Comparative Study of Machine Learning Algorithms for Traffic Sign Recognition

Kavita Kolpe[1], Sumit Tekam[2], Shubham Waghmare[3], Shantanu Yadav[4], Ashuthosh Talikhede[5]

[1]*Computer Engineering, Pimpri Chinchwad College of Engineering, Pune*
[2] *Computer Engineering, Pimpri Chinchwad College of Engineering, Pune*
[3] *Computer Engineering, Pimpri Chinchwad College of Engineering, Pune*
[4] *Computer Engineering, Pimpri Chinchwad College of Engineering, Pune*
[5] *Computer Engineering, Pimpri Chinchwad College of Engineering, Pune*

---***---

Abstract - The increasing integration of intelligent transportation systems has brought forth a critical need for accurate and efficient traffic sign detection algorithms. In this research, we conduct a comprehensive comparative analysis of four prominent convolutional neural network (CNN) architectures—LeNet, ResNet, VGG, and a standard CNN model—for their performance in traffic sign detection. The evaluation is conducted on a diverse dataset, utilizing the German Traffic Signs Recognition Benchmark (GTSRB). Our study encompasses an in-depth examination of accuracy metrics and performance under varying conditions, such as challenging lighting, occlusion, and environmental complexities. The results demonstrate nuanced differences in accuracy and computational efficiency among the algorithms. LeNet exhibits notable improvements in real-time processing, while ResNet showcases enhanced accuracy, and VGG demonstrates robustness in challenging conditions. The findings provide valuable insights into the strengths and weaknesses of each algorithm, aiding researchers and practitioners in selecting the most suitable model for specific traffic sign detection applications. This research contributes to advancing the field of computer vision in intelligent transportation systems, with implications for enhancing road safety and efficiency.

Key Words: **traffic sign recognition, deep learning, Res-Net, VGG19, CNN, LeNet.**

1.INTRODUCTION

In the ever-expanding domain of computer vision and pattern recognition, the pursuit of adept algorithms for traffic sign detection has catalyzed a myriad of advancements and comparisons. This research embarks on a comprehensive exploration, meticulously examining four prominent algorithms—Convolutional Neural Network (CNN), LeNet, ResNet, and VGG—specifically tailored for the critical task of traffic sign recognition. The central objective is to conduct an in-depth comparative analysis, spotlighting the unique strengths, weaknesses, and accuracies inherent in each algorithm. As the landscape of road infrastructure undergoes constant transformation, the imperative for precise and swift traffic sign interpretation persists. Through a thorough examination of these algorithms across diverse conditions and scenarios, the study endeavors to unravel insights that contribute to the ongoing evolution of traffic sign recognition technology. By dissecting the intricacies of algorithmic performance, the research aims to provide nuanced guidance for the selection and implementation of these models in broader computer vision applications, with a specialized emphasis on traffic sign detection.

II. RELATED WORK

The road traffic sign recognition algorithm based on the Improved VGG (IVGG) convolutional neural network demonstrates significant advantages. The IVGG model surpasses the VGG_16 model, exhibiting notable improvement in recognition rates, particularly evident in the German Traffic Sign Recognition Benchmark dataset. The inclusion of max-pooling, dropout operations, and strategic use of data augmentation and transfer learning contributes to a robust model with efficient feature extraction, ultimately reducing training time.

Nevertheless, challenges persist, primarily in scenarios of low light and motion blur, leading to a decline in recognition accuracy. The IVGG model, while effective, may necessitate further refinement to address these limitations.

From a methodological standpoint, the study meticulously designs the IVGG model, incorporating specific layers and operations tailored for traffic sign recognition. The utilization of data augmentation and transfer learning effectively addresses issues related to data scarcity. Through comprehensive experiments on the German Traffic Sign Recognition Benchmark dataset, the study provides a thorough evaluation of the model's accuracy, recognition rates, and training time. While the IVGG model demonstrates significant advantages, the study recognizes the imperative for future enhancements, particularly in addressing challenges associated with dark backgrounds and motion blur in traffic sign images.[1]

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

This paper delves into traffic sign recognition using pre-trained deep Convolutional Neural Network (CNN) models—VGG16, VGG19, AlexNet, and Resnet50. The research addresses challenges in computer vision and intelligent transportation systems, leveraging the efficiency of deep learning in image-related tasks. These pre-trained models, chosen for their success in feature extraction and classification, undergo customization to optimize performance for traffic sign recognition. The study recognizes challenges such as lighting variations and occlusion but does not explicitly address issues like motion blur or low-light conditions.

Methodologically, the research employs transfer learning, adjusting hyperparameters for each model. The German Traffic Sign Recognition Benchmark dataset (GTSRB) serves as the standardized benchmark for training and testing. Evaluation metrics include accuracy, precision, recall, and F1-score, ensuring a comprehensive assessment of the models' proficiency in recognizing traffic signs. While the study provides a robust methodology, addressing advantages, challenges, and employing relevant metrics, it might benefit from a more explicit consideration of real-world complexities.[2]

The academic paper titled "Traffic Sign Detection and Recognition Based on Convolutional Neural Network" by Ying Sun, Pingshu Ge, and Dequan Liu, affiliated with Dalian Minzu University in China, proposes a systematic approach to traffic sign detection. The methodology involves image preprocessing with enhancement and color space conversion, Hough Transform for sign detection based on color and shape, and a Convolutional Neural Network (CNN) for classification. The paper demonstrates high accuracy, exceeding 98.2%, particularly in identifying circular symbols in German datasets. The effectiveness of CNNs in various computer vision tasks is highlighted. Additionally, the integration of the Hough Transform contributes to accurate and timely circular road sign detection. Despite these advantages, the absence of a publication year and specific disadvantages, along with a limited literature survey, leaves some gaps in the overall understanding of the research context and potential drawbacks in the proposed methodology.[3]

The research introduces a novel approach to Traffic Sign Detection and Recognition using a modified Mask R-CNN model. This model incorporates additional features such as shape detection, color probability based on lighting conditions, histogram matching, and optical character recognition (OCR). The proposed methodology exhibits robust performance across diverse conditions, demonstrating resilience to variations in light, orientation, and scale. Integration of data augmentation techniques enhances the model's adaptability to real-world scenarios, while real-time dataset customization ensures relevance to Indian road conditions. However, challenges arise in the model's sensitivity to image quality variations, particularly with dirty or unclean traffic signs, and its limited applicability in scenarios with wide viewing angles. Strategies to mitigate errors, including the use of stereo cameras, are suggested but may add complexity and cost to implementation. Notably, the paper lacks specific accuracy metrics, and a comprehensive literature survey should be conducted to assess the strengths and weaknesses of existing Traffic Sign Recognition models and techniques in the field.[4]

This research delves into advanced methodologies for traffic sign detection using computer vision, tackling issues such as sensor errors and dataset imbalances. Employing techniques like brightness histograms, resampling, and data augmentation ensures noise reduction and improved dataset quality. Real-time image processing through OpenCV enhances overall image quality, while color-based detection and circular object recognition contribute to effective traffic sign identification, particularly focusing on red-circle delineation. The neural network architecture, a Convolutional Neural Network (CNN) with four layers, adeptly captures spatial and temporal dependencies for efficient image classification, utilizing SoftMax activation functions. While the research acknowledges limitations in generalizing to diverse sign characteristics and real-time capture challenges in adverse weather conditions, its strengths lie in smart weather condition analysis and a recognition system tailored to global sign diversity, emphasizing the importance of a generalized model in this domain.

This study focuses on the advancement of traffic sign recognition systems, vital for ensuring road safety. The researchers propose an improved Convolutional Neural Network (CNN) architecture based on the LeNet-5 model, aiming to enhance the accuracy of road sign classification. The methodology incorporates various image processing techniques, including grayscale conversion and histogram equalization, to improve image quality. The model achieves notable accuracy rates of 99.84% on the German Traffic Sign Recognition Benchmark (GTSRB) and 98.37% on the Belgian Traffic Sign Data Set (BTSD). The architectural enhancements, such as increased filter numbers, modified fully connected layers, and the incorporation of Leaky Elu activation functions, contribute to a lightweight design, essential for embedded applications. The research emphasizes the advantages of high accuracy, reduced parameters, and efficient training times, positioning the proposed model as a robust solution. While acknowledging challenges like dataset imbalances and real-time classification complexities, the study effectively contextualizes its contributions within existing literature. Overall, the research presents a substantial contribution to the field of traffic sign recognition, offering a blend of theoretical and practical advancements with promising applications in real-world scenarios, as demonstrated through successful webcam-based road sign classification.

The research presents an enhanced LeNet-5 model for the detection and recognition of road traffic signs, aiming to address challenges arising from intricate outdoor conditions. The improvements are centered on optimizing the LeNet-5 architecture by incorporating a 5×5 convolution kernel size, ReLU activation function, and dropout mechanism to enhance real-time performance and prevent overfitting. A notable innovation is the introduction of a multi-layer feature fusion method, effectively leveraging features from various layers of the convolutional neural network (CNN) to boost recognition accuracy. Utilizing the German Traffic Signs Data Set (GTSRB), the model achieves a recognition accuracy of 98.82%, outperforming traditional LeNet-5 and a single-scale feature approach. The methodology involves meticulous preprocessing of the dataset, addressing imbalances, and incorporating data augmentation. The results underscore the robustness and efficiency of the proposed model in real-world traffic sign recognition scenarios, contributing to the advancement of intelligent transportation systems. Despite these accomplishments, potential limitations, such as the necessity for diverse dataset validation and further exploration of real-world conditions, are acknowledged, suggesting avenues for future research and refinement of the proposed methodology. The study marks a significant advancement in enhancing the practical applicability of classical CNN architectures for intelligent transportation systems and advanced driver assistance systems.[7]

The paper titled "Application of Improved LeNet-5 Network in Traffic Sign Recognition" by Wenlong Li et al. tackles the challenges associated with complex convolutional neural network (CNN) architectures for traffic sign recognition (TSR) by enhancing the simplicity of LeNet-5. The study introduces modifications, including increased convolution kernel numbers in specific layers, the incorporation of Rectified Linear Unit (ReLU) activation function, and the replacement of mean pooling with maximum pooling. Additionally, a support vector machine (SVM) classifier is used in the output layer to reduce computation time. The proposed improved LeNet-5 achieves a high identification accuracy of 98.12% on the German Traffic Sign Recognition Benchmark (GTSRB), demonstrating real-time efficiency with an identification time of 0.154s. The research contributes to the field by combining the advantages of LeNet-5's simplicity with performance improvements, making it suitable for TSR applications. However, the paper lacks a detailed literature survey, and while it highlights the accuracy and efficiency benefits, a comprehensive discussion on potential disadvantages or limitations of the proposed methodology is not provided. Further exploration of the trade-offs and comparisons with other state-of-the-art methods would enhance the paper's contribution to the field.[8]

The paper "Real-time Traffic Sign Recognition System with Deep Convolutional Neural Network" tackles the crucial challenge of Traffic Sign Recognition (TSR) within Advanced Driver Assistance Systems, utilizing a LeNet-5 convolutional neural network (CNN) for image classification. Addressing the need for diverse traffic sign recognition in situations where digital map data may be inadequate, the authors leverage the multilayer perceptron design of LeNet-5 with shared weights and biases for its effectiveness in image classification. The methodology incorporates a light-weight color-based segmentation algorithm and Hough transform to extract candidate regions, aiming for real-time performance. Despite successful on-road tests and integration into an autonomous vehicle, the paper lacks a comprehensive literature survey on LeNet-5's accuracy in TSR, its advantages, and potential disadvantages. Architectural details, including the CNN's layer count and configurations, remain insufficiently detailed. A more thorough analysis of the model's strengths and limitations, along with a comparison to alternative approaches, would enhance the paper's contribution to the field.

The paper "Traffic Sign Recognition Based on ResNet-20 and Deep Mutual Learning" introduces a robust methodology for traffic sign recognition (TSR) leveraging ResNet-20 models and Deep Mutual Learning. The methodology yields impressive results, achieving a Top-1 accuracy of 99.612% and Top-5 accuracy of 99.952%, showcasing superior performance in comparison to mainstream TSR methods. The incorporation of Deep Mutual Learning during training enhances individual sub-networks' performance, leading to models with heightened robustness. Furthermore, the model fusion technique optimizes prediction accuracy without significantly increasing the parameter count, resulting in a lightweight model with 271,867 parameters. Despite these strengths, the paper has limitations limited exploration of future works, potentially hindering a comprehensive understanding of the proposed approach's broader context and future advancements. Additionally, considerations regarding computational resources, training complexity, and real-time implementation challenges associated with ResNet-20 models and Deep Mutual Learning are not extensively discussed. Addressing these aspects and conducting more extensive experiments, especially in dynamic TSR scenarios, would contribute to the methodology's applicability and scalability in real-world, autonomous driving systems.[10]

III. DATASET DESCRIPTION

The German Traffic Sign Benchmark, hosted at the International Joint Conference on Neural Networks (IJCNN) in 2011, extends an inclusive invitation to researchers from diverse fields, eliminating the need for specialized domain knowledge. This benchmark sets the stage for a single-image, multi-class classification challenge involving over 40 distinct classes and a substantial dataset exceeding 50,000 lifelike images. The competition's large, realistic database ensures its suitability for broad participation. It provides a valuable opportunity for researchers to showcase their skills and contribute innovative solutions to the dynamic field of traffic sign recognition. With its extensive dataset, the German Traffic Sign Benchmark emerges as a pivotal testing ground for advancements in image classification techniques. This platform fosters collaborative exploration and encourages breakthroughs in the domain, making it an ideal space for researchers to collectively propel the field forward.

Dataset Link:

https://www.kaggle.com/datasets/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign

Fig-1: Sample images of Dataset

IV.METHODOLOGY

1.Convolutional Neural Network (CNN): A Convolutional Neural Network is a class of deep neural networks specifically designed for image recognition and processing. It is inspired by the human visual system and utilizes a series of convolutional layers to automatically and adaptively learn spatial hierarchies of features from input images. The key components of a CNN include convolutional layers, pooling layers, and fully connected layers. Convolutional layers apply filters to the input data, capturing different features through convolution operations. Pooling layers reduce spatial dimensions, preserving essential information. Fully connected layers make predictions based on the learned features. CNNs excel in image-related tasks, such as object detection, classification, and segmentation, and have been widely used in various computer vision applications.

2.VGG19: The part of the Visual Geometry Group network, is characterized by its simplicity and uniform architecture. It comprises 19 layers, including 16 convolutional and 3 fully connected layers. VGG uses small 3x3 convolutional kernels consistently, ReLU as the activation function, and applies max pooling in its pooling layers. VGG networks have proven versatile and successful in various computer vision tasks.

3. LeNet: The architecture, developed by Yann LeCun, is a pioneering Convolutional Neural Network (CNN) consisting of seven layers. It includes three convolutional layers, two subsampling (pooling) layers, and two fully connected layers. The primary activation function used in LeNet is sigmoid, and average pooling is employed in its subsampling layers. Originally designed for handwritten digit recognition, LeNet laid the foundation for more complex CNN architectures.

4.ResNet: introduced by Kaiming He et al., brought the concept of residual learning, incorporating skip connections or shortcuts to facilitate information flow across layers. The key feature is the residual block, which mitigates the vanishing gradient problem. ReLU is the chosen activation function within each residual block. ResNet has been influential in enabling the training of very deep networks.

Comparative Analysis: The comparative analysis of the trained algorithms (CNN, LeNet, ResNet, VGG19) on the GTSRB dataset for 32 epochs reveals distinct performance characteristics.

ResNet achieved the highest accuracy of 96.20%, showcasing its proficiency in capturing intricate patterns and features within traffic sign images. This result aligns with ResNet's known ability to handle deep networks effectively, making it well-suited for complex image recognition tasks.

CNN also demonstrated strong performance with an accuracy of 93.40%, highlighting its effectiveness in learning relevant features from the dataset. CNNs are known for their capability to automatically learn hierarchical features, making them suitable for image classification tasks.

LeNet, while performing reasonably well with an accuracy of 89.91%, showed a slightly lower accuracy compared to ResNet and CNN. LeNet, being an early convolutional network architecture, may benefit from more sophisticated structures for improved performance.

VGG19, however, exhibited a comparatively lower accuracy of 72.35%, indicating potential challenges in capturing and recognizing complex patterns within the GTSRB dataset. VGG architectures are known for their simplicity, and this result suggests that for this specific task, a more intricate model may be necessary.

In summary, ResNet and CNN outperformed LeNet and VGG19 in terms of accuracy on the GTSRB dataset. Further investigations could explore fine-tuning and optimization strategies for each model to enhance their performance in traffic sign recognition.

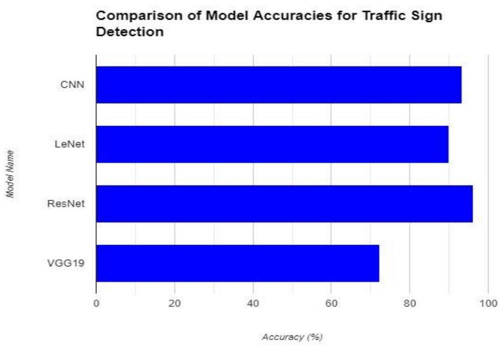

Fig-2: Bar Graph

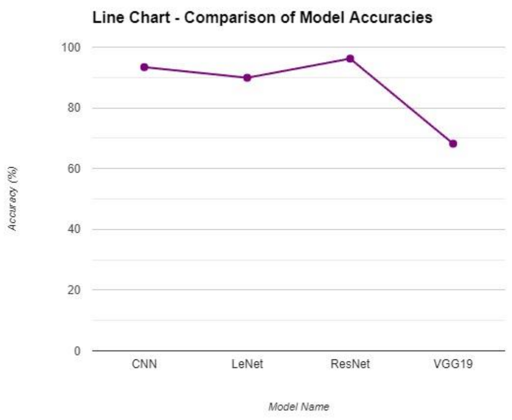

Fig-3: Line Graph

V. CONCLUSION

Conclusively, after evaluating the performance of four algorithms (CNN, ResNet, LeNet, VGG19) trained on the GTSRB dataset for 32 epochs, it is evident that ResNet achieved the highest accuracy of 96.20%. Following ResNet, CNN attained an accuracy of 93.40%. LeNet showcased competitive performance with an accuracy of 89.91%, while VGG19 yielded the lowest accuracy at 72.35%. This underscores ResNet's effectiveness in classifying traffic sign images, possibly attributed to its deep architecture with skip connections, facilitating better training of deeper networks by addressing the vanishing gradient problem. These findings highlight the importance of selecting an appropriate neural network architecture tailored to the specific dataset and task at hand to achieve optimal performance in image classification tasks.

VI. REFERENCES

[1] Zhou, S., Liang, W., Li, J. and Kim, J.U., 2018. Improved VGG model for road traffic sign recognition. *Computers, Materials & Continua*, *57*(1), pp.11-24.

[2] Khandalkar, S.A., 2022. *Traffic Sign Detection Using Deep Learning Algorithms* (Doctoral dissertation, Dublin, National College of Ireland).

[3] Sun, Y., Ge, P. and Liu, D., 2019, November. Traffic sign detection and recognition based on convolutional neural network. In *2019 Chinese automation congress (CAC)* (pp. 2851-2854). IEEE.

[4] Megalingam, R.K., Thanigundala, K., Musani, S.R., Nidamanuru, H. and Gadde, L., 2023. Indian traffic sign detection and recognition using deep learning. *International Journal of Transportation Science and Technology*, *12*(3), pp.683-699.

[5] Singh, K. and Malik, N., 2022. CNN based approach for traffic sign recognition system. *Advanced Journal of Graduate Research*, *11*(1), pp.23-33.

[6] Bouti, A., Mahraz, M.A., Riffi, J. and Tairi, H., 2020. A robust system for road sign detection and classification using LeNet architecture based on convolutional neural network. *Soft Computing*, *24*(9), pp.6721-6733.

[7] Zaibi, A., Ladgham, A. and Sakly, A., 2021. A lightweight model for traffic sign classification based on enhanced LeNet-5 network. *Journal of Sensors*, *2021*, pp.1-13.

[8] Li, W., Li, X., Qin, Y., Song, W. and Cui, W., 2019, December. Application of improved LeNet-5 network in traffic sign recognition. In *Proceedings of the 3rd International Conference on Video and Image Processing* (pp. 13-18).

[9] Jung, S., Lee, U., Jung, J. and Shim, D.H., 2016, August. Real-time Traffic Sign Recognition system with deep convolutional neural network. In *2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)* (pp. 31-34). IEEE.

[10] Huo, T., Fan, J., Li, X., Chen, H., Gao, B. and Li, X., 2020, November. Traffic sign recognition based on resnet-20 and deep mutual learning. In *2020 Chinese Automation Congress (CAC)* (pp. 4770-4774). IEEE.