Transport, Alexandria, P.O.B. 1029, Egypt*

A R T I C L E I N F O

Keywords:
Deep Learning Internet of Things (IoT)
Residual Neural Networks (ResNet)
Arabic Traffic Sign (ArTS)
Smart Devices
Convolutional Neural Networks (CNN)

A B S T R A C T

Car manufacturers around the globe are in a race to design and build driverless cars. The concept of driverless is also being applied to any moving vehicle such as wheelchairs, golf cars, tourism carts in recreational parks, etc. To achieve this ambition, vehicles must be able to drive safely on streets stay within required lanes, sense moving objects, sense obstacles, and be able to read traffic signs that are permanent and even temporary signs. It will be a completely integrated system of the Internet of Things (IoT), Global Positioning System (GPS), Machine Learning (ML)/Deep Learning (DL), and Smart Technologies. A lot of work has been done on traffic sign recognition in the English language, but little has been done for Arabic traffic sign recognition. The concepts used for traffic sign recognition can also be applied to indoor signage, smart cities, supermarket labels, and others. In this paper, we propose two optimized Residual Network (ResNet) models (ResNet V1 and ResNet V2) for automatic traffic sign recognition using the Arabic Traffic Signs (ArTS) dataset. Additionally, the authors developed a new dataset specifically for Arabic Traffic Sign recognition consisting of 2,718 images taken from random places in the Eastern province of Saudi Arabia. The optimized proposed ResNet V1 model achieved the highest training and validation accuracies of 99.18% and 96.14%, respectively. It should be noted here that the authors accounted for both overfitting and underfitting in the proposed models. It is also important to note that the results achieved using the proposed models outperform similar methods proposed in the extant literature for the same dataset or similar-size dataset.

1. Introduction

Traffic Sign Detection (TSD) is the detection of traffic signs. Traffic Sign Recognition (TSR) goes a step beyond TSD by detecting the signs, then recognizing and therefore interpreting their meaning. Recent developments and interest in self-driving cars have greatly increased interest in these fields. Automated systems can navigate through traffic on both open roads and intersections. For instance, traffic sign detection in Mazda cars uses a camera on top of the windshield to detect and display road signs related to the speed limit, stopping, and no entry. It displays these signs on the driver's dashboard to warn the driver of these road instructions. The TSR helps in avoiding the driver from getting into costly accidents or traffic violation fees, particularly when the driver is distracted or tired with significantly lower attention capability. TSR is useful in busy or dangerous road situations when the driver has to keep his eyes on the road ahead and not on the road signs. The TSR in smart cars can avoid traffic fatalities by automatically lowering speed and warning drivers of no-overtaking signs, slippery road signs, and maximum speed limits. Besides displaying warnings on dashboards, smart cars can orally announce the warning, sound special musical

* Corresponding authors at: Department of Electronics and Communications, College of Engineering and Technology, Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt (G. Latif); Computer Science Department, Prince Mohammad bin Fahd University, Khobar, Saudi Arabia (M H. Aly).

E-mail addresses: glatif@pmu.edu.sa (G. Latif), engdania.ce@gmail.com (D.A. Alghmgham), maheshh3@rediffmail.com (R. Maheswar), jaafar.alghazo@und.edu (J. Alghazo), sibai.f@gust.edu.kw (F. Sibai), drmosaly@gmail.com, mosaly@aast.edu (M.H. Aly).

https://doi.org/10.1016/j.aej.2023.08.047
Received 8 July 2022; Received in revised form 12 July 2023; Accepted 16 August 2023
Available online 24 August 2023

notes, recommend the driver to stop to rest or consume a caffeine-rich drink, spray special scents to raise the driver's attention, or recommend that the driver takes control of the car and disables the cruise control driving under dangerous road conditions and situations.

Both TSD and TSR use cameras and video or image analysis as done, for instance, in automatic license plate recognition systems [1,2]. TSD and TSR challenges are poor visibility during fog, rain, or time of the day when sunrays obscure vision and worn out or damaged signs. In [3], paper list some of the potential difficulties of traffic sign detection, the road video quality and the minimization of false alarms in keeping the driver hooked on the automated TSR are very important.

Traffic signs differ by shape and color and generally fall into five main classes: warning signs usually shaped as red triangles, prohibition signs usually shaped as red circular signs, reservation signs usually shaped as blue rectangular signs, mandatory signs shaped as blue circles, and temporary signs appearing as yellow triangular signs. The main difference between Arabic traffic signs and other traffic signs is in the instructions written in Arabic within the signs and the use of Indian numbers rather than Arabic numbers. Additionally, in some cases, both Arabic commands and their equivalent English translations appear within the signs. Also, in some cases, both Arabic and Indian numbers are displayed within the same sign. Some Arabic letters have dots above them or below them, which, when missed, change the letter and, therefore, the meaning. This means that these dots are part of the meaning and should not be missed or omitted by the TSR system.

The cost-saving and life-saving benefits of developing an Arabic TSR system are apparent. The traffic sign recognition system is gaining momentum in recent years as one of its main applications will be in autonomous vehicles. The technology of autonomous vehicles in fact has made recent advances but still lacks the optimal universal recognition models not only for traffic signs but also for the environment of the autonomous vehicles. Therefore, this being the main application will require researchers to continuously develop models for traffic sign recognition models that could be put to practical use in such applications. It should be noted that traffic sign recognition has many other applications which may include the use by pedestrians who are visually impaired. The visually impaired have an advanced walking stick and may have a guide dog but would still need to be informed of the traffic signs orally. In addition, the same technology can be integrated for the visually impaired for other signage recognition. The motivation for such work, as presented in this paper, is thus apparent and could be useful for many applications and has increased benefits which we will not detail here. Whether applied to autonomous vehicles or visually impaired applications, researchers from various countries and languages should work on the traffic sign from the various languages because the applications are not bound by geographical location. The system can also be helpful in alerting drivers to certain signage that they may have overlooked to avoid accidents.

The main objective of this work is to develop a modified model for classifying and recognizing Arabic Traffic sign images. The ultimate goal is to develop a model that can be applied in practical applications and useful to mankind.

Deep learning is used to develop an automatic Arabic TSR system with the capability of recognizing the traffic sign image through Deep ResNet networks. Car accidents occur mostly due to human error by drivers who either do not observe a sign visually or they observe the sign but do not follow the directions set by the sign. For example, a certain sign may set the speed at a maximum of 120 km/hr while the driver exceeds the maximum speed. Thus, the proposed system is developed to automatically recognize Arabic traffic signs with improved robustness and efficiency and is equipped with a smart system that can communicate to the driver certain instructions to avoid accidents. It is of the utmost importance that the recognition process and classification be done with minimal errors.

This research paper is organized such that Section 2 details the latest literature review, Section 3 contains the description of the ResNet, Section 4 shows the details of the proposed system, Section 5 contains the description of the dataset and experimental setup, and Section 6 displays and discusses the results. The conclusion and future work are highlighted in Section 7, followed by the list of all references.

2. Literature review

Traffic sign detection and recognition (TSDR) is an important topic for automating driving or providing assistance to drivers. Based on image processing, TSDR detects an image and analyzes its specific characterizing features to recognize a specific traffic sign. Visuals or audio can then alert the driver or control the self-driving vehicle. As such, it greatly enhances the quality of driving and relieves pressure off the driver, in particular in situations the driver's eyes must stay on the road, and not to the side of the road where the traffic signs are to prepare for anticipated difficult driving situations. It can also help when the driver's attention or alertness drops, at times of great distraction, or when visibility is poor such as the sun rays hit the driver's eyes.

In [3], the authors overviewed traffic sign detection and classification methods, including deep learning methods. In [4], the authors offered a review of the vision-based system for traffic sign recognition and classification.

Surveyed artificial neural network methods scored the highest accuracies. In particular, deep learning methods such as convolutional neural networks (CNN) scored even higher. Some of the traffic sign detection and recognition system challenges are variable lighting conditions, fading, blurring, and visibility due, for instance, to fog or rain, and obscured signs. Wali et al. advised that merging detection and classification under one step and using large image databases may be two tips for improving accuracies. The same authors employed RGB color segmentation with a Support Vector Machine (SVM) classifier, obtaining 95.71% accuracy [5]. Tabernik and Skocaj used deep learning with CNN on 200 traffic sign categories to achieve an error rate below 3% [6]. A traffic sign detection and recognition system for Indian roads used an RGB color saliency algorithm to distinguish the sign from other surrounding objects, morphological filters to capture the sign shape, and nearest-neighbor matching compares the extracted features from those stored in a database. This system achieved 98.66% accuracy during the day and 97.83% accuracy at night [7]. Image processing techniques are used for detecting the sign and a set of CNNs for the recognition of the traffic sign achieving a 98.11% accuracy for triangular signs and 99.18% for circular signs in [8]. Boujemma et al. explored the same problem using two ways: a color segmentation method with CNN, and a fast region-based CNN [9]. The authors achieved 93%-95% accuracy in the first method, while in the second method, they achieved 94.8% accuracy.

Farhat et al. [10] developed an algorithm for traffic sign detection and recognition, using the Maximally Stable Extremal Regions (MSER) method for extracting regions of interest and the Oriented FAST and Rotated BRIEF features for recognizing the signs. This algorithm was implemented on the Xilinx Zynq platform. It was able to identify traffic sign shapes using the MSER method with 94% recall and 95% precision rates and a mean recognition accuracy of 97%. Lim et al. [11] presented a GPU-based real-time traffic sign detection and recognition method that is robust against illumination changes and performs region detecting and recognition using a hierarchical model. This proposed method achieved an F1-score of 0.97 on the chosen dataset. Chung et al. [12] proposed an attention-based convolutional pooling neural network employing attention mechanism to feature maps to obtain key features and convolutional pooling to improve recognition accuracy in harsh environments. On the German traffic sign recognition benchmark with various noises, this method achieved 66.981% – 83.198% accuracies with CNN networks. Zhu et al. [13] used two deep learning networks, a fully convolutional network for identifying the coarse regions of traffic signs extracted by EdgeBox, and a deep CNN for object classification, sign detection, and recognition. Using the Swedish Traffic Signs Dataset

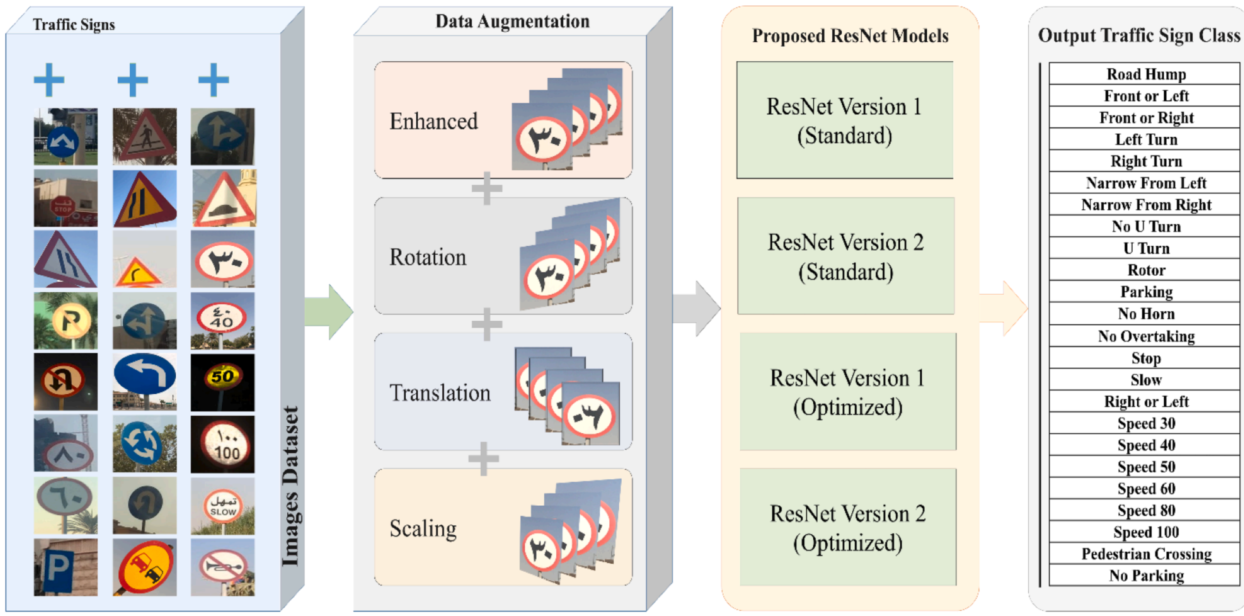

Fig. 1. Workflow of the methodology for traffic sign recognition.

(STSD), this method achieved 97.69% precision and 92.9 % recall.

Swathi and Suresh employed the multilayer perceptron trained with a histogram of oriented gradient features, achieving 97.14% and 95.71% recognition accuracies on the circular shape and triangular shape in real-time videos, respectively [14]. Dhar et al. also used CNNs to classify Bangladeshi signs after filtering grayscale versions of the images with Gabor wavelets, emphasizing the edges, extracting regions of interest using SVM, achieving 97% classification accuracy, and much lower accuracies with SVM, decision trees, KNN, and ensembles [15]. Fleyeh et al. reported the various techniques used for TSDR from ART1, ART2, Hopfield, Cellular neural networks, and fuzzy sets and classifiers [16]. Jingwen Feng [17] built a TSDR system using Histogram Orientation Gradients (HOG) and SVM and then replaced the latter 2 with a Maximally Stable Extremal Region (MSER) method and GPS method. The average cost of the HOG-based system was 2.4 times the cost of the MSER-based system. The GPS method was measured to be 3.8 times faster than the MSER method. Loy and Barnes [18] extended the fast radial symmetry transform to detect polygons at any orientation achieving 95% accuracy. Li et al. [19] combined color invariants image segmentation and a pyramid of HOGs for matching the sign shapes, and SVM to classify the images.

Another study [20] used HOG for detecting images and CNN for classifying them based on German and Italian traffic sign data scoring above 93% accuracy. Møgelmose et al. overviewed accuracy rates of 4 TSDR approaches using SVM classifiers, CNN, and K-d trees and random forests, ranging between 96% and 99%. The ones with CNN scored the

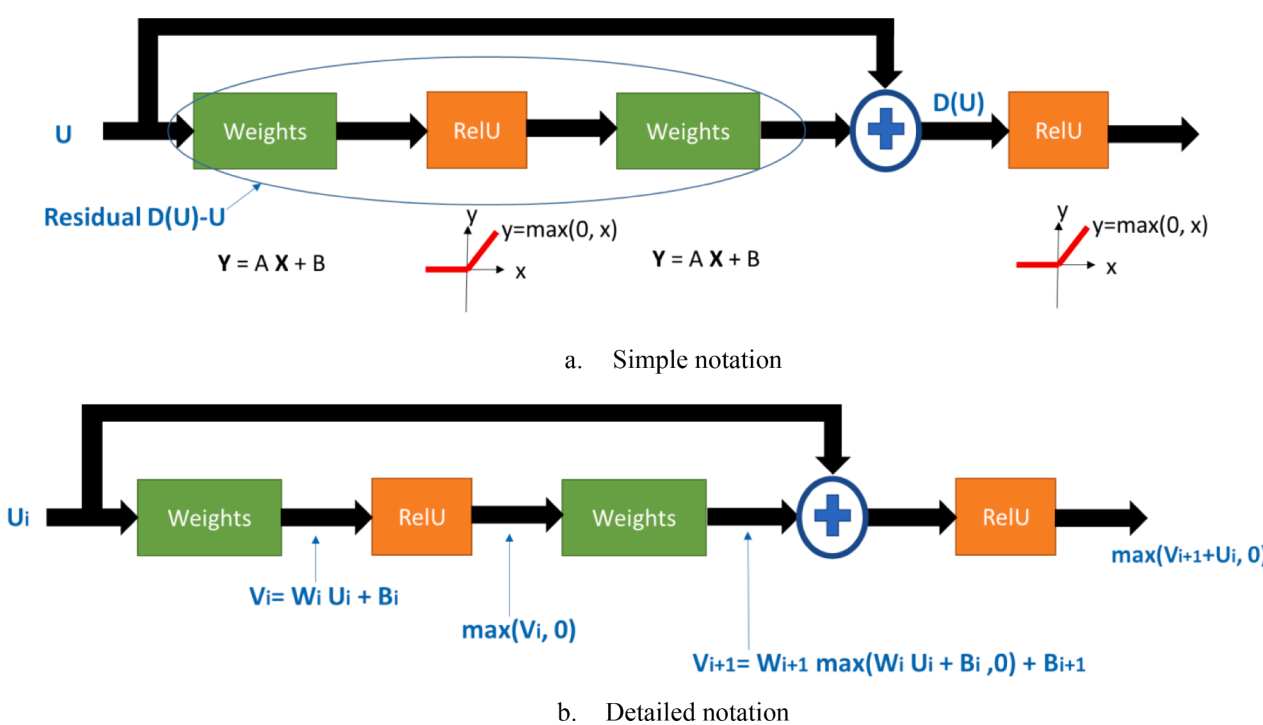

Fig. 2. ResNet block diagram.

Fig. 3. ResNet architecture.

highest on accuracy [21].

The TSDR studies rely on image databases for training and testing their systems. Among the popular TSDR image databases available online are the German Traffic Sign Recognition Benchmark (GTSRB) and the German Traffic Sign Detection Benchmark (GTSDB) [22], Chinese Traffic Sign Database [23], Tsinghua [24], Belgian [25], UK [26], and LISA [27] databases.

In this context, it should be noted that researchers are developing ML and DL-based systems for other aspects important to humans. The idea is to automate tasks that can be done by trained humans with the added benefit that the automation reduces cost and eliminates some fallbacks such as human errors. Several examples can be seen in exploring ML and DL methods in the field of medicine [28], agriculture [29], banking & education [30], geography [31], special education [32], and many others. DL and ML systems are being explored in all fields of science and knowledge and the future smart cities are envisioned to be fully automated.

3. Proposed methodology

The proposed methodology in this work is illustrated in Fig. 1. The first phase involves preprocessing the traffic dataset to prepare it for the deep learning models. The dataset is then augmented using techniques such as enhancement, rotation, translation, and scaling to address the class size and complexity. This results in an augmented dataset consisting of 57,078 images, which serves as the input to four deep learning algorithms. These algorithms, referred to as ResNet 1 (Standard), ResNet 2 (Standard), ResNet 1 (Optimized), and ResNet 2 (Optimized), in this

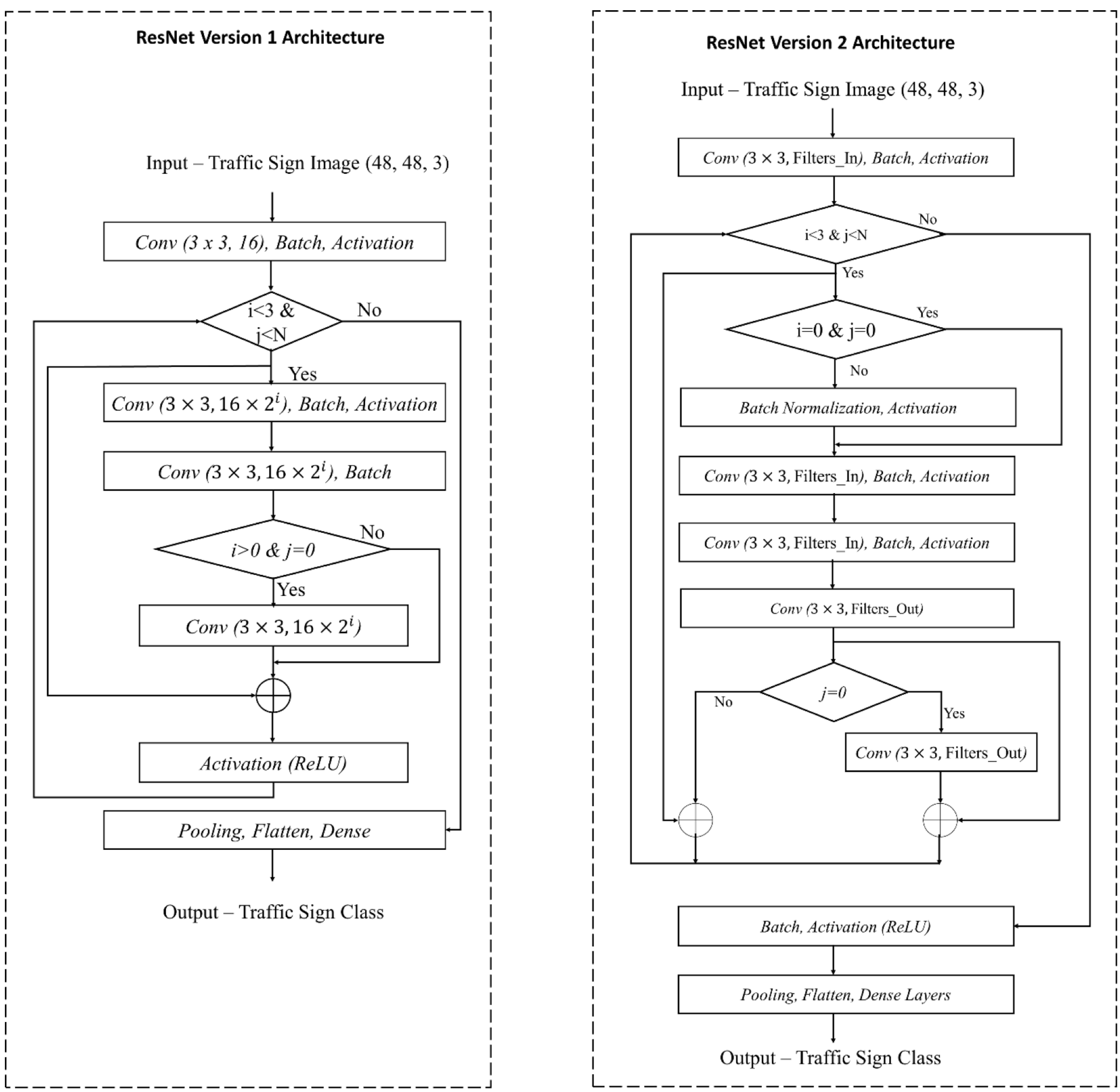

Fig. 4. Comparison of the ResNet V1 and ResNet V2 architectures.

#	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
1	Front or Right	الإتجاه الى الأمام او اليمين		144	13	Slow	تمهل		107
2	Front or Left	الإتجاه الى الأمام او اليسار		144	14	Stop	قف		117
3	Road Hump	مطب صناعي		124	15	No Overtaking	التجاوز محظور		112
4	Left Turn	منحنى الى اليسار		118	16	Right or Left	المرور على احد جانبي الطريق		104
5	Narrow From Right	الطريق يضيق من اليمين		103	17	Pedestrian crossing	عبور المشاة		111
6	No U Turn	ممنوع الدوران للخلف		104	18	No Parking	ممنوع الوقوف		110
7	Parking	مواقف		126	19	Speed 30	ممنوع تجاوز السرعة 30 في الساعة		132
8	No Horn	ممنوع التزمير		132	20	Speed 40	ممنوع تجاوز السرعة 40 في الساعة		114
9	Right Turn	منحنى الى اليمين		117	21	Speed 50	ممنوع تجاوز السرعة 50 في الساعة		110
10	Narrow From Left	الطريق يضيق من اليسار		105	22	Speed 60	ممنوع تجاوز السرعة 60 في الساعة		109
11	U Turn	الإتجاه الى الأمام او الخلف		102	23	Speed 80	ممنوع تجاوز السرعة 80 في الساعة		101
12	Rotor	الإتجاه مستدير		102	24	Speed 100	ممنوع تجاوز السرعة 100 في الساعة		80

Fig. 5. List of most common traffic signs with their total number of collected images [29].

paper, have a detailed architecture presented below. The output of the models is the classification of traffic signs into various classes, including Road Hump, Left turn, Right turn, and others.

It should be noted here that the methodology in Fig. 1 in itself is not a contribution; however, the main contributions of this paper are the two optimized deep learning models, the dataset, and the augmented dataset.

3.1. Residual neural network (ResNet) proposed model

The improvements in deep learning networks have been spectacular, from Lenet to AlexNet, to VGG, GoogleLeNet, and ResNet. The network depth has risen from under 10 layers to 50 or more layers breaking the 100 layers barrier. Deeper networks are investigated, aiming for better accuracies. However, accuracy is not proportional to network depth, and other factors play a role. At hundreds of layers and at some point, the training error and accuracy stop falling and start rising-falling as a result of the neural network degradation problem causing the effect of the early layers to be diluted. This problem stops the indefinite network depth increase. Moreover, before the rise of ResNets, it was difficult for the network to approximate the identity mappings of added nonlinear layers. With the introduction of ResNets, these barriers were overcome [33].

The basic building blocks of ResNet are shown in Fig. 2.a. with weight blocks that multiply the input matrix by a weight matrix then add a bias, and use the ReLU function, which keeps positive inputs as is but replaces negative inputs with a zero, and an adder which adds the two inputs. A more detailed ResNet block is pictured in Fig. 2.b. The ResNet employs a skip connection (top) to pass the identity to the output simplifying the generation of the input at the output. In addition, it is easier for the network to learn the residual function D(U)-U, where D(.) is the desired function to train than it is for the learn D(U). When it is desired to pass the input U to the output D(U), the ResNet block consisting of 2 wt blocks, separated by a ReLU is simply trained to produce the 0 matrix, which when added to U at the addition block, produces U. This is how ResNet helps achieve greater neural network depths, lessening the vanishing gradient problem commonly encountered in deep networks of tens of layers. The skip connection essentially allows the network to use fewer layers at the start of the training as some layers are skipped; expediating the training process. As the training progresses, the

Fig. 6. Samples of the newly created dataset for ArTSs [29].

ResNet network begins using the skipped layers.

Building deeper ResNet networks is made possible by cascading the network of Fig. 2.a. several times to reach the desired accuracy and reduce the training error as shown in Fig. 3. Some of these blocks may be pooling blocks, as when some Convolution or Weights layers generate V matrices of a different size than the U matrix, U is resized to match the size of the V matrix so that U can be properly added to V at the adder block. This resizing materializes by adding (at the adder block) V_{i+1} to (T. U_i), where T is a matrix zero-padded in the rows and columns which are missing in the original U_i. Fig. 1 is a cascaded architecture with building blocks from Fig. 1.b.

We propose two ResNet models with modified image parameters and a different number of layers for traffic sign classification and compare them with the performance of other CNN architectures. The ResNet V1 is the model that contains an additional modification of adding the shortcut identity to each block of two 3 × 3 filters as shown in Fig. 4. The identity mapping is used in all the shortcut connections and the projection shortcut is applied in situations of mismatch between the input and output dimensions. Batch normalization and nonlinear activation are used for the shortcut connection to avoid vanishing gradient and degradation problems. Eq. (1) shows the calculation of the depth of the convolutional layers of ResNet V1.

$$V1_{Network\_Depth} = N \times 6 + 2 \tag{1}$$

where the residual blocks are designated N. The number of stages is represented by the variable i.

The bottleneck connections are introduced in the ResNet V2 and with the depth calculation equation modified to multiply the residual blocks by 9 (increased of three compared to V1) filters as shown in Fig. 4. Convolutional layers of size 1x1, 3 × 3, and 1 × 1 are the three layers inside the residual block. Input dimensions are controlled using the 1 × 1 layer while smaller dimensions use the 3 × 3 as the bottleneck. The depth of the ResNet V2 is calculated using Eq. (2).

$$V2_{Network\_Depth} = N \times 9 + 2 \tag{2}$$

ResNet 1 starts by dividing the feature maps into two parts, while simultaneously doubling the filter maps. ResNet 2 introduces a bottleneck connection, where the filter size is determined according to Fig. 4. Furthermore, the skip connection's block size is tripled. Within each residual function block, there are three convolutional layers: one with a size of [1 × 1], another with [3 × 3], and a final one with [1 × 1]. The 1 × 1 layer handles the increase and decrease of input dimensions, while the 3 × 3 layer serves as a bottleneck, reducing the dimensions.

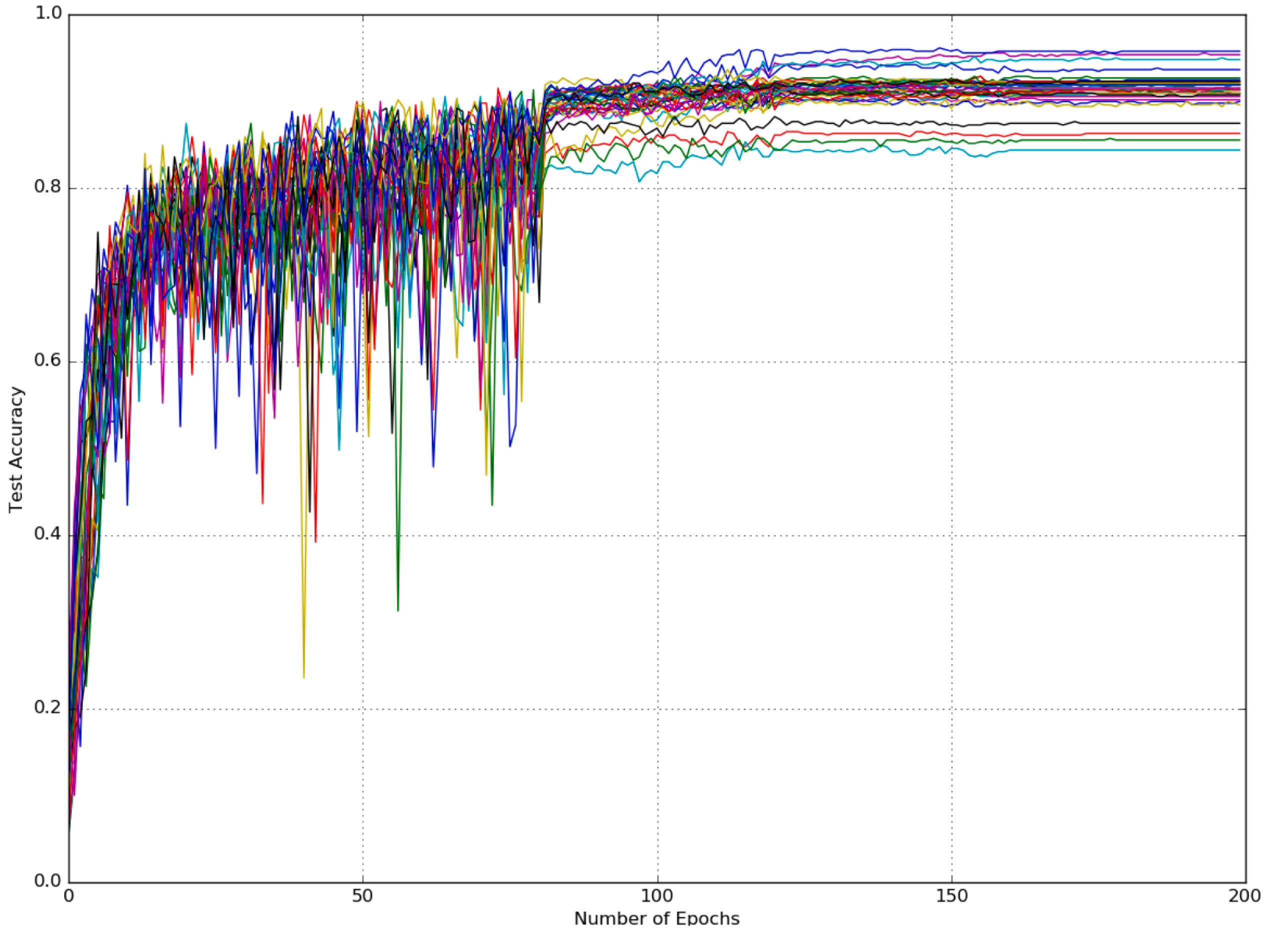

Fig. 7. Validation accuracies with different parameters of ResNet V1 and V2.

3.2. Dataset description

The ArTS is a newly developed labeled dataset of Arabic Traffic Sign images. The dataset is developed at Prince Mohammad Bin Fahd University, AlKhobar, Kingdom of Saudi Arabia, and is made publicly accessible for the researchers. The dataset is useful for all types of applications related to autonomous vehicles, traffic monitoring, and traffic safety. It can also be useful for developing applications for persons with visual impairments who are unable to see the signs. In [22–27], examples of datasets related to the dataset developed are presented. The high potential for the use of this dataset exists in the development of automated applications as well as the use in research for increasing the classification and recognition of the ArTSs.

A new dataset for ArTSs is developed for the selected most common 24 Arabic traffic signs. The dataset consists of 2,718 real captured images, as shown in Fig. 5 [34]. The images are captured from two major cities in the Eastern province of Saudi Arabia, namely, Alkhobar and Dammam. The newly developed dataset consists of 2,718 real images randomly partitioned into an 80% training set (2,200 images) and a 20% testing set (518 images).

3.3. Data augmentation and preprocessing

Due to the different dimensions of the RGB images in the dataset, there was a need for pre-processing before inputting them into the network. The total number of images per traffic sign varies, however, the pre-processed dataset consists of 57,078 augmented images. The newly developed dataset is stored as 3 channel RGB images with different dimensions and variations. As stated, the dataset was randomly portioned into 80% training and 20% testing. The training portion of the dataset was further partitioned to include 20% validation. The number of classes 24 corresponding to the 24 most common Arabic Traffic signs were numbered to range from 0 to 23, with each representing a class. The number of images in each class differs. Fig. 6 shows the different classes with their corresponding labels, names, and the number of images.

To make it more feasible for interested researchers to access the new ArTS dataset, the authors in [34] have generated, labeled, and published the ArTS dataset. The dataset is comprehensive and is proven to be sufficient for training and classification [35]. The current version of the published dataset can be used as-is, while future versions may include more image variations and numbers. The ArTS dataset is not without limitations. The insufficient light and noise variation images, as well as the non-inclusion of all the ArTS classes, have limitations that will be addressed in future versions of the dataset.

4. Experimental results and discussion

We used the ArTS dataset in running several experiments using the proposed ResNet V1 and ResNet V2 by varying the parameters to find the best configurations of the proposed dataset that will achieve the highest accuracy. The experiments are made using a high-end processing hardware machine with four NVIDIA Tesla V100 graphical processing units, 256 GB memory, and an Intel Xeon having 56 cores of 2.10 GHz while the software codes are written in the python programming language. Fig. 7 shows the validation accuracies for the different parameters with epochs incremented to a maximum of 200 epochs. As can be observed in Fig. 7, the validation accuracy is fluctuating up to approximately 80 epochs meaning that it will not predict new values of the dataset correctly using approximately 80 epochs or less. However, as the increase above 80 epochs, the fluctuation decreased dramatically,

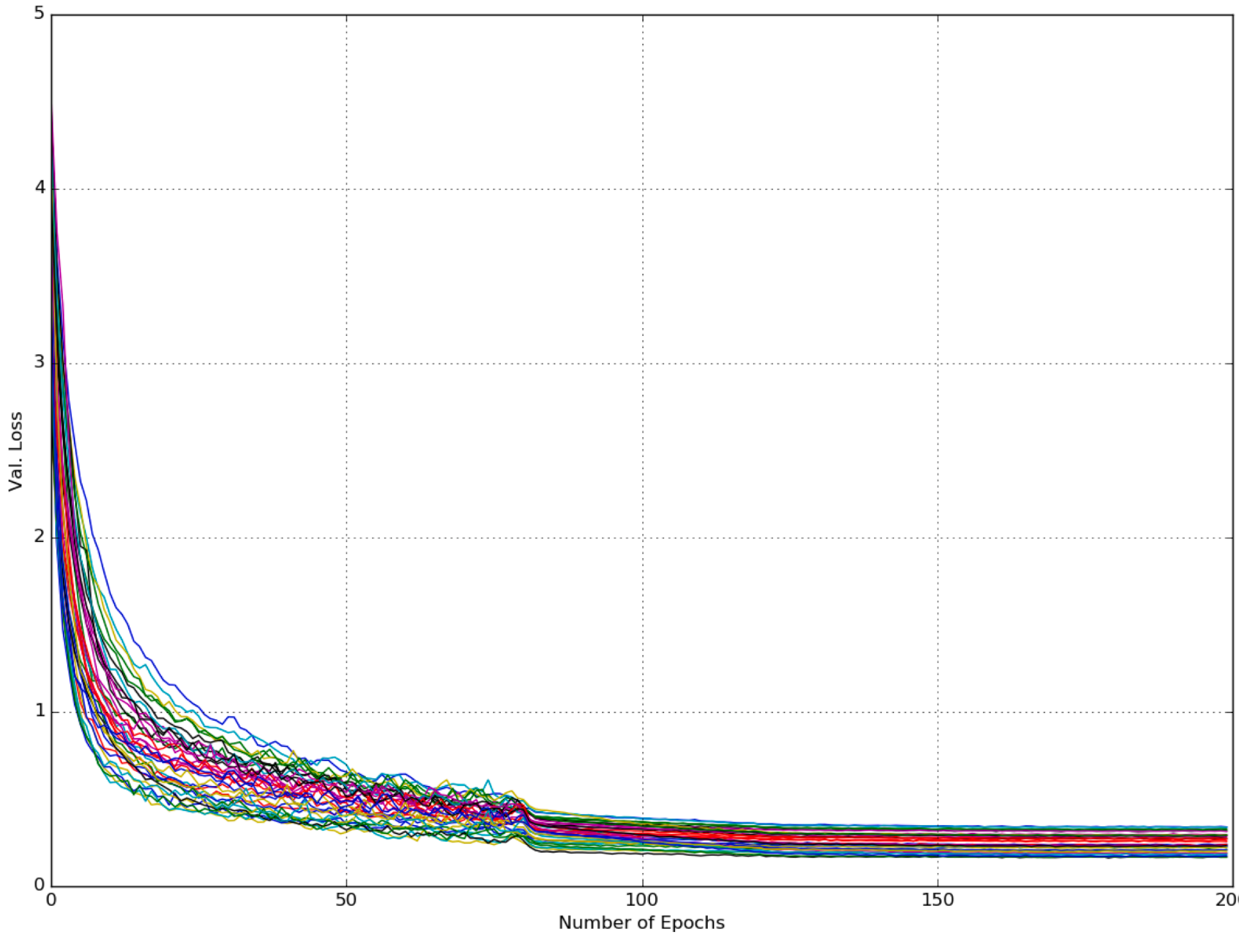

Fig. 8. Validation loss with different parameters of ResNet V1 and V2.

ensuring proper prediction power. The fluctuation becomes almost invisible at approximately 120 epochs and above, which indicates that the configuration with 120 epochs would have the best prediction accuracies.

The validation loss is also calculated through experimentation on the proposed ResNet V1 and ResNet V2 using the ArTS dataset with varying configuration parameters and the number of epochs. Fig. 8 that the validation loss converges to zero beyond 100 epochs which indicates that both ResNet V1 and ResNet V2 fit perfectly and are neither overfitting nor underfitting. This also indicates the power of the two versions to well generalize the predictive model.

In Fig. 9, the curves show a comparison of the training and validation accuracies of the proposed ResNet V2 using the optimized parameters and different epochs. It is seen that both the validation and training have reached almost similar values. This means that the gap between the training and validation is minimal, ensuring that the proposed model fits perfectly with no overfitting or underfitting.

Using the ArTS dataset, the proposed ResNet V1 model with optimized parameters and varying epochs is now used in the experiments. Fig. 10 shows a comparison between the training and validation accuracies of the experiments. It is noticed that both values are close to 1 and that the gap (difference) between the two values is very low, indicating that the model fits well with no overfitting and no underfitting.

The training and validation loss for the proposed ResNet V2 with optimized parameters and various epochs was then calculated and shown in Fig. 11. The convergence of both values toward zero indicates that the model is fitting well with no overfitting and underfitting. This also indicates the generalization power of the model as a predictive model.

The training and validation loss for the proposed ResNet V1 with optimized parameters and varying epochs is then calculated and the obtained results are displayed in Fig. 12. The convergence of both values towards zero indicates that the model fits well with no overfitting and no underfitting. This also indicates the generalization power of the model as a predictive model.

Some experiments are performed to predict the performance of the proposed optimized ResNet models V1 and V2 using the ArST dataset. Table 1 summarizes the results of the experiments and compares the performance of the standard ResNet Models V1 and V2 and the proposed optimized ResNet models V1 and V2 detailed in this paper. The training accuracy of the standard ResNet V1 is 95.73%, while the proposed optimized ResNet V1 achieved an average training accuracy of 99.18%, with an increase of 3.45%. The results show an average training accuracy of 94.82% for the standard ResNet V2, while the proposed optimized ResNet V2 achieved an average training accuracy of 97.55%, which increased by 2.73%. The proposed optimized ResNet V1 model achieved the best accuracy but it comes at the cost of training time as the proposed optimized ResNet V1 model took 8425 s for training which is the highest training time among all four models. The validation accuracy is also highest for the proposed optimized V1 model, with a validation accuracy of 96.14%. Both proposed models have been shown to fit well, and results indicated that there is no overfitting and no underfitting. Thus, both proposed models can generalize well and classify new data. A comparison with other research models will not be of benefit as this research is done on a newly developed dataset and once research is done on the dataset then a comparison can be done on future research.

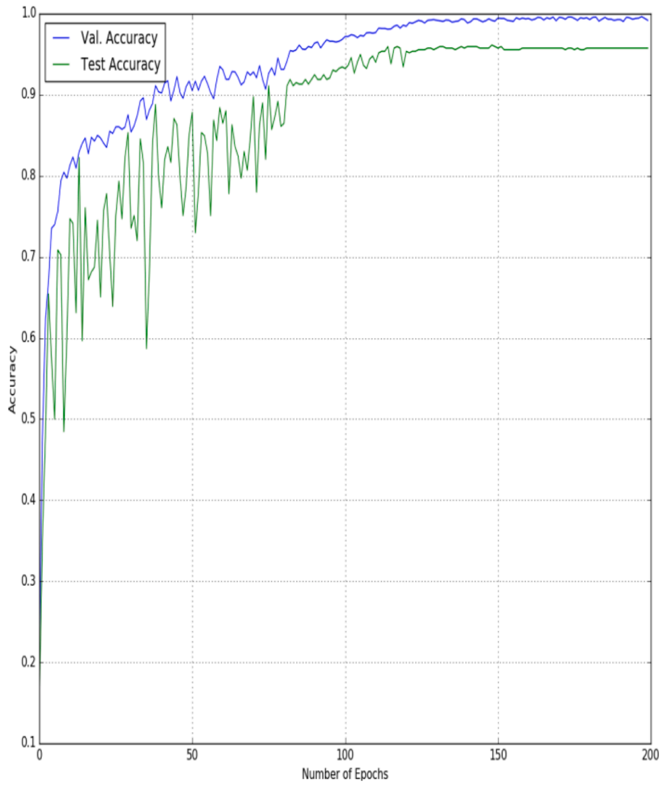

Fig. 9. Comparison between training and validation accuracies for ResNet V2 with optimized parameters.

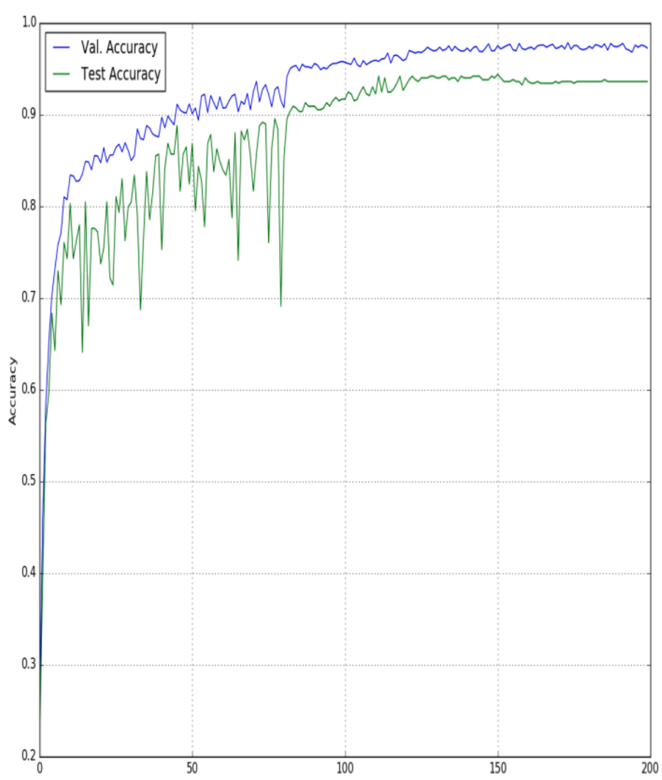

Fig. 10. Comparison between training and validation accuracies ResNet *V*1 with optimized parameters.

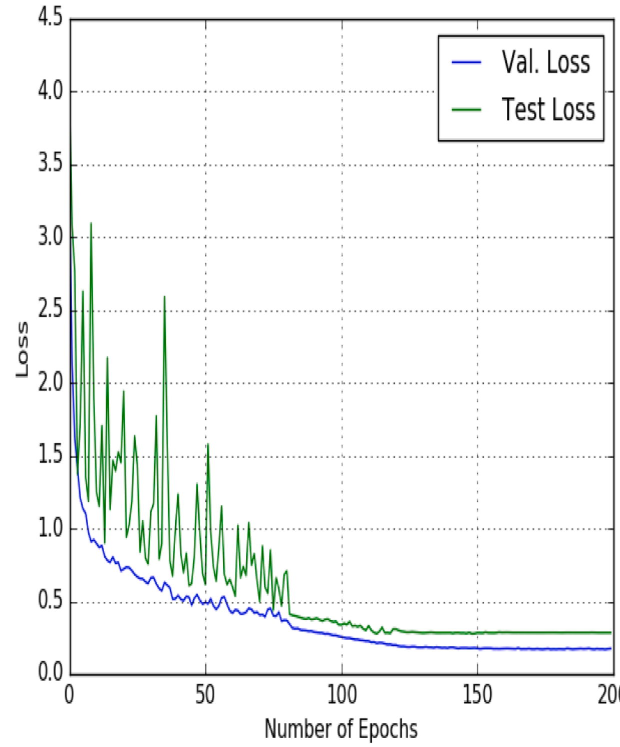

Fig. 11. Comparison between training and validation loss ResNet *V*2 with optimized parameters.

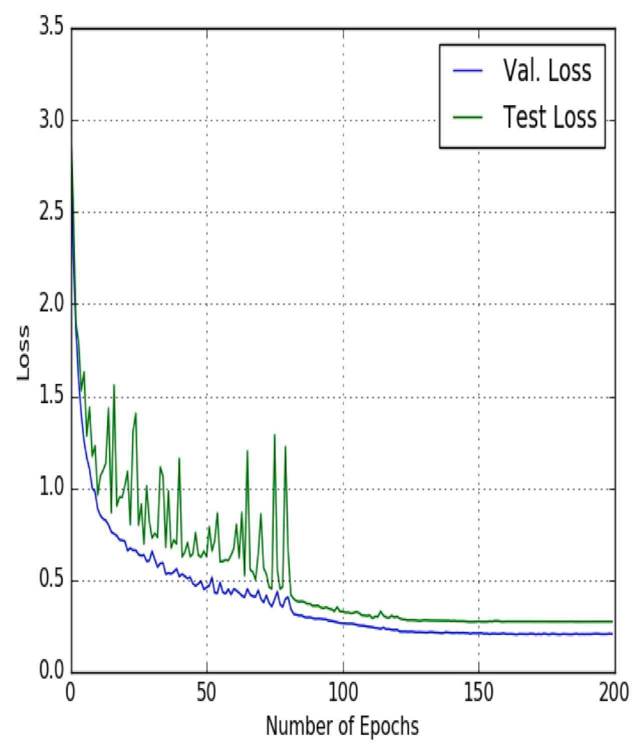

Fig. 12. Comparison between validation and test loss ResNet *V*1 with optimized parameters.

5. Conclusion

In this paper, a vital application of Automatic Arabic Traffic Sign Recognition is extensively studied. Residual Networks are used as a starting point to tackle the problem of automatic Arabic traffic sign

Table 1
Comparison of accuracies for the optimized ResNet models along with the computational cost.

Version	Layers	Val. Acc.	Val. Loss	Test Acc.	Test Loss	Training Time
V1 (Standard)	8	95.73%	0.1666	87.45%	0.3162	1758 s
V2 (Standard)	11	94.82%	0.2623	86.29%	0.4750	1210 s
V1 (Optimized)	56	99.18%	0.1774	96.14%	0.2765	8425 s
V2 (Optimized)	29	97.55%	0.2113	94.40%	0.2757	5023 s

recognition. Two optimized ResNet models are proposed version 1 and version 2 and are detailed in the paper. The ArTS dataset is used in this study. The best average training accuracy and average validation accuracy of 99.18% and 96.14%, respectively, are achieved using the proposed optimized ResNet V1. Though, the accuracy is extremely high and outperforms other methods reported in recent literature, yet, the researchers aspire to continue research on this topic to increase the accuracy to nearly 100% because the potential use of traffic sign recognition in driverless cars would require extremely high recognition rates for safety purpose. The researchers will continue to develop a more comprehensive complex dataset of Arabic Traffic Signs to benefit from a larger dataset taken under different conditions. We require images taken at different seasons, different weather condition especially during sandstorms which are frequent in Saudi Arabia, and signs that have suffered from degradation. Additionally, the authors would like to use the proposed models on embedded systems in real world scenarios to ensure that the proposed algorithms are capable to do the functions in real-time with not computational delay. Future work will include expanding the work to include indoor signage as a method to assist individuals with disabilities, such as individuals with visual impairments.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Tang, J., Wan, L., Schooling, J., Zhao, P., Chen, J., & Wei, S. (2022). Automatic number plate recognition (ANPR) in smart cities: A systematic review on technological advancements and application cases. *Cities*, *129*, 103833. Driss, M., Almomani, I., Al-Suhaimi, R., & Al-Harbi, H. (2022).
[2] A.S.A.L.P. Detection, R.U.D.C.N. Networks, In advances on intelligent informatics and computing: Health informatics, intelligent systems, data science and smart computing, Springer International Publishing, Cham, 2022, pp. 3–15.
[3] X.R. Lim, C.P. Lee, K.M. Lim, T.S. Ong, A. Alqahtani, M. Ali, Recent advances in traffic sign recognition: Approaches and datasets, Sensors 23 (10) (2023) 4674.
[4] M.I. Pavel, S.Y. Tan, A. Abdullah, Vision-based autonomous vehicle systems based on deep learning: A systematic literature review, Applied Sciences 12 (14) (2022) 6831.
[5] S.B. Wali, M.A. Hannan, A. Hussain, S.A. Samad, An automatic traffic sign detection and recognition system based on colour segmentation, shape matching, and svm, Mathematical Problems in Engineering (2015).
[6] D. Tabernik, D. Skočaj, Deep learning for large-scale traffic-sign detection and recognition, IEEE Transactions on Intelligent Transportation Systems 21 (4) (2019) 1427–1440.
[7] A. Alam, Z.A. Jaffery, Indian traffic sign detection and recognition, International Journal of Intelligent Transportation Systems Research 18 (1) (2020) 98–112.
[8] A. Vennelakanti, S. Shreya, R. Rajendran, D. Sarkar, D. Muddegowda, P. Hanagal, January). Traffic sign detection and recognition using a cnn ensemble, in: In IEEE International Conference on Consumer Electronics (ICCE), 2019, pp. 1–4.
[9] K.S. Boujemaa, I. Berrada, A. Bouhoute, K. Boubouh, Traffic sign recognition using convolutional neural networks, In IEEE International Conference on Wireless Networks and Mobile Communications (WINCOM) (2017) 1–6.
[10] W. Farhat, H. Faiedh, C. Souani, K. Besbes, Real-time embedded system for traffic sign recognition based on ZedBoard, Journal of Real-Time Image Processing 16 (5) (2019) 1813–1823.
[11] K. Lim, Y. Hong, Y. Choi, H. Byun, Real-time traffic sign recognition based on a general purpose GPU and deep-learning, PLoS One1 12 (3) (2017) e0173317.
[12] J.H. Chung, D.W. Kim, T.K. Kang, M.T. Lim, Traffic sign recognition in harsh environment using attention based convolutional pooling neural network, Neural Processing Letters 51 (3) (2020) 1–23.
[13] Y. Zhu, C. Zhang, D. Zhou, X. Wang, X. Bai, W. Liu, Traffic sign detection and recognition using fully convolutional network guided proposals, Neurocomputing 214 (2016) 758–766.
[14] M. Swathi, K.V. Suresh, Automatic traffic sign detection and recognition: A review, in: IEEE International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET), 2017, pp. 1–6.
[15] P. Dhar, M.Z. Abedin, T. Biswas, A. Datta, Traffic sign detection—A new approach and recognition using convolution neural network, In IEEE Region 10 Humanitarian Technology Conference R10-HTC (2017) 416–419.
[17] H. Fleyeh, M. Dougherty, Road and traffic sign detection and recognition, in: In *16th MIni-EURO Conference and 10th MeetIng of EWGT*, 2005, pp. 644–653.
[18] F. Feng, traffic sign detection and recognition system for intelligent vehicles. MSc thesis, school of electrical engineering and computer science, University of Ottawa, Canada, 2014.
[19] G. Loy, N. Barnes, Fast shape-based road sign detection for a driver assistance system, In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2004) 70–75.
[20] H. Li, F. Sun, L. Liu, L. Wang, A novel traffic sign detection method via color segmentation and robust shape matching, Neurocomputing 169 (2015) 77–88.
[21] A. Youssef, D. Albani, D. Nardi, D.D. Bloisi, Fast traffic sign recognition using color segmentation and deep convolutional networks, in: International Conference on Advanced Concepts for Intelligent Vision Systems, Springer, Cham, 2016, pp. 205–216.
[22] J. Stallkamp, M. Schlipsing, J. Salmen, C. Igel, The German traffic sign recognition benchmark: A multi-class classification competition, In IEEE International Joint Conference on Neural Networks (2011) 1453–1460.
[23] R. Qian, B. Zhang, Y. Yue, Z. Wang, F. Coenen, Robust chinese traffic sign detection and recognition with deep convolutional neural network, in: In IEEE 11th International Conference on Natural Computation (ICNC), 2015, pp. 791–796.
[24] Tsinghua database, https://cg.cs.tsinghua.edu.cn/traffic-sign/, Accessed Jan. 2022.
[25] Belgian dataset, https://btsd.ethz.ch/shareddata/, Accessed Jan. 2022.
[26] UK data set, https://www.gov.uk/guidance/traffic-sign-images, Accessed Jan. 2022.
[27] LISA data set, UCSD, http://cvrr.ucsd.edu/LISA/lisa-traffic-sign-dataset.html, Accessed Jan. 2022.
[28] G. Latif, F. Yousif Al Anezi, D.A. Iskandar, A. Bashar, J. Alghazo, Recent advances in classification of brain tumor from MR images-state of the art review, Current Medical Imaging, 2022 from 2017 to 2021.
[29] A. Kamilaris, F.X. Prenafeta-Boldú, Deep learning in agriculture: A survey, Computers and electronics in agriculture 147 (2018) 70–90.
[30] J.M. Alghazo, G. Latif, L. Alzubaidi, A. Elhassan, Multi-language handwritten digits recognition based on novel structural features, Journal of Imaging Science and Technology 63 (2019) 1–10.
[31] G. Grekousis, Artificial neural networks and deep learning in urban geography: A systematic review and meta-analysis, Computers, Environment and Urban Systems 74 (2019) 244–256.
[32] Latif, G., Alghazo, J., Mohammad, N., & Alghazo, R. (2021, July). Communicating with the Deaf and Hard of Hearing through Automatic Arabic Sign Language Translator. In *Journal of Physics: Conference Series* (Vol. 1962, No. 1, p. 012055). IOP Publishing.
[33] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition, In IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016) 770–778.
[34] Latif, G., Alghazo, J., Alghmgham, D. A. and Alzubaidi, L. (2020), ArTS: Arabic Traffic Sign Dataset, Mendeley Data, v1, https://doi.org/10.17632/4tznkn45mx.1, Accessed Dec. 2021.
[35] D.A. Alghmgham, G. Latif, J. Alghazo, L. Alzubaidi, Autonomous traffic sign (ATSR) detection and recognition using deep CNN, Procedia Computer Science 163 (2019) 266–274.