AVE Trends in Intelligent Computing Systems

AVE TRENDS PUBLICATIONS

Real-Time Hand Gesture Recognition for Stroke Rehabilitation Using Deep Learning Approach

S. Karthik*
Department of Computer Science and Engineering, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India. karthiks1087@gmail.com

S. Bhaggiaraj
Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India. ktsbhaggiaraj@srec.ac.in

Edwin Shalom Soji, L. Sarala Deve and S. Silvia Priscila
Department of Computer Science, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India.
edwinshalomsoji.cbcs.cs@bharathuniv.ac.in, saraladeve.cse@bharathuniv.ac.in, silviaprisila.cbcs.cs@bharathuniv.ac.in.

S. Suman Rajest
Department of Research and Development and International Student Affairs, Dhaanish Ahmed College of Engineering, Chennai, Tamil Nadu, India. sumanrajest414@gmail.com

*Corresponding author

Abstract: Developing patients' ability to recognize hand gestures is one of the benefits of deep learning generation. Intending to develop a stroke rehabilitation device that is low-cost, non-invasive, and environmentally friendly, this study intends to provide stroke patients with a mechanism that allows them to operate assistive devices via hand gestures. When bringing hand motions to life, superior photographs and preprocessing have been crucial in teaching a deep MediaPipe and OpenCV architecture. With the ability to enable patients to use assistive devices by applying hand gestures, the recommended machine shows promise as a valuable instrument for stroke rehabilitation. This is because the utilization of the recommended machine results in caution. Additional motions might be made to materialize on the device, which would assist stroke patients experiencing various limits. The fact that this is the case demonstrates how much knowledge of the past can improve the accuracy and efficiency of hand gesture popularity systems for stroke victims.

Keywords: Deep Learning; Hand Gesture Recognition; Stroke Rehabilitation; MediaPipe and OpenCV; Real-Time Processing; Machine Learning; Gesture Detection; Patient Engagement; Motor Function Recovery.

Cite as: S. Karthik, S. Bhaggiaraj, E. S. Soji, L. S. Deve, S. S. Priscila and S. S. Rajest "Real-Time Hand Gesture Recognition for Stroke Rehabilitation Using Deep Learning Approach," *AVE Trends In Intelligent Computing Systems*, vol. 1, no. 3, pp. 157 –172, 2024.

Journal Homepage: https://avepubs.com/user/journals/details/ATICS

Received on: 03/03/2024, **Revised on:** 01/05/2024, **Accepted on:** 22/06/2024, **Published on:** 01/09/2024

1. Introduction

Within artificial intelligence (AI), machine learning is a subfield whose definition is generally that of a machine's capacity to replicate intelligent human behaviour. Particularly in the context of healthcare applications, where it has the potential to transform patient care and rehabilitation processes [1], this branch of artificial intelligence has recently attracted a lot of interest [2]. Machine learning aims to create algorithms capable of learning from data, spotting trends, and generating predictions or judgments without explicit programming for every task [3]. In machine learning, historical data provides input to algorithms generating models reflecting the information using patterns [4]. Fundamentally, machine learning combines computational and statistical ideas from several fields, including biology [5], cognitive science, and statistics [6]. Leveraging these ideas helps machine learning algorithms anticipate based on fresh, unprocessed data, improving their accuracy over time [7]. Machine learning is used in many different disciplines, including medicine, where it is applied for patient monitoring, tailored treatment plans, and diagnosis predictions [8]. Among medical fields, rehabilitation for stroke victims is one of the most exciting uses. Particularly in the hands, stroke frequently causes decreased motor skills, greatly restricting a patient’s capacity to engage in daily tasks [9]. Conventional rehabilitation techniques might not always produce the best results and might be resource-intensive [49]. This emphasizes the need for creative ideas for patients that are non-invasive, reasonably priced, and more interesting [44].

1.1. Value of Gesture Recognition for Rehabilitation

Technology for hand gesture recognition presents a revolutionary method of stroke rehabilitation [12]. Encouragement of patients to use hand gestures to communicate with assistive technologies can help to create a more interesting and motivating rehabilitation process [13]. Controlling smart home devices, interacting, and engaging in virtual reality settings designed for rehabilitation activities [14], among other uses, can be made easier by gesture recognition. This improves stroke survivors’ quality of life and promotes independence, essential for psychological rehabilitation [41]. With recent computer vision and machine learning developments, especially with MediaPipe, real-time hand gesture recognition is now possible [16]. Designed by Google, MediaPipe is an open-source framework providing a complete pipeline for creating multimedia apps incorporating hand tracking and pose estimation [17]. MediaPipe’s capacity to instantly identify hand landmarks offers a strong basis for creating interactive systems capable of precisely and responsively interpreting patient motions [48].

1.2. Synopsis of MediaPipe

MediaPipe comprises pre-trained algorithms capable of tracking and identifying hand landmarks, generating 21 critical points for every hand [19]. These sites let one fully grasp hand motions since they include the fingers, palm centre, and knuckles. The framework's architecture offers hardware acceleration, improving its performance and efficiency and making it fit for real-time applications [20]. MediaPipe’s versatility lets builders design tailored gesture detection systems that are fit for more general rehabilitation initiatives [21]. Furthermore, MediaPipe’s simplicity of integration lets researchers and developers quickly prototype and test new concepts, facilitating fast development cycles [22]. In the framework of stroke rehabilitation, where patient needs might vary greatly and solutions must be flexible to fit various degrees of motor functioning impairment [23], this is especially crucial.

1.3. Problems and Future Directions

Even though gesture recognition and machine learning hold great potential for stroke rehabilitation, some difficulties still exist [24]. The variety of motions patients may need to do presents one major challenge, varying depending on personal ability and inclination [25]. Maximizing the value of gesture recognition technology in rehabilitation depends on a system that can precisely identify a broad spectrum of gestures and adjust to user-specific needs [26]. Furthermore, patient involvement is very important for rehabilitation success. Thus, any gesture detection system must be accurate but also interesting and pleasant to operate [28]. Including gamification, feedback systems, and personalization in the rehabilitation process can help patients to be much more motivated and follow therapeutic recommendations [27]. Furthermore, how well current healthcare systems incorporate machine learning models is considered [29]. Working with software developers, doctors, and healthcare managers will help guarantee that the solutions are user-friendly, clinically appropriate, and in line with healthcare practices [30].

Ultimately, study and application have an interesting horizon at the junction of machine learning, gesture detection, and stroke rehabilitation [45]. The suggested approach seeks to create a low-cost, non-invasive mechanism allowing stroke victims to operate assistive devices via hand gestures, improving their rehabilitation experience [15]. This initiative aims to provide insightful analysis and tools for enhancing the quality of life for stroke survivors by using deep learning methods and MediaPipe frameworks. Expanding the gesture recognition capacities of the system, improving user involvement, and investigating the integration of these technologies into more general rehabilitation frameworks will be the main priorities of the next work. Maintaining the patient’s demands first in development will help us ensure our solutions are technologically advanced but also useful and practical for end users.

2. Related Work

Superior deep-learning skeleton collection techniques have allowed researchers to extract skeleton points from the body more recently than they could have using expensive gear like motion capture devices or depth cameras [50]. Scholars can gather

RGB datasets and deduce 3D skeletons [4] or 3D skeleton [3] information using common cameras such as online, CCTV, or mobile. Sometimes, the skeleton data is produced by other channels, such as WiFi signals [46]. Researchers are more interested in creating skeleton-based action and hand gesture recognition systems due to this portable, less costly system and device, which also helps generate large datasets connected to resources [11]. We can often discriminate between two separate portions from the existing skeleton-based hand gesture recognition research effort. One can employ machine learning algorithms [47] to extract new features from the skeleton sequence [42]. In order to improve temporal and spatial contextual information and the generated feature and classification [3], the second kind is capable of designing and developing a deep neural network model [9].

The viewpoint-invariant global motion characteristics are often found in the skeleton-based dataset, which is generally regarded as a good representation [23]. Too little effectiveness in a feature is the primary flaw of research projects focused solely on features. Without considering global motion characteristics, many other researchers concentrated on the perspective invariant frame-based feature [43]. Zhang et al. [19] have suggested using the DD-Net CNN network in conjunction with global motion and non-global motion variables to identify gestures [51]. Their model's primary flaw is its lack of potential and duplicated features, which result in subpar performance. To identify hand movements, several researchers employed 1D CNN, CNN [23], RNN [28], and the Graph attention model [3]. Here, we suggest using a chosen skeleton point-based feature and attention-based temporal feature extraction approach to increase the model's performance accuracy while speeding up its speed.

Particularly for uses in rehabilitation, recent developments in deep learning and computer vision have greatly improved the discipline of gesture detection. Particularly for stroke patients, who typically need customized rehabilitation solutions, several research studies have investigated several approaches to increase the accuracy and efficiency of gesture-detecting systems. One interesting method uses convolutional neural networks (CNNs) for gesture recognition. CNNs have been shown by several studies to be rather good at identifying dynamic and stationary hand movements from pictures and video streams. For human-computer interaction, Zhang et al. [19], for example, used CNN architecture to attain great accuracy in identifying gestures, hence underlining the possibilities of deep learning in real-time applications [32]. In order to improve gesture categorization from sequential video data, Ranjith et al. [33] also included a hybrid model combining CNNs and recurrent neural networks (RNNs), therefore producing better results in recognition accuracy.

There is great interest in the use of gesture recognition in stroke recovery. Studies by Gnanaguru et al. [34] have concentrated on developing interactive rehabilitation activities utilizing motion capture and gesture recognition technologies to improve patient involvement and motivation. Their results imply that encouraging active engagement, including gesture-based technologies in rehabilitation, will help stroke victims achieve better recovery results. Moreover, for real-time hand tracking and gesture identification, the combination of frameworks such as MediaPipe has revolutionized the remarkable detecting powers of MediaPipe and made it possible for scientists to create real-time, precise tracking of hand motion devices. For instance, Gowthami and Priscila [35] investigated MediaPipe's efficacy in giving patients real-time feedback during exercises through hand gesture detection in rehabilitation settings. Their research highlights how responsive systems could inspire patients and improve their rehabilitation experience [36].

Furthermore, gamification techniques have boosted patient adherence to rehabilitation guidelines and motivation. Chen et al. [18] presented a gamified gesture recognition system that motivates stroke survivors to interact with rehabilitation activities through interactive games, improving patient outcomes and therapy session adherence. This strategy is consistent with the idea that improved rehabilitation experiences might result from user-friendly and interesting interfaces.

3. Dataset

Using eleven skeleton-based hand-gesture datasets, including SHREC'17 [15], DHGD [14], JHMBD [10], MSRA [3], ICVL, NV Gesture, NYU, NTU, UCF-Kinetic, UTKinetic, Florence 3-D action [3] we assessed the model. These datasets are primarily regarded as benchmarks for hand gesture recognition. Our goal was most closely resembled by the two datasets we selected. Based on the following equations (1), we viewed the Skeleton dataset as a 3D dataset that may be represented.

$$\text{D}= (Q1, Q2, Q3, \dots , Qn)^T \qquad (1)$$

Where D stands for the dataset sequence and Q_j

$$\text{Qj(t)} = (\text{X(i)},\text{Y(i)},\text{Z(i)}) \qquad (2)$$

$$Qi = (Xi, Y\ , Zi) \in \mathbb{R}^3, \forall i \in \qquad (3)$$

3.1. SHREC'17 Dataset

One of the best benchmark datasets for skeleton-based hand motions is SHREC'17 [15]. This collection has fourteen right-hand gestures, finger configuration, and metadata. Two distinct approaches to class label setup were considered when recording this dataset from 27 individuals using an Intel RealSense Camera [37]. They gathered 2800 video sequence datasets, with 20–70 frames each. There are 22 skeletal points in each of the 32 frames from each film that we took into consideration for this project. They use two methods, coarse gesture and fine gesture, to customize the class label setting. Both methods primarily rely on finger spelling. The dataset's 22 hand skeleton points are shown in Fig. 2.

3.2. DHGD Dataset

With 14 rig hand motions and finger spelling types, this freely accessible dataset employs a skeleton-based dynamic hand gesture dataset [29]. This dataset contains 2800 films that were gathered from 20 different sources. Each of the 32 frames in each film we examined in our study comprises 22 hand key points, including 3D coordinates [38]. Moreover, this dataset is fine and coarse according to finger spelling. The dataset's label setup also adhered to the coarse and fine methods. The dataset's 22 essential points are displayed in Fig. 2.

3.3. JHMBD Dataset

The JHMBD (Joint Hand and Motion Body Dataset) is a complete skeleton-based hand gesture dataset designed to facilitate the development and assessment of gesture recognition algorithms [39]. It captures thorough skeleton representations of hand and body movements by including several hand gestures executed by several people. Joint locations and motion data are tagged with every gesture in the collection, thereby enabling an in-depth study of hand movements' spatial and temporal dynamics [40]. Particularly helpful for human-computer interaction, animation, and robotics, the JHMBD dataset attempts to close the distance between hand and total body motions. The collection is a vital source of high-quality skeletal data for academics and developers working on gesture detection and associated technologies.

3.4. MSRA Dataset

Designed to forward gesture detection and interpretation, the MSRA (Microsoft Research Asia) skeleton-based hand gesture collection is a great tool. This dataset consists of a comprehensive collection of hand gestures carried out by several individuals, each gesture shown by thorough skeletal data recording the positions of important joints in the hand and arm. The MSRA dataset guarantees variation in both performance and context by including a wide spectrum of gestures, enabling the training of strong machine-learning models from both sides. Every gesture is painstakingly documented and provides necessary data for hand movement dynamics analysis. This dataset provides researchers with a strong basis for building and testing creative gesture detection algorithms, benefiting human-computer interaction, virtual reality, and robotics applications.

3.5. ICVL Dataset

Comprising a complete collection intended to support hand gesture detection and analysis, the ICVL (Imperial College Visual Computing Lab) skeleton-based hand gesture dataset is It consists of several participants doing a variety of hand gestures, each of which is shown using thorough skeleton data that precisely positions and moves important joints in the hands and arms. The dataset offers a strong basis for training and assessing machine learning models since it comprises several gestures in several situations and illumination conditions. Every gesture is painstakingly documented with pertinent metadata so that academics may investigate hand movement dynamics' subtleties. Serving as a fundamental instrument for improving the discipline of gesture recognition technology, the ICVL dataset is especially important for applications in human-computer interaction, robotics, and sign language recognition.

3.6. NV Gesture Dataset

Designed to further studies in gesture identification and human-computer interaction, the skeleton-based hand gesture collection NV Gesture consists of various hand motions carried out by participants, each shown using thorough skeleton data reflecting the spatial locations of important joints in the hands and arms. Comprehensive analysis and model training are made possible by the variety of gestures in the dataset across many settings and conditions. Every movement is labelled with joint information and class labels, enabling the creation of strong machine-learning algorithms. Particularly useful for virtual reality, augmented reality, and assistive technologies, the NV Gesture dataset gives researchers a strong basis to progress the state of the art in gesture recognition systems.

3.7. NYU Dataset

A prominent skeleton-based collection meant to support gesture detection and computer vision research is the NYU hand gesture dataset. It captures thorough skeletal data that shows the locations of important joints in the hands and arms by featuring a variety of hand movements executed by several individuals. The dataset offers a useful background for building strong recognition models since it consists of many motions in many contexts and surroundings. Every gesture is painstakingly documented with joint coordinates and labels so that researchers can precisely examine hand motions' dynamics and subtleties. Applications in human-computer interaction, robotics, and augmented reality especially benefit from the NYU dataset, which is also a necessary tool for expanding the knowledge and acceptance of hand gestures in real-time applications.

3.8. NTU Dataset

Comprising a complete resource meant to further human action detection and gesture analysis, the NTU (Nanyang Technological University) skeleton-based hand gesture collection consists of a varied set of hand movements executed by several people, recorded with meticulous skeletal data tracking of important hand and body joints. Robust model training and evaluation are made possible by the large range of gestures in the dataset in many contexts and configurations. Every gesture is exactly marked with joint coordinates and action labels, allowing scientists to investigate hand motion spatial and temporal dynamics. The development of sophisticated gesture recognition systems depends on the NTU dataset since it is especially useful for human-computer interaction, robotics, and sign language recognition applications.

3.9. UCF-Kinetic Dataset

Designed to support study in gesture identification and motion analysis, the UCF-Kinetic dataset is a specific skeleton-based hand gesture dataset. It consists of a great range of hand movements executed by several people, each shown by comprehensive skeletal data capturing the positions of important joints in the hands and arms. The dataset offers a rich background for building strong machine-learning models since it comprises many motions in many contexts. Every gesture is painstakingly documented with joint coordinates and action labels so that the dynamics and subtleties of hand motions may be closely examined. Particularly useful for applications in human-computer interaction, sports analytics, and virtual reality, UCF-Kinetic is a vital tool for developing the area of gesture detection and thereby facilitating the creation of more interactive and intuitive systems.

3.10. UTKinetic Dataset

Comprising a complete skeleton-based hand gesture dataset used to improve studies in gesture detection and human motion analysis, the UTKinetic dataset shows a great range of hand movements carried out by several participants, recorded with meticulous skeletal data tracking of important hand and upper body joints. The dataset lets one explore gesture dynamics and interactions extensively by including a variety of motions in several contexts. Every movement is meticulously labelled with joint coordinates and action labels so that researchers may create and assess strong machine-learning models. Applications in human-computer interaction, robotics, and virtual reality especially benefit from UTKinetic, which is also a necessary tool for enhancing user involvement in interactive applications, thereby augmenting gesture detection systems' capabilities.

3.11. Florence 3-D Action Dataset

Aiming to further research in movement detection and gesture analysis, the comprehensive skeleton-based hand gesture collection Florence 3-D movement consists of a wide spectrum of motions carried out by several people, each shown by thorough skeletal data capturing the positions of important joints in the body and hands. The dataset offers important new perspectives on human motion dynamics by including several gestures and motions in three-dimensional space. Every motion is painstakingly documented with joint coordinates and action labels so that researchers may examine the subtleties of hand gestures in connection to complete-body movements. Particularly helpful for human-computer interaction, animation, and robotics applications, the Florence 3-D Action dataset is a vital tool for creating advanced gesture detection systems and enhancing knowledge of intricate human motions.

4. Proposed Methodology

In the proposed system, there may be numerous steps for hand gesture detection and recognition utilizing deep learning of MediaPipe. Hand gesture photos would first be gathered, labelled, and preprocessed into a dataset. Resizing, normalization, and augmentation are a few preprocessing techniques that could increase the dataset’s quality and diversity. The choice of a MediaPipe architecture would then be made based on the conditions and limitations of the application. For instance, a lightweight MediaPipe architecture could be selected if the system is meant to be used on mobile devices. Alternatively, a MediaPipe architecture could be selected if the system is designed for high accuracy. The proposed system may have numerous hand gesture detection and recognition steps utilizing MediaPipe and OpenCV. Hand gesture photos would first be gathered, labelled, and preprocessed into a dataset. Resizing, normalization, and augmentation are a few preprocessing techniques that could increase the dataset's quality and diversity.

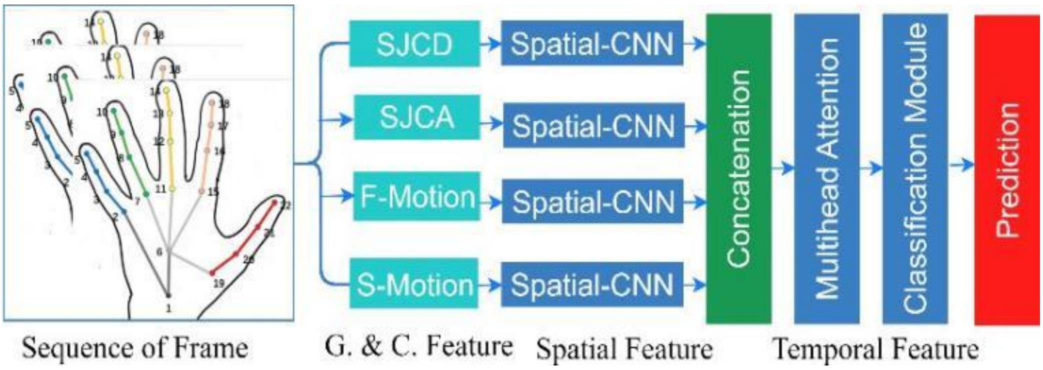

Figure 1: Proposed Working Flow Architecture [67]

4.1. Feature Extraction

Feature extraction is building informative features from a given dataset, enabling the subsequent learning and generalization steps in various machine learning domains. It converts raw data into numerical features that can be processed while maintaining the information in the original data set.

Selected Joint-Coordinate Distance (SJCD): We are considering the hand's tip and palm points to compute the distance and angle characteristics. For further clarification, each of the 32 frames of a movie may be characterized as a T, with N joints per frame in this example, N=22 in each frame. 3D coordinates for a certain frame.

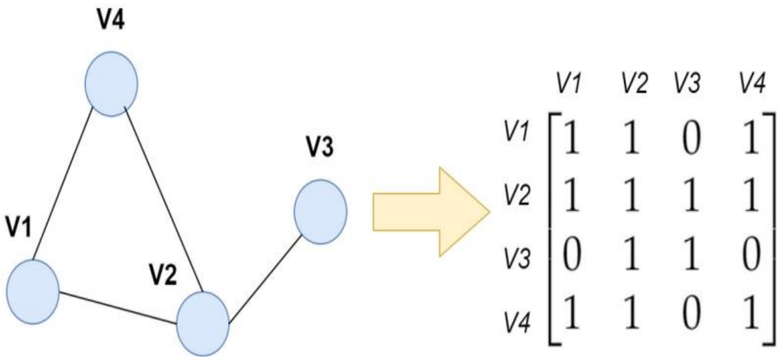

Figure 2: Selected Joint-Coordinate Distance (SJCD)

Here $\|\vec{J}^{T}_{i}\vec{J}^{T}_{j}\|$ $(j \neq j)$ represents the distance between J^{T}_{j} (4)

Because of the selection process, the final feature vector of the SJCD is formed as a dimension vector, which has a modest size. Equation 1 was used to determine the distance feature. We obtained the nine distances in the first raw by demonstrating the distance between each point and the palm or centre of the hand. After that, we determined the separation between each joint and every other point.

Global motion scales can differ throughout gestures, but they can also differ inside a single move. The temporal difference can have a quicker or slower scale [12]. Therefore, studying the global motion characteristics of slower and faster motion can be more beneficial. Here, we addressed both to address this problem, T.

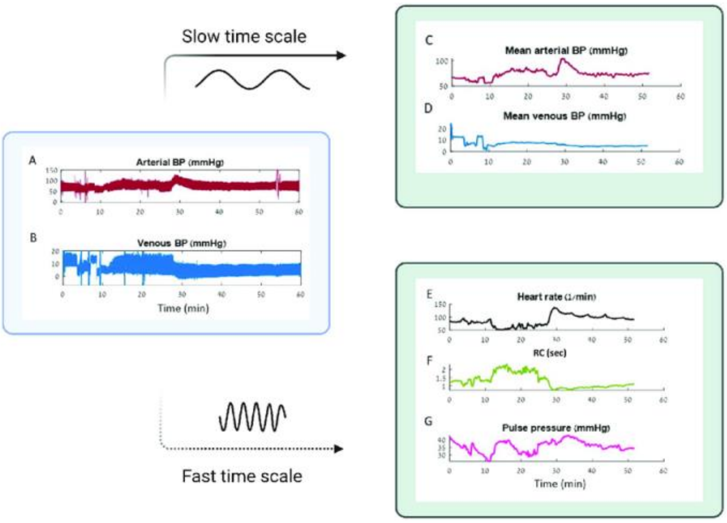

Figure 3: Motion Features

Here, GM^T, M^T_{Fast}, and M^T_{Slow} are represented by the global (5)

$$M^{[1,\ldots,T-1]}_{Slow}, M^{[1,\ldots,T/2]}_{Fast} \tag{6}$$

4.2. Spatial-CNN Model

The placement of the joint may change dynamically depending on the gesture; however, sometimes, it may be practically constant.

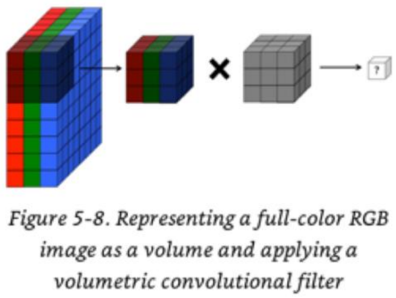

Figure 4: Temporal Context Enhancing CNN Module

The dynamic character of joint positions makes hand movements in stroke therapy difficult [51]. The individual gesture will affect the joint locations, which can vary greatly and provide difficulty in precisely identifying and characterizing these motions [52]. Although this is not always the case, particularly in rehabilitation situations where patients may have different degrees of mobility and coordination, traditional deep learning designs sometimes presume that all joints are connected [53]. We considered the dynamic interdependencies among joint locations [54]. We used a Spatial Convolutional Neural Network (CNN) model meant to improve the frame-wise representation of hand motions to address these problems efficiently [55]. The architecture is designed to automatically find correlations among joints through convolutional layers, allowing the model to fit individual patient variances [56].

Preparing and Augmenting Data: We extensively pre-processed our input data to guarantee its quality and consistency before model training. To improve the variety of the training set, this included standardizing the joint coordinates and using rotation, scaling, and flipping [57]. This augmentation is essential in rehabilitation environments since it enables the model to generalize more over several patient motions and compensatory techniques. Skeleton Representation & Noise Reduction: Noise, which can come from several sources, including differences in camera quality and lighting conditions, presents a major obstacle in gesture detection [58]. Our Spatial-CNN uses a skeleton-based representation of hand motions, which emphasizes the relative positions of joints instead of the raw pixel data from images, to help reduce these problems [59]. The model may efficiently lower noise by leveraging skeletal data while preserving important characteristics required for precise gesture classification [60]. Derived from the skeleton model, the final feature vector catches the spatial interactions among joints, enabling a strong description of gestures. Mathematically, we represent the concatenation of features over several frames as follows:

$$\text{Feature Vector}=\text{Joint}_1 \oplus \text{Joint}_2 \oplus \ldots \oplus \text{Joint}_N \quad (7)$$

Where $\oplus$ represents the concatenation operation, and N represents the total number of joints considered.

Attention Mechanism-Based Temporal Feature Learning: We proposed an attention strategy to learn the temporal aspects inherent in gesture sequences following building the spatial feature representation [61]. Attention processes let the model concentrate on particular joints and time steps, most likely to enable the recognition of a gesture. This is especially helpful in stroke therapy since patients' gestures' execution could differ in speed and fluidity [62]. Including attention in our model raises the interpretability and increases the recognition accuracy. Therapists can learn which elements of a patient's gesture performance could call for more attention during rehabilitation by examining the attention weights given to various joints and frames [63].

Model Assessment and Training: The model was trained on a large collection, including several hand movements by stroke victims. To reduce the loss function during training, we combined cross-entropy loss with an optimizer such as Adam [64]. We used k-fold cross-valuation to validate our method, guaranteeing that our model's performance was strong throughout several data subsets [65]. We assessed the model's performance using conventional criteria like accuracy, precision, recall, and F1-score. These measures give a whole picture of the accuracy with which the model can classify gestures and their dependability in a clinical environment [66]. Implementation for Real-Time Gesture Identification: Our model is included in an application allowing real-time gesture recognition once trained and certified. Stroke sufferers can use this program as a user-friendly interface to operate assistive tools with hand gestures. The system uses a trained Spatial-CNN model to read motions and camera input to gather joint data instantly.

Customer Comments and Ongoing Enhancement: The iterative development of the system depends on user feedback, including that of stroke victims and rehabilitation experts. We intend to extend the repertoire of gestures the system acknowledges by including user feedback systems to improve gesture recognition. We can improve the adaptability and efficiency of the model in assisting stroke recovery by always updating it with fresh data acquired from actual interactions. Our suggested approach emphasizes using cutting-edge deep learning methods, especially a spatial CNN paired with attention mechanisms, to improve hand gesture recognition for stroke recovery. We want to create a low-cost, efficient, non-invasive instrument that can greatly help stroke patient's rehabilitation by tackling the difficulties of dynamic joint placements and noise reduction using skeleton representations. This paper stresses the need for user-centred design in creating assistive technology and the possibilities of deep learning technologies in healthcare.

4.3. Multi-Head Attention Architecture

The query, key, and value of the input feature are the three completely linked layers utilized in tandem to generate the corresponding mapping information value. Moreover, represents the weight matrix of these mappings.

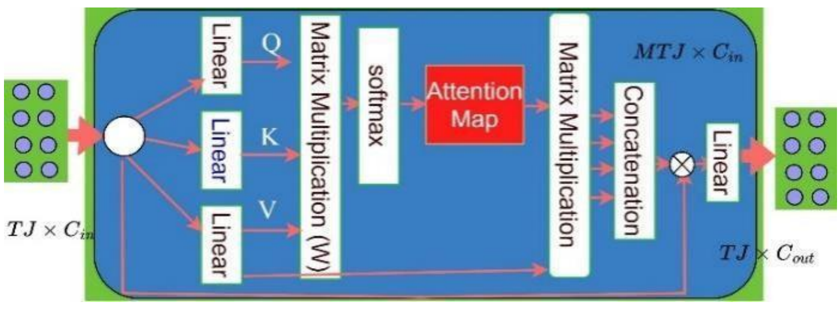

Figure 5: Proposed Attention Module [67]

Each stage operation is shown in detail in Fig. 5, where the dot product is operated using a query and key, followed by applying an activation function. Creating a value matrix and multiplying it by the attention map [3];[10];[28];[30];[31]. U represent the scale dot product between the query and key matrix, the inner matrix is indicated by ⟨., ⟩, and the dimension of the key vector is represented by d. The four output matrices were created similarly, and after concatenating the four features, we obtained the final feature. We propose a Multi-Head Attention design, as shown in Fig. 5. Effective capture of the complicated interactions among several joints during hand motion identification depends on this design. Especially in stroke rehabilitation, we may greatly increase the model's capacity to detect little variations in gestures by boosting the conventional attention mechanisms.

Multi-Head Attention's Conceptual Framework: The three basic components, the query, key, and value vectors, form the basis of the Multi-Head Attention method. We approach our concatenated feature vector as the first input feature for the attention model for every joint. Every head in the attention model independently analyzes these inputs to acquire unique representations, enabling the network to capture various facets of the input information.

Key, Query, and Value Computation: The three different representations, the query (Q), the key (K), and the value (V), are produced from the input feature. Whole connected layers help to accomplish these changes. Understanding how the position of one joint relates to the others during gesture execution depends on the interactions between several vectors.

Attractive Dot-Product Attention: Calculating the scaled dot-product attention, described numerically as follows, is the core of the attention mechanism:

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{8}$$

Here, the softmax function normalizes the scores to guarantee they are total to one, enabling us to understand them as probabilities. d_k is the dimension of the key vector.

Attention Weight Calculation: The dot product of the query and key vectors produces attention scores emphasizing the relevance of each joint's attributes about the others. This helps one to know which joints, most importantly, support a given gesture.

Generation of Output: Every head's output is formed from the weighted sum of the value vectors. Based on previous computed attention scores, this output catches the pertinent aspects. Concatenating the outputs from every head and then running a linear transformation yields the multi-head attention layer's final output. We use skip connections to concatenate the output of the attention model with the original feature input, hence strengthening the resilience of our model. This method stabilizes the learning process in addition to helping possibly lost data during attention processing be recovered. Retaining the original characteristics helps the model to learn the fundamental data from the input, and the subtle changes caught my attention.

Model Development and Instruction: The architecture integrates the Spatial-CNN mentioned with the Multi-Head Attention. Every attention layer inputs into a sequence of convolutional layers, which further handles the characteristics to identify the motions properly. The training approach aims to maximize a loss function that strikes a compromise between the computing efficiency needed for real-time applications and gesture recognition's accuracy.

Preparing a Dataset: To guarantee a broad range of motion and variety in execution speed, we compiled various hand motions executed by stroke patients. This dataset was extended to incorporate several environmental situations, strengthening the model's resilience.

Training Strategy: Using labelled data to reduce classification mistakes, the model was trained by combining many supervised learning methods. Given the model's complexity, we used early halting and dropout strategies to prevent overfitting.

Evaluation Tools: We used several criteria, including accuracy, precision, recall, and F1-score, to evaluate the success of our method. These measures revealed how well the model identified motions, especially for patients with different degrees of ability.

User Interface and Real-Time Implementation: The real-time application for stroke sufferers after training included the model. This useful tool for rehabilitation lets people operate assistive gadgets with hand gestures, therefore enabling control. Designed with accessibility in mind to serve people with perhaps restricted mobility or technological experience, the user interface is simple.

User Interaction: The program gathers video input constantly, runs it through the trained model, and converts identified movements into commands for the assistive tool. Users can witness the identified gesture and inspire participation using real-time feedback.

Continuous Learning: We created a feedback loop whereby users could comment on the accuracy of gesture detection, enhancing the model. Retraining and model fine-tuning will employ this information to guarantee the model adjusts to personal user needs and tastes.

Development of Gesture Recognition: The present approach emphasizes certain movements. Future research will broaden this repertory to incorporate more complicated movements, enabling a greater spectrum of interactions with assistive technologies.

Customization: Individual user customizing of the model could greatly increase accuracy. Among other methods, transfer learning might help modify the model depending on user-specific input.

Integration with Alternative Rehabilitation Strategies: Future system versions could interact with various rehabilitation modalities, such as gamified workouts or virtual reality, to improve patient involvement and recovery results. Finally, combining spatial-CNN characteristics with Multi-Head Attention architecture offers a strong basis for hand motion recognition in stroke rehabilitation. This method solves the complexity of combined interactions and improves system adaptability and efficiency. Our approach seeks to empower stroke victims by providing a non-invasive, effective tool for interacting with assistive equipment, aiding their recovery path.

4.4. Classification Module

Following the Multi-Head Attention network's generation of temporal features, we proceed to the classification module, which is essential in assuring correct gesture classification and improving the final feature representation. The module for categorization is meant to convert the output of the attention mechanism into definite class labels corresponding with certain hand gestures.

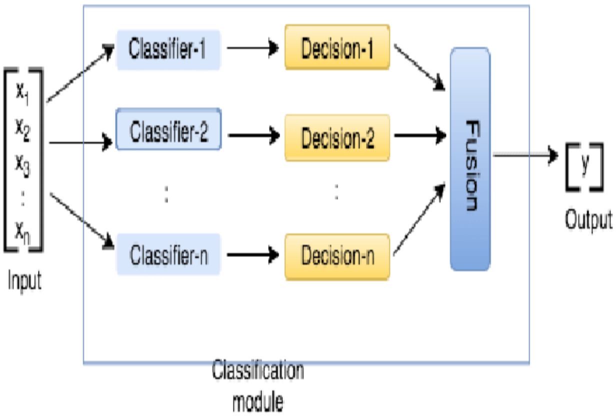

Figure 6: Classification Module

4.5. Structure of the Classification Module

As Fig. 6 shows, the classification module consists of several key components:

Temporal Feature Refinement: The categorization module receives temporal characteristics generated by the attention network. We use an average filter to guarantee these features are ready for classification. By helping to smooth out the temporal noise and fluctuations resulting from different gesture executions, this filter offers a more constant representation for classification.

CNN Layers (Convolutional Neural Network): The foundation of the classification module consists of a sequence of convolutional layers. Different spatial hierarchies and patterns in the feature maps are captured by variable filter sizes built into these layers. Several convolutional layers let the machine learn even more intricate hand gesture representations. Usually, ReLU (Rectified Linear Unit), an activation function, follows each convolutional layer to provide non-linearity in the model. This is crucial to let the network learn intricate mappings from the input features to the output classes.

Pooling Layers: Pooling layers, especially max pooling or average pooling, are scattered amongst the coevolutionary layers. Downsampling the feature maps depends on these layers, lowering the dimensionality and computational load while preserving the most salient characteristics. Applying pooling will help us concentrate on the most important features of the gesture data, facilitating more effective classification.

Completely Connected Layer: The output is flattened and processed through one or more fully connected layers following the series of convolutional and pooling layers. These layers do the last categorization and further combine the learnt characteristics. Designed with a softmax activation function, the last fully connected layer generates probability matching to every gesture type.

Loss Function and Optimization: Usually employing a suitable loss function, typically categorical cross-entropy, which measures the discrepancy between the predicted probability and the actual class labels, the classification module is trained. Based on the computed loss during backpropagation, an optimizer such as Adam or SGD (Stochastic Gradient Descent) is used to modify the network's weights.

Integration with the Attention Mechanism: Improving the model's performance depends mostly on integrating the Multi-Head Attention method with the classification module. While the classification module reduces this data into useful insights, the attention layer helps the model concentrate on the most relevant joints and frames. In stroke rehabilitation, where exact gesture recognition can greatly affect the efficacy of assistive equipment, this synergy is especially crucial.

4.6. Training the Classification Module

Dataset Preparation: Stroke patient's range of hand gestures makes up the training ground for the categorization module. This dataset is painstakingly annotated to guarantee that every gesture matches a given class. We use data augmentation methods like rotation, translation, and scaling to replicate real-world variances in gesture execution, strengthening the model.

Training Procedure: Through iteratively changing its parameters, the model learns to reduce the loss function throughout the training phase. We stabilize and speed up the training process using batch normalization, among other methods. Dropout layers can also be included to reduce overfitting and guarantee that the model generalizes to unprocessed input.

Test and Validation: We use a hold-out validation technique whereby a subset of the data is set aside to assess the model's performance during training. We compute accuracy, precision, recall, and F1-score metrics to evaluate gesture classification performance. Validation results help to fine-tune the model to maximize performance.

Real-Time Gesture Recognition Implementation: The categorization module is included in a real-time gesture recognition system once the training procedure is over. This system processes video frames to identify and classify hand motions on demand, so it functions well with the input from a camera or depth sensor.

User Interaction: Stroke patients use the system by making specific motions recorded in real-time. These gestures are handled by the categorization module, which also offers instantaneous visual or aural feedback. The engagement of patients in their rehabilitation activities depends on this interactive feedback loop.

Assistive Device Control: The identified gestures correspond to particular commands used to operate assistive tools. For instance, a "grab" gesture might be connected to turning on a robotic arm, while a "swipe" motion might be used to negotiate a user interface. A responsive user experience depends on this mapping.

Ongoing Learning and Improvement: We introduce a continuous learning framework to improve the model's adaptability. User comments help to pinpoint any misclassifications or difficulties they have. Retraining and model refinement depend much on this feedback, which guarantees that the model develops in line with user preferences and demands.

Gesture Expansion: Future versions of the classification module will seek to increase the gesture library to include more sophisticated gestures. Capture finer features in hand motions, and this will entail gathering more data and maybe including more sophisticated modelling approaches.

User-Centric Customization: Customizing on a personal level will be investigated to let the system fit particular consumers. Using transfer learning methods allows the model to be adjusted depending on a user's particular gesture patterns, enhancing accuracy and user satisfaction.

Integration with Multiple Modalities: We will look into combining our gesture detection technology with various therapeutic approaches, such as augmented reality or virtual reality. Such integration might improve patient involvement and help rehabilitation activities to be more fun and efficient. In essence, we concluded that the crucially important part of our proposed system is the classification module, which converts Multi-Head Attention network temporal data into useful insights for gesture detection. We seek great accuracy in identifying hand motions essential for stroke rehabilitation using convolutional layers, pooling techniques, and fully connected networks. This approach solves the technical problems of gesture detection and underlines the need for user involvement and adaptability in creating assistive devices for stroke patients.

5. Experimental Evaluation

The MediaPipe framework was trained with two classifiers, namely model 0 and model 1. Classifier 1 has produced the highest accuracy of 94 percent.

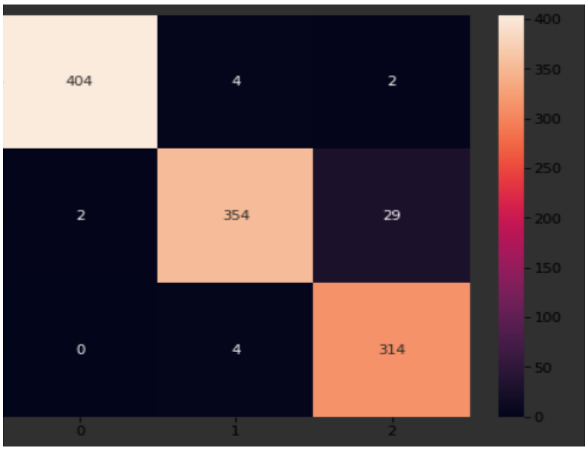

Figure 7: Performance Accuracy with Shrec'17 Dataset

Summary of Experimental Evaluation: We investigated real-time hand gesture detection for stroke rehabilitation using the MediaPipe framework and two classifiers, Model 0 and Model 1. Model 1 produced the best performance with a noteworthy 94% accuracy rate. Emphasizing performance measures across two separate datasets, the DHGD dataset and the SHREC'17 dataset, this part offers a comprehensive summary of our experimental evaluations.

5.1. Performance Evaluation with DHGD Dataset

We benchmarked our model against many current state-of-the-art techniques using a thorough evaluation leveraging the DHGD dataset. Figure 7 summarizes the data; using our suggested model, the accuracy is 92.21%. This outcome is a major development above many previous models.

Comparative Performance: Different models revealed different accuracy rates; many used skeletal and depth data and obtained an average performance of roughly 8.00%. Other approaches using joint similarities struggled to identify hand motions in real time properly and obtained a maximum accuracy of 83.35%.

Models Using RNNs: A small improvement was shown by some models using RNNs, reaching up to 84.68% correctness. Still, the performance of our suggested model exceeds these numbers, suggesting a more efficient recognition method.

High-Performing Models: While the Hif3d model virtually followed with 90.48%, the DG-STA model achieved 91.00% accuracy. Emphasizing the need for our model's exceptional accuracy, these findings show the competitive scene of gesture recognition systems. Our results imply that, especially for applications aiming at stroke rehabilitation, integrating advanced deep learning methods and the MediaPipe architecture greatly improves the accuracy of hand gesture recognition.

5.2. Performance Evaluation with SHREC'17 Dataset

We tested our model using the SHREC'17 dataset to confirm its resilience further and summarizes the outcomes and shows how our model performs very well, with an accuracy rate of 95.00%. This success emphasizes how well our method performs in several scenarios of gesture recognition.

Benchmarking Against Other Models: Our suggested approach differs from other models. The DG-STA model reported 94.40%; the STA-RES-TCN model attained an accuracy of 93.60%. Furthermore, as observed by Jiang et al., performance accuracy was 94.60%. These comparisons highlight how well our method beats current technology.

Significance of Results: Our model's high accuracy in both datasets shows its possible use in practical rehabilitation environments. The improved accuracy helps ensure more consistent interactions between stroke patients and assistive equipment, enhancing patient involvement and rehabilitation results.

6. Conclusion

MediaPipe and OpenCV, taken together, provide a potent hand gesture recognition solution with real-time detection and tracking of hand motions. This technology gives developers the tools and frameworks to include hand gesture detection features in their products. MediaPipe and OpenCV offer a strong framework for creating hand gesture recognition apps; this technology can transform human-computer interaction and enable consumers with fresh approaches to connect with digital systems. The experimental evaluation validates that our suggested gesture recognition model performs better than current state-of-the-art models. The outstanding accuracy measures achieved in the DHGD and SHREC'17 datasets emphasize the possibility of our method to enhance stroke rehabilitation procedures. We have created a solution that not only satisfies real-time gesture recognition but also tackles the particular difficulties experienced by stroke survivors by using deep learning approaches and MediaPipe framework features.

Deep learning, especially MediaPipe with OpenCV, showcases great promise for improving hand gesture detection in stroke therapy. We want to empower stroke victims by using a low-cost, non-invasive solution so they may communicate with assistive gadgets using natural hand motions. The great accuracy attained in our classification models underlines the efficiency of our suggested approach, which combines a strong attention-based architecture with sophisticated feature extraction methods. Expanding the gesture detection capabilities to fit a wider variety of motions and improving user engagement through gamified interfaces will be essential as we look ahead to allow for these developments, which will not only enhance the rehabilitation process but also enable stroke survivors more freedom. Our results imply that including these technologies in current healthcare systems might help to enable a more customized and successful rehabilitation procedure. Realizing the full possibilities of this creative approach in stroke recovery will depend on ongoing research and cooperation between technologists and healthcare professionals, enhancing the quality of life for patients and opening the path for more easily available rehabilitation solutions.

Acknowledgement: I am deeply grateful to all my co-authors for their invaluable contributions and unwavering support throughout this research endeavor

Data Availability Statement: This research is based on data related to Real-Time Hand Gesture Recognition for Stroke Rehabilitation using a Deep Learning Approach. The dataset includes key metrics such as views and timestamps, which were utilized for analysis.

Funding Statement: This manuscript and the associated research were prepared without any external funding or financial assistance.

Conflicts of Interest Statement: The author(s) declare no conflicts of interest concerning the publication of this manuscript. All citations and references have been properly acknowledged and documented.

Ethics and Consent Statement: Ethical approval for this study was obtained from the relevant institutional review board, and informed consent was secured from all participants and associated organizations involved in the data collection process.

References

1. F. Zhigang, "Computer gesture input and its application in human computer interaction," Mini-micro Systems, vol. 6, no.4, pp. 418-421, 1999.
2. S. Mitra and T. Acharya, "Gesture recognition: A survey," IEEE Trans. Syst. Man Cybern. Part C Appl. Rev, vol. 37, no. 3, pp. 311-324, 2007.
3. M. K. Ahuja and A. Singh, "Static vision based Hand Gesture recognition using principal component analysis," in 2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE), Amritsar, India, pp. 402-406, 2015.
4. R. K. Kramer, C. Majidi, R. Sahai, and R. J. Wood, "Soft curvature sensors for joint angle proprioception," in 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, United States of America, pp. 1919-1926, 2011.

5. E. Jesperson and M. R. Neuman, “A thin film strain gauge angular displacement sensor for measuring finger joint angles,” in Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Los Angeles, United States of America, 1988.
6. E. Fujiwara, M. F. M. dos Santos, and C. K. Suzuki, “Flexible optical fiber bending transducer for application in glove-based sensors,” IEEE Sensors Journal, vol. 14, no. 10, pp. 3631-3636, 2014.
7. A. S. M. Miah, M. A. M. Hasan, S.-W. Jang, H.-S. Lee, and J. Shin, “Multi-stream general and graph-based deep neural networks for skeleton-based sign language recognition,” Electronics (Basel), vol. 12, no. 13, pp. 1-15, 2023.
8. A. S. M. Miah, M. A. M. Hasan, J. Shin, Y. Okuyama, and Y. Tomioka, “Multistage spatial attention-based neural network for hand gesture recognition,” Computers, vol. 12, no. 1, p. 13, 2023.
9. A. Saleh Musa Miah, J. Shin, A. M. Md, and M. A. Hasan, “BenSignNet: Bengali Sign Language Alphabet Recognition Using Concatenated Segmentation and Convolutional Neural Network,” Applied Sciences, vol. 12, no. 8, p. 3933, 2022.
10. A. S. Musa Miah, J. Shin, M. A. M. Hasan, M. A. Rahim, and Y. Okuyama, “Rotation, translation and scale invariant sign word recognition using deep learning,” Computer Systems Science and Engineering, vol. 44, no. 3, pp. 2521–2536, 2023.
11. T. Aditya, T. Bertram, G. Frederic, S. Didier,“A probabilistic combination of CNN and RNN estimates for hand gesture based interaction in car. International symposium on mixed and augmented reality ISMAR-adjunct,” Washington, United States of America, pp. 1–6, 2017.
12. Y. Fuchang Hijian Fuchang and S. Hijian, “Research on static hand gesture recognition technology for human computer interaction system,” International conference on intelligent transportation; Piscatway, New Jersy, United States of America, pp. 459–463, 2016.
13. E. B. Marco, M. Cristhian, J. A. Z. Jonathan, G. J. Andrés, E. A. Carlos, Z. Patricio, S. Marco, B. P. Freddy, and P. María, “Real-time hand gesture recognition using the Myo armband and muscle activity detection,” in Second Ecuador Technical Chapters Meeting (ETCM), Salinas, Ecuador, 2017.
14. V. Viyas, “Hand pantomime apperception for robotic arm control”. International Conference of Electronics, communication and Aerospace Technology ICECA, Coimbatore, Tamil Nadu, India, pp. 120–125, 2017.
15. Q. D. Smedt, H. Wannous, and J.-P. Vandeborre, “Dynamic Hand Gesture Recognition using Skeleton-based Features,” in IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPRW, Las Vegas, Nevada, United States of America, 2016.
16. C. Xi, L. Lihua Xi, L. Chen, and T. Lihua, “Hand gesture recognition based on wavelet invariant moments,” in International Symposium on multimedia ISM; Taichung, Taiwan, pp. 459–464, 2017.
17. Y. Zhang, C. Cao, J. Cheng, and H. Lu, “EgoGesture: A new dataset and benchmark for egocentric hand gesture recognition,” IEEE Trans. Multimedia, vol. 20, no. 5, pp. 1038–1050, 2018.
18. S.-H. Chen, A. Hernawan, Y.-S. Lee, and J.-C. Wang, “Hand gesture recognition based on Bayesian sensing hidden Markov models and Bhattacharyya divergence,” in 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 2017.
19. Q. Zhang, M. Yang, Q. Zheng, and X. Zhang, “Segmentation of hand gesture based on dark channel prior in projector-camera system,” in 2017 IEEE/CIC International Conference on Communications in China (ICCC), Qingdao, China, 2017.
20. A. Laura, D. Paolo Laura, and M. S. Angelo, “A survey of glove-based systems and their applications,” IEEE Transactions on Systems, Man, and Cybernetics, vol. 38, no. 4, pp. 461–480, 2008.
21. J. Jinxing, Y. Jinxing, P. Jianhong, and L. Jun, “sEMG-based continuous hand gesture recognition using GMM-HMM and threshold model. International conference on robotics and biomimetics ROBIO; 2017,” Macau, Macao, pp. 1509–1514, 2017.
22. P. J. Mansfield and D. A. Neumann, “Essentials of Kinesiology for the Physical Therapist Assistant”. Elsevier Health Sciences, St. Louis, Missouri, United States of America, 2008.
23. Q. Sun, E. Gonzalez, and B. Abadines, “A wearable sensor based hand movement rehabilitation and evaluation system,” in 2017 Eleventh International Conference on Sensing Technology (ICST), Sydney, Australia, 2017.
24. A. A. A. Shehhi, R. F. A. Mamari, and P. Ranjith, "An Impact of Internet Technology on the Social Behavior Issues Among Children in Al Batinah Governorate During 2020-2021," International Journal of Information Technology, Research and Applications, vol. 1, no. 1, pp. 25-32, 2022.
25. A. J. Obaid, S. S. Rajest, S. S. Priscila, T. Shynu, and S. A. Ettyem, “Dense convolution neural network for lung cancer classification and staging of the diseases using NSCLC images,” in Proceedings of Data Analytics and Management, Singapore; Singapore: Springer Nature, pp. 361–372, 2023.
26. A. K. Singh, I. R. Khan, S. Khan, K. Pant, S. Debnath, and S. Miah, “Multichannel CNN model for biomedical entity reorganization,” Biomed Res. Int., vol. 2022, no.3, p.10, 2022.
27. C. Newton, M. Jayakkumar, and P. Ranjith, "xMIDURR: A Quality of Service Technique to Save Energy in Mobile Ad Hoc Networks," International Journal of Control Theory and Applications, vol. 9, no. 27, pp. 561-567, 2016.

28. C. Newton, P. Ranjith, and M. J. Aragon, "A Mathematical Model for LOCA-GAMNET," International Journal in IT and Engineering, vol. 4, no. 11, pp. 18-22, 2016.
29. C. Newton, P. Ranjith, and M. J. Aragon, "Location Aware Routing for MANET Using Genetic Algorithm," International Journal of Control Theory and Applications, vol. 9, no. 27, pp. 219-225, 2016.
30. C. Newton, P. Ranjith, and M. J. Aragon, "Mobility Prediction in MANET Routing Using Genetic Algorithm," International Journal of Innovations & Advancement in Computer Science, vol. 3, no. 10, pp. 28-33, Dec. 2014.
31. C. Newton, P. Ranjith, and M. J. Aragon, "Optimizing Routing in MANET Using Genetic Algorithms," International Journal of Applied Engineering Research, vol. 10, no.11, pp. 40124-40129, 2015.
32. C. Newton, P. Ranjith, and S. K. Mohideen, "OPTANT - Optimized Ant Colony Routing for Mobile Ad-Hoc Networks," International Journal of Advanced Research Trends in Engineering and Technology, vol. 4, no. 8, p.12, 2017.
33. C. P. Ranjith, K. Natarajan, S. Madhuri, M. T. Ramakrishna, C. R. Bhat, and V. K. Venkatesan, "Image Processing Using Feature-Based Segmentation Techniques for the Analysis of Medical Images," Engineering Proceedings, vol. 59, no. 1, p.11, 2023.
34. G. Gnanaguru, S. S. Priscila, M. Sakthivanitha, S. Radhakrishnan, S. S. Rajest, and S. Singh, “Thorough analysis of deep learning methods for diagnosis of COVID-19 CT images,” in Advances in Medical Technologies and Clinical Practice, IGI Global, USA, pp. 46–65, 2024.
35. G. Gowthami and S. S. Priscila, “Tuna swarm optimisation-based feature selection and deep multimodal-sequential-hierarchical progressive network for network intrusion detection approach,” Int. J. Crit. Comput.-based Syst., vol. 10, no. 4, pp. 355–374, 2023.
36. J. Selwyn and C. Prasanna Ranjith, "Towards Designing a Planet Walk Simulation in a Controlled Environment," International Journal of Data Informatics and Intelligent Computing, vol. 2, no. 1, pp. 70-77, 2023.
37. K. Sattar, T. Ahmad, H. M. Abdulghani, S. Khan, J. John, and S. A. Meo, “Social networking in medical schools: Medical student’s viewpoint,” Biomed Res., vol. 27, no. 4, pp. 1378-84, 2016.
38. M. J. Antony, B. P. Sankaralingam, S. Khan, A. Almjally, N. A. Almujally, and R. K. Mahendran, “Brain–computer interface: The HOL–SSA decomposition and two-phase classification on the HGD EEG data,” Diagnostics, vol. 13, no. 17, p. 2852, 2023.
39. M. S. Rao, S. Modi, R. Singh, K. L. Prasanna, S. Khan, and C. Ushapriya, “Integration of cloud computing, IoT, and big data for the development of a novel smart agriculture model,” presented at the 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023.
40. N. Rajesh and C. Prasanna Ranjith, "Analysis of Origin, Risk Factors Influencing COVID-19 Cases in India and Its Prediction Using Ensemble Learning," International Journal of System Assurance Engineering and Management, 2021. Press.
41. N. Rajesh, A. Irudayasamy, M. S. K. Mohideen, and P. Ranjith, "Classification of Vital Genetic Syndromes Associated with Diabetes Using ANN-Based CapsNet Approach," International Journal of e-Collaboration (IJeC), vol. 18, no. 3, pp. 1-18, 2022. doi: 10.4018/IJeC.3071331.
42. P. P. Anand, U. K. Kanike, P. Paramasivan, S. S. Rajest, R. Regin, and S. S. Priscila, “Embracing Industry 5.0: Pioneering Next-Generation Technology for a Flourishing Human Experience and Societal Advancement,” FMDB Transactions on Sustainable Social Sciences Letters, vol.1, no. 1, pp. 43–55, 2023.
43. P. Ranjith, B. M. Hardas, M. S. K. Mohideen, N. N. Raj, N. R. Robert, and P. Mohan, "Robust Deep Learning Empowered Real Time Object Detection for Unmanned Aerial Vehicles Based Surveillance Applications," Journal of Mobile Multimedia, vol. 19, no. 2, pp. 451–476, 2022. doi: 10.13052/jmm1550-4646.1925.
44. P. Ranjith, F. Sheikha, N. Noora, and F. Farma, "Smart Video-Based Sign Language App: Impacts on Communication for Deaf and Dumb Individuals," International Journal of Innovative Research in Advanced Engineering, vol. 8, no.10, pp. 277-282, 2021.
45. R. Natarajan, N. Mahadev, B. S. Alfurhood, C. Prasanna Ranjith, J. Zaki, and M. N. Manu, "Optimizing Radio Access in 5G Vehicle Networks Using Novel Machine Learning-Driven Resource Management," Optical and Quantum Electronics, vol. 55, no. 14, p.10, 2023.
46. R. Regin, Shynu, S. R. George, M. Bhattacharya, D. Datta, and S. S. Priscila, “Development of predictive model of diabetic using supervised machine learning classification algorithm of ensemble voting,” Int. J. Bioinform. Res. Appl., vol. 19, no. 3, pp. 151-169, 2023.
47. R. Venkatarathinam, R. Sivakami, C. Prasanna Ranjith, M. T. R., E. Mohan, and V. V. Kumar, "Ensemble of Homogenous and Heterogeneous Classifiers using K-Fold Cross Validation with Reduced Entropy," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 11, no. 8s, pp. 315–324, 2023.
48. S. Gayathri D., "Adapting the Effect of Asynchronous Impulsive Noise using Interleaving Based on Compound Signal Classification in Broad PLC," International Journal of Innovative Technology and Exploring Engineering, vol. 9, no. 2, pp. 2278–3075, 2019.

49. S. Gayathri D., "Optimal Impedance Matching System for Broadband PLC for Maximizing the Signal to Noise Ratio (SNR) and Data Rate," International Journal of Intelligent Systems and Applications in Engineering, vol. 12, no. 20s, pp. 222–229, 2024.
50. S. Khan and A. Alfaifi, “Modeling of coronavirus behavior to predict its spread,” Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 5, pp. 394-399, 2020. doi: 10.14569/IJACSA.2020.0110552.
51. S. Khan and S. Alqahtani, “Hybrid machine learning models to detect signs of depression,” Multimed. Tools Appl., vol. 83, no. 13, pp. 38819–38837, 2023.
52. S. Khan et al., “Transformer architecture-based transfer learning for politeness prediction in conversation,” Sustainability, vol. 15, no. 14, p. 10828, 2023.
53. S. Khan, “Artificial intelligence virtual assistants (chatbots) are innovative investigators,” Int. J. Comput. Sci. Netw. Secur., vol. 20, no. 2, pp. 93-98, 2020.
54. S. Khan, “Modern internet of things as a challenge for higher education,” Int. J. Comput. Sci. Netw. Secur., vol. 18, no. 12, pp. 34-41, 2018.
55. S. Khan, “Study factors for student performance applying data mining regression model approach,” Int. J. Comput. Sci. Netw. Secur., vol. 21, no. 2, pp. 188-192, 2021.
56. S. R. S. Steffi, R. Rajest, T. Shynu, and S. S. Priscila, “Analysis of an Interview Based on Emotion Detection Using Convolutional Neural Networks,” Central Asian Journal of Theoretical and Applied Science, vol. 4, no. 6, pp. 78–102, 2023.
57. S. S. Priscila and A. Jayanthiladevi, “A study on different hybrid deep learning approaches to forecast air pollution concentration of particulate matter,” in 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023.
58. S. S. Priscila and S. S. Rajest, “An Improvised Virtual Queue Algorithm to Manipulate the Congestion in High-Speed Network”,” Central Asian Journal of Medical and Natural Science, vol. 3, no. 6, pp. 343–360, 2022.
59. S. S. Priscila, D. Celin Pappa, M. S. Banu, E. S. Soji, A. T. A. Christus, and V. S. Kumar, “Technological frontier on hybrid deep learning paradigm for global air quality intelligence,” in Cross-Industry AI Applications, IGI Global, USA, pp. 144–162, 2024.
60. S. S. Priscila, E. S. Soji, N. Hossó, P. Paramasivan, and S. Suman Rajest, “Digital Realms and Mental Health: Examining the Influence of Online Learning Systems on Students,” FMDB Transactions on Sustainable Techno Learning, vol. 1, no. 3, pp. 156–164, 2023.
61. S. S. Priscila, S. S. Rajest, R. Regin, and T. Shynu, “Classification of Satellite Photographs Utilizing the K-Nearest Neighbor Algorithm,” Central Asian Journal of Mathematical Theory and Computer Sciences, vol. 4, no. 6, pp. 53–71, 2023.
62. S. S. Priscila, S. S. Rajest, S. N. Tadiboina, R. Regin, and S. András, “Analysis of Machine Learning and Deep Learning Methods for Superstore Sales Prediction,” FMDB Transactions on Sustainable Computer Letters, vol. 1, no. 1, pp. 1–11, 2023.
63. S. S. Rajest, S. Silvia Priscila, R. Regin, T. Shynu, and R. Steffi, “Application of Machine Learning to the Process of Crop Selection Based on Land Dataset,” International Journal on Orange Technologies, vol. 5, no. 6, pp. 91–112, 2023.
64. S. Silvia Priscila, S. Rajest, R. Regin, T. Shynu, and R. Steffi, “Classification of Satellite Photographs Utilizing the K-Nearest Neighbor Algorithm,” Central Asian Journal of Mathematical Theory and Computer Sciences, vol. 4, no. 6, pp. 53–71, 2023.
65. T. Shynu, A. J. Singh, B. Rajest, S. S. Regin, and R. Priscila, “Sustainable intelligent outbreak with self-directed learning system and feature extraction approach in technology,” International Journal of Intelligent Engineering Informatics, vol. 10, no. 6, pp.484-503, 2022.
66. V. Vinoth Kumar, U. Padmavathi, C. Prasanna Ranjith, J. Balaji, C. N. S. Vinoth Kumar, "An Elixir for Blockchain Scalability with Channel Based Clustered Sharding," Scalable Computing: Practice and Experience, vol. 25, no. 2, p.13, 2024.
67. A. S. Musa Miah, J. Shin, M. Al Mehedi Hasan, Y. Fujimoto and A. Nobuyoshi, "Skeleton-based Hand Gesture Recognition using Geometric Features and Spatio-Temporal Deep Learning Approach," 2023 11th European Workshop on Visual Information Processing (EUVIP), Gjovik, Norway, pp. 1-6, 2023.