iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79873

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

AI Based Real Time Transcription

I. Eswari M.E.[1], M. Riyasudeen[2], G. Thirisan[3], T. Deepak[4], H. Salim Mohamed Mahadeer[5]
Dept. of Computer Science and Engineering, MIET Engineering College, Tiruchirappalli, India

Abstract: ***Human beings usually rely on communication to express their feeling and ideas and to solve disputes among themselves. A major component required for effective communication is language. Language can occur in different forms, including written symbols, gestures, and vocalizations. It is usually essential for all of the communicating parties to be fully conversant with a common language. However, to date this has not been the case between speech-impaired people who use sign language and people who use spoken languages. A number of different studies have pointed out a significant gaps between these two groups which can limit the ease of communication. Therefore, this study aims to develop an efficient deep learning model that can be used to predict British sign language in an attempt to narrow this communication gap between speech-impaired and non-speech-impaired people in the community. Two models were developed in this research, CNN and LSTM, and their performance was evaluated using a multi-class confusion matrix. The CNN model emerged with the highest performance, attaining training and testing accuracies of 98.8% and 97.4%, respectively. In addition, the model achieved average weighted precession and recall of 97% and 96%, respectively. On the other hand, the LSTM model's performance was quite poor, with the maximum training and testing performance accuracies achieved being 49.4% and 48.7%, respectively. Our research concluded that the CNN model was the best for recognizing and determining British sign language***

Index terms: British Sign Language (BSL), Deep Learning, Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Sign Language Recognition, Real-Time Transcription, Computer Vision, Hand Gesture Recognition, Sign-to-Text Translation, Machine Learning.

I. INTRODUCTION

Sign language is a natural language used by deaf and hard-of-hearing individuals for communication. It involves hand movements, gestures, facial expressions, and body language to convey meaning. Sign languages are fully-fledged languages with their own grammar, syntax, and vocabulary. Sign language has existed for centuries, with documented evidence dating back to ancient civilizations. However, formal recognition and study of sign languages began in the 18th century with educators such as Charles-Michel de l'Épée in France and Thomas Gallaudet in the United States.

These pioneers established schools for the deaf and developed systems for teaching sign language. 3. Usage: Sign language is used primarily by deaf and hard-of-hearing individuals to communicate with each other and with hearing individuals who understand sign language. It is also used by hearing individuals who interact with the deaf community, such as interpreters, family members, and educators. Sign language facilitates communication in various contexts, including education, work, social interactions, and entertainment.

The paper is structured as follows. Section II reviews related work. Section III formally specifies the scope . Section IV describes the system analysis. Section V details the problem statement. Section VI reports and analyses the results. Sections VII through IX address Security Algorithms Used, Software Description, Test Results And Evaluation Metrics. Section X concludes..

II. RELATED WORKS

A. Exploring the Impact of Emotional Voice Integration in Sign-to-Speech Translators for Deaf-to-Hearing Communication

Emotional voice communication plays a crucial role in effective daily interactions. Deaf and Hard of Hearing (DHH) individuals, who often have limited use of voice, rely on facial expressions to supplement sign language and convey emotions. However, in American Sign Language (ASL), facial expressions serve not only emotional purposes but also function as linguistic markers that can alter the meaning of signs. This dual role can often confuse non-signers when interpreting a signer's emotional state. In this paper, we present studies that: (1) confirm the challenges non-signers face when interpreting emotions from facial expressions in ASL communication, and (2) demonstrate how integrating emotional voice into translation systems can enhance hearing individuals' understanding of a signer's emotional intent. An online survey with 45 hearing participants (non-ASL signers) revealed frequent misinterpretations of signers' emotions when emotional and linguistic facial expressions were used simultaneously.

©IJRASET: All Rights are Reserved

B. Effects of Text Release Features on Communication and Problem-Solving in Teams with a Deaf/Hard of Hearing and a Hearing Member Using Automatic Speech Recognition: Effects of Text Release Features

This study examined effects of three text release features used with Automatic Speech Recognition (ASR)-based automatic captioning and of hearing status on quality of performance on problem-solving tasks, number of words two team members exchanged with each other, and on member's judgments regarding the extent the ASR generated text contained errors. Each of 36 teams, consisting of one deaf/hard of hearing (DHH) and one hearing member, worked online to complete three mapping tasks. These mapping tasks required naming of map locations that were marked by numbers along a pathway indicated by a line with numerous turns. For each of three mapping tasks, the hearing member used a different ASR-based text release feature to produce captions for the DHH member to read, either: (a) controlled release; (b) automatic release with no editing; or (c) automatic release with editing. Results suggested that for DHH and hearing members to solve a problem using automatic captioning with ASR synchronously, it is important for the captioning technology to facilitate rapid exchange of words between members and, at the same time, facilitate dealing with the errors produced by the ASR. Consistent with this proposition, problem-solving performance was highest under the automatic release with editing feature.

C. Optimization of English Speech-to-Text Conversion Using Deep Reinforcement Learning

This paper studies the application of deep reinforcement learning (DRL) to the English speech-to-text (STT) task. Aiming at the limitations of traditional STT systems in terms of accuracy and efficiency, we propose an optimization framework based on Deep Q Network (DQN). This framework utilizes the dynamic decision-making capabilities of DRL to improve the accuracy of speech recognition through autonomous learning optimization strategies. The experiment used the publicly available LibriSpeech data set and used word error rate (WER) and sentence error rate (SER) as the main evaluation indicators. Experimental results show that compared with traditional baseline models, our DQN model achieves significant improvements in both WER and SER. Nonetheless, research also reveals challenges with DRL in terms of resource consumption and model robustness. Future work will focus on improving the training efficiency of the model, strategies for adjusting hyperparameters, and exploring ways to improve the model's performance in noisy environments. By comprehensively applying technologies such as semi-supervised learning and transfer learning, it is expected to further improve the performance and applicability of the STT system.

III. SCOPE OF THIS PAPER

The main goal of this research is to develop an efficient deep learning model that can be used to detect, understand, and translate British sign language to written text. Two models were developed, an LSTM model and a CNN model. Their performance was evaluated and the results compared to determine the best model.

Two approaches were applied in developing the models, differing based on the type of data used. The first approach involved importing preprocessed data from Kaggle, a platform where experiment datasets are freely available to the public. The second approach involved collecting data from a computer webcam using a computer vision algorithm and extracting key points, including hands, face, and pose. As a sequence, these key points can then be passed along to detect and decode actions and sign language. To achieve this, the model’s artifact provides an “h5” weight file for a model that is then applied in deploying and testing the model using images of different sign languages and producing a text output on a webcam.

To evaluate the objects of this research, we applied multilabelled classification. This supervised learning prototype involves assigning each data instance several labels from a predefined set of tags. This trending approach is used where the available dataset is too complex for each instance to have a specific class.

IV. SYSTEM ANALYSIS

The paper studies the workflow when a User provides Input, which inherently triggers the Character Prediction. This fundamental prediction process is directly supported by the Application , which provides the necessary backend logic to anticipate the user's keystrokes. Furthermore, Word Formation serves as an extension of character prediction, allowing the system to transition from identifying individual letters to assembling them into coherent, complete words.

Building upon these individual words, the Suggestions incorporates Word Formation to present viable linguistic choices to the user. This process is further sophisticated by Sentence Building, which extends basic suggestions by using contextual data managed again by the Application to construct entire phrases or sentences. The interaction loop concludes when the Sentence Building use case generates the final Output, which is then successfully delivered back to the User.

V. PROBLEM STATEMENT

Effective communication is the cornerstone of human interaction, yet a significant linguistic divide persists between speech-impaired individuals who utilize sign language and the hearing-abled majority who rely on vocalization. While sign languages are fully-fledged linguistic systems with complex grammar and syntax, they are not universally understood by the general public. This creates high communication barriers in critical areas such as public service institutions, healthcare settings, and various industrial sectors, often limiting speech-impaired individuals from accessing essential services or enjoying equal employment opportunities.

Current methods to bridge this gap, such as the use of human sign language interpreters, are often inefficient, costly, and lack the scalability required for real-time, everyday interactions. While advancements in machine learning and deep learning offer promising avenues for automated sign-to-text translation, the field remains in a state of active research. Specifically, it has not yet been definitively established which deep learning architectures—such as Convolutional Neural Networks (CNN) or Long Short-Term Memory (LSTM) networks—provide the optimal balance of accuracy, precision, and real-time processing capability for recognizing the intricate hand gestures and facial expressions inherent in British Sign Language (BSL). Consequently, there is an urgent need to develop and comparatively evaluate deep learning models to identify a robust, real-time transcription solution. Addressing this problem is essential to narrow the communication gap, promote social cohesion, and ensure that individuals with speech impairments can interact seamlessly with society at large.

VI. TEST AND ANALYSIS OF THE RESULTS

A feasibility study in sign language involves assessing the practicality of a project or initiative aimed at enhancing communication and accessibility for Deaf and Hard-of-Hearing individuals. This study evaluates key factors such as the needs of the target audience, the available technologies for sign language recognition or translation, and the potential impact on the community. It includes a market analysis to determine demand and competition, as well as an evaluation of the financial viability, considering budgets, funding sources, and potential revenue models.

A. Technical Feasibility

Technical feasibility in sign language involves assessing whether the necessary technologies and resources are available to develop and implement effective solutions for sign language communication and accessibility. This includes evaluating the potential of artificial intelligence and machine learning for accurate sign language recognition and translation, ensuring compatibility across various devices and platforms, and designing user interfaces that cater specifically to the Deaf and Hard-of-Hearing community. It also involves considering the infrastructure required for real-time processing, data storage, and efficient communication, as well as the availability of specialized hardware like cameras or sensors for capturing gestures.

B. Economic Feasibility

Economic feasibility in sign language evaluates whether a sign language-related project or initiative is financially viable and sustainable. This includes assessing the initial development costs, ongoing maintenance, and potential revenue streams. It involves determining the financial resources required for research, design, technology development and platform deployment. The study also examines potential funding sources, such as grants, donations, or partnerships with organizations serving the Deaf community, as well as considering the affordability of the technology for users.

C. Operational Feasibility

It examines whether the necessary infrastructure, personnel, and processes are in place to support the successful operation of the project. This includes evaluating the availability of skilled staff, such as sign language interpreters, educators, and technical experts, as well as the training needed for ongoing operations. Operational feasibility also involves considering the practicality of deploying the technology or service, ensuring that it aligns with the needs and preferences of the target audience, such as Deaf and Hard-of-Hearing individuals.

D. Educational Feasibility

The focuses on assessing whether a project or program aimed at teaching or promoting sign language can be effectively developed and implemented within educational settings. It involves evaluating the availability of appropriate curriculum materials, resources, and qualified instructors who can effectively teach sign language to diverse learners, including both Deaf and hearing individuals. This feasibility analysis also considers the adaptability of the program to different learning environments, whether in-person or online, and ensures that the content is engaging, accessible, and culturally sensitive.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

VII. SECURITY ALGORITHMS USED

To convey "model building" in sign language, you can use signs and gestures to represent the concept. Here's a suggested way to express "model building" in American Sign Language (ASL):

Model: Hold both hands in front of your body, palms facing each other, and use your dominant hand to form the shape of a small object or figure, indicating the concept of a model.

Building: Use both hands to mimic the action of building or constructing something, as if stacking blocks or assembling parts together. You can use upward movements to represent the act of building upwards or adding layers.

Combination: Combine the signs for "model" and "building" sequentially to convey the concept of model building. You can sign "model" first to establish the concept of a representation or prototype, followed by "building" to indicate the action of constructing or creating it. As with any sign language expression, it's important to consider regional variations and context when conveying specific concepts. Consulting with fluent signers or resources specific to the sign language you are using can help ensure accurate communication

VIII. SOFTWARE DESCRIPTION

A. Python

Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language. It was created by Guido van Rossum during 1985-1990. Like Perl, Python source code is also available under the GNU General Public License (GPL). This tutorial gives enough understanding on Python programming language. Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages. Python is a MUST for students and working professionals to become a great Software Engineer specially when they are working in Web Development Domain. Python is currently the most widely used multi-purpose, high-level programming language. Python allows programming in Object-Oriented and Procedural paradigms. Python programs generally are smaller than other programming languages like Java. Programmers have to type relatively less and indentation requirement of the language, makes them readable all the time. Python language is being used by almost all tech-giant companies like – Google, Amazon, Facebook, Instagram, Dropbox, Uber… etc. The biggest strength of Python is huge collection of standard library which can be used for the following: Machine Learning GUI Applications (like Kivy, Tkinter, PyQt etc.) Web frameworks like Django (used by YouTube, Instagram, Dropbox) Image processing (like OpenCV, Pillow) Web scraping (like Scrapy, Beautiful Soup, Selenium) Test frameworks Multimedia Scientific computing 26 Pandas pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.

B. Pandas

pandas is a Python package that provides fast, flexible, and expressive data structures designed to make working with "relational" or "labeled" data both easy and intuitive. It aims to be the fundamental high-level building block for doing practical, real world data analysis in Python. Pandas is mainly used for data analysis and associated manipulation of tabular data in Data frames. Pandas allows importing data from various file formats such as comma-separated values, JSON, Parquet, SQL database tables or queries, and Microsoft Excel. Pandas allows various data manipulation operations such as merging, reshaping, selecting, as well as data cleaning, and data wrangling features. The development of pandas introduced into Python many comparable features of working with Data frames that were established in the R programming language. The panda’s library is built upon another library NumPy, which is oriented to efficiently working with arrays instead of the features of working on Data frames.

C. Numpy

NumPy, which stands for Numerical Python, is a library consisting of multidimensional array objects and a collection of routines for processing those arrays. Using NumPy, mathematical and logical operations on arrays can be performed. NumPy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays.

D. Matplotlib

Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. Matplotlib makes easy things easy and hard things possible. Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

E. Seaborn

Seaborn is a library for making statistical graphics in Python. It builds on top of matplotlib and integrates closely with pandas data structures. Visualization is the central part of Seaborn which helps in exploration and understanding of data. Seaborn offers the following functionalities: Dataset oriented API to determine the relationship between variables.Automatic estimation and plotting of linear regression plots.It supports high-level abstractions for multi-plot grids. Visualizing univariate and bivariate distribution.

F. Scikit-Learn

scikit-learn is a Python module for machine learning built on top of SciPy and is distributed under the 3-Clause BSD license. 28 Scikit-learn (formerly scikits. learn and also known as sklearn) is a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support-vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

G. MySQL

MySQL tutorial provides basic and advanced concepts of MySQL. Our MySQL tutorial is designed for beginners and professionals. MySQL is a relational database management system based on the Structured Query Language, which is the popular language for accessing and managing the records in the database. MySQL is open-source and free software under the GNU license. It is supported by Oracle Company. MySQL database that provides for how to manage database and to manipulate data with the help of various SQL queries. These queries are: insert records, update records, delete records, select records, create tables, drop tables, etc. There are also given MySQL interview questions to help you better understand the MySQL database. MySQL is currently the most popular database management system software used for managing the relational database. It is open-source database software, which is supported by Oracle Company. It is fast, scalable, and easy to use database management system in comparison with Microsoft SQL Server and Oracle Database. It is commonly used in conjunction with PHP scripts for creating powerful and dynamic server-side or web-based enterprise applications. It is developed, marketed, and supported by MySQL AB, a Swedish company, and written in C programming language and C++ programming language. The official pronunciation of MySQL is not the My Sequel; it is My Ess Que Ell. However, you can pronounce it in your way. Many small and big companies use MySQL. MySQL supports many Operating Systems like Windows, Linux, MacOS, etc. with C, C++, and Java languages.

H. Wampserver

WampServer is a Windows web development environment. It allows you to create web applications with Apache2, PHP and a MySQL database. Alongside, PhpMyAdmin allows you to manage easily your database. WAMPServer is a reliable web development software program that lets you create web apps with MYSQL database and PHP Apache2. With an intuitive interface, the application features numerous functionalities and makes it the preferred choice of developers from around the world. The software is free to use and doesn’t require a payment or subscription.

I. Bootstrap 4

Bootstrap is a free and open-source tool collection for creating responsive websites and web applications. It is the most popular HTML, CSS, and JavaScript framework for developing responsive, mobile-first websites. It solves many problems which we had once, one of which is the cross-browser compatibility issue. Nowadays, the websites are perfect for all the browsers (IE, Firefox, and Chrome) and for all sizes of screens (Desktop, Tablets, Phablets, and Phones). All thanks to Bootstrap developers -Mark Otto and Jacob Thornton of Twitter, though it was later declared to be an open-source project. 30 Easy to use: Anybody with just basic knowledge of HTML and CSS can start using Bootstrap Responsive features: Bootstrap's responsive CSS adjusts to phones, tablets, and desktops Mobile-first approach: In Bootstrap, mobile-first styles are part of the core framework Browser compatibility: Bootstrap 4 is compatible with all modern browsers (Chrome, Firefox, Internet Explorer 10+, Edge, Safari, and Opera)

J. Flask

Flask is a web framework. This means flask provides you with tools, libraries and technologies that allow you to build a web application. This web application can be some web pages, a blog, a wiki or go as big as a web-based calendar application or a commercial website. Flask is often referred to as a micro framework. It aims to keep the core of an application simple yet extensible. Flask does not have built-in abstraction layer for database handling, nor does it have formed a validation support. Instead, Flask

supports the extensions to add such functionality to the application. Although Flask is rather young compared to most Python frameworks, it holds a great promise and has already gained popularity among Python web developers. Let's take a closer look into Flask, so-called "micro" framework for Python. Flask is part of the categories of the micro-framework. Micro-framework are normally framework with little to no dependencies on external libraries. This has pros and cons. Pros would be that the framework is light, there are little dependence to update and watch for security bugs, cons is that some time you will have to do more work by yourself or increase yourself the list of dependencies by adding plugins. In the case of Flask, its dependencies are: WSGI-Web Server Gateway Interface (WSGI) has been adopted as a standard for Python web application development. WSGI is a specification for a universal interface between the web server and the web applications

IX. FUTURE ENHANCEMENT

This paper points to the following recommendations. A live-trained CNN model can be developed based on insights from this research paper. This is because the CNN model has proven to perform better than LSTM based on this study. Such a model should convert, detect, interpret, or predict sign language from video and display it through a special screen. It should be able to pre-process data automatically. Sign language recognition and prediction should be effectively instantaneous. The model should be fitted with high-definition digital cameras to capture high-quality videos that the CNN model can process. Developers and sponsors can adopt the finding of this research on convolutional neural networks and modify it for the benefit of society at large. This could include building a more comprehensive model that accommodates the entire set of British Sign Language and using it to help speech-impaired people to more easily communicate in public offices or facilities. A further recommendation is that future research develop a CNN model that accommodates different sign languages and compare its effectiveness. Future research could additionally try to improve the performance of LSTM by combining it with other models and then comparing the performance of the LSTM composite models with CNNs

X. CONCLUSION

The primary approach of this research was to develop two deep learning models, a long short-term memory (LSTM) model and a convolutional neural network (CNN) model and compare their performance. The experiment involved collecting the required datasets, using them to develop the models, training and testing the two models, and applying a multi-class confusion matrix to evaluate their performance. The parameters used for the comparison included training and testing accuracy and the systems' respective precision and reliability/consistency in predicting sign language. The approach was then divided into two categories; the first used pre-processed data to predict hand gestures for British numerical sign language, and the second used a key points dataset to indicate simple common messages (facial expressions combined with pose signs). In the first approach, both the CNN and LSTM models were developed.

The CNN model showed the best performance in all aspects, including accuracy, precision, and reliability, as stated in the research hypothesis. Furthermore, this model showed a positive correlation between training/testing accuracy and the length of the training period as determined by the number of iterations and images per dataset. This resulted in the CNN model attaining high accuracy. A more significant number of signs and more iterations could be applied to increase the training and testing accuracy; the model would then be applicable for accommodating more than one type of sign language, making it more efficient. On the other hand, the LSTM model showed very poor performance in both categories of the experimental approach. This model attained very low accuracy, precision, and consistency in predicting the correct sign based on the multi-class confusion matrix.

The most reasonable explanation for the poor performance of the LSTM was due to certain limitations that were pointed out in the literature review. For instance, this model can be difficult to train, as it requires a memory-bandwidth-bound computation which has hardware limitations. LSTM models depend on more complex frameworks to achieve good performance compared to CNN model. This research found that while LSTM models are better in classification of text data, for image data sets more input parameters may be needed. The performance of the LSTM model could be improved by integrating it with other models to curb its limitations. It was not possible for a CNN model to be developed in the second approach, as the dataset used was incompatible with the requirements of a CNN. However, based on the results of the first 34 approach as well as on the literature review, it can be assumed that the CNN model would have performed well in this second approach.

Therefore, this research concludes that convolutional neural networks perform better in recognizing and predicting British sign language than LSTM models. In addition, this research further concludes that the CNN model could be used to accommodate more than one set/type of sign language recognition prediction. The findings of this research answer the question of which deep learning models perform better in attempting to narrow the gap between speech impaired people and the general public.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

XI. ACKNOWLEDGMENT

The authors express their gratitude to the Department of Computer Science and Engineering, MIET Engineering College, Tiruchirappalli, for making computational resources and institutional support available for this research. The authors also thank Dr. I. Shahanaz Begum M.E., Ph.D., for her mentorship and encouragement throughout the development of this work.

REFERENCES

[1] Official P. Barve, N. Mutha, A. Kulkarni, Y. Nigudkar, Y. Robert, Application of deep learning techniques on sign language recognition—a survey, in Data Management, Analytics and Innovation ed. by N. Sharma, A. Chakrabarti, V.E. Balas, A.M. Bruckstein. Lecture Notes on Data Engineering and Communications Technologies, vol. 70 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-16-2934-1_14

[2] A. Patil, A. Kulkarni, H. Yesane, M. Sadani, P. Satav, Literature survey: sign languagerecognition using gesture recognition and natural language processing, in Data Management, Analytics and Innovation, ed. by N. Sharma, A. Chakrabarti, V.E. Balas, A.M. Bruckstein. Lecture Notes on Data Engineering and Communications Technologies, vol. 70 (Springer, Singapore, 2021). https://doi.org/10.1007/978-981-16-2934-1_13

[3] Y. Robert, Y. Nigudkar, A. Kulkarni, N. Mutha, P. Barve, Literature survey: application of machine learning techniques on static sign language recognition, in Innovations in Bio-inspired Computing and Applications. IBICA 2020, ed. by A. Abraham, H. Sasaki, R. Rios, N. Gandhi, U. Singh, K. Ma. Advances in Intelligent Systems and Computing,vol. 1372 (Springer, Cham, 2021). https://doi.org/10.1007/978-3-030-73603-3_16

[4] M. Taskiran, M. Killioglu, N. Kahraman, A real-time system for recognition of American sign language by using deep learning, in 2018 41st International Conference on Telecommunications and Signal Processing (TSP) (2018), pp. 1–5. https://doi.org/10.1109/TSP.2018.8441304

[5] M.R. Islam, U.K. Mitu, R.A. Bhuiyan, J. Shin, Hand gesture feature extraction using deep convolutional neural network for recognizing American sign language, in 2018 4th International Conference on Frontiers of Signal Processing (ICFSP) (2018), pp.115–119. https://doi.org/10.1109/ICFSP.2018.8552044

[6] S. Stoll, N.C. Camgoz, S. Hadfield, R. Bowden, Text2Sign: towards sign language production using neural machine translation and generative adversarial networks. Int. J. Comput. Vision **128**(4), 891–908 (2020) 60

[7] R.D. Raj, A. Jasuja, British sign language recognition using HOG, in 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) (2018), pp. 1–4. https://doi.org/10.1109/SCEECS.2018.8546967

[8] J. Zamora-Mora, M. Chacón-Rivas, Real-time hand detection using convolutional neural networks for Costa Rican sign language recognition, in 2019 International Conference on Inclusive Technologies and Education (CONTIE), 180–1806 (2019). https://doi.org/10.1109/CONTIE49246.2019.00042

[9] C.O. Sosa-Jiménez, H.V. Ríos-Figueroa, E.J. Rechy-Ramírez, A. Marin-Hernandez, A.L.S. González-Cosío, Real-time Mexican sign language recognition, in 2017 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC) (2017), pp. 1–6

[10] Sugandhi, P. Kumar, S. Kaur, Indian sign language generation system. Computer **54**(3), 37–46 (2021). https://doi.org/10.1109/MC.2020.2992237

[11] R. San-Segundo, R. Barra, R. Córdoba, L.F. D'Haro, F. Fernández, J. Ferreiros, J.M. Lucas, J. Macías-Guarasa, J.M. Montero, J.M. Pardo, Speech to sign language translation system for Spanish. Speech Commun. **50**(11–12), 1009–1020 (2008)

[12] B. Saunders, N.C. Camgoz, R. Bowden, Everybody sign now: translating spoken language to photorealistic sign language video (2020). ArXiv preprint arXiv:2011.09846

[13] N. Vasani, P. Autee, S. Kalyani, R. Karani, Generation of Indian sign language by sentence processing and generative adversarial networks, in 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) (2020), pp. 1250– 1255. https://doi.org/10.1109/ICISS49785.2020.9315979

[14] P. Kapoor, R. Mukhopadhyay, S.B. Hegde, V. Namboodiri, C.V. Jawahar, Towards automatic speech to sign language Generation. ArXiv preprint arXiv:2106.12790 (2021)

[15] A.C. Duarte, Cross-modal neural sign language translation, in Proceedings of the 27th ACM International Conference on Multimedia (2019), pp. 1650–1654 61

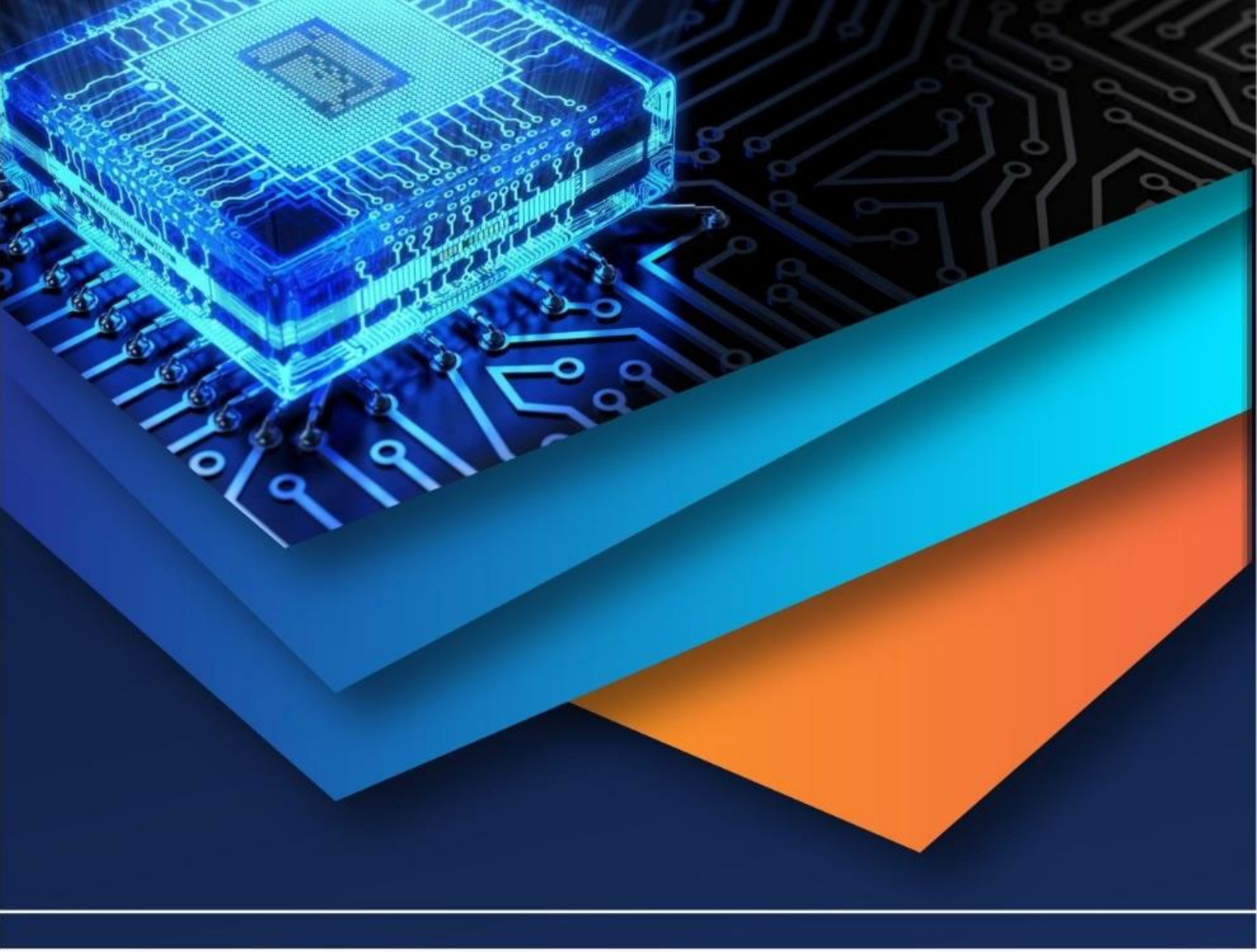
[16] T. Veale, A. Conway, Cross modal comprehension in ZARDOZ, an English to signlanguage translation system, in 4th International Workshop on Natural Language Generation (1994), pp. 67–72

[17] B.D. Patel, H.B. Patel, M.A. Khanvilkar, N.R. Patel, T. Akilan, ES2ISL: an advancement in speech to sign language translation using 3D avatar animator, in 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) (IEEE, 2020), pp. 1–5

[18] N. Mehta, S. Pai, S. Singh, Automated 3D sign language caption generation for video. Univ. Access Inf. Soc. **19**(4), 725–738 (2020)

[19] Sugandhi, P. Kumar, S. Kaur, Sign language generation system based on Indian sign language grammar. ACM Trans. Asian Low-Resour. Lang. Inf. Process. **19**(4), Article 54, 26 (2020). https://doi.org/10.1145/3384202

[20] T. Dasgupta, A. Basu, Prototype machine translation system from text-to-Indian sign language, in Proceedings of the 13th International Conference on Intelligent User Interfaces (2008), pp. 313–316.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)