JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

SIGN LANGUAGE BASED INFORMATION RETRIEVAL SYSTEM

[1]Sanjay KP, [2]Mithil Gunaga, [3]Manoj M, [4]Chinmayananda R M, [5]Prof. Sinchana
1,2,3,4 Student, 5 Assistant Professor.
1,2,3,4,5 Computer Science and Engineering.
1.2.3.4.5.6, Srinivas Institute of Technology, Mangalore, India

Abstract:People who are deaf and also hearing often struggle to interact with mainstream information retrieval systems, which typically require text input or spoken language. This research introduces an inclusive digital solutions that helps users to perform searches using sign language gestures. The system combines the both computer vision, deep learning, and natural language processing to accurately interpret visual signs. A convolutional neural network (CNN) processes spatial features from the video feed, while a recurrent neural network (RNN) captures the temporal patterns of motion. Recognized hand gestures or signs are then translated into specific commands or search queries, which the system uses to access relevant information from either online APIs or internal databases. A user-friendly interface ensures accessibility on both desktop and mobile platforms. Testing demonstrates that the system effectively interprets common signs and delivers accurate results, offering improved access for users who rely on sign language.

Index Terms – *Sign language recognition; information retrieval; convolutional neural network; accessibility; human–computer interaction; gesture recognition.*

I. INTRODUCTION

Although digital advancements have improved accessibility for many, individuals who rely on sign language remain underserved by platforms centered on textual or spoken interaction. Traditional search engines and data retrieval platforms typically do not accommodate gesture-based communication, creating a difference in usability. To bridge this divide, we propose the system that leverages sign language for querying information, integrating advanced techniques in advanced computer vision, natural language processing (NLP), and machine learning. The platform captures hand movements via a camera, converts these gestures into text commands, and executes searches across databases or web-based sources. By offering a seamless interface tailored to sign language users, the system promotes digital inclusivity and empowers the deaf, people who are not able to ear and hard-of-hearing community to engage more effectively with online content. The sections that follow review related literature, detail the proposed system's architecture, evaluate experimental outcomes, and explore avenues for further development.

II. LITERATURE REVIEW

Early systems for recognizing the sign language were often limited in scope, typically supporting only the small set of gestures and relying on basic recognition techniques. These models frequently utilized static gesture libraries and rudimentary pattern-matching algorithms, which struggled to interpret gestures accurately—especially those influenced by context or subtle variations. Moreover, many of these systems were unable to support advanced functions or deliver dynamic responses, limiting their practicality in real-world scenarios. They also lacked important features such as instant feedback and also the adaptability, which are essential for accessible user experiences.
With the evolution of deep learning, modern systems have become significantly more accurate and flexible. CNNs are now widely used for capturing visual features from gesture images, while RNNs are effective in analyzing gesture sequences. Studies by researchers like Koller and Huang have demonstrated the robustness of these models in handling large-scale gesture datasets. However, most existing solutions focus solely on recognition or translation, with little emphasis on integrating search functionality. This study addresses that shortcoming by linking gesture interpretation directly with an information retrieval engine.

III. METHODOLOGY

The proposed system is composed of four main components: a gesture capture unit (using live video input), a recognition engine (based on deep learning), a query interpretation module, and an information retrieval backend. The process begins with real-time video capture through a standard webcam. Preprocessing some methods, such as resizing and normalization, are applied to create the image frames for analysis.

At the heart of the system lies a hybrid deep learning model combining a convolutional neural networks (CNN) and a recurrentt neural network (RNN). The CNN extracts the features from each frame by applying multiple layers of convolution and pooling. These features are flattened and processed by dense layers that generate initial gesture classifications. Subsequently, the output is passed to an RNN—such as an LSTM layer—to account for temporal dynamics and improve recognition accuracy over sequences of frames.

Once a gesture is successfully recognized, it is mapped to a predefined search intent or command. For instance, a specific sign like "weather" triggers a query related to weather forecasts. The system then forwards the query to an internal database or external API, depending on the context, and retrieves relevant information. A user interface is included, providing features such as a text input option for manual queries and a live video panel for gesture feedback, thereby enhancing usability and accessibility.

Fig 1 shows the architectural diagram depicts a gesture classification process. It involves input by the user, preprocessing via a CNN, prediction.

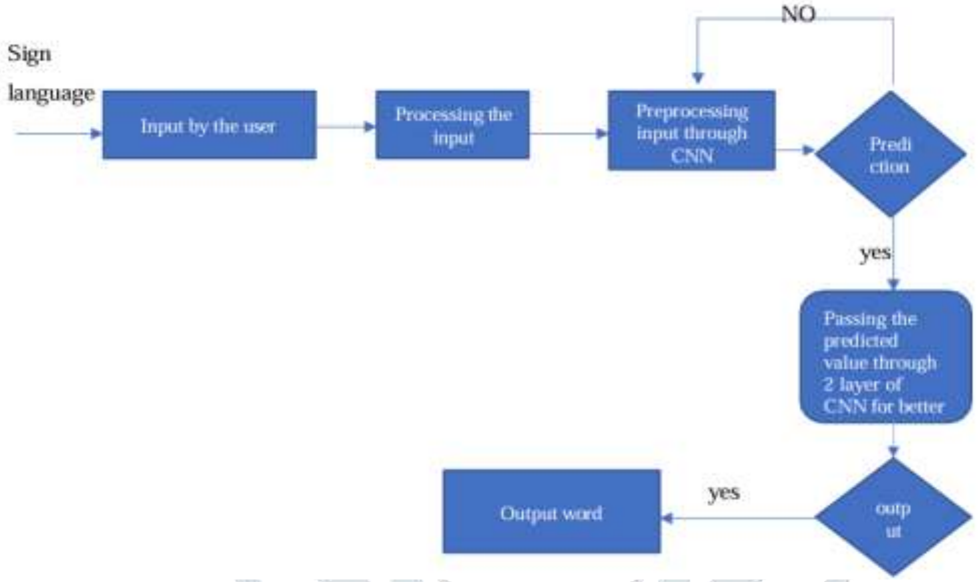

Fig.1: The proposed system

The architectural diagram illustrates the flow of information from gesture input to the retrieval of relevant data. It begins with the user capturing a sign gesture through a webcam, which is then processed and classified using a deep learning model. This classified gesture is mapped to a corresponding query via a natural language processor. The mapped query is sent to the retrieval engine, which fetches results from the database or web source. Finally, the system presents the information to the user through an accessible interface.

Fig 2 Represents the CNN architecture diagram illustrates the progression of an input image through convolution, max pooling.

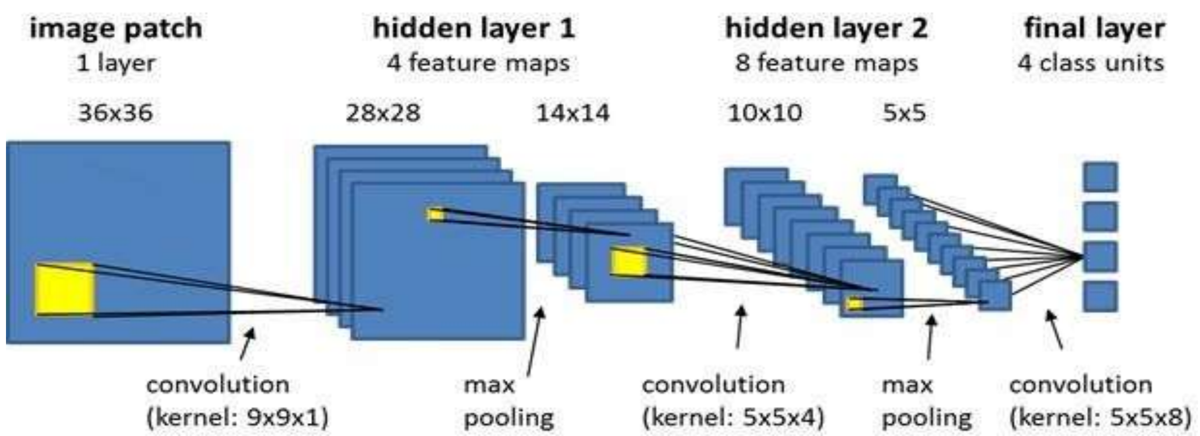

Fig.2: CNN Construction Diagram

This diagram illustrates the structure of the CNN used for gesture recognition. Input frames are processed through the multiple convolutional layers those are detect hand shapes and spatial features. Max pooling layers reduce the feature map size while retaining significant patterns. The resulting data is then flattened and passed through dense layers. Activation functions such as ReLU are applied to optimize learning, and the final output layer assigns a label to the gesture.

Fig 3 describes the depicting the interaction flow between the user interface, processing modules, query processor, and database to retrieve information.

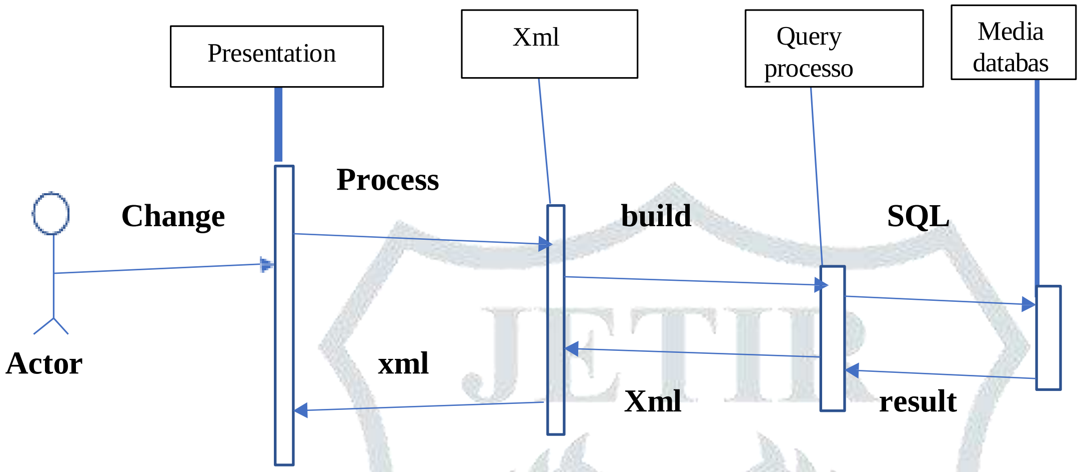

Fig 3: Sequence diagram proposed system

The sequence diagram represents the interaction flow between the user, system modules, and the database. It begins with the user performing a gesture, which is captured and processed by the vision module. The gesture is then classified using a trained model, and the result is passed to the NLP component to generate a relevant query. This query is sent to the information retrieval module, which accesses the database and returns the appropriate response. Finally, the system displays the output to the users in real-time env.

IV. RESULTS AND DISCUSSION

The system was implemented using Python with TensorFlow for gesture recognition, and a lightweight web interface for user interaction. During initial testing, frequently used gestures such as “Thank you,” “Sorry,” and “See you” were successfully identified by the system. Each recognized gesture triggered a relevant textual query that was used to fetch corresponding information—such as definitions or search results—from connected sources.

The recognition engine achieved processing speeds of approximately 10–15 frames per second on a standard laptop configuration, offering smooth real-time feedback. The user interface was found to be intuitive and accessible, particularly for first-time users. These early results indicate both strong classification accuracy and potential for real-world application, especially in environments where conventional input methods are not that practical.

V. CONCLUSION AND FUTURE WORK

This project shows a working sign language-based information retrieval system designed to address communication barriers for deaf and hard-of-hearing individuals. By combining real-time gesture recognition with an automated query system, the platform enhances accessibility to digital resources across sectors such as education, healthcare, and beyond for sign language users. Future developments will aim to expand the system’s scope and precision, including the incorporation of regional sign languages and real-time speech translation in multiple languages

REFERENCES

[1] Koller, O. et al., “Real-Time Sign Language Recognition Using CNNs,” 2016.
[2] Pigou, L. et al., “Sign Language Recognition with Recurrent Neural Networks,” 2014.
[3] Huang, J. et al., “A Deep Learning Approach to Sign Language Recognition,” 2015.
[4] Ghotkar, A. S., “Hand Gesture Recognition System Using CNN,” 2013.
[5] Manning, C. D. et al., “An Introduction to Information Retrieval,” 2008.
[6] **Moryossef, A., Goldberg, Y., & Oron, S. (2020).**
"Real-Time Sign Language Translation Using Human Pose Estimation."
In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP). [Link: https://aclanthology.org/2020.emnlp-main.673]
[7] **Pigou, L., Dieleman, S., Kindermans, P. J., & Schrauwen, B. (2015).**
"Sign Language Recognition Using Convolutional Neural Networks."
In: European Conference on Computer Vision (ECCV).
A foundational work on using CNNs for isolated sign recognition.
[8] **Zhou, B., & Yuan, X. (2021).**
"Vision-Based Continuous Sign Language Recognition Using Transformer Networks."
In: IEEE Transactions on Multimedia.
Describes an attention-based approach, useful for referencing novel architectures beyond CNN/RNN.
[9] **Sincan, M., & Keles, H. Y. (2020).**
"Autonomous Sign Language Recognition System Using Neural Networks."

In: Journal of Visual Communication and Image Representation.
Demonstrates a full autonomous system using vision and deep learning.

[10] **Camgoz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2018).**
"Neural Sign Language Translation."
In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Covers continuous sign language recognition and translation using sequence-to-sequence models.