Human Activity Diagnosis System Based on the Internet of Things

Mohammed Khammas Jabar and Ali Kadhum M. Al-Qurabat

Department of Computer Science, College of Science for Women, University of Babylon, Babylon, Iraq.

E-mail: alik.m.alqurabat@uobabylon.edu.iq

Abstract. The Cognitive Internet of Things (CIoT) is the next step in enhancing the accuracy and reliability of the Internet of Things (IoT) technology used for Cognitive Computing, which plays a main role in health and disease diagnosis. The study was suggested a diagnosis method to sound sensitivity by developing a framework with IoT and cloud based on a facial expression recognition system. It was achieved through the creation of a cognitive IoT hardware-based environment and the elements of the programs that are implemented to test the behaviour of people suffering from sound sensitivity, this operation is done by using a camera and image processing and it used Convolutional Neural Network (CNN) as a facial recognition software to track human facial emotions through live video. The sentiment values were analysed that they were collected and stored in a cloud using Transmission Control Protocol (TCP) protocol. These emotions were categorized as abnormal or normal. Normal states represented by happy or natural feelings that give the impression that the environment is suitable for people's senses and that they do not suffer from discomfort towards this environment, therefore; the system operates on a mechanism to increase the volume in this environment by using an Arduino microcontroller. Although the device operates automatically in the event of hypersensitivity detection, it reduces this severity. The obtained results showed the efficiency of the proposed system in recognizing facial emotions with 80% accuracy.

1. Introduction
The Significant advancements in telecommunications networks and the advent of the Internet of Things (IoT) paradigm have resulted in remarkable improvements in the everyday usage of computer resources and applications. The IoT is an evolving model of communication in which computers, cameras and sensors act as objects or 'things' capable of sensing their environment, communicating with each other, and sharing knowledge over the Internet [1, 2]. IoT key purpose is to allow us to define, signify, navigate and monitor items individually at any time and everywhere by the use of the Internet. A large number of intelligent and autonomous applications and resources will result from interconnected system networks, bringing substantial personal, technical and economic benefits [3, 4].

More recently, the Internet of Things paradigm has been used to build information ecosystems such as smart communities and diagnostic networks for human operation, with various functional fields and similar facilities in the medical industry, for example. The aim of creating such smart environments is to make human life more prosperous and convenient by resolving the problems of living conditions, energy use and industrial needs [5]. This purpose is clear, and is reflected in the significant increase in
applications for open and IoT-based services across various networks. IoT technology can be used in order to create a human behaviour diagnostic system. In order to provide the medical personnel with information, good medical diagnosis and control, proper care and comfort. The use of IoT technology puts together doctors and patients in new health care applications for real-time tracking and hospital management. One of the core technologies of IoT improvements in the healthcare monitoring system is the emotions recognition system based on the Internet of Things [6].

In this paper, a Human Activity Diagnosis System is proposed. The system consists of three essential entities: first, physical elements, smart sensors and actuators; secondly, a communication network (wired/wireless) typically interconnecting the physical compounds; third, intelligent processing of information, for the management and control of the human activity diagnostic system, using the Artificial Intelligence algorithms (e.g. CNN).

The contributions made by this paper are as follows:

- Cognitive IoT for smart diagnosis systems: this paper provides prototyping of smart diagnosis systems (hardware and software) for measuring sound sensitivity.
- Designing a face and emotion recognition system: the paper using Convolution neural network (CNN) as a software to detect faces and emotions signs.
- TCP protocol-based sound management system: the paper includes a control system that uses Arduino to publish and subscribe to cloud server features using the TCP protocol.

2. Related Works

The Human Activity Diagnosis System Based on the Internet of Things has become a highly developed research in the field of healthy intelligence, various approaches were proposed for designing such systems.

Söhsten and Murilo, in 2013, proposed performance evaluations on a neighboring Emgu CV multi-facial detection system and versions communicating with Windows Azure, calculating in real time the amount of facial information processed and the classification element. The greater the number of faces processed is better in terms of reliability [5].

Diego Castro et al., in 2017, presented a new method for Human Activity Recognition (HAR) based on the Internet of Things (IoT) by remotely tracking vital signs. To assess the activities conducted within four pre-established categories (lie, sit, walk, and jog), they use machine learning algorithms. Meanwhile, using a remote monitoring component with remote visualization and programmable alarms, it is able to provide input during and after the operation is completed [6].

Martin Magdin, et al., in 2019, presented the study of reactions recorded by the face analysis method. The experiment was performed on a group of 50 students from universities. 100 random images were presented to each student and the student's response to the image was registered. The reactions reported were subsequently compared to the predicted image response. Several imperfections in the face recognition method have been illustrated by the findings of the experiment. The machine has trouble classifying gestures and is unable to identify and understand inner emotions that an individual may experience when the picture is seen. Only feelings conveyed externally on a face through hormonal changes in some areas of the face can be identified through facial recognition systems [7].

Stefan Oniga and József Sütő in 2012, discussed studies using IoT technology for autonomous everyday life support for aged people or people with disabilities. The aim is to build a system that makes it possible to remain in a familiar atmosphere for as long as possible. The larger spread of assistive devices and the Internet of Things (IoT) would make this possible. In order to build a dynamic assistive device with functional capabilities and learning behavior, the authors plan to bring
together the current achievements in the area of the Internet of Things and Assistive Technology. In order to be eventually evaluated for a medical diagnosis, they used IoT technology to track the status of a patient in real-time or to get critical details [8].

Fadhil and Mandeel, in 2018, a remote tracking system for children with autism has been developed to help experts track the success of their patients and interpret the analysis of data. This method monitors the cases' reactions to different emotions such as sorrow, enthusiasm, nervousness, appetite, trepidation and relaxation, and then logs this detail [9].

3. Proposed System
In this paper, a diagnosis method to sound sensitivity by developing a framework using IoT and cloud computing for a facial expression recognition system is suggested. In the following subsections the face recognition and emotion detection algorithms are discussed in detail.

3.1. System Parameters
The proposed system depends on many parameters:

- **Emotion**: It is a case of mindfulness link with the nervous system. The emotions can be categorized as normal (i.e. happy) and abnormal (i.e. sad). Emotions parameter is represented by $e_i$, where ($i = 1,2$).
- **Sound Intensity**: It is the power that sound waves hold per unit area in a direction vertical to that area, International System of Units defined intensity of sound as watts per square meter ($\text{watts/m}^2$). $s_i$ represents the parameter of sound intensity where ($i = 1,2, ..., m$), and $s_0$ and $s_m$ denote the initial and maximal sound intensity respectively.
- **Levels of Sensitivity**: Experience many levels to find a human sensitivity to sound. The sensitivity level is placed through the symbol $n_i$ where ($i = 1,2, ..., m$), $n_0$ is the first sensitivity level and $n_m$ is the highest sensitivity level.

3.2. System Model
The mechanism of the system can be specifically created based on the diagnostic process performed at multiple sensitivity levels. Figure 1 shows the diagram of the diagnostic system. The sentiment detection system can interpret the captured data that represent the video frames and produce the emotions to describe the person's reaction to the sound. Sentiment collected from the previous stage, which was processed in the patient's computer to determine if the person is in a normal state or abnormal and choose the appropriate volume level for each condition. The results are then sent to the cloud that uses Internet protocols. Objects Encapsulated in a TCP Datagram. The results uploaded in the cloud can be accessed by the other party for checking if the emotions values are normal or not.
3.3. System Work

The value of sound intensity $s_i$, which decreases or increases depending on the equations and conditions described below, is calculated by the control system of the diagnostic system for acoustic sensitivities as follows:

1. The system of diagnosis determines the range of the levels of sensitivity $n_i$.
2. The maximal intensity of sound $s_m$ is set by the diagnostic system and the test begins with $s_0$, the initial intensity of sound.
3. The intensity of sound difference value $\Delta s$ which increases/decreases at each stage can be determined from the following equation:
   \[
   \Delta s = \frac{s_m - s_0}{n_m}
   \] ..... (1)
4. Diagnosis system regulates the $e_i$ at each level.
5. $e_i$ normality is an indicator to raise the intensity of sound and vice versa.
   1. The intensity of sound is determined from the following equation for normal emotions:
      \[
      s_i = s_{i-1} + \Delta s
      \] ..... (2)
   2. The intensity of sound is determined from the following equation for abnormal emotions:
      \[
      s_i = s_{i-1} - \Delta s
      \] .......... (3)
6. When it notices abnormal feelings or when $n_i$ equals $n_m$, the system stops.

The flowchart for the diagnostic sound sensitivity method displayed in Figure 2.

3.4. Facial Emotion Detection

In human life, emotions play a significant role. Human face represents at different times or moments how he feels or in what mood he is. Emotional factors have an enormous effect on social intelligence,
such as understanding interaction, making decisions, and helping to understand human behavioural aspects. The processing of facial emotion recognition passes through three stages [10]:

1. Facial Detection: Face position in any photo can be identified, offline video or live video.
2. Facial Recognition: Comparing various faces to define which faces belonging to the same individual. This is achieved by comparing the embedding vectors of the face.
3. Emotion Detection: Classifying the emotion as being on the forehead happy, angry, sad, neutral, surprise, disgust or fear as shown in Figure 3.

In this paper, facial emotions recognition will be used to support the faces and emotions detection of autism people and patients with sensitivity to sound.

![Sound Sensitivity Diagnostic Flowchart](image)

**Figure 2.** Sound Sensitivity Diagnostic Flowchart.
The mechanism used for emotion detection in this project is based on Deep learning. Convolutional Neural Networks (CNN) is one of the deep learning algorithms specialized in the images that have the best high resolution. The CNN, as well as other hidden layers, it consists of an input and output node. These layers are usually categorized into four forms: convolution, max pooling, flattening and full connection (short for fully-connected) [11] as shown in Figure 4.

Figure 4. CNN sequence to classify handwritten digits [11].

4. Design and Implementation
The software/hardware components of the system and the detailed description of its implementation is given in this section. The components that need to build the proposed system are Arduino Uno, Arduino WiFi Shield, Servomotor, Micro USB cable, breadboard, jumper wires, Smartphone, two laptops and sound generator. Figure 5 illustrates the system components.
4.1. Live Video Processing System

A smartphone (Redmi note 7) was utilized to recognize the patient's face during live video capturing mode for the purpose of detecting and examining his emotions. DroidCam software was used for both interfaces of emotion detection system (smartphone and laptop) that are connected via Wi-Fi using Wi-Fi IP and DroidCam port 4747. DroidCam software was connected to reprogrammed software written by python 2.7 language using PyCharm Community Edition 2020.2 x64 for face and emotion detection in laptop.

4.2. Emotions Detection

After taking the picture through the smartphone the image enters the analysis algorithm (as shown in Figure 6) which is as we mentioned in the third section convolutional neural network in which the patient's condition is (happy, angry, sad, neutral, surprise, disgust or fear). Currently, only these cases are taken into account due to limited time.

This algorithm relies on training the network by entering more than 1792 emotional images for training and nearly 20 emotional images to test the neural network, the dimensions of the input image are 48 * 48, it is implemented using Kiras Library, which is a Python compiler used in TensorFlow library including machine learning tools and algorithms. Inserting images into the convolution layer, whose size is the size of the input image, which creates very important features and by applying the filter that determines the edges of the image where the
images wanted for training were entered into the convolution layer four times to find the desired, then enter the Max pooling layer which is 24 * 24 in the first layer and continues to decrease each time, this layer is repeated four times as well. Then enter the Max pooling layer. The size is 24 * 24 in the first layer then continues to decrease each time and repeat this layer four times also, after that flattening layer which flattens the matrices emerging from the previous two layers. We have repeated this layer three times then the image or arrays enter the neural network for training. 

The emotions results are extracted from the input image as follows:

1. If the percentage of normal features inside the output matrix increases, the output value is normal
2. If a decrease in the percentage of normal features and an increase in abnormal features is observed within the trained network, the output will be abnormal. As shown in Figure 7.

![Confusion matrix](image)

**Figure 7. Output matrix of the CNN.**

### 4.3. Cloud Implementation

After completing the processing on the patient's computer, the edge node (i.e. PC) sends the complete results to the cloud, which can be accessed from anywhere. The results and patient photos are displayed on the cloud as shown in Figure 8.

![Uploaded results on Cloud](image)

**Figure 8. Uploaded results on Cloud.**
4.4. Devices of Sound Control
The sound control system contains the sound control knob (micro kigdom Q280) Arduino, servo engine, board and wires integrating to control the rotation of the sound volume control knob. The sound frequency of the sound generator should be between 0 dB and 70 dB. The control knob of sound volume is connected to the servomotor in order to increase or decrease the volume of sound through rotating it clockwise or counter-clockwise. To connect the Arduino Uno (low-level language) with Python (high-level language) we need to use the Firmata firmware from Arduino Uno.

4.5. System Implementation
After the completion of the configuration and installation of the software/hardware components of the system will look as shown in Figure 9. The practical implementation of the system will be as follows:

![Figure 9. Circuit Connections for the proposed system.](image)

1. The patient sits in front of the mobile camera that connected to the first computer through the Wi-Fi network.
2. Providing internet service to ensure that the results are sent to the cloud, which in turn reaches the doctor in any place.
3. When clicking from the controller's computer, an instruction is sent to the patient's computer on the second party via the TCP protocol to take a picture of the patient.
4. The facial recognition algorithm determines the patient's condition.
5. And then the computer sends the results to the cloud, at the same time the computer sends an instruction to the motor to move according to the results.

5. Results
Each case was examined to determine his/her sensitivity to sound using Nostalgia music for Yanni (it is adjusted in terms of duration to be appropriate for examination) for sensitivity ranging from 1 to 5 levels system works to increase this intensity within certain periods ranging from 106 sec to 636 sec depending on the patient sensitivity and examination situation. At each sensitivity level the case was examined to detect his/her emotions to find out if he is in a normal situation or suffers from sensitivity toward this sound intensity. This intensity was automatically increased when the case was confirmed to be normal and stable and vice versa if he turns out to be disturbed.

In figure 10 four cases are examined. Case 1 in Figure 10-A displays a normal emotional reaction for all levels of the heard sound. While the other cases in Figure 10 (B, C and D) illustrate a
hypersensitivity to a specific level of sound intensity. After the patient suffers from hypersensitivity, the sound intensity is reduced automatically to the previous level.

![Graphs showing sound sensitivity levels over time for different cases](image)

**Figure 10.** The sound sensitivity examination.

6. **Conclusion**

This paper introduced a smart diagnostic device to detect sound sensitivity to the effects of the surrounding atmosphere using cloud computing and cognitive IoT. The system prototype can be adopted as an effective system in medical centres dealing with the diagnosis of hypersensitivity toward environmental parameters. Several experiments have been conducted using different algorithms as fuzzy algorithm. These experiments showed that a convolutional neural network is 80% more accurate than as fuzzy algorithm but requires more time in training. Also, this system applies the idea of social spacing, especially in the context of the spread of the new Coronavirus. The proposed system is effective in diagnosing and responding to autistic patients with sound sensitivity.

As a suggestion for future works, hands and head motion detection can be added to the proposed system to support emotion detection. Also, with the training of many networks and the introduction of
all aspects of body language, can be converted into a polygraph system in criminal cases, with many factors such as light and sound and manipulation of guilty room temperature and other factors.

7. References

[1] Idrees A.K., Al-Qurabat A.K.M., Energy-Efficient Data Transmission and Aggregation Protocol in Periodic Sensor Networks Based Fog Computing. *J Netw Syst Manag*. 2020, 29(1):1-24.

[2] Al-Qurabat A.K.M., Idrees A.K., Data gathering and aggregation with selective transmission technique to optimize the lifetime of Internet of Things networks. *Int J Commun Syst*. 2020, 33(11):e4408.

[3] Al-Qurabat A.K.M, Idrees A.K., Abou Jaoude C. Dictionary-Based DPCM Method for Compressing IoT Big Data. In 2020 International Wireless Communications and Mobile Computing (IWCMC). 2020 (pp. 1290-1295). IEEE.

[4] Al-Qurabat A.K.M, Abou Jaoude C, Idrees A.K., Two tier data reduction technique for reducing data transmission in IoT sensors. In *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. 2019 (pp. 168-173). IEEE.

[5] von Söhsten D, Murilo S. Multiple face recognition in real-time using cloud computing, Emgu CV and Windows Azure. In *2013 13th International Conference on Intelligent Systems Design and Applications*. 2013 (pp. 137-140). IEEE.

[6] Castro D, Coral W, Rodriguez C, Cabra J, Colorado J. Wearable-based human activity recognition using an iot approach. *Journal of Sensor and Actuator Networks*. 2017 Dec;6(4):28.

[7] Magdin M, Benko Ľ, Koprda Š. A case study of facial emotion classification using affdex. *Sensors*. 2019;19(9):2140.

[8] Oniga S, Sütő J. Human activity recognition using neural networks. In *Proceedings of the 2014 15th International Carpathian Control Conference (ICCC)*. 2014 (pp. 403-406). IEEE.

[9] Fadhil TZ, Mandeel AR. Live Monitoring System for Recognizing Varied Emotions of Autistic Children. In *2018 International Conference on Advanced Science and Engineering (ICOASE)*. 2018 (pp. 151-155). IEEE.

[10] Deshmukh R, Scholar ME. A Comprehensive Survey on Techniques for Facial Emotion Recognition. *International Journal of Computer Science and Information Security*. 2017;15(3).

[11] Simonyan, K., and Zisserman, A. (2014) “Very deep convolutional networks for large-scale image recognition”. arXiv preprint arXiv:1409.1556.