Journal of Telecommunication, Electronic and Computer Engineering
ISSN: 2180 – 1843 e-ISSN: 2289-8131 Vol. 16 No. 3 (2024) 19-26
jtec.utem.edu.my
DOI: https://doi.org/10.54554/jtec.2024.16.03.003

Kurdish Sign Language Recognition Using Convolutional Neural Network (CNN)

Sarkhel H.Taher Karim*, Muhammed Latif Mahmood, Siva Sabir Abdulla, Shano Ali Abdulla
Computer Science Department, College of Science, University of Halabja, Halabja 46018, Kurdistan Region, Iraq.

Article Info

Article history:
Received Oct 19th, 2023
Revised Sep 1st, 2024
Accepted Sep 27th, 2024
Published Sep 30th, 2024

Index Terms:
CNN
Convolution neural network
MediaPipe
Kurdish Sign language
Machine learning
Neural networks

Abstract

The present study examines the obstacles encountered by the deaf population, with a specific emphasis on the growing importance of sign language in facilitating effective communication. The main mode of communication for deaf individuals is Sign Language (SL), which conveys meaning visually and expressively through facial expressions, hand movements, and body gestures. The objective of this project is to automate the recognition of sign language to improve accessibility and reduce reliance on interpreters. Specifically, this work focused on developing an alphabet recognition system for Kurdish Sign Language (KSL). Due to its many intricacies and resemblances to the Arabic script, KSL requires a robust recognition model. The proposed method utilizes Convolutional Neural Networks (CNN) trained on a real-world dataset to accurately recognize both numerical values and alphabetic characters in the Kurdish Sign Language (KSL). The real-time operation of the system enables rapid recognition of hand gestures, providing immediate textual output. The dataset used for training comprises 132,000 hand images, including 33 alphabetic signs and numeral signs from 0 to 9. The use of MediaPipe, a method for processing 3D images, significantly improves the efficiency of gesture detection. Multiple methodologies were investigated, and the integration of Convolutional Neural Networks (CNN), TensorFlow, and MediaPipe resulted in a remarkable accuracy of 99.87% with negligible dropout rates. This study establishes a foundation for enhanced communication and independence for the deaf community, representing a significance advancement in the automation of sign language recognition.

*Corresponding Author: Sarkhel.kareem@uoh.edu.iq

I. INTRODUCTION

Nowadays, according to the World Health Organization (WHO), approximately 5% of the global population (equivalent to 1 in every 20 people) suffers from hearing issues. Researchers assume that by 2050, the prevalence of hearing loss will rise significantly, with 1 in every 10 people affected [1]. People can experience complete deafness for various reasons, including prolonged exposure to loud noise, genetic conditions, traumatic brain injuries, or illness [2]. One of the primary means of communications for those are deaf is sign language. However, there is no universal sign language that is understood worldwide, as different forms of sign language have evolved naturally within distinct societies. The use of sign language in Western societies dates back to the 17th century, in which it allows individuals to communicate letters, words, phrases, or even entire ideas through gestures [3]. Hearing people typically grow up learning spoken languages, whereas non-hearing individuals acquire sign language as their primary form of communication. Deaf individuals, who have never experienced sound, cannot fully comprehend what sound is or how to speak verbally. Instead, they rely on visual cues, such as hand movements, to express their thoughts and communicate with others [4]. The structure of sign language (SL) is complex which makes it challenging for hearing people to understand and interact with deaf individuals. Addressing these challenges is crucial to improve the lives of deaf people, as they often face inequality and depression due to communication barriers. Respecting and recognizing their sign languages is a key step in acknowledging their rights and enhancing their quality of life.

Unfortunately, there are relatively few schools and educational resources dedicated to teaching SL, which contributes to a lack of communication and language skills among deaf individuals. Moreover, not all deaf individuals have the opportunities to attend such such specialized schools. Sign language users employ a combination of manual and non-manual methods to coney their messages, using parameters such as gestures, finger shapes, hand orientation, hand movements, body movements, facial expressions and lip patterns. These elements visually and manually transmit meaning, serving as a replacement for sound. Like any language, SL has specific syntax and vocabulary, hence, mastering it requires time and practice [5]. Sign language serves as a bridge between normal-hearing individuals and those with hearing disabilities, enabling communication and fostering understanding [6]. With advancements in technologies, particularly in computer science and artificial intelligence (AI), solutions are emerging to address significant challenges, including the detection and recognition of sign languages. These innovations aim to facilitate communication between deaf, mute, and hearing individuals. SL recognition systems can convert signs into speech or text in real-time, enhancing the efficiency and accuracy of communication. For these systems to be effective, they must be reliable, efficient and usable by both sides to meet their needs in communication [7].

The major goal of this project is to create a resilient sign language recognition system, specifically designed for Kurdish Sign Language, which remains significant underrepresented in current literature and technology. While notable advancements have been made in the recognition of languages such as American Sign Language (ASL) and Arabic Sign Language, Kurdish Sign Language still lacks substantial focused research and technological solutions. This gap highlights the critical need for comprehensive studies and the significance of investigating the distinct features and intricacies of Kurdish Sign Language. This study is motivated by the desire to enhance communication accessibility for Kurdish-speaking deaf communities and to pioneer sign language recognition systems that are adaptable to various linguistic and cultural settings.

The objective of this work is threefold. First, it aims to determine the optimal machine learning models and methodologies for producing precise recognition of Kurdish sign language motions. Second, its purpose is to explore how the integration of technologies such as MediaPipe, TensorFlow, and Convolutional Neural Networks can improve the precision and effectiveness of recognizing Kurdish sign language. The third objective is to investigate the obstacles and constraints involved in creating a reliable recognition system for Kurdish Sign Language. Specifically, this includes identifying effective approaches for collecting, preparing, and optimizing the training data, which are crucial steps in overcoming the existing obstacles. The study seeks to address these objectives to contribute meaningfully to the field of sign language recognition, which has a particular focus on underrepresented languages. This research intends to offer valuable insights for future developments in sign language recognition, thereby advancing knowledge and solutions within this domain.

II. LITERATURE REVIEW

The recognition of hand gestures is a critical task in machine learning, particularly for the development of sign language recognition systems. These systems rely on both supervised and unsupervised classification techniques to detect either static or dynamic hand movements. In such systems, both sign-level modeling and subunit sign-level modeling are possible. This discussion focuses on the research algorithms, accuracy rates, techniques, mechanisms, sign languages used, and the datasets employed for these purposes.

Tolga and Kamil [9] worked on the recognition of Turkish Sign Language (TSL). They trained and tested images derived from videos taken in a studio environment, featuring three different signers and using a green screen background to simplify images processing. Images containing signs were extracted from these video files and used for training and tested. Their dataset includes 217,500 images for training and 43,500 images for testing. The same test set of testing data was used for model evaluation. MATLAB’s video labeling tool was utilized for annotate the images, and a Region-based Convolutional Neural Network (R-CNN) was used to train the model for object detection. After training the R-CNN object detector, the dataset was divided into 29 classes for classification. Transfer learning was utilized with AlexNet, chosen due to its validated performance in similar applications. The system demonstrated high accuracy, achieving a 99.7% recognition rate for TSL characters [8][9].

Batool, Mohamed, and Ouiem developed a model for recognizing Arabic Sign Language (ArSL) using EfficientNet CNN. Their work specifically addressed the issue of Arabic Sign Language recognition. EfficientNet, a new family of Convolutional Neural Networks (CNN) released by Google in 2019, has proven to have fewer parameters and Floating-Point Operations Per Second (FLOPS) while producing competitive accuracy. They specifically adapted EfficientNet models to classify Arabic Sign Language gestures and developed lightweight deep learning models. The dataset used in this study was collected from over twenty volunteers, totaling 5,400 images. The EfficientNet-Lite 0 architecture with a label smoothing loss function, produced the best results. Their model achieved an accuracy of 94%, effectively handling variations in background settings [10].

To improve the recognition of 32 hand gestures in Arabic Sign Language, Yaser [11] developed a model using deep convolutional neural networks (CNN) optimized through transfer learning. The approach entails using architectures similar to VGG16 and ResNet152, incorporating pre-trained model weights into each network’s layers, and adding a soft-max classification layer at the end. The objective was to show the benefits of fine-tuning pre-trained models for improving network training and producing higher accuracy compared to other methods. The dataset consisted of 54,049 images across 32 different classes of Arabic sign language gestures. The system model is designed to handle 2D image data under varying backgrounds and lighting conditions. The model achieved a recognition accuracy of 99% [11].

Kurdish Sign Language (KSL) requires a robust processing method due to its unique characteristics and specific linguistic nuances. Recent research on KSL has faced challenges, such as limited datasets availability and the diverse linguistic features inherent in KSL. Karwan created a real-time model for recognizing KSL using a Convolutional Neural Network (CNN). The model was trained and evaluated on the KuSL2022 dataset, using various activation functions over multiple training epochs. The dataset consists of 71,400 images representing the 34 Kurdish sign language alphabets, sourced from two different databases. The main goal of the study was to identify Kurdish Sign Language alphabets by recognizing hand shapes using elements from ASL alphabet and the Ar2018 dataset. These datasets were updated to adapt to the Kurdish alphabet, effectively capturing and processing Kurdish-Arabic script through hand gestures captured by a camera. The model achieved an average training accuracy of 99.91%, demonstrating a significant improvements in both classification and prediction capabilities for recognizing Kurdish hand signs. [12].

Mayyadah, Adnan, and Zeynep [13] used a method called RTDHGRS to track and recognize ten Kurdish sign language words. The RTDHGRS system processes test images, which are then classified by an Artificial Neural Network (ANN) using two lines of features, compared with pre-stored features from the training dataset vectors. Each training dataset vector was created from fifty distinct features per line, leading to a simple yet efficient technique for recognizing dynamic hand gestures. The dynamic words were captured using frames 16 and 30 from video sequences. The dataset consisted of ten selected words from ten participants, resulting in a total of 200 images for training. The ANN-based classifier was able to identify 98% of the hand gestures during training and 20% during testing. The images were recorded in Bitmap Image File Format (.bmp), and the features were stored as vectors of 100 bits or pixels. The system achieved an average accuracy of 98% [13].

Another significant contribution was made by Abdulla Dlshad and Fattah Alizadeh [14], who focused on developing a letter-based Kurdish sign language recognition system. They developed a cross-platform, webcam-based program capable of translating Kurdish Sign Language (KuSL) into Kurdish-Arabic script. Three algorithms were tested to recognize Kurdish sign language: two well-known feature extraction techniques, Scale-Invariant Feature Transform (SIFT) and Speeded-Up Robust Features (SURF), along with a new algorithm introduced in their study called the grid-based gesture descriptor. The grid-based gesture descriptor featured a streamlined algorithm with reduced computational complexity, representing the main innovation of their work. The proposed algorithm achieved 67% accuracy in recognizing sign hands, outperforming the other two well-known algorithms (SIFT and SURF), which each reached an accuracy rate of only 42% [14].

The goal of this research is to develop a system capable of recognizing Kurdish Sign Language alphabets and digits using a Convolutional Neural Network (CNN) and incorporating advanced technologies, such as MediaPipe, to enhance accuracy. MediaPipe, a cutting-edge tool in gesture recognition, is expected to improve both the efficiency and precision of the system significantly.

III. RESEARCH METHODOLOGY

This work presents the development of a sign language recognition system using deep learning methodologies. The procedure started with meticulous data collection, which is essential for constructing a resilient model. A vision-based methodology was used, wherein images were captured using a camera under various conditions (including changes in lighting, background, and distance) to replicate real-world settings. This methodology enables the collection of a wide range of data, which is crucial for the rigorous training of a deep learning model capable of reliably identifying hand movements.

A. Data Acquisition

The dataset comprised 66,000 images representing Kurdish Sign Language (KSL) signs, covering 33 alphabet letters and digits from 0 to 9. These images were captured under various conditions to ensure the robustness of the recognition system. The images were resized to 256×256 pixels to maintain consistency and uniformity. Each sign category contained between 700 to 1500 images Additionally, data augmentation techniques were employed, effectively doubling the dataset size to 132,000 images. This augmentation enhanced the dataset variety, thereby improving the model’s ability to generalize.

B. Statistical data processing

The preprocessing stage included multiple steps aimed at improving data quality. To ensure consistency, all images were resized to dimensions of 256×256 pixels. Subsequently, data augmentation methods, including rotation, flipping, and brightness modifications, were used to enhance the diversity of the dataset, thereby enhancing the model's ability to generalize across different conditions. In addition, the images were normalized to ensure that the pixel values fell within a consistent range, which is crucial for enhancing the efficiency of the model's learning process.

C. Development of the Model:

A Convolutional Neural Network (CNN) was employed to perform the recognition tasks, featuring several layers:

- Input Layer: Receives the preprocessed images for analysis.
- Convolutional Layers: Utilizes multiple filters to extract relevant features, with increasing depth across successive layers to capture more complex patterns.
- Pooling Layers: Reduce dimensionality to mitigate overfitting.
- Fully Connected Layers: Combines extracted features to make the final classification.

The model underwent fine-tuning by adjusting several hyperparameters, such as the learning rate, batch size, and number of training epochs. Multiple experiments were conducted with different learning rates (e.g., 0.001, 0.0001) to determine the optimal value that minimized loss without causing model divergence. A batch size of 32 was chosen as it provided a suitable balance between memory usage and training speed. The model underwent 50 training epochs, during which early stopping was implemented to mitigate overfitting in cases where validation accuracy showed no further improvement after a certain number of epochs.

The dataset was partitioned into 80% for training and 20% for testing. During training, the model learned to recognize hand gestures by minimizing categorical cross-entropy loss using the Adam optimizer. Model performance was evaluated using the test set, with accuracy serving as the main performance metric. Furthermore, confusion matrices were used to analyze the model's performance across different categories, highlighting any instances of misclassification.

One notable challenge was distinguishing between visually similar letters, such as (þ) and (α). To address this issue, additional data augmentation was introduced to further diversify the dataset, and modifications were made to the model’s design to enhance its ability to differentiate subtle variations in hand movements.

Through this approach, a resilient sign language recognition system was developed, capable of accurately identifying Kurdish alphabet letters and number signs. By leveraging a CNN model along with comprehensive data preparation and strategic hyperparameter tuning, a high-performance system was developed that is well-suited for practical applications in recognizing Kurdish Sign Language.

Figure 1. Samples of Kurdish Alphabet and Number Signs

IV. TESTING AND ANALYSIS

Neural Networks have made significant strides in numerous fields due to their flexibility and scalability. These models hold immense potential for enhancing data modelling and analysis across various issues in the learning and behavioral sciences. Common neural models such as Stochastic Gradient Descent (SGD), Multilayer Perceptron (MLP), and Convolutional Neural Network (CNNs)), all of have been used till now [15]. For this particular system, CNN was used due to its effectiveness as a Deep Learning algorithm that can recognize differences between input images [16]. CNN is the most widely used deep learning technique for computer vision tasks, and it has three different kinds of layers as shown in figure 2.

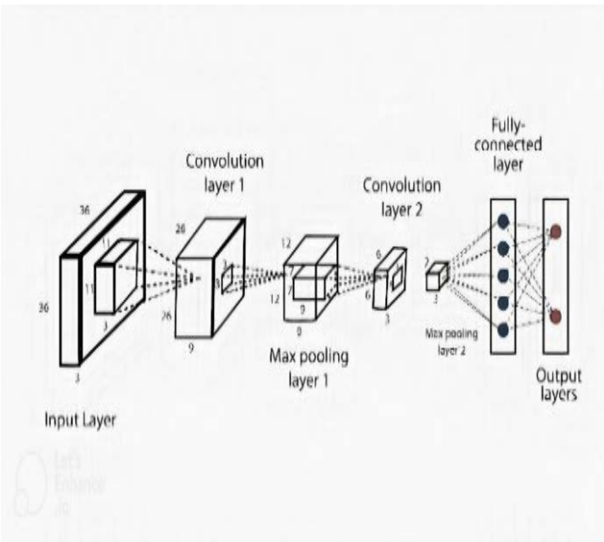

Figure 2. Convolutional Neural Network

A. Convolutional Layer

The convolutional layer is a key component in feature extraction within a CNN. Neurons in this layer are connected to other neurons in adjacent layers via a kernel, a small matrix, typically of size 3*3, though it can also be 5*5 or 7*7. The convolutional layer is responsible for detecting spatial features in the input data by sliding the kernel over the image and performing convolution operations.

B. Pooling Layer

The pooling layer is used to reduce the spatial dimensions of the feature maps, helping to manage computational complexity and prevent overfitting. This kind of layer, which receives many feature maps and applies the pooling operation on each of them, is often used between two layers of convolution.

C. Fully-Connected Layer

The fully-connected layer is not a characteristic of CNN because it is always the last layer of the neural network, it may be convolutional or not.

For this system, several packages were used, including TensorFlow, LSTM (Long Short-Term Memory) OpenCV, and MediaPipe. TensorFlow is designed to build neural networks for deep learning, making it simpler to write code that runs efficiently on both GPUs or CPUs in a distributed manner [17]. LSTM networks are a type of Recurrent Neural Network (RNN) used to learn long-term dependencies, especially useful in sequence prediction problems [18]. In this research the MediaPipe package was used for hand detection. MediaPipe's "landmark" feature identifies key points (21 key points) on the hand, enabling precise recognition of the hand's structure, as shown in figure 3.

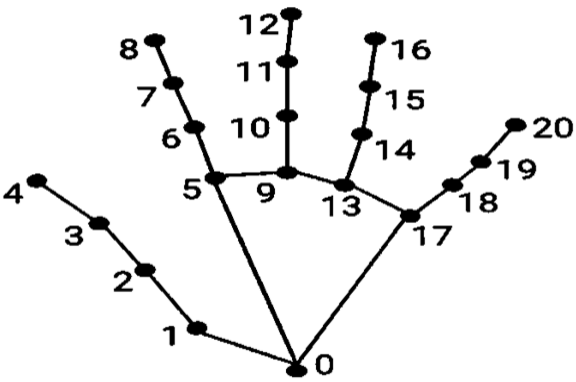

Figure 3 Hand Landmarks Detection

Before training the model, the dataset underwent optimization to ensure it was ready for efficient use during the training process. The following optimization functions were applied:

prefetch (): This function allows the system to read the next data sample during each training step, allowing simultaneous processing of the current and next batch of data. The tf. data. AUTOTUNE parameter was used to automatically determine the optimal size of the prepared data.

Shuffle (): Shuffling helps reduce potential correlations between consecutive images, which is important for generalization. Given the large dataset size, full randomization would be memory-intensive, so only a defined number of consecutive samples were shuffled to maintain randomness.

Cache (): This function allows the dataset to be loaded into working memory during the first epoch, keeping it cached for faster access during the whole training process.

The model underwent training through three approaches: First, the model was trained using the original dataset without any flipped data. As shown in Figure 7, there were minor discrepancies between training and validation accuracy, but not significant. In the second approach, the model was trained solely on flipped versions of the original images. As shown in Figure 8, this also resulted in discrepancies between training and validation accuracy. In the final step, the model was trained using both original and flipped datasets. As shown in Figure 9, this approach achieved highly similar training and validation accuracy, which indicates that the system performed well and is capable of producing accurate results in real-time applications.

Table 1
The Structure of the Proposed Model Layers

Layer Type	Output Shape	Parameter
Dropout_54(Dropout)	(None, 42)	0
Dense_81(Dense)	(None, 200)	8600
Dropout_55(Dropout)	(None, 200)	0
Dense_81(Dense)	(None, 10)	2010
Dense_81(Dense)	(None, 33)	363

Table 2
Summary of Model Parameters

Total Parameter	10973
Trainable Parameter	10973
Nontrainable Parameter	0

Finally using an effective optimizer is essential, as it updates the model's parameters, such as biases and weights, during training to minimize losses. In this work, the Adam optimizer was used for optimization, as it effectively handles noisy problems with sparse gradients and supports learning rate decay.

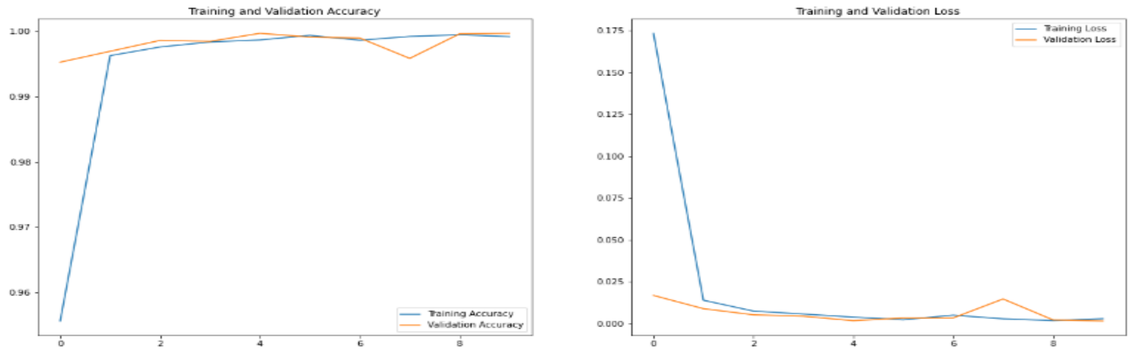

Figure 4. Line Plots of Accuracy and Loss using Original Data

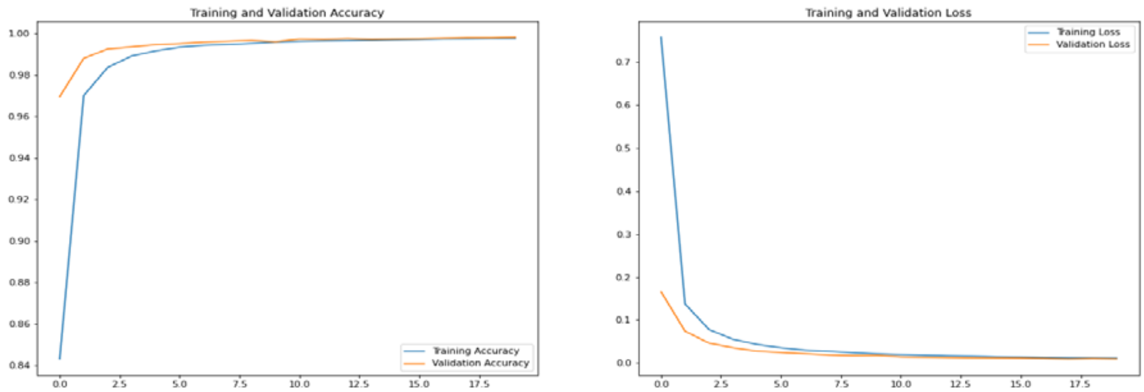

Figure 5. Line Plots of Accuracy and Loss using Flipped data

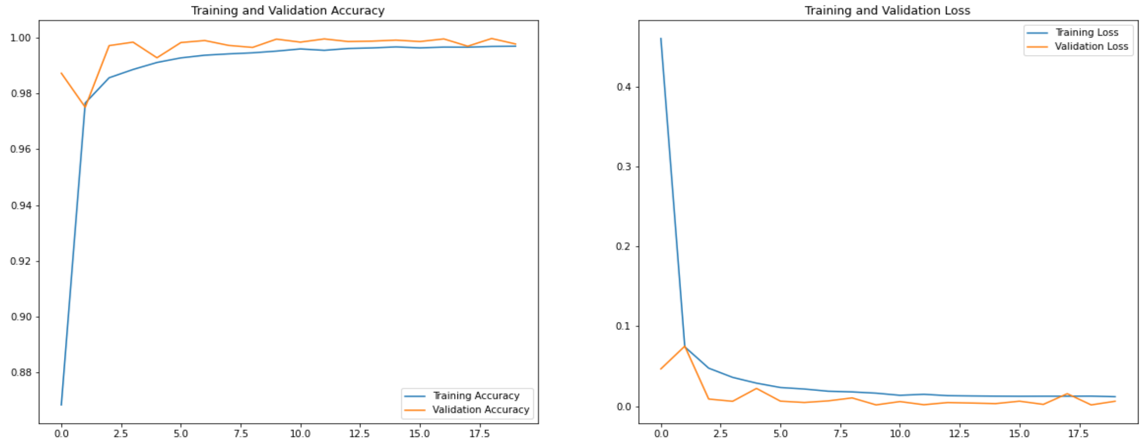

Figure 6. Line plots of Accuracy and using both original data and flipped data

V. RESULT AND DISCUSSION

The objective of this research paper is to develop a model capable of converting Kurdish signs into written text by employing several algorithms designed to extract distinguishing features from images. The optimal strategy depends on factors such as the characteristics of the data, the level of intricacy, and the required image resolution. To determine the most effective technique, multiple methodologies were evaluated to identify the ideal model parameters, even for experienced researchers.

After training the model, its accuracy was verified using a validated dataset. During the testing phase, the model was extended to a novel dataset to further verify its effectiveness. Comparing the predicted output to the actual output enhanced the accuracy of the model. The system takes an image as input and categorizes it into one of 43 distinct classes.

In this study, multiple models were assessed, including combinations of Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), MediaPipe, and TensorFlow. Each image was categorized into one of 43 classes, and the classification was based on the model that demonstrated the greatest level of accuracy.

1) The average test accuracy for several models examined in this work is shown in Table 3. MediaPipe and TensorFlow were evaluated through a series of experiments, in which these frameworks were used to analyze input images and extract features. To achieve real-time hand landmark detection, a critical aspect of sign language recognition, the MediaPipe framework, designed for constructing multimodal machine learning pipelines, was used. TensorFlow, a widely used deep learning framework, was then used to train and test the models using the extracted features.
2) **Long Short-Term Memory (LSTM), MediaPipe, TensorFlow (76.65% Accuracy):** The configuration included Long Short-Term Memory (LSTM) networks for sequence prediction, MediaPipe for hand landmark recognition, and TensorFlow for model training. The accuracy of this model is indicative of its ability to correctly identify sequences of signals, which is advantageous for applications requiring continuous sign language identification.
3) **MediaPipe, TensorFlow (92.85% Accuracy):** This approach used only MediaPipe for feature extraction and TensorFlow for training a simple model. Without the added complexity of LSTM, the model focused on recognizing static signs, resulting in higher accuracy due to simpler model requirements.
4) **CNN, MediaPipe, TensorFlow (99.99% Accuracy):** This setup integrated Convolutional Neural Networks (CNNs) with the MediaPipe and TensorFlow frameworks. The Convolutional Neural Networks (CNNs) show exceptional efficacy in pattern recognition within images; thus, substantially enhancing accuracy in comparison to other approaches. The use of MediaPipe ensured precise hand tracking, which is crucial for effective sign language recognition.
5) **CNN, MediaPipe, TensorFlow with Dropped Layers (99.73% Accuracy):** This version of the CNN model incorporates the removal of certain layers to reduce complexity and training time. Notwithstanding this simplification, the model retained a high level of accuracy, indicating that the condensed model still incorporated the crucial features necessary for effective sign recognition.

The findings, summarized in Table 3, indicate that the optimization of CNN, MediaPipe, and TensorFlow yields the highest accuracy for the recognition of Kurdish sign language. The selection of the model and setup is contingent upon the specific requirements of the application, such as the need for real-time processing or the ability to handle uninterrupted sequences in sign language.

Table 3
Comparing Different Models Accuracy

Model	Average Test Accuracy
LSTM, MediaPipe, TensorFlow [19]	84%
MediaPipe, TensorFlow [20]	92.85%
CNN, MediaPipe, TensorFlow [16]	98.99%
CNN, MediaPipe, TensorFlow with Dropped Layers [12]	99.73%

Table 4 presents the performance of several model configurations, demonstrating the efficacy of integrating Convolutional Neural Networks (CNNs) with MediaPipe and TensorFlow for the purpose of sign language recognition. The findings emphasize the importance of selecting the appropriate model architecture and tools based on the specific requirements of each application.

Table 4
Comparing Different Models Accuracy

Research Work	Sign Language or Task	Algorithm/Model Used	Accuracy	Key Features
Tolga and Kamil (2022)	Turkish Sign Language (TSL)	R-CNN, AlexNet	99.7%	Utilized green screen background; transfer learning with AlexNet for robust feature extraction [21].
Batool, Mohamed, and Ouiem (2023)	Arabic Sign Language (ArSL)	EfficientNet CNN	94%	Developed lightweight deep learning models; focused on reducing parameters and FLOPS for efficiency [10].
S. Yaser (2022)	Arabic Sign Language (ArSL)	VGG16, ResNet152 with transfer learning	99%	Fine-tuning of pre-trained models; adapted to varied lighting conditions and backgrounds [11].
H. R. Karwan (2022)	Kurdish Sign Language (KSL)	CNN	99.91%	Real-time recognition model; trained on KuSL2022 dataset; used updated datasets for Kurdish alphabet [12].
Mayyadah, Adnan, and Zeynep (2022)	Kurdish Sign Language (KSL)	Artificial Neural Network (ANN)	98%	Used RTDHGRS for dynamic gesture recognition; simple, fast technique with 98% training accuracy [13].
Abdulla Dlshad and Fattah Alizadeh (2022)	Kurdish Sign Language (KSL)	SIFT, SURF, Grid-Based Gesture Descriptor	67%	Developed a new algorithm for gesture recognition; focused on reducing computational complexity [14].
S. K. G. Meena et al. (2024)	Monkeypox Recognition	Deep Transfer Learning-based Neural Networks	98%	Applied transfer learning for disease recognition from images; new in the field of medical image analysis [22].
R. K. Singh et al. (2023)	Image-Based Sentiment Analysis	InceptionV3 Transfer Learning Approach	99.5%	Used InceptionV3 for sentiment analysis; focused on leveraging deep learning for sentiment prediction [23].
R. R. Choudhary, K. Jisnu, and G. Meena (2022)	Image Dehazing	Deep Learning Techniques	N/A	Developed deep learning-based models for improving visibility in hazy images; significant in image processing [24].
R. K. Yadav et al. (2024)	Devanagari Handwritten Character Recognition	Modified Lenet-5 Deep Neural Network	99.21%	Applied modified Lenet-5 for recognizing handwritten characters; tailored for script-specific challenges [25].

The Kurdish Sign Language (KSL) recognition conducted in this study was compared to other works. Tolga and Kamil's study in Turkish Sign Language (TSL) used Region-based Convolutional Neural Networks (R-CNN) and AlexNet for transfer learning, using 217,500 training images and 43,500 testing images to achieve 99.7% accuracy [21]. Similarly, Batool, Mohamed, and Ouiem utilized EfficientNet CNN to recognize Arabic Sign Language (ArSL) with an accuracy of 94% based on 5,400 pictures [10].

Additionally, research from 2023 and 2024, which leveraged deep learning and transfer learning for visual identification, was included for comparison. These studies focuses on topics such as Monkeypox recognition using deep transfer learning-based neural networks [10], sentiment analysis on images using different transfer learning models image dehazing using deep learning [23], and off-line Devanagari handwritten character recognition using a modified Lenet-5 deep neural network [25] are notable examples.

In the present study, the KuSL2022 dataset, consisting of 71,400 images across 34 classes, was used to train a convolutional neural network (CNN) for real-time Kurdish Sign Language detection. Our technique outperformed others in classification and prediction tasks, with an average training accuracy of 99.91%. The use of new technologies like MediaPipe further improved accuracy and robustness, especially in the recognition of KSL alphabets and digits.

A comparative analysis of these studies, presented in tabular form, highlights their datasets, methodologies, and achieved accuracies. This comparison showcases how our technique differentiates itself from previous work and advances the field of Kurdish Sign Language recognition.

VI. CONCLUSION

In this work, a resilient Kurdish Sign Language (KSL) recognition system was successfully developed by integrating Convolutional Neural Networks (CNN) with MediaPipe, TensorFlow, and other data preparation methods. The system exhibits remarkable precision in recognizing both Kurdish alphabetic letters and digits, achieving an impressive accuracy of 99.87%. This accomplishment highlights the effectiveness of using deep learning techniques to tackle the difficulties related to sign language recognition, especially for underrepresented languages such as KSL.

The meticulous data augmentation and preparation procedures ensured the model's ability to generalize well to diverse environmental conditions, thereby improving its reliability in practical applications. The incorporation of MediaPipe has enhanced the accuracy of hand gesture recognition, contributing to the system's exceptional performance. Notwithstanding the intricacies and distinctive features of KSL, the system successfully surmounted these obstacles, providing a valuable tool for enhancing communication accessibility for the Kurdish deaf community.

This study establishes a strong foundation for future advancements in KLS language recognition and highlights the potential of integrating sophisticated machine learning methods with state-of-the-art technology to address linguistic variability in sign language recognition systems. Future study could focus on expanding the dataset, exploring alternative machine learning models, and refining real-time implementation to further enhance the practicality and scalability of the system.

CONFLICT OF INTEREST

Authors declare that there is no conflict of interests regarding the publication of the paper.

AUTHOR CONTRIBUTION

The authors confirm contribution to the paper as follows:

Study conception and design: Sarkhel H. Taher Karim, Muhammed Latif Mahmood. Data collection: Siva Sabir Abdulla, Shano Ali Abdulla. Analysis and interpretation of findings: Muhammed Latif Mahmood, Sarkhel H. Taher Karim, Siva Sabir Abdulla. Draft manuscript preparation: Sarkhel H. Taher Karim, Siva Sabir Abdulla.

REFERENCES

[1] R. Kushalnagar, "Deafness and hearing loss," *Web accessibility: A foundation for research,* pp. 35-47, 2019.

[2] J. E. Lieu, M. Kenna, S. Anne, and L. Davidson, "Hearing loss in children: a review," *Jama,* vol. 324, no. 21, pp. 2195-2205, 2020.

[3] B. Woll, "The history of sign language linguistics,", Accessed: 1 Oct 2024, https://academic.oup.com/edited-volume/28195/chapter-abstract/213121853?redirectedFrom=fulltext, 2013.

[4] R. Rastgoo, K. Kiani, and S. Escalera, "Sign language recognition: A deep survey," *Expert Systems with Applications,* vol. 164, p. 113794, 2021.

[5] W. F. o. T. Deaf. "Our mission, our values, our people." Accessed: 1-10-2024, https://wfdeaf.org/who-we-are/our-philosophy/, 2024.

[6] C. Ochoa, "Advancing the language of human rights in a global economic order: an analysis of a discourse," *BC Third World LJ,* vol. 23, p. 57, 2003.

[7] J. Sudthipadh and S. Pumrin, "Development of Thai Sign Language Interpretation with MediaPipe," in *2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*, 2022, IEEE, pp. 841-843.

[8] I. A. Adeyanju, O. O. Bello, and M. A. Adegboye, "Machine learning methods for sign language recognition: A critical review and analysis," *Intelligent Systems with Applications,* vol. 12, p. 200056, 2021.

[9] T. Yirtici and K. Yurtkan, "Regional-CNN-based enhanced Turkish sign language recognition," *Signal, Image and Video Processing,* vol. 16, no. 5, pp. 1305-1311, 2022.

[10] B. Y. AlKhuraym, M. M. B. Ismail, and O. Bchir, "Arabic sign language recognition using lightweight cnn-based architecture," *International Journal of Advanced Computer Science and Applications,* vol. 13, no. 4, 2022.

[11] Y. Saleh and G. Issa, "Arabic Sign Language Recognition through Deep Neural Networks Fine-tuning," in *Proceedings of the International Conference on Artificial Intelligence and Computer Science*, 2020.

[12] K. H. Rawf, A. Abdulrahman, and A. Mohammed, "Effective Kurdish sign language detection and classification using convolutional neural networks," *Research Square*, 2022.

[13] M. R. Mahmood, A. M. Abdulazeez, and Z. Orman, "Dynamic Hand Gesture Recognition System for Kurdish Sign Language Using Two Lines of Features," in *2018 International Conference on Advanced Science and Engineering (ICOASE)*, Duhok, Iraq, 2018.

[14] M. R. Mahmood and A. M. Abdulazeez, "Different model for hand gesture recognition with a novel line feature extraction," in *2019 International Conference on Advanced Science and Engineering (ICOASE),* 2019: IEEE, pp. 52-57.

[15] A. D. Hashim and F. Alizadeh, "Kurdish sign language recognition system," *UKH Journal of Science and Engineering,* vol. 2, no. 1, pp. 1-6, 2018.

[16] P. Bo , N. Erik , and N. W. Ying "Deep Learning With TensorFlow: A Review," *SAGE Journals,* vol. 45, no. 2, September 10 2019.

[17] S. Saha, "A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way," *Towards Data Science*, 2018.

[18] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Zheng, X. (2016). *TensorFlow: A System for Large-Scale Machine Learning*. OSDI, 265-283.

[19] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). *LSTM: A Search Space Odyssey*. IEEE Transactions on Neural Networks and Learning Systems, 28(10), 2222–2232.

[20] Z. Zhang, H. Zhou, X. Xu, and Y. Liu, "Real-time hand tracking and gesture recognition with MediaPipe and TensorFlow," *2022 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2022, pp. 1527-1532.

[21] T. Yirtici and K. Yurtkan, "Regional-CNN-based enhanced Turkish sign language recognition," *Signal, Image and Video Processing,* vol. 16, pp. 1305–1311, 2022, doi: 10.1007/s11760-021-02082-2.

[22] G. Meena, K. K. Mohbey, and S. Kumar, "Monkeypox recognition and prediction from visuals using deep transfer learning-based neural networks," *Multimedia Tools and Applications,* pp. 1-25, 2024.

[23] G. Meena, K. K. Mohbey, S. Kumar, R. K. Chawda, and S. V. Gaikwad, "Image-based sentiment analysis using InceptionV3 transfer learning approach," *SN Computer Science,* vol. 4, no. 3, p. 242, 2023.

[24] R. R. Choudhary, K. Jisnu, and G. Meena, "Image dehazing using deep learning techniques," *Procedia Computer Science,* vol. 167, pp. 1110-1119, 2020.

[25] B. Yadav, A. Indian, and G. Meena, "Recognizing Off-line Devanagari Handwritten Characters Using Modified Lenet-5 Deep Neural Network," *Procedia Computer Science*, vol. 235, pp. 799-809, 2024, doi: 10.1016/j.procs.2024.04.076.