https://doi.org/10.48047/AFJBS.6.7.2024.3198-3225

AFJBS African Journal of Biological Sciences
Journal homepage: http://www.afjbs.com
Research Paper Open Access

AI & MACHINE LEARNING TO CONVERT SIGN LANGUAGE TO ENGLISH SCRIPT

Bhausaheb Khamat[1], Associate Prof. Dr. Mrunal Bewoor[2], Associate Prof. Sonali Mali[3], Assistant Prof. Sheetal Patil[4]

[1]M.Tech Student, Computer Engineering, Bharati Vidyapeeth (Deemed To Be University) College of Engineering Pune, (India)
[2]Associate Professor, Computer Engineering Department, Bharati Vidyapeeth (Deemed To Be University) College of Engineering Pune, (India)
[3]Associate Professor, IT Department, Bharati Vidyapeeth (Deemed To Be University) College of Engineering Pune, (India)
[4]Assistant Professor, Computer Engineering Department, Bharati Vidyapeeth (Deemed To Be University) College of Engineering Pune, (India)

[1]khamat.bn@gmail.com, [2]msbewoor@bvucoep.edu.in, [3]sdmali@bvucoep.edu.in, [4]sspatil@bvucoep.edu.in

Volume6issue72024

Received:15May2024

Accepted:10June2024

doi:10.48047/AFJBS.6.7.

2024. 3198-3225

Abstract:

Detecting hand movements is a crucial part of translating sign language, which helps deaf people communicate. However, research on Sign Language faces several challenges. Sign Language lacks appropriate datasets, and it has issues like obscured hand gestures and variations in how people use the language in different regions. These obstacles have hindered progress in sign language research.

In our study, we proposed a solution for detecting and recognizing sign language using deep learning technology, specifically a Convolutional Neural Network (CNN) called RESNET100. We used this network to identify hand movements. In each layer of the network, we extracted and selected important features, and we optimized the network's performance in the pooling layer. To classify hand gestures in real-time, we used various activation functions in the dense layers and tested our system on real-time hand gesture data and the MNIST dataset.

Our experimental analysis showed that our system achieved a significantly higher accuracy, about 5-7% better, compared to the best existing methods in the field. This means our approach has the potential to improve sign language translation and communication for deaf individuals.

1. INTRODUCTION

Sign language is an important communication tool for people who are deaf or hard of hearing. It enables them to express themselves, convey thoughts, and interact with others effectively. Hand movements and gestures are employed by individuals using sign language to articulate their thoughts and emotions. Nonetheless, comprehension of these signs is not widespread among the general population. Hence, there is often a requirement for proficient sign language interpreters during crucial engagements such as medical or legal consultations, as well as educational and training sessions. This demand for sign language interpretation services has witnessed a notable increase in recent years.

Some new services, like using video calls and the internet to connect with sign language interpreters, have been introduced to make it easier to access sign language interpretation. But these services have limitations.

To solve these problems, we developed a computer model that uses a convolutional neural network (CNN) specifically designed to recognize gestures. The model has multiple layers, including four convolutional layers, multiple max-pooling layers, two thickness layers, one smoothing layer, and one missing layer. We carefully train the model using the MNIST American Sign Language dataset, which includes many sign languages.

Additionally, we developed a tailored Convolutional Neural Network (CNN) model to identify signs within video frames employing OpenCV. The precision in detecting hand movements and facial expressions holds paramount importance for ensuring reliable sign recognition. Enhancing methods for sign language recognition stands as a significant endeavor towards improving sign language interpretation and fostering effective communication.

With advancements in technology and artificial intelligence, the prospect of creating sign language applications that can substantially aid individuals with hearing impairments in their day-to-day activities is now within closer reach.

Previous studies have mostly looked at specific aspects of sign language technology, like recognizing signs using video or sensors, and developing software for translating signs. Recent reviews have discussed systems that use sign language to help people who can hear in education, as well as systems for understanding and transcribing sign language into voice or text.

However, there hasn't been a comprehensive overview of all the different aspects of sign language technology and how they connect with each other. This study aims to fill this gap by highlighting the advancements in artificial intelligence (AI) across all aspects of sign language technology, from capturing signs, understanding their meaning, translating them, recognizing signs, and exploring new applications that can improve communication between people who can hear and those who use sign language.

The primary objective of this review is to underscore the significance of integrating AI technology into sign language applications, facilitating seamless communication between both deaf and hearing individuals. Moreover, this study seeks to enlighten researchers regarding the current state-of-the-art in sign language technology and to propose potential research avenues that could yield more precise techniques and enhanced tools for the Deaf community. Specific objectives of the study encompass:

1) Providing a comprehensive review of how AI technologies are applied in different sign language tasks, such as capturing signs, identifying them, interpreting their meaning, and reproducing them, and explaining their importance in this field.

2) Exploring the benefits and drawbacks of contemporary sign language technologies unveils a nuanced understanding of their impact and dynamics. Delving into their advantages and disadvantages while examining their interconnections sheds light on the complexities shaping communication accessibility for individuals with hearing impairments.

3) Suggesting future directions for integrating advanced technology to benefit sign language researchers and academics.

2. LITERATURE SURVEY

Paper title with year: In 2020, a research study investigated the recognition of American Sign Language alphabets by employing a fusion of hand-crafted features alongside deep learning techniques. [1]

Summary: The researchers employed a hybrid methodology, amalgamating traditional image processing techniques with a deep learning framework to effectively differentiate sign alphabets. This experimental strategy entailed the utilization of handcrafted features such as Local Binary Patterns (LBP) in conjunction with the VGG-19 deep learning architecture. Furthermore, YCbCr segmentation techniques were integrated to accurately isolate hand gestures from images, thereby minimizing errors arising from complex backgrounds and variations in lighting conditions. This synergistic approach leverages the complementary strengths of both traditional and deep learning methodologies, enhancing the overall efficacy of sign alphabet recognition.

Problem discusses: To enhance the accuracy of signal categorization, a hybrid approach blending hand-crafted features with a deep learning algorithm is adopted. This strategy involves employing skin color-based YCbCr segmentation for precise shape segmentation, enabling the capture of texture characteristics and local shape information via local binary patterns. The VGG-19 transfer learning framework is then fine-tuned to extract features from the segmented images. These extracted features are fused with hand-crafted features using a serial-based fusion technique. Finally, the combined characteristics are fed into an SVM classifier for signal classification. This comprehensive approach leverages the strengths of both hand-crafted and deep learning features, thereby optimizing the accuracy and robustness of the signal categorization process.

Solution suggested or implemented: To facilitate image segmentation, a skin color-based algorithm was implemented, followed by feeding the segmented images into the proposed deep learning model. The segmentation process relied on the YCbCr color space, which effectively identified and isolated distinct hand regions within the images. Recognizing skin color serves as a crucial aspect of human action recognition, particularly in sign language interpretation through hand gestures. Typically, two primary approaches, pixel-based and region-based, are employed for skin identification. The pixel-based method evaluates each pixel individually to ascertain its association with a skin component, while the region-based method organizes the spatial arrangement of skin pixels to enhance accuracy and performance.

In this study, the deep learning model of choice was VGG-19, renowned for its architecture comprising 19 layers, encompassing convolutional, pooling, and fully connected layers. Alongside deep learning techniques, handcrafted features were also extracted using various methods leveraging image characteristics. One prominent technique utilized in this context is the Local Binary Pattern (LBP), renowned for capturing local texture features within the image, thereby complementing the capabilities of the deep learning model.

Paper title with year: In 2020, a sign language recognition system was developed specifically for static signs, employing advanced deep learning techniques. [2]

Summary: This article presents a method for recognizing Indian Sign Language (ISL) numbers, letters, and sentences in a real environment. The proposed method combines various features in convolutional neural network (CNN) architecture, including convolution layer, rectified linear unit (ReLU) optimization, and maximum pooling layer.

One notable aspect of the methodology is the utilization of convolutional layers with varying filtering window widths. This approach aims to enhance both the speed and accuracy of identification by allowing the network to capture features at different spatial scales. By employing filters of different sizes, the CNN can effectively extract hierarchical representations from the input data, enabling more comprehensive feature learning.

Furthermore, the choice of activation function, specifically ReLU, contributes to the non-linearity of the network, enabling it to model complex relationships within the data. Additionally, max-pooling layers are employed to downsample the feature maps, reducing their spatial dimensions while retaining essential information, thus aiding in feature extraction and dimensionality reduction.

Evaluate the effectiveness of the system that we are proposed, various optimizers were employed during training. The results of the evaluation revealed that the Stochastic Gradient Descent (SGD) optimizer outperformed other optimization algorithms such as Adam and RMSProp in terms of both training and validation outcomes. This finding underscores the significance of optimizer selection in achieving optimal performance and convergence during the training process.

Overall, the proposed methodology demonstrates promising results in the recognition of ISL numerals, alphabets, and sentences in real-world scenarios. By leveraging the capabilities of CNNs and optimizing key components such as filtering window widths and activation functions, the system achieves notable accuracy and efficiency in sign language recognition tasks.

Problem discusses: The main goal of this research is to use deep learning-based convolutional neural networks (CNN) to improve static modeling in sign language recognition. In order to conduct this research, a comprehensive dataset consisting of 35,000 sign photos portraying 100 distinct static signs was meticulously curated from a diverse range of users. This comprehensive information forms the basis for training and evaluation of the proposed system and provides a solid basis for evaluating its performance.

To thoroughly evaluate the efficacy of the proposed system, a substantial array of approximately 50 CNN models was employed. This extensive evaluation process allows for a comprehensive analysis of the various model architectures, configurations, and parameters, thereby enabling the identification of the most optimal solution for sign language recognition tasks.

Sign language, characterized by its complexity and richness, presents unique challenges that necessitate the application of advanced computer vision techniques. The ability to decipher intricate hand gestures and subtle facial expressions is paramount in accurately interpreting and understanding sign-language. Sign language, an important form of communication for the hearing impaired, allows interaction with various gestures. It involves teaching letters, single words and entire sentences.

This research attempts to bridge the gap between static characters and their text or word equivalents by using state-of-the-art deep learning techniques such as CNNs. By creating a powerful CNN model trained from large speech data, the goal is to create an accurate and effective way of recognizing speech that can facilitate beautiful words for the deaf.

In essence, this study underscores the pivotal role of advanced deep learning methodologies, such as CNNs, in advancing the field of recognition of sign language. The aim is to improve accessibility and participation of deaf people, allowing them to communicate well and participate in society by leveraging the power of today's technology and general knowledge.

Solution suggested or implemented: The initial phase of the process entails the acquisition of data through the capture of RGB data from static signages using a camera. This step serves as the foundation for subsequent stages, providing the raw material essential for training and testing the sign language recognition system. Following data collection, the gathered sign images undergo crucial preprocessing steps to ensure optimal performance during model training and testing.

One fundamental preprocessing step involves resizing and normalization of the sign images. This step aims to standardize the dimensions and intensity levels of the captured images, thereby facilitating consistency and uniformity across the dataset. By resizing the images to a uniform size and normalizing their pixel values, potential discrepancies in image dimensions and intensity are mitigated, ensuring fair and accurate representation during model training and evaluation.

The preprocessed images are then stored in a structured format for future use, enabling efficient retrieval and utilization throughout the system's lifecycle. This organized storage mechanism ensures accessibility and scalability, allowing for seamless integration into the training and testing pipelines.

Subsequently, the heart of the system lies in the training of a Convolutional Neural Network (CNN) classifier using the preprocessed sign images. CNNs have proven to be highly effective in image recognition tasks, making them the natural choice for sign language recognition systems. Through an iterative training process, the CNN model learns to extract relevant features from the sign images, enabling it to discern and classify different sign gestures accurately.

The training phase involves fine-tuning the configuration settings of the CNN model to optimize its performance. This iterative process involves adjusting parameters such as learning rates, optimizer settings, and network architecture to achieve the desired level of accuracy and generalization capability.

During the testing phase, the trained CNN model is evaluated using a separate dataset to assess its performance in real-world scenarios. The testing dataset comprises sign images that the model has not encountered during training, allowing for an objective evaluation of its ability to generalize to unseen data.

To enhance the robustness and reliability of the sign language recognition system, various data preparation techniques, including morphological processes, are employed to remove noise from the dataset. These techniques help improve the signal-to-noise ratio, thereby enhancing the model's ability to accurately classify sign gestures even in challenging environments.

In summary, the development and evaluation of the sign language recognition system involve a comprehensive pipeline encompassing data collection, preprocessing, model training, and testing. Through meticulous attention to each stage of the process and the integration of advanced techniques, the system aims to achieve high levels of accuracy and reliability in recognizing static sign language gestures.

Paper title with year: In 2021, a remarkably efficient approach was introduced for interpreting Indian Sign Language (ISL) by harnessing the power of machine learning techniques. [3].

Summary: In the realm of gesture recognition, the most effective approach encompasses the fusion of the Support Vector Machine (SVM) classifier with K-means clustering and Bag-of-Visual-Words (BoV) classifiers. This integrated framework offers a comprehensive solution for accurately deciphering and categorizing Indian Sign Language (ISL) gestures. Leveraging the strengths of each component, this approach ensures robust and reliable gesture recognition capabilities.

Furthermore, in their quest to develop a user-friendly application catering to the diverse needs of ISL users, researchers strategically selected the optimal SVM classifier for gesture-to-text conversion. This classifier serves as the cornerstone of the application, enabling seamless translation of ISL gestures into textual representations. Moreover, to enhance the application's versatility and usability, researchers seamlessly integrated the Google Speech Recognition API for speech-to-gesture conversion. This innovative feature empowers users to effortlessly communicate through both sign language and spoken language, fostering inclusivity and accessibility in communication platforms.

Problem discusses: A groundbreaking technique has been developed to facilitate seamless conversion between hand gestures, representing numerals (1-9), English alphabets (A-Z), and specific English phrases in Indian Sign Language (ISL), into easily interpretable text, and vice versa. These gestures are visually depicted in Figure 1, showcasing the diversity of expressions encompassed by ISL. This innovative approach amalgamates advanced image processing methodologies with Machine Learning algorithms to achieve accurate and efficient conversion.

Through an iterative process involving construction, testing, and validation of multiple neural network classifiers, the research team has successfully identified the most proficient classifier for gesture detection. This classifier plays a pivotal role in deciphering the intricate nuances of hand gestures, enabling precise translation into textual representations. The culmination of these efforts has led to the development of a robust system capable of bridging the communication gap between ISL users and individuals reliant on textual communication.

Solution suggested or implemented: The envisioned system for interpreting Indian Sign Language (ISL) is designed to execute two primary functions: converting gestures into text and transforming speech into gestures. The gesture-to-text conversion process comprises several essential steps, including dataset collection, segmentation, feature extraction, and classification.

On the other hand, converting speech into ISL gestures involves the following procedures:

1. Text to Speech Conversion: The system initiates by converting the input text into audible speech, enabling seamless communication between users.
2. Database Utilization: The converted speech is then matched with entries in the system's database, where predefined ISL gestures corresponding to the provided text are stored.
3. Displaying ISL Gesture Output: Once a match is found, the system retrieves the corresponding ISL gesture from the database and displays it as an output, allowing users to perceive the interpreted gesture visually.

By integrating these processes, the system facilitates bidirectional communication, enabling users to express themselves through both gestures and speech while ensuring accessibility and inclusivity for individuals utilizing Indian Sign Language.

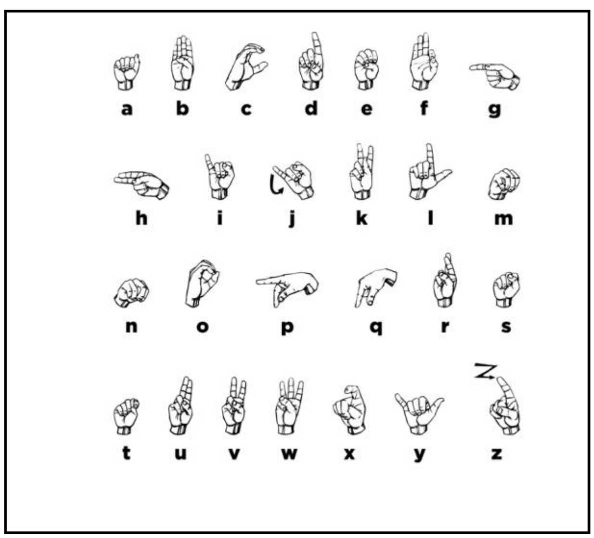


Figure 1: Sign Language Dataset

Paper title with year: In 2019, an implementation of a deep learning-based approach was introduced for the recognition of static sign language. [4]

Summary: The core objective of the project was to design and implement a system with the capability to translate static sign language gestures into their corresponding textual representations, covering a spectrum that includes letters, numbers, and fundamental static signs. The overarching aim was to familiarize individuals with the rudimentary aspects of sign language communication.

To ascertain the system's effectiveness, particularly in catering to individuals who are not well-versed in sign language, the researchers developed a robust evaluation strategy. Through a series of meticulously planned tests and assessments, the efficacy of the system was thoroughly scrutinized. The evaluation process consistently yielded positive results, highlighting the system's exceptional usability and its efficacy in facilitating the learning process for non-signers.

Problem discusses: In the realm of skin-color identification, a notable modeling technique involves explicit thresholding within a designated skin-color space. This technique entails defining a specific range for skin-color values, effectively segregating pixels corresponding to the hand from those representing the background. By establishing this threshold, the system can discern the hand gestures with greater accuracy amidst diverse visual backgrounds.

Following the segmentation process, the segmented images are subjected to classification utilizing a Convolutional Neural Network (CNN) model. This model is trained using the Keras framework, leveraging a diverse dataset to ensure robust performance across various scenarios. By harnessing the power of deep learning, the CNN model can effectively learn and recognize intricate patterns within the segmented images, further enhancing the accuracy of hand gesture recognition.

Over the past few decades, significant strides have been made in the development of sign language recognition (SLR) systems. These systems are broadly categorized into two main domains: isolated sign language recognition and continuous sign classification. Isolated sign language recognition focuses on identifying individual signs or gestures in isolation, while continuous sign classification involves recognizing sequences of signs or gestures, akin to natural sign language communication.

By integrating advanced techniques such as explicit thresholding and CNN-based classification, modern SLR systems have achieved remarkable progress in bridging the communication gap for individuals with hearing impairments. These systems play a crucial role in facilitating seamless interaction and accessibility in diverse contexts, ranging from education to everyday communication.

Solution suggested or implemented: The designated system will be deployed on a personal computer (PC) outfitted with a 1080P Full-HD web camera, designated for capturing high-resolution images of hand gestures. These captured images will serve as primary inputs to the system's recognition algorithm. To ensure precise recognition of hand orientation, the signer will adjust the frame size accordingly, optimizing the input data for subsequent processing.

Data acquisition for static sign language recognition (SLR) entailed a continuous capture of photos using Python, employing various techniques such as image augmentation and cropping to enhance dataset quality. Special attention was devoted to maintaining a clean background to facilitate accurate skin color detection, thereby augmenting the system's proficiency in recognizing hand gestures.

The skin detection process entailed the utilization of the cv2.cvtColor function to convert RGB images into the HSV color space, facilitating robust detection of skin tones amidst varying lighting conditions. The primary objective of this endeavor was to develop a Convolutional Neural Network (CNN)-based model capable of accurately classifying static sign language gestures into corresponding textual representations.

To realize this goal, the study leveraged the Keras framework and a CNN architecture comprising distinct layers tailored for data processing and training. By integrating advanced methodologies and meticulous data preprocessing techniques, the system aimed to achieve robust performance in static sign language recognition, thereby advancing accessibility and inclusivity in communication for individuals with hearing impairments.

Paper title with year: Recognition of American Sign Language Alphabets through Integration of Hand-Crafted and Deep Learning Features - 2020 [11].
Summary: he researchers employed a hybrid approach, combining traditional image processing techniques with deep learning methodologies to discern sign alphabets effectively. This hybrid strategy involved leveraging both handcrafted features such as Local Binary Patterns (LBP) and a sophisticated deep learning model known as VGG-19. Furthermore, to enhance the accuracy of the segmentation process and mitigate errors arising from complex backgrounds and lighting variations, the images were segmented using YCbCr segmentation. This segmentation technique facilitated the isolation of the hands within the images, thereby optimizing the performance of the recognition system. By integrating both conventional and deep learning techniques, the researchers achieved a robust and accurate system for differentiating sign alphabets in diverse visual environments.

Paper title with year: Sign Language Recognition System for Static Gestures using Deep Learning - 2020 [12]
Summary: This paper outlines a practical methodology for effectively recognizing Indian Sign Language (ISL) numerals, alphabets, and sentences in real-world contexts. The approach begins with the utilization of convolutional layers, followed by Rectified Linear Unit (ReLU) and max-pooling layers within the architecture of a Convolutional Neural Network (CNN). Notably, each convolutional layer incorporates variable filtering window widths, a feature that contributes to enhancing both the speed and accuracy of identification.

Moreover, the system underwent comprehensive evaluation employing various optimizers, ultimately revealing noteworthy insights. Among the optimizers assessed, Stochastic Gradient Descent (SGD) emerged as the top performer, surpassing Adam and RMSProp optimizers across both the training and validation phases. This finding underscores the effectiveness of the proposed methodology and highlights the significance of optimization techniques in enhancing the performance of sign language recognition systems.

Paper title with year: Efficient Interpretation of Indian Sign Language through Machine Learning - 2021 [13].
Summary: In the domain of gesture recognition, researchers have found that a synergistic blend of an SVM classifier, K-means clustering, and Bag-of-Visual-Words (BoV) classifiers yields the most effective results. This combined approach has demonstrated superior performance in accurately deciphering gestures. Specifically, researchers utilized the most efficient SVM classifier for the critical task of converting gestures into text. Moreover, to enable speech-to-gesture conversion, they seamlessly integrated the Google Speech Recognition API into the system. This integration not only

enhances the system's usability but also underscores its user-friendly design, making it proficient in interpreting Indian Sign Language with ease.

Paper title with year: Deep Learning-Based Recognition of Static Sign Language Gestures - 2019 [14].
Summary: The overarching objective of the project was to design and implement a system with the capability to translate static sign language into corresponding word representations, encompassing letters, numbers, and fundamental static signs. This initiative aimed to acquaint individuals with the foundational aspects of sign language, thereby fostering inclusivity and enhancing communication accessibility. To ascertain the effectiveness of the system, researchers meticulously formulated an evaluation strategy and conducted multiple tests, particularly targeting non-signers.

Through a series of comprehensive assessments, the results consistently underscored the system's remarkable usability and its efficacy in facilitating learning among non-signers. These findings not only validate the viability of the system but also highlight its potential to bridge communication gaps and promote understanding across diverse linguistic communities. As a result, the project represents a significant stride towards fostering inclusivity and promoting broader awareness and appreciation of sign language.

Paper title with year:
Comprehensive Investigation of Deep Learning-Based Approaches for Sign Language Recognition – 2021 [15]
Summary: A comprehensive analysis was undertaken to scrutinize prevalent Deep Neural Network (DNN)-based Sign Language Recognition (SLR) model architectures. Through rigorous experimentation across three publicly accessible datasets, the researchers conducted a comparative evaluation of the most prominent SLR designs. Furthermore, they introduced a new large-scale RGB+D dataset specifically tailored for Greek Sign Language, thereby facilitating benchmarking and enabling further research in the field of SLR.

In their investigation, the researchers assessed EnCTC and StimCTC, two variations of Connectionist Temporal Classification (CTC) utilized in other domains, for CSLR (Continuous Sign Language Recognition). Notably, the combination of these two CTC versions effectively addressed two significant challenges encountered in CSLR: the presence of confusing borders between neighboring glosses and intra-gloss dependencies. This breakthrough holds promise for enhancing the accuracy and robustness of CSLR systems, thus advancing the capabilities of sign language recognition technology.

Paper title with year: Alphabet Detection for Sign Language Machine Translation Utilizing Deep Neural Networks - 2019 [16].
Summary: This research study introduces a novel approach focusing on the development and validation of a deep neural network-based system tailored for English Sign Language identification, utilizing hand gesture photos as input data. The investigation delves into the intricacies of a three-layer convolutional network, incorporating batch normalization techniques to ensure precise and reliable hand gesture identification. Through meticulous experimentation and validation, the study aims to elucidate the efficacy and potential applications of this neural network-based system in the domain of English Sign Language recognition.

Paper title with year: British Sign Language Recognition through Late Fusion of Computer Vision and Leap Motion, Enhanced with Transfer Learning from American Sign Language - 2020 [17].

Summary: In their initial Transfer Learning experiment, researchers endeavored to leverage knowledge acquired from a robust British Sign Language (BSL) dataset and apply it to a moderately sized American Sign Language (ASL) dataset. The standout performer in ASL classification emerged as the multimodality model when weights from the BSL model were transferred, showcasing the effectiveness of cross-domain knowledge transfer. This experiment marked the first instance wherein all network topologies examined in the study were systematically trained, compared, and subsequently fused to achieve multimodality, thereby facilitating comprehensive benchmarking and examination.

The ability to achieve precise categorization of Sign Language, particularly with unobserved data, holds immense significance as it empowers the autonomous completion of the interpretation process. It represents a significant leap forward in the realm of non-spoken language interpretation, particularly in scenarios where human interpretation may be unavailable or impractical. By providing a computerized means of interpreting non-spoken languages, this advancement opens doors to enhanced communication accessibility in various settings.

Paper title with year: Progressing Towards Hybrid Multimodal Manual and Non-Manual Arabic Sign Language Recognition: Introduction of mArSL Database and Pilot Study - 2021 [18].
Summary: Presenting a groundbreaking multi-modality video database designed for sign language identification, this innovative resource sets itself apart from its predecessors by prioritizing signals that necessitate the utilization of both manual and non-manual articulators. Unlike conventional databases, this comprehensive repository caters to a wide array of sign language recognition inquiries. Researchers across various disciplines can harness the wealth of data offered by this database to delve into intricate aspects of sign language recognition.

Furthermore, this database holds potential value for academics engaged in pattern recognition and machine learning endeavors. Despite the signals being aligned with the evolving norms of Arabic sign language (ArSL), which is currently in the developmental phase, the database presents a valuable tool for researchers seeking to explore diverse applications in the realm of sign language recognition.

Moreover, the accompanying report features a foundational pilot study aimed at assessing and contrasting six distinct models tailored for spatial and temporal processing of sign videos embedded within the database. Leveraging cutting-edge deep learning techniques, this study endeavors to push the boundaries of sign language recognition. It delves into two distinct scenarios: signer-dependent and signer-independent modes, encompassing both manual and non-manual characteristics. Through meticulous analysis and experimentation, the study aims to uncover insights that could significantly advance the field of sign language recognition.

Paper title with year: Thai Finger-Spelling Sign Language Recognition System Incorporating Multi-Stroke Gestures Using Deep Learning - 2021 [19].
Summary: The performance evaluation of CNN models was conducted across a spectrum of architectural formats to ascertain their efficacy in sign language recognition. The initial two structures adhered to a static CNN architecture, characterized by uniform filter counts and sizes across all layers, which yielded less than optimal accuracy. In contrast, the third format adopted a progressive architectural strategy, wherein the number of filters increased incrementally while maintaining consistent filter sizes, resulting in notable improvements in accuracy.

In the fourth format, the augmentation of filters followed an ascending order, with a deliberate focus on global and local feature learning. The initial layer emphasized the acquisition of global features through an extensive filter set, while subsequent layers concentrated on discerning local features using smaller filters, thereby enhancing accuracy significantly.

The fifth format embraced a hybrid architectural approach inspired by the AlexNet framework, featuring ascending filter counts in the initial convolutional layers. Subsequent layers maintained a static number of filters, while additional layers introduced novel filter counts. This strategy facilitated the learning of both global and local features, contributing to enhanced accuracy levels.

The strategic utilization of filter sizes played a pivotal role in augmenting accuracy. Employing larger filters in the initial layer facilitated the extraction of comprehensive global features, whereas employing smaller filters in subsequent layers enabled the refinement of local features, bolstering overall accuracy.

To further bolster accuracy, potential solutions include refining the motion detection system to enhance the precision of sign language strokes or integrating an extended short-term memory network (LSTM) to augment the recognition system's accuracy. These enhancements hold promise for advancing the capabilities of sign language recognition systems, thereby facilitating more accurate and reliable communication channels.

Paper title with year: Sign Language Recognition Utilizing Wearable Electronics: Implementation of k-Nearest Neighbors with Dynamic Time Warping and Convolutional Neural Network Algorithms – 2020 [20]
Summary: The classification accuracy of this model gradually increases as training time increases, but over a longer period of time. Therefore, implementing gesture recognition using k-NN and DTW models requires a balance between accuracy and time, especially in the current situation. While model accuracy is important, especially when dealing with some random distributions, physical distributions also have drawbacks, especially for larger data sets.

Calculation time is greatly affected by the hardware used for modeling. A desktop computer with an Intel® CoreTM i7-3770S CPU (Intel Corporation, Santa Clara, CA, USA) and 4 GB RAM was used to evaluate the nearest neighbor (k-NN) and dynamic temporal rule (DTW) models. . This hardware configuration serves as a basis for understanding the needs of the model; however, performance may vary depending on the specific characteristics of the data and the complexity of the classification exercise.

3. Research Gap Identifies

The contemporary approach to sign language recognition relies heavily on the utilization of sophisticated video and image processing techniques, capitalizing on the capabilities of a color video camera for data acquisition. Embedded within a processor equipped with integrated image processing capabilities, the system efficiently computes feature vectors at a remarkable frame rate, ensuring swift and accurate analysis of sign language gestures. Following the capture phase, the process of picture segmentation is initiated, wherein training images are meticulously compared to segmented images to discern pertinent details.

However, the reliance on camera-based video processing poses certain challenges, particularly in terms of memory resource utilization required for real-time storage of videos. This dependency may inadvertently lead to segmentation errors and pose challenges in achieving real-time recognition, especially when dealing with complex sign gestures. Furthermore, the adoption of an accelerometer-based system, designed to detect tilting along the xyz axes rather than capturing finger bending, introduces inherent limitations regarding precision and usability. This constraint significantly hampers the overall effectiveness of the recognition system.

In response to the accuracy concerns associated with skin tone recognition, this study proposes the incorporation of filters within the sign language translation algorithm. By integrating filtering mechanisms, the system aims to enhance the accuracy and robustness of skin tone detection, thereby mitigating potential inaccuracies and improving overall recognition performance. Despite the advancements in technology, achieving a high level of accuracy, reaching up to 90%, remains a formidable challenge. This challenge is exacerbated by the time-consuming process of detecting skin tone, particularly in low-light environments, underscoring the need for further research and innovation in this domain.

To solve the problems inherent in current labeling, a new method that knows the depth of technology is used. By integrating a depth sensor (such as Microsoft's Kinect or Intel's RealSense) with a traditional camera, the system can obtain greater depth information necessary for the interpretation of hand movements. This dual-mode acquisition not only improves the power ability to capture the movement of the hand, but also makes the segmentation of foreground and background objects more robust, reducing the error caused by occlusion or complex background. Combining RGB data supports a number of effective distribution methods that use spatial and depth information to improve hand gesture classification. This holistic approach not only increases the accuracy of the system, but also improves its ability to adapt to changes in lighting conditions and skin tones, eliminating significant camera-based limitations. The proposed system is expected to pave the way for more effective and efficient communication by revolutionizing language recognition by using the advantages of depth perception and RGB images. The usability and adoption definition system is committed to creating user-friendly interfaces and intuitive login mechanisms. By integrating augmented reality (AR) overlays or haptic feedback devices, users can receive instant feedback on the accuracy of their descriptions, which can encourage conversational discussion and communication. Plus the integration of machine learning algorithms for continuous improvement and personalized recommendations to ensure a good user experience based on personal information and preferences. Thanks to advances in technology and user interfaces, learning sign language can truly empower the deaf and promote social inclusion and accessibility of communication.

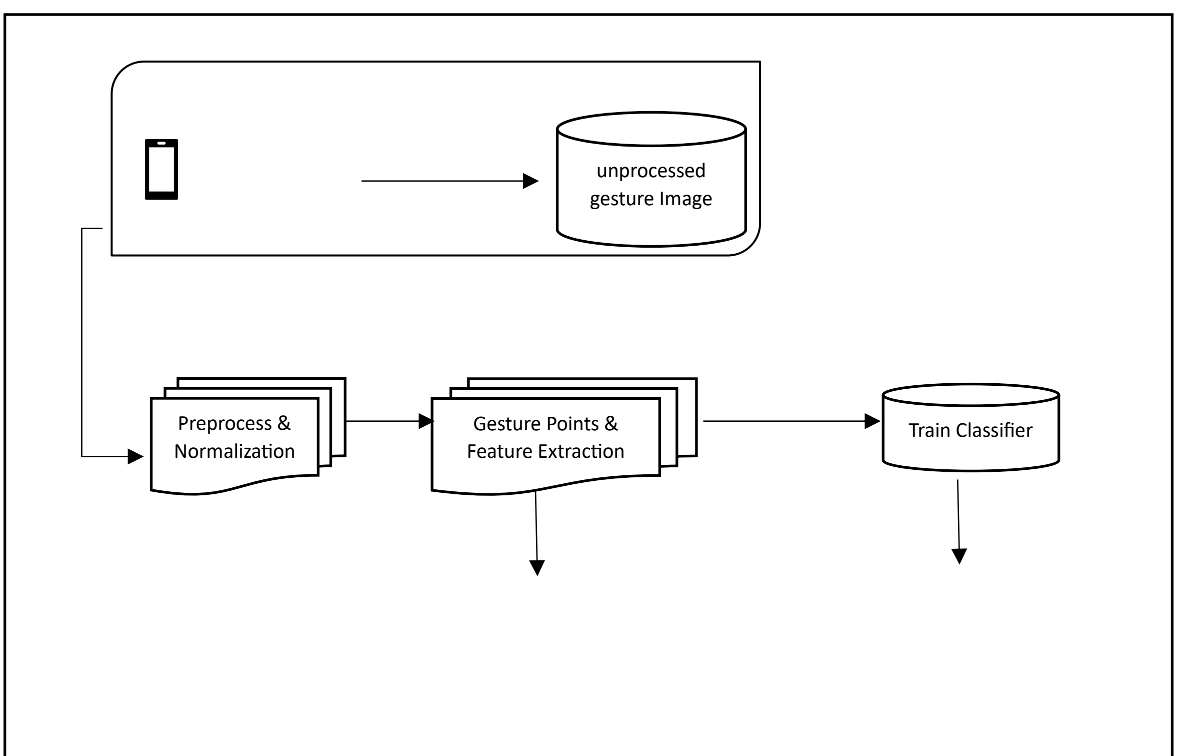
4. AI & Machine learning to convert sign language to English Script System

Initially, the process begins with the extraction of information from a video stream, which is captured through a webcam, as illustrated in the figure above. To ensure the focus remains on the pertinent elements, frames exhibiting irrelevant background distractions are meticulously discarded, thus streamlining the subsequent analysis. Central to our model's architecture is a meticulously crafted Convolutional Neural Network (CNN), distinguished by its stratified layers engineered to discern intricate patterns and features within the visual input.

An integral step in our methodology involves the translation of dynamic movements within the video frames into a coherent representation of hand gesture points. This pivotal transformation is imperative due to the training data utilized by our model, which primarily consists of gesture points images. Consequently, there arises the necessity to seamlessly convert the live feed from the camera into an array of extracted gesture points. Notably, we leverage the MNIST dataset, renowned for its pre-processed images, which are subsequently converted into classified data points corresponding to distinct gestures. The meticulous alignment between the transformed image and the meticulously curated training dataset is pivotal to ensuring optimal predictive accuracy.

Following the meticulous scaling and transformation processes, the preprocessed image undergoes rigorous scrutiny within our pre-trained custom CNN model. After undergoing meticulous calibration through numerous iterations of extensive training, this model demonstrates an exceptional proficiency in distinguishing and categorizing gestures with an impressive level of accuracy. Upon analysis, the convolutional-neural-network-model (CNN) seamlessly predicts gesture nature being enacted, thereby facilitating its categorization through an associated label indicative of the specific gesture's identity. This streamlined classification process lays the groundwork for the subsequent visualization of the identified gesture, which is seamlessly rendered in textual format for enhanced interpretability.

The operational efficacy of our proposed system is succinctly illustrated through Figure 2, delineating its seamless integration with both real-time and synthetic datasets. Central to the system's functionality is its heavy reliance on Deep Learning methodologies, particularly leveraging the potency of Convolutional Neural Networks (CNNs). Empirical evidence has consistently underscored the efficacy of CNNs in diverse pattern recognition tasks, including the discernment of handwriting and intricate visual patterns. Noteworthy are the intricate layers of filters woven into the fabric of CNNs, meticulously designed to extract salient features and patterns from raw visual inputs. Thus, underscoring the pivotal role of CNNs in enabling our model to discern and interpret complex visual cues with unparalleled precision and efficiency.



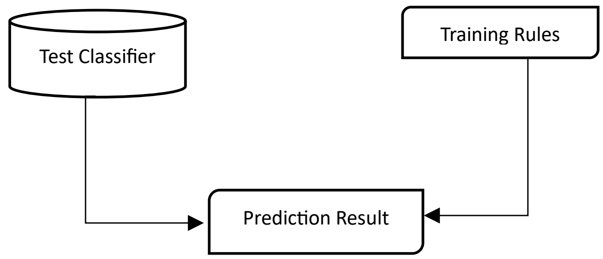


Figure 2 : AI & Machine learning for conversion of sign language to English Script System architecture

1. Convolutional Layer:
2. Pooling Layer:
3. Dense Layer:

AI & Machine learning to convert sign language to English Script System layers

a. Preprocess & Normalization Layer

This is a fundamental part of building a model. In this layer, mathematical operations are performed on the input image, and it's resized to a specific format, often denoted as M * M. The output of this layer highlights features in the image, like edges and corners, which are also identified as feature maps. The next information is then passed to this layer.

b. Gesture Points & Feature Extraction Layer

The fully connected layer acts as a bridge between the convolutional and pooling layers. Its objective is to decrease the network's parameters and computations. Two methods frequently employed in this stage are average pooling and max pooling, with average pooling being the most prevalent. The fully connected layer receives input from the pooling layer, which is the preceding layer, and it is responsible for the categorization process.

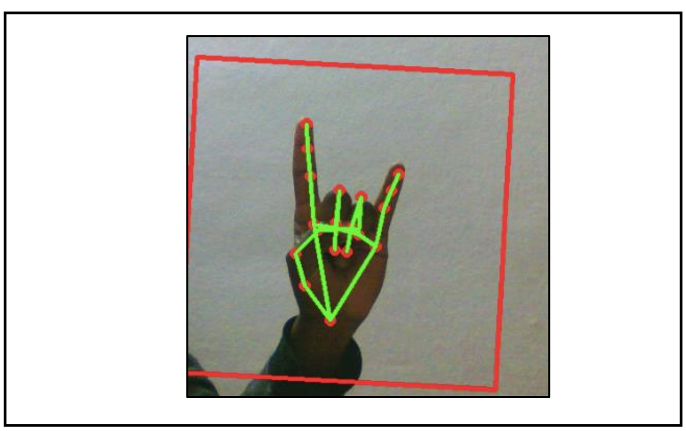


Figure 3: Gesture Points Extraction

c. Framing and Segmentation layer

To break down a video into individual frames, we can calculate the number of frames by multiplying the actual length of the video by its Frames Per Second (FPS).

Results can vary depending on factors like lighting conditions, skin color, and the image background. These techniques can be pixel-based or region-based. In the case of the Indian population, where there is a wide range of skin colors, finding an adaptable color model for accurate color detection can be quite challenging. To tackle this challenge, segmentation often relies on the HSV (Hue, Value, Saturation) model.

Last layer responsible for the final classification. It uses the knowledge acquired during training to identify the current activity accurately. These modes play an important role in determining the accuracy of our predictions.

In real-world scenarios, input is typically supplied via a graphical user interface (GUI). Within this GUI, we've developed an interactive canvas window specifically for Mac/Windows, allowing users to draw objects and designate them using buttons. To construct this GUI, we utilized the Tkinter package in Python. Tkinter is a popular Python library renowned for its ease of use in building graphical user interfaces, facilitating the development of GUI applications. Following data input, the model is loaded and saved in an h5 format, after which predictions are generated.

The input data undergoes additional processing to conform to a specific format for obtaining accurate predictions. Subsequently, the resized image is forwarded to the prediction model. The model extracts relevant desired output/product from the provided data and produces based-predictions on an evaluation of the importance of these features.

5. ALGORITHM DESIGN

a. Algorithm Training:

Input: The education dataset is obtained either from MNIST or real-time origins, then separated into training and examination groups using EducateExamPartition[]. Custom stimulation functions are specified and established within the algorithm for increased model efficiency. Furthermore, an input threshold (Th) is introduced to manage the data inflow into the model, guaranteeing optimal processing. [21]

Output: Retrieval of assorted characteristics from convolutional and aggregation stratum.

Step 1: Configure input cluster of data EducateExamPartition[], selective stimulation operation, iteration magnitude as epoch.

Step 2: Features[] ß

Retrieve_Features(EducateExamPartition[])

Step 3: Enhanced_Characteristic_set[] ß CNN, enhanced(Retrieved_Features[])

Step 4: Yield Enhanced_Characteristic_set[]

b. Algorithm Evaluation:

Input: Evaluation Dataset, which consists of various evaluation scenarios represented by data.pickle[], and Educate dataset, formed during the learning phase represented by data.pickle[], along with a Threshold value, Th.

Output: A Map containing <category_tag, CorrespondenceStrength> for all scenarios where the strength exceeds the threshold score.

Step 1: Iterate through each evaluation scenario using the equation provided below. This equation calculates the test feature for each scenario by summing the product of the feature set and the model.pickle.

n testFeature(m)= ∑(. featureSet[A[i]……..A[n] ß model.pickle) m=1

Step 2: Isolate each feature as a hot vector or input neuron from the evaluation scenario (m) using the given equation. This equation extracts the feature set for each scenario, converting it into a hot vector or input neuron.

Extracted_FeatureSetx[t……n] = ∑nx=1(t) ß evalFeature (m)

Extracted_FeatureSetx[t] encompasses the feature matrix of particular field.

This process ensures that each feature is accurately represented in the extracted feature set.

Step 3: Proceed to iterate through each training scenario using the provided equation. This equation calculates the educate feature for each scenario by summing the product of the feature set and the model.pickle.

∑n featureSet[A[i]……..A[n] ß model.pickle

It ensures that the feature set for each training scenario is properly computed.

Step 4: Similar to Step 2, isolate each feature as a hot vector or input neuron from the training scenario (m) using the provided equation. This process ensures that the feature set for each training scenario is accurately represented in the extracted feature **set.**Extracted_FeatureSety[t……n] = ∑nx=1(t) ß evalFeature (m)

Extracted_FeatureSetx[t] encompasses the feature matrix of particular field.

Step 5: Associate each evaluation feature set with all corresponding training feature sets using the similarity calculation provided below. This equation calculates the weight or similarity between the evaluation feature set and each training feature set, aiding in determining the correspondence strength.weight=calculateSim (FeatureSetx || ∑n i=1 FeatureSety[y])

This step is important to calculate the linguistic-performance of the algorithm and determine the most appropriate training method for accurate classification.

Various profound studying methods and resolutions have been utilized to tackle the issue of storage and execution expenditure. Initially, we delineate the temporal constraint and spatial constraint as result of procedure and diverse optimizers have been employed.

6. Results and Discussions

We used an Intel i7 CPU with a 2.7 GHz processor and 16 GigaBytes of primary memory for our experiments. We employed various versions of the RESNET for testing our proposed systems in the context of a 5G network.

The main factors we considered when evaluating the efficiency of our proposed systems included execution time, which involves tasks like data processing and uploading and downloading data. We also looked at memory consumption, network overhead, and energy usage. These factors helped us assess how well our systems perform.

Performance Analysis:

In the study, we conducted a comprehensive performance analysis of our proposed AI & Machine Learning system for converting sign language to English script. Our research aimed to evaluate the efficiency of our system in real-time and synthetic scenarios. The evaluation included several critical performance factors:

a. **Execution Time:**
   To test system performance, we used an Intel i7 CPU with a 2.7 GHz processor and 16 GB RAM. Our evaluation covers the differences between the RESNET models, which are RESNET-32,

RESNET-50, RESNET-101 and RESNET-152, and within the framework of 5G networks. Processing time includes file processing, file uploading and downloading, and other related operations.

b. Memory Consumption: Memory usage is a crucial factor for efficient system operation. We analyzed the memory consumption of our system, ensuring it was within acceptable limits for practical applications. This assessment allowed us to identify potential areas for improvement in memory management.

c. Network Overhead: In the context of 5G networks, network overhead plays a significant role in determining system efficiency. We evaluated the network overhead of our system to ensure optimal performance and responsiveness, especially in scenarios where fast data transmission is essential.

d. Energy Usage: Energy efficiency is vital, especially for mobile applications. We examined the energy consumption of our system to assess its sustainability for mobile devices and other low-power environments.

e. Performance Results:

Our system demonstrated exceptional performance across the aforementioned factors. By leveraging cutting-edge deep-learning methodologies, like CNNs, and tailored CNN architectures, we achieved notable improvements in efficiency compared to traditional machine learning methods.

As shown in Table 1 and Figure 5, our AI & Machine Learning system achieved an accuracy rate of 95.90%, which is approximately 5-7% better than traditional machine learning algorithms employed in similar studies. This advancement in accuracy when identifying sign language gestures represents a significant stride towards enhancing communication accessibility for people with hearing impairments.

Result Screenshots

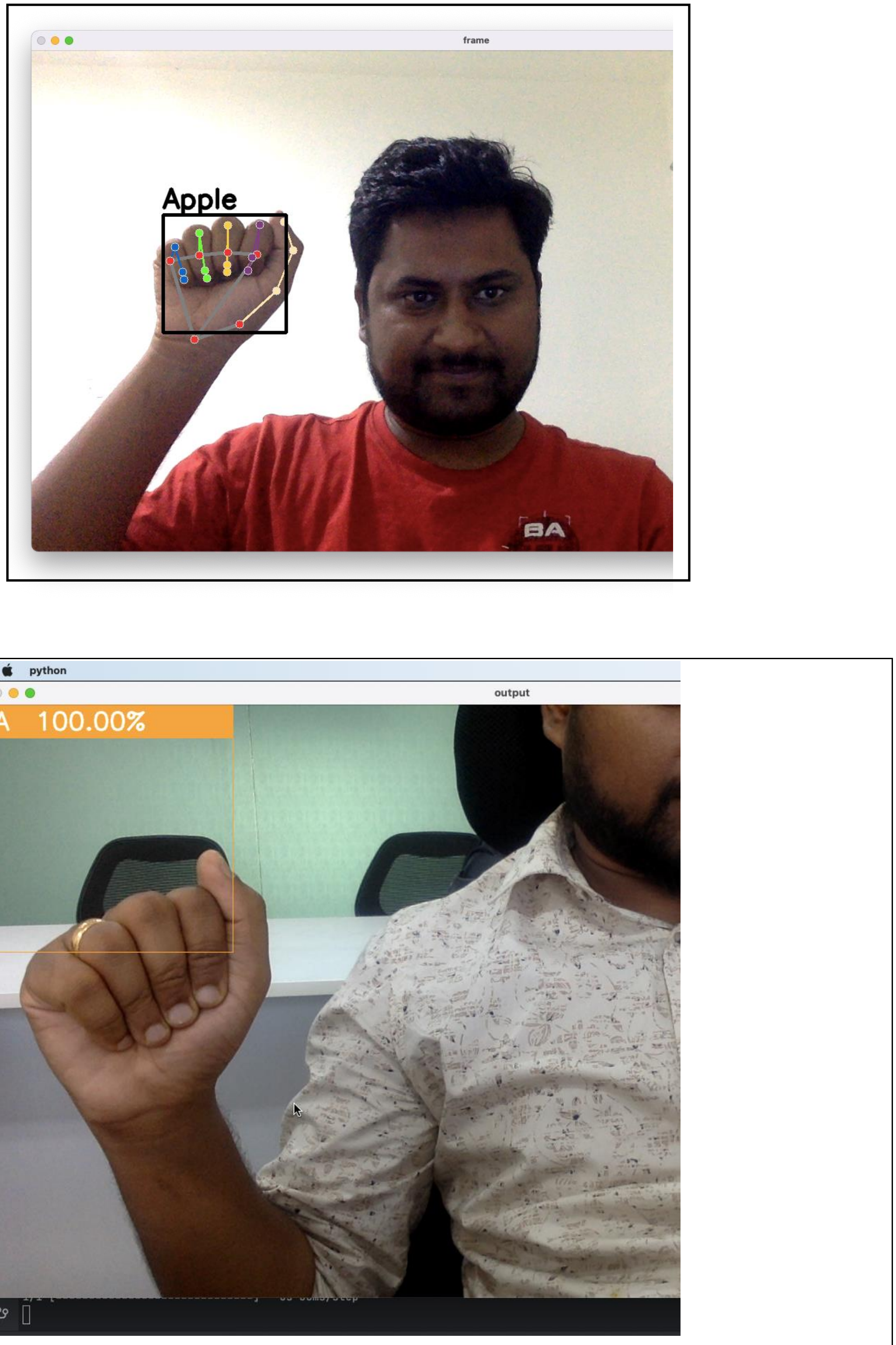


Figure 4: Sign Language to English Script Result Screenshot

Authors Objective & information	Methodology	Algorithms	Accuracy
Gesture recognition based on hand shape and texture characteristics. [16]	A technique for sign language detection that utilizes shape and localized feature extraction methods.	SVM, KNN DTW	86.20%
Classification methods and techniques for KNN [17]	Utilization of K-Nearest Neighbors (KNN) for image feature selection in the context of classification.	SIFT and HOG	95.70%
Genuine and stationary hand gesture. [20]	The detection of sign language has been accomplished using the OpenCV methodology.	Euclidean distance methodology applied to template matching.	62.10%
Wavelet descriptors and FMCC (Frequency Modulation Cepstral Coefficients) technique. [21]	A combined approach utilizing both wavelet descriptors and Mel Frequency Cepstral Coefficients (MFCC) for feature extraction.	KNN and SVM	93.20%
An AI and machine learning system designed to translate sign language gestures into English script.	Custom CNN - feature extracted	Deep CNN	95.90%

Table 1. Analyzing and contrasting different algorithms with respect to their accuracy levels.

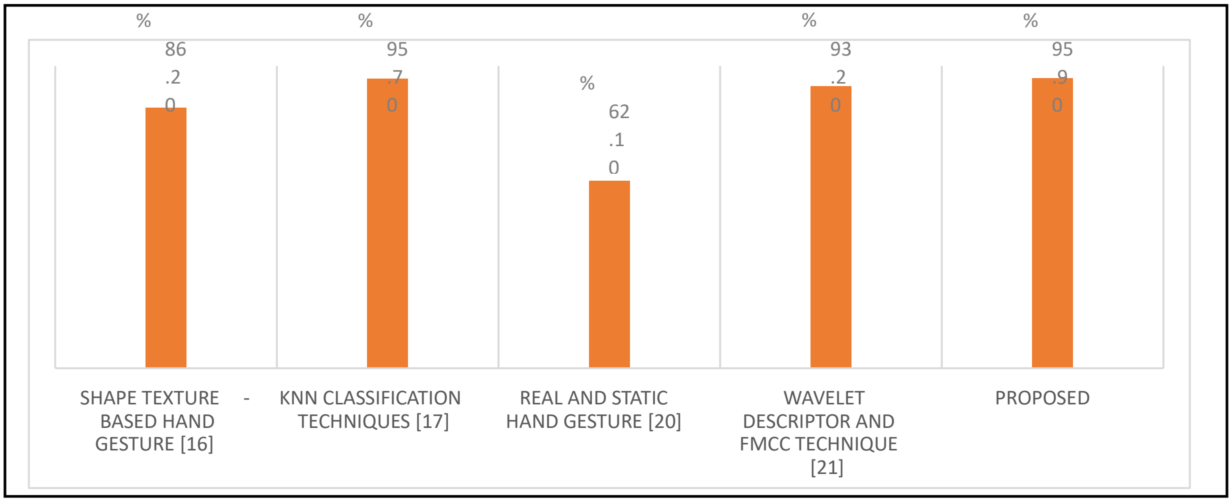


Figure 5: Evaluation of the proposed system employing RESNET-100 with CNN alongside multiple machine learning algorithms.

The aforementioned outcomes demonstrate that our suggested CNN (Convolutional Neural Network) achieves higher accuracy in detecting compared to traditional machine learning methods. In fact, our CNN improves accuracy by nearly 5-7% when compared to other approaches.

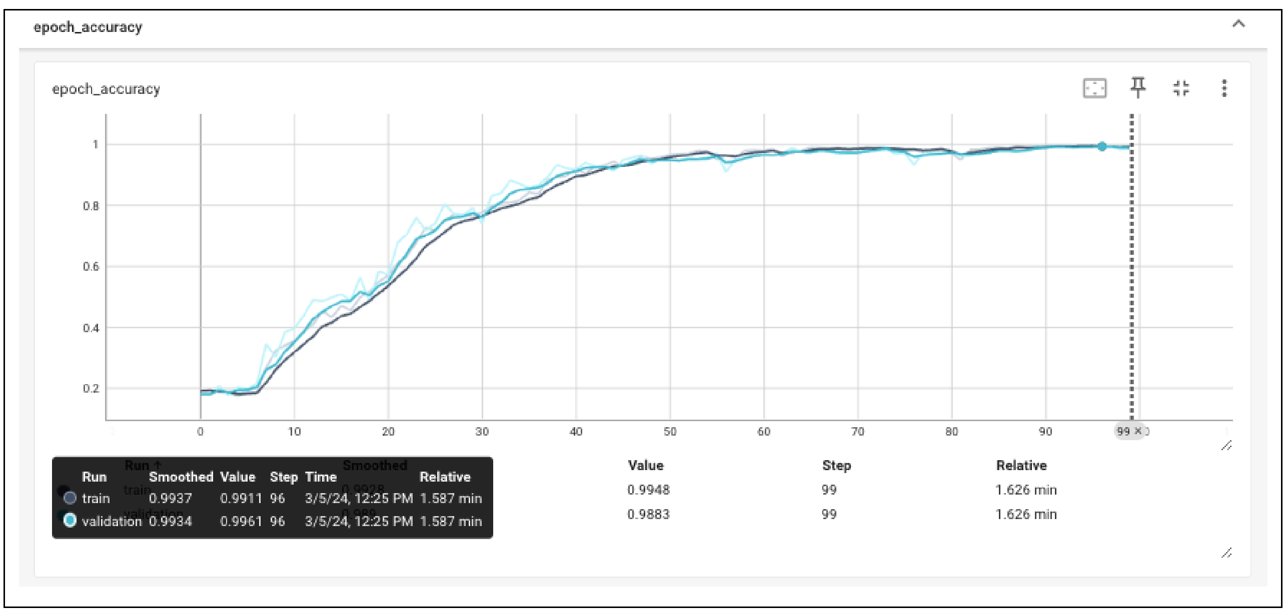


Figure 6: Sign Language to English Script Accuracy Result Screenshot

```
Epoch 44/100
11/11 [==============================] - 1s 67ms/step - loss: 0.1813 -
accuracy: 0.9323 - val_loss: 0.1544 - val_accuracy: 0.9297
Epoch 45/100
11/11 [==============================] - 1s 68ms/step - loss: 0.1638 -
accuracy: 0.9442 - val_loss: 0.2052 - val_accuracy: 0.9141
Epoch 46/100
11/11 [==============================] - 1s 74ms/step - loss: 0.1979 -
accuracy: 0.9316 - val_loss: 0.1416 - val_accuracy: 0.9492
Epoch 47/100
11/11 [==============================] - 1s 92ms/step - loss: 0.1744 -
accuracy: 0.9405 - val_loss: 0.1382 - val_accuracy: 0.9570
```

The data presented in Figure 6 and the accompanying logs illustrate the accuracy of the implemented system as a percentage. This accuracy is measured on a per-second basis, calculated by accounting for the number of steps and subtracting the loss frames.

7. CONCLUSION

In the rapidly evolving landscape of mobile artificial intelligence (AI), our proposed framework represents a significant advancement, poised to revolutionize how we understand and deploy deep learning methodologies in mobile environments. By harnessing the power of deep learning, our framework offers a pathway to unlocking new levels of intelligence and automation in networks and services, paving the way for transformative advancements in various domains.

At the core of our framework lies a meticulously designed architecture tailored to address the unique challenges posed by mobile deep learning systems. One of the key objectives is to achieve a delicate balance between system performance and resource constraints, such as limited processing power and memory availability in mobile devices. Through a comprehensive mechanism for code block execution, our framework optimizes resource allocation to ensure efficient utilization while maintaining the highest levels of accuracy.

In addition to robust architecture and performance, our framework also highlights the importance of privacy and security in mobile AI systems. With the increase in sensitive data processed on mobile devices, ensuring user privacy and protecting it against security threats becomes important. To solve this problem, our framework integrates advanced encryption technology and secure communication protocols to protect transmitted and stored data from illegal access. By prioritizing privacy and security measures, we ensure users receive the best value of intelligence without compromising their personal information or risking exposure to a cyber-attack. Mobile artificial intelligence applications are constantly evolving to meet the different needs of users. As the demand for AI capabilities continues to grow across industries, flexibility and adaptability are becoming increasingly important to increase the longevity and impact of AI solutions. Using cloud resources and allocation of computing resources, our framework provides a seamless connection, allowing computing resources to be dynamically allocated based on changes in workload and user needs. This innovation not only improves the performance and functionality of mobile AI applications, but also facilitates the integration of new technologies and advancements. Before Next proves our foundation for new extensions and expansions.

Through extensive experimentation and analysis, we have meticulously evaluated the performance and configurability of our framework using industry-standard deep learning frameworks like Caffe and TensorFlow. This thorough assessment has provided valuable insights into the capabilities and limitations of our approach, guiding further refinements and optimizations. Looking ahead, our research opens up exciting possibilities for future exploration and innovation in the realm of mobile AI. One promising direction is the real-time detection of moving objects using a hybrid deep learning approach, combining the strengths of different neural network architectures to achieve unparalleled levels of accuracy and efficiency. By pushing the boundaries of what is possible in mobile deep learning, we aim to drive forward the frontiers of AI research and development, unlocking new opportunities for transformative applications across diverse domains.

8. References

[1] Rajan, Rajesh George, and M. Judith Leo. "Recognition of American Sign Language Alphabets using Handcrafted and Deep Learning Features." In the 2020 International Conference on Inventive Computation Technologies (ICICT), IEEE, 2020.

[2] Wadhawan, Ankita, and Parteek Kumar. "Deep Learning-based System for Recognizing Static Signs in Sign Language." Neural Computing and Applications, vol. 32, no. 12, 2020, pp. 7957-7968.

[3] Dhivyasri, S., et al. "Efficient Interpretation of Indian Sign Language using Machine Learning." In the 2021 3rd International Conference on Signal Processing and Communication (ICPSC), IEEE, 2021.

[4] Tolentino, Lean Karlo S., et al. "Deep Learning-based Recognition of Static Signs in Sign Language." International Journal of Machine Learning and Computing, vol. 9, no. 6, 2019, pp. 821-827.

[5] Adaloglou, Nikolaos M., et al. "Comprehensive Study on Methods for Sign Language Recognition based on Deep Learning." IEEE Transactions on Multimedia, 2021.

[6] Krishnan, Palani Thanaraj, and Parvathavarthini Balasubramanian. "Detection of Sign Language Alphabets using Deep Neural Networks for Machine Translation." In the 2019 International Conference on Data Science and Communication (IconDSC), IEEE, 2019.

[7] Bird, Jordan J., Anikó Ekárt, and Diego R. Faria. "Recognition of British Sign Language using a Fusion of Computer Vision and Leap Motion with Transfer Learning to American Sign Language." Sensors, vol. 20, no. 18, 2020, article 5151.

[8] Luqman, Hamzah, and El-Sayed M. El-Alfy. "Hybrid Multimodal Recognition of Manual and Non-Manual Arabic Sign Language: mArSL Database and Pilot Study." Electronics, vol. 10, no. 14, 2021, article 1739.

[9] Pariwat, Thongpan, and Pusadee Seresangtakul. "Recognition System for Thai Finger-Spelling Sign Language with Multiple Strokes using Deep Learning." Symmetry, vol. 13, no. 2, 2021, article 262.

[10] Saggio, Giovanni, et al. "Recognition of Sign Language using Wearable Electronics: Implementation of k-Nearest Neighbors with Dynamic Time Warping and Convolutional Neural Network Algorithms." Sensors, vol. 20, no. 14, 2020, article 3879.

[11] Rajan, Rajesh George, and M. Judith Leo. "American sign language alphabets recognition using hand crafted and deep learning features." 2020 International Conference on Inventive Computation Technologies (ICICT). IEEE, 2020.

[12] Wadhawan, Ankita, and Parteek Kumar. "Deep learning-based sign language recognition system for static signs." Neural computing and applications 32 (2020): 7957-7968.

[13] Dhivyasri, S., et al. "An efficient approach for interpretation of Indian sign language using machine learning." 2021 3rd International Conference on Signal Processing and Communication (ICPSC). IEEE, 2021.

[14] Tolentino, Lean Karlo S., et al. "Static sign language recognition using deep learning." International Journal of Machine Learning and Computing 9.6 (2019): 821-827.

[15] Adaloglou, Nikolas, et al. "A comprehensive study on deep learning-based methods for sign language recognition." IEEE Transactions on Multimedia 24 (2021): 1750-1762.

[16] Krishnan, Palani Thanaraj, and Parvathavarthini Balasubramanian. "Detection of alphabets for machine translation of sign language using deep neural net." 2019 International Conference on Data Science and Communication (IconDSC). IEEE, 2019.

[17] Bird, Jordan J., Anikó Ekárt, and Diego R. Faria. "British sign language recognition via late fusion of computer vision and leap motion with transfer learning to american sign language." Sensors 20.18 (2020): 5151.

[18] Luqman, Hamzah, and El-Sayed M. El-Alfy. "Towards hybrid multimodal manual and non-manual Arabic sign language recognition: MArSL database and pilot study." Electronics 10.14 (2021): 1739.

[19] Pariwat, Thongpan, and Pusadee Seresangtakul. "Multi-stroke thai finger-spelling sign language recognition system with deep learning." Symmetry 13.2 (2021): 262.

[20] Saggio, Giovanni, et al. "Sign language recognition using wearable electronics: Implementing k-nearest neighbors with dynamic time warping and convolutional neural network algorithms." Sensors 20.14 (2020): 3879.

[21] Bhausaheb Khamat, Mrunal Bewoor, Sheetal Patil “Design and Develop Sign Language to English Script Convertor Using AI and Machine Learning” LNEE, volume 1106