An Ensembled Real-Time Hand-Gesture Recognition using CNN

Kavitha M N[1]
Department of Computer Science and Design
Kongu Engineering College
Erode, TamilNadu, India
kavithafeb1@gmail.com

Saranya S S[2]
Department of Computer Science and Engineering
PSG Institute of Technology and Applied Research
Coimbatore, TamilNadu, India
saranya.sowndarajan@gmail.com

Prasad M[3]
Controller of Examinations
K S Rangasamy College of Arts and Science (Autonomous)
Namakkal, TamilNadu, India
mprasadkrishna@gmail.com

Kaviyarasu S[4]
Bachelor of Science – Information Systems
Kongu Engineering College
Erode, TamilNadu, India
kaviyarasus.21bir@kongu.edu

Ragunath N[5]
Bachelor of Science – Information Systems
Kongu Engineering College
Erode, TamilNadu, India
ragunathn.21bir@kongu.edu

Rahul P[6]
Bachelor of Science – Information Systems
Kongu Engineering College
Erode, TamilNadu, India
rahulp.21bir@kongu.edu

2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT) | 979-8-3503-7024-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICCCNT61001.2024.10725948

***Abstract*— Hand sign recognition is a vital technology in the human-computer interaction, enabling individuals to communicate with machines naturally and effectively. An innovative approach for real-time hand sign identification with the help of CNN and OpenCV is introduced with the fusion of computer vision and deep learning that can accurately interpret and classify an extensive range of hand signs and gestures. This research contributes significantly to the fields of computer vision and human-computer interaction, offering a practical and efficient solution for hand sign recognition. The combination of CNN and OpenCV presents a promising avenue for enhancing accessibility and communication, especially in environments where verbal communication is limited or non-existent. The model is trained with multiple data so that the system can recognize the hand gestures more precisely. Pre-trained architectures like ResNet and MobileNet are combined with the CNN model using ensemble learning and the performance is improved when compared to all the three CNN architectures individually. The ensemble model provides better accuracy of 96%. The potential applications of this technology are vast, from assisting the hearing-impaired in understanding sign language to more immersive and intuitive interactions. Overall, the approach holds the promise of bridging the gap between human gestures and machine understanding, opening new doors for meaningful interactions between individuals and intelligent systems.**

Keywords—Hand Sign Recognition, OpenCV, Convolutional Neural Network (CNN), Deep Learning, MobileNet, ResNet, Ensemble learning

I. INTRODUCTION

Hand sign recognition, a technology rooted in the evolution of computer vision and machine learning, has come a long way since its inception. Early research on hand shape and gesture recognition took place between the 1960s and the 1990s. In the 1990s and 2000s, as machine learning became more popular, statistical methods such as Hidden Markov Models (HMMs) were developed to simulate gesture sequences. But the real revolution in the field occurred in the 2010s with the emergence of deep learning, which allowed for robust and flexible recognition in unstructured environments. Large-scale, annotated hand sign datasets were developed, which further speed up the process.

These days, hand sign recognition technology is widely used across many industries. It makes communication between sign language users and non-users more seamless by facilitating sign language translation. Digital device interaction is made easier for people with hearing problems because of improved accessibility. Virtual reality platforms, smart TVs, and game consoles may all be operated with hand gestures, which enhances user experience. In healthcare, it aids rehabilitation exercises, and in robotics, it improves human-robot interaction. It also has uses in security, education, and other areas, indicating that the field will likely continue to develop and innovate. In simpler terms, hand sign recognition is like teaching a computer to understand the secret language of hand gestures, making it easier for people to interact with machines in a more natural and intuitive way.

Deep learning is a subfield of machine learning and artificial intelligence that aims to develop artificial neural networks capable of performing tasks requiring human-like intelligence. Deep learning models, also called deep neural networks consists of multiple layers of linked nodes, or neurons that can automatically extract and represent structured information from data. These models are useful for solving difficult problems in many domains, including natural language processing, computer vision, and speech recognition. Its feature learning ability, which enables it to understand and decode data in a way that improves performance across numerous domains, is one of its fundamental advantages. Large datasets are also ideal for deep learning, which uses them to increase the accuracy and generalization of models. Models are able to adjust to themselves during training through the use of gradient descent and backpropagation, which produces improved outcomes.. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are popular deep learning architectures which are advanced fields like computer vision and natural language processing.

Furthermore, a broad range of applications may be accommodated by deep learning models- from recommendation systems and autonomous cars to picture identification and speech understanding - makes them highly versatile. Because of its versatility, constant evolution, and ongoing study, deep learning is regarded as a key element of today's artificial intelligence, pushing the limits of what AI systems are capable of. Deep learning techniques offers good accuracy and performance for various recognition techniques. This project focuses on hand-gesture recognition using deep learning techniques

Fig.4. Proposed Model for real-time hand gesture recognition

IV. Results

The proposed model was executed using Python in PyCharm environment with GPU support. The image extracted from the video is preprocessed as shown in the figure 6. These images are fed into the ensemble model which classifies it as 'thumbs' correctly as shown in figure 5. Other gestures like Ok, Fist, Not Ok, None, Five, etc were also classified accurately.

Fig. 5. Hand gestures captured through web camera classified as 'thumbs'

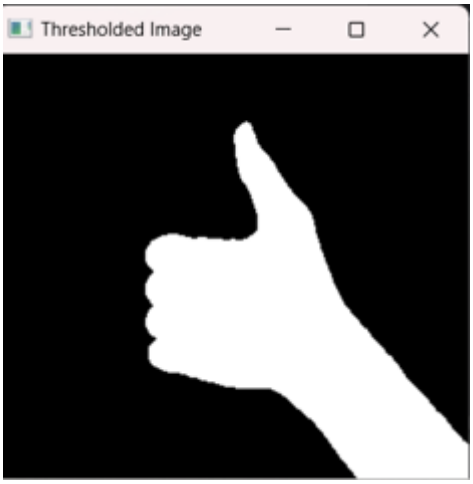

Fig.6. Preprocessed hand gesture image

The ensemble model produces better performance in terms of accuracy, precision, recall and f-measure and are compared and tabulated for the individual models as shown in the Table. 1.

Table 1. Comparison of performance metrics for CNN, ResNet, MobileNet and the Proposed Model

Model	Performance Metrics (in %)			
	Accuracy	Precision	Recall	F-measure
CNN	91.3	89.2	87.6	90.2
ResNet	91.2	91.3	89.5	92.3
MobileNet	92.4	91.5	91.2	92.5
Proposed Ensembled Model	96.5	92.4	94.5	95.5

V. Conclusion and Future Enhancement

Effective human-machine interaction can be enhanced by using a good hand gesture recognition system. Many researchers have came with different methods for hand gesture recognition. The proposed method using ensemble learning combines the results obtained from CNN, MobileNet and ResNet and offers a better model. This model out performs the individual model performance in terms of accuracy, recall, f-measure and precision as shown in the Table.1. The performance can be further enhanced by using different models in ensemble learning. Hyperparameter optimization also can be included to obtain better hand gesture classification.

References

[1] Munir Oudah , Ali Al-naji And Javaan Chahl,"hand Gesture Recognition Based On Computervision: A Review Of Techniques", 2020.

[2] X. Jiang And W. Ahmad, "hand Gesture Detection Based Real-time American Sign Language Letters Recognition Using Support Vector Machine", 2019, Pp. 380-385.

[3] D. S, K. H. K B, A. M, S. M, D. S And K. V, "an Efficient Approach For Interpretation Of Indian Sign Language Using Machine Learning", 2021, Pp. 130-133.

[4] Parvathy, P., Subramaniam, K., Venkatesan, G.k.d.p. Et Al. " Development Of Hand Gesture Recognition System Using Machine Learning". J Ambient Intell Human Comput 14 (suppl 1), 517 (2023).

[5] H. N. Saha, S. Tapadar, S. Ray, S. K. Chatterjee And S. Saha, "a Machine Learning Based Approach For Hand Gesture Recognition Using Distinctive Feature Extraction", 2018, Pp. 91-98.

[6] Neethu, P.s., Suguna, R. & Sathish, D. An Efficient Method For Human Hand Gesture Detection And Recognition Using Deep Learning Convolutional Neural Networks. Soft Comput 24, 15239–15248 (2020).

[7] Oyedotun, O.k., Khashman, A. Deep Learning In Vision-based Static Hand Gesture Recognition. Neural Comput & Applic 28, 3941–3951 (2017).

[8] Tan, Y.s., Lim, K.m., Tee, C. Et Al. Convolutional Neural Network With Spatial Pyramid Pooling For Hand Gesture Recognition. Neural Comput & Applic 33, 5339–5351 (2021).

[9] Indriani Moh.harris Ali Suryaperdana Agoes,"applying Hand Gesture Recognition For User Guide Application Using Mediapipe",2021.

[10] Sahoo, A. K. "indian Sign Language Recognition Using Machine Learning Techniques". 397(1), (2021).

[11] Pose Andre Jallen S.ong,melvin Cabatuan,janos Lance L. Tiberio, John Anthony Jose," Lstm-based Traffic Gesture Recognition Using Mediapipe",2022.

[12] S. Bouguezzi, H. Faiedh And C. Souani, "slim Mobilenet: An Enhanced Deep Convolutional Neural Network", 2021, Pp. 12-16.

[13] Hong, F.; Tay, D.w.l.; Ang, A. Intelligent Pick-and-place System Using Mobilenet. Electronics 2023, 12, 621.

[14] Tan, Y.s., Lim, K.m., Tee, C. Et Al. "convolutional Neural Network With Spatial Pyramid Pooling For Hand Gesture Recognition". Neural Comput & Applic 33, 5339–5351 (2021).

[15] Kavitha M.N., Saranya S.S., Pragatheeswari E., Kaviyarasu S., Ragunath N., Rahul P, "A Real-Time Hand-Gesture Recognition Using Deep Learning Techniques" Proceedings of 4th International Conference on Artificial Intelligence and Smart Energy (ICAIS 2024) Information Systems Engineering and Management, vol 3. Springer, Cham., (2024).