AI- Powered Deaf Companion System for Inclusive Communication between Deaf and Hearing Individuals

I. Janani, S. Brindhashree, M. Boomika
PSNA College of Engineering and Technology, Dindigul, India

M. Sahana Preethi
Assistant Professor, Department of Information Technology
PSNA College of Engineering and Technology, Dindigul, India

Abstract
Sign language is the usual means of communication used by speech and hearing disabled people, and Bengali Sign Language (BdSL) is especially difficult with its large vocabulary, alphabet, and diverse expressions. Conventional remedies like mastering BdSL or employing an interpreter tend to be expensive and hard to obtain. Machine translation of sign language provides a more accessible solution, and deep learning, especially computer vision, holds much promise. To meet this, we created KUNet, a CNN-based classifier optimized with Genetic Algorithms (GA) to classify BdSL gestures.
KUNet attained 99.11% accuracy on the KU-BdSL dataset and surpassed a number of state-of-the-art models. Explainable AI (XAI) was also utilized to explain the model's decision process, increasing its transparency. The study seeks to develop a BdSL machine translator, which will benefit the non-verbal and hearing-impaired community since it will facilitate easy communication and less dependency on expensive interpreters.
Keywords: Bengali Sign Language (BdSL), Classification, Computer Vision, Deep Learning, Machine Learning, Sign Language Recognition, Optical Flow

OPEN ACCESS

Volume: 1

Issue: 1

July 2025 to December 2025

E-ISSN: Applied

Introduction

Daily, we perform many simple activities in our routine daily life, such as brushing our teeth or even conversing with a person. However, for most, such simple tasks involve a great amount of effort as a result of disability. Anyone who suffers from an unusual physical or mental status that hampers the normal capability of a human being is disabled, and his situation is disability. There are those who are born disabled, but there are those who acquire this later in their lives. Disabled individuals have it difficult regardless of the time used in developing a disability. Disabled people are mostly classified into categories such as hearing disability, visual disability, and mental disability. One may experience one disability, or they could experience many disabilities. World Health Organization (WHO) estimates that 1 billion people (15% of the global population) are impacted by a physical, sensory (e.g., hearing impaired or vision impaired), intellectual, or mental health disability that significantly impacts their daily life.

Among these large populations, speech and hearing impaired people face numerous challenges when communicating with the general people. We practically all encounter. We want to help them as much as they want assistance from us. The only barrier is the communication medium, which is primarily sign language. Being so essential for this enormous community, it lacks universality. It varies from one country to another. Again, it can have diversity within a geographical boundary, for instance, Bengali sign language (BdSL). The two broad groups of BdSL are - static sign language and dynamic sign language. These two groups can further differ depending on the involvement of the number of hands. However, due to the complexity of double-handed gestures, single-handed ones were used for BdSL alphabet representation.

The existing solutions for understanding speech and hearing impaired people is to learn sign language or to seek help from someone who has learned. Both take courage and effort to learn sign language, which are the primary obstacles to learning sign language. Moreover, patience and practice are required to learn and achieve expertise. In comparison to human learning, computers can learn quite quickly. Hence, we can solve the dilemma of learning sign language by adopting deep learning. A large population in Bangladesh is suffering from the curse of hearing impairment, resulting in around 2.4 million BdSL users in Bangladesh. Being essential for this large community, the BdSL datasets are underdeveloped. The state-of-the-art BdSL datasets are not as advanced as other sign languages.

We have worked on BdSL to contribute towards the betterment of the hearing impaired and non verbal people, which involves creating a novel deep learning model named KUNet, a classification framework optimized by the GA that identifies the hand gestures. The model is comparatively shallow in nature, as it has a simple architecture with a few layers. A shallow model has a few trainable parameters, a few layers simple construction, and easy interpretability. Also, a shallow model requires low time complexity and space complexity. Hence, the KUNet model can reduce the computational time. The optimization part of the model has been accomplished by aGA. To achieve better accuracy, the GA chose the model's hyperparameters. The model attained better accuracy, yet we were using a model that lacks interpretability. The use ofa black-box model could limit the contribution towards the D&D society. Even newer regulations (for instance, European General Data Protection Regulation restrict the adoption of black-box models.

Thus, in order to maintain the accessibility and usability of this research, we have used explainable AI (XAI) to handle sensitive data and interpret the model. XAI demonstrates the pixels responsible for the prediction. We can decide whether to trust the model's prediction from the XAI results. The principal contributions of the research are illustrated below.

- Recognition of the BdSL consonant using KUNet that might alleviate the difficulty of learning BdSL by providing ML-based translation.
- Exploration of the utilization of GA optimization in BdSL recognition, which is yet to be discovered.
- Use of XAI for interpretation of the outputs to enhance machine learning (ML)-based translation of BdSL.

The sole purpose of this study is to ease the communication. with the D&D people, and to do so, we have organized. GA optimization result, comparison with related works and state-of-the-art model, and XAI results.

Review of Literature

Inclusive communication between deaf and hearing people using AI-powered systems exploits the capabilities of the most recent technologies, like automatic speech recognition, natural language processing, and intelligent tutoring system to improve interaction and learning among deaf and hard of hearing students (Alkahtani, 2024). Some new features are specific applications, such as the use of AI using captioning, interpretations, and assistive applications, which facilitate overcoming communication obstacles and inclusivity (Alkahtani, 2024).

AI technologies also have an essential role in creating sign language applications. They entail techniques of sign language capturing, recognition, translation and representation. This demand can be attributed to the fact that accurate automatic sign language recognition systems will enhance communication with individuals who do not know sign language, making them communicate better with others (Papastratis et al., 2021; Zaineldin et al., 2024).

Besides, AI-based solutions in education deal with accessibility and encourage inclusive practices. To give an example, voice recognition software and adaptive learning environments allow every student to have access to their needs, allowing more participation and academic success based on AI (Adeleye et al., 2024; Mehta et al., 2023). The communication and collaboration tools offered by these technologies increase inclusivity because it gives the students the opportunity to share their views and expertise (Adeleye et al., 2024).

The use of AI technologies in the communication system does not only make people with disabilities powerful but fosters the inclusion of society. AI-driven tools contribute to reducing the gap between people with disabilities and their peers as their educational, employment, and social activities can be improved and help them feel less alienated (Eziamaka et al., 2024). However, all the challenges of data privacy, algorithmic bias, and accessibility that still exist are issues that cannot be understated because continual study and ethical use of AI systems are essential to get all the best out of them (Eziamaka et al., 2024).

Methodology

The GA-optimized model was used for classification tasks, and we compared our findings with other state-of-the-art research our BdSL.

A. Dataset

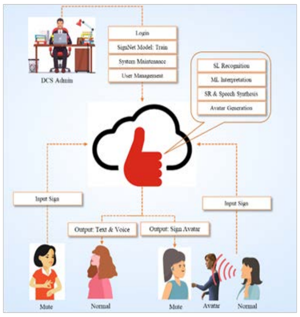
The dataset contains 30 classes with 50 images each, for a total of 1,500 512 × 512 pixel images, which were down sampled to 64 × 64 pixels to minimize processing time. The classes cover 38 Bengali consonants from the "banjonborno" alphabet. Multiple letters are represented by some hand gestures, thus the number of classes in BdSL is higher. Vertical and horizontal flipping, rotation, and contrast variation (50%, 20%, and 20%) were applied as image augmentation. The data was divided into 70% for KUNet training and 30% for testing.

B. Genetic Algorithm (GA)

Goldberg and Holland proposed the Genetic Algorithm (GA) for machine learning in 1988, simulating biological evolution. Since then, GA has been implemented in many optimization problems. For this research, GA was implemented to optimize A layer parameters of KUNet and the details of which are included. Models were trained to 25 epochs to minimize the cost of processing, and the accuracy was computed for every configuration of layers. Models were then sorted by accuracy, and the top 10% were selected for the next generation while the bottom 50% were discarded. The remaining 40% along with the top 10% took part in forming the next population. Two parent models were randomly chosen for crossover, mating, and mutation (C/M/M) operations.

Parameters that are Optimized in Different Layers

Layer	Parameter	Range of Parameter
Conv2D	No. of Filters, Kernel-size	32 to 2048 (multiple of 32), 1 to 5
Maxpooling	Pool-size	2 to 3
Dropout	Rate	5% to 50%
Activation	Function	ReLU or ELU or Softmax
Dense	Units	32 to 4096 (multiple of 32)



Structural Architecture

C. Explainable AI (XAI)

Explainable AI (XAI) or Interpretable AI allows humans to comprehend machine learning (ML) models. Local Interpretable Model Agnostic Explanation (LIME) is a well-known technique that is model-agnostic and applicable to images. LIME strives for "local fidelity," or the explanation will closely represent the model's behavior in the neighborhood of the instance being predicted. It achieves this by passing data through the model, manipulating different features, and generating a new local dataset. This data is utilized to train a surrogate model, which is weighted by the closeness of the samples to the instance of interest. For images, LIME randomly changes individual pixels, creating variants by dividing the image into superpixels and turning them on or off. Superpixels, a collection of connected pixels with the same colors, can be turned off by modifying the color of the pixels to a user-specified color, for example, black, with a specific probability of switching them off.

Result and Discussion

The KUNet model possesses the second-lowest number of parameters overall after ResNet50, with AlexNet possessing 5,183,328 more parameters than KUNet. KUNet takes 132.9 ms longer than AlexNet to classify 15 samples even though it has fewer parameters. ResNet50 is the quickest among the models, taking close to 4.5 times longer than KUNet, while VGG16 and VGG19 take close to six and seven times longer, respectively. KUNet also has a low classification time standard deviation (2.40ms), slightly more than AlexNet (2.07ms), and ResNet50 and VGG16 have nearly double, and VGG19 has 11 times the standard deviation. The KUNet model was trained with a Genetic Algorithm (GA) and enhanced its accuracy from 93.78% to 99.78% in the first 15 generations.

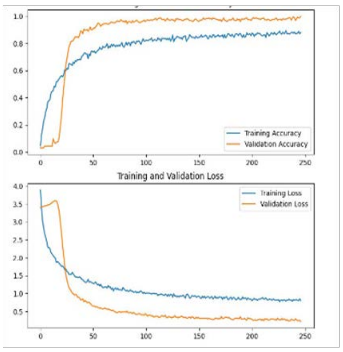


Figure Summary of the KUNet

The best optimized KUNet model, produced with generation 22 hyperparameters, was compared and performed better than other acclaimed models such as AlexNet, ResNet50, VGG16, and VGG19 when classifying images of the KU-BdSL database. The best accuracy, precision, and F-measure values were recorded in the model even when it exceeded AlexNet performance by a bit. A confusion matrix showed that the model performed close to perfect classification, with 100% recognition for all classes except one case where class 4 was wrongly classified. The training outcomes of accuracy and loss showed improved performance on test cases because of dropout layers in the network.

Sample and Responsible Features

The XAI results gave additional insights, indicating that the model was able to identify regions that were responsible for predictions, as evidenced by the LIME approach. For instance, the model was able to correctly predict a sample belonging to class 2480/2524/2525, detecting relevant image features. Nevertheless, in certain instances of the model was not able to learn the image features correctly, resulting in misclassification. In spite of these occasional breakdowns, the overall performance of the model in classifying KU-BdSL images far surpasses other deep learning models in terms of accuracy and effectiveness.

Conclusion

Sign language is crucial for communication among individuals who suffer from hearing and speech disabilities. Our novel KUNet classification model, developed with the help of deep learning, is a cost-effective and reliable solution, with a remarkable 99.78% accuracy. It surpasses popular deep learning-based models such as AlexNet, ResNet50, VGG16, and VGG19 in terms of primary performance metrics such as accuracy, recall, precision, and F1 score. The model's superior performance can be credited to the optimization of Genetic Algorithms (GA), an onetime exercise that makes the model both correct and trustworthy. Moreover, we applied Explainable AI (XAI) to explain the black-box nature of the model to ensure transparency and correctness in making decisions, particularly for the vulnerable Deaf and Dumb population.

Our work provides valuable contributions to the Bengali Sign Language (BdSL) community by bringing GA and explainable AI into the equation, which makes the model more reliable. It is the first piece of research on BdSL recognition to integrate these technologies, paving the way for limitless future expansions. Although the model provides encouraging results, there is a computational cost vs. optimization trade-off. When there are more training samples, there is a growing computational requirement that we intend to investigate further so that performance may be enhanced while not sacrificing precision.

References

Adeleye, O., Eden, C., & Adeniyi, I. (2024). Innovative teaching methodologies in the era of artificial intelligence: A review of inclusive educational practices. *World Journal of Advanced Engineering Technology and Sciences*, *11*(2), 69-79.

Alkahtani, B. N. (2024). The Impact of Artificial Intelligence on Quality of Life for Deaf and Hard of Hearing Students. *American Annals of the Deaf*, *169*(4), 329-347.

Eziamaka, N., Odonkor, T., & Akinsulire, A. (2024). AI-Driven accessibility: Transformative software solutions for empowering individuals with disabilities. *International Journal of Applied Research in Social Sciences*, *6*(8), 1612-1641.

Mehta, P., Chillarge, G. R., Sapkal, S. D., Shinde, G. R., & Kshirsagar, P. S. (2023). Inclusion of children with special needs in the educational system, Artificial Intelligence (AI). *AI-Assisted Special Education for Students With Exceptional Needs* (pp. 156-185). IGI Global.

Papastratis, I., Chatzikonstantinou, C., Konstantinidis, D., Dimitropoulos, K., & Daras, P. (2021). Artificial Intelligence Technologies for Sign Language. *Sensors*, *21*(17).

ZainEldin, H., Gamel, S. A., Talaat, F. M., Aljohani, M., Baghdadi, N. A., Malki, A., Badawy, M., & Elhosseini, M. A. (2024). Silent no more: A comprehensive review of artificial intelligence, deep learning, and machine learning in facilitating deaf and mute communication. *Artificial Intelligence Review*, *57*(7).