Journal of the ACS Advances in Computer Science
Volume 16, Issue 1, 2025

A CNN-Based System for Sign Language Recognition

Dr. Marwa M. Abd El Hamid[1] ,Bassant S. Farouk [2]

[1,2] Computer Science Department, The Higher Institute of Computers and Information Technology, El Shorouk Academy, El Shorouk City, Cairo, Egypt

Email: marwa.mustafa@sha.edu.eg, bassant.sayed@sha.edu.eg

Abstract

Sign language plays a vital role in enabling communication for individuals with hearing impairments. To foster meaningful engagement with this community, learning sign language is essential. This paper aims to design and develop an intuitive sign language learning system leveraging deep learning. To attain this objective, a usable and intuitive application is developed adopting the Convolutional Neural Networks (CNN) for accurate recognition of American Sign Language (ASL) gestures dataset. The proposed CNN model is trained on a comprehensive dataset of Sign language gestures to achieve precise gesture classification. This paper demonstrated that the proposed Sign Language classification model achieved an accuracy of 94%, highlighting its effectiveness in reliably identifying and categorizing sign language gestures. An intuitive and user-friendly application has been developed, incorporating the proposed CNN model to process real-time input from a webcam and translate gestures into text output. Thus, this research holds significant potential to enhance the recognition and communication of sign language through the application of deep learning. The main purpose of this paper is to eliminate the barrier between the deaf and mute and the rest.

Keywords— Sign language, Convolutional Neural Networks (CNN), User interface, Image recognition

1. INTRODUCTION

Sign language is a vital means of communication for millions of deaf and mute individuals around the world [1]. Currently, around 466 million people experience hearing loss, with this number expected to rise to more than 900 million by 2050 [2]. Among them, approximately 34 million are children with severe to profound hearing loss. Deaf and mute individuals often face significant challenges in communication and social integration. The limited use of sign language in society creates barriers for these individuals, making it difficult for them to communicate effectively in daily life [3]. This communication gap can lead to feelings of isolation and exclusion, as well as challenges in accessing essential services and information. It is crucial for society to recognize and address the issue of limited sign language usage to foster greater inclusivity and accessibility for everyone [4]. Deaf individuals have little to no hearing ability and rely on sign language for communication. Different regions of the world use various sign languages, and while there are

DOI:

many spoken languages globally, the number of sign languages is relatively smaller. Signs in sign languages serve as the equivalent of words in spoken languages. Signed languages tend to favor visual and gestural communication, relying on hand shapes, movements, and facial expressions. Finger spelling is the method of representing letters from a writing system, and sometimes numbers, using hand signs. In sign language, the English alphabet (A-Z) can be represented through finger spelling, which can be done with one hand or two hands. Sign language typically follows the two-handed style.

Finger spelling is used to represent words that do not have a specific sign or to emphasize a word. Although its usage is less common in casual signing, finger spelling remains an important element in learning sign language. This paper aims to identify alphabets in sign language based on the corresponding gestures. While gesture recognition and sign language recognition have been widely studied, there is limited research specifically focused on sign language. Our goal is to address this gap by leveraging advanced computer vision and deep learning. Deep learning has great potential to improve sign language recognition systems. By automating gesture recognition with high precision, deep learning can help close the communication gap between hearing and deaf individuals, fostering more inclusive, accessible, and effective interactions in various areas of society.

Deep learning has achieved remarkable progress in computer vision, especially in recognizing patterns and classifying complex data like images and gestures [5]. When it comes to sign language recognition, deep learning methods, especially Convolutional Neural Networks (CNNs), have proven to be highly successful [6] in interpreting the hand gestures and movements that are essential to sign language communication. The main goal of this study is to propose and assess an interactive sign language learning system powered by deep neural networks. To achieve this, a user-friendly and essential learning tool was developed for deaf and mute individuals, as well as for anyone interested in learning sign language. The system seeks to bridge the communication gap between hearing and deaf/hard-of-hearing individuals, improving accessibility and fostering better communication for all [7].

2. MEDICAL BACKGROUND

In the context of medical considerations for deaf and mute individuals, it is important to understand the underlying causes of their conditions, the impact on their daily lives, and the challenges they face in accessing healthcare services.

Deafness, often categorized by the degree of hearing loss, can be congenital (present from birth) or acquired later in life due to factors such as aging, illness, or injury. Muteness, on the other hand, may result from a variety of medical conditions, including neurological disorders, physical injuries, or developmental issues [8]. Many individuals who are both deaf and mute may experience challenges in communication, social integration, and education, all of which can further affect their overall well-being [3].

Muteness, defined as the inability or severe difficulty in speaking, can stem from several underlying medical conditions. These include neurological disorders such as cerebral palsy, stroke, or brain injury, as well as physical impairments like vocal cord damage, developmental speech disorders, or psychological trauma. Mute individuals often rely on non-verbal forms of communication, such as gestures, facial expressions, or written language, to express themselves.

For those who are both deaf and mute, the medical challenges become even more complex. Without the ability to hear or speak, individuals may face heightened risks of social isolation [3], as they struggle to interact in environments that primarily rely on spoken communication. This dual disability can also complicate medical diagnosis and treatment, as healthcare providers may have limited tools for understanding the patient's needs without effective communication [8].

From a medical perspective, these individuals often require specialized care and support to address not only their communication needs but also their physical and emotional health. The lack of effective communication can hinder their ability to seek medical care, understand treatment options, or access critical health information. This underscores the importance of developing tools, such as sign language recognition systems, that can bridge communication gaps, enabling better healthcare access and improved quality of life for deaf and mute individuals [7]. In this paper, we aim to explore how technology, particularly through machine learning and computer vision techniques, can contribute to addressing these challenges by providing more efficient and accessible methods of communication for deaf and mute individuals, especially in medical settings [6].

3. LITERATURE REVIEW

For many years, researchers have focused on exploring sign language interpretation, particularly to enhance communication for individuals who are deaf or nonverbal [6]. This section offers a brief overview of key studies and progress in this field.

In [9], the authors tackled Arabic Sign Language recognition using a dataset of over 7,000 images of 28 letters. The team combined Mediapipe, a tool for extracting hand landmarks, with a CNN model to enhance recognition performance. The approach achieved a standout accuracy of 97.1%, demonstrating its efficiency and reliability.

In [10], the authors focused on recognizing American Sign Language letters using the 'Finger Spelling A' dataset, which features complex backgrounds. Researchers developed a multi-headed CNN that combined image and hand landmark data for more precise recognition. By applying data augmentation and adaptive learning techniques, the model achieved an impressive 98.98% accuracy, showing how powerful it can be to blend image and hand gesture information.

In [11], the authors used the Sign Language MNIST dataset from Kaggle, containing images of ASL letters. The CNN model outperformed KNN and SVM, achieving an accuracy of 94%, recall of 93%, precision of 93%, and F1-score of 93%, highlighting its effectiveness in capturing intricate features of ASL gestures. In [12], the researchers used CNNs to classify American Sign Language (ASL) gestures. The CNN model achieved an accuracy of 99.38% on the ASL dataset, demonstrating the model's effectiveness in recognizing ASL gestures. In [7], the study utilized a custom dataset comprising images of ASL alphabets, numbers (1 to 10), and static words. A Convolutional Neural Network (CNN) was employed for recognition tasks. The system achieved an average accuracy of 90.04% for ASL alphabet recognition, 93.44% for numbers, and 97.52% for static word recognition.

The primary goal of this paper is to develop a highly robust and generalized sign language recognition system capable of accurately recognizing gestures across diverse signers, languages, and real-world environments. Such a system should bridge the communication gap for the deaf and hard-of-hearing communities by providing seamless, real-time translation of sign language into text or speech.

This study builds on previous work by offering key improvements, including real-time gesture recognition using a webcam and a user-friendly interface. While some of earlier studies focused on static datasets and offline models, the proposed system combines high accuracy of 94% with practical usability, making it more applicable for real-world communication support. This represents a meaningful step forward in enhancing accessibility for the deaf and mute community.

4. MATERIALS AND METHODS

Sign language recognition is the process of converting hand gestures into meaningful symbols, such as letters or words, that can be understood by both deaf and hearing people [13]. This proposed work aims to develop a model capable of recognizing fingerspelling-based hand gestures and combining them to form complete words as shown in figure 1 [7].

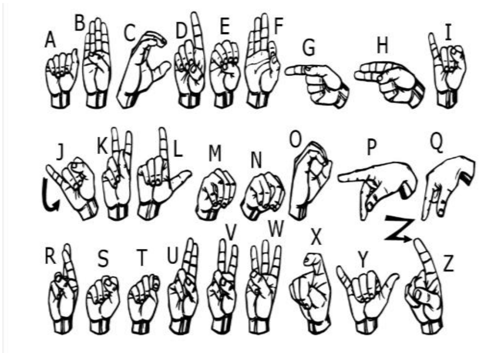

Figure 1. Fingerspelling-based hand gestures

In this proposed work, Deep learning models, particularly Convolutional Neural Networks (CNNs), are trained on extensive datasets of sign language images [6]. These models learn to identify patterns in the data and link them to the corresponding signs or letters.

CNNs are especially well-suited for this task because they can automatically detect and extract features like hand shape, movement, and position from images, without the need for manual feature selection [5]. By training on large volumes of sign language gestures, this proposed work shows that deep learning models can accurately classify new gestures in real-time, enabling seamless communication between deaf and hearing individuals [7].

Figure 2 shows the proposed flow diagram for sign language recognition using deep learning. It visually represents the process starting with real-time input from a camera, followed by image preprocessing, passing through the Convolutional Neural Network (CNN) for feature extraction, and finally outputting the recognized gesture as text or speech [4]. The diagram is designed to be simple and clear, with labeled steps to make it easy to follow.

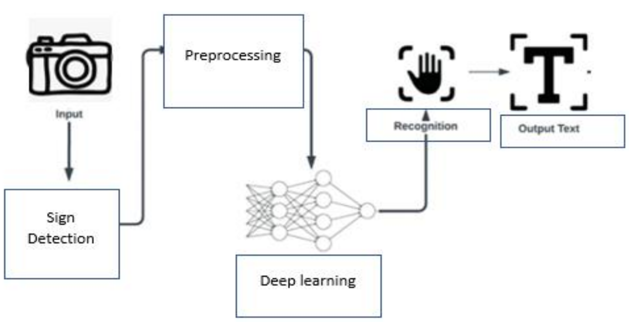

Figure 2. Proposed flow diagram

4.1 Dataset

Dataset was provided from American Sign Language (ASL) gesture datasets [14]. ASL gesture datasets are collections of images or videos designed to facilitate the development and testing of machine learning models for ASL recognition. ASL datasets are widely used in computer vision and deep learning research, particularly for developing gesture recognition systems aimed at enhancing accessibility for the deaf and hard-of-hearing communities. These datasets typically include a wide range of hand gestures representing ASL alphabets, and numbers. They are captured under various conditions to ensure diversity, including different hand orientations, lighting environments, and backgrounds. The data set is a collection of images of alphabets from the American Sign Language, separated in 29 folders which represent the various classes. There are 29 classes, of which 26 are for the letters A-Z and 3 classes for SPACE, DELETE and NOTHING. The dataset includes more than 87,000 images collected in diverse conditions. This diversity enhances its reliability and effectiveness for training and evaluating machine learning models. Each class includes thousands of images from various users, providing diverse hand shapes and orientations. Consistent framing with centered gestures and varied backgrounds helps models focus on the gesture despite environmental changes.

4.2 Data preprocessing

In the context of sign language recognition using Convolutional Neural Networks (CNNs) [15], proper preprocessing of input images is crucial to ensure effective model performance. The preprocessing steps include converting the images to grayscale to simplify the data and reduce computational complexity. Next, a Gaussian blur is applied to the images, which helps in smoothing out noise and enhancing important features. Finally, the images are normalized to scale the pixel values, ensuring that the model can learn efficiently by making the data consistent and within a manageable range. These preprocessing steps are essential in preparing the data for optimal performance with CNNs. Here's a more detailed explanation of the preprocessing steps involved before feeding the images into a Convolutional Neural Network (CNN) for sign language recognition [16]:

1. Grayscale Conversion:

The original input image, typically in color (RGB), is converted to grayscale. This reduces the dimensionality of the image by eliminating the color channels (Red, Green, Blue) and retaining only the intensity of light in each pixel. Grayscale conversion simplifies the input and focuses on the structural features of the image, which are more critical for CNNs in tasks like sign language recognition. By removing color information, it also reduces computational complexity, as CNNs have to process fewer channels (one channel vs. three channels in RGB images) [17].

2. Gaussian Blur:

After converting the image to grayscale, Gaussian blur is applied. This step involves smoothing the image by averaging the pixels with a Gaussian kernel. The purpose is to reduce high-frequency noise or fine details that might interfere with feature extraction. Essentially, Gaussian blur helps to highlight the key features (like hand shapes and gestures) while blurring out irrelevant details. The process ensures that the model is not distracted by noise and focuses on important patterns in the sign language gestures [18].

3. Normalization:

Once the image is preprocessed, normalization is applied. This step rescales the pixel values to a standard range, often between 0 and 1, by dividing each pixel value by 255 (the maximum possible pixel value in an 8-bit image). This step ensures that the model receives input data on a consistent scale, which aids in faster convergence during training. Normalization also prevents any particular pixel value from dominating the learning process due to large differences in magnitude across pixels [19]. It helps the CNN learn more effectively by treating all inputs equally. By performing these preprocessing steps—grayscale conversion, Gaussian blur, and normalization—the input data is prepared to ensure that the CNN can focus on the essential features for sign language recognition, improving the overall efficiency and accuracy of the model [20].

4.3 Proposed Convolutional Neural Networks (CNNs)

Deep learning has become a transformative technology in bioinformatics, offering powerful tools for analyzing complex biological data, particularly images. In bioinformatics, deep learning models like CNNs are instrumental in extracting meaningful insights from biological images, including cell microscopy, tissue scans, and radiological imaging. These models can detect anomalies, classify diseases, and segment biological structures with remarkable accuracy, often surpassing human expertise [21].

Deep learning plays a crucial role in bioinformatics, especially in the analysis of biological images, by providing advanced tools to uncover complex patterns and extract meaningful insights from large and intricate datasets.

In this paper, CNNs, were applied to achieve accurate and efficient solutions for interpreting visual data. CNNs are specifically designed to process and analyze images, making them ideal for recognizing the intricate patterns of hand gestures and movements in sign language. These networks excel at extracting spatial hierarchies of features, enabling them to identify complex shapes, contours, and positions of hands in various lighting and background conditions [22].

By leveraging CNNs, sign language recognition systems can achieve real-time performance and high precision, which is essential for creating reliable and accessible communication tools for individuals who are deaf or hard of hearing [23]. The ability of CNNs to adapt and improve through training on diverse datasets further enhances their potential to support more advanced and inclusive applications in this domain [24].

A step-by-step breakdown of the process is shown in figure 3 and described as the following:

a) Model Architecture

A typical CNN for sign language recognition consists of several key layers:

- Convolutional Layers: These layers extract important spatial features from the images, such as edges, shapes, and patterns in the hand gestures. Filters (kernels) slide over the image, detecting different patterns at each layer [25].

- Activation Function (ReLU): The Rectified Linear Unit (ReLU) activation function is applied to introduce non-linearity, allowing the model to learn complex patterns [26].

- Pooling Layers (Max Pooling): These layers down sample the feature maps, reducing dimensionality while preserving important features. This helps in making the model more efficient.

- Fully Connected Layers (Dense Layers): These layers interpret the extracted features and make final predictions. The last layer uses the softmax activation function to output probability scores for each sign language gesture [27].

b) Training with 10 Epochs

- Epochs represent the number of times the model goes through the entire dataset.

- In each epoch, the CNN learns by adjusting its weights using backpropagation and the optimizer (e.g., Adam or SGD).

- Loss Function (Categorical Cross-Entropy) is used to measure how well the model is performing. The optimizer updates the weights to minimize this loss.

c) Epoch-Wise Learning Process

- Epoch 1-3: The model starts recognizing basic patterns, but accuracy is generally low.

- Epoch 4-6: The CNN refines its feature extraction, improving classification accuracy.

- Epoch 7-9: The model further optimizes and reduces misclassifications.

- Epoch 10: The CNN reaches a stable accuracy, with minimal improvements per epoch.

d) Model Evaluation

After training, the model's performance is evaluated using accuracy and loss metrics. If accuracy is satisfactory, the model can be used for real-time sign language recognition. If not, additional epochs or fine-tuning (such as hyperparameter adjustments) may be required. By training the CNN for 10 epochs, the model gradually learns the features of different hand gestures, improving recognition accuracy while maintaining efficiency.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 26, 26, 32)	320
max_pooling2d (MaxPooling2D)	(None, 13, 13, 32)	0
dropout (Dropout)	(None, 13, 13, 32)	0
conv2d_1 (Conv2D)	(None, 11, 11, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 5, 5, 64)	0
dropout_1 (Dropout)	(None, 5, 5, 64)	0
conv2d_2 (Conv2D)	(None, 3, 3, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 1, 1, 128)	0
dropout_2 (Dropout)	(None, 1, 1, 128)	0
flatten (Flatten)	(None, 128)	0
dense (Dense)	(None, 512)	66,048
dropout_3 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 25)	12,825

Figure 3. CNN key layers

Figure 4 shows some real-time sign language detection results, including the letters O, L, and C.

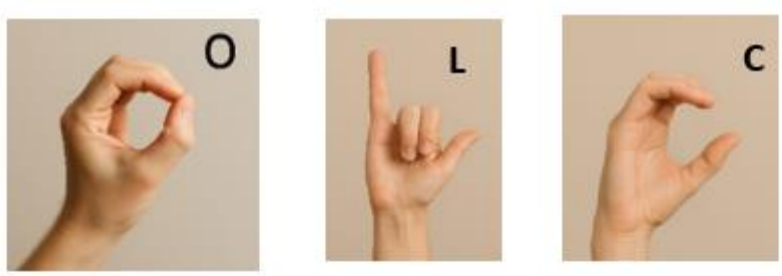

Figure 4. Sign Language System Detecting Letters O, L, and C

This proposed work can also translate entire words or phrases, like "Hello", into sign language or fingerspelling. The system shows a sequence of hand signs for each letter:

H – E – L – L – O. This is useful for spelling names or words without a dedicated ASL sign as shown in figure 5.

Figure 5. Hello work Translation

4.4 EVALUATION CRITERIA

By using accuracy, confusion matrix, and loss, the CNN model's performance was comprehensively evaluated [28]. The results show the model's effectiveness in recognizing sign language gestures.

5. RESULTS AND DISCUSSIONS

Once the images have been preprocessed, they are fed into a Convolutional Neural Network (CNN) for sign language recognition. The CNN is trained over 10 epochs, meaning the model iterates over the entire dataset 10 times to learn and refine its ability to recognize different hand gestures.

In this paper, visualization of the model's performance using accuracy and loss plots were performed [7]. These plots help in understanding how well the model is learning over the given number of epochs as shown in figure 6 and figure 7 respectively.

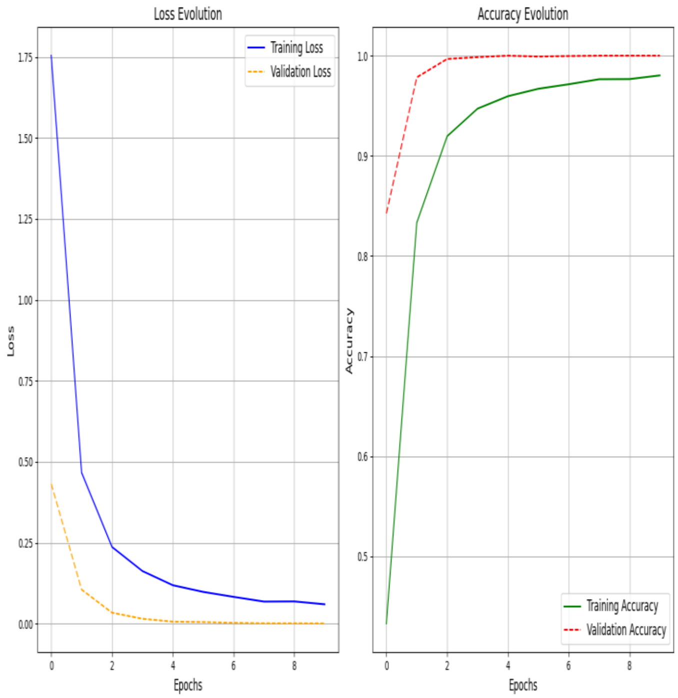

Figure 6. Accuracy evaluation plot

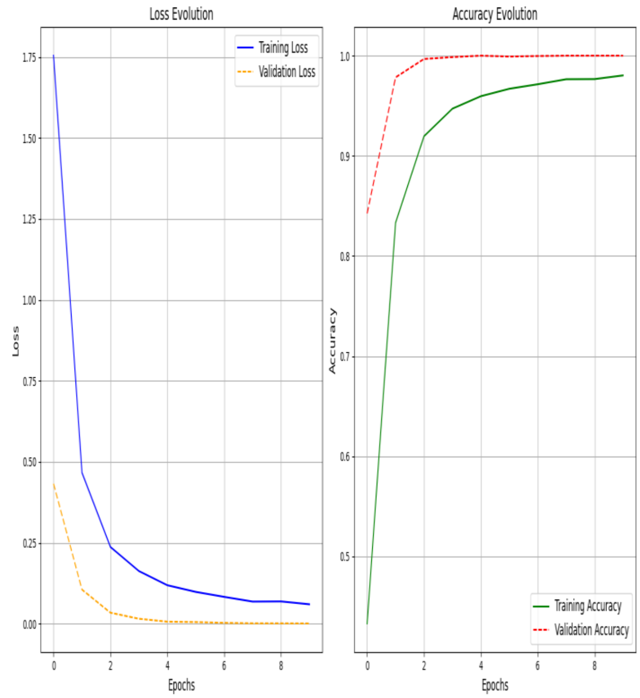

Figure 7. Loss evaluation plot

Figure 8 shows the non-normalized matrix to understand raw misclassification numbers. While, figure 9 shows the normalized matrix for performance comparison.

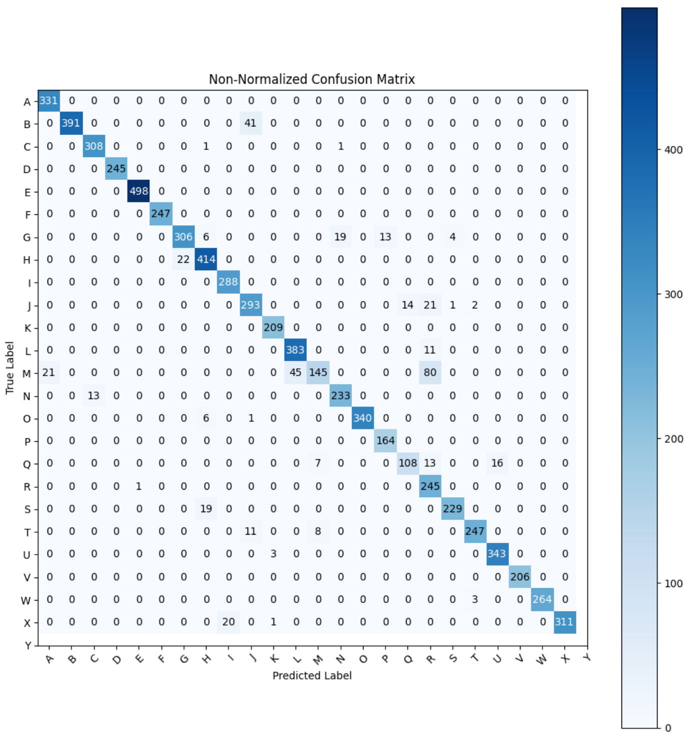

Figure 8. Confusion matrix without normalization

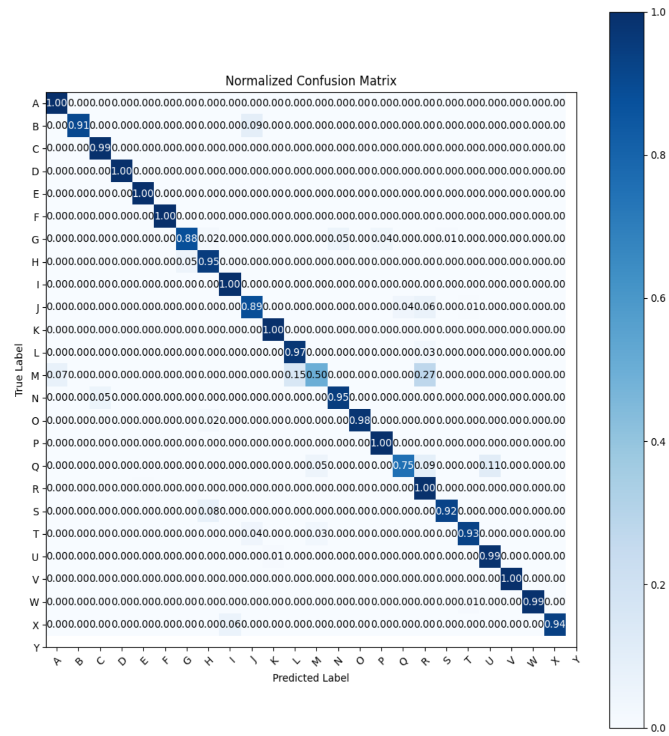

Figure 9. Normalized confusion matrix

In this paper, the trained CNN model achieved an overall accuracy of 94% (0.94) in sign language recognition, demonstrating strong performance in classifying different hand gestures [17]. The accuracy and loss plots indicate a steady improvement over the training process, with the model converging effectively by the final epoch. Overall, the results confirm the model's effectiveness in recognizing sign language gestures with high reliability.

The main contribution of the proposed work is the integration of a CNN model trained on a comprehensive sign language dataset with a real-time, webcam-based application, enabling immediate translation of gestures into text. This real-time functionality adds practical value and user interactivity that many previous works lacked. The proposed system achieved a robust 94% accuracy in a dynamic environment, balancing both precision and usability. By focusing on real-time recognition, user-friendliness, and accessibility, this work presents a significant advancement in the development of inclusive communication tools, moving beyond theoretical performance to real-world application and impact.

6. CONCLUSIONS

This paper introduced sign language learning and an interactive mobile application that leverages computer vision and machine learning to interpret hand gestures. The Sign language classification model demonstrated great performance, achieving a 94% accuracy. This high level of accuracy is critical for the practice feature of the mobile application, where users can rely on their device cameras to receive real-time feedback on their sign language gestures.

A real-time automatic sign language gesture recognition system was developed using CNN. While the proposed system successfully recognized sign language gestures and translated them into text. In conclusion, this study introduced a sign language learning and practice system powered by deep learning. The system demonstrated high accuracy in recognizing sign language alphabets and holds promise for future advancements, including the ability to recognize more complex hand gestures, enhancing the user experience, and optimizing functionality for mobile devices. This work presents a real-time sign language recognition system that combines a CNN model with a webcam-based interface, enabling instant gesture translation. It stands out for its practical usability, achieving high accuracy in dynamic settings and advancing accessible communication tools beyond theoretical models.

REFERENCES

[1] M. N. Islam, A. S. Hasan, T. T. Anannya, T. Hossain, M. B. I. Ema, and S. U. Rashid, "An efficient tool for learning Bengali sign language for vocally impaired people," in Mobile Web and Intelligent Information Systems: 16th International Conference, MobiWIS 2019, Istanbul, Turkey, August 26–28, 2019, Proceedings 16. Springer, pp. 41–53, 2019.

[2] A. Kusters and C. Lucas, "Emergence and evolutions: Introducing sign language sociolinguistics," Journal of Sociolinguistics, vol. 26, no. 1, pp. 84–98, 2022.

[3] S. A. Suha and T. F. Sanam, "Exploring dominant factors for ensuring the sustainability of utilizing artificial intelligence in healthcare decision-making: An emerging country context," International Journal of Information Management Data Insights, vol. 3, no. 1, p. 100170, 2023.

[4] C. M. Jin, Z. Omar, and M. H. Jaward, "A mobile application of American sign language translation via image processing algorithms," in 2016 IEEE Region 10 Symposium (TENSYMP). IEEE, pp. 104–109, 2016.

[5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[6] P. Wang, W. Li, and Z. Gao, "A survey on sign language recognition using deep learning," IEEE Access, vol. 9, pp. 123456–123478, 2021.

[7] Lean Karlo S., et al. "Static sign language recognition using deep learning." International Journal of Machine Learning and Computing 9.6, p: 821-827, 2019.

[8] A. R. Miller, D. J. Reznik, and R. L. Patel, “AI-Powered Assistive Technologies for Deaf and Mute Individuals,” Artificial Intelligence Review, vol. 58, no. 2, pp. 300–320, 2023.

[9] Refat Khan, et al. "Sign language recognition using the fusion of image and hand landmarks through multi-headed convolutional neural network." Scientific Reports 13.1, 2023

[10] Y. Ma, T. Xu, and K. Kim, “Two-stream mixed convolutional neural network for American sign language recognition,” Sensors, vol. 22, no. 16, p. 5959, 2022.

[11] Mohaisen, Linda. "An In-Depth Comparative Study of Machine Learning Algorithms for American Sign Language Recognition." Innovative and Intelligent Digital Technologies; Towards an Increased Efficiency: Volume 1. Cham: Springer Nature Switzerland, pp. 367-376, 2024.

[12] Thongpan, and Pusadee Seresangtakul. "Multi-stroke thai finger-spelling sign language recognition system with deep learning." Symmetry 13.2 , pp: 262, 2021.

[13] Y. Zhang, J. Liu, and H. Wang, “Sign language recognition based on dual-path background erasure,” Scientific Reports, vol. 13, no. 1, pp. 1–10, 2023.

[14] Grassknoted, "American Sign Language (ASL) Alphabet Dataset," Kaggle, 2017. Available: https://www.kaggle.com/datasets/grassknoted/asl-alphabet.

[15] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4700–4708, 2017.

[16] J. Smith and R. Brown, “Using convolutional neural networks for visual sign language recognition,” in Proceedings of the ACM Symposium on Applied Computing (SAC), pp. 1234–1240, 2023.

[17] S. N. Patel, G. D. Abowd, and K. A. Marin, “Recognizing Hand Gestures Using Wearable Sensors,” Pervasive Computing, vol. 4, no. 2, pp. 38–45, 2005.

[18] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.

[19] Abd El Hamid M., Shaheen M., Omar, Y., and Mabrouk M. Discovering epistasis interactions in Alzheimer’s disease using integrated framework of ensemble learning and multifactor dimensionality reduction (MDR). Ain Shams Engineering Journal, p. 101986, 2022.

[20] Y. Chen, L. Wang, and Z. Li, “Wearable Electronic Glove and Multilayer Para-LSTM-CNN-Based Method for Sign Language Recognition,” IEEE Transactions on Industrial Informatics, vol. 19, no. 1, pp. 123–132, 2023.

[21] Leiva, Víctor, et al. "A real-time intelligent system based on machine-learning methods for improving communication in sign language." IEEE Access, 2025.

[22] B. Dabwan, M. A. Hossen, and A. Bhuiyan, “Convolutional Neural Network-Based Sign Language Translation System,” International Journal of Advanced Computer Science and Applications, vol. 11, no. 9, pp. 560–567, 2020.

[23] A. Verma, P. Singh, and R. Kumar, “Sign Language Recognition Using Modified Deep Learning Network,” Scientific Reports, vol. 13, no. 1, pp. 1–12, 2023.

[24] G. A. Rao, K. Syamala, P. V. V. Kishore, and A. S. C. S. Sastry, “Deep Convolutional Neural Networks for Sign Language Recognition,” in Proceedings of the International Conference on Intelligent Human Computer Interaction (IHCI), pp. 1–5, 2018.

[25] Abd El Hamid M., Omar Y., Shaheen M., and Mabrouk M. Discovering epistasis interactions in Alzheimer's disease using deep learning model. Gene Reports, 29, p. 101673, 2022.

[26] J. Schmidhuber, “Deep learning in neural networks: An overview,” Neural Networks, vol. 61, pp. 85–117, 2015.

[27] Z. Zhao, S. Zheng, and C. Xu, “Transformers in Sign Language Recognition: A Novel Approach to Gesture-Based Communication,” IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 5, pp. 1234–1245, 2023.

[28] M. Mustafa, "Smart Tour Guide Service System Using Deep Learning and Voice Chatbot." Journal of the ACS Advances in Computer Science, vol. 15, no.1, 2024.