Artificial Intelligence Biosensing System on Hand Gesture Recognition for the Hearing Impaired

Kayal Padmanandam, BVRIT Hyderabad College of Engineering for Women, India

Rajesh M. V., Aditya Engineering College, India*

Ajay N. Upadhyaya, SAL Engineering and Technical Institute, India

K, Ramesh Chandra, Vishnu Institute of Technology, India

Chandrashekar B,, Sri Venkateswara College of Engineering and Technology, India

Swati Sah, Amity University, India

ABSTRACT

AI technologies have the potential to help deaf individuals communicate. Due to the complexity of sign fragmentation and the inadequacy of capturing hand gestures, the authors present a sign language recognition (SLR) system and wearable surface electromyography (sEMG) biosensing device based on a Deep SLR that converts sign language into printed message or speech, allowing people to better understand sign language and hand motions. On the forearms, two armbands containing a biosensor and multi-channel sEMG sensors are mounted to capture quite well arm and finger actions. Deep SLR was tested on an Android and iOS smartphone, and its usefulness was determined by comprehensive testing. Sign Speaker has a considerable limitation in terms of recognising two-handed signs with smartphone and smartwatch. To solve these issues, this research proposes a new real-time end-to-end SLR method. The average word error rate of continuous sentence recognition is 9.6%, and detecting signals and recognising a sentence with six sign words takes less than 0.9 s, demonstrating Deep SLR's recognition.

KEYWORDS

AI Biosensors, Artificial Intelligence (AI), Hand Gestures, Hearing-Impaired, Neural Networks Multi-Level Pronunciation, Sign Language Recognition (SLR), Wearable Computing,

INTRODUCTION

The major mode of communication between hearing-impaired persons and other populations is sign language (SL), which is represented through both manual and non-manual elements. The purpose for creating sign language tools to enhance communication in hearing-impaired people has long been recognised by the scholarly community. The implementation of applications can be difficult due to the great number of sign languages recent breakthroughs in AI and ML have helped to automate and improve such systems. The expansion of sophisticated ML algorithms that reliably identify human actions to isolated signs or continuous phrases is known as sign language recognition (SLR).

*Corresponding Author

Because of advancements in size and comfort, wearable sensors are becoming more common in applications to monitor health (Kim-Campbell, et al., 2019). Wearable biosensors can use ML algorithms for processing signal to deliver real-time monitoring of signals. The advantages of local (in-sensor) signal processing in lower communication connection bandwidth and radio power requirements are advantages of wirelessly streaming raw data to an external compute unit (Liu-Sacks, et al., 2017). Whenever the basic method of a classiðer fails to acknowledge a broad number of constraints, the model's classification accuracy degrades (Milosevic, Farella and Benaui, 2018). Furthermore, in-sensor model updates are not supported by systems capable of in-sensor training (Pancholi and Joshi, 2019).

A gesture is a physical movement of the hands, fingers, arms, and other parts of the human body that allows people to communicate meaning and information with one another. The data gloves method and the vision-based approach are two alternative approaches for human–computer interactions. The detection and classification of hand motions were among the investigations that looked into the vision-based approach. A One of the logical methods to create a convenient and adaptable interface between devices and users is to use hand gestures. HCI systems can use applications like virtual object manipulation, gaming, and gesture recognition. Hand tracking is a theoretical area of computer vision that deals with three key elements: hand segmentation, hand part identification, and hand tracking. Hand gestures are the best communicating approach and the most popular notion in a gesture recognition system. Hand gestures can be identified using one of the following methods: posture is a static hand form ratio without movement, and gesture is a dynamic hand motion with or without movement. Any camera may detect any form of hand gesture; keep in mind, however, that different cameras have varied resolution qualities. Most finger gestures can be detected by two-dimensional cameras in a continuous surface termed 2D.

One of the most common instances of a hand gesture system is sign language. It's a linguistic system that uses hand motions in addition to other motions. For example, most hearing-impaired people utilise universal sign language all across the world. The three basic components of sign language are word level sign vocabulary, non-manual characteristics, and finger spelling. Sign language is one of the most effective ways to communicate with hearing-impaired people.

Object detection and object motions were among the experiments given by the researchers. Three-Dimensional (3D) hand tracking is a hot topic in the gaming world. Recent film releases, such as Avatar, revolutionised cinema at the start of the decade by integrating content development and 3D technology with real performers, resulting in the birth of a new genre. Following the breakthrough of 3D movie, various electrical businesses concentrated their efforts on developing Three-Dimensional Television (3DTV) technology. The dome auto stereoscopic display was proposed by the researchers and is used to observe the position that is still constrained. Stereo and multi-view are two separate technologies that rely on the brain to merge the two views to produce the illusion of 3D.

The majority of studies follow a similar procedure for carrying out their experiments. Pre-processing is the initial phase employed in most research, and it basically involves preparing the image for the second phase. Following that, image processing prepares to receive the entire image so that it can be tracked with techniques like Wavelet Transform (WT) and Empirical Mode Decomposition (EMD). Many classifiers, such as Neural Network (NN) and Convolutional Neural Network (CNN), are released by artificial intelligence, each having the ability to categorise data based on its configuration and capabilities. The most capable tools for extracting visual features are the WT and EMD approaches. Feed forward is the type of ANN utilised in some experimental research for categorization. Apart from CNN, it is the most efficient classifier type for gesture recognition.

Hand gesture recognition applications can facilitate interaction with the non-impaired through sign language translation, which is the principal mode of communication for the partially deaf (Xu Zhang-Xiang Chen, et al., 2011). Hand gesture recognition has also showed promise in a number of upcoming applications, such as interactions with smartphones, virtual reality (VR) (Yang-Xu, et al., 2021) and then in selection control to prevent searching visually.

Sign language, which comprises of intricate gesture linguistics, is the major mode of communication for hearing-impaired people. The majority of people who aren't visually handicapped, on the other hand, don't understand sign language. As a result, the hearing impaired and the majority of the non-hearing impaired population have a considerable communication gap. Because hearing challenged people utilise sign language in everyday contexts, computer vision-based techniques are ineffective due to privacy concerns, lighting sensitivity, and higher energy usage (Pan-Tsai, et al., 2020). Artificial intelligence voice recognition (Sagayam and Hemanth, 2016), based on the current state of the aforesaid approaches. Previous studies have primarily focused on some particular sign language technologies (Zhao and Allison, 2019) and sign language translation (Badi and Hussein, 2014). The purpose of our research is to design the wide range of recent revolutionary recognition of hand gesture and wearable interfaces as well as the present problems that prevent practical application (Reyana-Krishnaprasath, et al., 2020).

Wearable devices that use surface electromyography to detect muscle activity could be useful in the development of hand motion recognition applications. For gesture classification, such devices often use machine-learning models, either locally or remotely. Most devices with local processing, on the other hand, are unable to train and update the machine-learning model while in use, resulting in inferior performance in real-world situations. We present a wearable surface electromyography biosensing system with in-sensor adaptive learning capabilities that is based on a screen-printed conformal electrode array. Our system uses a neuro-inspired hyperdimensional computing approach for real-time gesture classification, as well as model training and updating under multiple arm postures and sensor replacements.

Consider how difficult it would be to type on a computer without a keyboard, play a video game without a controller, or drive a car without a steering wheel. One of the purposes of a new device developed by engineers at the University of California, Berkeley, that recognises hand motions using electrical signals detected in the forearm is to do just that. Wearable biosensors and artificial intelligence (AI) are combined in this system, which could one day be utilised to control prostheses or interact with nearly any form of electronic device. Detection, Tracking, and Recognition are the three levels that make up the notion of recognising movements made with hands and/or other body parts. We employ customised interfaces to capture these movements, then use computer vision and deep learning algorithms to deduce the underlying pattern. There are currently various gesture-interface products on the market from Big Tech firms such as Intel, Apple, and Google for use in applications such as home automation, commerce, virtual/augmented reality games, consumer electronics, and navigation, among others.

Nature has endowed humans with a voice that allows them to engage and communicate with one another. Unfortunately, due to hearing and speaking impairments, not everyone has this capacity. Most people who are not familiar with sign language find it difficult to speak with the person without the use of an interpreter. As a result, there is a need to develop a technique that converts signals in hand gestures and the speaker's voice into simple text or audio that can aid real-time communication. As a result, we discover that a new strategy based on these new depth sensing devices, applied to machine learning, stochastic processes, and vision, is needed.

Wearable sensing electronic systems (WSESs) have sparked interest due to advancements in new materials and soft and stretchable circuits for a variety of applications, including health monitoring, disease diagnosis, personalised healthcare, on-demand treatment, assistive device, human–machine interface (HMI), and virtual and augmented reality. [The sensor unit, power unit, wireless communication unit, data collection/storage/transmission unit, and data processing unit are all common components of a WSES. Each of these components must be intelligent and intelligent in order for WSES to be used on a big basis. First, sensors must be sensitive, dependable, robust, and wearable, with high-quality sensing data being a requirement for WSES to work well. Second, there are three ways for powering devices currently available: self-powering, integrated battery, and wireless power. Third, data from the sensor system can be saved in the WSES' memory or wirelessly sent to an external device (i.e., tablet or cellphone). Note that in some circumstances, external sources (e.g., special equipment for spectroscopy data acquisition) are still used to capture sensing signals,

which compromises the wearability of the devices. Fourth, data processing can be configured in an online or offline mode within the WSES using external devices (i.e., cloud computing or cellphone).

On the one hand, optimising each component in the WSES on many aspects has resulted in numerous accomplishments, including carefully fitted materials, high permeability for wearing, long-term stability, and carefully constructed soft circuits, among others. The WSES, on the other hand, is still relatively unknown as a healthcare device or for other purposes. This is most likely because, despite the fact that a significant volume of raw data may be easily collected, the relevant information output from contemporary WSES still falls short of consumers' expectations. This is especially true when multiple sensor arrays with various data modalities are combined into a single system. As a result, within the paradigm of traditional data processing techniques, interpreting a vast volume of multimodal data becomes problematic. A promising strategy is to extract as much information as possible from the gathered raw data utilising current WSES without further complicating the device structure designs. Fortunately, with the rapid rise of artificial intelligence, new data processing algorithms could close the gap between customers' expectations and device performance (AI).

The combination of WSES data and AI techniques could be a game-changer, improving the current WSES' performance and revolutionising various applications in personal healthcare, public health, sports, and games. Machine learning (ML) has already been widely used in many domains of mathematics, physics, chemistry, engineering, and materials science as a subset of AI. In many ways, ML algorithms make the WSES more realistic for a future application as medical devices because they are a powerful tool for processing and evaluating raw data acquired from wearable devices. With the help of machine learning algorithms, a simple biosensor device made of gold nanoparticles can quickly screen for coronavirus illness 2019 (COVID-19). ML algorithms can quickly and correctly extract relevant information from the WSES, allowing for the assessment of several vital indications of health state. For type 1 diabetes, adequate disease management with appropriate dose adjustment has also been demonstrated. Other significant advances in recent years include sign language translation, human–machine haptic interaction, and brain–to–text communication, among others. The widespread use of machine learning algorithms in engineering and materials science research has already begun.

Wearable devices can collect a large amount of raw data from various physiological signs, which must then be processed and analysed. Many physiological signals from our daily activities, for example, are sensed as biochemical concentrations, biopotential patterns, and biophysical activity intensities. Users and doctors can alter treatments or take further steps more quickly and easily if they have a clear comprehension of the information given by these data from wearable sensors. Traditional data analysis paradigms, such as threshold limits, simple mathematical models (i.e., linear or polynomial regressions), or manual selection, are insufficient for large volumes of raw data handling due to a variety of factors, including data structure complexity and dimensionality, multiple modalities, and so on. As a result, machine learning algorithms have become important in establishing a new paradigm of data analysis that will aid in the advancement and practicality of smart and intelligent wearable gadgets. Useful information of various signal properties can be retrieved from raw data and utilised to the best extent possible by implementing appropriate ML algorithms, allowing these wearable devices to perform better in an intelligent manner. It is vital to highlight that choosing the right algorithm for different types of raw data is critical for establishing a correct and reliable correlation between sensing signals and physiological status. We cover significant data processing approaches reported in prior wearable electronics experiments in this section. Data preprocessing procedures like principle component analysis (PCA) and hierarchical cluster analysis (HCA), classification algorithms like support vector machine (SVM), decision tree (DT), and random forest (RF), and artificial neural network algorithms are all included (ANN). Here's a system that makes it easier for persons with disabilities to communicate. Computer recognition of hand gestures is a critical research subject for helping deaf and dumb individuals to communicate. The objective is to develop and build an intelligent system that takes visual inputs of sign language hand motions and generates easily recognised outputs utilising image processing, machine learning, and artificial intelligence ideas.

RELATED STUDIES

Speech recognition has made remarkable as well as the usage of huge data computing. People's daily lives have been impacted by multivoice technology. Amazon's Alexa, the Baidu signal-to-fly method, Apple's Sift, the Ding Ding intelligent sound box, and others are examples, with single-word Mandarin pronunciation recognition accuracy over 95%.

Chen et al. (Zhou-Chen, et al., 2020) also proposed utilising a wrist-mounted camera to detect background variations and derive finger movements. A hybrid technique combining the above sensing methods, as well as surface electromyography (sEMG), offers a promising option. Kudrinko (Sheng, 2019) just reviewed this topic, which has gotten a lot of interest.

The most basic target gesture sets in this application were 26 American sign language letters and the 10 American numerals (Kudrinko-Flavin, et al., 2021). Electromyography (EMG), sensor gloves and PPG are examples of biosensor-based systems. EMG sensors can record data indicating muscular activity. For collecting isolated gestures, Lu et al. coupled a sEMG sensor with an ACC signal, on the other hand, can only detect arm motions along vertical or horizontal axes.

Wearable sensors are used to capture sign language actions. In (Wang, Pan and Liu, 2018), Wang et al. used a two-armband system with both IMU and EMG sensors to collect hand locations. However, only three movements are recognised, and feature extraction is limited to finger locations rather than the complete hand.

Furthermore, the programme is not real-time. In contrast, Ozarkar et al. (Ozarkar-Chetwani, et al., 2020) developed a three-module smartphone application. The sound classification module recognised and classified input sounds, and vibrations notified the user. The Indian sign language video was detected by the gesture recognition module, which then transformed it to natural English. Furthermore, Paudyal et al. (Paudyal-Lee, et al., 2019) created a smartphone application that gives sign language feedback from the students based on the position, movement, orientation, and hand-shape signs.

Human hand gesture detection by AI systems has been a significant advancement in the recent decade, with applications in high-precision surgical robots, health monitoring equipment, and gaming systems. The use of inputs from wearable sensors has improved AI gesture recognition systems that were previously visual-only. This is known as 'data fusion.' Wearable sensors mimic the skin's detecting abilities, one of which is referred to as 'somatosensory.' The low quality of data received from wearable sensors, which is often owing to their bulkiness and poor contact with the user, as well as the impacts of visually occluded objects and poor lighting, continue to limit gesture detection precision.

AI-BIOSENSOR NETWORKS (AIBN)

The creation of biosensors has stimulated academic and corporate interest, and AI is a key factor in improving biosensor performance. As a result, AIBN has the ability not only to give early warning of many application scenarios but to modify our way of life through "smart" applications. Applications of AIBN can assist hearing challenged people in recognising hand gestures (Pardeshi, Sreemathy and Velapure, 2019).

Figure 1. Design of the proposed model
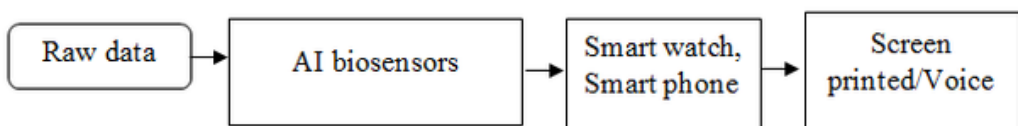
Thus, a compact and reliable SLR system is needed necessary to actually help the deaf people to interact with the normal individuals at any place. To overcome these issues, this study implements and designs a revolutionary end-to-end SLR system termed Deep SLR, as displayed in Figure 1. It continually converts sign language into voices in real-time so that individuals who are deaf or hard of hearing can comprehend what a hearing-impaired person is saying, even if they are unfamiliar with sign language (Moores, McIntyre and Weiss, 1973). Unlike other SLR systems, to gather sign signals on both forearms we use two armbands, each with an AI biosensor sensor and surface electromyogram (sEMG) sensors. Arm movements are captured by the AI biosensor, which consists of a gyroscope (GYRO) and an accelerometer (ACC); perfectly alright finger motions are captured by the sEMG sensors. Some basic hand gestures are shown in Figure 2. Which can be easily converted into voice recognition.

Figure 2. Basic sign language using hand gestures

Smartphone-Based Platforms

With the worldwide popularity, the smartphone-based sensing systems have earned a lot of interest. Smartphones are playing an increasingly essential role in AI-biosensors for sharing, cloud interaction, storage and data processing of various sensors and functionalities which are referred as communication and processing which will be helpful for hearing impaired people (Grover-Aggarwal, et al., 2021). Smartphones have additional hardware as Bluetooth, cameras, USB, and audio ports to improve accessibility to receive detection data and control the detection process. Data security will become increasingly important as smartphones collect more personal data from AI-biosensors.

Smartwatch Based System

Hand gesture recognition technology is also being used in smartwatches, which is an emerging field. Because the screens on smartwatches are very small for touching motions, hand motions for operation or typing could be a viable substitute. Popular hand gesture detection systems, such as sEMG and forcemyography (FMG), however, necessitate additional components and a substantial amount of space, which most smartwatches require (SAYEM, 2014). Existing smartwatch sensors are using recognition of hand gestures such as microphones, photoplethysmography (PPG) and bone-conducted sound sensing have been reported in new research. The PPG method has the potential to be implemented as a low-cost recognition method of gesture on commercial smartwatches. The PPG technique has several advantages, including being inexpensive, lightweight, and easily implemented on a smartwatch (Chen-Lv, et al., 2019).

METHODS

The high-level overview of Deep SLR is presented in this section. Data collection, data preprocessing and continuous recognition are the key steps of Deep SLR (Wang, Pan and Liu, 2018).

Data Collection

We employ two armbands on the forearms to collect both hands' real-time sign signals. Each armband has an AI biosensor and eight axes of sEMG sensors. The bio sensor records the angular velocity and acceleration of hand movements, while sEMG sensors record the muscle movements that correspond to hand actions (Reyana and Kautish, 2021).

Data Preprocessing

This step comprises of data cleaning and extraction features to normalise and eliminate noises in real-time signals. Given that the AI biosensors and sEMG sensors have various sample frequencies, we first use spline interpolation to normalise the gathered signals to the same length, and then to clear spike noise.

Continuous Recognition

We use a multi-channel CNN and an attention-based encoder-decoder model to achieve end-to-end continuous SLR without fragmentation for improving identification accuracy without division. Finally, we utilise a grammar-based classification models and a laser approach to infer the most likely sequence of words from the probability matrices, which we use as the final text phrase.

SIGN LANGUAGE RECOGNITION

The task of detecting sign language phrases from video feeds is known as sign language recognition (SLR). It is an essential research subject because it has the potential to overcome the communication gap between Deaf and hearing individuals, allowing hearing-impaired persons to participate more fully in society (Kudrinko-Flavin, et al., 2021). Furthermore, depending on whether the video streams sign language recognition can be characterised as CSLR and ISLR.

Continuous Sign Language Recognition

The purpose of Continuous Sign Language Recognition is to categorise signed videos into whole phrases. CSLR is a difficult job since it necessitates the detection of glosses from video streams without any prior information of sign limits. For feature extraction, most works use 2D or 3D-CNNs, for sequential modelling. The word error rate (WER) is a standard metric for CSLR performance.

The number of operations necessary to convert the projected sequence into the target classification is measured by WER (Wang-Zhao, et al., 2020; Lee, 2018; Lu-Chen, et al., 2014).

Isolated Sign Language Recognition

Isolated sign language recognition (ISLR) is the task of correctly identifying single sign motions from videos. It is typically approached in the same way as action and hand gesture recognition video process by extracting and understanding highly feature representations. Extraction of hand and mouth areas from video sequences is a frequent technique to the task of isolated sign language identification in the literature, in an effort to eliminate loud surroundings that can hamper classification results (CHU-LEE, et al., 2021; Samčović, 2020).

Sign Language Translation

Sign language videos are converted into spoken language by modelling not just the glosses but also the grammar and linguistic structure is known as sign language translation. It's a significant study field that helps Deaf and other communities communicate more effectively. Furthermore, due to the incorporation of linguistic norms and the representation of spoken languages, the SLT challenge is more difficult than CSLR.

PERFORMANCE EVALUATION

The performance of Deep SLR is presented in this section. We examine the model training parameters first, then assess the impact of each component in Deep SLR. Then we do a comparison between Deep SLR and ISLR. Following that, we perform a full evaluation of our approach, taking into account each individual and each phrase, and then we test the method's resilience by identifying phrases from additional participants. Finally, the performance of Deep SLR in real time is explored.

Evaluation Metrics

As an assessment parameter, we employ the word error rate (WER), which is extensively used in the recognition of speech, and CSLR. It calculates the smallest number of deletions, insertion and replacement operations required to convert a recognised text phrase to the ground truth. The WER of a recognised text phrase is represented as:

$$WER = \frac{\#\,sub \,+\, \#\,ins \,+\, \#\,del}{\#\,words}$$

Therefore #ins, #sub and #del represent the lowest number of insertion, substitution and deletion operations required to change the phrase and #words represents the set of words in the underlying data.

Parameter Analysis

The effects of several parameters on Deep SLR are investigated such as learning rate, hidden state size in the LSTM, and regulation strength. The assessment of these characteristics is shown in Figure 3. Without further explanation, WER is the average of 6 training sessons, and the error bars represent the variance. Because the learning rate is 0.2, we utilise 0.2 as the standard learning rate in our approach gets the lowest WER.

The regulatory strength is set to default levels, much like the size of the hidden state. Finally, we arrive at a customized model with 7,500 parameters, 9.6% WER on the training set, and 7% WER on the testing set. When we increase the label and double the training set, WER will reduce dramatically if we have enough time and enough personnel cost.

Figure 3. The illustration of parameter analysis
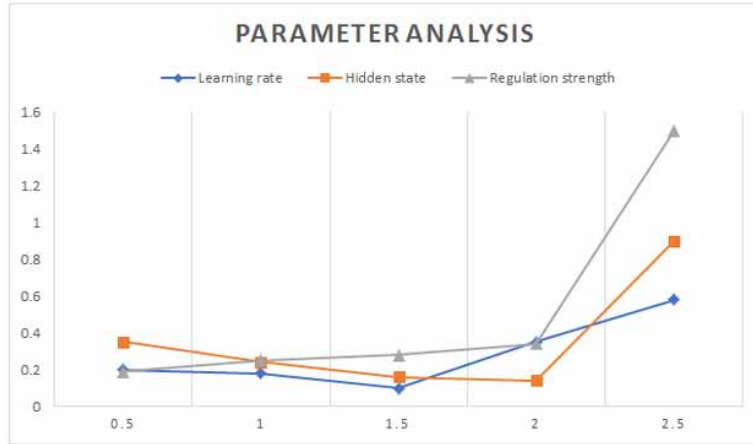
Effectiveness of Each Component

The influence of each component on Deep SLR recognition performance is established in this section.

Signals Impact

The AI-biosensor is only used by Sign Speaker for isolated fingerspelling recognition and CSLR. We gathered all the signs of hand gestures in addition to the whole test dataset. We can observe that when both sensors are used, the WER is always lower than when only the AI-biosensor is used, highlighting the relevance of sEMG signals in recording finger movements.

Hand Rotation Features Impact

All the data of signals with arm movements are selected in addition to the entire test set. We can observe the hand rotation characteristics of WER is always lower, implying that hand rotation is essential for accurate SLR.

Multi-channel CNN Impact

The recognition rate of Deep SLR is compared with and without the CNN to assess the scalability provided by the multi-channel CNN. It is constructed utilising examples from two people with very varying signal intensities, in addition to the test set. Clearly, the WER of Deep SLR with CNN is substantially lower. We also see that the WER of Deep SLR utilising CNN is nearly same, demonstrating that CNN improves SLR scalability.

Attention Mechanism Impact

By visualising the alignment function, we can further examine the efficiency of the attention process. Figure 4 depicts the display among one channel of sEMG signals from a sign phrase consisting of four Indian signs, as well as the accompanying context vectors c2, c3, c4, and c5. In the vector representation, a deeper square represents a greater value. We can observe that the greater results in the vector representation correlate to the locations of clear signals, implying that the approach has learnt to show additional "attention" to the relevant section in the signal obtained for each predictions, resulting in alignment.

**Figure 4. Alignments for signals with four Indian signs**
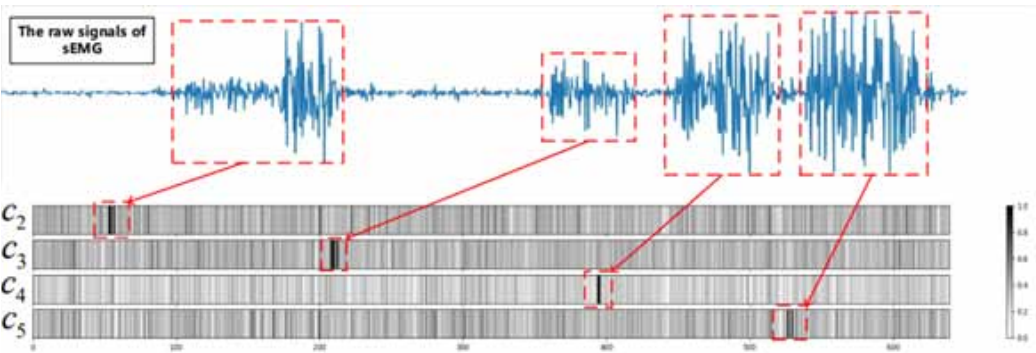
Comparison of AI Technologies

A comparison between AI technologies in terms of the suggested networks of CSL, ISLR and SLP's Accurateness, Requirements of Hardware, Existing datasets and Future potential is depicted in Figure 5.

Except for the real datasets, their results are determined on research for standard datasets. ISLR approaches great precision with low requirements of hardware, but they have been widely investigated with minimised future potential. CSLR and SLP approaches, on the other side, have substantial hardware and training standards as a significant impact on future researches.

**Figure 5. Comparison of AI technologies**
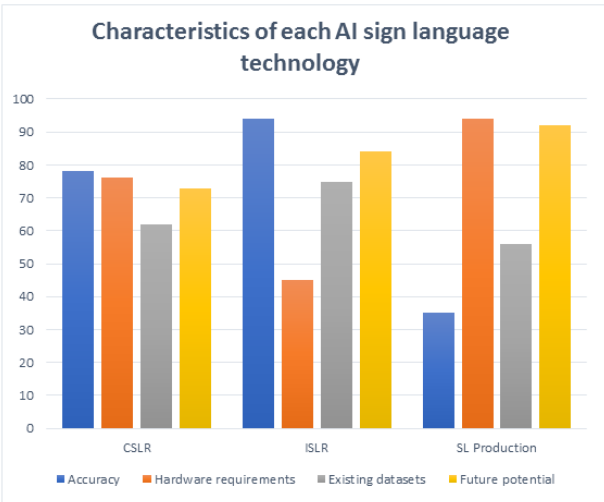
Real-time Performance

We examine Deep SLR's real-time performance from three different angles in this section: delay and recognition speed. The tests are carried out on three smartphones with low, medium, and high computational capabilities, namely the Redmi 9, MI note 10, and SAMSUNG Galaxy S21 ultra.

Delay

The time spent for signal identification and data preparation in Deep SLR is referred to as the delay. Figure 6(a) illustrates the time it takes for 10 participants to recognise sign texts of various lengths. We can observe that on devices with more Computational resources, the delay will be shorter. We also see that the delay gets greater as the phrases get lengthier. Note that the typical latency for processing a 6-word phrase for a smartphone with medium computational capabilities is merely 0.45 s, which has little bearing on SLR's actual results.

Figure 6. Performance of Real-time: (a) delay, (b) recognition speed
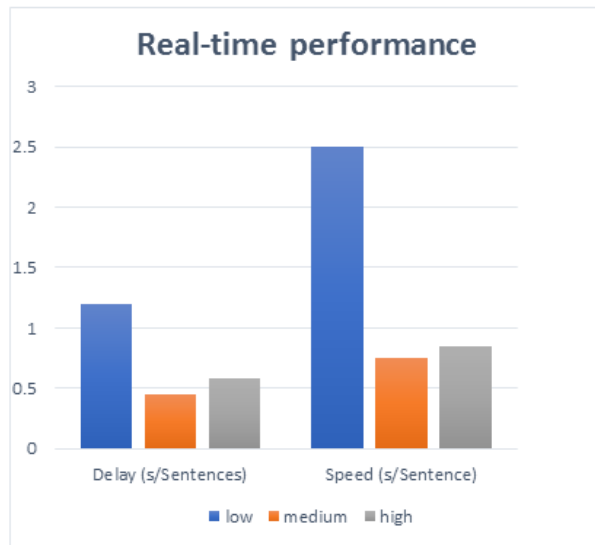
Recognition Speed

The avg time utilised for the recognition component, which is the real-time speed recognition of words over 10 individuals displayed in Figure 6(b). We can observe that the smartphone with weaker computer capability, as well as lengthier words, take longer. It's worth noting that the recognition accuracy speed of a phrase with six sign words for a smartphone with medium computational power is 0.65 seconds, demonstrating Deep SLR's actual capability.

CONCLUSION

In this research, we executed a real-time end-to-end CSLR system termed as Deep SLR to enable individuals "hear" sign language by translating it into speech. Hand movements and quite well finger gestures are captured using both sEMG and AI-biosensors. To provide efficient, adaptable, and end-to-end CSLR without sign division with encoder-decoder structure and multi-channel CNN was presented. Continuous sentence recognition has a WER of 9.6%, which is significantly lower than isolated methods, and it tends to take less than 0.9s to detect signs and recognise a phrase with six sign words, demonstrating the technique and limited ability of Deep SLR in real life situations. The experimental results also confirm Deep SLR's resilience and scalability. We intend to integrate our suggested technique in various wearable artificial intelligence biosensors in the future.

REFERENCES

A, R., & Kautish, S. (2021). Corona virus-related Disease Pandemic: A Review on Machine Learning Approaches and Treatment Trials on Diagnosed Population for Future Clinical Decision Support. *Current Medical Imaging: Current Medical Imaging Reviews, 17*. 10.2174/1573405617666210414101941

Badi, H. S., & Hussein, S. (2014). Hand posture and gesture recognition technology. *Neural Computing & Applications*, *25*(3–4), 871–878. doi:10.1007/s00521-014-1574-4

Chen, F., Lv, H., Pang, Z., Zhang, J., Hou, Y., Gu, Y., Yang, H., & Yang, G. (2019). WristCam: A Wearable Sensor for Hand Trajectory Gesture Recognition and Intelligent Human–Robot Interaction. *IEEE Sensors Journal*, *19*(19), 8441–8451. https://doi.org/10.1109/jsen.2018.2877978

Chu, W., Lee, K. T., Luo, W., Bhambri, P., & Kautish, S. (2021). Predicting the security threats of internet rumors and spread of false information based on sociological principle. *Computer Standards & Interfaces, 73*, 103454. 10.1016/j.csi.2020.103454

Grover, Y., Aggarwal, R., Sharma, D., & Gupta, P. K. (2021). Sign Language Translation Systems for Hearing/Speech Impaired People: A Review. *Proceedings of the IEEE 2021 International Conference on Innovative Practices in Technology and Management (ICIPTM)*, 10–14.

Kim, J., Campbell, A. S., de Ávila, B. E.-F., & Wang, J. (2019). Wearable biosensors for healthcare monitoring. *Nature Biotechnology*, *37*(4), 389–406. doi:10.1038/s41587-019-0045-y PMID:30804534

Kudrinko, K., Flavin, E., Zhu, X., & Li, Q. (2021). Wearable Sensor-Based Sign Language Recognition: A Comprehensive Review. *IEEE Reviews in Biomedical Engineering*, *14*, 82–97. doi:10.1109/RBME.2020.3019769 PMID:32845843

Kudrinko, K., Flavin, E., Zhu, X., & Li, Q. (2021b). Wearable Sensor-Based Sign Language Recognition: A Comprehensive Review. *IEEE Reviews in Biomedical Engineering*, *14*, 82–97. https://doi.org/10.1109/rbme.2020.3019769

Lee, B. G., & Lee, S. M. (2018). Smart Wearable Hand Device for Sign Language Interpretation System With Sensors Fusion. *IEEE Sensors Journal*, *18*(3), 1224–1232. https://doi.org/10.1109/jsen.2017.2779466

Liu, X., Sacks, J., Zhang, M., Richardson, A. G., Lucas, T. H., & van der Spiegel, J. (2017). The Virtual Trackpad: An Electromyography-Based, Wireless, Real-Time, Low-Power, Embedded Hand-Gesture-Recognition System Using an Event-Driven Artificial Neural Network. *IEEE Transactions on Circuits and Wystems. II, Express Briefs*, *64*(11), 1257–1261. doi:10.1109/TCSII.2016.2635674

Lu, Z., Chen, X., Li, Q., Zhang, X., & Zhou, P. (2014). A Hand Gesture Recognition Framework and Wearable Gesture-Based Interaction Prototype for Mobile Devices. *IEEE Transactions on Human-Machine Systems*, *44*(2), 293–299. https://doi.org/10.1109/thms.2014.2302794

Milosevic, B., Farella, E., & Benaui, S. (2018). Exploring Arm Posture and Temporal Variability in Myoelectric Hand Gesture Recognition. *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*, 1032-1037.

Moores, D. F., McIntyre, C. K., & Weiss, K. (1973). Gestures, Signs, and Speech in the Evaluation of Programs for Hearing Impaired Children. *Sign Language Studies*, *1002*(1), 9–28. doi:10.1353/sls.1973.0009

Ozarkar, S., Chetwani, R., Devare, S., Haryani, S., & Giri, N. (2020). AI for Accessibility: Virtual Assistant for Hearing Impaired. *Proceedings of the IEEE 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT),* 1–7.

Pan, T. Y., Tsai, W. L., Chang, C. Y., Yeh, C. W., & Hu, M. C. (2020). A Hierarchical Hand Gesture Recognition Framework for Sports Referee Training-Based EMG and Accelerometer Sensors. *IEEE Transactions on Cybernetics*, 1–12. doi:10.1109/TCYB.2020.3026571 PMID:32776885

Pancholi, S., & Joshi, A. M. (2019). Electromyography-Based Hand Gesture Recognition System for Upper Limb Amputees. *IEEE Sensors Letters*, *3*(3), 1–4. doi:10.1109/LSENS.2019.2898257

Pardeshi, K., Sreemathy, R., & Velapure, A. (2019). Recognition of Indian Sign Language Alphabets for Hearing and Speech Impaired People using Deep Learning. SSRN *Electronic Journal*. Published. 10.2139/ssrn.3430055

Paudyal, P., Lee, J., Kamzin, A., Soudki, M., Banerjee, A., & Gupta, S. (2019). *Learn2Sign: Explainable AI for Sign Language Learning*. IUI Workshops.

Reyana, A., Krishnaprasath, V., Kautish, S., Panigrahi, R., & Shaik, M. (2020). Decision-making on the existence of soft exudates in diabetic retinopathy. *International Journal of Computer Applications in Technology*, *64*(4), 375. doi:10.1504/IJCAT.2020.112684

Sagayam, K. M., & Hemanth, D. J. (2016). Hand posture and gesture recognition techniques for virtual reality applications: A survey. *Virtual Reality (Waltham Cross)*, *21*(2), 91–107. doi:10.1007/s10055-016-0301-0

Samčović, A. (2020). Accessibility of services in digital television for hearing impaired consumers. *Assistive Technology*, 1–10. 10.1080/10400435.2020.1757786

Sayem, A. (2014). Speech Analysis for Alphabets in Bangla Language: Automatic Speech Recognition. *International Journal of Engineering Research, 3*(2), 88–93. 10.17950/ijer/v3s2/211

Sheng, Ye. (2019). The realization of voice intelligent control based on Artificial Intelligence. *Science & Technology Information*, *17*(7), 19–21.

Wang, H., Pan, J., & Liu, C. (2018). Research development and forecast of automatic speech recognition technologies. *Telecommunications Science*, *34*(02), 1–11.

Wang, Z., Zhao, T., Ma, J., Chen, H., Liu, K., Shao, H., Wang, Q., & Ren, J. (2020). Hear Sign Language: A Real-time End-to-End Sign Language Recognition System. *IEEE Transactions on Mobile Computing*, 1. 10.1109/tmc.2020.3038303

Yang, K., Xu, M., Yang, X., Yang, R., & Chen, Y. (2021). A Novel EMG-Based Hand Gesture Recognition Framework Based on Multivariate Variational Mode Decomposition. *Sensors (Basel)*, *21*(21), 7002. doi:10.3390/s21217002 PMID:34770309

Zhang, X., Chen, X., Li, Y., Lantz, V., Wang, K., & Yang, J. (2011). A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors. *IEEE Transactions on Systems, Man, and Cybernetics. Part A, Systems and Humans*, *41*(6), 1064–1076. doi:10.1109/TSMCA.2011.2116004

Zhao, J., & Allison, R. S. (2019). Comparing head gesture, hand gesture and gamepad interfaces for answering Yes/No questions in virtual environments. *Virtual Reality (Waltham Cross)*, *24*(3), 515–524. doi:10.1007/s10055-019-00416-7

Zhou, Z., Chen, K., Li, X., Zhang, S., Wu, Y., Zhou, Y., Meng, K., Sun, C., He, Q., Fan, W., Fan, E., Lin, Z., Tan, X., Deng, W., Yang, J., & Chen, J. (2020). Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays. *Nature Electronics*, *3*(9), 571–578. doi:10.1038/s41928-020-0428-6

Kayal Padmanandam is an Associate Professor who has over a decade of credentials in the domain of Computer Science with wide exposure through teaching, research, and industry. She is appointed as a ACM Distinguished Speaker for computer community by ACM. She is passionate about research and specialized in Data Science and Machine Learning Algorithms in which she pursued her doctoral research. She has several publications especially related to machine-learning applications. Her research highly concentrates on supervised learning algorithms, measure optimization techniques, parameter selection & calibration for input-based model building, analytics on large predictor variables for any Intelligent Predictor System.

Swati Sah joined Amity University in in Tashkent as An Associate Professor in Dept of IT and Enggineering in Feb 2021 She has an excellent experience of 10 years in research and teaching. Her interest area of research are Artificial Intelligence, Machine Learning, Project Management etc. She was associated with Bharati Vidyapeeth Deemed University,Pune during 2010 to 2017. She got awarded her PhD degree in Computer Application from Bharati Vidyapeeth University, Pune, India, in 2016. The topic of her PhD research was "Evaluating Optimal Clustering Techniques for Efficient Storagre Retrieval Methods in Large DataBase Using Soft Computing Techniques". She did her MCA from UP Technical University, Lucknow, India. She has also been awarded the Master degree M.P.M from Birmingham City University, United Kingdom U. K. She has also been proactively involved with professional associations and professional bodies. She has attended and organized various FDP/ Seminar/workshop/Conferences. She has published her research papers in National and International journals.