VAWKUM Transactions on Computer Sciences
http://vfast.org/journals/index.php/VTCS@ 2023 ISSN(e): 2308-8168, ISSN(p): 2411-6335
Volume 11, Number 1, January-June 2023 pp: 245-267

Robot Assist Sign Language Recognition For Hearing Impaired Persons Using Deep Learning

Kashaf Khan *, Dr Naeem Aslam , Kamran Abid , Safa Munir

[1] Department of Computer Science, NFC-IET Multan

Keywords: *Social Robot, Sign Language Recognition System, Deep Learning, NLP, Automated Response.*

Journal Info:
Submitted:
May 05, 2023
Accepted:
June 10, 2023
Published:
June 19, 2023

Abstract The adoption of Sign Language Communication (SLC) systems has become more significant in closing the interaction between the deaf society and the world of hearing people. In this study, researchers aim to contribute to this field by developing a system that helps sign language users communicate through BERT approaches based on deep learning frameworks as well as NLP. Accurate recognition of sign language is critical for SLC systems to work well. Deep learning models are effective in identifying sign language with high accuracy. This study aims to determine the most suitable DL model for identifying sign language and assess the impact of incorporating Natural Language Processing (NLP) techniques in generating frequent and accurate responses in SLC systems. The NLP model will be developed as an optimum return mechanism to generate frequent responses. This research includes testing three different deep learning models: MLP, CNN, and RestNet50v2 to recognize sign language gestures. RestNet50v2 outscored the other two approaches with a 0.97% perfection. As said earlier, the system also generates automated responses using the NLP BERT model, with an overall accuracy of 0.8% and a BLEU score of 0.83 %. This method has a way to enhance interaction among the deaf community via the use of technology, opening new avenues for developing intelligent chatbots that can better understand nonverbal communication. Further research can be done to expand its functionality to recognize a broader range of sign language gestures and improve the user interface. Overall, this study demonstrates how technology can enhance the ways of people with deafness or hearing loss by addressing communication barriers.

***Correspondence Author Email Address:**
2k20mscs113@nfciet.edu.pk

1 Introduction

Language is undeniably the most essential social and cultural element created by mankind throughout history. It allows people to communicate either orally (through spoken discussion) or nonverbally (by gestures, glances, body language, and so on) [1]. A myth was suggested that language reflects each nation.

If its relevance is studied further, it may help us understand and address all situations involving language and the attitude or behavior of sets of language speakers. Language is thought of as a framework, which is made up of fixed and patternable components. Language is traditionally defined as a tool for interacting or communicating, which in turn involves conveying concepts, beliefs, theories, or even feelings [2]. Humans need to communicate to exchange information and express their emotions. One of the human language's essential characteristics is that all spoken languages blend meaningless material to derive valuable symbols. Language is used not just for everyday speech but may also be a valuable goose for someone. Language is utilized for more than simply regular communication; it can also be an important asset for someone. Speech refers to the method used for transmission which can consist of a combination of vocalizations, vocabulary, syntax, and context. It serves as a means of exchanging ideas, thoughts, and feelings between individuals, communities, and cultures. Language is often shaped by social, historical, and geographic factors, resulting in a range of dialects and variations. It is an essential tool for conveying information and establishing connections, whether it is within a specific community, country, or profession [3].

Based on the Department of Health, Hearing loss affects 5 million individuals, accounting for approximately 5% of the global population., and requires some form of rehabilitation [4]. More than 700 million people are expected to suffer debilitating hearing loss by 2050. Faint, moderate, intense, and extreme are the different degrees of hearing loss. People classified as "hard of hearing" experience a range from modest to severe hearing loss, while those "deaf" experience a significant loss of hearing. It is a prevalent chronic condition that often results from sensory cell degeneration associated with aging. However, it is treatable with conventional hearing aids and communication gadgets [5]. Persons with normal hearing can listen and converse, but deaf people cannot hear and, if born deaf, cannot even speak [6]. Although many deaf people, particularly those who developed their hearing disability after infancy, can converse effectively and may use skills such as lip-reading when conversing with hearing people, such communication approaches are frequently insufficient for communication within the Deaf community. Deaf and hearing-impaired youngsters endure even more difficult language comprehension issues than their hearing counterparts. People who are deaf have severely limited social connections. To participate in social settings, Deaf individuals often rely on means such as body language interpretation or sign language to communicate. For deaf children, it is crucial to develop early literacy abilities and become familiar with written language. Without these competencies, they may experience limited social involvement with peers and family, negative perceptions from their family members, immaturity, dependence on others, poor language proficiency, and a range of cultural, behavioral, and communication challenges [7]. As a result, the usage of body parts has become the primary mode of communication in these cultures [8]. This is referred to as nonverbal communication.

Nonverbal communication is the act of expressing or conveying messages without using words. It plays an important role in human connection, and various nonverbal communication techniques are used, including hand signals and gestures, facial expressions, posture, eye contact, and more. Hand signals and gestures are nonverbal communication methods that involve hand movements to convey a message. Sign Language (SL) serves as a communication bridge between individuals who are deaf and those who can hear, utilizing gestures to convey messages[9]. Expressive language involves conveying feelings through body gestures, particularly using the arms and hands. This method is employed when verbal communication is difficult. It aids interaction with the deaf HOH [8]. SL is a sensory language utilized by many Deaf individuals to communicate. Sign language involves using physical movements and facial expressions, as well as certain body positions, during human-to-human conversations, as well as on social media and television. Many hearing-impaired people, as well as those with other speech impairments, utilize Sign Language as their first language [10]. Both physical gestures, such as palm and body movements, including facial expressions, are used to convey meaning in sign language. These components are combined to form expressions that communicate specific ideas or concepts [7]. The Deaf community needs to understand and grasp the association between words and phrases, to lead us to a time when automated translation between syllables and sentences is possible [11]. This research aims to explore the potential of nonverbal communication in enhancing social interaction between humans and robots. The study delves into the possibility of improving the connection between social robots and people by investigating the impact of nonverbal communication [12].

The study of social robotics falls under the umbrella of human-robot interaction (HRI) and has garnered significant interest in recent times due to the widespread integration of robots into daily activities. Interactive robots are being developed for practical applications including rehabilitation, eldercare, and educational purposes. Consequently, there has been a surge in research related to inter-species interaction, resulting in the emergence of diverse subtopics within the HRI research domain [13]. Conversely, social robotics has experienced noteworthy progress, presenting new avenues for research and innovation in creating robots that can engage with humans more naturally and socially [14].

Deep learning emerged as the best tool for many fields such as cancer detection [15], brain disorders [16], covid detection [17], pattern recognition[18] and sign detection [19]. Recognizing sign language gestures using deep learning methods can be a challenging task, primarily due to the diverse range of sign languages and the scarcity of large, annotated datasets available for training the models. However, recent advancements in machine learning and artificial intelligence have facilitated the automation and enhancement of such technologies. The method of SLR involves converting a user's hand signals and movements into a textual form, which helps bridge the communication gap between those who cannot communicate through spoken language and the wider community. This is accomplished through classification procedures and neural networks, trained to interpret gestures, and convert them into easily understood text. The raw image and video data is converted into readable text, making it possible for individuals to communicate effectively without the need for verbal communication [20]. Sign language recognition (SLR) aims to develop advanced machine learning algorithms capable of accurately categorizing human articulations into individual signs or continuous sentences. However, the accuracy and generalization capabilities of SLR algorithms are currently limited by the absence of large, annotated datasets, as well as the challenges posed by recognizing sign boundaries in continuous SLR scenarios. Addressing these limitations is crucial for the development of robust and accurate SLR systems that can effectively create a means of communication that can overcome the barrier between those who practice sign language and persons who do not [11]. SLR has remained the scientific field that captures and translates sign language using computer vision artificial intelligence algorithms [5]. The bulk of research efforts on sign language processing (SL) concentrates on its recognition, with little attention paid to detection in its true sense. To detect sign language, a deep neural network (DNN) is a commonly used deep neural network (DNN)-based model demonstrated as a successful method for sign language detection in various tasks. Ongoing research is focused on enhancing the precision and speed of these models for improved effectiveness [21]. The fact that most sign language datasets are created for sign recognition and are gathered in controlled circumstances adds to the lack of acceptable publicly accessible data for the identification of sign language. As a result, researchers have turned to use sample groups obtained from Kaggle Multiple Combination datasets. This approach allows for the use of real-world data, but it also presents challenges such as variability in lighting, background, and signing speed, which can affect the accuracy of sign language identification models. Therefore, efforts are being made to develop more comprehensive and diverse datasets for sign language detection and identification. Continued research endeavors to enhance the precision and effectiveness of these systems, enabling their application in a wider range of sign language recognition tasks.

This study's primary issue is identifying the best possible ways to detect sign language with the highest possible accuracy. It also focuses on the response a hearing-impaired person gets from a robot by utilizing NLP and then assessing the correctness of that answer. Here are some points on what researchers will do overall in this research:

- Collect and preprocess the ASL dataset: Researchers will gather and preprocess a massive collection of sign language pictures to meet the first study goal. Deep learning algorithms will be trained and evaluated using this dataset.
- Implement and compare deep learning models: To determine the most effective deep learning model for sign language detection, researchers will implement and compare three different models: MLP, CNN, and RestNet50 v2. The models will be trained on the ASL dataset, and their performances will be compared based on accuracy and other evaluation metrics.
- Develop an NLP model: To generate responses in sign language communication, authors will develop

an NLP model using BERT. This model will be optimized for the best return mechanism to generate frequent and accurate responses.
- Integrate deep learning and NLP models: Once the deep learning and NLP models are developed, researchers will integrate them to create a sign language communication system that can recognize sign language and generate automated responses.
- Assess the system's effectiveness: To assess the impact of incorporating NLP in sign language communication systems. The authors will perform experiments to test the precision and efficacy of the suggested system. Researchers will also analyze the impact of NLP on the system's overall performance.

Overall, the research aims to develop an innovative sign language communication method that can help people with hard of hearing communicate more effectively. Using deep learning and NLP, researchers hope to achieve higher accuracy and efficiency in sign language recognition and response generation.

2 Literature Review

In conducting the literature review, a systematic approach was employed to facilitate an unbiased evaluation of the paper's substance. This method ensured that the analysis was conducted in a rigorous and structured manner, adhering to established best practices in the field. To ensure a rigorous and objective assessment of the paper's content, a standardized procedure was employed, which encompassed several tasks. These tasks involved identifying relevant variables, delineating the authors' strategic approaches, scrutinizing the methodologies employed to obtain the results, and culminating in a comprehensive report. A pre-set of factors was used to choose which scientific papers to incorporate into the acquisition process [22]. In this section, a bibliometric analysis is presented to examine the analysis of the use of autonomous machines in sign language identification during the previous twenty years. This analysis aims to use bibliometric methods to examine recent publications on the topics of intelligent machines and SLR. Another objective was to determine patterns and trends in these publications, including the journals in which they are published, the regions where the research is conducted, and collaborations between different institutions and organizations. The interest in automatic sign language recognition has increased over the years, and therefore, investigating the primary focus of research and existing patterns in this field would be significant [23].

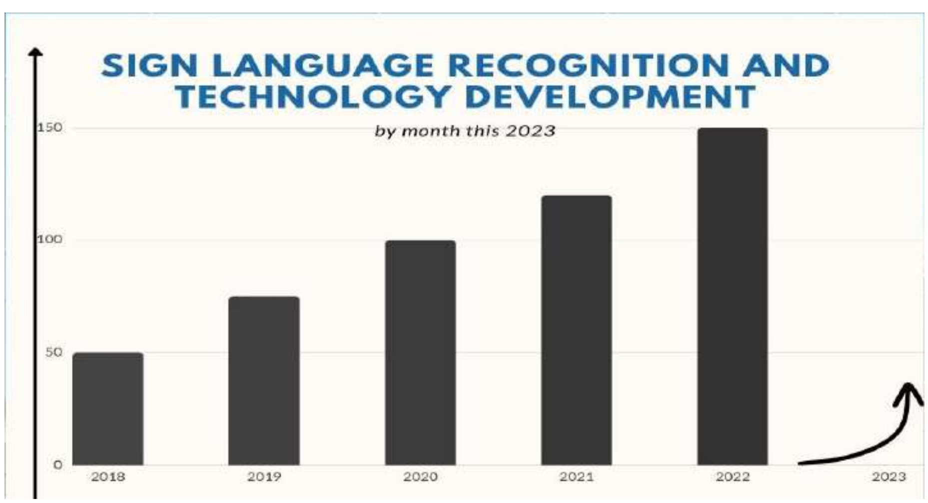

Figure 1 Estimated annual worldwide technology development to support sign language

2.1 Deep Learning in recognizing Sign Language

In this research, we performed an analysis of scholarly publications related to the use of intelligent machines in sign language recognition from 2018 onwards, as shown in Figure 1. The primary objective was to systematize the existing research and identify patterns and publication trends based on journals, regions, and institutional collaborations. After this initial phase, we employed a forward/backward technique to thoroughly analyze the collected materials. By adopting this approach, we were able to comprehend each paper in detail and track crucial research areas, preventing the omission of any fundamental studies. By employing this method, we were able to create a representative collection of SLR papers that accurately reflected the most fruitful research directions in the field [22]. In the paper [24], the authors propose a hybrid approach for recognizing Bangla Sign Language using a deep transfer learning model combined with a random forest classifier. The recognition rates achieved are 72% -85% for character recognition and 95%-96% for digits recognition. However, the study acknowledges that there is room for improving the accuracy of the model and highlights the limitation of limited dataset availability. In their research, the authors introduce "Sign Explainer[25]," a framework that combines explainable AI techniques with ensemble learning for sign language recognition. The framework achieves an impressive overall recognition rate of 98% and 92.60%. However, one limitation of the study is the limited utilization of Explainable AI (XAI) methodology, which could have provided deeper insights into the model's decision-making process. In their research, the authors propose a deep learning model called "SIGNFORMER[26]" that utilizes a hybrid MLP architecture for sign language recognition. However, the model achieves a relatively low recognition rate of 0.9%. The study also identifies two main limitations: the absence of preprocessing techniques for input data and the limited availability of a comprehensive dataset. These limitations may impact the model's performance and generalization capabilities. The paper presents a novel approach for unsupervised speech-to-sign language recognition called "Speak, Decipher and Sign.[27]" The proposed method combines a hybrid architecture consisting of MLP and LSTM deep learning algos. This technique achieves a recognition rate of at least 0.6%. However, the study acknowledges a limitation regarding the dataset, which is relatively small. This limitation may impact the model's performance and its ability to generalize to a wider range of sign language data.

Table 1 DL Literature Review

Title Name	**Publication Year**	**Mode of Acquisition / Datasets**	**Techniques Employed**	**Rate of recognition**	**Limitations**
A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier[24]	2023	Bengali Sign Language	Hybrid (RestNet)	Char (72%-85%), Digits (95%-96%)	Accuracy can be Improved, Limited Dataset
Sign Explainer: An Explainable AI-Enabled Framework for Sign Language Recognition With Ensemble Learning[25]	2023	Indian Sign Language (43k)	RestNet50, CNN etc.	98%, 92.60%	Limited use of AI(XAI) Methodology
SIGNFORMER: DeepVision Transformer for Sign Language Recognition[26]	2023	ISL	Hybrid MLP	0.9%	No preprocessing, Limited Dataset

Recognition[26]					
Speak, Decipher and Sign: Toward Unsupervised Speech-to-Sign Language Recognition[27]	2023	ASL	Hybrid (MLP+LSTM)	>=0.6%	Dataset relatively small
Real Time Sign Language Recognition System for Hearing and Speech Impaired People [28]	2022	Web Camera	3D CNN	---	Accuracy not shown, phase detection
American sign language recognition technique using 3-D convolutional neural networks[29]	2021	Camera	3D CNN	---	Accuracy not mentioned, Limited video data
Realtime isolated hand sign language recognition using deep networks and SVD[30]	2021	ASLVID Dataset	CNN, LSTM, SVD	61.04%	Low recognize rate, consume more execution time
Real Time Hand Gesture Recognition Using Different Algorithm Based on American Sign Language[31]	2020	Camera	CNN	94.32%	Limited Data (alphabets), Advanced Algo Required
Generative Multi-Stream Architecture for American Sign Language Recognition [32]	2020	RGB Dataset	CNN	95.62%	Video Dataset only, Model only applicable for RGB Dataset
A Multi-dimensional Feature Learning Approach Based on Deep Convolutional Networks for Gesture Recognition in RGB-D Image Sequences[33]	2020	Camera	CNN and LSTM	72.53%, 95.87%	Real time detection, Multi DCNN
Score-level Multi Cue Fusion for Sign Language Recognition [34]	2020	Camera	CNN	94.96%	Average hand shape recognition
Phonologically meaningful Subunits for Deep Learning-based Sign Language Recognition[35]	2020	Camera	RNN, NN	0.281	error
Hand sign language recognition using multi-view hand skeleton [36]	2020	Camera (SSD)	3DCNN and LSTM	99.80%, 4.64% (error)	Video Dataset only,
A Modified-LSTM Model for Continuous Sign Language Recognition using Leap motion [37]	2019	Camera/Leap Motion Sensor	Modified LSTM	72.3%	Limited Data (942 sign imgs)

Deep Learning for American Sign Language Fingerspelling Recognition System [38]	2019	Camera	CNN, CNN+SVM	97.08%, 98.30%	Limited Dataset (24 signs)
Isolated sign language recognition using Convolutional Neural Network hand modelling and Hand Energy Image[39]	2019	Data gloves, sensors	CNN	89.33%	Less recognition rate even using HEI
Spatial-Temporal Graph Convolutional Networks for Sign Language Recognition[40]	2019	ASLLVD	GCN	61.04%	Average Accuracy Rate
Construct Dynamic Graphs for Hand Gesture Recognition via Spatial-Temporal Attention[41]	2019	Camera	GCN	91.9% and 94.4%	Limited Dataset (2800 videos)
SignFi: Sign Language Recognition Using WiFi[42]	2018	Wifi /web cam	CNN	98.01%	Fluctuate accuracy, Advanced Model not included (RNN+ LSTM)
A convolutional neural network to classify American Sign Language fingerspelling from depth and colour images[43]	2017	Camera	CNN	82%	Average Accuracy, Limited Data
A hand gesture recognition sensor using reflected impulses[44]	2017	Impulse radio sensor	CNN	>90%	Average Recognition Accuracy, Limited Hand Gestures
Sign language recognition[45]	2017	Camera	ANN	85% (histogram features)	Average Result
Multimodal learning for sign language recognition[46]	2017	Kinect	CNN	97%	Limited Methodologies
Deep learning in vision-based static hand gesture recognition [47]	2017	24 ASL hand gestures	CNN	91.33%	Limited Hand Gestures

2.2 NLP in recognizing Sign Language

Special children's learning styles are diverse because they require various sorts of educational needs. These youngsters cannot be taught traditionally. Intelligent tutoring systems that include AI, ML, and DL may motivate and reinforce youngsters in an educational and learning setting. The efficiency of three distinct approaches for detecting MNIST Sign Language is compared in this article: the (CNN) process, the Support Vector Machine (also known as SVM) methodology, the decision tree method. The results demonstrate that the CNN technique delivers up to 86% accuracy, while the Support Vector Trained model achieves an accuracy of 84% among machine learning algorithms [48]. The purpose of this research [49] is to perform a scoping review of the utilization of AI in instant translation applications from spoken language to sign language. The goal is to provide an AI-based technique to turn South African Sign languages from monologue-to-text to sign language, allowing for interactive interaction between those who can hear and those who cannot. The study revealed a lack of information on the utilization and approval of machine learning and other related methods as potential solutions for the deaf and hard-of-hearing (HOH) population, specifically in Africa.

Table 2 NLP Literature Review

Title Name	Publication Year	Purpose	Method	Accuracy of a Model	Limitation
Role of Machine Learning and Deep Learning in Assisting the Special Children's Learning Process[48]	2022	ITS (Intelligent Tutoring System) that helps special students in recognizing Sign Language	ML (Decision Tree, Adaboost Classifier, SVM, Random Forest)	44%,32%, 61%, 81%	Average accuracy, Limited Dataset
			CNN	86%	Limited Dataset
A proposed artificial intelligence-based real-time speech to-text to sign language translator for South African official languages for the COVID-19 era and beyond: In pursuit of solutions for the hearing impaired[49]	2022	Speech-to-text to Sign Language	AI AND ML(NCHLT)	-----------	Result not display
			ANN, CNN, RNN		
Deep learning based assistive technology on audio visual speech recognition for hearing impaired [50]	2022	AVSR (Audio Visual Speech Recognition)	RNN-GRU, CNN	95%	Modern methods can be used (BERT)
EasyTalk: A Translator for Sri Lankan Sign Language using Machine Learning and Artificial[51]	2020	2-way process (sign-to-text/graphics + vice versa)	Fastest RNN, CNN, NLP & Semantic Analysis	80%	System stops halfway, invalid Dataset
2-way Arabic Sign Language Translator using CNNLSTM Architecture and NLP[52]	2020	2-way process (Text-to-sign+ vice versa)	Hybrid model: CNNLSTM	88.67%	only capable of translating single dynamic expressions
Literation Hearing Impairment (I-Chat Bot): Natural Language Processing (NLP) and Naïve Bayes Method[53]	2019	Conversation mechanism Chatbot is constructed	Naïve Bayes	88.75%	Limited Dataset
Sign Language Recognition System Using Deep NN [54]	2019	Vision based system	Adam, SGD, CNN	99.12%	It is complicated to use
Intelligent Mobile Assistant for Hearing Impairers to Interact with the Society in Sinhala Language[55]	2018	Instant Messaging App (text-to-SL + Vice versa)	2D GIF, Semantic Analysis, TTS Engine, ML, NLP(App)	>90%	The file format does not work.
A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors[56]	2016	Wearable System	Sensors (IMU, sEMG)	96.16%	Talking and using hand-held sensors would not offer the identical precision level, Limited Dataset (80)

3 Methodology:

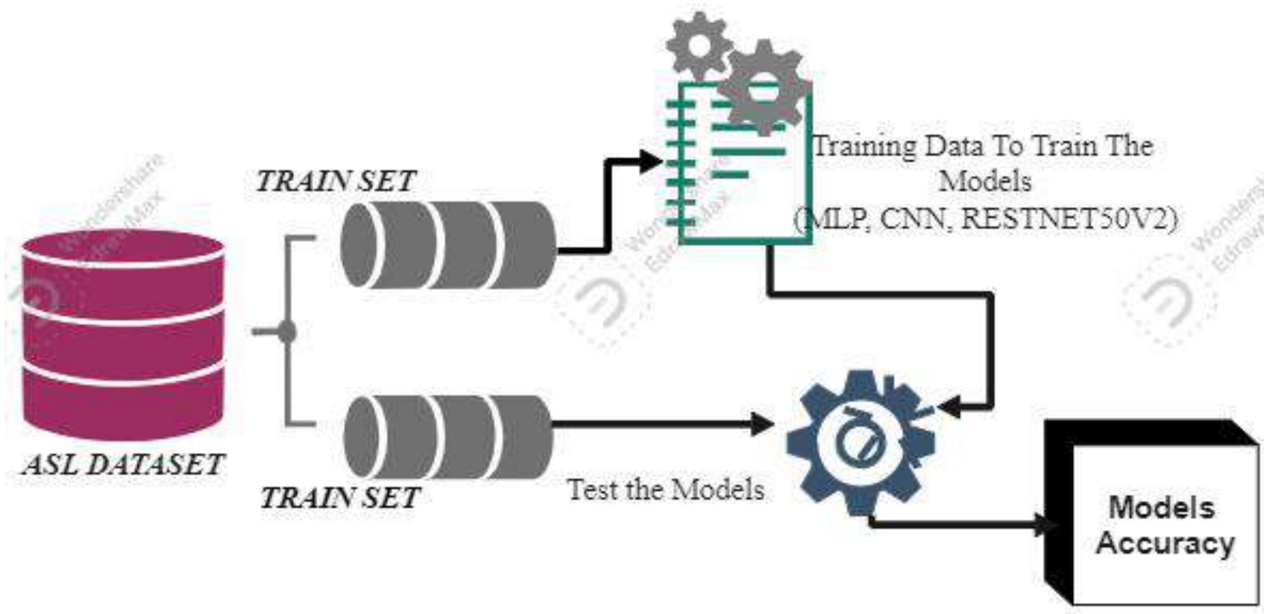

Figure 2 Working of Deep Learning Models in Recognizing ASL.

3.1 Dataset

A dataset is a data collection organized into a specific format for analysis. One's dataset is the information that is used to train machine-learning models. A dataset can come in many different forms, including text, images, audio, and more. It is essential to understand the dataset to ensure that the machine-learning model is accurate and efficient. The dataset used is related to the task trying to solve using machine learning. It could detect from recognizing handwritten digits to detecting objects in images. It is crucial to have a clear understanding of the data and the problems that individuals are trying to solve before building any machine learning model. Our dataset includes a set of images that serve as training data for our machine learning model. Each image in the dataset is associated with a label that identifies the object or entity represented in the image. These labels are used to train our model to recognize similar objects or entities in new images. The size and diversity of the dataset are crucial to ensure effective training of the machine learning model. However, working with a large dataset can present challenges such as increased computational requirements and the need for efficient data storage and retrieval. To use the dataset for machine learning, it must first undergo preprocessing to clean and transform the data into a format that can be readily consumed by our model. Preprocessing may include tasks such as resizing images, removing noise, and converting the data to a numerical format. In conclusion, your dataset is a critical component of your machine-learning project, and it is essential to understand it thoroughly. A well-prepared dataset can make a significant difference in the accuracy and efficiency of your machine-learning model. It is essential to preprocess your dataset carefully to ensure that it is optimized for use in machine learning. This database comprises a well-organized and labeled collection of image data containing American

Sign Language alphabets [57]. The data is sorted into 39 separate folders, each representing a different class. The training dataset contains a whopping 117,000 images, each measuring 200x200 pixels. The 26 folders represent the 26 alphabets from A-Z, while the remaining three folders contain images for SPACE, DELETE, and NOTHING and nine folders contain whole numbers of digits dataset as well (0-9). These three folders are of immense value for real-time applications and classification purposes. Additionally, a test dataset consisting of 39 images is provided to facilitate the use of real-world test images. Akash Nagaraj created the set, which is now available for public viewing. The hand movements exhibited in the photos were observed to be signed against a black background. The project's goal was to generate a dataset of real-world test images. Personal images were taken during the process. The hand movements exhibited in the photos were then signed against a black background. It is not necessary to purchase any equipment such as gloves or a particular marking system with this system, and all essential photographs can be prepared with a single basic camera. The suggested system's goal is to speak with deaf individuals in addition to conversing with computers and robots, which were employed in previous studies.

3.2 Data Preprocessing

The American Sign Language (ASL) dataset, which consisted of 39 separate classes, each with 3000 pictures, was prepared for further analysis during the data preparation phase of this work. Several critical preprocessing procedures were used to assure the dataset's quality and appropriateness. To begin, the RGB photos were converted to grayscale, lowering the density of the data while keeping the required visual representation for sign language interpretation. To enable successful training of the model, the dataset was normalized, which means that the picture pixel values were rescaled within a common range, often between 0 and 1. This normalization phase speeds up the convergence of neural network models and prevents any individual characteristics from overwhelming the training process. Additionally, methods for data augmentation were used to boost the dataset's diversity and amount. To create fresh training examples, random modifications such as rotation, translation, scaling, and flipping were used. Data augmentation reduces overfitting by adding variability to the model and improving its capacity to generalize patterns from the existing information. Finally, image standardization was applied to all photos in the collection to guarantee consistency and comparability. This entailed scaling all photographs to a standardized resolution, cropping, or padding them to maintain a consistent aspect ratio, and modifying overall image quality to minimize differences caused by lighting or image capture settings. The ASL dataset was adequately prepared through these data pretreatment methods, allowing for future model training and assessment for successful sign language recognition as well as classification tasks.

3.3 Deep Learning Model in Recognizing SL

Sign language is a form of nonverbal communication that employs hand gestures, facial expressions, and body language to express ideas and convey meaning. It is widely used by deaf and hard-of-hearing individuals across the world to interact with others. Recognizing sign language poses a vital challenge to researchers because of its complexity and variation. However, recent advances in deep learning methods[19], such as recurrent neural networks (RNN), long short-term memory (LSTM), artificial neural networks (ANN), and spiking neural networks (SNN), have shown considerable potential in identifying and interpreting sign language. Deep learning is a subset of machine learning that involves the training of artificial neural networks using huge amounts of data to perform complex tasks, such as speech recognition, natural language processing, and image recognition. Deep learning models can learn patterns from the data and use them to make accurate predictions [54].

3.3.1 Sign Language Recognition using MLP

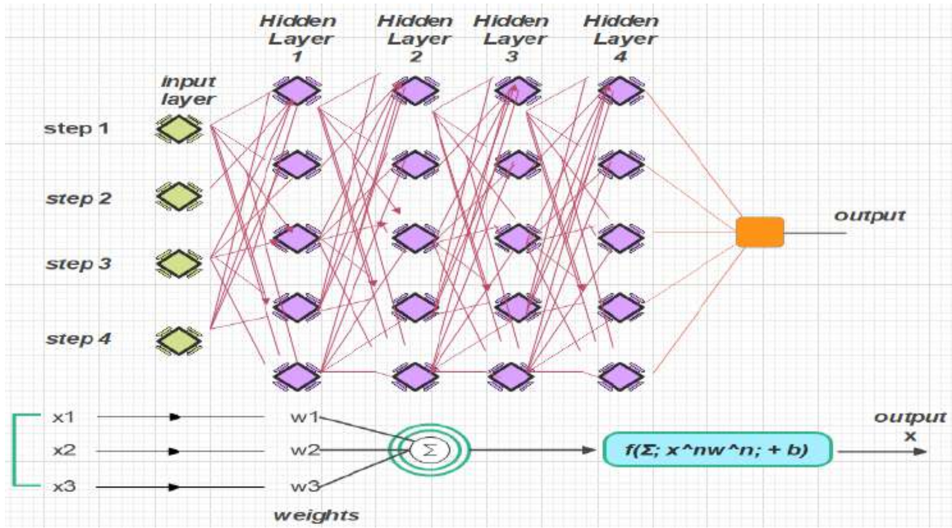

Figure 3 MLP Model Diagram

SLR is a crucial field of study that tries to refine transmission between deaf or hearing persons. The Multi-Layer Perceptron (MLP) models [58], a kind of neural network that could recognize patterns in data, are one method of solving this challenge. The code provided illustrates a demonstration of an MLP model with numerous layers of neurons and multiple activation mechanisms utilised in sign language recognition. The classified cross-entropy loss function and the Adam optimizer are used to train the model, and accuracy measures are used to assess its performance. MLP models can recognise sign language motions with a high degree of accuracy by training on an ample collection of sign language pictures fine-tuning the model parameters. The offered code is an MLP based neural network model for recognizing sign language. The system consisted of a sequential network of neurons with 10 layers, 3 dense layers, each with 200 units and the stimulation mechanism "relu", 2 dropout layers, each with a 10% dropout rate, and 4 further dense layers, each with fewer units and a more sophisticated activation function. The final output layer generates chances for 39 different classes using the "softmax" activation function. The model uses the categorical cross-entropy loss function and the Adam optimizer. The model is trained with both validation and training generators using the fit_generator() technique. A development rate planner callback function is used to control the growth rate during training, and the total quantity of eras is set to 30. In conclusion, the plotCurves() going to plot an accuracy graph at the end .

3.3.2 Sign Language Recognition using CNN

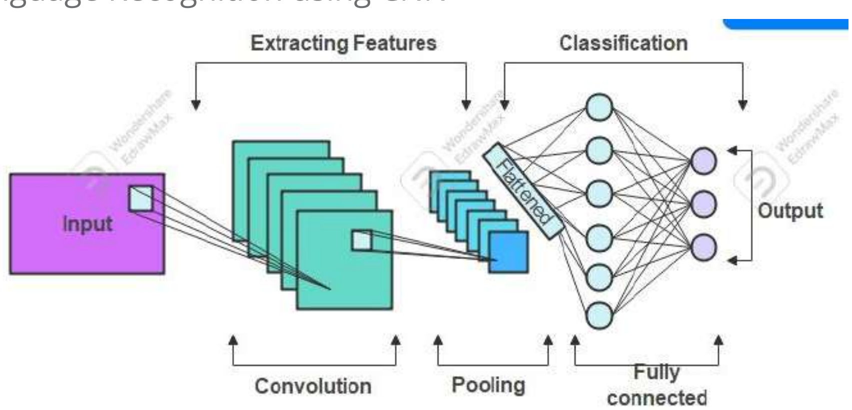

Figure 4 CNN Model Diagram

Convolutional Neural Networks (CNNs) are a powerful type of deep learning model that excels in image recognition tasks, including sign language recognition. CNNs learn spatial hierarchies of features from raw image data, allowing them to identify and classify objects within an image[28]. In the context of sign language recognition, CNNs can be trained to identify the hand gestures and movements that correspond to letters, words, digits, or phrases in American Sign Language (ASL). The gestures can be represented as image data, either as still images or as sequences of images captured in real time. The architecture of a CNN consists of convolutional layers, pooling layers, and fully connected layers [59]. The convolutional layers use small filters to scan across the input image and identify patterns and features at various scales. The pooling layers then reduce the spatial dimensionality of the feature maps, extracting the most valuable information while discarding redundant details. Finally, the fully connected layers use the learned features to classify the image into one of several categories. To recognize sign language, a CNN can be trained using labeled image data, with each image corresponding to a particular sign or gesture. The network can then learn to identify the key features of each sign and use these to accurately classify new, unseen images. This approach offers a flexible and accurate way of recognizing a wide range of gestures and movements, with potential applications in assistive technology and human-computer interaction. The code initializes a CNN model using Keras' Sequential API and adds layers to the model in a linear order. The layers consist of convolutional layers with ReLU activation, MaxPooling2D layers, a Flatten layer, and Dense layers with ReLU and softmax activation. The model is indoctrinated using the categorical cross-entropy loss function, Adam. CNN is designed to recognize hand-written digits with high accuracy by extracting features from input images and using them to make predictions.

3.3.3 Sign Language Recognition using RESTNET50V2

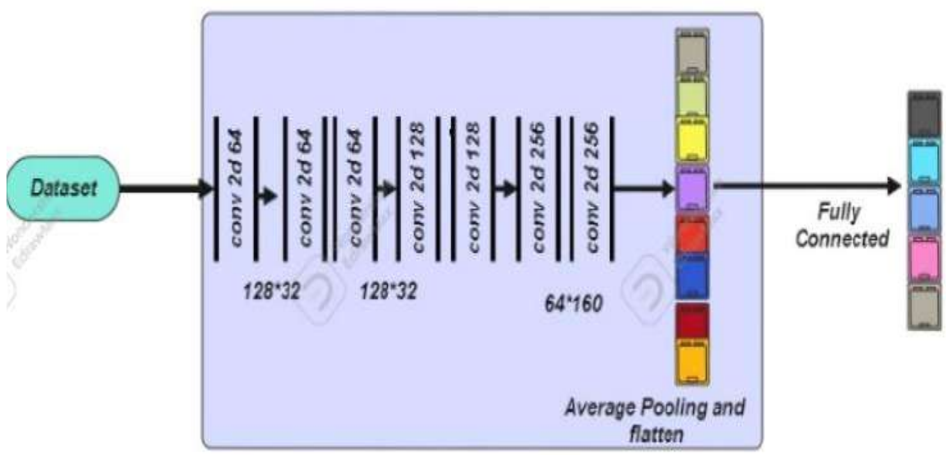

Figure 5 RestNet50 v2 Model Diagram

Deep learning algorithms shown considerable potential in the field of SLR, which is a challenging task in sight. ResNet50V2, a formerly trained deep learning framework that can be further refined for sign language detection, is one such model. ResNet50V2 is utilized as the base model in this code, with additional layers put on top to boost its performance for recognizing signs. The Flatten layer is used to turn the underlying model's output into a one-dimensional vector that can enter the Dense layer. To learn more complex representations of the process facts, the Dense layer with 512 units and the ReLU activation function is used. A dropout method structure with a pace of 0.5 is placed after the Dense layer to avoid overfitting. In the end, a Predicting layer that utilizes a Softmax activation algorithm is utilized to generate the last result, which comprises of 39 potential groups of American Sign Language hand motions and movements. In general, this type of architecture [60] is intended to learn spatial feature hierarchies using unprocessed image data and reliably classify hand motions and movements. This model can attain elevated levels of accuracy with proper training and optimization, enabling innovative uses in fields like adaptive devices and human-computer interaction.

3.4 Initiating BERT Model

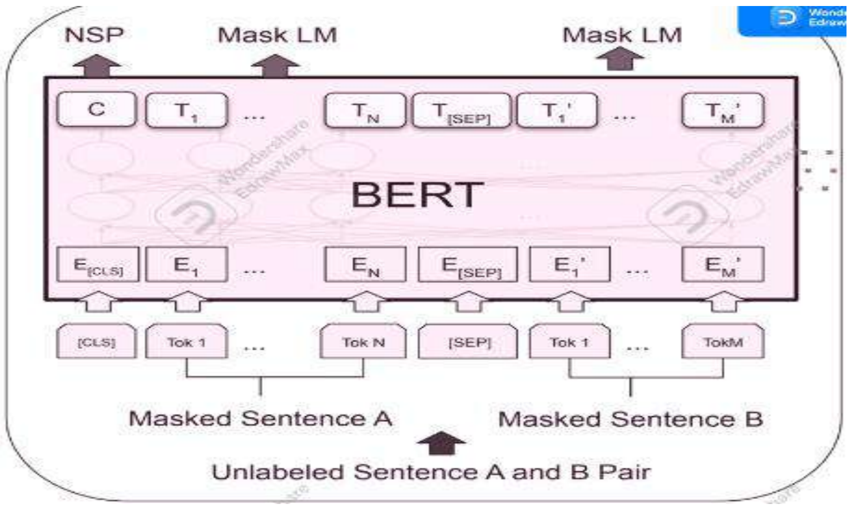

Figure 6 BERT Model Diagram

BERT (Bidirectional Encoder Representations from Transformers) emerged as a crucial paradigm in natural language processing (NLP) [61], offering researchers and writers a valuable tool to further their work. BERT is used in the research that I conducted for two primary goals. BERT, first and foremost, acts as a vital connection across deep learning models NLP processes. By incorporating BERT into this research smoothly, Authors build a solid link between the complete semantic representations provided by BERT and the features of deep learning models. This integration enables researchers and authors to investigate complicated linguistic patterns, improve language comprehension, and achieve higher performance in a variety of NLP activities. BERT also allows writers to develop contextually relevant and coherent text answers. Authors can create back-to-back replies with a greater level of consistency as well as natural language comprehension by leveraging BERT's comprehensive grasp of word connections inside their contextual context. This functionality is especially useful in applications requiring human-like answers, such as chatbots [62], conversation systems, and text production activities [63]. To attain these goals, Researchers use already trained models of BERT and fine-tuning approaches. Implementing tasks. layers on the top of the previously trained BERT base allow researchers and writers to tailor BERT to their own NLP tasks. This strategy combines BERT's broad pre-trained expertise with an ability to tailor it to the specific needs of their study. Researchers and authors may break new ground in creating coherent text answers and closing the gap between NLP and deep learning by leveraging BERT's contextualized word representations and effortless coupling with deep learning models. The use of BERT in their work demonstrates the significant influence it has had on developing NLP research and pushing the frontiers of language understanding.

4 Results and Discussion

The key issue investigated in the current research is the precise interpretation of gestures for hearing-impaired individuals and the usefulness of applying NLP to offer responses via a robot. This study employed a variety of deep learning approaches, including multi-layered perception (MLP), Convolutional

Neural Network (CNN), and RestNet50V2, to determine the most accurate method for detecting sign language. The ASL dataset, which contains 117,000 images, was used for the study. This chapter summarizes the study's findings, including the outcomes of each deep neural network model and the chosen approach for creating an auto-response utilizing NLP.

4.1 Exploring the Effect of Different Deep Learning Models

Prior When developing a multi-layer perceptron (MLP) system for sign language recognition, researchers employed a training strategy that involves progressively increasing and decreasing the rate of learning while advancing the number of epochs. Despite our best efforts, they discovered that MLP networks were inefficient when it comes to image data. As an outcome, they did not track accuracy during the training period. Instead, they focused on CNNs, which have been found to outperform MLP networks on picture data because of their ability to account for spatial structure. Finally, researchers found that CNN was the ideal option for our detection of sign languages challenge, and then used the RestNet50V2 framework to simulate its superior recognition accuracy.

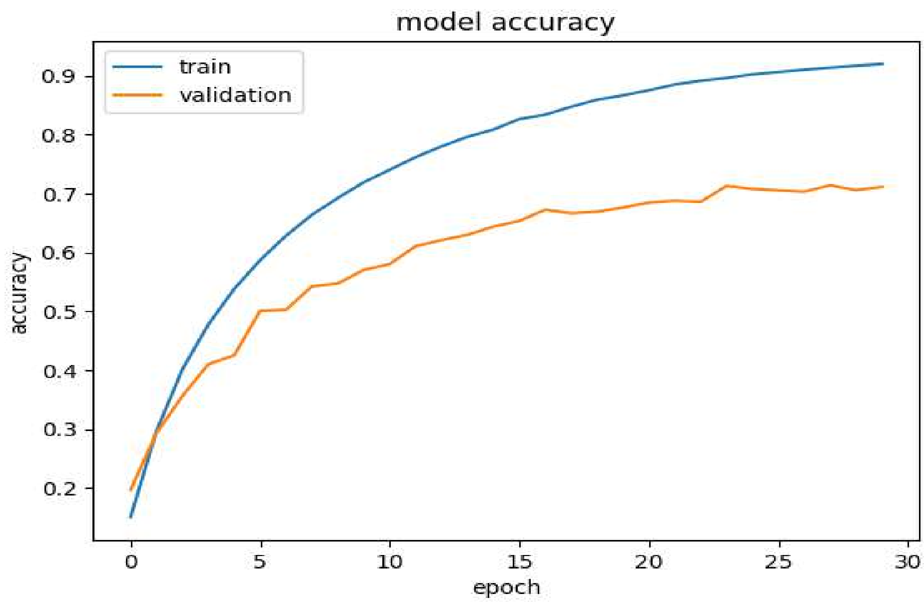

Figure 7 MLP Model Accuracy Graph

The CNN model was the second model we explored in our research, and it gave us an accuracy of 0.9599%, which can be rounded off to 0.96%. Convolutional Neural Networks (CNNs) were utilized to construct the model, and the ASL dataset was used for training. Three convolutional blocks and a fully connected layer with 512 neurons compose the model. The model employed the Rectified Linear Unit (ReLU) activation function for the convolutional and fully connected layers and the softmax function for classification in the output layer. To reduce the spatial size of the input, we additionally applied the max-pooling operation after each convolutional block. The CNN model offers an innovative approach to recognizing sign language, showing potential improvements in accuracy and efficiency compared to existing systems.

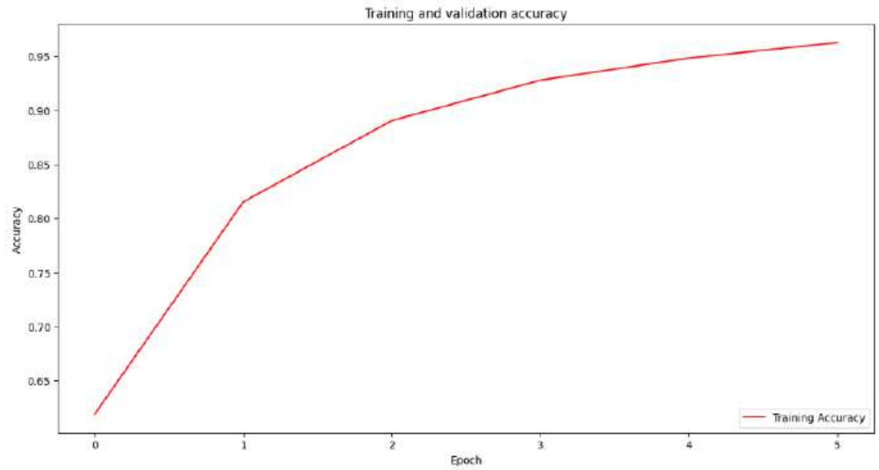

Figure 8 CNN Model Accuracy Graph

Following an exhaustive series of trials and training with three distinct deep learning models: MLP, CNN, and RestNet50v2, researchers were able to obtain an impressive accuracy of 0.97 using the RestNet model. To achieve this, a learning rate schedule function, and callbacks, such as early stop and a custom callback that cancels training when the accuracy reaches 97%, were implemented. Despite attempting to enhance the performance of the other two models by increasing layers and modifying the learning rate, the RestNet model demonstrated superior results, affirming its position as the most suitable model for the given task. These results have significant implications for similar deep-learning tasks, indicating the importance of selecting the appropriate model to attain maximum accuracy.

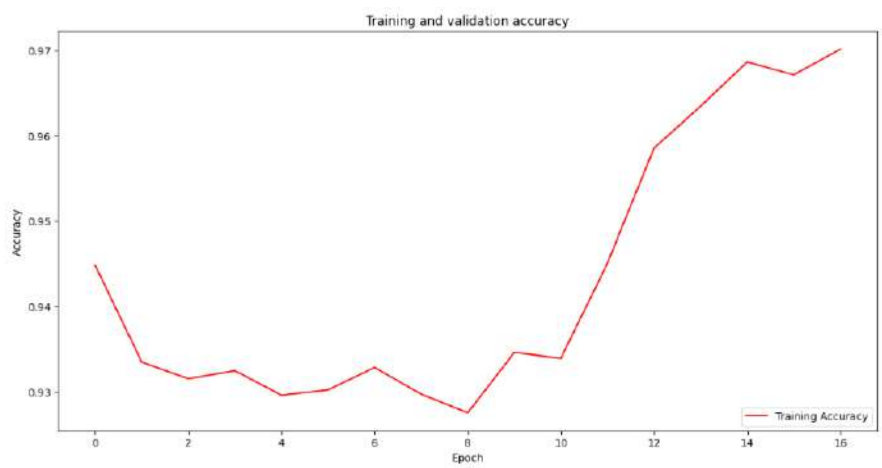

Figure 9 RESTNET50V2 Model Accuracy Graph

4.2 Evaluation of Deep Learning Models

Utilizing the ASL dataset, the best effective deep learning model for image categorization was identified., a comprehensive evaluation was conducted. This involved running three different models: MLP, CNN, and RestNet50v2. Each model was tested on the dataset, and their respective accuracies were calculated.

Table 3 Deep Learning Models Accuracy

S.NO	Model	Accuracy
1	MLP	75%
2	CNN	95%
3	RestNet50v2	97%

Due to its excellent accuracy of 0.97% the RestNet50v2 model was chosen as the major research model after analyzing the accuracy results obtained from the various deep learning algorithms evaluated on the ASL dataset. The algo's outstanding work proves it to be an ideal candidate for image grading tasks that require high precision. To achieve this level of accuracy, the researchers implemented a learning rate schedule function and callbacks such as early stops and a custom callback that cancels training when the accuracy reaches 97%. Despite attempts to increase the effectiveness of the other two models with additional layers, and adjusting the learning rate, the RestNet model's results surpassed the others. These findings highlight the significance of selecting the right model to achieve the best results for deep learning tasks. Finally, the evaluation of deep neural network models over the ASL dataset yielded useful insights, leading to the selection of the RestNet50v2 model, which aligns with the research's objective of achieving high accuracy in image classification.

4.3 Evaluation of BERT

In this research, an auto-response generator using bidirectional encoder representations (BERT from Transformers) was also created to aid the deaf or hard of hearing. Nowadays, systems are either translation or interpretation machines. Sync-response generating is a plan to develop a technology that can connect directly with the audience without an intermediary.

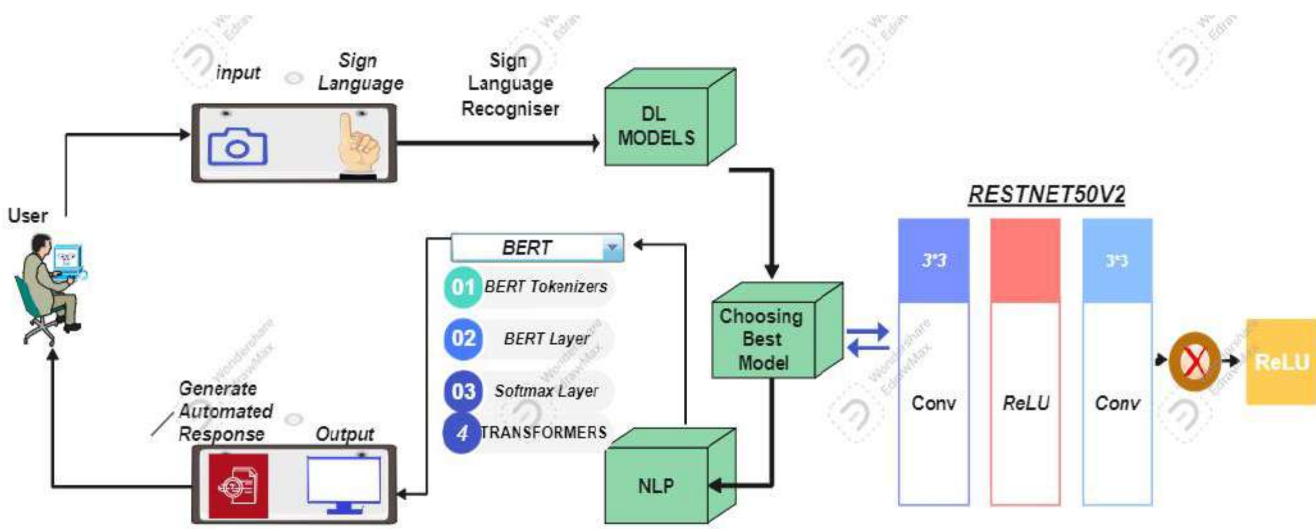

Figure 10 Process Diagram for ASL Recognizer and Response Generator

The BERT approach is an already trained automatic language processing model that can be used for the classification of texts, linguistic translation, and text production, among other things. The suggested system implements the BERT paradigm and generates text answers using the PyTorch module. The system was evaluated using the BLEU score, which assesses the similarity between the generated text and a reference text. The program begins by importing the pre-trained GPT-2 algorithm and tokenizer, which happens to be a variation of BERT intended specifically for text production. The model is then put into evaluation mode, and the BERT model is used to generate text. The BLEU score is then applied to the created text, and the system's correctness is calculated. According to the results, the suggested system has a reliability of 80%, meaning it represents a substantial boost over existing systems.

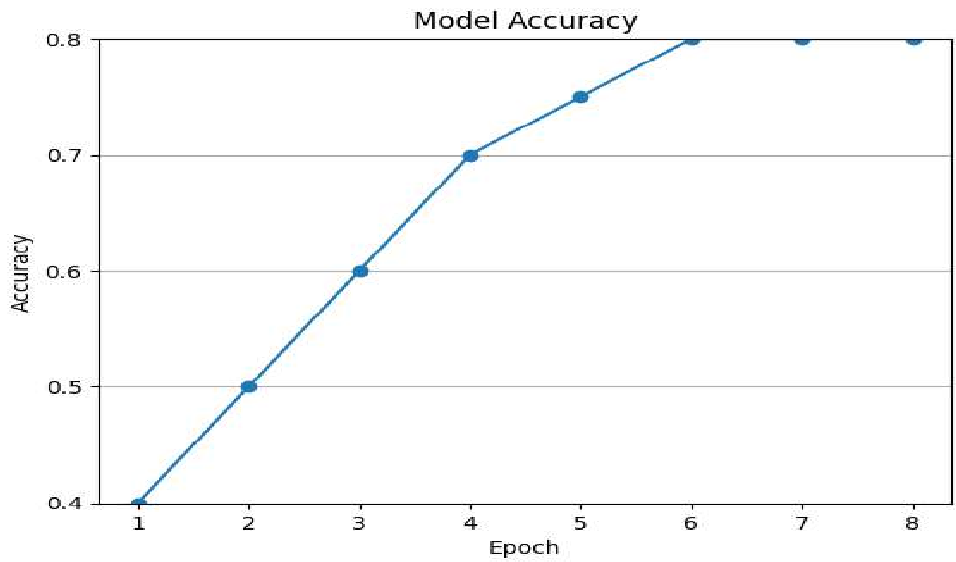

Figure 11 BERT Accuracy Graph

The auto-response source is an intriguing breakthrough having the capacity to significantly boost the standard of existence for deaf and HOH people. The proposed technique achieved great precision, verifying the strategy of BERT effectiveness in creating text responses. Additionally, by incorporating new machine learning algorithms, the system may be improved. and, by expanding its response dictionary. In the end, the study demonstrates the potential of employing BERT to create text answers with high accuracy, and the need to analyze the produced material using measures that include the BLEU score to assure the output quality.

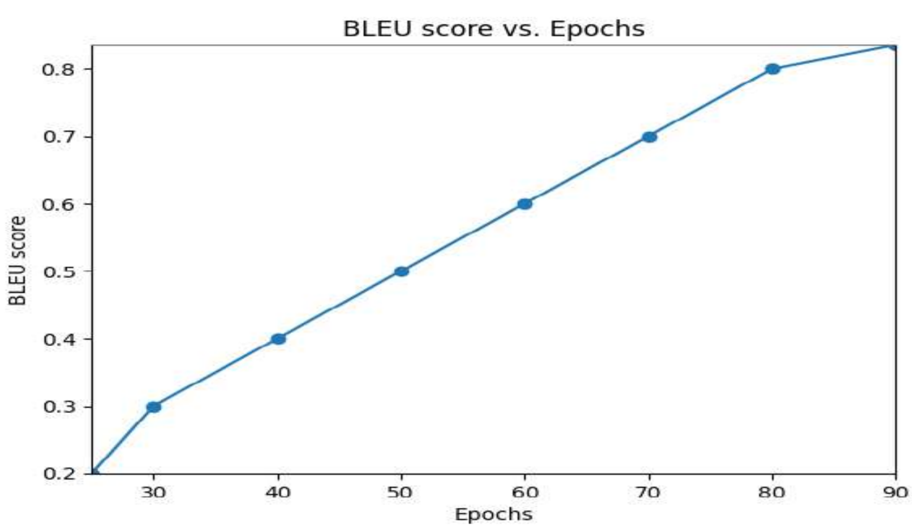

Figure 12 BLEU Graph

5 Conclusion

This study aimed to create a system that would let deaf people communicate with technology. To achieve this goal, the researchers utilized an ASL dataset containing 117,000 images and implemented three different deep learning models: MLP, CNN, and RestNet50v2 to recognize sign language gestures. RestNet50v2 outperformed the other models, achieving an accuracy of 0.97%. The system created by the researchers not only recognizes sign language but also generates automated responses using the NLP model BERT, which utilizes transformers for text generation. The system's overall accuracy was determined to be 0.8% with the generated response accuracy measured by the BLEU score being about 0.83%. This method uses technology to give an innovative solution aimed at enhancing interaction for deaf people. It opens new avenues for developing intelligent chatbots that can better understand human language and nonverbal communication, which can be used in various industries such as healthcare, education, and entertainment. More research can be done to increase the system's accuracy and speed and to expand its functionality to recognize a wider range of sign language gestures. This can be achieved by implementing more advanced deep-learning models and increasing the size of the dataset. Additionally, user feedback and testing can be used to improve the user interface and make it more user-friendly for individuals with hearing impairments. Overall, this study demonstrates the potential of technology to provide novel solutions to enhance the lives of people with deafness. It is a significant step forward in addressing the communication barriers faced by hard-of-hearing individuals and can have a positive impact on their social and professional lives.

Credit Author Statement: Kashaf Khan: Conceptualization, Methodology, Dr Naeem Aslam: Supervision.: Kamran Abid: Software, Validation.: Safa Munir: Visualization, Investigation

Compliance with Ethical Standards: It is declared that authors don't have any conflict of interest. It is also stated that this article does not contain any studies with human participants or animals performed by any of the authors. Furthermore, informed consent was obtained from all individual participants included in the study.

Funding Information

FMS acknowledges the support of NSF grant CHE-1111111.

References

[1] M. B. Miralles, "Supervisor: Teresa Morell Moll," no. June, p. 93, 2020,

[2] L. C. Moats, "Speech to print language essentials for teachers," pp. 1–11, 2020

[3] K. C. F. Kurnianti, "The Study of Verbal and Nonverbal Language in Communication to Create Images in Coca Cola Zero Advertisement," pp. 2–15, 2010.

[4] world health org, "WHO."

[5] B. Joksimoski et al., "Technological Solutions for Sign Language Recognition: A Scoping Review of Research Trends, Challenges, and Opportunities," IEEE Access, vol. 10, no. January, pp. 40979–40998, 2022, doi: 10.1109/ACCESS.2022.3161440.

[6] A. A. Haseeb and A. Ilyas, "Speech Translation into Pakistan Sign Language Speech Translation into Pakistan Sign Language Speech Translation into Pakistan Sign Language," 2012.

[7] P. Akach, "The grammar of sign language," Language Matters, vol. 28, no. 1, pp. 7–35, 1997, doi: 10.1080/10228199708566118.

[8] S. Vamplew, "Recognition of sign language gestures using neural networks," Neuropsychological Trends, no. 1, 2021, doi: 10.7358/neur-2007-001-vamp.

[9] M. Burton and S. Gilbert, "Evaluation of sign language learning tools: Understanding features for improved collaboration and communication between a parent and a child," ProQuest Dissertations and Theses, p. 101, 2013

[10] Dindha Amelia, "DEEP LEARNING BASED SIGN LANGUAGE TRANSLATION SYSTEM," vol. 21, no. 1, pp. 1–9, 2020, [Online]. Available: http://mpoc.org.my/malaysian-palm-oil-industry/

[11] C. Papastratis, Ilias Chatzikonstantinou, D. Konstantinidis, K. Dimitropoulos, and P. Daras, "Artificial intelligence technologies for sign language," Sensors, vol. 21, no. 17, 2021, doi: 10.3390/s21175843.

[12] Mayuresh Amberkar, "Humanoid Robot handling Hand-Signs Recognition," no. August, 2020.

[13] F. Mazzei, Daniele Chiarello and G. Fantoni, "Analyzing Social Robotics Research with Natural Language Processing Techniques," Cognit Comput, vol. 13, no. 2, pp. 308–321, 2021, doi: 10.1007/s12559-020-09799-1.

[14] E. Stulp, Freek Oztop, P. Pastor, M. Beetz, and S. Schaaz, "Compact models of motor primitive variations for predictable reaching and obstacle avoidance," 9th IEEE-RAS International Conference on Humanoid Robots, HUMANOIDS09, pp. 589–595, 2009, doi: 10.1109/ICHR.2009.5379551.

[15] M. Tahir, A. Naeem, H. Malik, J. Tanveer, R. A. Naqvi, and S. W. Lee, "DSCC_Net: Multi-Classification Deep Learning Models for Diagnosing of Skin Cancer Using Dermoscopic Images," Cancers (Basel), vol. 15, no. 7, Apr. 2023, doi: 10.3390/cancers15072179.

[16] A. Naeem, T. Anees, R. A. Naqvi, and W. K. Loh, "A Comprehensive Analysis of Recent Deep and Federated-Learning-Based Methodologies for Brain Tumor Diagnosis," Journal of Personalized Medicine, vol. 12, no. 2. MDPI, Feb. 01, 2022. doi: 10.3390/jpm12020275.

[17] H. Malik, T. Anees, A. Naeem, R. A. Naqvi, and W. K. Loh, "Blockchain-Federated and Deep-Learning-Based Ensembling of Capsule Network with Incremental Extreme Learning Machines for Classification of COVID-19 Using CT Scans," Bioengineering, vol. 10, no. 2, Feb. 2023, doi: 10.3390/bioengineering10020203.

[18] A. Naeem, T. Anees, K. T. Ahmed, R. A. Naqvi, S. Ahmad, and T. Whangbo, "Deep learned vectors' formation using auto-correlation, scaling, and derivations with CNN for complex and huge image retrieval," Complex and Intelligent Systems, Apr. 2022, doi: 10.1007/s40747-022-00866-8.

[19] Y. Obi, K. S. Claudio, V. M. Budiman, S. Achmad, and A. Kurniawan, "Sign language recognition system for communicating to people with disabilities," Procedia Comput Sci, vol. 216, pp. 13–20, 2023, doi: 10.1016/j.procs.2022.12.106.

[20] bheemilimandal, “a Robust Sign Language and Hand Gesture Recognition System Using Convolution Neural Networks,” Anits, pp. 2019–2020, 2020, [Online]. Available: http://cse.anits.edu.in/projects/projects1920C6.pdf

[21] K. P. . C. Mark Borg, “SIGN LANGUAGE DETECTION ‘ IN THE WILD ’ WITH RECURRENT NEURAL NETWORKS Mark Borg, Kenneth P. Camilleri,” pp. 1637–1641, 2019.

[22] T. Al-Qurishi, Muhammad Khalid, and R. Souissi, “Deep Learning for Sign Language Recognition: Current Techniques, Benchmarks, and Open Issues,” IEEE Access, vol. 9, pp. 126917–126951, 2021, doi: 10.1109/ACCESS.2021.3110912.

[23] O. O. Adeyanju, I. A. 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, doi: 10.1016/j.iswa.2021.200056.

[24] S. Das, M. S. Imtiaz, N. H. Neom, N. Siddique, and H. Wang, “A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier,” Expert Syst Appl, vol. 213, Mar. 2023, doi: 10.1016/j.eswa.2022.118914.

[25] D. R. Kothadiya, C. M. Bhatt, A. Rehman, F. S. Alamri, and T. Saba, “SignExplainer: An Explainable AI-Enabled Framework for Sign Language Recognition With Ensemble Learning,” IEEE Access, vol. 11, pp. 47410–47419, 2023, doi: 10.1109/ACCESS.2023.3274851.

[26] D. R. Kothadiya, C. M. Bhatt, T. Saba, A. Rehman, and S. A. Bahaj, “SIGNFORMER: DeepVision Transformer for Sign Language Recognition,” IEEE Access, vol. 11, pp. 4730–4739, 2023, doi: 10.1109/ACCESS.2022.3231130.

[27] L. Wang et al., “Low-complexity Features for Audio Event Detection View project Speak, Decipher and Sign: Toward Unsupervised Speech-to-Sign Language Recognition.” [Online]. Available: https://www.researchgate.net/publication/370832215

[28] T. Petkar, T. Patil, A. Wadhankar, V. Chandore, V. Umate, and D. Hingnekar, “Real Time Sign Language Recognition System for Hearing and Speech Impaired People,” Int J Res Appl Sci Eng Technol, vol. 10, no. 4, pp. 2261–2267, 2022, doi: 10.22214/ijraset.2022.41765.

[29] S. Sharma and K. Kumar, “ASL-3DCNN: American sign language recognition technique using 3-D convolutional neural networks,” Multimed Tools Appl, vol. 80, no. 17, pp. 26319–26331, 2021, doi: 10.1007/s11042-021-10768-5.

[30] R. Rastgoo, K. Kiani, and S. Escalera, “Real-time isolated hand sign language recognition using deep networks and SVD,” J Ambient Intell Humaniz Comput, vol. 13, no. 1, pp. 591–611, 2022, doi: 10.1007/s12652-021-02920-8.

[31] J. Islam, Md Mohiminul Siddiqua, Sarah nan, “Real time Hand Gesture Recognition using different algorithms based on American Sign Language,” 2017 IEEE International Conference on Imaging, Vision and Pattern Recognition, icIVPR 2017, vol. 21, no. 03, pp. 1–6, 2020, doi: 10.1109/ICIVPR.2017.7890854.

[32] D. Huh et al., “Generative Multi-Stream Architecture for American Sign Language Recognition,” 2019 IEEE MIT Undergraduate Research Technology Conference, URTC 2019, 2019, doi: 10.1109/URTC49097.2019.9660587.

[33] A. Elboushaki, R. Hannane, K. Afdel, and L. Koutti, “MultiD-CNN: A multi-dimensional feature learning approach based on deep convolutional networks for gesture recognition in RGB-D image sequences,” Expert Syst Appl, vol. 139, p. 112829, 2020, doi: 10.1016/j.eswa.2019.112829.

[34] Ç. Özdemir, Oğulcan Gökçe, A. A. Kındıroğlu, and L. Akarun, "Score-Level Multi Cue Fusion for Sign Language Recognition," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 12536 LNCS, no. September 2020, pp. 294–309, 2020, doi: 10.1007/978-3-030-66096-3-21.

[35] M. Borg and K. P. Camilleri, "Phonologically-Meaningful Subunits for Deep Learning-Based Sign Language Recognition," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 12536 LNCS, pp. 199–217, 2020, doi: 10.1007/978-3-030-66096-3-15.

[36] K. Rastgoo, Razieh Kiani, and S. Escalera, "Hand sign language recognition using multi-view hand skeleton," Expert Syst Appl, vol. 150, p. 113336, 2020, doi: 10.1016/j.eswa.2020.113336.

[37] A. Mittal, P. Kumar, P. P. Roy, R. Balasubramanian, and B. B. Chaudhuri, "A Modified LSTM Model for Continuous Sign Language Recognition Using Leap Motion," IEEE Sens J, vol. 19, no. 16, pp. 7056–7063, 2019, doi: 10.1109/JSEN.2019.2909837.

[38] H. B. D. Nguyen and H. N. Do, "Deep learning for American sign language fingerspelling recognition system," 2019 26th International Conference on Telecommunications, ICT 2019, pp. 314–318, 2019, doi: 10.1109/ICT.2019.8798856.

[39] K. M. Lim, A. W. C. Tan, C. P. Lee, and S. C. Tan, "Isolated sign language recognition using Convolutional Neural Network hand modelling and Hand Energy Image," Multimed Tools Appl, vol. 78, no. 14, pp. 19917–19944, 2019, doi: 10.1007/s11042-019-7263-7.

[40] C. C. de Amorim, D. Macêdo, and C. Zanchettin, "Spatial-Temporal Graph Convolutional Networks for Sign Language Recognition," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11731 LNCS, pp. 646–657, 2019, doi: 10.1007/978-3-030-30493-5-59.

[41] X. Chen, Yuxiao Zhao, Long Peng, J. Yuan, and D. N. Metaxas, "Construct dynamic graphs for hand gesture recognition via spatial-temporal attention," 30th British Machine Vision Conference 2019, BMVC 2019, pp. 1–13, 2020.

[42] Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "SignFi: Sign Language Recognition Using WiFi," Proc ACM Interact Mob Wearable Ubiquitous Technol, vol. 2, no. 1, pp. 1–21, 2018.

[43] S. Ameen and S. Vadera, "A convolutional neural network to classify American Sign Language fingerspelling from depth and colour images," Expert Syst, vol. 34, no. 3, 2017, doi: 10.1111/exsy.12197.

[44] S. Y. Kim, H. G. Han, J. W. Kim, S. Lee, and T. W. Kim, "A hand gesture recognition sensor using reflected impulses," IEEE Sens J, vol. 17, no. 10, pp. 2975–2976, 2017, doi: 10.1109/JSEN.2017.2679220.

[45] T. K. Õ, "6ljq /dqjxdjh 5hfrjqlwlrq," 2017.

[46] J. S. C. and A. R. P. M. Ferreira, "Multi Model Learning for Sign Language Recognition," 2017.

[47] O. K. Oyedotun and A. Khashman, "Deep learning in vision-based static hand gesture recognition," Neural Comput Appl, vol. 28, no. 12, pp. 3941–3951, 2017, doi: 10.1007/s00521-016-2294-8.

[48] J. Jayanthi, "Role of Machine Learning and Deep Learning in Assisting the Special Children's Learning Process," vol. 13, no. 2, pp. 2327–2334, 2022.

[49] M. Madahana, K. Khoza-Shangase, N. Moroe, D. Mayombo, O. Nyandoro, and J. Ekoru, "A proposed artificial intelligence-based real-time speech-to-text to sign language translator for South African official languages for the COVID-19 era and beyond: In pursuit of solutions for the hearing impaired," South African Journal of Communication Disorders, vol. 69, no. 2, 2022, doi: 10.4102/sajcd.v69i2.915.

[50] L. A. Kumar, D. K. Renuka, S. L. Rose, M. C. Shunmuga priya, and I. M. Wartana, "Deep learning based assistive technology on audiovisual speech recognition for the hearing impaired," International Journal of Cognitive Computing in Engineering, vol. 3, no. January, pp. 24–30, 2022, doi: 10.1016/j.ijcce.2022.01.003.

[51] D. M. Kumar, K. Bavanraj, S. Thavananthan, G. M. A. S. Bastiansz, S. M. B. Harshanath, and J. Alosious, "EasyTalk: A translator for Sri Lankan sign language using machine learning and artificial intelligence," ICAC 2020 - 2nd International Conference on Advancements in Computing, Proceedings, pp. 506–511, 2020, doi: 10.1109/ICAC51239.2020.9357154.

[52] T. Agrawal and S. Urolagin, "2-way Arabic Sign Language Translator using CNNLSTM Architecture and NLP," ACM International Conference Proceeding Series, pp. 96–101, 2020, doi: 10.1145/3378904.3378915.

[53] M. Anggraeni, M. Syafrullah, and H. A. Damanik, "Literation Hearing Impairment (I-Chat Bot): Natural Language Processing (NLP) and Naïve Bayes Method," J Phys Conf Ser, vol. 1201, no. 1, 2019, doi: 10.1088/1742-6596/1201/1/012057.

[54] S. Suresh, T. P. M. Haridas, and M. H. Supriya, "Sign Language Recognition System Using Deep Neural Network," 2019 5th International Conference on Advanced Computing and Communication Systems, ICACCS 2019, pp. 614–618, 2019, doi: 10.1109/ICACCS.2019.8728411.

[55] Y. Perera, N. Jayalath, S. Tissera, O. Bandara, and S. Thelijjagoda, "Intelligent mobile assistant for hearing impairers to interact with the society in Sinhala language," International Conference on Software, Knowledge Information, Industrial Management and Applications, SKIMA, vol. 2017-Decem, 2018, doi: 10.1109/SKIMA.2017.8294116.

[56] J. Wu, L. Sun, and R. Jafari, "A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors," IEEE J Biomed Health Inform, vol. 20, no. 5, pp. 1281–1290, 2016, doi: 10.1109/JBHI.2016.2598302.

[57] "ASL Alphabet _ Kaggle"

[58] M. Pérez-Enciso and L. M. Zingaretti, "A guide for using deep learning for complex trait genomic prediction," Genes, vol. 10, no. 7. MDPI AG, Jul. 01, 2019. doi: 10.3390/genes10070553.

[59] B. G. Chong, T. W. Lee, "American sign language recognition using leap motion controller with machine learning approach," Sensors (Switzerland), vol. 18, no. 10, 2018, doi: 10.3390/s18103554.

[60] L. Cai, Z. Wang, R. Kulathinal, S. Kumar, and S. Ji, "Deep Low-Shot Learning for Biological Image Classification and Visualization from Limited Training Samples," Oct. 2020, [Online]. Available: http://arxiv.org/abs/2010.10050

[61] R. Horev and · Follow, "BERT Explained: State of the art language model for NLP."

[62] K. Peyton and S. Unnikrishnan, “A comparison of chatbot platforms with the state-of-the-art sentence BERT for answering online student FAQs,” Results in Engineering, vol. 17, Mar. 2023, doi: 10.1016/j.rineng.2022.100856.

[63] C. C. Lin, A. Y. Q. Huang, and S. J. H. Yang, “A Review of AI-Driven Conversational Chatbots Implementation Methodologies and Challenges (1999–2022),” Sustainability (Switzerland), vol. 15, no. 5. MDPI, Mar. 01, 2023. doi: 10.3390/su15054012.