2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS) | 979-8-3315-0546-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/CSITSS64042.2024.10816986

NLP Enabled Autonomous Indoor Navigation Robot for Kannada in RHCs

Kunal A Kathare
Dept. of Computer Science
PES University
Bengaluru, India
kunalkathare1200@gmail.com

Kushal C G
Dept. of Computer Science
PES University
Bengaluru, India
kushalganesh2@gmail.com

Kevin Philip Kottukapally
Dept. of Computer Science
PES University
Bengaluru, India
kevinphilip2002@gmail.com

Kshitij Sharma
Dept. of Computer Science
PES University
Bengaluru, India
mailmekshitij.sharma@gmail.com

Vadiraja Acharya
Dept. of Computer Science
PES University
Bengaluru, India
vadiraja@pes.edu

***Abstract*—In Regional Healthcare Centers (RHCs) in many of rural areas of India, due to the language barriers between the patients and the RHC staff, communication difficulties often intensifies the patient's stress due to their inability to understand the workings of a RHC. This research paper presents a medical service robot which is designed to overcome these issues. The medical service robot can understand the patient's symptoms which are described to it in Kannada language using speech recognition and then using NLP techniques, it will be able to predict the appropriate department to guide the patient to. The robot then uses ROS2 and its navigation techniques to guide the patient to the appropriate department location. The robot is also attached with a camera at the back so that it can track the patient and see to it that the patient follows the robot. This paper discusses the technical aspects of the robot, highlighting its potency in improving the patient experience. This innovation has the capability to amplify the healthcare accessibility and patient care in rural areas of India, reducing linguistic and navigational challenges.**

***Index Terms*—Autonomous Robot, NLP Enabled Robot, Image Processing**

I. Introduction

Healthcare providers may face language barriers, but these can be more of a challenge for individuals residing in rural areas who have limited access to healthcare services. When patients seeking medical attention do not share the same language as their caregivers, they are likely to misinterpret medical instructions, diagnoses and treatment alternatives leading to confusion and misunderstanding. This is particularly troublesome because it means there is no understanding on how RHCs should work. It may become difficult for patients to use the RHC, get medical assistance or describe their health issues. Consequently, this leads to increased tension amidst patients already in pain due to their ailments with subsequent poor health outcomes.

Autonomous robot navigation is an exciting technology that can facilitate various activities within RHCs. These robots are equipped with sensors such as LiDAR and artificial intelligence which helps them find their way around indoor and outdoor environments. Additionally, these robots can direct people across different environments thereby ensuring that they reach their destinations without getting lost.

In this research, the primary concern is to resolve the language barriers and enhance patients' experience in RHCS. Our imaginative angle emphasizes on a robot that provides medical services by having it equipped with different technologies in RHCs for patients. This healthcare service robot has been programmed to understand Kannada, a regional language that regards the patient's symptoms by way of speech recognition. Furthermore, it applies Natural Language Processing techniques to grasp what exactly the patient is going through and which department within the RHC may cater him according to his symptoms.

This medical service robot uses Robot Operating System 2 (ROS2) and its navigation approaches to act as guiding light for the patients, leading them to the correct location in RHC, ensuring they receive the required care. Also, it tracks patients using Image Processing and Computer Vision (IPCV) while following the robot so that they are not lost along their way. When the patient ceases to follow, the robot stops and waits for him or her to catch up. We use this method not only to overcome language barriers but also cope with some navigation problems that big RHC usually have to address when serving patients.

This paper discusses in depth about the technical aspects of the medical service robot, taking a look at its design, functionality, and integration of various modules. By reducing the linguistic barriers and simplifying the navigation process, this innovation has the ability to transform the healthcare services provided to patients in rural areas. In the pages that follow, we will see how this transformative technology can reshape the healthcare services, making sure that everyone, regardless of language or location, can receive the care they deserve.

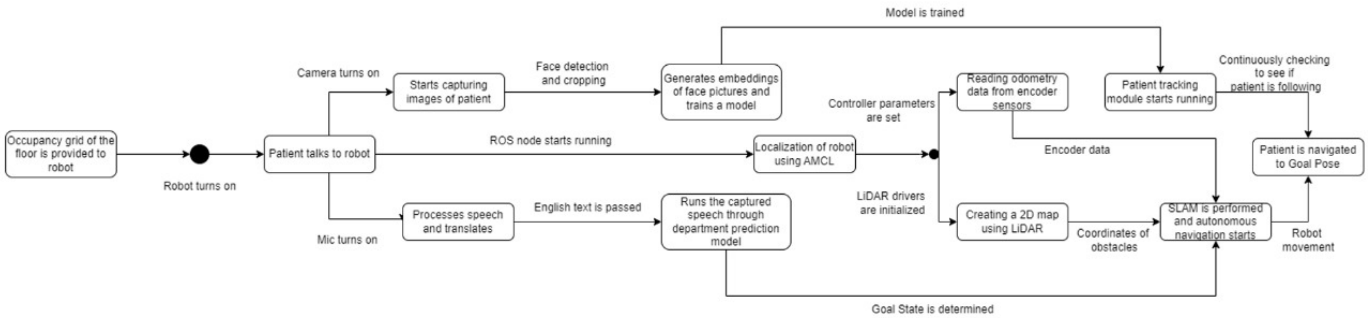

Fig. 1. Workflow of the robot.

II. RELATED WORK

A. Autonomous Robot Navigation

Autonomous robot navigation within indoor environments has seen significant research interest in recent years, due to its potential applications in various fields such as healthcare, manufacturing, and logistics. The ability of robots to navigate autonomously through complex areas can greatly improve their productivity, efficiency, and safety. However, achieving robust and precise autonomous navigation indoors poses several challenges, including the requirement for accurately sensing the surroundings, planning paths effectively, as well as being able to perform obstacle avoidance.

To address these challenges, researchers have explored multiple approaches. One method uses deep convolutional neural networks to learn path planning [3], depth perception, and surface normal detection [4]. Although this approach requires extensive training [3] and is quite computationally expensive, but it has shown promising outcomes in indoor environments [4]. Another strategy [3] combines visual simultaneous localization and mapping (SLAM) with object detection using pretrained convolutional neural networks. This method [5] uses a Kinect V2 RGB-D camera and YOLO V3 tiny model for object detection and RTAB-Map for visual SLAM. Nevertheless, additional experiments are essential to assess the YOLO V3 model's robustness across diverse and complex environments.

B. Path Planning Algorithms

Path planning is an essential component of robotics, particularly in the task of optimizing navigation from a starting point to a destination point within mobile robot systems. The purpose is to navigate in an efficient way through cluttered environments and find the most optimal route from the starting position to the target position. To achieve this goal, quite a few techniques have been developed, including deterministic, non-deterministic, and evolutionary algorithms. Among these various techniques the [9][11] A* algorithm and the artificial potential field (APF) method are two commonly used techniques for path planning. The A* algorithm is a method that uses heuristics to find the optimal path from the source to the destination.

Recent research [9][11] has combined the A* algorithm and the APF method to improve the efficient of path planning.

C. Image Processing

In recent years, there has been an increasing demand for service robots that have the capability to autonomously follow humans. To achieve this, the robots must have the ability to track and identify individuals using sensors such as monocular cameras [13][14], laser rangefinders [14], and depth cameras [12]. Consequently, several research studies have been conducted to develop algorithms and frameworks that can help to accurately track and identify individuals in real-time. One such study [14] proposed a human-following algorithm designed for an autonomous mobile robot equipped with a 2D laser rangefinder and a monocular camera. The algorithm used data fusion techniques to combine the data from both the sensors to track the participants in cluttered environments.

D. NLP

The first paper [15] focuses on part-of-speech tagging for Kannada and Hindi languages through a mix of machine learning and deep learning techniques. The study results help to conclude that deep learning techniques provide better accuracy, but require larger datasets, while machine learning techniques may be more feasible in case of smaller datasets. The next paper [17] introduces a healthcare chatbot system that uses AI to diagnose diseases and provide basic medical information, aiming to improve access to medical knowledge and reducing healthcare expenses. However, the sparsity of feature spaces represented by N-gram models is a major drawback. The third research [18] suggests building an Automatic Speech Recognition (ASR) system for the Kannada language. The system utilizing the Kaldi toolkit, extracts MFCC features from audio data to create monophonic and triphone models. The triphone model performs better than the monophone model, but there is still further room for improvement.

E. Integrating NLP with Robot

In recent years, there has been an increasing interest in developing intelligent humanoid robots that are able to engage in natural language interactions with humans. The paper[2], "Anita: Intelligent Humanoid Robot with Self-Learning Capability Using Indonesian Language," aims to create a robot that can communicate with people in their native language, specifically Indonesian. The researchers used Google speech-to-text to convert user input into tokens, and Word2vec method is used to find the most relevant terms in the question asked to provide an appropriate answer.

III. METHOD

A. Design and structure of the robot

The robot is made up of 2 layers, the bottom layer houses the batteries, Arduino and the motor driver. All of the wheels along with the motors are attached below the bottom layer. The top layer houses the Jetson Nano and the Li DAR. A PVC pipe is place attached to the bottom layer and passes through the top layer. This pipe allows us to attach a camera equipped with a microphone to it. A differential drive system is used with the front 2 motors and the free moving castor wheel at the back. This allows for movement in all directions and will not require a lot of space in order to turn.

Fig. 2. The entire bot along with the camera and the pipe attached to it.

B. Odom generation

Odometry is essential for autonomous navigation, it allows the robot to know the amount of distance it has covered and it which direction allowing it to generate its own path and guess its position. This is done with the help of the motor encoders sensors. This is mainly done in order to reduce error accumulation and it is also crucial for Simultaneous Mapping And Localization(SLAM). This will facilitate environment mapping and precise positioning. Odometry feedback is vital for closed-loop control, allowing real-time adjustments to optimize robot movements.

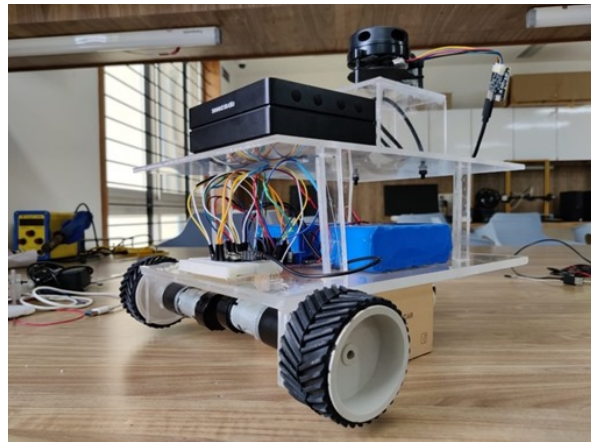

Fig. 3. The bot only containing the bottom 2 layers (Does not include the camera and microphone)

C. Simultaneous Localization and Mapping

Autonomous navigation is enabled through SLAM technique. Adaptive Monte Carlo Localization(AMCL) is a probabilistic localization system which allows to robot to understand its own position in the map. The robot plans its paths and moves around the environment with the help of a cost map which is in grayscale, where black represents an obstacle and white represents free space, any other shade in between shows that the robot is not sure about that location and that it has not been mapped correctly.

D. Autonomous Navigation

Autonomous Navigation involves the process of generating paths from source node to goal node based on the occupancy grid and the SLAM data that is generated from the map and the encoder readings. This is employed with the help of a nav2fn stack which is a popular navigation package provided by ros. It uses Dijkstra's algorithm in order to create these paths. 2 paths are generated, a local path and a global path this allows the robot to avoid obstacles that maybe present along the global path, by dynamically changing the local path during navigation.

E. Face Detection

A USB Webcam is used to capture images of the patients, these images are then stored in designated folders for analysis. To achieve precise face detection, the OpenCV Deep Neural Netword (DNN) face detection module is utilized, ensuring accurate identification of facial features in the captured images. The identified faces are then cropped and stored separately for further processing. After the completion of every trip, the system is reset and the images of the patients faces are deleted allowing each new patient undergoes the same procedure when they use the system.

F. Facial Embeddings Generation and training

The OpenFace pre-trained model was chosen for its efficacy in generating facial embeddings, numerical representations that capture distinctive features of faces, deemed essential for recognition purposes. The embeddings are then systematically extracted from the images that have been captured by the USB webcam with the help of OpenCV DNN face detection module. These embeddings are stored for training the model. Due to the limited training data, a Random Forest Classifier was selected for live recognition since it is very adaptable to such scenarios.

G. Live Recognition

The USB Webcam sends a video stream in which the faces are detected and the embeddings are generated from these faces. The trained model is then applied to these generated embedding for live facial recognition. This approach allows for the continuous processing of real-time video streams. The model can promptly and accurately recognize faces based on their unique embeddings.

H. Creation of ROS2 Package

ROS2 packages have been developed so that the facial recognition system can be integrated into the robotic environment. This package allows communication between the different nodes that are setup for each of the system components. This allows for effective deployment of the system within ROS2.

I. Speech Recognition and speech to text

In the methodological approach employed for this research, robust speech recognition was achieved through the utilization of the Google Speech Recognition API, coupled with the mtranslate library for Kannada to English translation. For the purposes of this reasearch Google Speech Recognition API is used to achieve robust speech recognition and the mtranslate library is used for Kannada to English translation.

J. NLP model creation

The NLP model that has been implemented uses Hugging Face Transformers, specifically Distil BERT. A fine-tuning has been done with the custom dataset of symptoms for eight medical departments. A custom data set is generated for the eight medical departments with the help of Python Faker library. Achieved 99.33 accuracy on the testing data.

K. Integration of NLP with ROS

The Goal Pose Publisher Node played an essential role in integrating human language with robotic actions by converting natural language instructions generated through NLP into commands executable by the robotic system. The synchronization of the NLP predictions with ROS was designed with focusing on efficiency, ensuring real-time adaptability of the robotic system to continuously changing and evolving user inputs. This integration while not only advancing the capabilities of the system in understanding and responding to natural language commands but also helped to enhance its overall adaptability and responsiveness in real-time scenarios.

L. Implementation Details

The robot has 2 layers and a PVC pipe that house all of the hardware components required for the project. It utilizes odometry data and SLAM techniques to perform autonomous navigation. Facial recognition is performed with the help of OpenCV DNN and the NLP is handled by Distil BERT. Speech Recognition is done with the help of Google API and kannada to english translation is done through mtranslate library. A Random Forest Classifier is used to identify the patients after the embeddings are generated from their facial features. The packages allow for seamless communication between The face recognition module and ROS. The Goal Publisher Node performs the task of converting NLP commands to goal poses in ROS.

IV. RESULTS

A. Mapping result

The successful mapping of the environment and the autonomous navigation of the robot to reach the required goal pose represent noteworthy milestones in the robotics system. This capability not only demonstrates the robot's spatial awareness but also its ability to convert this awareness into deliberate and precise movements, improving its overall efficiency in navigating through healthcare environments.

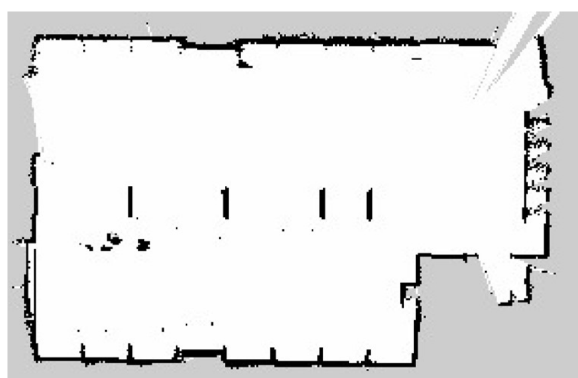

Fig. 4. The result of map obtained using the robot is shown.

B. Face Recognition Accuracy:

Under optimal lighting conditions, the facial recognition component showcased satisfactory performance. The component could greatly benefit from further training with more diverse datasets, accounting for changes in lighting, as well as considering additional facial recognition algorithms. These steps could help achieve further improvements in accuracy.

TABLE I
PATIENT TRACKING MODULE

Class	Precision	Recall	F1-score	Support
Person	0.9841	0.9118	0.9466	68
Unknown	0.9318	0.9880	0.9591	83

TABLE II
PATIENT TRACKING MODULE

Fold1	Fold2	Fold3	Fold4	Fold5
0.9337	1.00	0.9006	0.9337	0.9466

C. NLP Delay

The observed delay in processing of the natural language instructions, which is attributed to the internal workings of the Google API calls and server response time, introduces a consideration for optimizing the system's latency. The latency of the module could further be reduced by exploring alternate speech recognition options or optimizing the current infrastructure. In conclusion, each challenge encountered during both the development and integration processes provides valuable insights and opportunities for refinement.

TABLE III
DEPARTMENT PREDICTION MODULE

Model	Accuracy	Precision	Recall	F1-score
Pre-trained Distil Bert	93.33	0.94	0.93	0.93

V. PERFORMANCE

A number of factors and limitations significantly affect the robot's performance. Firstly, achieving accurate facial recognition requires the availability of optimal lighting condition, the absence of which would impair the robot's capacity for accurate identification. The robot can only be effective in tracking patients if their height falls within a particular range that it can capture using its camera; otherwise, precise tracking may not occur. Furthermore, its short operational runtime of 30 minutes and large recharge time of 1 hour places some constraints that require proper scheduling for its strategic deployment and charging. For the robot to be able to effectively communicate with the patients, he/she must speak clearly and coherently to help enable understanding of commands and responses. Moreover, the robot's motors which help it in moving are capped at a speed of sixty revolutions per minute, which may restrict its responsiveness in dynamic environments. While it is equipped with obstacle avoidance capabilities, the robot's effectiveness is limited to level surfaces, with unpredictable moving obstacles or sensor constraints posing challenges. Another thing to take note of is that low internet bandwidth can disrupt internal communication affecting the overall performance.

VI. CONCLUSIONS

The ROS integrated patient symptom analysis system has successfully demonstrated the potential for NLP in determining treatment departments and converting these forecasts into actionable objectives for a robot. The patient's symptoms are analyzed which determines the department and the robot is able to track the patient and guide them to the destination effectively. This integration of several components indicates how healthcare can be better automated in a manner that reduces alignment issues in care of patients.

Fig. 5. The patient following the robot as well as performing static obstacle avoidance.

Fig. 6. The robot stopping as the patient stops following it.

VII. FUTURE WORK

User Interface for Manual Control: User interface implementation for manual control of the robot can be an invaluable choice. This function is important in cases where human supervision is necessary or during unpredictable moments that need immediate decision making.

Map Selection Functionality: Multiple maps enable the bot to adapt to different environmental conditions. This attribute is particularly helpful in large healthcare facilities with diverse arrangements or when moving from one department to another within a hospital that has different maps.

Improving Robot Aesthetics and Sturdiness: Physical modifications made on the robot's body can increase its aesthetic and ensure it's durability.

Continuous Model Training: Have a system that provides for continuous model training, so that NLP keeps up with the latest medical terminologies and treatment protocols.

VIII. Acknowledgment

The authors would like to express their sincere appreciation to Professor Vadiraja A, Department of Computer Science and Engineering, PES University for his priceless guidance, unwavering support as well as constant motivation throughout the development of this research shown in UE20CS390A-Capstone Project Phase – 2.

The authors also acknowledge Dr. Mamatha HR who is the Chairperson of the Department of Computer Science and Engineering at PES University because she was knowledgeable and gave an unwavering support towards it.

Gratitude is therefore expressed to Chancellor, Dr. M.R Doreswamy, PES University, Prof. Jawahar Doreswamy, Pro Chancellor - PES University, Dr.Suryaprasad J, Vice-Chancellor, & Dr.B.K.Keshavan, Dean Faculty - PES University who without which valuable opportunities at each step could not have been provided till enlightenment in this case.

References

[1] K. Vamsi, P. Alle, T. Brichpuria and P. Malarvizhi, "ROS Based Autonomous Disinfectant Mobile Robot for Hospitals," 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2021, pp. 94-100, doi: 10.1109/ICECA52323.2021.9675920.

[2] V. Andreas, A. A. S. Gunawan and W. Budiharto, "Anita: Intelligent Humanoid Robot with Self-Learning Capability Using Indonesian Language," 2019 4th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS), Nagoya, Japan, 2019, pp. 144-147, doi: 10.1109/ACIRS.2019.8935964.

[3] Sleaman, Walead & Yavuz, Sırma. (2020). Indoor mobile robot navigation using deep convolutional neural network. Journal of Intelligent & Fuzzy Systems. 39. 5475-5486. 10.3233/JIFS-189030.

[4] S. Gatesichapakorn, J. Takamatsu and M. Ruchanurucks, "ROS based Autonomous Mobile Robot Navigation using 2D LiDAR and RGBD Camera," 2019 First International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP), Bangkok, Thailand, 2019, pp. 151-154, doi: 10.1109/ICA-SYMP.2019.8645984.

[5] M. Kulkarni, P. Junare, M. Deshmukh and P. P. Rege, "Visual SLAM Combined with Object Detection for Autonomous Indoor Navigation Using Kinect V2 and ROS," 2021 IEEE 6th International Conference on Computing, Communication and Automation (ICCCA), Arad, Romania, 2021, pp. 478-482, doi: 10.1109/ICCCA52192.2021.9666426.

[6] S. F. Ramadhan, M. Taufik, D. Novita and A. Turnip, "Design of 2D LiDAR-based Indoor SLAM for Medical Robot Covid-19," 2021 International Conference on Artificial Intelligence and Mechatronics Systems (AIMS), Bandung, Indonesia, 2021, pp. 1-5, doi: 10.1109/AIMS52415.2021.9466015.

[7] D. Hutabarat, M. Rivai, D. Purwanto and H. Hutomo, "Lidar-based Obstacle Avoidance for the Autonomous Mobile Robot," 2019 12th International Conference on Information & Communication Technology and System (ICTS), Surabaya, Indonesia, 2019, pp. 197-202, doi: 10.1109/ICTS.2019.8850952.

[8] Joao Guerreiro, Daisuke Sato, Saki Asakawa, Huixu Dong, Kris M. ˜ Kitani, and Chieko Asakawa. 2019. CaBot: Designing and Evaluating an Autonomous Navigation Robot for Blind People. In Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19). Association for Computing Machinery, New York, NY, USA, 68–82. https://doi.org/10.1145/3308561.3353771

[9] Faiza Gul, Wan Rahiman & Syed Sahal Nazli Alhady — Kun Chen (Reviewing editor) (2019) A comprehensive study for robot navigation techniques, Cogent Engineering, 6:1, DOI: 10.108

[10] Xiao, Xuesu & Liu, Bo & Stone, Peter. (2020). Agile Robot Navigation through Hallucinated Learning and Sober Deployment.

[11] Y. Jin, "Research on Path Planning of Airport VIP Service Robot Based on A* Algorithm and Artificial Potential Field Method," 2021 5th Asian Conference on Artificial Intelligence Technology (ACAIT), Haikou, China, 2021, pp. 485-489, doi: 10.1109/ACAIT53529.2021.9731342.

[12] S. Flores and J. Jost, "Person Re-Identification on a Mobile Robot Using a Depth Camera," 2022 IEEE 31st International Symposium on Industrial Electronics (ISIE), Anchorage, AK, USA, 2022, pp. 115-122, doi: 10.1109/ISIE51582.2022.9831515.

[13] Koide, Kenji & Miura, Jun & Menegatti, Emanuele. (2020). Monocular Person Tracking and Identification with On-line Deep Feature Selection for Person Following Robots. Robotics and Autonomous Systems. 10.1016/j.robot.2019.103348.

[14] Chebotareva, Elvira & Hsia, Kuo-Hsien & Yakovlev, Konstantin & Magid, Evgeni. (2020). Laser Rangefinder and Monocular Camera Data Fusion for Human-Following Algorithm by PMB-2 Mobile Robot in Simulated Gazebo Environment. 10.1007/978-981-15-5580-0 29.

[15] V. Advaith, A. Shivkumar and B. S. Sowmya Lakshmi, "Parts of Speech Tagging for Kannada and Hindi Languages using ML and DL models," 2022 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2022, pp. 1-5, doi: 10.1109/CONECCT55679.2022.9865745.

[16] M. V, D. Manur, M. F. Farook, C. G. K S and S. M. Desai, "Doctor Patient Assistance System Using Artificial Intelligence," 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2021, pp. 534-538, doi: 10.1109/ICESC51422.2021.9532856.

[17] L. Athota, V. K. Shukla, N. Pandey and A. Rana, "Chatbot for Healthcare System Using Artificial Intelligence," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2020, pp. 619-622, doi: 10.1109/ICRITO48877.2020.9197833.

[18] T. N. M. Kumar, A. Jayan, S. Bhat, M. Anvith and A. V. Narasimhadhan, "Monophone and Triphone Acoustic Phonetic Model for Kannada Speech Recognition System," 2022 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India, 2022, pp. 336-340, doi: 10.1109/WiSPNET54241.2022.9767115

[19] G. S. Pavan, N. Kumar, K. Karthik N. and J. Manikandan, "Design of a Real-Time Speech Recognition System using CNN for Consumer Electronics," 2020 Zooming Innovation in Consumer Technologies Conference (ZINC), Novi Sad, Serbia, 2020, pp. 5-10, doi: 10.1109/ZINC50678.2020.9161432.

[20] Kulkarni, P., 2021. Named Entity Recognition on Kannada Low Resource Language using Deep Learning Models (Doctoral dissertation, Dublin, National College of Ireland).

[21] Dharwadkar, Rashmi & Deshpande, Neeta. (2018). A Medical ChatBot. International Journal of Computer Trends and Technology. 60. 41-45. 10.14445/22312803/IJCTT-V60P106.

[22] Nev´ eol, Aur ´ elie & Dalianis, Hercules & Velupillai, Sumithra & Savova, ´ Guergana & Zweigenbaum, Pierre. (2018). Clinical Natural Language Processing in languages other than English: Opportunities and challenges. Journal of biomedical semantics. 9. 12. 10.1186/s13326-018-0179-8.

[23] J. N. Jothi, S.Poongodi, V.Chinnammal, L.Kannagi, M.Panneerselvam and R. T. Prabu, "AI Based Humanoid Chatbot for Medical Application," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022, pp. 1135-1140, doi: 10.1109/ICOSEC54921.2022.9951910.

[24] T. Kakusa and M. Hsiao, "Converting Natural Language Text to ROSCompatible Instruction Base," 2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), Taichung, Taiwan, 2018, pp. 219-226, doi: 10.1109/AIVR.2018.00051.