Journal of
Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7766169

# TRAFFIC SIGN IDENTIFICATION AND RECOGNITION SYSTEM USING DEEP NEURAL NETWORK

**[1]R.Gayathri*, [2]S.Hemamalini, [3]K.Harini, [4]N.G.Bangayar Selvi, [5]I.Akshaya, [6]Dr.Deepa Natesan***

Department of Electronics and Communication Engineering

[1,2,3,4,5,6]Sri Manakula Vinagayar Engineering College, Puducherry, India.

[1]gayathridevee@gmail.com, deepa.bme@smvec.ac.in

**ABSTRACT** – The traffic sign identification and recognition system are the process of detecting the traffic signs present on the roadside. At present, the world is developing with new technologies but the accident rate is highly increasing. The main cause is driver's failure in interpreting the traffic sign boards. In this paper, a new method of traffic sign identification and recognition system is proposed to reduce the accident rate. The implementation of the proposed system uses Deep Neural Network (DNN) and an object detection algorithm called YOLO V5 (You Only Look Once) algorithm. This method is trained with dataset from Kaggle called traffic signs dataset in YOLO format. This system enhances the driver interactivity while driving a vehicle. As a result, it improves the attentiveness of the driver by displaying signs by producing alarm sound with improved precision rate compared to the traditional methods. A comparison of simulation results shows the effectiveness of the system.

Keywords: Traffic sign identification and recognition system, Object detection algorithm, DNN, YOLO V5 algorithm, Kaggle.

*** Corresponding Author:**

Dr.N.Deepa, Assistant Professor, Sri Manakula Vinayagar Engineering College, Madagadipet, Puducherry, India-605107. Mail: deepa.bme@smvec.ac.in. Mobile: 91-9940811780.

## INTRODUCTION

In daily life, most people often use vehicles for transportation. Drivers are taking more risks when driving as a result of the growth in automobiles, which may also cause accidents. Every year, the number of accidents increases significantly (Lopez-Montiel et al.,2021). The primary cause of these accidents is the driver's failure to interpret all of the visual information accessible to them while operating a vehicle. Thus, a sophisticated driver assistance system is introduced to reduce the accident rate (Liu et al.,2017). Traffic signs are classified into three types namely mandatory signs, cautionary signs and informatory signs.

Deep Learning (DL) algorithms designed for real-time applications are currently undergoing rapid development. Deep learning is a branch of machine learning and artificial intelligence (AI) that models how people gather specific types of information (Stallkamp et al.,2012), (Logisvary et al.,2022). The various deep learning networks are Convolutional Neural Networks (CNNs), Deep Neural Networks (DNN), Long Short-Term Memory Networks

(LSTM), Recurrent Neural Networks (RNN), Multilayer Perceptron (MLP), Self-Organizing Maps (SOMs) and so on. These algorithms are arranged in a hierarchy of increasing complexity and abstraction, as opposed to conventional machine learning algorithms which are linear (Wali et al.,2019). In the study of conventional computer vision algorithms, geometric aspects were used to analyze images for the visual problem of recognizing traffic signs. However, these algorithms lacked the necessary power to achieve the desired results (Cheng et al.,2018), (Wang et al.,2019), (Hatcher et al.,2018). On the other hand, the internal conditions are the parameters that the algorithm may modify response time, detection accuracy, adaptability and hardware dependence (Everingham et al.,2009).

The Region-Based Convolutional Neural Networks (R-CNN) is the first deep learning-based effort that served as the foundation for the two-stage approaches. The most recent approach is cascaded R-CNN can produce more accurate detections (Zhang et al.,2020). But it requires more computing, it is less suitable for realtime applications. However, the one-stage and two-stage techniques have distinct drawbacks as its processing capacity differs.

Using the algorithms of DL, which have evolved steadily over time and are becoming more effective at a variety of tasks such as object recognition, segmentation, and classification, the recognition task can be solved. The YOLO (You Only Look Once) model is an essential one-step technique (Jeong et al.,2018). YOLO is a technique that uses neural networks to recognize objects in real time. The efficiency and speed of this algorithm are the reason for its popularity. It has been used in a variety of ways to recognize animals, people, parking meters, and traffic signs. In this method, the image is divided into N grids, each of which has an equal area of SxS. Each of these N grids is responsible for detecting and locating the object it contains. The YOLO V5 algorithm is a regression-based method that predicts classes and bounding boxes in a single pass for the entire image, not just the main part. Modern real-time object recognition techniques have benefited greatly from it. The algorithm has been further developed into a new YOLO V5 version that increases the accuracy of object detection.

(M. Lopez-Montiel et al.,2021) presented an autonomous scheme for recognizing traffic signs based on deep learning. It cannot be applied in real-time environments and provides a preliminary analysis for such applications. (J.Zhang et al.,2020) refers to multiscale cascaded R-CNN. The drawback is that it often led to false detection. (Lee et al.,2021 refers end-to-end deep learning model and CNN-MT model as it's technique. The main drawback is that it does not encode the position and orientation of an object. G. Piccioli et al.,2020 proposed a method that analyzes the edges extracted from a single monochromatic image. The drawback is that the single image analysis should increase the robustness of the system. (R.Ayachi et al.,2020) refers to convolutional neural network-based deep learning technique. The drawback is that the CNN requires a significant amount of training data and it fails. (X.Bangquan et al.,2019) refers to traffic sign classification network and traffic sign detection. TSR is to ensure its efficiency, adequate accuracy, generalization and speed in real time. (J.Liu et al., 2019) paper explains the deep learning models on a mobile device. The drawback is that the performance characterization is critical for designing and implementing deep learning models on mobile devices. (M.Swathi et al., 2017) refers to feature matching or template matching algorithms and machine learning approaches. The disadvantage of this technique is the detection may suffer due to vibration and motion.

In this section, the traffic sign identification and recognition system, YOLO architectures, and various methods are discussed. Section 2 describes a proposed methodology for evaluating the traffic sign identification and detection system. Section 3 describes the experimental results and analysis of the proposed system. Section IV describes conclusions and future opportunities.

## PROPOSED SYSTEM

In this section, a new traffic sign identification and recognition method is proposed to improve the speed and accuracy by using YOLO V5 algorithm and DNN. The block diagram of the traffic sign identification and recognition system is shown in Fig. 2.1. It consists of two phases, the preprocessing phase and the recognition phase.

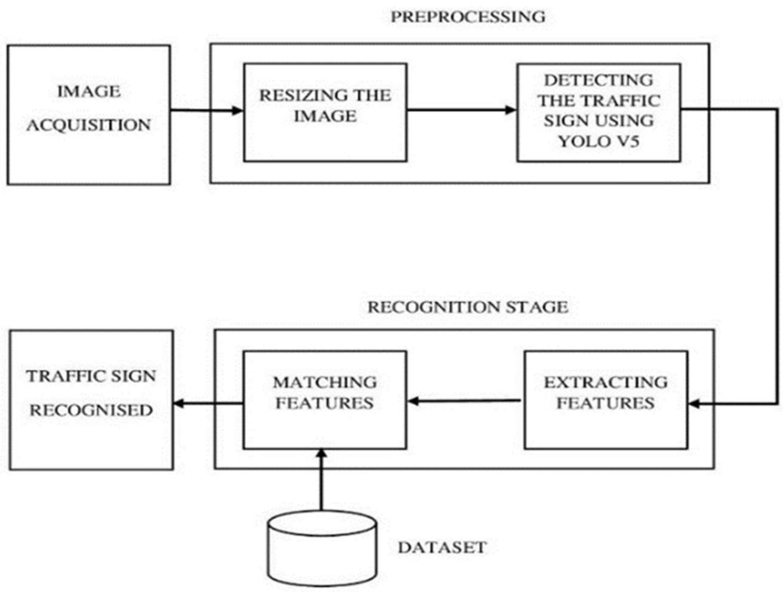


Fig. 2.1 Block Diagram of Traffic Sign Identification and Recognition System

From the block diagram, it is clear that the video is given as input to the hardware connected to CPU. The execution of the program is done using Microsoft Visual Studio. To identify the traffic signs, the image is resized so that all images have the same size. Then the resized image is processed to detect the traffic sign using YOLO V5 algorithm. In the recognition phase, the recognized traffic sign is processed by extracting the features as boundary. Then, the traffic sign features are matched with the dataset containing trained data. In this way, the traffic sign is detected with rectangular boundaries indicating the name of the traffic sign with its precision value.

### Phase 1 - Preprocessing

Preprocessing may refer to the manipulation of data before it is used to improve the performance. Analysis of data that have not been carefully checked for such problems can lead to misleading results. Therefore, the presentation and quality of the data is the most important factor before any analysis is performed.

### (i) Image Acquisition

The input video from the environment captured by the camera is shown in Fig. 2.2. The video has a frame size of 24 frames per second. It is forwarded to the central processing unit (CPU) and the hardware part connected to the system. Hardware acceleration is a technique that forces the hardware of a computer to work faster. Every digital computer system has a CPU which consists of the main memory, the control unit and the arithmetic-logic unit. The Node MCU microcontroller is a low-cost system-on-a-chip (SoC) with an open-source software and hardware development environment. The image is then pre-processed in next stage.


**Fig. 2.2 Input video for image acquisition**

(ii) Resizing the Image

This is the first stage of pre-processing. To identify the traffic signs, the image is resized, bringing all images to the same size. The original image size is 1360x800 pixels as shown in fig. 2.3(a). It is reduced to 416x416 pixels as shown in fig. 2.3(b).

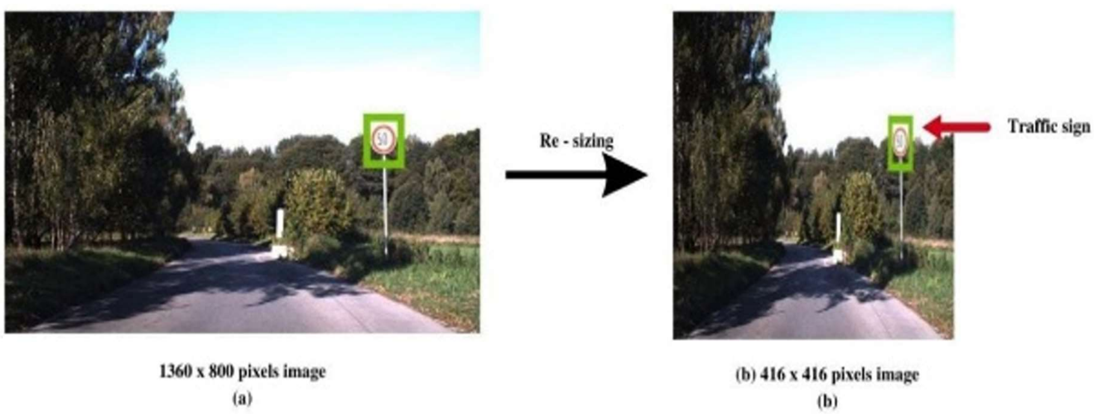


**Fig. 2.3** Resized output image

**(iii) Detecting the traffic sign using YOLO V5**

The resized image is processed to detect the traffic sign using the YOLO V5 algorithm. The detected traffic sign output is shown in fig. 2.4. This is a processing stage that identifies the presence of traffic signs when the name of the traffic sign is not recognized. The' ability of the model to classify high-quality images in the first layers of the network was the most important factor considered. Deep learning is a category of machine learning algorithms that use numerous layers to progressively extract higher-level information from input. In image processing, lower layers detect edges, while higher layers can identify things important to humans, such as numbers, letters, or faces. With an accuracy of 95.55%, the improved YOLO V5 has significantly improved recognition accuracy.

Fig. 2.4 Detected traffic sign output

The YOLO algorithm provides a high frame rate for real-time use with significantly better performance on all metrics. Therefore, the YOLO V5 algorithm is chosen. Deep neural networks are Artificial Neural Networks (ANNs) with more hidden layers between the input and output layers than ANNs normally have. When the system generates high-level functions from the supplied data using multiple layers of nodes. This means that the data is transformed to become a more imaginative and abstract element.

**Phase 2 - Recognition Stage**

The goal of the recognition phase is to obtain feature distributions of detected objects, compare them with the model database, and find the closest match. To avoid excessive computation time during recognition, only a small subset of all available features is used.

**(iv) Feature Extraction**

The recognized traffic sign is then processed in the recognition phase. The feature extraction method used here is the SURF (Speeded Up Robust Features) technique. This is a fast and robust approach for local similarity invariant image representation and comparison. The main attraction of the SURF technique is the fast computation of operators using box filters, which enables real-time applications such as tracking and object detection. It is used to extract image features and the advanced visual significance algorithm is used to evaluate the image and extract the region of significance. The extracted image output is shown in fig. 2.5.


Fig. 2.5 Extracted image output

**(v) Matching Features**

In this phase, the extracted feature image is used as input. It starts with comparing the extracted image with the trained model of the dataset. Data selection is the most important step in the selection and implementation of DL algorithms. The image of matched feature is shown in fig. 2.6.

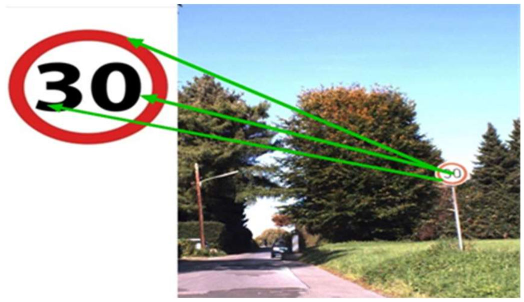


Fig. 2.6 Matched features

**(vi) Dataset**

The amount of data must be considered, because if the dataset is too small, the training suffers as a function of the number of samples per class. The traffic sign dataset in YOLO format consists of 43 classes of traffic signs that can be found in the meta folder of Kaggle. These samples are composed of both colour and grayscale examples. The colour examples are considered in the initial phase because they contain more information about the feature extraction phase of the DL models.

**(vii) Recognised Traffic Sign**

The traffic sign identification and recognition system using the YOLO V5 algorithm and DNN method is recognized and its name is displayed on the monitor as shown in fig. 2.7.


Fig. 2.7 Input Traffic Sign and Corresponding Output as Right_Curve

The following are the evaluation measures used to evaluate each detection phase of the models of DL. In the detection phase, bounding box, recall and precision metrics are used to evaluate the sensitivity and precision of the TSD systems. In the detection task, the prediction of the model is evaluated by the bounding box measure, where the overlap ratio between the predicted bounding box (Bp) and the ground truth box (Bgt) is calculated. A correct detection is given if the overlap ratio Intersection over Union (IoU) exceeds 0.5 using the equation (1).

$$\text{IoU} = \frac{area(Bp \cap Bgt)}{area(Bp\ U\ Bgt)} \tag{1}$$

The classification of traffic signs results was divided into three groups: true positive (x), false positive (y) and false negative (z). x are examples correctly labeled as positives, y are negative examples incorrectly labeled as positive, and z are positive examples labeled incorrectly as negative. Let precision be α, loss be β, recall be μ, mean Average Precision(mAP) be ∂, n be number of terms.

$$\alpha = \frac{x}{x+y} \tag{2}$$

Equation (2) shows the precision calculation. Precision is defined as the quality of positive prediction level performed by the model.

$$\beta = 1 - \alpha \tag{3}$$

Equation (3) shows the calculation of loss. The consequence of an inaccurate prediction is known as loss. It represents the degree of error in the model's prediction. A zero value indicates a flawless prediction, while any other value indicates a degree of inaccuracy.

$$\mu = \frac{x}{x + z} \tag{4}$$

Equation (4) shows the calculation of recall. The recall is a measurement of the model's ability to detect positive samples.

$$\partial = \frac{1}{N}\sum_{i=1}^{i=n} APi \tag{5}$$

Equation (5) shows the calculation of Mean Average Precision (mAP). It is a performance measure used to evaluate the accuracy of search engines and other information retrieval systems. It is calculated by averaging the Average Precision (AP) value obtained from the average of the precision scores obtained at each relevant result.

## EXPERIMENTAL RESULT AND ANALYSIS

The traffic sign identification and recognition system using the YOLO V5 algorithm and DNN has been successfully implemented. This algorithm provides higher accuracy than the others. The Node MCU microcontroller is integrated with CPU and programmed in Visual Studio to display the output on LCD. In the existing system, SSD were used for traffic sign detection. In the proposed system, YOLO V5 algorithm and DNN are used. After real-time testing, it is found that this is one of the most effective methods for recognising traffic signs. Moreover, this system can be used in daily life to facilitate the lifestyle.

### (i) Output Analysis

The output of traffic sign identification and recognition system is detected and recognized as pedestrian walk and its precision value are also indicated as shown in fig. 3.1.


Fig. 3.1 Input Video with Corresponding Output as Pedestrian Walk

The output of traffic sign identification and recognition system is detected and recognized as maximum_60 and its precision value are also indicated as shown in fig. 3.2.


Fig. 3.2 Input Video with Corresponding Output as Right Curve

**(ii) Tabular Analysis**

The various results of mean Average Precision (mAP) from the previous and the proposed methods are shown in table 1. The mAP value is calculated by averaging the Average Precision (AP) values.

**Table 1 Results obtained from various methodologies**

| Sl.No. | Reference | Method | Classes | AP/mAP |
|---|---|---|---|---|
| 1 | M. Lopez-Montiel et al., [1] | Feature Pyramid Network (FPN) + Single Shot Detector (SSD) | Prohibitory | 0.8075(AP) |
| | | | Mandatory | 0.8(AP) |
| | | | Informatory | 0.9308(AP) |
| | | | All | 0.8461(mAP) |
| 2 | A. Mogelmose et al., [32] | Integral Channel Feature (ICF) | Prohibitory | 0.8732(AP) |
| | | | Mandatory | 0.9109(AP) |
| | | | Informatory | 0.9603(AP) |
| | | | All | 0.9148(mAP) |
| 3 | A. Mogelmose et al., [32] | Aggregate Channel Feature (ACF) | Prohibitory | 0.9898(AP) |
| | | | Mandatory | 0.9617(AP) |
| | | | Informatory | 0.9611(AP) |
| | | | All | 0.9708(mAP) |
| 4 | Proposed System | YOLO V5 + Deep Neural Network (DNN) | Prohibitory | 0.9769(AP) |
| | | | Mandatory | 0.9892(AP) |
| | | | Informatory | 0.9783(AP) |
| | | | All | **0.9814(mAP)** |

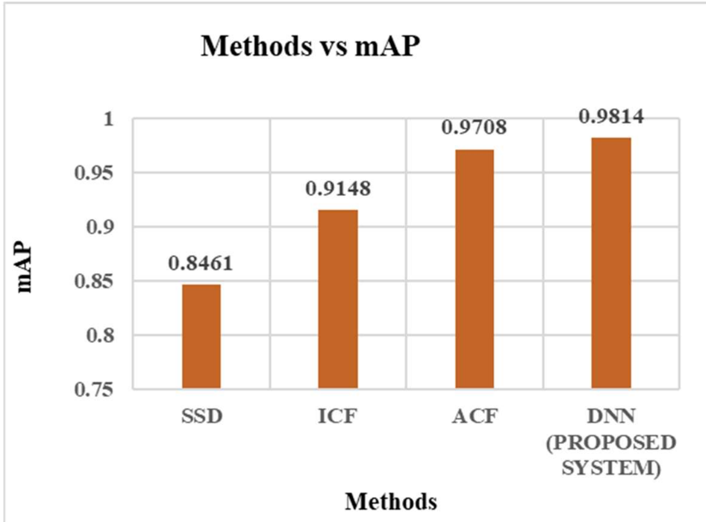


Fig. 3.3 Plotting of various methods and mAP

Fig. 3.3 shows the different methods with the mean average precision (mAP). From this, it can be inferred that the proposed method using DNN achieves better mAP compared to the previous methods.

### (iii) Graphical Analysis

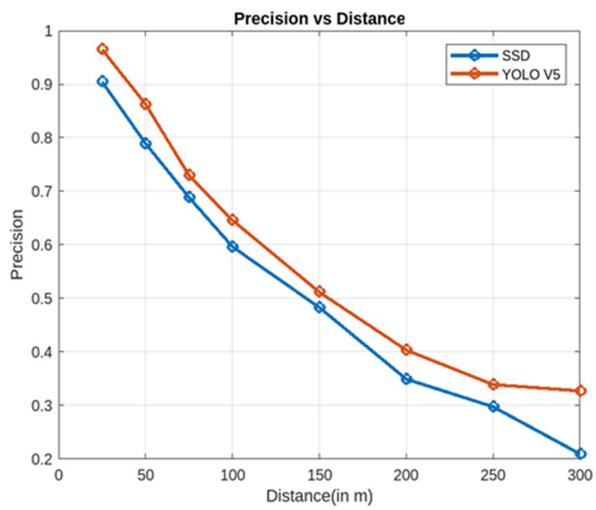


Fig. 3.4 Variation of Precision and Distance Between Proposed System and Previous System

Fig. 3.4 shows that the proposed traffic sign identification and detection system using YOLO V5 and DNN has higher precision in terms of distance than previous methods.

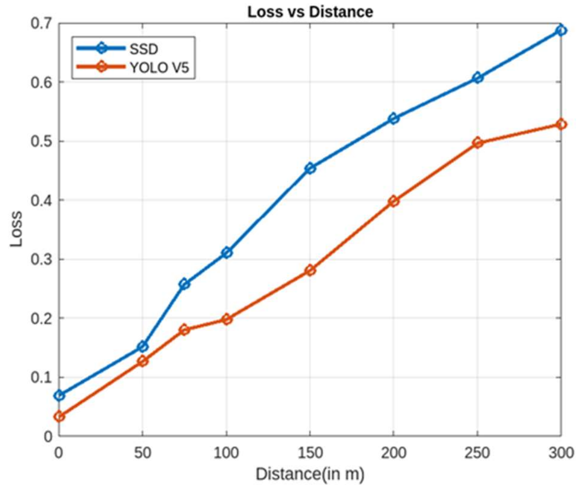


Fig. 3.5 Variation of Loss and Distance Between Proposed System and Previous System

Fig 3.5 shows that the proposed traffic sign identification and detection system using YOLO V5 and DNN has a lower loss in distance than previous methods.

In this section, the results of using DNN and YOLO V5 for traffic sign detection are analysed. The YOLO V5 model provides more accurate results compared to the traditional methods. The SSD technique provides less accurate results and there is a delay in traffic sign detection. Compared to the proposed system, the YOLO method provides higher accuracy than the SSD method. YOLO V5 is used to anticipate speed and accuracy to some extent.

## CONCLUSION

The traffic sign identification and recognition system using DNN and YOLO V5 was successfully implemented. By using CPU as a hardware accelerator, high speed of traffic sign recognition is achieved. The drivers received an accurate description of the output instruction from the input image of the traffic sign. This helped drivers avoid collisions caused by the incorrect recognition of traffic signs. In the future, this traffic sign identification and recognition system may also be used in autonomous self-driving cars.

## REFERENCES

M. Lopez-Montiel, U. Orozco-Rosas, M. Sánchez-Adame, K. Picos and O. H. M. Ross (2021), “Evaluation Method of Deep Learning-Based Embedded Systems for Traffic Sign Detection”, IEEE Access, Vol. No. 9, pp. 101217-101238.

W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi (Apr.2017), “A survey of deep neural network architectures and their applications”, Neurocomputing, Vol. 234, pp.11-26.

J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel(Aug.2012), “Man Vs. Computer: Benchmarking Machine Learning Algorithms for Traffic Sign Recognition”, Neural Netw., vol. 32, pp. 323-332.

Logisvary.V, Pushparaj.S, Dr.N.Jothy, Girijaadevi (2022), “Efficient and Deployable Smart Farm Processing System Using Machine Learning”, International Conference on Emerging Innovative Technologies in Engineering (ICEITE-22).

S. B. Wali, M. A. Abdullah, M. A. Hannan, A. Hussain, S. A. Samad, P. J. Ker, and M. B. Mansor(May 2019), “Vision-based traffic sign detection and recognition systems: Current trends and challenges”, Sensors, Vol. 19, no. 9, pp. 2093.

J. Cheng, P.S.Wang, G. Li, Q.-H. Hu, and H.-Q. Lu (Jan. 2018), “Recent Advances In Efficient Computation of Deep Convolutional Neural Networks”, Frontiers Inf. Technol. Electron. Eng., Vol. 19, no. 1, pp. 64-77.

J. Wang, S. Jiang, W. Song, and Y. Yang(2019), “A comparative study of small object detection algorithms”, in Proc. Chin. Control Conf. (CCC), pp. 8507.

W. G. Hatcher and W. Yu (2018), “A survey of deep learning: Platforms, applications and emerging research trends”, IEEE Access, Vol. 6, pp. 24411-24432.

M. Everingham, L.Van Gool, C. K. I.Williams, J.Winn, and A. Zisserman (Sep. 2009), “The Pascal visual object classes (VOC) challenge”, Int. J. Computer Vis., Vol. 88, no. 2, pp. 303-338.

J. Zhang, Z. Xie, J. Sun, X. Zou, and J. Wang (2020), “A cascaded R-CNN with multiscale attention and imbalanced samples for traffic sign detection”, IEEE Access, Vol. No. 8, pp. 29742–29754.

H.-J. Jeong, K.-S. Park, and Y.-G. Ha (Jan. 2018), “Image preprocessing for efficient training of YOLO deep learning networks”, in Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp),.

Y. E. Wang, G.-Y. Wei, and D. Brooks(2019), “Benchmarking TPU, GPU, and CPU platforms for deep learning”, arXiv:1907.10701.

W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu (2016), “SSD: Single shot multi-box detector in Computer Vision”—ECCV, Cham, Switzerland: Springer, pp. 21-37.

T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan and S. Belongie (Jul. 2017), “Feature pyramid networks for object detection”, Proc. IEEE Conf. Comp. Vision Pattern Recognition. (CVPR), pp. 936-944.

A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam (Jul. 2017), “Mobile Nets: Efficient convolutional neural networks for mobile vision applications”.

D.-H. Lee, K.-L. Chen, K.-H. Liou, C.-L. Liu, and J.-L. Liu (Jan. 2021), “Deep learning and control algorithms of direct perception for autonomous driving”, Int. J. Speech Technol., Vol. No. 51, no. 1, pp. 237–247.

G. Piccioli, E. De Micheli, and M. Campani (2020), “A robust method for road sign detection and recognition in Computer Vision”, - ECCV, Berlin, Germany: Springer, pp. 493-500.

R. Ayachi, M. Afif, Y. Said, and M. Atri (Feb. 2020), “Traffic signs detection for real-world application of an advanced driving assisting system using deep learning”, Neural Process. Lett., Vol. 51, No. 1, pp. 837–851.

S. You, Q. Bi, Y. Ji, S. Liu, Y. Feng, and F. Wu (Oct. 2020), “Traffic sign detection method based on improved SSD,’’ Information, Vol. 11, No. 10, pp. 475.

F. J. Belmonte, S. Martin, E. Sancristobal, J. A. Ruiperez-Valiente, and M. Castro (Feb. 12, 2020), ‘‘Overview of embedded systems to build reliable and safe ADAS and AD systems,’’ IEEE Intell. Transp. Syst. Mag., early access.

X. Bangquan and W. X. Xiong (2019), “Real-time embedded traffic sign recognition using efficient convolutional neural network”, IEEE Access, Vol. 7, pp. 53330-53346.

M. M. William, P. S. Zaki, B. K. Soliman, K. G. Alexsan, M. Mansour, M. El-Moursy, and K. Khalil (Dec. 2019), “Traffic signs detection and recognition system using deep learning”, in Proc. 9th International Conf. on Intelligence Computer Information System (ICICIS).

J. Liu, J. Liu, W. Du, and D. Li (Jul. 2019), “Performance analysis and characterization of training deep learning models on mobile devices”.

J. Hanhirova, T. Kamarainen, S. Seppala, M. Siekkinen, V. Hirvisalo, and A. Ylä-Jääski (Jul. 2018), “Latency and Throughput Characterization of Convolutional Neural Networks for Mobile Computer Vision”.

M. Swathi and K. V. Suresh (2017), “Automatic traffic sign detection and recognition: A review”, International Conference on Algorithms, Methodology, Models, and Applications in Emerging Technologies, pp. 1-6.

Jothy.N, Gayathri.R, Logisvary.V (2022), “Detection and Recognition of Unclear Vehicle Number Plate Using YOLO V5 Algorithm”, International Conference on Emerging Innovative Technologies in Engineering (ICEITE-22), Vol. 1, pp-25.

J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, and K. Murphy (Jul. 2017), “Speed/accuracy trade-offs for modern convolutional object detectors”, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 7310-7311.

A. Mogelmose, M. M. Trivedi, and T. B. Moeslund (Dec. 2012), ‘‘Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey,’’ IEEE Trans. Intell. Transp. Syst., vol. 13, no. 4, pp. 1484–1497.

K. He, X. Zhang, S. Ren, and J. Sun (Jun. 2016), ‘‘Deep residual learning for image recognition”, in Proc. IEEE Conf. Comput. Vis. Pattern Recognition (CVPR).

A. Mogelmose, D. Liu, and M. M. Trivedi (Sep. 2015), “Detection of U.S. traffic signs”, IEEE Trans. Intell. Transp. Syst., Vol. 16, no. 6, pp. 3116–3125.

J. Pradeep, E. Srinivasan and S. Himavathi, "Diagonal based feature extraction for handwritten character recognition system using neural network," 2011 3rd International Conference on Electronics Computer Technology, Kanyakumari, India, 2011, pp. 364-368.

Deepa, N & Pandiaraja, P, “Electronic healthcare system data privacy preserving efficient file retrieval from the cloud service provider using attribute based file encryption”, in Journal of Ambient Intelligence and Humanized Computing, DOI: 10.1007/s12652-020-01911-5, (2020)

Deepa, N & Pandiaraja, P, “Hybrid context aware recommendation system for e- health care by merkle hash tree from cloud using evolutionary algorithm”, in journal of Soft Computing (Springer), vol. 24, issue 10, pp. 7149–7161, (2020).

Deepa, N & Pandiaraja, P "A novel data privacy-preserving protocol for multi-data users by using genetic algorithm," in journal of Soft Computing (Springer), vol. 23, issue 18, pp. 8539–8553, (2019).