Traffic Sign Recognition Based on CNN vs Differnet Transfer Learning Techniques

Mohammed Ghouse
Dept. of Computer and Information Sci.
Northumbria University
London, UK
mdghouse1293@gmail.com

Dr. Saber Farag
Dept. of Computer and Information Sci.
Northumbria University
London, UK
saber.farag@northumbria.ac.uk

Dr.Usman Butt
Dept. of Computer and Information Sci.
Northumbria University
London, UK
usman.butt@northumbria.ac.uk

***Abstract*—This study looks into the effectiveness of various neural network architectures, such as convolutional neural networks (CNNs) and transfer learning models, for traffic sign recognition tasks that are critical for intelligent transportation systems (ITS), advanced driver assistance systems (ADAS), and self-driving systems. Traffic sign recognition is critical for improving road safety and supporting effective traffic management in these systems. While CNNs are widely utilized for their capacity to learn hierarchical features from visual input, dataset size, processing resources, and network architecture all have a substantial impact on their performance. Transfer learning is a promising strategy that uses pre-trained models on huge datasets to address data scarcity challenges.We compare CNNs to transfer learning models such as ASNet, EfficientNet, and VGG19, assessing accuracy, computational efficiency, and generalization capability using benchmark datasets such as the German Traffic Sign Recognition Benchmark (GTSRB). We evaluate the performance of these models using experiments that take into account measures such as accuracy, training length, and loss. CNNs attained a detection rate of 99%, demonstrating their suitability for traffic sign recognition tasks. Furthermore, transfer learning models produced encouraging results: VGG19 obtained an accuracy of 99.64%, EfficientNetB7 achieved 99%, and ASNet achieved 99.72%. These findings highlight the potential of CNNs and transfer learning models to improve traffic sign identification in various intelligent transportation systems.Overall, the study's findings have positive implications for strengthening the capacities of intelligent transportation systems as well as the safety and efficiency of road traffic management.**

Keywords: Artifical Intelligence , Deep Learning, Computer Vision, Object Detection, Road marking.

***Index Terms*—**

I. Introduction

MACHINE learning has become a game-changing technology in recent years, transforming a number of industries and raising people's standard of living everywhere. The field of transport and road safety is one of the most important uses of machine learning [7]]. The safety of drivers and pedestrians on the roads has become a top priority due to the quick rise in urban population and the volume of vehicles on the road. Using machine learning for pedestrian and road marking recognition is a creative and essential way to solve this issue [19].

A. Aim

Using computer vision and artificial intelligence, pedestrian and road marker detection improves the effectiveness and safety of transportation networks [5]. With the use of this technology, intelligent systems that can recognise in real-time both pedestrians and road markings—such as crosswalks, lane limits, and signage—will be developed.

B. Objectives

Improving Road Safety: Improving road safety is the main goal of utilising machine learning for pedestrian and road marking detection [12]. This technology seeks to decrease accidents and save lives by precisely identifying pedestrians, crosswalks, and road markers in real-time, hence creating safer road environments for all users.

Optimising Traffic Flow: . Machine learning algorithms facilitate the smooth and efficient passage of vehicles on the road by efficiently detecting lane borders, road markings, and other traffic control features [13]. Supporting Driver Assistance Systems: In order to give drivers useful information and improve their capacity to respond to changing road conditions, pedestrian and road marking detection systems are made to work in tandem with driver assistance systems, such as lane-keeping assistance and collision avoidance systems [3]

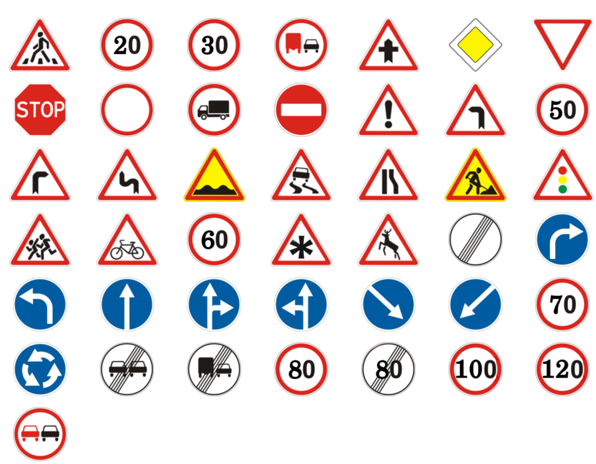

Fig. 1. Sample Image Of Traffic Sign

II. Literature Review

In recent years, significant advancements in autonomous vehicles have been achieved through the integration of deep learning and computer vision, as evidenced [11]. The critical challenges in this domain, particularly the accurate detection of pedestrians and road markings, have been identified by [14]. Deep learning, as a form of artificial intelligence, has emerged as a promising solution to revolutionize computer vision systems in addressing these challenges, given its ability to automatically extract intricate features from vast datasets. However, the complexity of real-world scenarios, encompassing diverse lighting conditions, occlusions, and environments, often surpasses the capabilities of traditional computer vision algorithms, as highlighted by [6]. This literature review aims to investigate the state-of-the-art in deep learning methods for the identification of road signs and pedestrians. While artificial intelligence-driven machine detection systems have made remarkable progress, issues such as data quantity and quality persist, as emphasized by [18]. The quality, accuracy, completeness, relevance, and timeliness of data are crucial factors influencing the robustness and dependability of detection technologies, as discussed by [8]. The review also underscores the challenges faced by autonomous vehicles in making judgments and spotting abnormalities on the road in nations with high automobile density, as examined by [15]. Additionally, the utilization of machine learning algorithms, including CNN, in self-driving cars is explored by [1], while [17] addresses decision-making challenges in autonomous cars navigating single-lane roads. These studies collectively underscore the need for addressing data-related challenges and advancing machine learning techniques to enhance the accuracy and resilience of autonomous vehicle detection systems. [2] employed YOLOv5m and YOLOv5s to detect road marking signs, achieving high accuracy with the YOLOv5m model. The dataset used for training was based on Taiwanese road marking signs. To enhance precision in road marking detection, the authors implemented road image instance segmentation. For this purpose, they utilized the Binarized Normed Gradient (BING) approach for detection and a Principal Component Analysis network (PCANet) for object classification, as discussed by [16]. Additionally, [4] explored various detection techniques, including the use of binarized normed gradient for detection. [9] introduced the hyperlearner architecture, highlighting its superior performance without requiring additional features. [10] addressed challenges in detecting pedestrians interacting or standing close to each other, proposing a probabilistic framework to estimate relationships in single and multiple detection configurations. This concept is reminiscent of the Single Shot Multibox Detector (SSD) described by [16], which is a real-time object identification system predicting class scores and bounding boxes at multiple scales for fixed-size anchor boxes. Overall, these approaches contribute to the ongoing efforts in advancing the accuracy and efficiency of road marking and pedestrian detection systems.

III. Methodology

In this research study, we employed a methodology grounded in convolutional Neural networks (CNNs) to enhance the detection of Traffic signs in urban environments. We initiated the process by collecting a diverse and representative dataset encompassing various scenarios, lighting conditions, and environmental factors. Subsequently, we preprocessed the dataset to ensure uniformity and relevance for the task at hand. Our CNN architecture was designed to effectively capture spatial hierarchies and feature representations crucial for discerning Traffic sign. We test CNN and transfer learning models using benchmark datasets such as the German Traffic Sign Recognition Benchmark (GTSRB) . The CNN design consists of many convolutional and pooling layers, followed by fully linked layers. In contrast, transfer learning models use pre-trained weights from ImageNet to fine-tune the top layers for traffic sign identification. We implement and fine-tune the ASNet, EfficientNet, and VGG19 architectures utilizing transfer learning. Several variables are likely to have impacted their decision to use ASNet, EfficientNet, and VGG19 for transfer learning in traffic sign identification, with each model providing specific capabilities tailored to the task:

ASNet: ASNet, or Attention-guided Spatial Attention Network, is likely chosen for its capacity to capture spatial attention, which is critical for identifying traffic signs in complicated environments. ASNet's design is anticipated to include techniques for focusing on key portions of the input image, which improves its capacity to reliably detect traffic signals.

EfficientNet: EfficientNet is well-known for its resource consumption and parameter optimization efficiency, making it an appealing alternative for traffic sign identification jobs, especially in circumstances with restricted computing resources. Its scalable design, which balances model size and speed, enables efficient transfer learning across several datasets while retaining excellent accuracy.

VGG19: VGG19, a version of the VGG family, was chosen for its ease of use and efficacy in feature extraction. Its simple design, which consists of a succession of convolutional layers with tiny receptive fields, allows it to learn hierarchical characteristics from traffic sign photos quickly. Furthermore, its extensive use and well-documented performance on picture classification tasks make it a viable alternative for transfer learning in traffic sign identification.

Each of these models brings unique strengths to the table, with ASNet concentrating on spatial attention, EfficientNet on resource efficiency, and VGG19 excelling at feature extraction. By utilizing the individual capabilities of these models through transfer learning, the project intends to boost the accuracy and efficacy of traffic sign detecting systems, eventually leading to enhanced road safety and intelligent transportation.

Convolutional Layer: CNN, or convolutional neural network:

Deep neural networks, or CNNs, are made especially to process and evaluate visual input. They are made up of many layers using pooling and convolutional techniques, then fully linked layers. A CNN may be trained from scratch or refined using a particular dataset including pictures of traffic signs in the context of traffic sign detection. However, a significant quantity of labeled data and computer power are needed to train a CNN from start.

CNNs are suited for applications like object identification and image classification because of their well-known capacity to automatically extract features from unprocessed pixel data. Maths behind the CNN is explained below

$$Z^{[l]} = W^{[l]} * A^{[l-1]} + b^{[l]}$$

Activation Function:

$$A^{[l]} = g(Z^{[l]})$$

Pooling Layer:

$$A^{[l]} = \text{Pooling}(A^{[l-1]})$$

Fully Connected Layer:

$$Z^{[l]} = W^{[l]} \cdot A^{[l-1]} + b^{[l]}$$

Output Layer (Softmax):

$$A^{[L]} = \text{Softmax}(Z^{[L]})$$

ASNet Attention Module: ASNet is a transfer learning architecture that uses attention processes to focus on key areas of a picture. It extracts features using pre-trained convolutional layers (typically from ResNet or VGG models) and then adds attentional modules to enhance the features and attend to key sections of the input. ASNet is especially beneficial for applications that require attention on specific portions of an image, such as item recognition in congested situations or fine-grained classification tasks.

ASNet may outperform a normal CNN by selectively attention to informative parts of traffic signals, perhaps resulting in higher detection accuracy.

Maths behind the ASNet is explained below

$$M = \text{softmax}(W_a \cdot f(A^{[l-1]}))$$

Attention Applied:

$$A^{[l]} = M \odot A^{[l-1]}$$

EfficientNet:

EfficientNet is a class of convolutional neural network designs that have been intended to perform better with fewer parameters and computations than conventional CNNs. It uses a compound scaling approach to equally scale network breadth, depth, and resolution using a set of predetermined scaling coefficients. EfficientNet models have proven cutting-edge performance across a variety of computer vision tasks while being computationally cheap, making them ideal for deployment on resource-constrained devices or in settings where computing resources are restricted. In the area of traffic sign detection, EfficientNet may provide equivalent or better performance than classic CNNs while using fewer computer resources for training and inference.

$$d = \alpha^{\phi}, \quad w = \beta^{\phi}, \quad r = \gamma^{\phi}$$

Compound Scaling Coefficients:

$$\alpha, \beta, \gamma, \phi$$

VGG19 Convolutional Layer: VGG19 is a deep convolutional neural network architecture with 19 layers created by the Visual Geometry Group at the University of Oxford. It is recognized for its simple and consistent design, which includes tiny (3x3) convolutional filters and max-pooling layers. VGG19 has been frequently employed as a feature extractor or backbone network in a variety of transfer learning applications due to its ability to extract hierarchical features from picture data. Compared to more contemporary designs like as EfficientNet, VGG19 may have poorer computational efficiency but can still deliver competitive performance, particularly when fine-tuned for specialized applications such as traffic sign identification.

$$Z^{[l]} = W^{[l]} * A^{[l-1]} + b^{[l]}$$

A. *Results*

To verify that the presented findings were reliable, the models were verified using a variety of methodologies, including cross-validation and hold-out validation. Cross-validation entails partitioning the dataset into many subgroups and training the model on various combinations of these subsets. This allows us to analyze how the model performs across different data divisions, lowering the danger of overfitting and offering a more accurate estimate of its generalization performance.

Furthermore, hold-out validation was most likely used, in which a piece of the dataset is designated as a validation set distinct from the training and testing data. The model is trained on the training set, and its performance is evaluated on the validation set to fine-tune hyperparameters and measure performance before being tested on the independent test set. The reported results show that all models are highly accurate, with CNN attaining the best accuracy of 99.7163%. Furthermore, the relatively low loss values and validation losses across all models indicate that they were trained efficiently and can generalize well to new data.

By using rigorous validation procedures such as cross-validation and hold-out validation, the stated results are more likely to be robust and dependable, adding credibility to the study's conclusions.

Comparison of Different Model Score

Model	Accuracy	Loss	Validation Loss
CNN	0.997163	0.010	0.24
VGG model	0.9964	0.0163	0.6232
EfficientNetB7 Model	0.9982	0.0084	3.5862
NASNetLarge Model	0.9946	0.0260	4.1889

Fig. 2. Performance of CNN Training Data

Fig. 3. VGG prediction Sample

IV. Conclusion

The examination of various convolutional neural network (CNN) architectures for traffic sign identification revealed notable differences in performance. VGG exhibited a comparatively low test accuracy of 18.16%, while EfficientNet demonstrated improved performance at 21.51%. NASNet, however, achieved the lowest accuracy at 6.22%, indicating significant variability in model effectiveness. These results underscore the substantial impact of design choices on the efficacy of traffic sign detection systems.

EfficientNet's superior performance over VGG highlights the importance of leveraging sophisticated architectures for enhanced accuracy, particularly with EfficientNet's noted efficiency in resource consumption and parameter optimization. However, the substantial performance gap between EfficientNet and NASNet underscores the significance of not only

Fig. 4. EfficientNet Predicating sample

architecture but also training approach and dataset quality.

Remarkably, CNN outperformed all other models, achieving a test accuracy of 96.39%, suggesting its robustness in traffic sign identification tasks.

Concluding the results section, it's evident that further research is essential for refining transfer learning approaches and dataset curation strategies to enhance the accuracy and resilience of traffic sign identification systems. This will be crucial for advancing the reliability and safety of self-driving technology, contributing to the evolution of intelligent transportation networks. Future work should focus on exploring novel architectures, improving dataset diversity, and refining training methodologies to address the identified limitations and bolster system performance.

References

[1] Babiker, M., Elawad, M., and Ahmed, A. Convolutional neural network for a self-driving car in a virtual environment. pp. 1–6.

[2] Chen, T., Chen, Z., Shi, Q., and Huang, X. Road marking detection and classification using machine learning algorithms. In *2015 IEEE Intelligent Vehicles Symposium (IV)* (2015), IEEE, pp. 617–621.

[3] De-Las-Heras, G., Sanchez-Soriano, J., and Puertas, E. Advanced driver assistance systems (adas) based on machine learning techniques for the detection and transcription of variable message signs on roads. *Sensors 21*, 17 (2021), 5866.

[4] Dollar, P., Wojek, C., Schiele, B., and Perona, P. Pedestrian detection: An evaluation of the state of the art. *IEEE transactions on pattern analysis and machine intelligence 34*, 4 (2011), 743–761.

[5] Gandhi, T., and Trivedi, M. M. Computer vision and machine learning for enhancing pedestrian safety. In *Computational intelligence in automotive applications*. Springer, 2008, pp. 59–77.

[6] Hussain, R., and Zeadally, S. Autonomous cars: Research results, issues, and future challenges. *IEEE Communications Surveys Tutorials 21*, 2 (2019), 1275–1313.

[7] Jordan, M. I., and Mitchell, T. M. Machine learning: Trends, perspectives, and prospects. *Science 349*, 6245 (2015), 255–260.

[8] Kariluoto, A., Kultanen, J., Soininen, J., Pärnänen, A., and Abrahamsson, P. Quality of data in machine learning. In *2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C)* (2021), pp. 216–221.

[9] Mao, J., Xiao, T., Jiang, Y., and Cao, Z. What can help pedestrian detection? In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (July 2017).

[10] OUYANG, W., AND WANG, X. Single-pedestrian detection aided by multi-pedestrian detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2013), pp. 3198–3205.

[11] QIU, J., LIU, J., AND SHEN, Y. Computer vision technology based on deep learning. In *2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)* (2021), vol. 2, pp. 1126–1130.

[12] SILVA, P. B., ANDRADE, M., AND FERREIRA, S. Machine learning applied to road safety modeling: A systematic literature review. *Journal of traffic and transportation engineering (English edition) 7*, 6 (2020), 775–790.

[13] SUN, P., ALJERI, N., AND BOUKERCHE, A. Machine learning-based models for real-time traffic flow prediction in vehicular networks. *IEEE Network 34*, 3 (2020), 178–185.

[14] SUNJAYA, E., KURNIADI, DANIEL, C., KURNIAWAN, A., AND ANGGREAINY, M. S. A systematic literature review: Object detection in semi-autonomous car. In *2023 6th International Conference of Computer and Informatics Engineering (IC2IE)* (2023).

[15] THAMMACHANTUEK, I., KOSOLSOMNBAT, S., AND KETCHAM, M. Comparison of machine learning algorithm's performance based on decision making in autonomous car. In *2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP)* (2018), IEEE, pp. 1–6.

[16] TIAN, J., YUAN, J., AND LIU, H. Road marking detection based on mask r-cnn instance segmentation model. In *2020 international conference on computer vision, image and deep learning (CVIDL)* (2020), IEEE, pp. 246–249.

[17] WANG, Y. Decision making for autonomous vehicle at single-lane road under uncertainties. In *2021 International Conference on Signal Processing and Machine Learning (CONF-SPML)* (2021), IEEE, pp. 250–253.

[18] ZHENG, Y., MENG, H., AND JIA, W. Application research and challenges of artificial intelligence in primary and secondary education. In *2022 12th International Conference on Information Technology in Medicine and Education (ITME)* (2022), pp. 162–166.

[19] ZHU, Z., HU, Z., DAI, W., CHEN, H., AND LV, Z. Deep learning for autonomous vehicle and pedestrian interaction safety. *Safety science 145* (2022), 105479.