SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I11P101
Volume 11 Issue 11, 1-11, November 2024

Original Article

Intelligent Robotic Medical Assistive Device for Elderly Individuals Support

Chandrashekhar Kumar[1], T. Muthumanickam[2], T. Sheela[3]

[1,2,3]*Department of Electronics and Communication Engineering,*
Vinayaka Mission's Kirupananda Variyar Engineering College, Tamil Nadu, India.

[1]*Corresponding Author : chandrashekharbandhu@gmail.com*

Received: 01 September 2024 Revised: 03 October 2024 Accepted: 01 November 2024 Published: 03 December 2024

Abstract *- This paper discusses trends in technology like Artificial Intelligence and machine learning algorithms in developing intelligent robotic systems. It focuses on attribute-picking points, classification, and fuzzy rule-based decision-making in settings for robot actuation planning. The system uses a Natural Language Processing-based User Interface, cameras, and image processing modules. It proposes two new feature selection and classification algorithms, Intelligent Voice to Text Conversion and Fuzzy Temporal Rule-based Semantic Analysis Algorithm. The system also introduces three new algorithms for object detection and grasping. Robots with assistive technology may help with senior or geriatric care, but their ability to track objects, estimate motion, and estimate poses a barrier. Although researchers have suggested real-time posture estimates as a dependable option, traditional tracking techniques still highly value stance. The volume of data, extensive processing duties, and start-up all contribute to the complexity of real-time tracking. An innovative mobile robot system designed to assist has been put forth to enable elderly individuals to live longer, safer lives in their homes. The study aims to address these concerns and develop technology that meets the needs of senior citizens and geriatrics.*

Keywords *- Artificial Intelligence, Machine Learning, Medical Assistive, Intelligent Systems, Robotics, Elderly Support.*

1. Introduction

This paper focuses on applying machinery artificial intelligence-based algorithms in developing intelligent robotic systems. It focuses on attribute selection, classification, and fuzziness in rule-based judgment in robot actuation planning environments [1].

Robots are increasingly used in medical assistance systems, particularly for geriatric patients. Non-industrial robots, such as assistive and non-assistive robots, play a crucial role in societal service provision [2].

1.1. Assistive Robots

The Indian population is expected to increase three times in the next few years due to the increasing number of elderly people. With nearly 104 million aged persons aged 60 and above, many require assistance for their physical and intellectual activities. This issue also affects developed countries, where medical and robotics researchers focus on home and community-based healthcare services [3-5]. In the United States, 1 50 000 people use Activities of Daily Living (ADL) assistance Wheelchair robots. As the elderly face major health problems, global countries appoint special care providers and companions. Service robots are used in various daily activities, including healthcare, personal assistance, physical recovery, and rehabilitation. Social robots, assistive robots, and service robots are essential services for the elderly population [6].

1.2. Telerobotics – Kinova Jaco^2 Robot

Tele-robots are robots controlled remotely using rules and heuristics. They are used in various fields, such as space construction, disaster relief, and hazardous waste management. Tele-operation robots, on the other hand, operate in unstructured remote environments. Cameras and monitors provide visual feedback to these robots. Tele-operation robots are useful for special care to geriatric patients and medical services.

A voice and vision-based control system, the Kinova Jaco2 robot shown in Figure 1, has been developed with AI and ML-based controls for intelligent robotic arm movement. The robot is designed for people with mobility restrictions on an electrically powered wheelchair, allowing them to perform daily activities like drinking, eating, and picking up objects. However, object identification and arm movement improvements are needed, and new algorithms are needed for more effective handling [7].

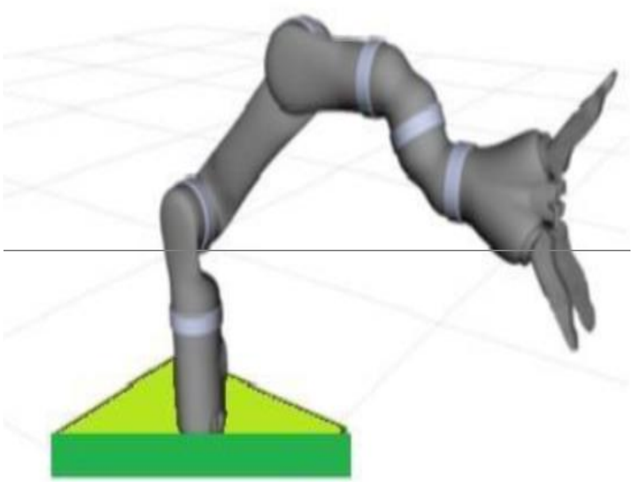

Fig. 1 3-D model of Kinova Jaco2 robot manipulator

1.3. Need for Natural Language Processing

Speech and vision-based object recognition systems enable contactless interfacing devices, making NLP best for geriatric human-to-rats communication. This technology has gained popularity due to its use of techniques like Hidden Markov Models (HMM), Artificial Neural Networks (ANN), Fast Fourier Transform (FFT), Learn Vector Quantization (LVQ), and Neural Networks. Natural Language Processing (NLP) is an important contribution of AI as it helps computers understand human language through various phases involved in processing. In machine intelligence technology, which can have its own control programmed, NLP is more important for speech capturing and observing than articulated robotic arm control systems, making it a prominent technology for medical robot design [8-10]. Speech recognition systems are designed with two phases: dimension reduction or information selection phase and discriminative estimation phase. The first phase extracts useful features from users' speech signals based on training and reduces its dimensionality based on specific work knowledge. Techniques such as Gaussian Mixture Models (GMM), FFTs, and Porter stemming algorithms are used for feature extraction. Deep learning methods have been used in vision systems to improve speech recognition. Convolution Neural Networks (CNN) was introduced in 1994, providing massive advances in the field of Deep Learning. CNN can directly depict acoustic features in spoken phonemes and is more efficient and accurate for specific applications than traditional multilayer perceptions-based back propagation neural networks. CNN is the most preferred choice for learning features from speech, working in two stages: feature learning and classification [11-14].

1.4. Stereo Vision System

Researchers have used stereo vision tracking systems for robot applications to calculate 3-D perception or depth sensation. Stereo vision involves capturing images from two cameras and extracting 3-D information. Position and orientation estimation is crucial in stereo vision, as it requires accurate tracking of the robot and the object's depth. Digital Camera fine-tuning and physique order transfiguration are necessary for initialization, and there are combo tracking modules: stationary and pan tilt or moving cameras. An image processing feedback control technique is proposed to control stereo-vision functioned robots with visual information. Cameras are needed to capture video frames in real-time or with low latency, and automatic movement on path layout is used to develop accurate object largeness and concavity. The study uses two monocular autofocus cameras to frame the stereo vision camera and capture images for processing. Camera calibration methods are based on the radial distortion method, and the image coordinate system is used in robotics [15-17].

1.5. Object Grasping

The Jaco2 robot uses a depth sensor to acquire objects and store them in a database as virtual scenes. A segmentation algorithm and image processing with deep learning techniques determine the robot's grasping action point. Artificial intelligence algorithms for robotics are fine-tuned as autonomous grasping robots need to recognize and grasp objects in unknown environments with reduced time and high accuracy. The research focuses on object detection, localization, classification, and detection using machine learning and deep learning methods. The study considers 12 objects, including water bottles, tumblers, spoons, tablets, black keys, syringes, and plates. Deep learning algorithms like Convolution Neural Network (CNN) are used for accurate classification and localization. Fuzzy temporal logic and genetic algorithms are combined to perform optimization in feature selection, reducing classification time and increasing accuracy. Linear classifiers are trained using Stochastic Gradient Descent (SGD) algorithms and objective function formulations [18].

2. Literature Survey

This paper presents the literature survey conducted for this research, aiming to understand various research works published in Robotics, Computer Vision, and Machine Learning journals and peer-reviewed international conferences.

2.1. Machine Learning

The advancement of robotics and computer technology has significantly impacted human culture and life, providing a range of services to humans. However, the design of assistive robots faces challenges due to their complex operations and the need for Machine Learning (ML) algorithms to solve these problems. Neural networks are used in robotics to solve inverse kinematics problems, trajectory planning, sensor data mapping, intelligent control, and task planning. Researchers have spent years reorganizing digital images, identifying context, and detecting emotions. Convolution Neural Network (CNN) is the most extensively pre-owned algorithm in machine perception technology, capable of analyzing data in depth and handling all types of datasets. CNN can be divided into three types of layers: Convolution Layer (CL), Pooling Layer (PL), and Fully Connected Layer (FCL). Deep learning models are used in object classification, using RGB depth data and object position estimation algorithms [19-21]. Deep learning algorithms have excellently performed in various domains like computer vision, natural language processing, and audio filtering. Support Vector Machine (SVM) is the most important algorithm in supervised learning methods, and it is used in applications such as medical diagnosis, text classification, credit card malfunction, and face recognition systems. Machine vision learning algorithms include the k-Nearest Neighbour (KNN) algorithm and Decision Tree (DT) to speed up confluence and improve reliability. Deep learning algorithms outperform traditional machine learning algorithms in areas such as image processing, speech recognition, natural language processing, intelligent question-answering systems, cross-lingual systems, and machine translation. They also offer superior predictability and accuracy. Researchers in the fields of artificial intelligence, machine learning, intrusion detection systems, and image processing have also looked into the Genetic Algorithm (GA) as an optimization technique. GA reproduces offspring by choosing two parents at a time following the survival of the fittest model [22-25].

2.2. Image Feature Selection/Pre-processing

Researchers increasingly focus on visualization and articulated machines to capture and analyse images more effectively. Technological advances have made it possible to shift from 2D to 3D technologies, allowing for faster classification algorithms. Feature selection is crucial in this process, as it reduces data dimension by eliminating irrelevant features and assigning weights. Learning algorithms focus on data characteristics for future prediction and analysis. Feature extraction in visualization systems can provide huge attributes, increasing classification time. Identifying the most significant and impactful features is imperative to improve decision-making.

Optimization techniques, such as evolutionary search, rule fine-tuning approaches, and linearly formulation of code data, are used for feature reduction. This paper proposes a unique protocol called Fuzzy Rules and Information Gain Ratio based Feature Selection Algorithm (FIGRFSA), which works rooted in the principles of indistinct brainy networks. The algorithm uses fuzzification and defuzzification to convert fuzzy outputs into crisp outputs. The Intelligent Agent-based Optimal and Incremental Feature Selection Algorithm (IAOIFSA) is proposed to increase decision accuracy based on image classification from the robot execution environment [26].

2.3. Image Classification Techniques

Photo classifications are a crucial research era involving large numbers of images for decision-making. Classifiers are learning models that form rules based on the dataset's characteristics and are used in supervised and unsupervised classification problems. The paper proposes a new technique called the FIGRFSA algorithm, based on supervised machine learning, for effective feature selection. The selection of contributing features depends on the high predicate nature of the attribute for the class rather than predicate with one another among features. The research focuses on computer vision and machine learning, focusing on image classification and testing the algorithms on large-scale datasets. The FRNGCA and the NGFTRCA are used for classification, with the latter utilizing an enhanced form of Convolutional Neural Networks [27-30].

2.4. Works on Assistive Robots

The invention of machine intelligence and implementation in life care support are common in countries like Japan, the US, Germany, Italy, and Spain. Many studies discuss the design and implementation of medical assistive robots, with most being controlled from remote locations called tele-robotics. Assistive technology has been initiated to help geriatric people by incorporating voice recognition without an internet connection. Researchers have developed various robot control interfaces, such as touch screens, Brain Control Interfaces, electromyography, eye gaze detection, and virtual joysticks. The recent technology of the smart Internet of Things (IoT) provides more intelligent devices, facilities, and possibilities for the elderly and disabled persons. These advanced and smart IoT devices incorporate features for assisting the elderly and disabled persons, offering smart infrastructure to senior citizens. Assistive robots are usually classified as rehabilitation intelligence machines and social assistive machines. Rehabilitation robots focus on physical assistance technology features, while social assistive robots are classified into companion bots and service bots. Care bots serve and care for human beings' tasks and routine intake and outtake activities. Research areas for service robots include wheelchair robots and mobility aids, physically disabled manipulator arms, rehabilitation robots, companion robots, and educational robots. Some researchers have developed nurse assistant robots like Pearl to assist nursing staff and elderly people. The aging problem has been analysed by many researchers, with Joyce et al. (2015) advocating for the use of robots as ambient assistive technology. Gianluca Giuffrida et al. (2019) planned a low-cost manipulator to realize simple tasks controlled by three different graphical HMI. The You Only Look Once (YOLO) v2 device developed with CNN considers the video stream generated by a robotic arm manipulator mounted with a camera, helping the user interact and recognize objects. Despite the presence of existing systems, new intelligent assistive robotic systems have been developed, such as fuzzy compensation methods to grasp an object, track, and guide robot manipulation using a stereo vision system. The fuzzy error compensation method standardizes gripper position to reduce error and ensures acceptable thresholds. Social robots are used in treatment, social assistance, and home companionship, but ethical concerns arise for custodians. Telemedicine healthcare robots collect user data and track movements, while George Mois and Jenay (2020) survey how robots support older people with physical, social, and cognitive support. A new service robot software architecture for manipulator services in human environments, autonomously performing various processes to grasp objects was developed [31-34].

2.5. Works on Natural Language Processing

Speech recognition systems use Hidden Markov Models (HMM) to match words in audio signals. Rabiner and Jaung (1993) introduced the first speech recognition system. Robot assistance systems are used to assist people with cerebral palsy, spinal cord injuries, and old age in daily tasks. The mapping range is obtained by combining rule-based expert system methods and Recurrent Neural Network (RNN) by extending context-free grammars with RNN. Kinova-Jaco2 robot systems have been developed for assistive robotic arms, which require object handling and manipulation for geriatric people. Recent works have proposed new techniques to control 6 DOF robots mounted on wheelchairs or tables using voice or speech recognition. Amin Atrash et al. (2009) developed a speech recognition system for wheelchair control using open-source recognition systems HTK and CMU's Sphinx-4. These systems are user-independent and perform continuous speech recognition. Terrin Babu et al. (2018) developed an oral command system to assist articulated arm bots in working on movement tasks, which is commercially available with a joystick control Jaco2 robot arm. Ali Bou Nassif et al. (2019) reviewed statistical analysis techniques for performing deep learning in speech applications. The latest Town Management System (TMS), named ROS-TMS 5.0, engages ROS to operate various scale-level sources in a care facility. The voice control system with open-source software platforms helps with focus. This paper develops a new assisting robot system, allowing users to interact with the robot using Python speech recognition API and Pocket Sphinx to convert the robot operator's speech commands to text form [35].

2.6. Augmented Reality in Machine Bot Eyesight

Elevated reality environments are beneficial for creating intelligent robotic systems, and reviewing previous works helps understand the features of existing systems.

2.7. Works on Augmented Reality

Bejczy et al. (1990) introduced augmented reality into teleoperation, introducing virtual predictive simulation results into normal displays. Rastogi et al. (1996) proposed a telerobotic AR system that displayed 3-D views in monitors, using location tracking via mechanical and optical sensors for camera adjustment; Takashi Okuma et al. (2004) proposed a 3-D object tracking model using new see-through techniques merging database handling techniques for effective video data analysis. Portela Sotelo et al. (2012) developed a 3D model-based multiple-object video tracking for treatment room supervision, using fuzzy logic systems for feature selection and effective integration of heterogeneous data. A markerless method for high accuracy and low latency without device tracking systems was presented by Akash Bapat et al. (2016) [36].

2.8. Works on Stereo Vision System

The human-machine interface is crucial in telerobotics, as it allows communication with telerobots. Computer vision algorithms have been studied in the field, with early research focusing on computational stereo vision. Stereo vision systems help calculate disparity from scene depth maps and image registration frameworks. Lucas and Kanade (1981) proposed an invaluable image registration method using the Newton-Raphson iteration method. Roger Y. Tsai (1987) provided theoretical and experimental models for future 3-D machine vision metrology and 3-D robot vision research. Don Murray and Cullen Jennings (1997) developed a stereo vision-based mapping and navigation model for mobile robots, integrating grid mapping, robot path planning, mapping, and navigation systems. Roland and Aural (2015) implemented robot arm control using colour recognition algorithms, allowing physically reachable 3-D spaces to be moved. Photogrammetry is a 2D image reconstruction system used in engineering, manufacturing, geology, and archaeology. Grigorescu et al. (2011) proposed a robust real-time camera pose and scene structure estimation system. Taihú Pire et al. (2017) proposed a real-time feature-based stereo system for video processing and robot localization. Wong and Majji (2018) used recursive linear least squares algorithms to estimate relative pose and pose rate, and kinematic motion models were utilized for target tracking in their relative navigation filters for localization and mapping applications utilizing stereo vision sensors [37].

2.9. Intelligent Robotics Design Methods Using a Learning Approach

Machine learning algorithms, such as Neural Networks (NNs) with backpropagation and Support-Vector Regression (SVR), are used to improve classification-based rules in computers. Deep learning, which increases the number of hidden layers, is a recent trend in machine learning. NN models work with approximations and have non-linear activation functions and bias functions. Deep learning architectures often use deep neural networks. Feng et al. (2012) proposed a machine learning-based robotic system using inverse kinematic expressions and gradient-based learning algorithms for robot arm control. Their approach reduced computation time and improved precision. CNN is commonly used for object detection and classification and to understand surroundings through visual cues. Researchers have also explored image segmentation and Region-Based Object Recognition (RBOR) methods [38].

2.10. Image Feature Selection and Classification

Researchers have developed various feature selection, classification, and augmented reality techniques. The accuracy of feature selection depends on the technique used for feature selection. Vision techniques are classified into Active and Passive vision, with passive vision categorized into time of light, structured light, light coding, and laser triangulation methods. Anchal Solio and Siddharth (2013) proposed a model for content-based retrieval, which uses media items as input for a given query. Prior to inferring information about the contents of images, Content Based Image Retrieval (CBIR) systems seek to extract relevant information from raw data. Deepu Rani and Monica Goyal (2014) proposed a method for image retrieval using CBIR and Support Vector Machine (SVM). Barbara Zahradnikova et al. (2015) reviewed Image Mining, emphasizing the need for automated processing and classification to obtain specific information from an image collection. Hassaballah et al. (2016) evaluated basic notation and mathematical concepts for detecting and describing image features, exploring properties of perfect features and discussing existing detection and description methods. Pavithra and Sivaranjani (2016) proposed a new algorithm for Content-Based Image Retrieval (CBIR) that automatically extracts low-level visual image features like colour, texture, and shape for image descriptions and indexing purposes. Their algorithm reduced the elapsed time of the system for all target search methods and produced good results in terms of accuracy. Lei Zhu et al. (2017) proposed extracting rich semantics from auxiliary texts of images to improve visual hashing performance. Feature selection (or preprocessing) consists of detecting relevant attributes from the data set and discarding irrelevant ones, with advantages such as performance improvement, data understanding, and cost reduction. Feature selection is useful in intrusion detection systems as it reduces classification time and helps capture differences between intruders and normal users. Krizhevsky et al. (2012) achieved the best data set results in the large Image Net dataset subset ILSVRC2010 and 2012 by training with CNN.

2.11. Works on Fuzzy Classification

A method for object edge detection from images was proposed by Nitin et al. (2018). It involved altering the fuzzy membership function to prioritize multi-focus fusion over fuzzy edge detection. To contribute to mathematical problem-solving, Omar Abu Arqub (2015) also introduced a numeral technique for solutions for fuzzy-based divergent equations using Hilbert sites, which can be applied to robotic applications for accurate inferences through fuzzy reasoning. Fuzzy Min-Max neural network classifiers are used for visual pose estimation in robotic manipulators. Aung Myat San et al. (2018) described the pictorial posture projection of an indeterminate robotic manipulator using the Artificial Neuro-Fuzzy Inference System (ANFIS) with double un-calibrated imaging tools. Hind Rustum Mohammed et al. (2016) proposed an improved fuzzy c-means algorithm for faster cluster detection. This paper proposes new feature selection and classification algorithms for more accurate augmented image classification [39-41].

2.12. Works on Object Grasping

Since 1994, researchers have developed various techniques for object grasping, including robot image motion analysis, algorithms using colour, texture, and edge information, and real-time manipulation. Miller et al. developed an automatic grasp planning system, which models objects and uses sensor feedback to calculate the optimum grasp. Kofman et al. presented a teleoperation robot using non-contact vision-based methods, while Clothier and Shang focused on the self-governing location of an articulated robot arm. Bone et al. developed an online silhouette and structured light 3-D object modeling approach using images from multiple viewpoints from a wrist-attached video camera. Zhang (2018) proposed a new robot that produces a force to close grasps and outputs for identifying the gripper's position and orientation for grasping the object.

2.13. Research Works On Jaco2 Robot

Kinova Robotics, a Canadian organization, developed the JACO2 robotic arm for assistive or rehabilitation purposes. The JACO Rehab Edition was introduced in 2010 for people with reduced mobility or upper limb impairments. In 2012, an upgraded version was prepared for the scientific community, known as the JACO2 Research Edition. Clinical trials have shown that the assistive equipment is easy to use in regular jobs and reduces essential caregiving time. Yoonseok Pyoa developed a service robot system for real environment detection, fetching, and task execution, using ROS-TMS to provide real-time environment information. The FRIEND series used visual servo technology to control the robotic arm motion, but this was limited to hybrid visual serving. Ka et al. developed a semi-autonomous Assistive Robotic Manipulation Assistance (ARoMA) system, controlling the JACO2 robotic arm using machine eye vision and sound perception.

3. Proposed Work

The literature survey identifies research gaps in object grasping methods, including the need for a deep learning-based approach and a new robot arm-handling method. Existing methods have higher error rates in robot arm positioning, resulting in inaccurate joint position values.

The paper aims to introduce smart JACO2 robots to assist elders, propose new image-based commanding algorithms, feature selection algorithms, deep learning-based classification techniques, and a grasping algorithm using fuzzy rules for precise object identification. These improvements aim to enhance the fidelity of object grasping.

The proposed system shown in Figure 2 uses a Natural Language Processing-based User Interface to provide commands to cameras, which capture video and convert it into frames. The images are extracted by the camera management module and processed by the image processing module, which performs feature extraction, feature selection, and classification to identify object locations. Two new feature selection and classification algorithms, FIGRFSA and IAOIFSA, are proposed for robot action planning using image processing tasks. The Intelligent Voice Text Conversion (IVTC) technique converts voice input into text form, and the Fuzzy Temporal Rule-based Semantic Analysis Algorithm (FTRSAA) is used for semantic analysis.

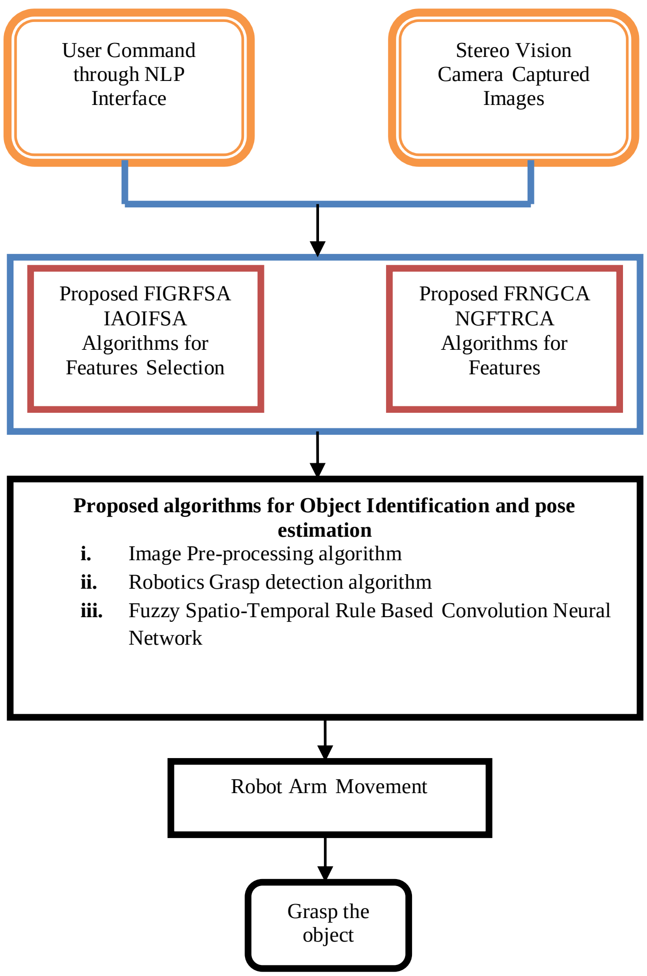

Fig. 2 Flow diagram of the proposed system

The system also proposes three new algorithms for object detection and grasping, including an Image Preprocessing Algorithm, Object Grasping using a Fuzzy Spatio-Temporal Rule Based Convolutional Neural Network Classification Algorithm, and a Robotic Grasp Detection Algorithm. Assistive technology robots can potentially improve elder or geriatric care, but they face challenges in object tracking, motion estimation, and pose estimation. Researchers have proposed real-time pose estimation for reliable solutions, but traditional tracking methods still require strong priority on pose. The complexity of real-time tracking is due to initialization, large processing tasks, and data volume.

A new assistive mobile robot system has been proposed to help geriatric people live safe, better, and longer lives in their homes. This paper aims to address these challenges and develop technology that meets the needs of older adults and geriatrics.

4. Research Methodology

4.1. Fuzzy Rules and Information Gain Ratio-based Feature

The Fuzzy Rules and Information Gain Ratio-based Feature Selection Algorithm (FIGRFSA) is a new feature selection algorithm that finds how many features are best for a classification algorithm. It computes the knowledge gain ratio and selects the finest number from 25 extracted features, including shape, statistical, temporal, colour, texture, and spatial. This reduced set of features reduces classification time and increases accuracy.

4.2. Intelligent Agent-Based Optimal and Incremental Feature

The Intelligent Agent-based Optimal and Incremental Feature Selection Algorithm (IAOIFSA) is a feature selection algorithm that uses intelligent agents to sensitize the environment and form rules for finding prime features. It reduces classification time and enhances accuracy. The algorithm extracts key frames from a video sequence and uses instant comparison, interval comparison, boundary analysis, information gain ratio, fuzzy temporal rules, and de-fuzzification. The identified facets are then used in the grouping technique.

4.3. Fuzzy Rule-based Neuro-Genetic Classification Algorithm

The Fuzzy Rule-based Neuro-Genetic Classification Algorithm (FRNGCA) is a new categorization algorithm designed to enhance image classification accuracy for robot arm control and elder medical assistance. It uses fuzzy rules, triangular membership function values, and Genetic Algorithms (GAs) for weight adjustments and activation functions. The neural network has one input layer, two hidden layers, a sigmoidal activation function, and a bias function. Error functions are used to reduce errors and optimize classification. This innovative classifier enhances decision-making for object identification.

4.4. Neuro-Genetic Fuzzy Temporal Rule-Based Deep Learning

The Neuro-Genetic Fuzzy Temporal Rule-based Deep Learning Classification Algorithm (NGFTRCA) uses Convolutional Neural Networks for classification, IFT rules, and ST. The algorithm has ten convolution levels and eleven max-pooling levels for effective data set classification. Genetic algorithms optimize the search space during training, improving classification accuracy and reducing error rates.

4.5. Object Grasping Algorithms

This paper proposes three new algorithms for accurate object grasping: Image preprocessing, object Grasping using Fuzzy Spatio-Temporal Rule Based Convolutional Neural Network Classification Algorithm, and Robotic Grasp detection. The image pre-processing algorithm removes noise, normalizes data, and accurately identifies object centres. The classification algorithm uses fuzzy logic, forward-chaining inference, and temporal rules. The robotic grasping detection algorithm generates rules for gripping objects accurately. Experiments with a Kinova Jaco robot were conducted to test these algorithms. Results showed improved decision accuracy and reduced grasping time. This paper presents a survey of related works in image and signal processing, pre-processing, classification, AI and ML methods, fuzzy rules, natural language processing, and assistive robot design. It analyses existing techniques, identifies research gaps, and proposes new techniques. The remaining Papers will discuss the new techniques and experimental results.

4.5.1. System Architecture

This paper presents the architecture, which is shown in Figure 3, of a proposed intelligent robotic system for elder assistance, consisting of ten major modules: connecting with the user, voice capturing and Voice Text Conversion, Instruction Validation Device, Data Repository, Judgement Strategies, and Robot with Object Plan, Picture Analysis System, and Image Capture segment.

4.5.2. User Interface Block

The medical assistance robot uses a user interface module, a laptop-based microphone, and camera-connected display systems to communicate with users. The module uses voice commands to recognize medical items and hand them over to elderly people. The camera display system provides feedback on the robot's position and allows the robot arm to move based on user commands to the next target object.

4.5.3. Speech Recognition Module

Speech recognition is a process where a microphone captures voice or acoustic signals and converts them into words. Two modules identify speech series and words from them. The module accepts user commands and performs pre-processing using noise discrimination models to remove noise. It uses a Support Vector Machine (SVM) classification algorithm to identify spoken commands. The module has two phases: training to recognize commands and real-time prediction testing.

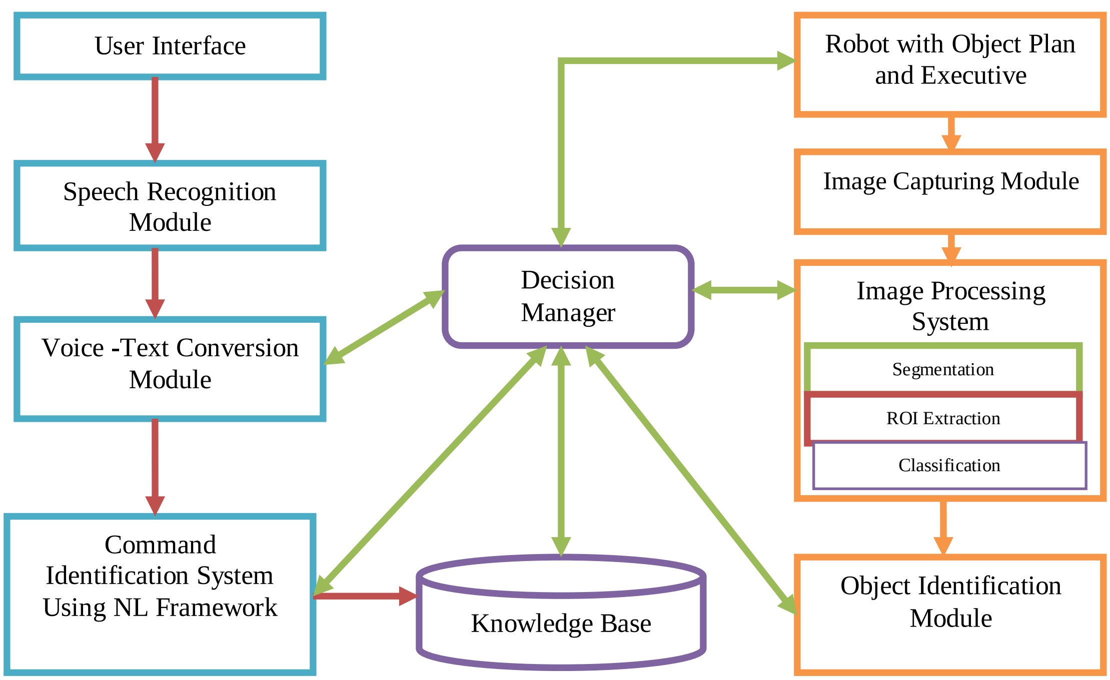

Fig. 3 System architecture

4.5.4. Voice-to-Text Conversion Module

The IVTCA is a pre-processing algorithm that converts user voice into Mel-Frequency Cepstral Coefficients (MFCCs) features. This module accurately approximates user-to-system responses and identifies voices in numerical terms. IVTCA offers better service and accurate conversion through deep learning techniques than other voice-to-text conversion tools, making it suitable for machine learning algorithms.

4.5.5. Command Identification System

Commands are processed by a voice-to-text conversion module, converting them to text for further analysis before execution by the robot. This includes morphological analysis, parts of speech tagging, syntax analysis, and semantic analysis to understand accurate command meaning.

4.5.6. Knowledge Base

The Knowledge base is a tool used by decision managers to understand user speech sentences, particularly those with ambiguity. It stores and retrieves rules from training models, which are applied during testing for accurate decision-making. The base consists of facts and rules, represented in natural English, and are used for more accurate decision-making.

4.5.7. Decision Manager

The DM governs every module in the proposed system, with interactions shown in the architecture diagram. It monitors voice capturing, voice-to-text conversion, data visualization, robot planning and execution, image capture, object identification, and knowledge base. The decision manager guides the entire process by applying rules from the knowledge base, making decisions, selecting features, and making decisions in classification, robot pose, and entity detection.

4.5.8. Robot Planning for Object Grasping

The robot's main module, the plan and execute model, generates a path plan from the current robot's location to the object's and user's locations. It uses a hierarchical planning algorithm and heuristic search techniques to reach the object optimally. The robot’s inclusion parts are divided into sub-areas; as distance increases, it backtracks to the previous position and finds an alternate path. The robot then executes movements through the identified optimal route.

4.5.9. Image Capturing Segment

The dual audio eyesight device uses a double autofocus recorder image to take pictures of objects and robots. It uses planning modules to reach objects and learns the environment through rules and past moves. The robot image-capturing module extracts frames from videos for better image capture. The captured images are archived employing R-tree archiving and captured in the storage device, retrieved as needed and sent to the image processing module for analysis.

4.5.10. Image Processing Module

The image capturing module captures an import object, which is then sent to the image processing module for segmentation, extraction of regions of interest, feature selection, and classification. This process uses new algorithms, increasing detection accuracy and reducing classification time. The module also performs pre-processing activities.

4.5.11. Object Identification Module

The stereo vision camera uses database information to accurately measure object depth and environment area. It uses an object recognition module with fuzzy temporal rules for effective object identification. The system handles incomplete data through fuzzification, temporal constraints, and de-fuzzification. The rules are fired using forward chaining inference, and discriminant networks are formed for rule firing and matching. This research contributes to natural language processing, image capture, pre-processing, segmentation, feature selection, and classification using deep learning algorithms.

This paper presents the structure of the medical assistive robotic system proposed in this paper, illustrating the interaction among modules and explaining their functionalities. The architecture diagram illustrates the overall flow of the robotics system designed and implemented.

4.5.12. System Implementation

This research uses a stereo vision camera, monitor, and microphone to control a robot arm in the Robotic Operating System (ROS), using Ubuntu 18.04 for Linux-oriented implementation. The ROS control Jaco2 robot arm implements communication, real-time environment management, and real-time task planning functions. Its architecture consists of three main modules: Stereo vision image processing, Voice/Speech Recognition, and Robotic Arm Control Model. These functions ensure efficient communication and user interaction.

4.5.13. System Description

This Paper explores the design of intelligent assistive robots for elderly patients, focusing on the 6 DOF Jaco2 robots. The robot is interfaced with a laptop via USB and controlled by ROS's spot velocity conduct mode. The JACO2 robot is chosen due to its modularity, efficient payload-to-weight ratio, user-friendly interface, safe handling of objects, and 6 degrees of freedom. The articulated robot is powered by a 24V battery or standard electrical plug, and its lightweight, safety standards, portability, and usability make it an ideal choice for research.

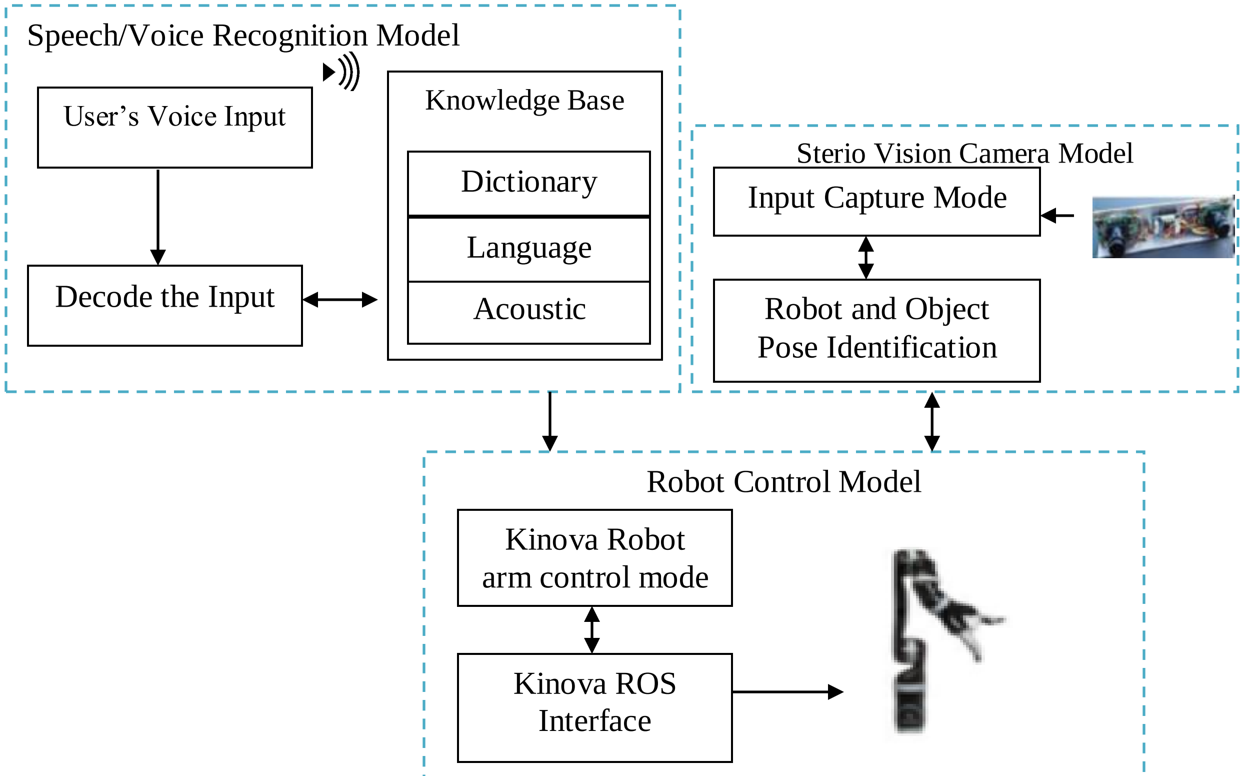

Fig. 4 Overall structure of the Jaco2 robotic system

The proposed system shown in Figure 4 uses a laptop with an in-built or external microphone for user convenience and limited mobility. It uses a speech recognition system with predefined commands in natural language and English based on suggestions from nurses, caregivers, and geriatrics. The system controls a robot arm in 3-D Cartesian space, with a range of movements defined by user commands.

5. Conclusion and Future Work

This work proposes new algorithms for object detection and grasping using machine learning. Three key contributions include speech-to-text conversion, feature selection and classification algorithms, and pre-processing, classification, and object-grasping algorithms. These techniques improve object identification and classification, making the robot model more effective in object grasping. This work proposes a new Image pre-processing algorithm for robotic arm movement using stereo cameras.

The algorithm removes noise, performs feature extraction and selection, and assigns selected features to the Fuzzy Rule-based Neuro-Genetic Classification Algorithm. The techniques are applied to regions of interest identification, image segmentation, and object grasping. The algorithm has been tested on objects like water bottles, tumblers, and spoons. The deep learning-based classification approach is used, and the proposed approach increases grasping accuracy by 5% and reduces time by over 5 milliseconds. The system performs robotic actions more accurately, with a better user interface, faster response, and accurate decisions. There are numerous opportunities to expand this work and add features that will benefit the elderly population more.

Intelligent agents can be placed in various locations where the objects are stored and trained to communicate with the robot to quickly aid in the deduction process. Specialized sensors can be created for this use. Using higher order logics, such as situation 153 logic and description logics, to create efficient robotic plans that can carry out detective inference and aid in developing more precise judgments regarding object grabbing is another avenue for future research.

This work introduces a new methodology for controlling a robot arm to grasp user-commanding objects using Convolutional Neural Networks. The robot arm joint position values are calculated based on the object's position from the robot end effectors.

The research focuses on JACO2 robot design, image-based commanding algorithms, feature selection algorithms, and classification techniques. It also proposes three new algorithms for image data pre-processing, classification algorithms with fuzzy temporal rules, and object grasping using effective image analysis. The goal is to provide medical assistance to elders by enhancing the robot arm's accessibility and accuracy.

Acknowledgments

Regarding acknowledgements of this research, work has been done by Author 1, Chandrashekhar Kumar, for manuscript preparation and software robotic section and electronics design work, and Author 2 and the author 3 have been verified with this research and guided by manuscript design guide. Author 1 and Author 2 contributed equally to this work.

References

[1] Ossama Abdel-Hamid et al., "Convolutional Neural Networks for Speech Recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no.10, pp. 1533-1545, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[2] Abdulla Mohamed, Chenguang Yang, and Angelo Cangelosi, "Stereo Vision Based Object Tracking Control for a Movable Robot Head," *IFAC-PapersOnLine*, vol. 49, no. 5, pp. 155-162, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[3] Prabha Adhikari, "Geriatric Health Care in India - Unmet Needs and the Way Forward," *Architecture Medical Health Science*, vol. 5, no. 1, pp. 112-114, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[4] Akash Bapat, Enrique Dunn, and Jan-Michael Frahm, "Towards Kilo-Hertz 6-DoF Visual Tracking Using an Egocentric Cluster of Rolling Shutter Cameras," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 11, pp. 2358-2367, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[5] Alexandre Campeau-Lecours et al., "Kinova Modular Robot Arms for Service Robotics Applications," *International Journal of Robotics Applications and Technologies*, vol. 5, no. 2, pp. 49-71, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[6] Ali Bou Nassif et al., "Speech Recognition Using Deep Neural Networks: A Systematic Review," *IEEE Access*, vol. 7, pp. 19143-19165, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[7] Amin Atrash et al., "Development and Validation of a Robust Speech Interface for Improved Human-Robot Interaction," *International Journal of Social Robotics*, vol. 1, no. 4, pp. 345-356, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[8] Anchal A. Solio, and Siddharth A. Ladhake, "A Review of Query Image in Content-Based Image Retrieval," *International Journal of Advanced Research in Computer Engineering and Technology*, vol. 2, no. 4, pp. 1619-1622, 2013. [Google Scholar]

[9] James A. Anderson, *An Introduction to Neural Networks*, Cambridge: MIT Press, pp. 1-650, 1995. [Google Scholar] [Publisher Link]

[10] Archana Kale, and Shefali Sonavane, "Optimal Feature Subset Selection for Fuzzy Extreme Learning Machine Using Genetic Algorithm with Multilevel Parameter Optimization," *Proceeding of the IEEE International Conference on Signal and Image Processing Applications*, Kuching, Malaysia, pp. 445-450, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[11] Aung Myat San, Wut Yi Win, and Saint Saint Pyone, "ANFIS-Based Visual Pose Estimation of Uncertain Robotic Arm Using Two Uncalibrated Cameras," *International Journal of Wireless Communications and Mobile Computing*, vol. 6, no.1, pp. 20-30, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] Ronald T. Azuma, "A Survey of Augmented Reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, pp. 355-385, 1997. [Google Scholar] [Publisher Link]

[13] R. Azuma et al., "Recent Advances in Augmented Reality," *IEEE Computer Graphics and Applications*, vol. 21, no. 6, pp. 34-47, 2001. [CrossRef] [Google Scholar] [Publisher Link]

[14] Aaron Bangor, Philip T. Kortum, and James T. Miller, "The System Usability Scale (Sus): An Empirical Evaluation," *International Journal of Human-Computer Interaction*, vol. 24, no. 6, pp. 574-594, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[15] Barbora Zahradnikova, Sona Duchovicova, and Peter Schreiber, "Image Mining: Review and New Challenges," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 7, pp. 242-246, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[16] Emanuele Bastianelli et al., "Effective and Robust Natural Language Understanding for Human-Robot Interaction," *European Conference on Artificial Intelligence*, pp. 57-62, 2014. [Google Scholar] [Publisher Link]

[17] Emanuele Bastianelli et al., "Structured Learning for Spoken Language Understanding in Human-Robot Interaction," *The International Journal of Robotics Research*, vol. 36, no. 5-7, pp. 660-683, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[18] A.K. Bejczy, W.S. Kim, and S.C. Venema, "The Phantom Robot: Predictive Displays for Teleoperation with Time Delay," *IEEE International Conference on Robotics and Automation*, Cincinnati, OH, USA, vol. 1, pp. 546-551, 1990. [CrossRef] [Google Scholar] [Publisher Link]

[19] Linda M. Beuscher et al., "Socially Assistive Robots," *Journal of Gerontology Nursing Diagnosis*, vol. 43, no. 12, pp. 35-43, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[20] A. Bicchi, and V. Kumar, "Robotic Grasping and Contact: A Review," *Proceedings 2000 ICRA. Millennium Conference IEEE International Conference on Robotics and Automation. Symposia Proceedings*, San Francisco, CA, USA, vol. 1, pp. 348-353, 2000. [CrossRef] [Google Scholar] [Publisher Link]

[21] Jonathan Bidwell, Alexandra Holloway, and Scott Davidoff, "Measuring Operator Anticipatory Inputs in Response to Time Delay for Teleoperated Human-Robot Interfaces," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, pp. 1467-1470, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[22] Gary M. Bone, Andrew Lambert, and Mark Edwards, "Automated Modeling and Robotic Grasping of Unknown Three-Dimensional Objects," *IEEE International Conference on Robotics and Automation*, Pasadena, CA, USA, pp. 292-298, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[23] Cynthia Breazeal, *Designing Sociable Robots*, MIT Press, Cambridge, pp. 1-282, 2004. [Google Scholar] [Publisher Link]

[24] Joost Broekens, Marcel Heerink, and Henk Rosendal, "Assistive Social Robots in Elderly Care: A Review," *Gerontechnology*, vol. 8, no. 2, pp. 94-103, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[25] Charles G. Burgar et al., "Development of Robots for Rehabilitation Therapy: The Palo Alto VA/Stanford Experience," *Journal of Rehabilitation Research and Development*, vol. 37, no. 6, pp. 663-673, 2002. [Google Scholar] [Publisher Link]

[26] Filippo Cavallo et al., "Development of a Socially Believable Multi-Robot Solution from Town to Home," *Cognitive Computation*, vol. 6, no. 4, pp. 954-967, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[27] Jun-Wei Chang et al., "Implementation of an Object-Grasping Robot Arm Using Stereo Vision Measurement and Fuzzy Control," *International Journal Fuzzy System*, vol. 17, no. 2, pp. 193-205, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[28] Changjoo Nam et al., "A Software Architecture for Service Robots Manipulating Objects in Human Environments," *IEEE Access*, vol. 8, pp. 117900-17920, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[29] Chaobin Liu, Yuexiang Yang, and Chuan Tang, "An Improved Method for Multiclass Support Vector Machines," *ACM International Conference on Measuring Technology and Mechatronics Automation*, pp. 504-508, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[30] Yixin Chen, and J.Z. Wang, "Support Vector Learning for Fuzzy Rule-Based Classification Systems," *IEEE Transactions on Fuzzy Systems*, vol. 11, no. 6, pp. 716-728, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[31] Yuehui Chen, Ajith Abraham, and Bo Yang, "Feature Selection and Classification using Flexible Neural Tree," *Neurocomputing*, vol. 70, no. 1-3, pp. 305-313, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[32] Ching-Yu Chiu et al., "A Wireless Steady State Visually Evoked Potential-Based BCI Eating Assistive System," *International Joint Conference on Neural Networks*, Anchorage, AK, USA, pp. 3003-3007, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[33] Chotirat Ann Ratanamahatana, and Dimitrios Gunopulos, "Feature Selection for the Naive Bayesian Classifier Using Decision Trees," *Applied Artificial Intelligence*, vol. 17, no. 5-6, pp. 475-487, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[34] Cheng-Shiu Chung et al., "Performance Evaluation of a Mobile Touchscreen Interface for Assistive Robotic Manipulators: A Pilot Study," *Topics in Spinal Cord Injury Rehabilitation*, vol. 23, no. 2, pp. 131-139, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[35] Dan Ciregan, Ueli Meier, and Jürgen Schmidhuber, "Multi-Column Deep Neural Networks for Image Classification," *IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, USA, pp. 3642-3649, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[36] Kurt E. Clothier, and Ying Shang, "A Geometric Approach for Robotic Arm Kinematics with Hardware Design, Electrical Design, and Implementation," *Journal of Robotics*, vol. 2010, no. 1, pp. 1-10, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[37] Ulysse Côté-Allard et al., "Transfer Learning for sEMG Hand Gestures Recognition Using Convolutional Neural Networks," *IEEE International Conference on Systems, Man, and Cybernetics*, Banff, AB, Canada, pp. 1663-1668, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[38] M. Cowling, and R. Sitte, "Analysis of Speech Recognition Techniques for Use in a Nonspeech Sound Recognition System," *Digital Signal Processing for Communication Systems*, vol. 703, no. 1, pp. 31-46, 2002. [Google Scholar]

[39] S. Davis, and P. Mermelstein, "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 28, no. 4, pp. 357-366, 1980. [CrossRef] [Google Scholar] [Publisher Link]

[40] Steven B. Davis, and Paul Mermelstein, "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences," *Readings in Speech Recognition*, Elsevier, pp. 65-74, 1990. [CrossRef] [Google Scholar] [Publisher Link]

[41] Deepu Rani, and Monica Goya, "A Research Paper on Content-Based Image Retrieval System Using Improved SVM Technique," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 4, no. 10, pp. 683-686, 2014. [Google Scholar]