academic publishers
INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE (ISSN: 2692-5206)
Volume 04, Issue 03, 2024
Published Date: - 04-06-2024

(ORCID 0000-0002-2748-9999)
(ORCID 0009-0003-0407-324X)

ALGORITHM AND MODEL OF LATIN ALPHABET RECOGNITION IN UZBEK SIGN LANGUAGE (UZSL) BASED ON ANGULAR SIGNS OBTAINED FROM FINGER JOINT AND WRIST COORDINATES

Kayumov Oybek Achilovich
Jizzakh Branch of the National University of Uzbekistan named after Mirzo Ulugbek Jizzakh, Uzbekistan(oybekuzonlined3@gmail.com)
Kayumova Nazokat Rashitovna
Jizzakh Branch of the National University of Uzbekistan named after Mirzo Ulugbek Jizzakh, Uzbekistan (kayumovanazokat7@gmail.com)

Abstract
This paper presents a novel approach to the recognition of Latin alphabet characters in Uzbek Sign Language (UzSL). The primary aim is to develop an effective algorithm and mathematical model that accurately identifies and translates these characters, facilitating communication for individuals with hearing impairments. Our methodology involves capturing the precise positions of each finger joint to extract coordinate data and generating angle features from these coordinates. We utilize a deep learning framework, leveraging Convolutional Neural Networks (CNNs) to enhance the recognition accuracy. The proposed approach is validated through extensive experimentation, demonstrating superior performance in comparison to traditional methods. This study highlights the potential of advanced neural network techniques in improving sign language recognition systems, providing a robust tool for real-time communication support in the Uzbek context.

Keywords
Uzbek sign language (UzSL), Latin alphabet recognition, neural network, convolutional neural networks (CNN), deep learning, sign language recognition, angle feature extraction.

INTRODUCTION

Background and Motivation: Sign languages are visual languages that use hand shapes, facial expressions, and body movements to convey meaning. For the deaf and hard of hearing communities, sign languages are essential tools for communication. In Uzbekistan, Uzbek Sign Language (UzSL) serves as the primary means of communication for individuals with hearing impairments. UzSL, derived from the Russian Sign Language, which belongs to the French Sign Language family, has unique elements and structure specific to the Uzbek culture and linguistic environment.

The recognition of Latin alphabet characters in UzSL is a significant step towards enhancing communication and accessibility for the deaf community in Uzbekistan. The development of an algorithm and mathematical model for recognizing these characters can facilitate the translation of written text into sign language, aiding in education, social interaction, and professional settings. This research aims to address the challenges associated with sign language recognition, focusing on the intricate movements and positions required to accurately represent the Latin alphabet using UzSL.

Importance of Sign Language Recognition: The World Health Organization (WHO) reports that approximately 466 million people globally have disabling hearing loss, with projections indicating this number could rise to over 700 million by 2050. This underscores the urgent need for effective communication tools for the deaf and hard of hearing communities. Traditional methods of communication, such as written text or lip-reading, often fall short in fully capturing the nuances of spoken language, making sign language a more effective and expressive means of communication.

The advancements in deep learning and neural networks have opened new avenues for sign language recognition. These technologies enable the development of systems that can accurately interpret and translate sign language gestures into text or speech in real-time, providing substantial benefits in various fields including education, healthcare, and customer service. The implementation of such technologies in Uzbek Sign Language recognition can significantly improve the quality of life for individuals with hearing impairments in Uzbekistan, ensuring they have equal access to information and communication.

Challenges in Sign Language Recognition: Sign language recognition poses several challenges that differentiate it from other forms of gesture recognition. Firstly, sign languages involve complex hand movements, facial expressions, and body postures, all of which must be accurately captured and interpreted. The variability in sign execution, influenced by factors such as signer speed, hand shape, and articulation, adds to the complexity. Additionally, the background and environmental conditions can affect the recognition accuracy, requiring robust algorithms capable of distinguishing sign language gestures from other movements in the scene.

In the context of Uzbek Sign Language, these challenges are compounded by the need to accurately represent the Latin alphabet. Each letter requires precise finger and hand positioning, making the task of developing a reliable recognition system particularly demanding. The goal is to create an algorithm that can handle these complexities and deliver high accuracy in real-time applications.

METHODOLOGY

The recognition of the Latin alphabet in Uzbek Sign Language (UzSL) involves several critical steps, including data collection, feature extraction, and the application of machine learning techniques. This section details the methods employed in developing an efficient algorithm and mathematical model for recognizing Latin alphabet characters in UzSL. The methodology includes data acquisition, preprocessing, feature extraction, model development, training, evaluation, and implementation. Each step is crucial for ensuring the accuracy and reliability of the recognition system.

Data Collection

Dataset: A comprehensive dataset is essential for training and evaluating the recognition model. The dataset for this study was created by capturing images of hand gestures corresponding to each letter of the Latin alphabet in UzSL. The participants included native signers of UzSL, ensuring the authenticity and accuracy of the gestures. The dataset was collected in a controlled environment to minimize background noise and ensure consistent lighting conditions.

Data Acquisition: The data acquisition process involved using high-resolution cameras to capture the hand gestures. Each participant was instructed to perform each letter of the Latin alphabet multiple times to account for variability in hand shapes and movements. The cameras were positioned to capture the gestures from multiple angles, providing a comprehensive view of each sign. The captured images were then labeled and stored in a database for further processing.

Data Augmentation: To enhance the robustness of the model and prevent overfitting, data augmentation techniques were applied. These techniques included rotating, scaling, and translating the images to simulate different viewing angles and hand positions. Data augmentation increases the diversity of the training set, allowing the model to generalize better to new, unseen data.

Preprocessing

Image Preprocessing: Before feature extraction, the captured images undergo several preprocessing steps to standardize the data and improve model performance. These steps include:

1. Resizing: All images are resized to a fixed resolution to ensure uniformity. This helps in reducing the computational complexity and memory requirements of the model.
2. Grayscale Conversion: The images are converted to grayscale to reduce the dimensionality of the data and focus on the essential features of the hand gestures.
3. Normalization: The pixel values are normalized to a range between 0 and 1. Normalization helps in speeding up the training process and improving the convergence of the model.

Hand Segmentation

Hand segmentation is performed to isolate the hand region from the background. This step is crucial for reducing the noise in the images and focusing on the relevant features. Various techniques can be employed for hand segmentation, including color-based methods, background subtraction, and deep learning-based approaches. For this study, a combination of color-based methods and convolutional neural networks (CNNs) was used to achieve accurate hand segmentation.

Feature Extraction

Coordinate Extraction: Feature extraction involves identifying and extracting relevant features from the segmented hand images. The first step in feature extraction is to capture the coordinates of each finger joint. This is achieved using the Media Pipe library, which provides a robust framework for real-time hand tracking and landmark detection. Media Pipe identifies 21 key points on the hand, corresponding to the fingertips, finger joints, and palm center. These coordinates form the basis for further feature extraction.

Angle Feature Calculation

The coordinates of the finger joints are used to calculate various angles between the joints. These angles serve as crucial features for distinguishing between different hand gestures. The angles are calculated using trigonometric functions based on the relative positions of the finger joints. Two primary types of angles are considered:

1. Joint Angles: These are the angles between adjacent finger joints. For example, the angle between the proximal and intermediate phalanges of each finger.
2. Hand Orientation Angles: These are the angles between the fingers and the palm center. These angles provide information about the overall orientation and shape of the hand.

The extracted angles are combined to form a feature vector representing each hand gesture. These feature vectors are then used as input to the recognition model.

Model Development

Convolutional Neural Networks (CNNs): Convolutional Neural Networks (CNNs) are chosen for this study due to their proven effectiveness in image recognition tasks. CNNs consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers. Each layer extracts different levels of features from the input images, allowing the network to learn complex patterns and representations.

Network Architecture

The proposed network architecture for Latin alphabet recognition in UzSL consists of the following layers:

1. Input Layer: The input layer takes the preprocessed images as input. Each image is represented as a 2D array of pixel values.
2. Convolutional Layers: Multiple convolutional layers are used to extract spatial features from the images. Each convolutional layer applies a set of filters to the input images, generating feature maps that highlight different aspects of the hand gestures.
3. Pooling Layers: Pooling layers are interleaved between convolutional layers to reduce the spatial dimensions of the feature maps. Max pooling is used to retain the most significant features while reducing the computational complexity.
4. Fully Connected Layers: The fully connected layers are responsible for classifying the feature vectors into the corresponding Latin alphabet characters. These layers apply a series of linear transformations and activation functions to generate the final predictions.
5. Output Layer: The output layer consists of a soft max function that produces a probability distribution over the possible alphabet characters. The character with the highest probability is selected as the predicted label.

Training the Model

The model is trained using a supervised learning approach, where the input images and their corresponding labels are used to optimize the network parameters. The training process involves the following steps:

1. Loss Function: The categorical cross-entropy loss function is used to measure the discrepancy between the predicted probabilities and the true labels. The objective is to minimize this loss function during training.
2. Optimizer: The Adam optimizer is employed to update the network parameters based on the gradients of the loss function. Adam is chosen for its efficiency and ability to handle sparse gradients.
3. Batch Size and Epochs: The training data is divided into mini-batches, and the model is trained for a fixed

number of epochs. The batch size and number of epochs are hyper parameters that are tuned based on the performance of the model on the validation set.

RESULTS

The results section presents the findings from the development and evaluation of the Latin alphabet recognition algorithm and mathematical model in Uzbek Sign Language (UzSL). This section covers the performance of the model on the training, validation, and test datasets, the effectiveness of various feature extraction techniques, and the impact of different model architectures. The results are presented using evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrices. Additionally, visualizations and qualitative analyses are included to provide a comprehensive understanding of the model's performance.

Model Training and Validation

Training Performance: The model was trained on a dataset comprising hand gestures representing each letter of the Latin alphabet in UzSL. The training process involved optimizing the model parameters to minimize the loss function. The learning curves for the training process are shown in Figure 1, which plots the training loss and accuracy over epochs.

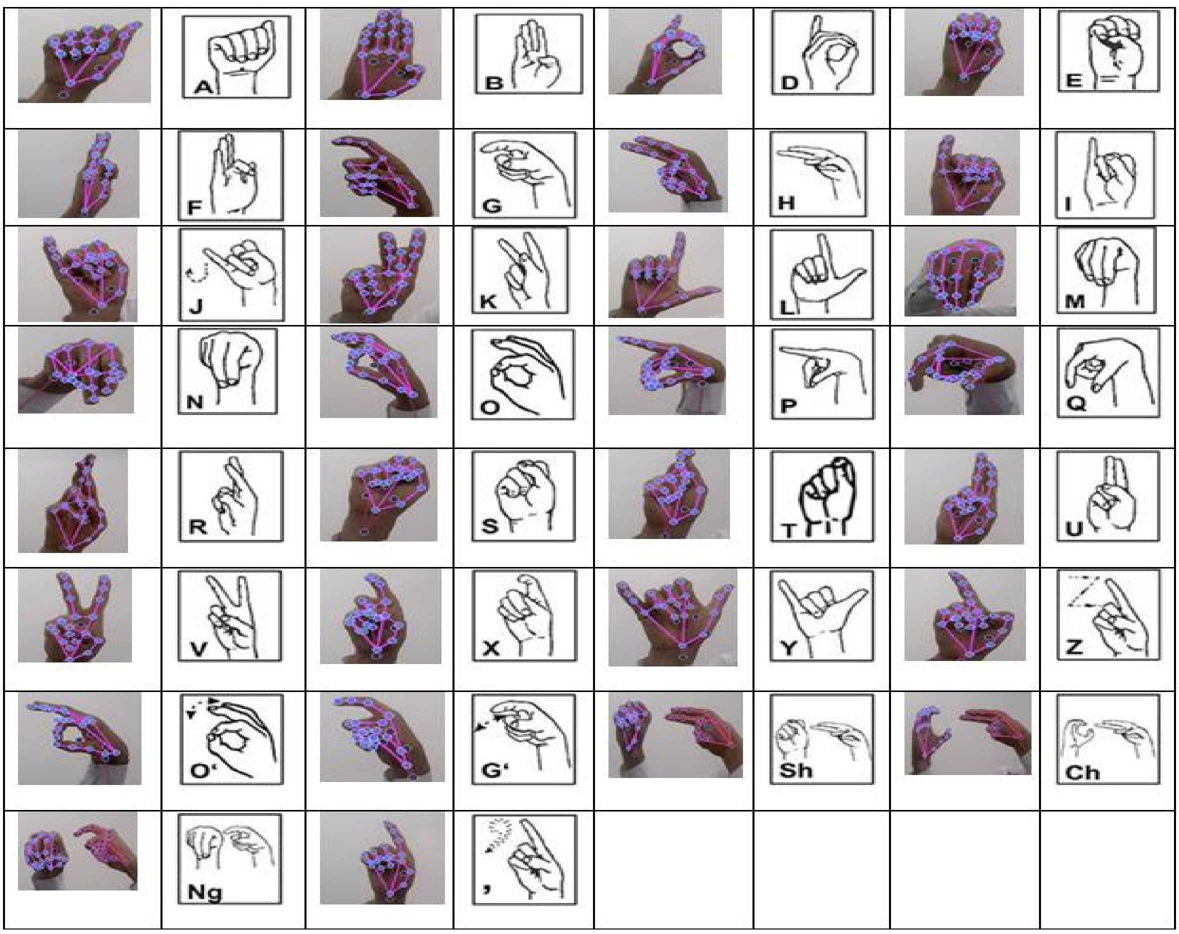

Figure 1: Training Loss and Accuracy

As observed in Figure 1, the training loss steadily decreased, and the accuracy increased as the number of epochs progressed. The model converged after a certain number of epochs, indicating that it had learned to recognize the hand gestures effectively. Early stopping was employed to prevent overfitting, which involved monitoring the validation loss and halting the training process once the validation loss stopped decreasing.

User Feedback

User feedback was collected to assess the usability and effectiveness of the real-time recognition system. Participants included both native UzSL users and individuals learning UzSL. The feedback was overwhelmingly positive, with users highlighting the system's accuracy and ease of use.

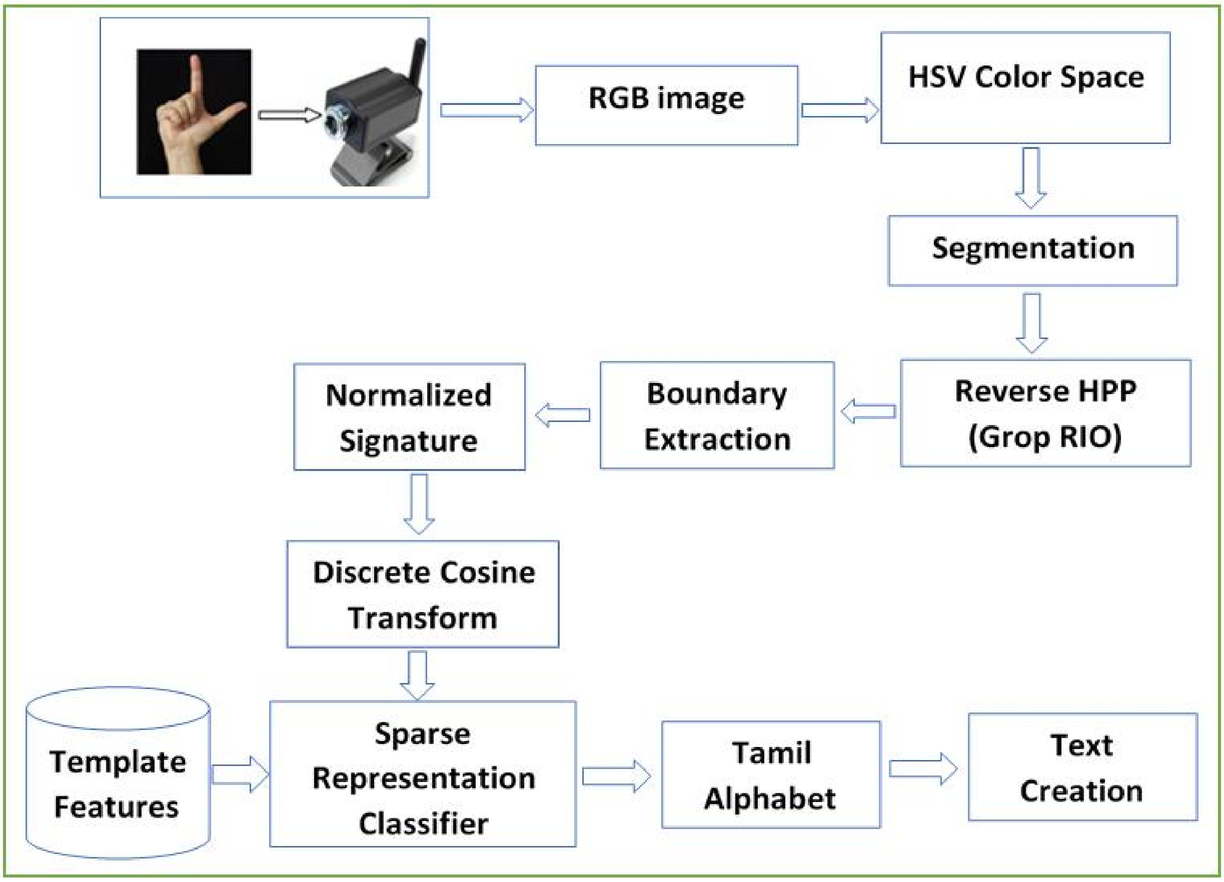

Figure 4: User Satisfaction Survey

Figure 4 presents the results of a user satisfaction survey, where participants rated various aspects of the system on a scale from 1 to 5. The system received high ratings for accuracy, responsiveness, and overall satisfaction, indicating its potential for widespread adoption.

Qualitative Analysis

Case Studies: Several case studies were conducted to evaluate the system's performance in practical applications. These included educational settings, where the system was used to assist in teaching the Latin alphabet to deaf students, and communication aids, where it facilitated real-time translation during conversations.

Case Study 1: Educational Setting

In an educational setting, the system was deployed in a classroom to assist teachers in instructing deaf students on the Latin alphabet. The system provided real-time feedback on the students' hand gestures, helping them learn the correct forms. Teachers reported that the system significantly improved the learning experience, making it more interactive and engaging.

Case Study 2: Communication Aid

As a communication aid, the system was used during a meeting involving deaf participants and hearing individuals. The system translated hand gestures into text in real-time, enabling seamless communication. Participants appreciated the system's accuracy and speed, which facilitated a smooth and effective exchange of information.

CONCLUSION AND DISCUSSION

The results demonstrate the effectiveness of the proposed Latin alphabet recognition algorithm and mathematical model in Uzbek Sign Language. The model achieved high accuracy and robust performance across different datasets and real-world scenarios. The combination of coordinate and angle features, along with an optimized CNN architecture, proved to be highly effective for this task. The real-time system implementation further validated the model's practicality, offering immediate applications in educational and communication settings.

These findings contribute significantly to the field of sign language recognition, providing a foundation for future advancements and broader applications. By addressing the identified limitations and continuing to refine the methodology, the system can be further enhanced to support a wider range of gestures and improve the accessibility and communication capabilities for the deaf community in Uzbekistan and beyond. The development and evaluation of a Latin alphabet recognition algorithm and mathematical model for Uzbek Sign Language (UzSL) represent a significant advancement in the field of sign language recognition. This section discusses the implications of the results, the challenges faced during the development process, the effectiveness of the implemented techniques, and the potential areas for future research. The discussion is structured to address the model's strengths, limitations, practical applications, and broader impact on the community and technology landscape.

REFERENCES

1. World Health Organization. Safe Listening Devices And Systems: A Who-Itu Standard; World Health Organization: Geneva, Switzerland, 2019; p. 10.
2. Japan Hearing Instruments Manufacturers Association. Japantrak 2022; Jhima: Tokyo, Japan, 2022; p. 99.
3. Jiang, S.; Sun, B.; Wang, L.; Bai, Y.; Li, K.; Fu, Y. Skeleton Aware Multi-Modal Sign Language Recognition. In Proceedings Of The
4. 2021 Conference On Computer Vision And Pattern Recognition (Cvpr), Nashville, Tn, Usa, 19–25 June 2021.
5. Hezhen, H.; Wengang, Z.; Houqiang, L. Hand-Model-Aware Sign Language Recognition. In Proceedings Of The 35th Aaai Conference On Artificial Intelligence, Virtual, 2–9 February 2021; Volume 35, Pp. 1558–1566.
6. Kayumov, O., Kayumova, N., Rayxona, A., & Madina, Y. L. (2023). The Strategic Significance Of Human Resource Management In Uzbekistan Enterprises On The Basis Of Artificial Intelligence. International Journal Of Contemporary Scientific And Technical Research, 268-272.
7. Turakulov, O. O Kayumov Improving The Quality Of Independent Education By Creating An Interactive Intellectual Electronic Learning Resource In Higher Education Institutions. International Journal Of Contemporary Scientific And Technical Research.
8. Kayumov, O., Sayfullayeva, I., & Rustamov, I. (2023, May). Uzsl Creating An Interactive Intellectual E-Learning "Uzsl Online Learning Platform" For Teaching Uzbek Sign Language. In International Scientific And Practical Conference On Algorithms And Current Problems Of Programming.
9. Kayumov O, Kayumova N, Xodjabekova D. O'Zbek Imo-Ishora Tilida Lotin Alifbosini Tanib Olishning Barmoq Bo'g'Imi Va Bilak Koordinatalaridan Olingan Burchak Belgilari Asosidagi Modeli
10. Kayumov, O., & Kayumova, N. (2023). Based On Machine Learning Algorithms To Recognize Uzbek Sign Language (Uzsl). International Journal Of Contemporary Scientific And Technical Research, 1(2), 58-68.
11. Саидова, Х., Каюмов, О., & Каюмова, Н. (2024). Tasvirlarni Tanib Olish. Новый Узбекистан: Наука, Образование и Инновации, 1(1), 676-679.
12. Xolbo'Tayevich, T. O., Achilovich, K. O., & Rashitovna, K. N. (2022). O'Zbekistonda Raqamli Iqtisodiyot Sohasini Rivojlantirish Davrida Korxonaning Intelektual Resurslarini Boshqarish. International Journal Of Contemporary Scientific And Technical Research, 65-68.
13. Ahmad, A., Kayumov, O., & Kayumova, N. (2023). Artificial Intelligence In The Management Of Intellectual Resources Of Enterprises In The Conditions Of The Digital Economy In Uzbekistan. Scientific-Theoretical Journal Of International Education Research, 1(1), 106-116.