AMERICAN Journal of Science on Integration and Human Development
Volume 3, Issue 8, 2025 ISSN (E): 2993-2750

Sign Language to Text and Voice Conversion for Speaking Impaired People Using Convolutional Neural Network

K. K. Sreedeve[*1], K. Jayashree[2], Rincya Mabujan[3], B. Sneha[4], N. Selvam[5], M. Mohamed Sameer Ali[6]

[1,2,3,4,5,6]Department of Computer Science and Engineering, Dhaanish Ahmed College of Engineering, Padappai, Chennai, Tamil Nadu, India
*Correspondence: sreedeve@dhaanishcollege.in

Abstract: Sign language, a visual-gestural form of communication predominantly utilised by individuals with hearing impairments, has distinct problems in routine encounters with non-signers. To solve this problem with communication, we suggest a complete Sign Language Detection and Translation System (SLDTS) that uses Convolutional Neural Networks (CNN) to recognise sign language and turn it into audio. It also uses keyword recognition to make communication more efficient. The SLDTS is built on deep learning methods, notably CNN-based models, that can accurately find and recognise sign language motions. These models learn quickly how to recognise different hand shapes, movements, and gestures by being trained on massive, annotated datasets of sign language gestures. By analysing input video streams or images in real time, the CNN-based sign language detection component of the SLDTS can efficiently recognise and understand sign language motions, creating the foundation for effective communication between sign language users and non-signers. When the SLDTS sees sign language gestures, it uses text-to-speech (TTS) synthesis to turn them into spoken words. This translation procedure takes the recognised signs and turns them into text, then combines them into audio output. The SLDTS uses powerful TTS technology to make sure that the translated audio is clear and sounds natural, so that those who don't know sign language may understand what the signer is trying to say. This feature lets people with and without hearing problems talk to each other in real time, which helps them understand each other better and interact more.

Keywords: Text-To-Speech (TTS), Convolutional Neural Networks (CNN), Sign Language Detection and Translation System (SLDTS)

Introduction

Sign language is a vital means of communication for individuals with hearing impairments, enabling them to express thoughts, emotions, and ideas through hand gestures and movements. However, the communication barrier between sign language users and non-signers poses significant challenges in facilitating effective interaction and understanding. In light of this, we propose a novel Sign Language Detection and Translation System (SLDTS) that harnesses the power of Convolutional Neural Networks (CNNs) to recognise sign language gestures and translate them into spoken language [31]. Complemented by keyword recognition functionality, the SLDTS aims to enhance communication efficiency and promote inclusivity in various contexts. By seamlessly bridging the gap between sign language users and non-signers, the SLDTS represents a significant advancement in fostering inclusive communication and accessibility for individuals with hearing impairments [32]. This paper presents the design, implementation, and potential impact of the SLDTS, highlighting its role in promoting mutual understanding and interaction between individuals of diverse linguistic backgrounds. Through the integration of cutting-edge technologies and innovative approaches, the SLDTS aims to revolutionise communication accessibility and contribute to the creation of a more inclusive and connected society [33].

Develop a robust Sign Language Detection and Translation System (SLDTS) that utilises Convolutional Neural Networks (CNN) to recognise sign language gestures accurately. Implement real-time sign language recognition functionality capable of processing input video streams or images to detect and interpret hand shapes, movements, and gestures [34]. Integrate text-to-speech (TTS) synthesis capabilities into the SLDTS to translate recognised sign language gestures into spoken language audio output [35]. Enhance communication efficiency by incorporating keyword recognition functionality to identify and highlight specific sign language gestures representing keywords or phrases. Promote inclusivity and accessibility by enabling seamless communication between sign language users and non-signers in various settings, including educational institutions, workplaces, and public spaces [36]. Evaluate the performance and usability of the SLDTS through rigorous testing and user feedback, ensuring its effectiveness in facilitating mutual understanding and interaction among individuals with and without hearing impairments. Contribute to the advancement of assistive technology solutions aimed at breaking down communication barriers and fostering a more inclusive society for individuals with diverse linguistic needs [37].

Image processing is a way to improve images or extract valuable information from them by applying various operations on them [38]. This form of signal processing takes an image as input and produces either another image or some associated features and characteristics as output. One area of technology that is booming right now is image processing [39]. In the fields of computer science and engineering, it is also an important subject of study. There are three stages to image processing: Analogue and digital image processing are the two main approaches to this task. Printouts, images, and other physical copies can benefit from analogue image processing. Image analysts employ a range of basic interpretations when working with these visual techniques [40]. Techniques for digital image processing allow for the manipulation of digital images through the use of computers. When working with digital methods, every data must go through the same three stages: pretreatment, augmentation, presentation, and information extraction [41].

The use of digital computers for image manipulation is known as digital image processing. In recent decades, its usage has skyrocketed. Its uses span the gamut from geological processing and remote sensing to entertainment and medical. Digital image processing is foundational to multimedia systems, which are essential to today's information society [42]. The processing of digital images involves working with these values with a finite degree of accuracy [43]. Enhancement, restoration, analysis, and compression are the four main categories into which digital image processing falls. To improve an image's readability for human eyes, image enhancement makes use of several manipulations, most commonly heuristics, [44]. The use of computers to alter visual data is known as digital image processing. Applying a set of procedures to a numerical representation of an item in order to produce an intended outcome is the definition of digital image processing [45]. Processing digital images entails digitising them from their physical counterparts and then, using a variety of techniques, extracting useful information from them [46].

Pattern recognition: To separate items from photos, you need to use pattern recognition technology. Then, you can use statistical decision theory to identify and sort these things. When an image involves more than one object, pattern recognition happens in three steps, as shown in Figure 1.

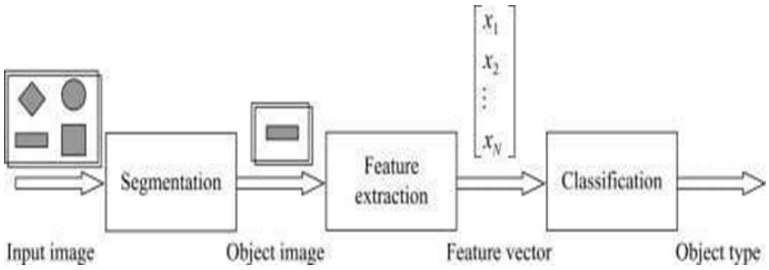

Figure 1. Phases of pattern recognition.

Image segmentation and object separation are part of the initial phase. Objects are identified and isolated from their surroundings during this stage [47]. Extracting features is the second step. Measurements are taken at this stage. During feature extraction, a collection of features is merged to form a feature vector, and the measuring feature provides quantitative estimations of several significant object attributes [48]. Classification is the third step. At this stage, all that is produced is a choice about the category to which each object is assigned [49]. Consequently, pictures are the inputs for pattern recognition, and the outputs are the kinds of objects and the structural analysis of those images [50]. In order to properly comprehend and evaluate the crucial details of a picture, structural analysis provides a description of the image [51].

Hand gestures, together with facial emotions and body postures, are an integral feature of this language [53]. The majority of its users are those who have hearing impairments. American, British, and Indian sign languages are only a few examples of the numerous sign languages in use today. For both BSL and ASL users, the other's language can be difficult to understand [54]. The inattentive could be able to interact with non-signing people without an interpreter if a sign recognition system were to work. It has the potential to produce text or voice, empowering the deaf and hard of hearing to live more independently. So yet, no system has been able to accomplish these feats [55]. During the course of this project, we intend to create a system that reliably determines signing. Though its grammar differs from English, American Sign Language (ASL) is a full-fledged natural language sharing many linguistic features with spoken languages. Sign language is communicated by facial and hand gestures [56]. Some hearing people use it, and many deaf and hard of hearing North Americans rely on it as their main language [52].

The process of translating the user's hand signals and motions into text is known as sign language recognition. For the speechless, it is a means of connecting with the outside world. In order to associate the gesture with the correct text in the training data, image processing techniques and neural networks do the mapping [57]. The result is text that may be read and comprehended from the original photos or videos [58]. In most societies, persons who are considered dumb are unable to participate in everyday conversations. They have a hard time communicating with regular people using their gestures because most people only recognise a small fraction of them. People who are deaf or hard of hearing rely on nonverbal cues such as body language and facial expressions for the majority of their communication needs [59]. For those who are deaf or hard of hearing, sign language is their main form of communication. Its grammar and vocabulary are similar to those of any language, but its mode of communication is visual [60]. The issue emerges when deaf or stupid people attempt to communicate with others using the grammar of sign language. Reason being, the average person doesn't know this grammar [61]. The outcome is that those who are deaf or hard of hearing can only communicate with those who share their language [62].

More and more, public support and funding for global initiatives are going towards initiatives that highlight sign language. But scientists have been working on it for a while, and thus far, their findings are encouraging [63]. Speech recognition is an exciting area of research and development, but there is yet no marketable solution for sign recognition [64]. Create human-computer interfaces (HCIs) that are easy for people to use by teaching computers to understand human language. One step in that direction is teaching computers to recognise human speech, facial expressions, and gestures. Communication using gestures is a kind of non-verbal communication [65]. Numerous gestures can be executed simultaneously by a single individual.

Computer vision researchers are very interested in human gestures since they are seen visually.

Motivation

About 1.3 million persons in India had "hearing impairment" according to the 2011 census. The National Association of the Deaf in India, on the other hand, says that about 18 million people, or 1% of the Indian population, are deaf. These numbers were what drove us to do our endeavor [69]. There is a need for a system because people who are deaf or have speech problems need a way to talk to regular people. Not everyone who is normal can comprehend how to sign for persons who are disabled. So, the goal of our project is to turn sign language movements into text that anyone can read [70].

Problem Statement

People who can't talk use hand signs and gestures to talk to each other. Normal individuals have trouble understanding their language. So, we need a system that can recognise diverse signals and gestures and provide the information to regular people. It connects people who have physical disabilities with people who don't [71].

Neural Networks: A neural network is a set of algorithms that tries to find hidden connections in a data set by imitating how the brain functions [68]. In this way, neural networks are groups of neurons, whether they are real or made by people. Neural networks can adjust to new inputs and still produce the best possible result without having to change the output criteria [72]. Neural networks, which come from AI, are quickly becoming popular for building trading systems [67]. A neural network works in a way that is similar to how the brain's neural network works. A "neuron" in a neural network is a mathematical function that sorts and stores information in a certain way. The network is quite similar to statistical techniques like regression analysis and curve fitting [66].

ANN is being utilised in the following fields. It implies that ANN employs an interdisciplinary methodology in its development and applications. Speech Recognition: Speech plays a big role in how people talk to each other. So, it's only natural for humans to want to be able to talk to computers [74]. People still need complex languages that are hard to learn and use to talk to machines in this day and age. A simple way to get over this communication barrier would be to talk to the machine in a language it can understand [75]. There has been a lot of progress in this sector, but these systems still have trouble with limited vocabulary or grammar and having to retrain the system for various speakers in different situations [76].

ANN is a big part of this. After that, ANNs have been employed to recognise speech. The Kohonen Self-organising feature map is the best network for this because it takes small parts of the speech waveform as input [77]. This is termed the feature extraction technique, and it will map the same kinds of phonemes as the output array [78]. After taking out the features, it will recognise the utterance with the help of some acoustic models as back-end processing. ANN is a big part of this. The following ANNs have been used to recognise speech [73].

Character Recognition

It is an intriguing issue that pertains to the broader field of Pattern Recognition. Numerous neural networks have been created to identify handwritten characters, letters, or numbers. Here are a few ANNs that have been used to recognise characters [79]. Even though backpropagation neural networks have a lot of hidden layers, the way they link is limited to one layer at a time [80]. Like this, recognition also includes numerous hidden layers, and for these kinds of applications, it is trained one layer at a time [81].

Application for Signature Verification: One of the best ways to prove someone's identity and give them permission to do something is to use their signature [82]. The method for checking signatures does not rely on vision. For this application, the initial step is to extract the feature or the geometrical feature set that represents the signature [83]. We need to use an efficient neural network approach to train the neural networks with these feature sets. This trained neural network will check the signature to see if it is real or fake [84].

Face recognition: First, all of the incoming photographs need to be analysed. Then, you need to make that picture smaller. Finally, a neural network training procedure must be used to classify it [86]. We utilise the following neural networks using preprocessed photos for training. They are fully connected multilayer feed-forward neural networks that were trained using a backpropagation algorithm [85].

Deep Learning: Deep-learning networks are different from the more conventional single-hidden-layer neural networks because they are deeper and have more node layers that data has to go through in a multistep pattern recognition process [91]. The initial perceptrons and other early neural networks were shallow, with only one input layer, one output layer, and, at most, one hidden layer [92]. Deep learning is when there are more than three layers, including the input and output layers. Deep isn't just a jargon to make algorithms appear like they read Sartre and listen to bands you haven't heard of yet. It is a very specific word that denotes more than one concealed layer [90].

Convolutional neural networks (CNNs) are a type of artificial neural network that Yann LeCun came up with in 1988. Some parts of the visual cortex are used by CNN [93]. One of the most common uses for this architecture is to sort images. Facebook uses CNN to automatically classify photographs, Amazon uses it to suggest products, and Google uses it to search through users' photos [89]. The computer sees a grid of pixels instead of the picture. Suppose the picture is 300 by 300, for example [88]. The array will be 300x300x3 in this example. Where 300 is the width, 300 is the height, and 3 is the RGB channel values. For each number, the computer gives it a value between 0 and 255. This number tells you how bright the pixel is at each position [87].

You may also use convolutional neural networks to look at documents [94]. This is helpful for analysing handwriting and is a big part of recognisers [95]. A machine must carry out over a million commands per minute in order to scan someone's writing and compare it to its large database [96]. People say that CNNs and other models and algorithms have made the mistake rate go up, see Figure 2.

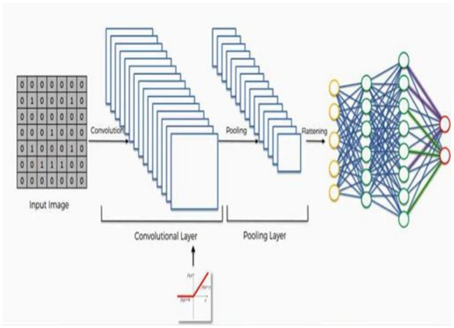

Figure 2. Layers involved in CNN.

Literature Survey

People who have trouble hearing and speaking have a lot of trouble talking to other people. The suggested effort helps those who can't hear or speak, or who are blind, talk to other people. The suggested approach offers speech/text output for sign input, creating a user-friendly platform [2]. Most people don't know the sign language they use because they are a minority. So, the proposed idea is a system that changes American Sign Language (ASL) into text and speech. ASL says that this technique uses convolutional neural networks (CNN) to find useful hand features that can be used to figure out hand movements. The suggested model is 88% accurate. This technique helps identify the hand motions of some people and turns them into text and speech so that they can talk to normal people more easily [1].

People need to talk to one another a lot. People who are "specially abled," meaning they have speech or hearing problems, or who are "mute" or "deaf," rely on visual communication all the time. Individuals who don't know sign language sometimes have trouble helping individuals who need it because they can't see or hear. They understand sign language well and utilise it to talk to each other [11]. To allow two-way communication between people with disabilities and the general public, a system that can turn gestures into text and speech is needed. Translating these sign languages into text and voice can be very helpful since it will connect the deaf and mute with the rest of society by making it possible for them to talk to each other [12]. This study report suggests a way to record hand movements, which is how people use sign language. It leverages the MediaPipe platform, which correctly identifies the hand and gets the information it needs to feed into the model [13]. After the model understands the sign language, it puts words and letters together to make a phrase, which is then turned into speech. The proposed framework has shown good accuracy with LSTM. A poll found that CNN and LSTM give better outcomes in terms of quality and accuracy [3].

Deep learning techniques have changed a lot over the years, leading to big improvements in Sign Language (SL) identification, translation, and video production [16]. Nonetheless, despite these advancements, significant obstacles remain, particularly in attaining high recognition accuracy and superior visual quality in generated content. A thorough examination of the current literature uncovers a variety of methodologies, ranging from traditional machine learning techniques to advanced deep neural networks, all designed to tackle the intricacies associated with SL-related activities. Researchers have historically investigated numerous techniques to address the distinctive issues of sign language recognition and translation. Traditional machine learning techniques, including decision trees, support vector machines (SVMs), and Hidden Markov Models (HMMs), formed the foundation of early research in this sector. These methods were crucial steps in the right direction, but they often had trouble getting the precision and reliability needed for real-world use [4].

The emergence of deep learning led to a significant transformation in SL research, addressing these constraints. Convolutional Neural Networks (CNNs) became a strong tool for SL gesture detection because they can automatically learn hierarchical characteristics from raw input data [15]. CNN-based architectures, transfer learning, and data augmentation techniques have considerably increased recognition accuracy by capturing subtle spatial patterns inherent in sign gestures [16]. Also, using recurrent neural networks (RNNs), especially Long Short-Term Memory (LSTM) networks, has changed the way SL translation jobs are done. LSTM-based models have shown amazing skill at translating spoken language into sign language representations by taking advantage of the sequential character of sign language and capturing temporal relationships in gestures [14]. LSTM networks are good at capturing the subtle grammar and syntax of sign languages because they can remember context over long runs [5].

Recent progress in generative modelling has opened up new ways to improve the visual quality of sign language videos, in addition to recognition and translation [17]. Generative Adversarial Networks (GANs) have demonstrated potential in generating realistic sign language motions from textual descriptions [18]. In an adversarial training paradigm, GANs may learn to create very realistic and coherent sign language sequences by putting a generator network against a discriminator network. This solves the problem of visual fidelity in created content [6].

The literature emphasises the continuous endeavour to create SL systems that are more precise, resilient, and visually appealing [19]. Deep learning has made a lot of progress in the last several years, but there are still some problems that need to be solved, such as a lack of data, adapting to new domains, and real-time performance. Future research endeavours are set to expand upon these foundations, utilising innovations in model architectures, multimodal learning, and human-computer interaction to facilitate the emergence of the next generation of SL technologies [20]. A two-stage deep learning solution for recognising Arabic sign language continuously using word count prediction and motion images, Tamer Shanableh, IEEE Access, Year: 2023. The difficulty of recognising continuous sign language arises from the intrinsic complexity of capturing and interpreting the fluidity of signing, where the quantity of words in a phrase and their delimitations remain uncertain during the recognition phase [21]. To solve this problem, recent research has suggested a two-step method for predicting how many words are in a sign language sentence and then breaking the sentence into equal halves for analysis. This method is very different from how things are usually done, which often involves trouble with the changing length and time of sign language sequences [7].

The first step in the suggested technique is to train a model to guess how many words are in a sentence in sign language. This prediction is an important part of the segmentation process because it lets you break the input sequence down into separate units at the right time [22]. Once divided into segments, each unit is shown as a single image using a new method that combines the variations between frames, utilising motion estimate and compensation. This makes moving graphics that show the different ways that each sign language word moves. Transfer learning convolutional neural networks (CNNs) get useful information out of these motion pictures. Transfer learning lets CNN use pre-trained models on big datasets to find important features in motion images [23]. We then utilise these extracted features to teach two different deep learning models: one to guess how many words are in a sentence and the other to understand what the sign language sentences mean. One of the best things about this method is that it can place words in context by picking up on traces of phrases that came before or after each motion image. This contextual information makes sign language identification more accurate overall by helping the model comprehend how signing works in a sequence [24]. Using bidirectional Long Short-Term Memory (LSTM) layers also makes the word prediction and sentence recognition phases more accurate [8].

In conclusion, the suggested two-stage strategy is a new way to deal with the problems of recognising sign language continuously. The approach gets very accurate results when it comes to recognising words and sentences by guessing how many words are in each sentence and using motion images to get information about the context [25]. These results show how deep learning could help improve sign language recognition and make communication tools more accessible to everyone [9].

Sign Language Recognition (SLR) systems are very important for helping people who are hard of hearing and people who are not deaf talk to each other. But the fact that there isn't a single sign language (SL) makes things worse because many countries have their own cultural SLs, such as Korean, American, and Japanese sign language [26]. Current SLR systems are great at recognising their own cultural SLs, but they often have trouble understanding other SLs or multi-cultural sign languages (McSL). This work introduces GmTC, a novel end-to-end SLR system designed to translate McSL into similar text for enhanced comprehension, thereby addressing these problems. GmTC uses a Graph and General deep-learning network as two-stream modules to get useful features [27]. The first stream uses superpixel values and the graph convolutional network (GCN) to build a feature based on a graph. This method tries to find distance-based complicated relationship characteristics between the superpixels, which will help the model better understand the subtleties of McSL [28]. In the second stream, attention-based contextual information is used to get long-range and short-range dependence features. This information goes via multi-stage, multi-head self-attention (MHSA) and convolutional neural network (CNN) modules [29]. By putting these features together, GmTC makes final feature representations that are passed to the categorisation module for interpretation. Hared trials on five cultural promotion 205 HQ datasets, show that GmTC is more accurate than other state Basin models in each spa. This confirmation of superiority and generalisability highlights GmTC's ability to solve the problems that McSL presents [30]. GmTC plays a big role in making communication easier for people from different cultural and language backgrounds by giving McSL detection and translation a strong basis [10].

Existing System

Support Vector Machine (SVM) is a supervised machine learning method for both regression and classification. We also use the term "regression problems," but it works best for categorisation. The goal of the SVM algorithm is to find a hyperplane in an N-dimensional space that clearly separates the data points into different groups. Haar cascade is an algorithm that can find things in pictures, no matter how big or small they are or where they are in the picture. The research utilises several input sensors, gesture segmentation, feature extraction, and classification methodologies. The goal of this research is to look at and compare the many ways that SLR systems work and the ways that they classify things. It additionally proposes the most viable approach for the forthcoming study. Because classification methods have improved a lot lately, many of the most recent proposed works focus on them, including the hybrid method and Deep Learning. This study examines the categorisation methodologies employed in previous Sign Language Recognition systems. Based on our assessment, previous research has thoroughly examined HMM-based techniques, including various variations. This study utilises several input sensors, gesture segmentation, feature extraction, and classification techniques. This work seeks to analyse and evaluate the methodologies utilised in the classifications of SLR systems and to propose the most trustworthy strategy for future research endeavours. Recent improvements in classification methods have led to the emergence of numerous newly proposed works that primarily focus on classification techniques, including hybrid methods and Deep Learning. Based on our assessment, previous research has looked into HMM-based methods a lot, including changes to them. Hybrid CNN-HMM and completely Deep Learning methodologies have demonstrated encouraging outcomes and present avenues for additional investigation [97].

In this study, we suggested various ways to make it easier for individuals to recognise signs so they can talk to each other. The result of such symbols and signs will be turned into text. We use a camera to collect hand gestures and turn them into a black-and-white image for this project [98]. The Otsu thresholding technique is used to break up the grey-scale image of a hand gesture into parts. There are two classifications for the whole image level: one for the hand and one for the background. The best threshold value is calculated by computing the ratio between class variance and total class variance. The Canny edge detection method is used to discover the edges of hand gestures in the picture. We used edge-based and threshold-based segmentation in Canny edge detection. Then, Otsu's algorithm is utilised since it is easy to calculate and stable. This approach doesn't work when the goal and background are very different around the world.

Recognising sign language by computers is an essential scientific subject that will help people who are deaf or hard of hearing communicate. This study presents an efficient and rapid algorithm for determining the count of fingers extended in a gesture symbolising a letter of the Binary Sign Language. The system doesn't need the hand to be completely in line with the camera. The research employs an image processing system to figure out what sign language deaf individuals use to talk to one another, specifically English alphabetic sign language. The main goal of this project is to make a smart computer system that will let people who are deaf or have intellectual disabilities talk to everyone else using their natural hand movements. The concept involved creating an intelligent system that uses image processing, machine learning, and artificial intelligence to take visual inputs of sign language hand motions and turn them into outputs that are easy to understand. So, the goal of this project is to construct an intelligent system that can translate between sign language and spoken language in real time and make it easier and more effective for people who are deaf or hard of hearing to talk to others who are not. We are putting in place a system for Binary sign language that can find any sign language after processing the image.

One of the worst things about our culture is that it makes it hard for people with disabilities to meet and talk to people who don't have disabilities. The only way to send a message or share our ideas is through communication. Still, someone who is deaf or stupid has problems communicating with people who are normal. Sign language is the principal way that many deaf and hard-of-hearing people converse with one another. The purpose of sign language recognition (SLR) is to let deaf and hard-of-hearing people talk to hearing people by having computers automatically interpret sign language. We aim to build a system that will help the person who teaches deaf and hard-of-hearing people how to utilise sign language or hand gesture recognition to talk to other people. This system uses the SURF method and image processing to locate and pull out aspects of hand movements. This task is done with the help of the MATLAB program.

A speech impediment is a disability that makes it hard for a person to hear and talk. People like them use sign language to talk to each other. People who don't know sign language still have problems talking to people who can't speak. The objective of this work is to develop an application that translates sign language into English, both in text and audio formats, thereby facilitating communication with those who utilise sign language. The application uses the computer's webcam to acquire visual data. Then it uses a combination algorithm to process the input and recognise it using template matching. After that, the written translation is transformed into sound. There are 6000 photos of English letters in the database of this system. We used 4800 photographs to train and 1200 pictures to test. The system is right 88% of the time. This research presents a prototype system engineered to detect hand movements from regular humans to enhance communication with specific persons. The previous study examines the challenge of gesture recognition within the deaf community employing sign language. Digital Image Processing and methods like Colour Segmentation, Skin Detection, Image Segmentation, Image Filtering, and Template Matching are all used to solve the problem. This system can understand American Sign Language (ASL) gestures, like the alphabet and parts of words.

Proposed System

The newest computer vision algorithms are used in YOLOv8 to develop a sign language recognition system that helps people who are deaf or hard of hearing talk to each other better. The first thing to do is get a lot of information. This has a lot of sign language motions that are well marked with boundary boxes and labels. The YOLOv8 model learns from this dataset by applying transfer learning to alter its pre-trained structure so that it can better recognise sign language. The model has been fine-tuned to recognise hand motions with high accuracy and reliability by evaluating hyperparameters and data augmentation techniques over and over. After training, the YOLOv8 model can find sign language movements in real time from video streams or live webcam feeds.

The model uses complicated techniques like non-maximum suppression to detect and mark hand areas in the input frames. It then paints boxes around them so you can see them. The system also employs a different classifier, which might be based on convolutional neural networks (CNNs) or recurrent neural networks (RNNs), to turn the hand motions it finds into text. Gesture classification is a key aspect of the system because it allows it to turn gestures it sees into text or voice that makes sense. By connecting recognisable motions to their spoken equivalents, the system makes it easy for persons who use sign language and those who don't to converse with one another.

Adding text-to-speech (TTS) synthesis also enables individuals to hear feedback immediately, which makes it easier for people with diverse ways of talking to utilise. We make sure that the system works well and is easy to use by testing and evaluating it thoroughly. We utilise a distinct test dataset to check things like accuracy, precision, recall, and F1-score. This offers us numbers that tell us how well the system performs. You can also get a qualitative sense of how well the system functions in the real world and how straightforward it is to use by asking end users for feedback, especially those who are deaf or hard of hearing, see Figure 3.

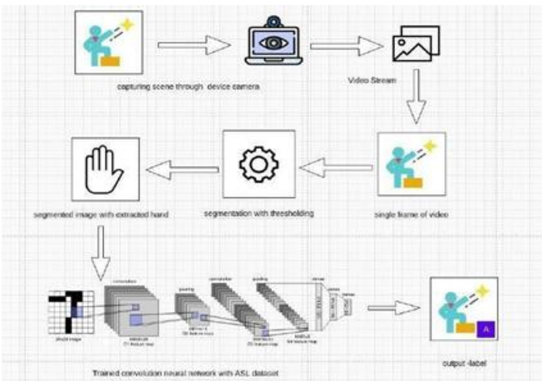

Figure 3. Architecture of Sign Language Recognition System.

Methodology

Supervised machine learning: This is one technique to teach a model by giving it input and expected output data. To make a model like this. It is important to go through these steps:

- Model training
- Model testing
- Model evaluation

Model construction: It depends on how machine learning works. It was neural networks in this project. An algorithm like this looks like this:

- Begin with its object: model = Sequential ()
- Then, there are layers with their types: model. add(type_of_layer()
- After adding a sufficient number of layers, the model is compiled. Before model training, it is important to scale the data for further use.

After building the model, it's time to train it. During this stage, the model learns from training data and the expected output for that data. It looks like this: model. fit(training_data, expected_output). When the script starts, you may see progress on the console. It will tell you how accurate the model was in the end.

Model Testing: In this step, a second set of data is loaded. The model has never seen this data set before; therefore, we will check its accuracy. Once the model training is done and the model is known to give the proper answer, it may be saved with model.save("name_of_file.h5"). The stored model can finally be used in the actual world. This step is called model evaluation. This indicates that the model can be used to check new data.

Image scaling

When you resize a digital image in computer graphics and digital imaging, that's called image scaling. In video technology, upscaling or resolution augmentation is the process of making digital data bigger, see Figure 4. You can use geometric transformations to scale a vector graphic image without losing any quality. This means that the visual primitives that make up the image can be scaled. To scale a raster graphics image, you have to make a new image with more or fewer pixels. When you reduce the number of pixels (scale down), the quality usually goes down. From the perspective of digital signal processing, the scaling of raster graphics is a two-dimensional instance of sample-rate conversion, which entails the transformation of a discrete signal from one sampling rate (specifically, the local sampling rate) to another.

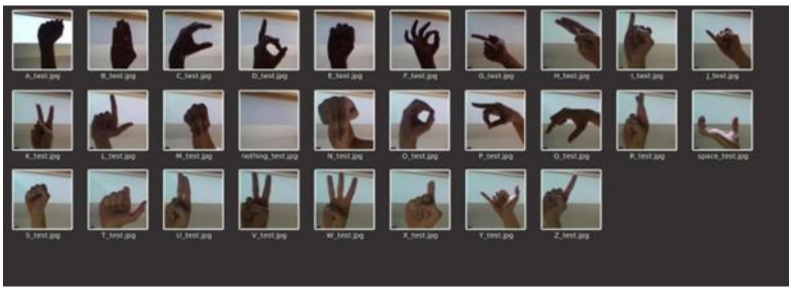

Figure 4. Sample pictures of training data

Algorithm

How to Calculate a Histogram: A histogram is a collection of counts of data that are put into a set of preset bins. When we talk about data, we don't just mean intensity value. You can collect any information that helps you explain your image. Let's look at an example. Think of a Matrix as a way to store information about an image, like how bright it is (0–255).

What do we do if we wish to count this data in an orderly way? We can break our range into smaller pieces, called bins, since we know that the range of information value for this situation is

256 values. For example:

[0,255]=[0,15] [16,31]

[240,255]range=bin1

bin2 bin=15

We can keep track of how many pixels are in each bin's range. Back Propagation: The main part of training a neural net is backpropagation. This is how you adjust the weights of a neural net based on the error rate from the last epoch (iteration). By boosting the model's generalisation, you can lower the error rates and make it more reliable by properly tweaking the weights. The abbreviated term "backwards propagation of errors" is "backpropagation." It is a common way to teach artificial neural networks. This method helps find the gradient of a loss function with respect to all the weights in the network.

Optimiser (Adam): Adam is like RMSprop and Stochastic Gradient Descent with momentum all in one. It scales the learning rate like RMSprop by using the squared gradients. It also employs momentum by using the moving average of the gradient instead of the gradient itself, like SGD with momentum. Adam is a method for adaptive learning rates that calculates different learning rates for each parameter. The term comes from adaptive moment estimation, which is what Adam does to change the learning rate for each weight in the neural network. It employs estimates of the first and second moments of the gradient to do this.

How Image Segmentation Works

Image segmentation is the process of turning an image into a set of pixel sections that are shown by a mask or a labelled picture. You can only process the crucial parts of an image by breaking it up into segments instead of processing the whole thing. One frequent method is to search for sudden changes in pixel values, which usually show edges that define a region. Another frequent way to do this is to look for similarities in different parts of an image. Region expanding, clustering, and thresholding are some methods that use this idea. Different methods for picture segmentation have been created that use knowledge from a specific field to tackle segmentation challenges in that field.

Classification: Convolution Neural Network

When you do image classification, you take an input (such as a photo) and give it a class or a chance that it belongs to that class. Neural networks are used in these steps: Encode the data with one-hot. You can use a one-hot encoding on the integer representation. This is where the integer-encoded variable is taken out, and a new binary variable is inserted for each integer value that is different. Define the model: In its simplest form, a model is just a function that takes in specific inputs, does its best to learn from them and then forecasts or classifies them, and then gives the right output. Put the model together: The optimiser sets the rate at which you learn. Our optimiser will be "Adam." In most circumstances, Adam is a good optimiser to utilise. During training, the Adam optimiser changes the learning rate. The learning rate controls how quickly the model's best weights are found. A lower learning rate may give you more precise weights, but it will take longer to figure them out.

Teach the model: To train a model, you need to learn (figure out) the best values for all the weights and the bias from samples that have been tagged. In supervised learning, a machine learning algorithm looks at a lot of samples and tries to find a model that minimises loss. This is termed empirical risk minimisation. A convolutional neural network utilises 2D convolution layers to combine learnt features with input data. Convolution Operation: Convolution is a mathematical function that combines two other functions through integration to show how one form changes when the other shape is added to it.

Convolution formula

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$

System Design

Dataflow Diagram: The DFD is often called a bubble chart. It is a simple graphical way to show how a system works, including the input data, the processing that happens on that data, and the output data that the system produces. It shows how information moves through a process or system and how inputs and outputs are handled. It shows where data comes from, goes, and is stored, as well as the paths between each destination, using set symbols such as rectangles, circles, and arrows. You can use them to look at an old system or come up with a new one. A DFD may often show things that are hard to put into words, and they work for both technical and non-technical reasons.

Conclusion and Future Scope

These days, apps need different types of photos to help them understand and analyse things. To do different things, you need to pull out a few features. Degradation happens when a picture is changed from one form to another, like when it is digitised, scanned, sent, or stored. The produced image needs to go through a process called image enhancement, which is a set of processes that try to make an image look better. Image enhancement basically informs human viewers on the interpretability or awareness of information in images and offers improved input for various automated image processing systems. Then, the image is processed in several ways to extract features that make it easier for the computer to read. A sign language recognition system is a very useful tool for gathering expert knowledge, finding edges, and putting together wrong information from diverse sources. The suggested sign language identification system for identifying sign language letters can be enhanced to include the recognition of gestures and facial expressions. Instead of letter labels, sentences should be shown as better translations of the language. This also makes it easier to read. You can make diverse sign languages more useful. You can add more training data to make it easier to find the letter. This idea might perhaps be expanded to turn the signs into speech.

References

[1] K. Alvi, M. Y. B. Azhar, M. Usman, S. Mumtaz, S. Rafiq, R. U. Rehman, and I. Ahmed, ‘‘Pakistan sign language recognition using statistical template matching,’’ Int. J. Inf. Technol., vol. 1, no. 3, pp. 765–768, 2007.

[2] S. Kausar, M. Y. Javed, and S. Sohail, ‘‘Recognition of gestures in Pakistani sign language using a fuzzy classifier,’’ in Proc. 8th Conf. Signal Process., Computer. Geometry Artif. Vis., 2008, National University of Sciences and Technology (NUST), Pakistan, 2008.

[3] Z. Halim and G. Abbas, ‘‘A Kinect-based sign language hand gesture recognition system for hearing- and speech-impaired: A pilot study of Pakistani sign language,’’ Assistive Technol., vol. 27, no. 1, pp. 34–43, 2015.

[4] Z. Ren, H. Li, C. Yang, and Q. Sun, ‘‘Multiple kernel subspace clustering with local structural graph and low-rank consensus kernel learning,’’ Knowledge-Based Syst., vol. 188, no. 1, p. 105040, 2020.

[5] I. Lauriola, C. Gallicchio, and F. Aiolli, ‘‘Enhancing deep neural networks via multiple kernel learning,’’ Pattern Recognit., vol. 101, no. 5, p. 107194, 2020.

[6] S. Jadooki, D. Mohamad, T. Saba, A. S. Almazyad, and A. Rehman, ‘‘Fused features mining for depth-based hand gesture recognition to classify blind human communication,’’ Neural Comput. Appl., vol. 28, no. 3, pp. 3285–3294, 2017.

[7] C. Koneti, G. C. Saha, and E. Howard, "End-to-End Visibility in Global Supply Chains: Blockchain and AI Integration," European Economic Letters, vol. 14, no. 4, pp. 545–555, 2024.

[8] M. Madanan, P. Patel, P. Agrawal, P. Mudholkar, M. Mudholkar and V. Jaganraja, "Security Challenges in Multi-Cloud Environments: Solutions and Best Practices," 2024 7th International Conference on Contemporary Computing and Informatics (IC3I), Greater Noida, India, 2024, pp. 1608-1614.

[9] P. Agrawal, N. Marathe, H. Byeon, and S. K. Singh, Machine Learning: Application and Challenges, p. 222, Xoffencer international book publication house, Chhetak Puri, Gwalior, 2024.

[10] P. Agrawal, R. Arora, W. C. Dietrich, R. L. Haecker, R. Hazeu, and S. Singh, "Method, system, and computer program product for implementing automated worklists," U.S. Patent 8,326,864, Dec. 4, 2012.

[11] R. Ingle, Dr. S. Donthu, M. H. K. Kochha, P. Agrawal, Dr. A. M. Kulkarni, and B. Viyyapu, "Fake news detection in social media management using deep learning and AI," Indian Patent Application 202441050770, 2024.

[12] V. Samatha N. Praba, P. Agrawal, P. Tripathi, N. Jain, and B. Kanwer, "Data security and privacy concerns in cloud-based HRM systems," J. Informatics Educ. Res., vol. 4, no. 3, pp. 1374-1380, 2024.

[13] P. K. Aggarwal, D. H. Rakesh, R. Arya, P. Agrawal, P. Kumar, and H. Y. S., "Chatbots and virtual assistants: Revolutionizing customer service and engagement in marketing," J. Informatics Educ. Res., vol. 4, no. 3, pp. 2044-2049, 2024.

[14] Md S. Miah and Md S. Islam, "Big Data Analytics Architectural Data Cut-Off Tactics for Cyber Security and Its Implication in Digital Forensic," 2022 International Conference on Futuristic Technologies (INCOFT), Belgaum, India, 2022, pp. 1-6.

[15] M. Abu Obaida, Md S. Miah, and Md A. Horaira, "Random Early Discard (RED-AQM) Performance Analysis in Terms of TCP Variants and Network Parameters: Instability in High-Bandwidth-Delay Network," International Journal of Computer Applications, vol. 27, no. 8, pp. 40-44, 2011.

[16] M. Katamaneni, P. Agrawal, S. Veera, A. K. Sahoo, K. Singh Sidhu and M. F. Hasan, "AI-Based Risk Management in Financial Services," 2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES), Lucknow, India, 2024, pp. 1-5.

[17] Kumar, B., Agrawal, P., Uike, D., Lourens, M., Bhati, I., & Peter, M.C. (2024, December). ML Techniques for Employee Performance Prediction. In 2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp.1-6). IEEE.

[18] L. N. R. Mudunuri, M. Hullurappa, V. R. Vemula, and P. Selvakumar, “AI-powered leadership: Shaping the future of management,” in Advances in Business Strategy and Competitive Advantage, IGI Global, USA, pp. 127–152, 2024.

[19] M. Hullurappa and M. Kommineni, “Integrating blue-Green Infrastructure into urban development: A data-driven approach using AI-enhanced ETL systems,” in Advances in Public Policy and Administration, IGI Global, USA, pp. 373–396, 2024.

[20] M. Hullurappa, “Uniting Quantum Computing and Artificial Intelligence: Exploring New Frontiers,” FMDB Transactions on Sustainable Computer Letters., vol. 2, no. 2, pp. 120–130, 2024.

[21] M. Hullurappa, "Fairness-Aware Machine Learning: Techniques for Ensuring Equitable Outcomes in Automated Decision-Making Systems," Int. J. Adv. Eng. Res., vol. 28, no. 5, pp. 9, 2024.

[22] M. Hullurappa, "Natural Language Processing in Data Governance: Enhancing Metadata Management and Data Catalogs," Int. Sci. J. Res., vol. 6, no. 6, pp. 1-22, 2024.

[23] D. B. Acharya, B. Divya, and K. Kuppan, "Explainable and fair AI: Balancing performance in financial and real estate machine learning models," IEEE Access, vol. 12, no.10, pp. 154022–154034, 2024.

[24] K. Kuppan, D. B. Acharya, and B. Divya, "Foundational AI in insurance and real estate: A survey of applications, challenges, and future directions," IEEE Access, vol. 12, no. 12, pp. 181282–181302, 2024.

[25] D. B. Acharya, K. Kuppan and B. Divya, "Agentic AI: Autonomous Intelligence for Complex Goals—A Comprehensive Survey," in IEEE Access, vol. 13, no.1, pp. 18912-18936, 2025.

[26] M. Hullurappa, "Exploring Regulatory Dimensions in Computing and Artificial Intelligence through Comprehensive Analysis," FMDB Transactions on Sustainable Computing Systems., vol. 2, no. 2, pp. 74–83, 2024.

[27] M. Hullurappa, "Intelligent Data Masking: Using GANs to Generate Synthetic Data for Privacy-Preserving Analytics," Int. J. Inventions Eng. Sci. Technol., vol. 9, no. 1, pp. 9, 2023.

[28] R. K. Dahal, "Customer satisfaction in Nepalese cellular networks," Tribhuvan University Journal, vol. 33, no. 2, pp. 59–72, 2019.

[29] R. K. Dahal, "Contemporary management accounting techniques and organizational performance," Pravaha, vol. 26, no. 1, pp. 177–185, 2020.

[30] R. K. Dahal, "Assessing social and environmental performance," Academy of Accounting and Financial Studies Journal, vol. 25, no. 6, pp. 1–9, 2021.

[31] R. K. Dahal, "Performance score as a measure of organizational effectiveness," Pravaha, vol. 27, no. 1, pp. 131–138, 2021.

[32] R. K. Dahal, "Customers' perspectives on the Nepalese cellular telecommunications services' technological and innovation capabilities," International Journal of Social Sciences and Management, vol. 9, no. 1, pp. 41–47, 2022.

[33] Pothu, A. R., "Celery Trap: A Browser and Email-Based Extension for Proactive Phishing, Spearphishing, and Web Threat Detection," SSRN, Oct. 10, 2024. [Online]. Available: https://ssrn.com/abstract=4983399.

[34] M. A. Raj, M. A. Thinesh, S. S. M. Varmann, A. R. Pothu, and P. Paramasivan, "Ensemble-Based Phishing Website Detection Using Extra Trees Classifier," AVE Trends In Intelligent Computing Systems, vol. 1, no. 3, pp. 142 –156, 2024.

[35] S. Chundru, "Harnessing AI's Potential: Transforming Metadata Management with Machine Learning for Enhanced Data Access and Control," International Journal of Advances in Engineering Research, vol. 27, no. 2, pp. 39-49, 2024.

[36] S. Chundru, "Beyond Rules-Based Systems: AI-Powered Solutions for Ensuring Data Trustworthiness," International Transactions in Artificial Intelligence, vol. 7, no. 7, p. 17, 2023.

[37] S. Chundru, "Seeing Through Machines: Leveraging AI for Enhanced and Automated Data Storytelling," International Journal of Innovations in Scientific Engineering, vol. 18, no. 1, pp. 47-57, 2023.

[38] E. Geo Francis and S. Sheeja, "Enhanced intrusion detection in wireless sensor networks using deep reinforcement learning with improved feature extraction and selection," Multimedia Tools and Applications, 2024.

[39] E. Geo Francis and S. Sheeja, "Bi-Level Intrusion Detection in IoT Networks Using Ensemble Method and A-GRU-RNN Classifier," Electric Power Components and Systems, 2024.

[40] E. G. F., S. Sheeja, John, A., and J. Joseph, “Intrusion detection system with an ensemble DAE and BiLSTM in the fog layer of IoT networks,” Journal of Applied Research and Technology, vol. 22, no.6, pp. 846–862, 2024. https://jart.icat.unam.mx/index.php/jart/article/view/2485

[41] E. Geo Francis, S. Sheeja, E.F. Antony John and Jismy Joseph, “An Efficient Intrusion Detection System using a Multiscale Deep Bi-Directional GRU Network to Detect Blackhole Attacks in IoT based WSNs,” Journal of Multiscale Modelling, vol. 15, no. 3, 2024.

[42] E. Geo Francis, S. Sheeja, E. F. Antony John and J. Jismy, "IoT Network Security with PCA and Deep Learning for Unmasking Anomalies," 2024 IEEE 13th International Conference on Communication Systems and Network Technologies (CSNT), Jabalpur, India, 2024, pp. 322-328.

[43] H. Mistri, A. Ghosh, and M. Dandapathak, "Bidirectional triple-band truly incident angle insensitive polarization converter using graphene-based transmissive metasurface for terahertz frequency," Frequenz, vol. 78, no. 11-12, pp. 569-579, 2024.

[44] A. Ghosh, A. Ghosh, and J. Kumar, "Circularly polarized wide-band quad-element MIMO antenna with improved axial ratio bandwidth and mutual coupling," IEEE Antennas Wireless Propag. Lett., vol. 23, no. 12, pp. 4718-4722, 2024.

[45] H. Mistri, A. Ghosh, A. R. Sardar, and B. Choudhury, "Performance enhancement of graphene-based linear to circular polarization converter for terahertz frequency using a novel parameter prediction methodology," Plasmonics, pp. 1-15, 2024.

[46] S. Nej, S. K. Bairappaka, B. N. V. Sai Durga Sri Raja Ram Dinavahi, S. Jana, and A. Ghosh, "Design of a high order dual band MIMO antenna with improved isolation and gain for wireless communications," Arab. J. Sci. Eng., pp. 1–18, 2024.

[47] S. K. Bairappaka, A. Ghosh, O. Kaiwartya, A. Mohammad, Y. Cao, and R. Kharel, "A novel design of broadband circularly polarized rectenna with enhanced gain for energy harvesting," IEEE Access, vol. 12, pp. 65583–65594, 2024.

[48] A. Gupta, N. Mahesh, S. K. Bairappaka, and A. Ghosh, "Comparison of the performance of L and Pi matching networks for design of a 2.4 GHz RF-DC rectifier for RF energy harvesting," in Proc. 2024 IEEE 4th Int. Conf. Sustainable Energy and Future Electric Transportation (SEFET), Hyderabad, India, 2024, pp. 1–5.

[49] S. Genikala, A. Ghosh, and B. Roy, "Triple band single layer microwave absorber based on closed loop resonator structures with high stability under oblique incidence," AEU-Int. J. Electron. Commun., vol. 164, Art. no. 154629, 2023.

[50] K. Mazumder, A. Ghosh, A. Bhattacharya, S. Ahmad, A. Ghaffar, and M. Hussein, "Frequency switchable global RFID tag antennae with metal compatibility for worldwide vehicle transportation," Sensors, vol. 23, no. 8, p. 3854, 2023.

[51] B. Leena and A. N. Jayanthi, "Brain Tumor Segmentation and Classification-A Review," Annals of the Romanian Society for Cell Biology, vol. 25, no. 4, pp. 11559-11570, 2021.

[52] N. T. J. Thirumurugan, T. Thirugnanam, L. Bojaraj, and L. R, "Formulation of a two-level electronic security and protection system for malls," International Journal of Electronic Security and Digital Forensics, vol. 16, no. 1, pp. 63-72, 2024.

[53] B. Leena, "Deep Learning-Based Convolutional Neural Network with Random Forest Approach for MRI Brain Tumour Segmentation," in System Design for Epidemics Using Machine Learning and Deep Learning, 2023, pp. 83-97.

[54] L. Bojaraj and A. Jayanthi, "Hybrid Feature Extraction with Ensemble Classifier for Brain Tumor Classification," International Journal of Pattern Recognition and Artificial Intelligence, vol. 36, no. 10, pp. 2250031, 2022.

[55] D. Farah, J. Pandey, N. Kumar, S. Singh, S. Sen, D. Purohit, Y. Dave, H. Mahapatra, U. Sharma, and S. Mandal, "The Role of Wireless Telemedicine and e-Health in Modern Healthcare: Challenges and Innovations," International Journal of Communication Networks and Information Security, vol. 16, no. 3, pp. 313-322, Sep. 2024.

[56] H. Mahapatra and M. Khuntia, "Does positive family history increase the likelihood of early-onset diabetes?," International Journal of Diabetes Development Countries, vol. 35, no. Suppl. 4, pp. S541, Dec. 2015.

[57] H. Mahapatra and M. Khuntia, "Familial transmission of type 2 diabetes," International Journal of Diabetes Development Countries, vol. 35, no. Suppl. 4, pp. S541, Dec. 2015.

[58] H. Mahapatra and M. Khuntia, "Health seeking behavior in type 2 diabetes in Odisha: Experience from a single centre," International Journal of Diabetes Development Countries, vol. 35, no. Suppl. 4, pp. S537, Dec. 2015.

[59] H. Mahapatra and M. Khuntia, "High percentage of Odia diabetic subjects achieving ADA goal for glucose," International Journal of Diabetes Development Countries, vol. 35, no. Suppl. 4, pp. S537, Dec. 2015.

[60] H. Mahapatra and M. Khuntia, "Prevalence of painful diabetic neuropathy among Odia subjects," International Journal of Diabetes Development Countries, vol. 35, no. Suppl. 4, pp. S583, Dec. 2015.

[61] H. Mahapatra and M. Khuntia, "Study of nephropathy using estimated glomerular filtration rate (eGFR) in type 2 diabetes subjects in Odisha," International Journal of Diabetes Development Countries, vol. 35, no. Suppl. 4, pp. S582, Dec. 2015.

[62] H. Mahapatra, B. Rusha, V. Halkurke, S. Shivashankaraiah, S. F. Hamid, M. Khuntia, and S. Yadav, "The Role of Nurse Practitioners in Optimizing Surgical Diabetes Care," Journal of Population Therapeutics and Clinical Pharmacology, vol. 31, no. 10, pp. 329-335, 2024.

[63] H. Mahapatra, L. Mahapatra, M. Khuntia, A. Sahoo, and S. Sucharita, "Dapagliflozin improves transaminitis in non-alcoholic fatty liver disease among adults with type 2 diabetes," Endocrine Abstracts, vol. 73, 2021.

[64] H. Mahapatra, L. Mahapatra, M. Khuntia, A. Sahoo, and S. Sucharita, "Vildagliptin is a safe and powerful oral antidiabetic drug for elderly adults with type 2 diabetes," Endocrine Abstracts, vol. 73, 2021.

[65] H. Mahapatra, L. Mahapatra, M. Khuntia, and A. Sahoo, "Dapagliflozin Improves Renal Function in Obese Indian Diabetics with Microalbuminuria," Diabetes, vol. 70, Supplement_1, p. 793-P, 2021.

[66] H. Mahapatra, L. Mahapatra, M. Khuntia, and A. Sahoo, "Dapagliflozin is Superior to Metformin as Monotherapy in Improving HbA1c and BMI in Drug Naive Type 2 Diabetic Adults," Endocrine Practice, vol. 27, no. 6, Supplement, S 56, 2021.

[67] H. Mahapatra, M. Khuntia, A. K. Sahoo, K. J. B., B. Jena, and S. R. Barik, "Are sleep disorders common in diabetes?," ESICON 2017 Abstracts, Indian Journal of Endocrinology & Metabolism, vol. 21, no. Suppl. 2, p. 29, 2017.

[68] H. Mahapatra, M. Khuntia, A. K. Sahoo, K. J. B., S. Mishra, and A. R. Jena, "Prevalence and risk factors of peripheral arterial disease in subjects with diabetic foot ulcer in Eastern India," ESICON 2017 Abstracts, Indian Journal of Endocrinology & Metabolism, vol. 21, no. Suppl. 2, pp. 4–79, 2017.

[69] H. Mahapatra, M. Khuntia, and A. R. Jena, "Efficacy and safety of vildagliptin as add-on therapy in patients uncontrolled with secretagogue and metformin combination: A single centre experience from Odisha," International Journal of Diabetes Development Countries, vol. 35, no. Suppl. 4, pp. S571, Dec. 2015.

[70] H. Mahapatra, M. Khuntia, and A. Sahoo, "Remogliflozin Reduces Gout Risk in Type 2 Diabetes," Endocrine Abstracts, vol. 70, 22nd European Congress of Endocrinology, Sep. 2020.

[71] A. Kulkarni, "Generative AI-Driven for SAP Hana Analytics," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 12, no. 2, pp. 438-444, 2024.

[72] Kulkarni, "Natural Language Processing for Text Analytics in SAP HANA," International Journal of Multidisciplinary Innovation and Research Methodology, vol. 3, no. 2, pp. 135-144, 2024.

[73] Kulkarni, "Enhancing Customer Experience with AI-Powered Recommendations in SAP HANA," International Journal of Business, Management and Visuals, vol. 7, no. 1, pp. 1-8, 2024.

[74] Kulkarni, "Digital Transformation with SAP Hana," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 12, no. 1, pp. 338-344, 2024.

[75] Kulkarni, "Supply Chain Optimization Using AI and SAP HANA: A Review," International Journal of Research Radicals in Multidisciplinary Fields, vol. 2, no. 2, pp. 51-57, 2024.

[76] Kulkarni, "Image Recognition and Processing in SAP HANA Using Deep Learning," International Journal of Research and Review Techniques, vol. 2, no. 4, pp. 50-58, 2024.

[77] R. S. Gaayathri, S. S. Rajest, V. K. Nomula, and R. Regin, "Bud-D: enabling bidirectional communication with ChatGPT by adding listening and speaking capabilities," FMDB Transactions on Sustainable Computer Letters, vol. 1, no. 1, pp. 49–63, 2023.

[78] R. C. Komperla, K. S. Pokkuluri, V. K. Nomula, G. U. Gowri, S. S. Rajest, and J. Rahila, "Revolutionizing Biometrics with AI-Enhanced X-Ray and MRI Analysis," in Advancements in Clinical Medicine, P. Paramasivan, S. Rajest, K. Chinnusamy, R. Regin, and F. J. Joseph, Eds. USA: IGI Global, 2024, pp. 1–16.

[79] A. S. Mohammed, A. R. Neravetla, V. K. Nomula, K. Gupta, and S. Dhanasekaran, "Understanding the Impact of AI-driven Clinical Decision Support Systems," in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Mandi, Himachal Pradesh, India, 2024, pp. 1–6.

[80] R. Neravetla, V. K. Nomula, A. S. Mohammed, and S. Dhanasekaran, "Implementing AI-driven Diagnostic Decision Support Systems for Smart Healthcare," in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Mandi, Himachal Pradesh, India, 2024, pp. 1–6.

[81] V. K. Nomula, A. S. Mohammed, A. R. Neravetla, and S. Dhanasekaran, "Leveraging Deep Learning in Implementing Efficient Healthcare Processes," in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Mandi, Himachal Pradesh, India, 2024, pp. 1–6.

[82] S. S. Ramesh, A. Jose, P. R. Samraysh, H. Mulabagala, M. S. Minu, and V. K. Nomula, "Domain Generalization and Multidimensional Approach for Brain MRI Segmentation Using Contrastive Representation Transfer Learning Algorithm," in Advancements in Clinical Medicine, P. Paramasivan, S. Rajest, K. Chinnusamy, R. Regin, and F. J. Joseph, Eds. USA: IGI Global, 2024, pp. 17–33.

[83] P. S. Venkateswaran, F. T. M. Ayasrah, V. K. Nomula, P. Paramasivan, P. Anand, and K. Bogeshwaran, "Applications of artificial intelligence tools in higher education," in Advances in Business Information Systems and Analytics, USA: IGI Global, 2023, pp. 124–136.

[84] H. Mahapatra, M. Khuntia, L. Mahapatra, S. Das, B. Jena, S. Mishra, A. R. Jena, R. R. Khatua, A. Patwari, S. K. Pradhan, S. S. Biswal, S. Nayak, U. S. Mishra, and A. K. Sahoo, "Long term efficacy and safety of canagliflozin in type 2 diabetes: A 52-week Indian

experience," Diabetes, vol. 68, no. Suppl. 1, Jun. 2019.

[85] H. Mahapatra, M. Khuntia, R. K. Padhi, B. Jena, S. R. Barik, and B. B. Kumar, "Pattern of pulse pressure distribution and its determinants in various stages of diabetic nephropathy," ESICON 2017 Abstracts, Indian Journal of Endocrinology & Metabolism, vol. 21, no. Suppl. 2, p. 29, 2017.

[86] H. Mahapatra, M. Khuntia, R. Mahapatra, B. Jena, A. K. Sahoo, and S. Sambit, "Impact of short term yoga practice on metabolic parameters in diabetes," Endocrine Review, ENDO 2018 Abstract, Chicago, USA.

[87] E. Geo Francis and S. Sheeja. “IDSSA: An Intrusion Detection System with Self-adaptive Capabilities for Strengthening the IoT Network Security,” Advances in Computational Intelligence and Informatics (ICACII), Hyderbad, India, 2024, Lecture Notes in Networks and Systems, vol 993, pp. 23-30.

[88] E. Geo Francis and S. Sheeja. “Chaotic Resilience: Enhancing IoT Security Through Dynamic Data Encryption,” Intelligent Informatics. (ISI), Bangalore, India, 2024, Smart Innovation, Systems and Technologies, vol 389, pp 331–344.

[89] S. Chundru, "Cloud-Enabled Financial Data Integration and Automation: Leveraging Data in the Cloud," International Journal of Innovations in Applied Sciences & Engineering, vol. 8, no. 1, pp. 197-213, 2022.

[90] S. Chundru, "Leveraging AI for Data Provenance: Enhancing Tracking and Verification of Data Lineage in FATE Assessment," International Journal of Inventions in Engineering & Science Technology, vol. 7, no.1, pp. 87-104, 2021.

[91] S. Chundru, "Ensuring Data Integrity Through Robustness and Explainability in AI Models," Transactions on Latest Trends in Artificial Intelligence, vol. 1, no. 1, pp. 1-19, 2020.

[92] M. Hullurappa, "Anomaly Detection in Real-Time Data Streams: A Comparative Study of Machine Learning Techniques for Ensuring Data Quality in Cloud ETL," Int. J. Innov. Sci. Eng., vol. 17, no. 1, pp. 9, 2023.

[93] M. Hullurappa, "The Role of Explainable AI in Building Public Trust: A Study of AI-Driven Public Policy Decisions," Int. Trans. Artif. Intell., vol. 6, no. 6, pp. 1-17, 2022.

[94] Anand S, Rajan M, Venkateshbabu N, Kandaswamy D, Shravya Y, and Rajeswari K., “Evaluation of the Antibacterial Efficacy of Azadirachta indica, Commiphora myrrha, Glycyrrhiza glabra Against Enterococcus faecalis using Real Time PCR,” Open Dent. J., vol. 10, pp. 160–165, May 2016.

[95] Chum J. D., Lim D. J. Z., Sheriff S. O., Pulikkotil S. J., Suresh A., and Davamani F., “In vitro evaluation of octenidine as an antimicrobial agent against Staphylococcus epidermidis in disinfecting the root canal system,” Restor. Dent. Endod., vol. 44, no. 1, pp. e8, Feb. 2019.

[96] Kandaswamy D., Venkateshbabu N., Arathi G., Roohi R., and Anand S., “Effects of various final irrigants on the shear bond strength of resin-based sealer to dentin,” J. Conserv. Dent., vol. 14, no. 1, pp. 40–42, Jan. 2011.

[97] Nagendrababu V., Jayaraman J., Suresh A., Kalyanasundaram S., and Neelakantan P., “Effectiveness of ultrasonically activated irrigation on root canal disinfection: a systematic review of in vitro studies,” Clin. Oral Investig., vol. 22, no. 2, pp. 655–670, Mar. 2018.

[98] Nagendrababu V., Pulikkotil S. J., Suresh A., Veettil S. K., Bhatia S., and Setzer F. C., “Efficacy of local anaesthetic solutions on the success of inferior alveolar nerve block in patients with irreversible pulpitis: a systematic review and network meta-analysis of randomized clinical trials,” Int. Endod. J., vol. 52, no. 6, pp. 779–789, Jun. 2019.