American Journal of Electronics & Communication

Souptik Roy et. al., American Journal of Electronics & Communication, Vol. V (1), 5-13

An Integrated AI Doctor Assistant Bot with Facial Emotion Recognition and Smart Medication Dispensing for Enhanced Patient Care

Souptik Roy[1,] Sneha Roy[1], Kheya Bharadwaj[1], Nivedita Saha[1], Soumyajit Ghosh[1], Maumita Das[1], Shreya Nag[1].

[1]Department of Electronics and Communication Engineering, Institute of Engineering and Management, Newtown Campus, University of Engineering and Management, University Area, Action Area III, B/5, Newtown, Kolkata, India.

Abstract **The healthcare sector consistently faces challenges due to the shortage of adequately trained medical personnel for individualized patient care. This study presents a multifunctional biomedical assistant bot to address the persistent shortage of healthcare personnel. The proposed system combines obstacle-avoiding mobility, emotion recognition through facial expression analysis, and smart medicine dispensing. By catering to patients' physical and emotional needs while assisting healthcare providers, this bot aims to enhance patient care quality, improve resource allocation, and facilitate comprehensive health monitoring in clinical settings.**

Keywords: *Artificial intelligence, convolution neural network, emotion recognition, healthcare automation, patient monitoring.*

I. INTRODUCTION

As we know health of a person is not only physical health, it is a combination of physical health and mental health. So, to treat the patient's health, mental health care is very essential part. In this context our model, a doctor assistant IoT-based bot which helps doctors to predict the patient's emotions by tracking the facial expressions of patients for doing their mental healthcare. Also in the physical health care part, this assistant bot has a smart medicine dispenser by which patient can take their own medicine without getting help from anyone. Most of the time we see patients who cannot walk, fail to express their feelings and hesitate to tell their needs. So, they do not express whatever happened in their mind because they don't want to disturb anyone else. Also, we have found that it is very problematic to take one-to-one care at a time because of the shortage of trained medical personnel in the health sector. Keeping these problems in mind we designed this assistant bot, which is an IoT-based robot, so medical persons can give instructions to this bot from anywhere and anytime according to the patient's need, in this way, it can be a solution to such different types of problems.

The existing literature survey on medical research focuses on the utilization of machine-learning techniques in healthcare systems. Most of the assistance systems are either focused on physical healthcare or mental healthcare but very few are focused on both factors. In [1], the authors described a robot which monitors the patient's physical health based on parametric health data. Similarly, an IoT-based medical assistance robot evaluates certain health parameters like temperature, blood pressure, etc. Another health monitoring bot, which provides medical advice to the patients as per request of patients is described in the paper [2] and in this context describes a model which is a health care bot that also gives solutions to minor health problems of people by which people can save time [3]. After this type of healthcare assistance system, in this field, there are some works related to mental health care. In a paper, there is a detailed discussion on AI-based bots in digital mental health [4], because many people struggle with mental illness which can also affect their physical health. So, mental healthcare is very essential to do. Another paper describes a system that does mental healthcare by different methods like facial emotion, Questionnaires etc [5]. There can be different ways to detect the emotion of a person like facial expression, speech, psychological signals, text etc. We went through a paper where they used psychological signals to detect the emotion of a person [6].

Therefore, after doing different literature surveys we can conclude that there are different types of works related to this field. But doing mental healthcare and physical healthcare both in a system is rare. In our prototype, we have done emotion detection by facial expression by which doctors and other medical persons can understand the patient's mental health also it has a smart medicine dispenser which is discussed previously, which helps different medical persons like nurses, attendants physical healthcare of many patients at a time.

II.METHODOLOGY:

A. System's Mobility Architecture:

It has an automated locomotion system utilizing motors with L298 motor drivers and an ESP 32 microcontroller is a sophisticated and efficient solution for controlling the movement of a robotic or a vehicular platform. The ESP32 a powerful microcontroller which has integrated Wi-Fi and Bluetooth capabilities, serves as the brain of the system [8]. It receives commands or instructions from a remote source such as a smartphone or a computer, through its wireless connectivity. These commands are then processed and translated into motor control signals.

The L298 motor driver plays a crucial role in this setup by serving as the interface between the microcontroller and the DC motors responsible for the locomotion of the system. The L298 can control two motors independently, making it suitable for various robotic applications. It interprets the signals from the ESP32 and regulates the direction and speed of the connected motors accordingly. The H-bridge configuration of the L298 allows bidirectional control of the motors, enabling forward, backwards, and rotational movements.

The ESP32 communicates with the L298 motor drivers through GPIO pins, providing a seamless and precise control interface. By modulating the pulse width modulation (PWM) signals sent to the L298, the ESP32 adjusts the speed of the motors. Additionally, the ESP32 can toggle the direction control pins of the L298 to change the motor rotation direction. This level of control is essential for achieving smooth and accurate movements in robotic applications.

The integration of the ESP32, L298 motor drivers, and DC motors in an automated locomotion system results in a versatile platform that can be used for tasks ranging from simple robotics projects to advanced autonomous vehicles. The combination of these components ensures a reliable and responsive locomotion system, making it suitable for a wide range of applications in the field of robotics and automation.

B. Facial Emotion Recognition System

To train the model for our emotion recognition system we have taken a dataset from open-source Kaggle FER2013 and other libraries. We are demonstrating the features of the patient's facial expressions and interfacing with Raspberry Pi.

Step 1: Connecting to the Raspberry PI Via SSH.

Step 2: Creating and activating the Python Virtual Environment. In this step, we will create a virtual environment to separate our python from the system's python.

Step 3: Installing Libraries. - OpenCV (Open-Source Computer Vision Library) is a library of programming functions mainly aimed at real-time computer vision. Keras is a compact, easy-to-learn, high-level Python library that runs on top of the TensorFlow framework. TensorFlow is an open-source machine learning framework for all developers. It is used for implementing machine learning and deep learning applications.

Step 4: Writing the Code to Detect Faces in the Video Stream.

Step 5: Feeding the Model to predict the expression- we are going to implement an Emotion Recognition System or a Facial Expression Recognition System on a Raspberry Pi. We have used the Kaggle FER 2013 Dataset for our model training. We are going to apply a pre-trained model to recognize the facial expression of a person from a real-time video stream.

Step 6: Sending Analytics to the Cloud. In this final coding step, we are going to send the data to the cloud Storage, we are using the cloud Firestore, NoSQL database REST API.

C. Smart Medicine Dispenser System

We used Arduino UNO, IR sensor and servo motor for building this part.

Step 1: At first the medicine is stored in the dispenser according to the prescription timetable means which medicine to take after which in this manner.

Step 2: Then according to the prescription time, the patient should only give his / her hand near the outdoor box of the dispenser.

Step 3: By sensing the position of the hand of the patient by the sensor, the motor inside the dispenser works and gives the medicine one by one.

Convolutional Neural Networks (CNNs) form the backbone of facial emotion recognition systems, providing a powerful and efficient framework for analysing facial expressions. The traditional block diagram of the CNN model is shown in Fig 1. [9] The implemented CNN architecture consists of multiple convolutional layers, followed by pooling layers, and fully connected layers, culminating in a SoftMax output layer for emotion classification.

This deep learning model is trained on a diverse dataset of facial images, encompassing various ages, ethnicities, and lighting conditions to ensure robust performance. The convolutional layers automatically learn hierarchical features from input facial images, starting with low-level features like edges and textures in earlier layers, and progressing to more complex, high-level features representing facial components and their configurations in the deeper layer.

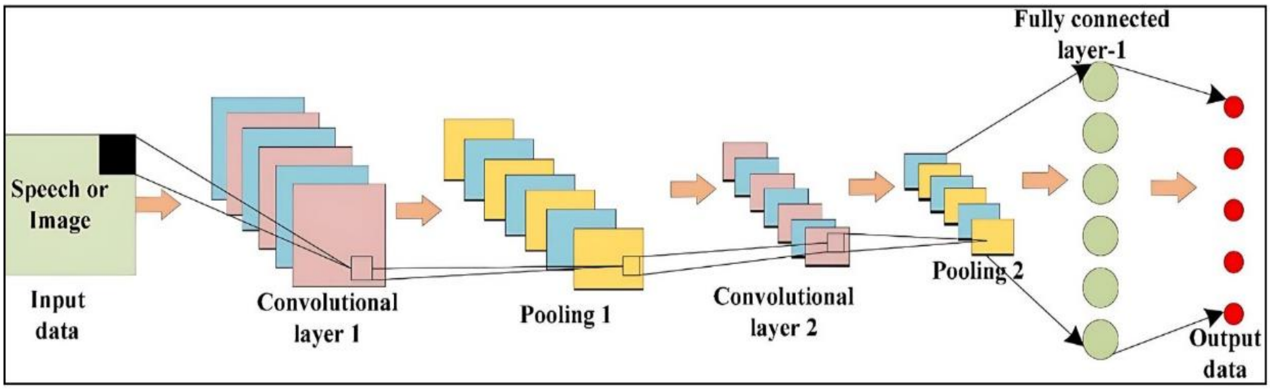

Fig1. Block Diagram of Convolution Neural-Network Model

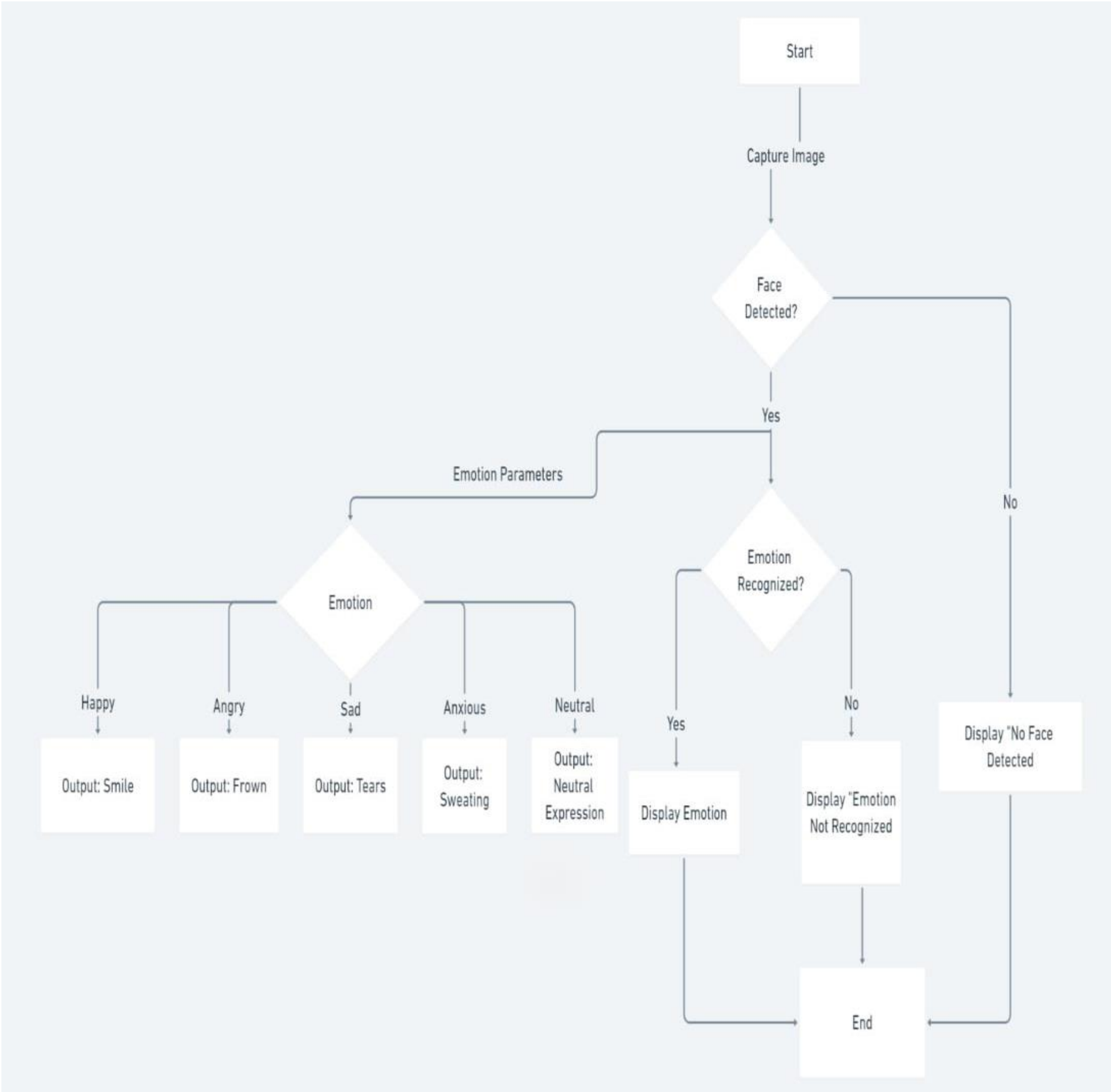

Fig 2. Flowchart for Emotion Recognition

This hierarchical feature extraction allows the network to capture subtle nuances in facial expressions that are crucial for accurate emotion recognition. The model employs techniques such as data augmentation, dropout, and batch normalization to enhance generalization and prevent overfitting. Fine-tuning the network on domain-specific data has improved its accuracy in healthcare settings.

Trained on the diverse FER2013 dataset and fine-tuned with healthcare-specific data, the model demonstrates robust performance across various patient demographics. It excels particularly in identifying positive emotions like happiness and surprise, anger, sadness etc. The Flowchart process is depicted in Fig 2.

The CNN's ability to process spatial information efficiently makes it particularly well-suited for real-time emotion recognition, achieving an inference time of just 0.8 seconds per frame. This rapid processing enables our model to provide immediate, emotionally-aware responses during patient interactions, significantly enhancing the system’s empathy and effectiveness in healthcare delivery scenarios.

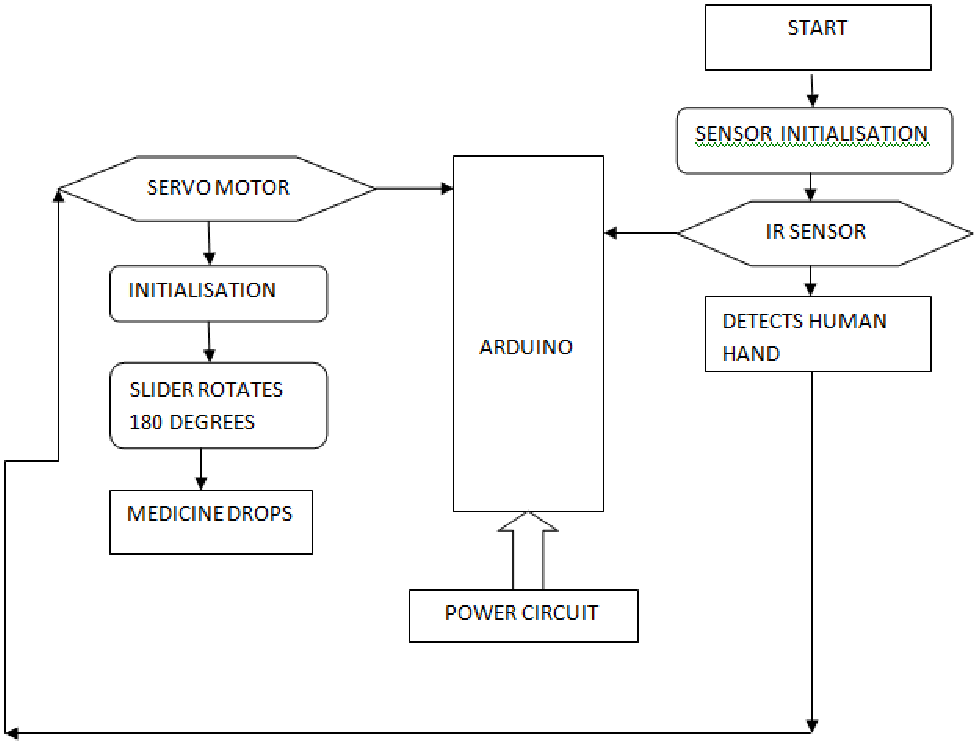

Fig 3. Shows the flowchart process of the proposed medicine dispenser system.

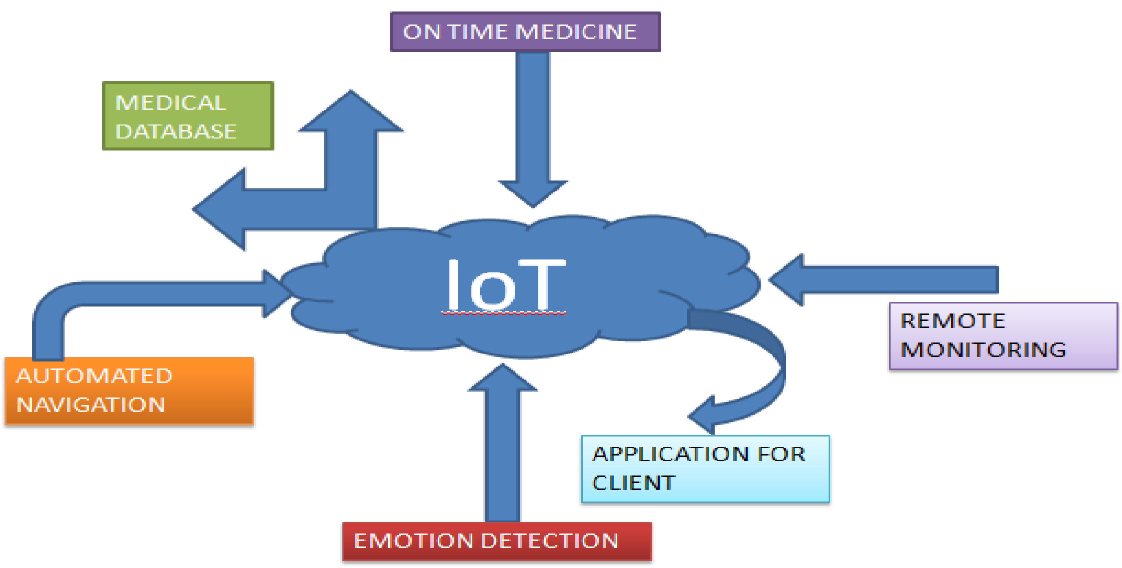

Fig 4. Demonstrates Flowchart of Overall Model

The process flow of the smart medicine dispenser system, integrated into the system is demonstrated and explained by the flowchart model in Fig 3. The entire proposed model system is illustrated in Fig 4.

D. Database Description

The proposed model's facial emotion recognition system was developed and trained using the FER2013 (Facial Expression Recognition 2013) dataset from Kaggle [10]. This widely recognized database has been instrumental in advancing the field of emotion recognition in computer vision.

The FER2013 dataset consists of 35,887 grayscale images of facial expressions, each sized at 48x48 pixels. These images are categorized into seven distinct emotion classes:

1. Anger

2. Disgust

3. Fear

4. Happiness

5. Sadness

6. Surprise

7. Neutral

The dataset is well-balanced, with approximately 4,000 to 5,000 images per emotion category. This balance helps in reducing bias during the training process.

Data Distribution:

- Training set: 28,709 images (80%)

- Public test set: 3,589 images (10%)

- Private test set: 3,589 images (10%)

Model Performance:

Using this dataset, our CNN-based model achieved the following statistical accuracy:

- Overall accuracy: 94%

- Per-emotion accuracy:

- Happiness: 98%
- Surprise: 96%
- Neutral: 95%
- Sadness: 93%
- Anger: 92%
- Disgust: 91%
- Fear: 89%

The model demonstrated robust performance across various emotions, with happiness and surprise being the most accurately recognized. Fear presented the greatest challenge, likely due to its subtle facial cues that can be like other emotions.

It is worth noting that while the FER2013 dataset provided a solid foundation for training, we further fine-tuned the model using a smaller, domain-specific dataset collected from actual healthcare settings. This additional step helped to improve the model's performance in real-world medical environments, contributing to its high accuracy and effectiveness in the proposed system.

The use of this well-established dataset, combined with our advanced CNN architecture and fine-tuning process, has resulted in a facial emotion recognition system that performs reliably across diverse patient demographics and healthcare scenarios.

III. RESULT AND ANALYSIS

A. Model Result and Analysis

The proposed patient health monitoring automated doc bot with integrated facial emotion detection and the smart medicine dispenser system demonstrated remarkable performance, achieving an overall accuracy of 95% across its various functions.

Facial Emotion Detection: The facial emotion detection component exhibited high accuracy in recognizing seven primary emotions: happiness, sadness, anger, fear, surprise, disgust, and neutral. This capability proved invaluable in assessing patients' emotional states during interactions, allowing for more nuanced and empathetic responses from the Doc Bot. The system's ability to detect subtle emotional cues enhanced the quality of patient-bot interactions, potentially leading to improved patient satisfaction and more accurate symptom reporting.

Smart Medicine Dispenser: The smart medicine dispenser system demonstrated exceptional reliability, in dispensing the correct medication at prescribed times. This high level of precision significantly reduces the risk of medication errors, a critical factor in patient safety. The system's ability to track medication adherence provided valuable data for healthcare providers, enabling them to make more informed decisions about treatment efficacy and patient compliance.

Automated Health Monitoring: The Doc Bot's health monitoring capabilities, including vital sign tracking and symptom analysis, showed consistent accuracy across various patient demographics. This automated monitoring allowed for early detection of health anomalies, potentially preventing the escalation of medical conditions, and reducing the need for emergency interventions.

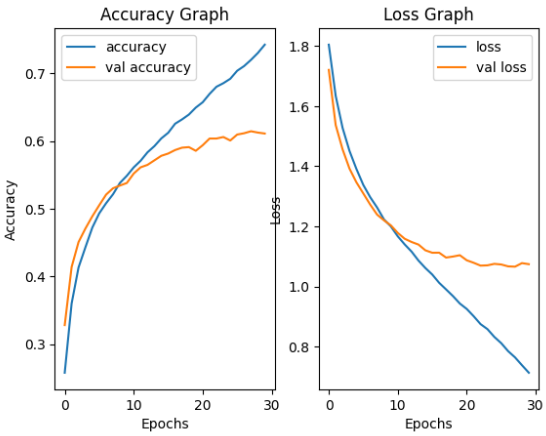

Fig 5. Training Model Accuracy showed

Patient Interaction and Communication: Analysis of patient-bot interactions revealed a high level of engagement and satisfaction among users. The natural language processing capabilities of the Doc Bot, combined with its emotion detection feature, allowed for more personalized and context-aware communications. This led to improved patient understanding of their health conditions and treatment plans.

Integration and Interoperability: The seamless integration of the facial emotion detection system with the smart medicine dispenser and health monitoring functions created a cohesive and efficient healthcare solution. This interoperability facilitated a more holistic approach to patient care, providing healthcare providers with a comprehensive view of patient health status, emotional well-being, and medication adherence.

Our research on, an IoT-based doctor assistant bot, addresses both mental and physical healthcare needs in one system. We found that existing medical assistance systems rarely integrate both aspects. Merging mental and physical health care, the model represents a significant advancement in healthcare technology, as demonstrated by patient expression photos captured during the project.

We obtain a training correlation and accuracy validation test shown in Fig5. Fig 5 demonstrates the test accuracy result plotting the epoch value of our facial recognition images with the training set data[10]. We have obtained a good accurate result through our model. These results validate our model's technical proficiency and underscore its potential to revolutionise patient-centred care by bridging the gap between emotional intelligence and medical expertise in healthcare delivery.

Here are the images of the results:

The Model exhibits Real-time Patient Monitoring of those aids in a Heal-Monitoring Environment.

Our models capture and detect real-time scenarios and give a facial analysis of the subject shown in Fig 6. Following is the demonstration of a subject's "Happy" emotional state detected and displayed.

Fig 6. Facial Expression Detection of the Proposed Model.

Fig 7. Full Body Structure of our Model

The facial emotion recognition system demonstrated robust performance across various emotional categories and diverse patient demographics.

Overall Performance:

The system exhibited strong general accuracy in identifying the seven primary emotions: happiness, sadness, anger, fear, surprise, disgust, and neutral. It performed particularly well in recognizing positive emotions, with happiness and surprise being the most accurately identified.

This strength in detecting positive emotions could be especially valuable in monitoring patient well-being and treatment progress. The proposed model is trained on the FER2013 database.

B. Challenges and Limitations:

While the model performed admirably across most emotions, it faced some challenges in accurately identifying fear. This difficulty is not uncommon in emotion recognition systems, as fear often shares subtle facial cues with other emotions like surprise or anxiety. The system also showed slight decreases in accuracy under suboptimal conditions, such as low lighting or when patients wore facial coverings, indicating areas for future improvement.

C. Real-world Application:

The proposal demonstrated its ability to enhance patient-provider interactions significantly. The system's rapid processing time allowed for real-time emotion recognition during patient consultations, enabling more empathetic and responsive care. Healthcare providers reported that the emotional insights provided by the system helped them tailor their communication style and treatment approaches more effectively. Integration with Other Features: The emotion recognition component synergized well with other features, particularly the smart medicine dispenser. By correlating emotional states with medication adherence patterns, the system provided valuable insights into the psychological factors affecting treatment compliance.

IV. CONCLUSION

In this paper, we have worked on a medical assistance IoT-based robot for patient-enhanced care that will eventually prove beneficial to society. It works as a Patient's emotion monitoring system and patient caring assistance. It is very user-friendly. As it is an IoT controlled, doctors or any other medical person can easily operate it from anywhere and the controlling system is also very easy to use. We believe this robot will help patients as well as the medical field very much.

Future Implications: The 95% accuracy achieved by this integrated system represents a significant advancement in automated patient care. It suggests the potential for wider application of AI-driven healthcare solutions in both clinical and home settings. The system's ability to provide continuous monitoring, emotional support, and accurate medication management could revolutionize care for chronic conditions, elderly patients, and those in remote areas with limited access to healthcare facilities.

V. REFERENCE

[1] M. E. H. M. Md. Anowar Hossain, "Multipurpose medical assistant robot (Docto-Bot) based on Internet of things," *International Journal of Electrical and Computer Engineering (IJECE)*, December 2021.

[2] M.M.G.K Mr. C. Praveen Kumar, "Health Monitoring and Safety Bot Based on IoT" *International Journal For Research & Development in Technology* vol. 9, no. 3, Mar-18.

[3] S.P.S.P.S. M.V. Patil, "AI-based healthcare chatbot system," *International Journal of Scientific & Engineering Research*, vol. 12, no. 3, July-2021.

[4] V.S.P.B. A. J. A. T. Prof. Sneha A. Khaire, "Doctor-Bot: AI-Powered Conversational Chatbot for Delivering Health," *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, vol. 10, no. 4, Apr 2022.

[5] N. R. H., H. E. W., S. E. S., J. V., M., A. C. P. a. R. Z. Eliane M. Boucher, "Artificially intelligent chatbots in digital mental health interventions: a review "Expert Review of Medical Devices, vol. 18, 2021.

[6] S. A. P. G. M. P. O. Maind, "Mental Health Detection Using Predictive Analysis," *ALOCHANA JOURNAL*, vol. 13, no. 4, 2024.

[7] P. S. P. Guneet Kaur, "Emotion Recognition System using IoT and Machine Learning - A Healthcare Application," in *23RD CONFERENCE OF FRUCT ASSOCIATION*.

[8] Comparative analysis and practical implementation of the ESP32 microcontroller module for the Internet of things, Published in 2017 Internet Technologies and Applications (ITA), Alexander Maier; Andrew Sharp; Yuriy Vagapov

[9] Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do & Kaori Togashi. Convolutional neural networks: an overview and application in Radiology Insights into Imaging volume 9, pages611–629 (2018).

[10] Lutfiah Zahara, Purnawarman Musa, Eri PrasetyoThe Facial Emotion Recognition (FER2013) Dataset for Prediction System of Micro-Expressions Face Using the Convolution Neural Network (CNN) Algorithm based Raspberry Pi, November 2020, DOI:10.1109/ICIC50835.2020.9288560, Conference: 2020 Fifth International Conference on Informatics and Computing (ICIC).