Emotion Based Home Automation

Aishwarya S[1] , Apeksha S[2]
Student, Dept. of Information Science and Engineering,
Sir M. Visveswaraya institute of Technology,
Bangalore, India [1,2]

Shivani N S[3] , Sirisha J C[4]
Student, Dept. of Electronics and communication Engineering,
Sir M. Visveswaraya institute of Technology,
Bangalore, India [3,4]

Abstract:- The use of 'Internet of Things' in homes has grown rapidly over the past few years. However, these systems have always been operated via a remote control, voice control or through user applications. In this paper, we are proposing a system design that provides a way for identifying a wide range of human emotions by employing a hybrid deep learning CNN algorithm to index the emotional state and send the same over a ZigBee wireless module to an Arduino. In order to reflect the user's sentiment, the music and lighting are altered by the Arduino. The system is also equipped with an LCD device that displays the current emotion and room temperature as measured by the temperature sensor.

Keywords:- *Home Automation, Neural Networks, CNN-LSTM, Facial expression recognition, IoT, Arduino, Zigbee*

I. INTRODUCTION

Up until recently, only bigger residential and commercial structures had automated central control of all systems throughout the building. Building automation typically involved lighting and temperature systems, rarely offering more than the most basic control, monitoring and scheduling capabilities which were only available from specific control locations within the building itself.

With the growing need to comprehend the social requirements of the users, developing socio-technical or people-oriented systems, such as smart home technology, has never been more difficult. Users have emotional requirements such as the need to be independent, cared for, and secure. Because these requirements cannot be easily translated into functional or non-functional requirements using current software engineering approaches, they are frequently neglected or trivialized in contemporary software development processes. These social requirements, also known as "users' emotional goals", are essential for a system's adoption and cannot be disregarded [1]. Facial expression is one of the most important means to express human emotions in social communication. Automatic facial expression recognition has become a "favorite" topic in the research area.

Home automation is a step toward what is referred to as the "Internet of Things" and can be monitored remotely. Several methods have been proposed in the design of such systems using sensors, biometrics and face detection. The emotion-based home automation system is an AI-powered system which detects human emotions and sets the household environment according to those emotions. This study suggests a technique for identifying human emotions that takes into consideration the entire face examination. It suggests that a wide range of emotions can be reliably identified by examining the eyes, nose, and lips individually.

The goal is to extract the user's live image from a camera and interfere with the user's surroundings by analyzing their emotion. In this system we are considering six emotions that will be examined: happiness, sadness, fear, anger, surprise and neutral. We now extract the corresponding facial landmarks and examine various features and models for predicting human mood. The Arduino board will now work in tandem with home automation to control the lighting and music of the house.

II. PROPOSED SYSTEM

We are proposing a system that will take into account the emotion of the person at the time he is entering the house and then change the lighting and music of the house according to that. This designed system provides a method for identifying human emotions that considers the entire facial analysis, implying that the emotions may be reliably recognized by evaluating the eyes, nose, and lips separately, thus encompassing a wide spectrum of emotions which is done using CNN algorithm to index the six emotional states. Based on the emotion the algorithm had detected, the Zigbee module would transmit the user's mood to the home's IoT gadgets, which would then modify the lighting and music in the space. This would allow the user to unwind and enjoy a stress-free atmosphere as soon as they enter the house. The designed system is integrated with an LCD that displays the current emotion as well as the room's temperature, allowing the user to adjust the AC to his preferences.

This project focuses on accurately predicting the emotion from the user's face expression while entering the test environment using neural networks and implementation of IoT to set the music and lighting of the environment appropriately to the emotion predicted.

The project progressed through 4 phases:

1. Face recognition through camera
2. Using the facial expression to interpret the emotion
3. Wireless communication of the result through Zigbee Module
4. Lighting and music control through home automation

ISSN No:-2456-2165

A. System Architecture:

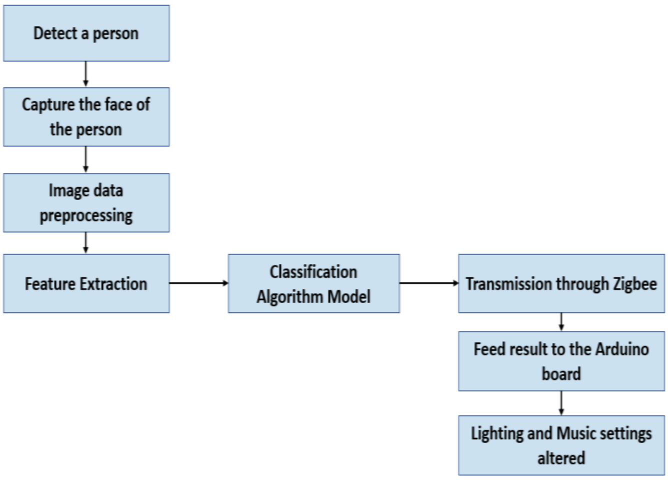

Fig 2.1:- Stages in the working of AI and IoT together as a System

The basic flow of the modules that make up the system architecture is as in Fig. 2.1. The technology initially detects the presence of a person and then captures the individual's face with the help of a camera.

The captured info is then decoded from JPEG to RGB grids of pixels and subsequently into Grayscale grids of pixels using the image data preprocessing technique. This is followed by normalization, which involves dividing the image into 255x255 pixels and then cropping it to 48x48 pixels to match the data set used. The feature extraction procedure is implemented to identify regions of interest such as the eyes, nose, and mouth for emotion recognition. The Convolutional Neural Network (CNN) technique is used to classify the image and assign a value to each of the six emotions. Considering different cultural backgrounds and people's subjective feelings for understanding emotions, facial expressions are divided into six categories, including happiness, anger, surprise, fear, neutral, and sadness [4].

The value, which corresponds to an appropriate emotion, is wirelessly transferred via the Zigbee module. This value is conveyed to the Arduino Board through the serial ports on reception at the receiving Zigbee end. The relevant set of embedded C code is uploaded onto the Arduino that modifies the music and lighting controls to reflect the user's mood.

B. Software Model:

The emotion recognition algorithm was coded in Python and the Seaborn library was utilized to plot graphs and matrices. To create our hybrid deep learning CNN-LSTM model, we used CNN, a convolutional neural network technique.

Convolutional Neural Networks belong within the domain of Deep Learning, and the reason for preferring Deep Learning over Machine Learning is that these algorithms can assess whether an estimate is true or not on their own [3].

Three-dimensional, high-resolution RGB images are captured in real-time by our model. The images are read in an array format and the accuracy of the model is hampered by the higher dimensions of data. Therefore, the image data was reduced into 48 x 48-pixel single-dimensional grayscale images. The images from the training dataset FER 2013 and ck+, which were likewise greyscale, are shown below in Fig. 2.2.

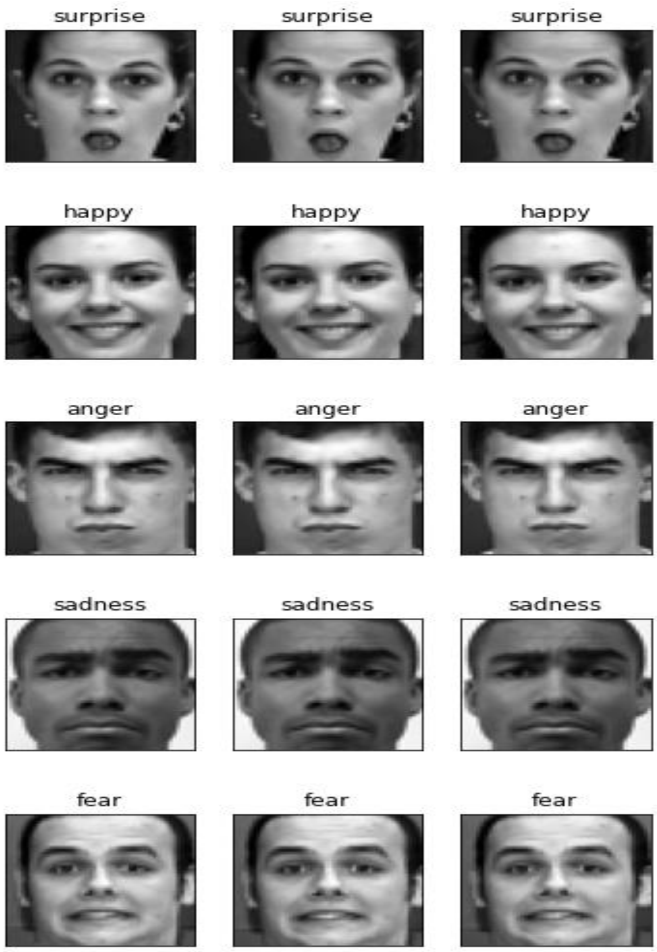

Fig 2.2:- FER2013 Dataset images

More than 35,000 image data was split at a ratio of 7:3 for training and testing respectively. A convolution layer was created for feature extraction from the images. Initially, the RNN approach with a simple short volatile memory was deployed.

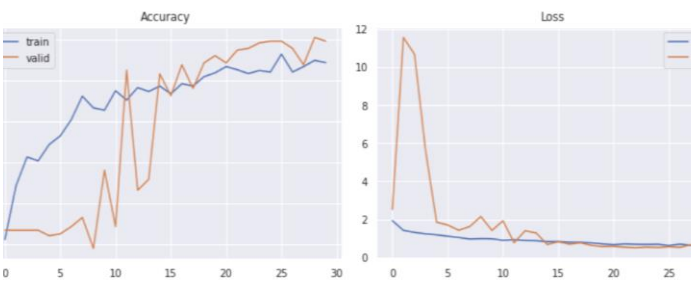

Fig.2.3:- Scikit plot for the RNN approach

Scikit-plot learning allowed us to visualize the training and testing curves of a classifier as in Fig.2.3 and Fig.2.4, which helped in better understanding of the model's behavior.

According to a study, when the gap between relevant information and the point where it is needed is particularly wide, RNN does not seem to connect the information. Compared to a conventional feedforward neural network, an advanced variation of RNN, LSTM is able to memorize the knowledge about prior outputs. [2]. The LSTM outperforms the RNN strategy with its improved feedback loop which promotes the retention capabilities in LSTM.

A typical LSTM cell is represented in Fig. 2.4. It consists of an input gate, a forget gate, an output gate, a memory state, a hidden state and a candidate layer.

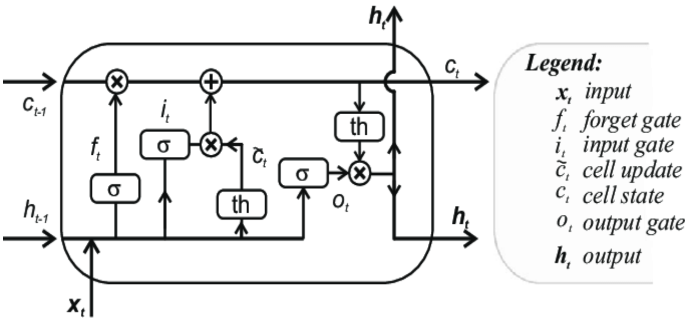

Fig 2.4:- A Typical LSTM cell

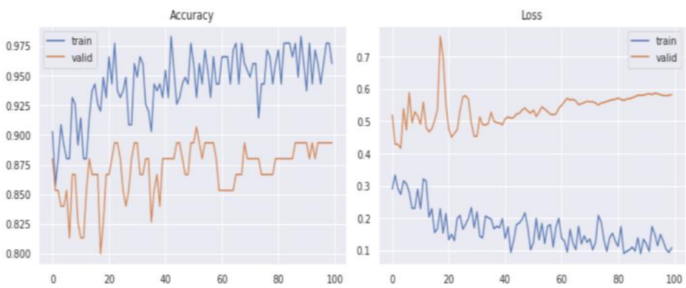

Fig 2.5:- Scikit plot for the LSTM approach

Thus, we employed the LSTM technique as in Fig.2.5. On evaluating the results acquired from the RNN (RestNet 5.0) approach and the LSTM technique, with respect to accuracy aspect in the form of a bar plot as depicted in Fig.2.6, we were able to conclude that LSTM delivered superior results.

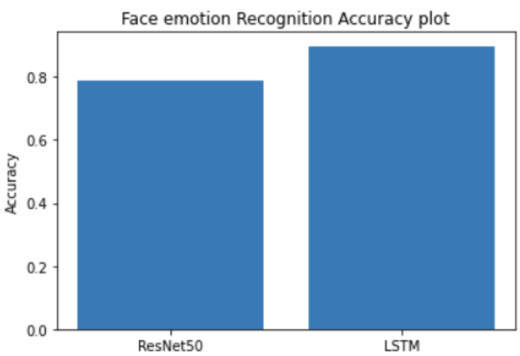

Fig 2.6:- Comparison plot

We used bidirectional LSTM cells, which can learn in both forward and reverse orientations. The second LSTM cell recalls only what the first LSTM cell sent through the memory gate, cleansing the extracted features such that we are left with the most influential elements of the face that define the person's emotion. The LSTM cell recalls the preceding four to five frame-sequences that influence the categorization outcome. To prevent overfitting in our model, the LSTM cell excludes the final pixels that have no influence on the outcome. Following the extraction of LSTM features, batch normalization is performed with the filter size set to 512, and using three channels. We now include callbacks, early stopping, and checkpoints to prevent data overfitting and overlearning during batch normalization. The loss is evaluated every 12 epochs, and if it declines in all 12 cases, the model terminates training to avoid overfitting. Convolution layers in the Convolutional Neural Network model convolve the input image through numerous filters to create a feature map.

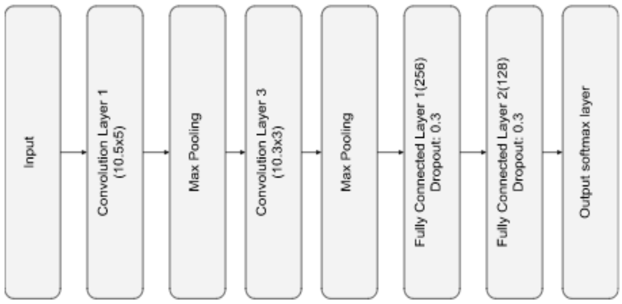

Fig 2.7:- CNN hybrid deep learning model

Depending on the outcome of the SoftMax algorithm, each feature map is integrated into fully linked networks to identify the facial expressions as belonging to a specific class.

C. Working:

The face recognition algorithm is the heart of the system. Fig. 2.8 depicts a block diagram of the transmitter. Initially, the computer receives input in the form of a picture via the camera. The algorithm is used by the system to recognize the observed emotion, which is then assigned an ASCII value. For example, "A" signifies "Happy," "B" represents "Sad," and so on. This computer byte value is mapped to its corresponding UART value for transmission using a USB to UART converter which is linked to a Zigbee module that transmits data to the receiver.

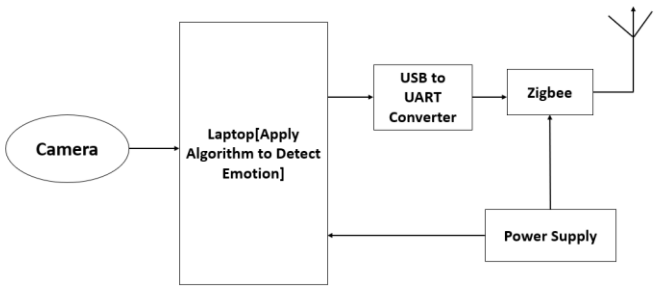

Fig 2.8:- Block diagram of transmitter

ISSN No:-2456-2165

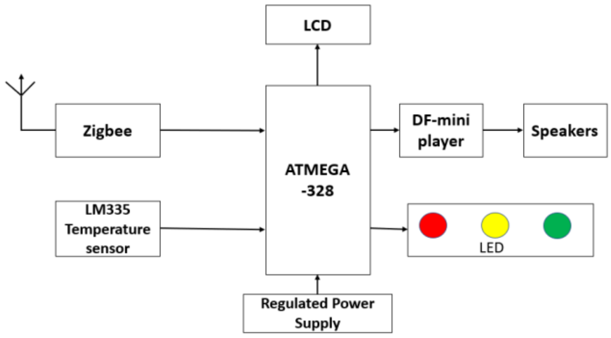

Fig 2.9:- Block diagram of Receiver

The Zigbee module receives wireless data from the transmitter and transfers it to the Arduino board via the serial port. The ATMega328 CPU controls the audio through the DF-mini player, which is coupled to a speaker and the audio files in .mp3 format reflect each emotion saved on an SD card. The LEDs, which are in common-cathode mode, are likewise controlled by the processor. Depending upon the received token, power is supplied to or cut down from the corresponding device [5]. The LM335 Temperature sensor measures the ambient room temperature in real time. The temperature sensor's temperature is reflected on the 16x2 LCD display along with the emotion sensed by the AI system. Safe power levels are ensured by incorporating a Regulated Power Supply (RPS) in the setup.

III. RESULTS

Based on the accuracy and confusion matrix scores, we assessed the performance of our model.

MODEL	ACCURACY (%)
CNN-RestNet5.0	79.55%
CNN-LSTM	89.8%

Table 1:- Model assessment

According to Table 1, the proposed CNN-LSTM method achieves better performance than CNN. The accuracy rate of the hybrid CNN-LSTM model improved by 10.25 %, from 79.55% to 89.8%.

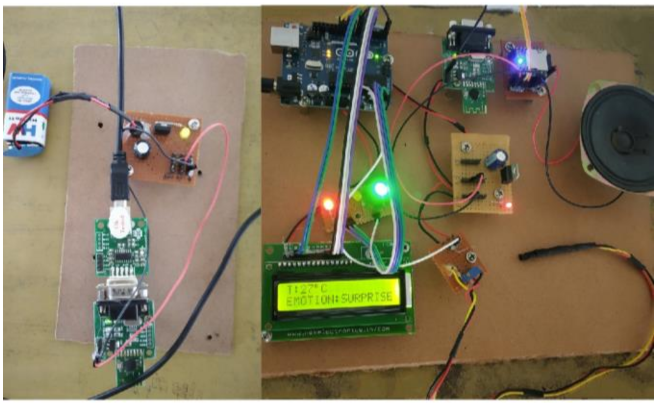

Fig 3.1:- The transmitter and receiver side setup

Fig. 3.1 shows the working prototype of transmission of the algorithmic resultant and reception of the same at the Arduino. The received emotion is displayed on the LCD along with the room temperature sensed by the LM335 sensor and the output is indicated by the LED (red and green for the emotion surprise) and the appropriate music is played by the speakers.

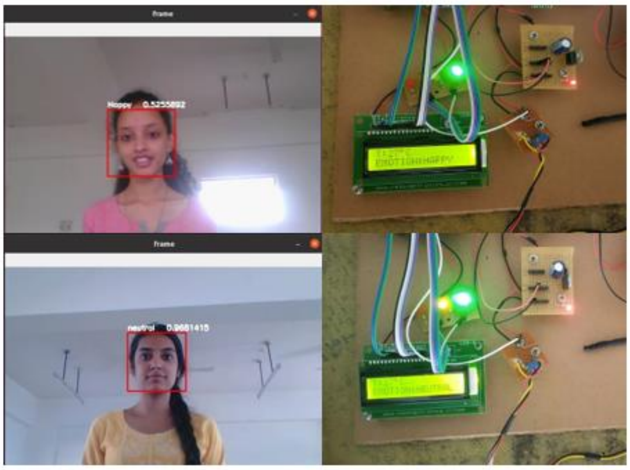

Fig 3.2:- Emotions recognized in the frame and corresponding output obtained on LED, Speaker, and LCD.

A successful integrated implementation of the AI model and IOT system for recognizing emotions and accordingly adjusting light and music settings is demonstrated in Fig.3.2.

IV. CONCLUSION

In this research, we offer a method for automating the management of the music and lighting configuration of the room using the Arduino Uno, based on the facial expressions identified by our hybrid deep learning CNN-LSTM model trained on the FER 2013 and ck+ dataset.

Our algorithm is currently about 95% accurate in identifying an emotion. Without any effort on the part of the user, the system aims to understand the facial expression as soon as they appear in front of the camera.

Typically, when we sense and measure physical quantities, we transmit the data to a computing device so it can be processed, interpreted, and used to deliver services. In this instance, the IOT device handles service by changing our actual environment in accordance with the mood the camera senses.

Our experiment was set up as a straightforward point-to-point network, but by incorporating many nodes into the Zigbee modules, we can create a mesh network that will allow us to connect more devices and even span greater distances. We can also offer the customer more comfort and integrity by including an automated door opening and closure mechanism. By adding other features that enable the system to classify users based on a variety of other criteria, such as location, we may improve the emotion algorithm and advise the user to visit that location while playing music that suits their mood.

Customized cross-domain solutions can also be developed to determine what kind of physical environment is best for the occupants.

REFERENCES

[1]. Maheswaree Kissoon Curumsing, Niroshinie Fernando, Mohamed Abdelrazek, Rajesh Vasa, Kon Mouzakis, John Grundy, "Emotion-oriented requirements engineering: A case study in developing a smart home system for the elderly", Journal of Systems and Software, Volume 147, 2019, Pages 215-229, ISSN 0164-1212, https://doi.org/10.1016/j.jss.2018.06.077.

[2]. Raghvendra Kumar, Nguyen Ho Quang, Vijender Kumar Solanki, Manuel Cardona, Prasant Kumar Pattnaik, "Research in Intelligent and Computing in Engineering, Select Proceedings of RICE 2020," Springer Nature Singapore Pte Ltd. 2021

[3]. A. Deopa, A. Sinha, A. Prakash and R. K. Sinha, "Facial Expression Recognition using Convolutional Neural Network and SoftMax function on Captured Images," 2019 International Conference on Communication and Electronics Systems (ICCES), 2019, pp. 273-279, doi: 10.1109/ICCES45898.2019.9002524.

[4]. P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," J. Personal. Soc. Psychol., vol. 17, no. 2, pp. 124–129, Feb.1971.

[5]. D. L. Kamath, R. Jackson, R. Jacob, C. C. Chacko, C. M. Soumya and P. R. Krishnan, "A novel method for automatic device control using emotion recognition," 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), 2017, pp. 1173-1176, doi: 10.1109/ICICICT1.2017.8342736.