GESTURE BASED HOME AUTOMATION USING CNN AND IOT

JOYMERLINE C.G., RICHARD NOVAS and ROSHNI PRIYA

Dept. of Electronics and Communication Engineering
Madha Engineering College, Kundrathur, Chennai- 69, Tamil Nadu, India

ABSTRACT

Home Automation model using convolution neural networks is designed for providing ease of control of home appliances to the people, especially elderly or those who are physically unable to efficiently perform day-to-day activities. As technology is raising more advancements are made in making the life of these people easier by providing methods easy to monitor and manage. The previous methods have seen the use of accelerometers, which are fixed on the hand as they are considered to be the best tool for carrying out such practices. These models are accurate but not flexible and portable. Our system is based on remote control home automation where-in our remote is an application. In our system, the web camera is used to give gestures as input to the gesture recognition model to filter and predict the gesture. Web-based automation communicates directly with nodemcu esp8266 to control home appliances. The proposed method, allows users to flexibly and portably control multiple household appliances with simple gestures.

Keywords: Gesture, Automation, Home appliances and Technology.

INTRODUCTION

The hand gesture is a nonverbal form of communication. It consists of linguistic content that carries a large amount of information in sign language. It also plays a pivotal role in human-computer interaction (HCI) systems (1). Therefore, automatic hand gesture recognition is in high demand. Since the end of the last century, this field has attracted the attention of many researchers. The importance of automatic hand gesture recognition has increased due to (1) the growth of the deaf and hard-of-hearing populations and (2) the extended use of vision-based and touchless applications and devices such as video games, smart TV control, and virtual reality applications (2).

Robust hand gesture recognition is required as a part of sign language interpretation to help hearing-impaired people. There is a significant communication gap between people who can hear and hearing-impaired people (3). A translation system between gestural language and verbal language will bridge this communication gap. This translation system will facilitate the lives of hearing-impaired people and help them to integrate with society. Unlike sign language translation, hand gesture recognition techniques involve HCI (Human Computer Interaction) to a great degree. Today, HCI has a wide range of applications from video games to tele surgery. As with all time-varying signals, hand gestures cannot be directly compared in Euclidean space because of their temporal dependency. This dependency indicates important discriminative features (4-5).

Temporal misalignment, in addition to massive irrelevant regions in every frame, makes it very hard to extract representative hand-engineered features for hand gestures. For conventional classifiers to perform well, the extracted features should implicate vigorous descriptors. These descriptors code enough information for the inter-frames temporal dependency, as well as the hand position, shape and orientation in each frame. The computed features should be able to minimize the effect of different circumstances like background clutter and occlusions. Therefore, we employed deep learning in this paper as a promising solution (6).

In recent years, many researchers have efficiently exploited convolutional neural networks (CNNs) deep architectures for feature engineering. CNNs have shown excellent performance in fields such as object and speech recognition, image classification, and edge distribution, and human activity recognition. The existence of large datasets that comprise millions of annotated samples is the main reason behind such excellent performance. We propose a well-adapted deep architecture for automatic hand gesture recognition for home automation (7).

Convolution Neural Network: A convolutional neural network, or CNN, is a deep learning neural network designed for processing structured arrays of data such as images. Convolutional neural networks are widely used in computer vision and have become the state of the art for many visual applications such as image classification, and have also found success

in natural language processing for text classification. Convolutional neural networks are very good at picking up on patterns in the input image, such as lines, gradients, circles, or even eyes and faces. It is this property that makes convolutional neural networks so powerful for computer vision. Unlike earlier computer vision algorithms, convolutional neural networks can operate directly on a raw image and do not need any preprocessing (5).

Convolutional neural networks contain many convolutional layers stacked on top of each other, each one capable of recognizing more sophisticated shapes. With three or four convolutional layers it is possible to recognize handwritten digits and with 25 layers it is possible to distinguish human faces. The usage of convolutional layers in a convolutional neural network mirrors the structure of the human visual cortex, where a series of layers process an incoming image and identify progressively more complex features (8).

MATERIALS AND METHODS

IOT Architecture: Internet of Things or IoT is nothing but an evolved version of the Internet, which includes sensors, consumer electronic devices and other embedded systems connected to it besides computers, smartphones and tablets to collect and exchange data with one another. Our system consists of nodemcu, various household electrical devices connected to the nodemcu board via the relay and a computer connected to the Internet (Fig 1).

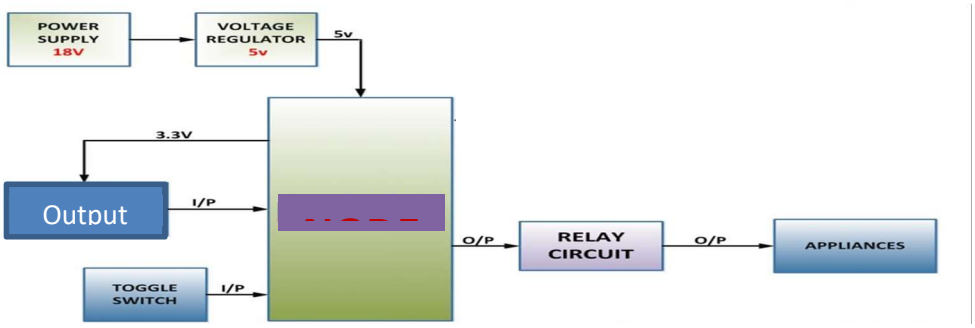


Fig 1: Internet of things architecture

In this proposed system we can give the input as a hand gesture to the web camera. This hand gesture is processed using the background subtraction morphology, and threshold. After the hand is detected on the web camera, the particular gesture is detected. This detected gesture is

converted into the binary value and sent that value as a signal using the Internet. The receiver site receives that gesture signal through the internet. This signal is processed in the nodemcu. Nodemcu connects with the device; according to the signal it controls the device like on the device or off the device.

Material and software: NodeMCU is an open-source Lua based firmware and development board specially targeted for IoT based Applications. It includes firmware that runs on the ESP8266 Wi-Fi SoC from Espressif Systems, and hardware which is based on the ESP-12 module Fig 2).

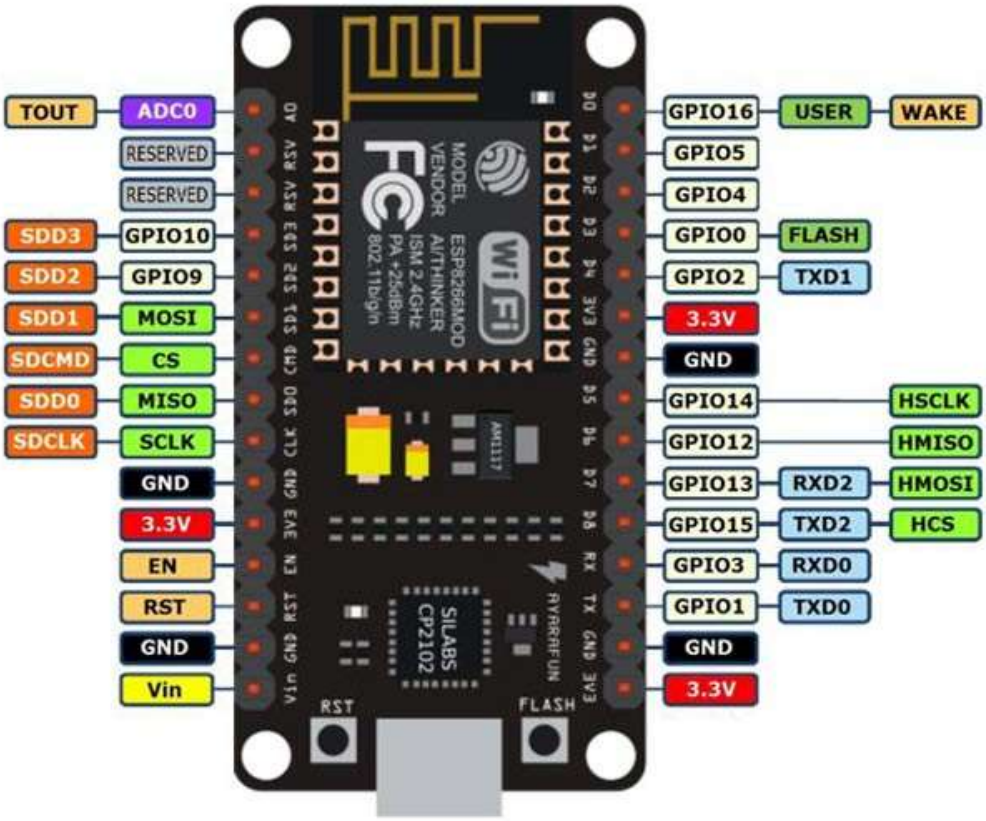


Fig 2: NodeMCU

Four-channel relay module: The four-channel relay module contains four 5V relays and the associated switching and isolating components, which makes interfacing with a microcontroller or sensor easy with minimum components and connections. The contacts on each relay are specified for 250VAC and 30VDC and 10A in each case, as marked on the body of the relays.

Components Present on A Four-Channel Relay Module: Following are the major components present on the four-channel relay module, we will get into the details of this later in the article. 5V relay, terminal blocks, male headers, transistors, optocouplers, diodes, and LEDs.

Four-Channel Relay Module Specifications

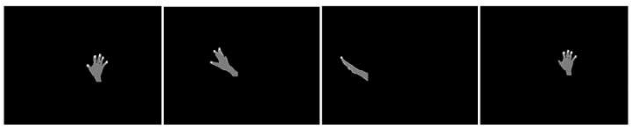
- Supply voltage – 3.75V to 6V
- Trigger current – 5mA
- Current when the relay is active - ~70mA (single), ~300mA (all four)
- Relay maximum contact voltage – 250VAC, 30VDC
- Relay maximum current – 10A

20bn jester v1 datasets: The proposed approaches were implemented in Python. Our experiments were conducted on a machine with a Nvidia GTX 980 ti GPU. In this section, we study the proposed approaches performance on the 20bn jester v1 datasets detailed in section. The 20BN-JESTER dataset is a large collection of labeled video clips that show humans performing pre-defined hand gestures in front of a laptop camera or webcam. The dataset was created by a large number of crowd workers. It allows for training robust machine learning models to recognize human hand gestures.

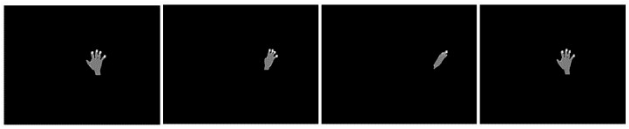
Firebase: Firebase provides a quick way to persist sensory data collected at the device level, and it works great with the APIs. A lot of mobile and device programmers that I have come across struggle with server-side programming. Firebase can really help bridge that gap and make it easier. It will be interesting to see developers use its offline features. If you are new to IoT or in general any device that collects data and needs to transmit it over networks, the golden rule to be assumed is that network connectivity cannot be assumed. As a result, you will need to collect the data offline and when network is available, transmit this over to your server. Firebase with its offline feature can really make this simple for a lot of developers. Firebase has a ton of features including Real-time Database, Authentication, Cloud Messaging, Storage, Hosting, Test Lab and Analytics but I'm only going to use Authentication, Real-time Database.

RESULTS AND DISCUSSION

In order to test the proposed system in a real application, 4 gestures were selected from the 24 gestures dataset that collected previously. Those gestures selected based on the convenience uses in the application. The 4 gestures included 4 dynamic gestures and Figure 3 represents the sample gesture types. The gestures as seen in Figure 3 are “swipe up” “swipe down” “right swipe” and “left Swipe”.

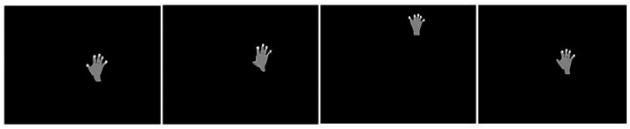


(a) Swipe left



(b) Swipe right

(c) Swipe down



(d) Swipe up

Fig 3. Gestures

Neural Network Training: Flowchart of neural network training is as shown in Fig. 4. First step in training a network using deep learning for an application is to prepare an appropriate dataset and make Train-Test Split depending on the available data. Suitable network is designed or selected (in case of Transfer Learning) for training [9]. Validation Loss is monitored throughout the training process to produce a very less constant value after few epochs, if not then the hyper parameter tuning is performed on model to give lowest possible validation loss values. Model with best validation loss is saved and tested on real world dataset. The model is said to be good if a descent precision and recall values are obtained for new datasets else the model needs to be trained on enhanced dataset for increased performance.

Single Object Detection: Fig. 5. Shows flow chart of single object detection, necessary libraries are imported first and training data is given input via the Google drive. Google-Colab, an online simulation tool for python and Tensor-flow algorithms was used. The algorithm then compiles data and learns form it in a supervised manner [10].

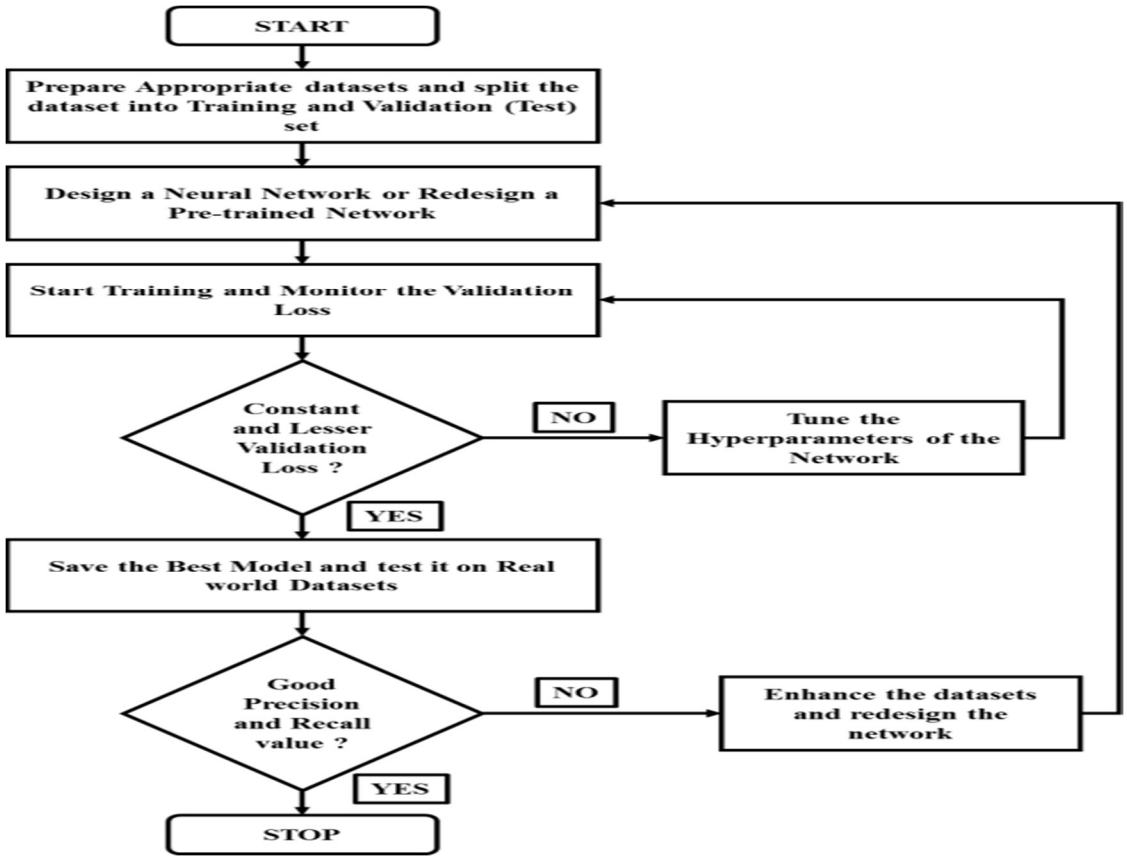


Fig 4. Flowchart of the neural network

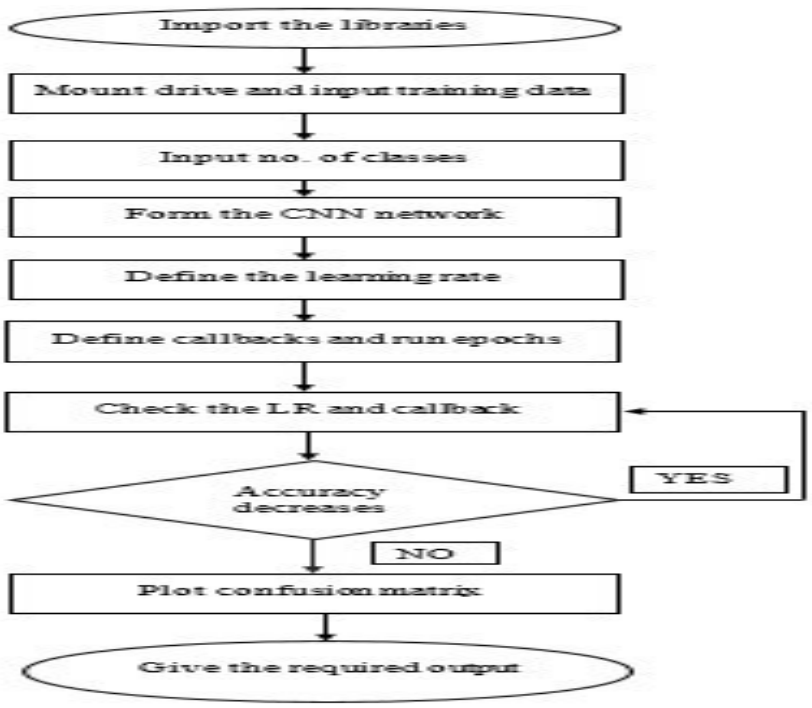


Fig 5. Flowchart of object detection

Figure 6 illustrates the accuracy of the classification of 3DCNN model on the train and validation sets. The plots suggest that the model has a good fit on the problem.

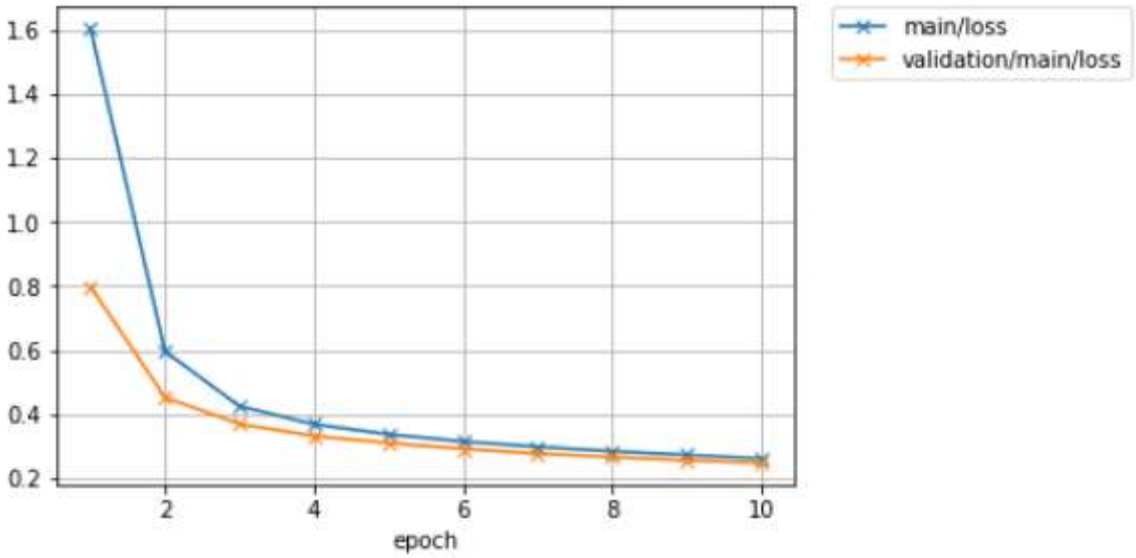


Fig 6. The accuracy of 3DCNN model during training

SUMMARY AND CONCLUSION

The Internet of things is the concept of basically connecting any device with an on and off switch to the Internet. IoT technology can also be applied to create a new concept and huge development space for smart homes to provide intelligence, comfort to improve the quality of life. Home automation is the control of any or all Electrical devices in our home, whether we are there or away. Home automation is also controlling using the gesture (11-12).

Gestures are a natural way of communication, which can be interfaced with a computer to establish human machine interface. Gesture recognition is suffering from the recognition accuracy. In my proposed approach I solve the recognition accuracy and control the home device which is more useful to people for better quality of life. This system will allow the user to control it in a way that reduces the gap between the physical world and the digital world with an output more intuitive.

REFERENCES

1. Chhowa, M. A. Rahman, A. K. Paul, and R. Ahmmed, “A Narrative Analysis on Deep Learning in IoT based Medical Big Data Analysis with Future Perspectives,” in 2019 International Conference on International Journal of Engineering Trends and Technology Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1–6.
2. Somnath Singh, Debajyoti Saha, Pragya Khaware, Suman Das, Dayanidhi Raj, Subhabrata Das, Chandra Sekhar Nandi ``Home Automation and Internet of Things" International Advanced Research Journal in Science, Engineering and Technology. Vol. 3, Issue 6, June 2016.
3. Ruchi Manish Gurav, Premanand K. Kadbe ``Real time Finger Tracking and Contour Detection for Gesture Recognition using OpenCV" 2015 International Conference on Industrial Instrumentation and Control (ICICI) College of Engineering Pune, India. May 28-30, 2015
4. RenqiangXie and Juncheng Cao “Accelerometer-Based Hand Gesture Recognition by Neural Network and Similarity Matching” IEEE sensors journal, Vol. 16, No. 11, June 1, 2016.

5. Angelo Cenedese, Gian Antonio Susto, Giuseppe Belgioioso, Giuseppe Ilario Cirillo, and Francesco Fraccaroli ‘’Home Automation Oriented Gesture Classification From Inertial Measurements’’ IEEE transactions on automation science and engineering, Vol. 12, No. 4, October 2015.
6. Alexander Nelson, Gurashish Singh, Ryan Robucci, Chintan Patel, and Nilanjan Banerjee “Adaptive and Personalized Gesture Recognition Using Textile Capacitive Sensor Arrays” IEEE transactions on multi-scale computing systems, Vol. 1, No. 2, April-June 2015.
7. Priyanka D. Hatwar, Neha AWhile, Isha M. Padiya, “Hand Gesture Recognition using ANFIS”,Engineering Research (IJETER), Volume 5, Issue 3,March 2017
8. Rafiqul Zaman Khan, Noor Adnan Ibraheem, “Hand Gesture Recognition”, International Journal of Artificial Intelligence & Applications (IJAIA), Volume 3, Issue 4, July 2012
9. Rishad E K, Vyshakh C B, Shameer U Shahul, “Gesture Controlled Speaking Assistance for Dump and Deaf”, National Conference on Discrete Mathematics & Computing (NCDMC), 2017.
10. Jagdish L. Raheja, A. Singhal, Sadab, Ankit Chaudhary, “Android based Portable Hand Sign Recognition System”, International Journal of Science Engineering and Technology Research (IJSETR), Volume 3, Issue 5, May 2014.
11. Chandan et.al., 2018. “Real Time Object Detection and Tracking Using Deep Learning and Open CV”, International Conference on Inventive Research in Computing Applications (ICIRCA).
12. Raghunandan et. al., 2018. "Object Detection Algorithms for Video Surveillance Applications," International Conference on Communication and Signal Processing (ICCSP), 2018.