Smart Hat: Design and Implementation of a Wearable Learning Device for Kids Using AI and IoTs Techniques

I-Hsiung Chang[1], Huan-Chao Keh[1], Bhargavi Dande[1], Diptendu Sinha Roy[2]
[1] Department of Computer Science and Information Engineering, Tamkang University, Taiwan
[2] Department of Computer Science and Engineering, National Institute of Technology Meghalaya, India
elite5931.tw@gmail.com, hckeh@mail.tku.edu.tw, 606785037@s06.tku.edu.tw, diptendu.sr@nitm.ac.in

Abstract

Internet of Things (IoT) and Artificial Intelligence (AI) have received much attention in recent years. Embedded with sensors and connected to the Internet, the IoT device can collect massive data and interact with a human. The data collected by IoT can be further analyzed by applying AI mechanisms for exploring the information behind the data and then have impacts on the interactions between human and things. This paper aims to design and implement a Smart Hat, which is a wearable device and majorly applies the IoT and AI technologies, aiming to help a kid for exploring knowledge in a manner of easy, active, and aggressive. The designed Smart Hat can identify objects in the outside environment and give output as an audio format, which adopts the IoT and AI technologies. The learning Smart Hat intends to help kids aid them in the primary learning task of identifying objects without the supervision of the third party (parents, teachers, others etc.,) in real life. This Smart Hat device provides a sophisticated technology to kids for easy, active, and aggressive learning in daily life. Performance studies show that the obtained results are promising and very satisfactory.

Keywords: Artificial Intelligence (AI), Convolutional Neural Networks (CNN), Internet of Things (IoTs), OpenCV

1 Introduction

Nowadays, computer technologies play a major role in human life. The use of technologies to improve the quality of human life is becoming a common aspect of modern society. Recently, IoT [1-4] has received much attention and considered as the appealing technology that allows people and things to connect anytime and anywhere. It offers a platform for sensors and smart devices to be connected continuously within a smart environment in order to provide advanced and knowledgeable services for human beings. The IoT is a network of physical objects embedded with sensors, electronics, software and network connectivity that enables these objects to collect and exchange data. This means that a user can remotely monitor or even control the states of home appliances. In general, most smart devices are embedded with sensors to detect the environmental status, machine- operating status or to collect behavior data and can respond to users or machines since it connects to the Internet and can be interactive with humans or machines.

AI is an emerging technology in the world, which makes all other technologies to work autonomously and helps humans to get more perfection and accuracy in daily life. Research in AI has been primarily dominated by impressive advances in Machine Learning, with a strong emphasis called Deep Learning framework. It has allowed considerable achievements such as human- level performance in a highly complex game of Alpha Go [5], image and speech recognition [6] and description [7], and even in Video games. Nowadays, Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in object detection, visual object recognition, speech recognition etc.

A small system needs training with huge data, so Integrated IoT to collect data is used as an input to the training. AI is the best solution to manage huge data flows and storage in the IoT network. IoT nowadays becoming more and more popular with the inventions of high-speed Internet networks and many advanced sensors that can be integrated into a micro controller. Since everything will be connected to the Internet, huge data, including sensors data, user data as well as the rich data sent from home appliances and a variety of wearable devices will be collected automatically. With the increase in the number of smart devices and more and more sensors, a learning model can be constructed by machine learning algorithms to make good use of the huge data and used for predictions such as Image Classification, Facial Recognition, Speech Recognition, and Object Detection etc. For these reasons, we are proposing a novel prototype implementation by integrating IoT and AI technologies.

Traditional learning can be conducted in various

*Corresponding Author: Huan-Chao Keh; E-mail: hckeh@mail.tku.edu.tw
DOI: 10.3966/160792642020032102026

ways, including textbooks, children's intelligent robot, teachers and parents, smart devices etc. All these types of traditional learning to have some disadvantages, like single source knowledge, limited learning space and time, lack of interaction between traditional textbooks and kids as well as the health issue raised due to using electronic devices for a long time. Some other informal learning adopts mobile learning technique to shift the learning environment from the traditional classroom to museums or outside environment [8]. However, the learning contents are predefined and the learning models are usually fixed. Furthermore, mobile learning is mainly designed for children, which is not suitable for kids. This paper is motivated by all these types of traditional learning. Our designed Smart Hat is portable and therefore ideal to simplify the interaction with the smart environment. The main effort is to provide these kids with a Smart Hat (see Figure 1), in order to explore the outside learning within the smart environment by using just their Smart Hat. Our goal is to explore the advantages of IoT and AI techniques to provide the kid's learning with good features including active, interactive and mobile. However, this goal cannot be achieved several years ago because AI is not highly developed and it is not so intelligent to timely recognize the objects and returns an accurate result. But nowadays the accurate object recognition can be achieved by using deep learning architectures. In addition to AI, the new IoT technique also makes it possible to achieve our goal. This occurs because of the supports of new IoT techniques including the involution of high-speed Internet networks and energy saving and smaller size sensors and related hardware. Therefore, good learning features can be achieved in the right way by applying IoT and AI techniques. The contributions of this paper are itemized as follows.

Figure 1. The four-year-old kid using our Smart Hat

Integration of IoT and AI. Generally, IoT collects the data and interacts with people. AI helps the IoT to think very smartly and has a decision to complete the work very easily. The IoT with AI helps the kids easily carry the Smart Hat and also easily identify the objects. In our proposed work, IoT devices like (Arduino, Raspberry Pi) acts as input and output devices but the decision of identifying the object is done by AI which acts as a brain or CPU. Hence, in the proposed work, we integrated IoT with AI.

High interactions between kid, learning objects and learning contents. We aim to make every kid not just seeing the objects but also knowing or learning about them without extra efforts. This could be discussed in several aspects. (a)Independence of learning: The kids learn to recognize an object by interacting with the Smart Hat, without the need of guidance from parents or teachers. (b)The mobility of learning: Our proposed Smart Hat belongs to portable learning technology which kid can wear and move easily in the outdoor environment. Without any interference, kids can learn whatever they want since they have the knowledge database in their Smart Hat.

Active, mobile and exploring learning. The kids have many disadvantages with traditional learning like single source knowledge, lack of interaction and limited space and time. Therefore, this inconvenience gives the motivation to develop a Smart Hat based on the IoT and AI technologies. This research is to use the Smart Hat as wearable learning device such that kids may actively learn about the objects in the outside environment and to arouse their interest to learn.

Being the bridge of AI and IoT to learning technologies. AI makes the machine learn from its experiences and manage with new data. Similarly, IoT is the combination of many devices or sensors over the Internet. The proposed system adopted IoT devices and sensors to generate a picture, apply the Deep learning technology to recognize the picture and finally get the information from the Google Cloud Server.

The remainder of this work is organized as follows. Section 2 reviews the related works. Section 3 illustrates an overview of the proposed system. Section 4 provides the design and implementation of the proposed system. Section 5 presented the programming design of the system. The experimental study is investigated in Section 6. Finally, Section 7 offers the conclusion.

2 Related Works

Kids learning is always the most important and has received much attention from their parents and teachers. There have been a lot of previous studies proposed learning tools, models and policies for improving the learning performance. The existing learning aids for preschool kids are categorized into traditional learning, mobile learning as well as the intelligent robot. The traditional learning, including parents teaching kids at home and teachers teaching kids at preschools, have considered the interaction between the kid and the learning tool, but they do not consider the interaction between the kid, learning contents as well as the learning environment. Mobile learning, which falls in the second category, uses electronic devices to teach kids in preschool. Kids are forced only to learn the contents that are prefixed and there is no dynamic change according to the environment and scenario. The third learning type is the intelligent robot, which considers the interaction of the kid. However, they have not considered the outside learning. Our Smart Hat learning, fall in the category of mobile learning, integrates the technologies IoT and AI. With the help of our Smart Hat, kids can self-explore knowledge and have freedom and independence of learning. The following presents the related works of traditional learning, mobile learning, and intelligent robot.

Study [9] shows that parents were an excellent source of knowledge about their kid's skills, learning behavior and intelligence. However, knowledge from parents is the single source and there is a lack of outdoor interaction learning. The parents of the competitive society today are unable to dedicate time for kids learning needs and are not up-to-date with the change of technology. Now a day's parents leave their kids in preschools for their early childhood education before they begin compulsory education at primary school. The teachers in the preschools teach kids through creative activities such as games, paintings, videos, images and other tools [10]. However, the contents prepared by preschool teachers are fixed and there is no dynamic change as the environment and scenario change. This type of learning lacks interaction and autonomy to learn from exploration.

A. S. Drigas et al [11] presented a mobile learning for preschool education. This paper shows the complete overview of mobile applications used for preschoolers. They concluded that Handheld mobile technologies are emerging in classrooms and enable the kids to use it anywhere and anytime. Research in mobile technology is considered as supplement teaching tool for children, as they seem to support children while increasing the motivation to learn. Another study [12], proposed iPads as a literacy-teaching tool in early childhood. They conclude that kids were very eager to use the device and ensure that they stay on task without distracting. This is because of the iPad's quick and easy access to information for the kid. The portability allows it to be easily moved around the school and classroom. Studies [11-12] aimed only to grab the attention of kids, but they did not consider the disadvantages caused by electronic devices. Learning from digital devices is an interactive type of learning but it has many disadvantages, such as myopia and eye fatigue, because of blue light emission from digital devices.

Study [13] proposed a humanoid social robot as teaching aid in preschool classes. It is a learning and playful robotic companion for preschool kids. The robots are programmed to deliver six lessons over a span of 3 months. Robots have built-in sensors and accessories such as cameras, microphones, LED lights, speakers etc. It supports several features, including multi-modal interaction, reaction to kid's gestures and movements, facial expressions and communicates using voice, lights, and inside-out projections on its body. However, they have limited space and cannot support kid's self-exploration.

All aforementioned types of learning emphasized the interaction of kids or aimed to cope with the kid's attention. However, most of them do not consider outdoor learning. Table 1 summarizes the comparison of our related work. This paper proposes a Smart Hat, which helps the kids between the age groups 2-8 to enhance their learning experience by providing them with the ability to learn themselves. This can support kid's self- exploration and knowledge building. This research focused on investigating how to support self-learning and outdoor learning of modern-day preschool kids while keeping their attention continuously.

Table 1. Comparison of our related work

	Active learning	Interactive learning	Outdoor learning
[9]	×	O	O
[10]	O	×	×
[11] & [12]	O	×	×
[13]	O	O	×
Ours	O	O	O

3 Smart Hat System Overview

This section introduces the key components of the Smart Hat system and discusses how they interact with each other to identify the objects in the outside environment. The proposed system consists of four parts: (i)IoT Controller (ii)IoT Hardware Platform, (iii)AI Software Platform and (iv)Google Cloud Platform. Figure 2 shows the block diagram of the proposed system.

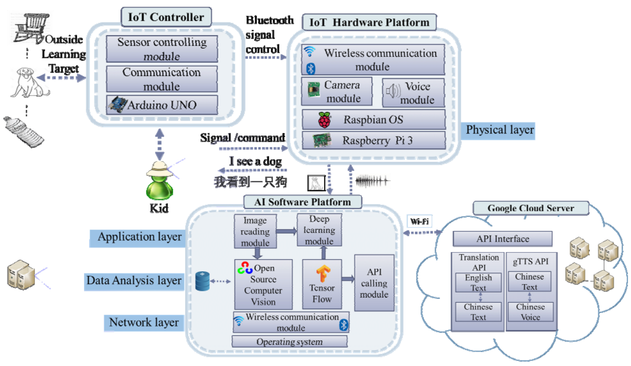


Figure 2. The proposed architecture for Smart Hat

3.1 IoT Controller

The IoT controller consists of Arduino UNO, a wireless communication module and a sensor-controlling module. A handheld button is integrated with the Smart Hat, which makes the kids to easily interact with the object. When the kid initiates the learning process by pressing the button in the hand, the Bluetooth chip which has been embedded in the handheld Arduino will send a command to the Raspberry Pi to capture the image.

3.2 IoT Hardware Platform of Learning Hat

The IoT hardware platform mainly embedded in the Smart Hat as a wearable device. We used Raspberry Pi 3 to implement the IoT hardware platform of learning hat. When the kid presses the button, the command signal will be sent to the Raspberry Pi using Bluetooth wireless communication. Then the camera on the Raspberry Pi captures the image and sends it to the Deep learning server for further classification. The result is sent to the Raspberry Pi from the Deep learning server using Wi-Fi wireless communication. The speaker module on Raspberry Pi then delivers the voice of the received content.

3.3 AI Software Platform of Learning Hat

The AI software platform is developed using Python language. The AI software platform consists of the Image reading module, Deep learning module, API calling module, and the wireless communication module.

When the captured image is sent to the AI software platform from the Raspberry Pi, the Image reading module uses OpenCV (Open Source Computer Vision) to read and denoise the received image before sending to the Deep learning module. Then the Deep learning module uses Deep learning framework called TensorFlow to perform the image classification. Specifically, the object recognition leverages the Inception-v3 pre-trained model, which was built by Google. The TensorFlow takes the images as input and gives the output as text. Once the classification is done, the API calling module uses a wireless transmission module to call the Google Cloud Server to transform the information to voice.

3.4 Google Cloud Server

Google Cloud Server consists of Google Translate API and Google Text to Speech (gTTS) API. Google Translate API module is used to convert the English Text to Chinese Text and gTTS API module is used to convert the Chinese Text to the voice. When the API calling module in the software platform calls the Google Cloud Server, Google Translate API takes the input as English Text and gives output as Chinese Text. Then the gTTS API takes the input of the Chinese Text and converts it to voice and then sends to Deep learning server to send it to the Raspberry Pi. The Raspberry Pi then sends actual results via Speaker. We used gTTS API because it provides a more humanoid aspect to the speech output. The gTTS API returns the string result to the Raspberry Pi via WiFi wireless communication. This module is also responsible for the system coordination. Then the kids can identify the name of the object only by listening to the audio file through the speakers.

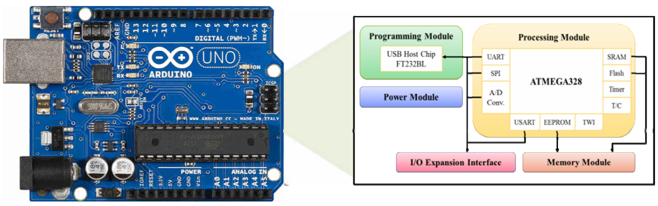
4 Design and Implementation

This section initially introduces the features of each hardware component. Then the design and implementation of the hardware and software platforms of Smart Hat are presented.

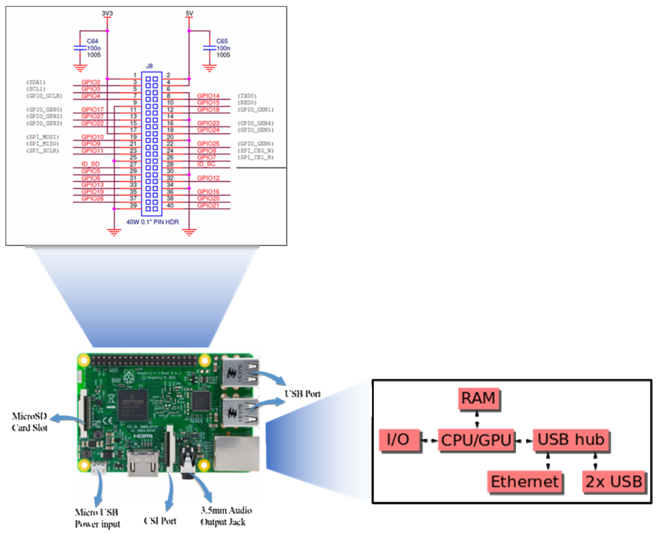
4.1 Features of Hardware Components

4.1.1 Main Key Hardware Platform

In order to drive the camera, the kid holds a controller where a pressure sensor is embedded in it. As shown in Figure 3(a), Arduino is an I/O interface control board developed based on open source code and has a development environment similar to Java and C. The Arduino control board provides fourteen digital I/O and six Analog I/O.



(a) Arduino hardware component



(b) Raspberry Pi 3 hardware component

Figure 3. Key components of smart hat system

Figure 3(a) depicts the functions of Arduino hardware and some connected hardware components for implementing the handheld device of our learning hat system. The following first describes the functionalities of Arduino hardware. Then the related hardware components, including a Bluetooth chip and button sensor will be illustrated.

A microcontroller ATmega 168 is a whole computer system, including a microcontroller, memory, a clock oscillator, and I/O on a single integrated circuit chip. The ATmega168/V supports a real Read-While-Write Self-Programming function. UART is the communication interface between the Arduino board and another device which is the USB host chip. The power module is the external power supply which takes an input voltage from 7-12V. The SRAM and EEPROM are the memory modules which have 2 KB and 1KB memory. The I/O expansion interface has 14 digital pins. The number of Analog input pins is 6 while the clock speed is 16 MHz.

The button sensor is used to trigger the camera connected to Raspberry Pi in the Hat device. The triggering signal is sent through the Bluetooth chip which is connected to Arduino UNO. The button is connected to the Key pin of the Bluetooth device BT (HC-05) and the ground pin (12th Digital pin) of Arduino UNO. Then the Bluetooth chip BT (HC-05) is connected to the 9 and 10 digital pins of Arduino Uno and button sensor, respectively. Therefore, this connective module of a button, Arduino and Bluetooth help to complete the function of pressing a button to send a signal to the HAT device for capturing an image.

The Raspberry Pi is a hardware component that includes a circuit with simple I/O capabilities and Linux software. It can read a large number of switches and sensor signals and accordingly control the lights, cameras, and its various physical devices. Raspberry Pi can also be used as a computer to develop many embedded devices, perform a variety of software and applications on a Linux PC and communicates with other devices in a wireless manner.

The various styles of the Raspberry Pi board have brought developers many advantages, whether professional or amateur. The Raspberry Pi 3 is equipped with 1GB RAM and 400MHz Video Core IV GPUs. The size of the entire motherboard is the same as that of the Raspberry Pi 2. It has built-in Bluetooth 4.0 and 802.11n WiFi and supports one to four USB ports. The architectural improvements and processor upgrades of Raspberry Pi 3 have increased its clock frequency by 33%. Compared to the 32-bit Raspberry Pi 2, the performance of Raspberry Pi 3 has improved by 50 to 60%. Raspberry Pi 3 has a processor speed ranging from 700 MHz to 1.4 GHz and onboard RAM memory ranging from 256 MB to 1 GB. Secure Digital (SD) cards are used to store SDHC or Micro SDHC-sized operating systems and programs. For video output, HDMI and composite video are supported with a standard 3.5 mm jack for audio output. Lower-level outputs are provided by many GPIO pins that support common protocols such as I2C.

As shown in Figure 3(b), the key to the success of Raspberry Pi 3 is the standard hardware interface, which supports the 40-pin interface, 28 GPIO pins, I2C, SPI and UART connections. In addition to the GPIO pins, the Raspberry Pi’s standard interface also provides 3.3 V, 5 V, and ground. Thanks to this universal pin configuration, Raspberry Pi users can find a variety of add-on boards from third-party vendors and use this standard interface as a basis for their construction.

4.1.2 Wireless Communication Module

The WiFi wireless module, is an independent embedded wireless network module supporting 802.11b/g communication protocol. It provides 10 Digital and 8 Analog output/input ports. Modules can be issued commands via external devices. Because of the MOSI and MISO of the SPI communication interfaces are independent, it thus allows data packet transmission and reception functions to be performed at the same time.

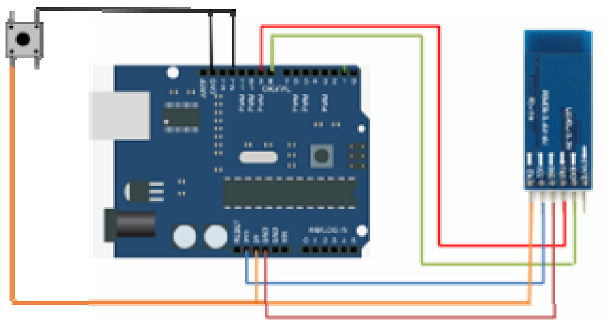
The Bluetooth wireless module, is a wireless standard used to allow fixed and mobile devices to exchange data between short distances to form a personal area network (PAN). It uses short-baud high frequency (UHF) radio waves to communicate via the 2.4 to 2.485 GHz ISM band. The Bluetooth transmission module uses Cambridge Silicon Radio’s BC417143 chip, which supports Bluetooth 2.1+EDR. Specification.

4.2 Design and Implementation of Hardware Platform

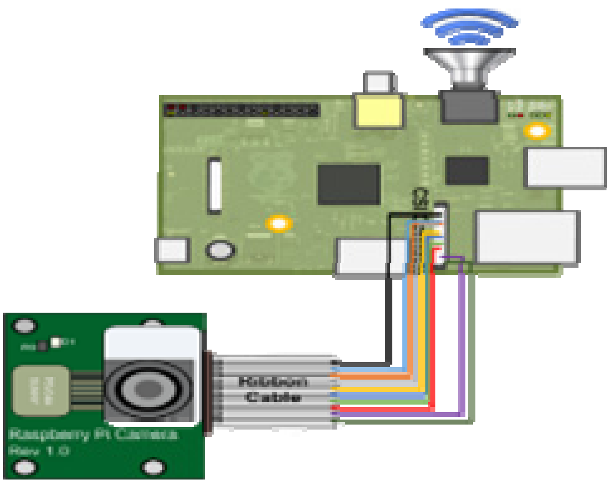
This subsection presents the design and implementation of the Smart Hat hardware platform. We implemented a Smart Hat using Raspberry Pi, which employs functions including photographing, wirelessly transmitting photos, receiving voices as well as playing learning contents. In addition, we also implemented the kid handheld button device, which can trigger the action of the Smart Hat by using Arduino hardware and button sensor.

4.2.1 Handheld Button Device

We connect the button sensor and Bluetooth wireless transmission module to the Arduino hardware, as shown in Figure 4(a). We connect one pin of the button sensor to the Arduino’s 5V pin and another pin to the signal line and GND of the Arduino. When the kid pushes the button, the corresponding pin of Arduino will be set to HIGH. Otherwise, the corresponding pin of Arduino will be set to LOW when the button sensor is not triggered. For the Bluetooth module, it’s RX and TX pins are connected to the TX and RX pins of the Arduino board, respectively. As a result, the handheld button device can communicate with the Smart Hat wirelessly. The VCC pin and GND pin of the Bluetooth module is connected to the corresponding pins of the Arduino. When the Raspberry Pi module is powered on, we set the Bluetooth ID and address of the transmitter and receiver so that the kid can press the button to send a Bluetooth signal to the Smart Hat. Consequently, the handheld button device can trigger the speaker of the Smart Hat to broadcast the learning content.



(a) Bluetooth button wiring diagram



(b) Raspberry Pi camera module

Figure 4. The circuit design of control button and Raspberry Pi camera modules

4.2.2 Hardware Platform of Smart Hat

On the Raspberry Pi board, we embed the Raspberry Pi camera module on the CSI Connector, as shown in Figure 4(b). After that, we changed the permission of the camera in the Raspberry Pi module to control the camera's time and a number of shots. In addition, we also use the program to define Wi-Fi settings of the Raspberry Pi.

As a result, the designed hardware platform of the Smart Hat can connect to the server where the Deep learning is performed. The Smart Hat can use Wi-Fi to transfer photos to Deep learning server and receive the voice results from the server. Finally, the speaker of the 3.5mm jack outputs the voice of the received content.

4.3 Design and Implementation of Software Platform

This section presents the design and implementation of the Smart Hat software platform. We use OpenCV as image processing module because OpenCV library has more than 2500 optimized algorithms and predefined functions which is helpful in Image Processing [14]. In addition, we apply Deep learning framework called TensorFlow as image classification module.

4.3.1 OpenCV

The OpenCV is used as an Image Processing module. We install OpenCV on the Deep learning server. When the Raspberry Pi sends the image to the AI software platform, OpenCV reads the image using cv2.imread() function. The function cv2.IMREAD_ COLOR() is called to load the color images and the function cv.fastNIMeansDenoisingColored() is called to denoise the color image. Finally, the processed image can be fed to TensorFlow to work on classification. In short, when we use OpenCV, we can directly call the OpenCV functions, then add our own written programs. It not only reduces the difficulty of the development process but also shortens the development cycle of the relevant procedures.

4.3.2 Tensorflow

Google released a model called Inception-v3 with TensorFlow.The Inception-v3 [15] with the TensorFlow is a pre-trained Deep convolutional neural network architecture for recognizing objects in images. The node lookup class will process the label_ lookup_path and uid_lookup_path and give back a human-readable string for each classification result. The TensorFlow takes an image as an input, classifies the object in the image and then sends back a text of the object name. This object name is the major content for kid's learning.

4.3.3 Google Cloud Server

We used Google Translate API 2.3.0 and gTTS API 1.1. In terms of implementing the Smart Hat, TensorFlow successfully generates the English name of the object after identifying the object in the image. The Google Translate API can convert the English object name to a Chinese object name and gTTS API converts the Chinese Text to voice as it provides a more humanoid aspect to the speech output. Finally, Google Cloud Server returns the string result to the Deep learning server via WiFi wireless communication. The Deep learning server then sends back the voice file to the Raspberry Pi. Then the Raspberry Pi broadcasts actual results via Speaker. Figure 5 depicts the physical prototype of our proposed Smart Hat implementation.

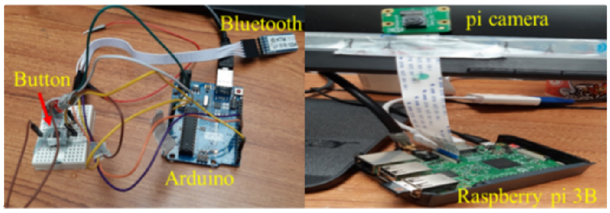


Figure 5. Smart Hat prototype

5 Programming Design for Deep Learning System

The programming design for Deep Learning System is divided into three modules: (i) Image reading module (ii) Image classification module and (iii) Speech module. The dependencies of our project are OpenCV 3.2, TensorFlow 1.0 and gTTS API 1.1, Google Translate API 2.3.0, Numpy 1.12 modules. Programs are completely written in Python language and Jupyter notebook is used for compilation.

Image reading with OpenCV is the first module. In the AI software platform, we installed OpenCV and TensorFlow with the anaconda environment in Deep learning server. We call the cv2.IMREAD_COLOR() to read the incoming image. After reading the image we denoise the image using cv.fastNlMeansDenoising Colored() function. The Image classification with Inception-v3 is the second module. A node lookup class is created for getting the human-readable string from the result. Node lookup class is to load the mapping of integer node ID (target class) to a human-readable string. It will process the label _lookup_path and uid_lookup _path and give back a human-readable string for each classification result. Label_lookup_path is to load the mapping of integer node ID (target_class) to target_class_string and uid_ lookup_path is to load the mapping from target_class_string to the human-readable string. Finally, TensorFlow gives the output as human_readable_string (text).

Our final module is the Speech module. In the Google Cloud Server, we used Google Translate API to convert English Text (umbrella) to Chinese Text (). Then we used gTTS API to convert human_ readable_string () to the voice format using tts= gTTS (text= “” + human_readable_string). Then we save it as an mp3 file using tts. Save (“human_string_filename. mp3”) and then send to Deep learning server. The Google Cloud Server sends back the result to the Raspberry Pi using the WiFi. Finally, the Raspberry Pi sends the actual result () via the speaker.

6 Performance Evaluation

This section compares the performances of our proposed system in terms of accuracy, time delay and efficiency index. The MATLAB simulator is used as the simulation tool.

The subsequent illustrates the parameters considered in the simulation environment. Our proposed system comprises seven stages, noted by T_1 to T_2, for implementing the Smart Hat. In the first stage $T_1$the kid will initialize the process by pressing the handheld button device which is embedded in Arduino. Then the Arduino will send a request packet to Raspberry Pi via Bluetooth radio to notify Raspberry Pi camera. In the second stage T_2 the Raspberry Pi captures the image of the object in front of the camera. In stage T_3 the captured image of the object is sent to the OpenCV server using Wi-Fi wireless communication module. In stage T_4 the classification process is carried out. The image from the OpenCV is classified by TensorFlow in the software platform and the output is the text representing the object. In stage T_5 the language translation process is carried out by translating the English Text to Chinese Text, which uses the Google Translate API, provided by Google Cloud Server. Inthe stage T_6 the conversion from text to voice is done. The Chinese text is transformed to Chinese voice by using gTTS API service, which is offered by Google Cloud Server. Then in the final stage T_7 the audio signal obtained from gTTS is sent to the Raspberry Pi to send the result of the classification to the kid through the speaker integrated in Raspberry Pi. The various time stages are summarized in Table 2.

Table 2. Time stages of proposed system

Time stages	Descriptions
T_1	Bluetooth Transmission from the handheld button device to the Raspberry Pi
T_2	Capturing of the image by Raspberry Pi camera
T_3	Sending the captured image to the AI software platform.
T_4	Classification process by TensorFlow
T_5	Conversion from English Text to Chinese Text using Google Translate API
T_6	Conversion from Chinese Text to voice using gTTS API
T_7	Sending the voice to the Raspberry Pi

One important parameter is the pixel density, which might impact the required processing time for some stages. The considered pixel density in the experiment is controlled ranging from 100 dpi to 900 dpi. The maximal allowable time delay is 1, 3, 5 and 7 seconds. This indicates that the total time, including T_1 to T_7 cannot exceed the maximal allowable time delay. Experiments were conducted on the kids of various age groups, comprises 2 to 5 years old. The above-mentioned parameters are summarized in Table 3.

To further investigate the performance of the proposed mechanisms, three scenarios are considered in the experiments, including HAND, HAT, and SHOULDER scenarios, as depicted in Figure 6(a), Figure 6(b) and Figure 6(c), respectively. In the HAND scenario, the camera is placed in the hands of the kid. In the HAT scenario, the camera is placed in Smart Hat (our proposed system) of the kid. Finally, in the SHOULDER scenario, the camera is placed on the shoulder of the kid.

Table 3. Parameters considered in the paper

Parameter	Value
Simulator	Matlab 2018b
Pixel density	100 dpi- 900 dpi
Allowed time delay	1 sec- 7 sec
Kids of age	2 years- 5 years

(a) HAND Scenario (b) HAT Scenario (c) SHOULDER Scenario

Figure 6. Three scenarios considered in the experiments

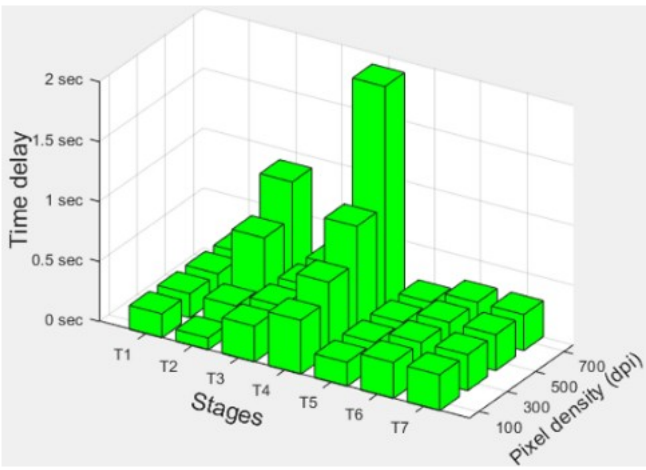


Figure 7. Comparison of “Time delay” for seven stages in terms of pixel density

Figure 7 compares the time delay of seven stages by varying the pixel density ranging from 100 dpi to 700 dpi. As shown in Figure 7, stages T_1 have constant time delay because the pixel density does not impact Bluetooth transmission. At stage T_2, time delay slightly increases as the pixel density increases because the capturing an image with more pixels consumes more time. At stage T_3 the time delay is constant because image reading with OpenCV is not affected by pixel density. At stage $T_4$the time delay goes on increasing as the pixel density increases. This occurs because, the stage T_4 is the classification stage of TensorFlow and hence requires time based on the pixel density of the image. This increases the time delay of stage T_4 as the pixel density increases. Stages T_5 and T_6 have constant time delay even though the pixel density is increasing. This occurs because that the Google Translate API and gTTS API don't related to the pixel density. Finally, stage T_7 also has constant time delay because wireless transmission operation doesn't impact by the pixel density.

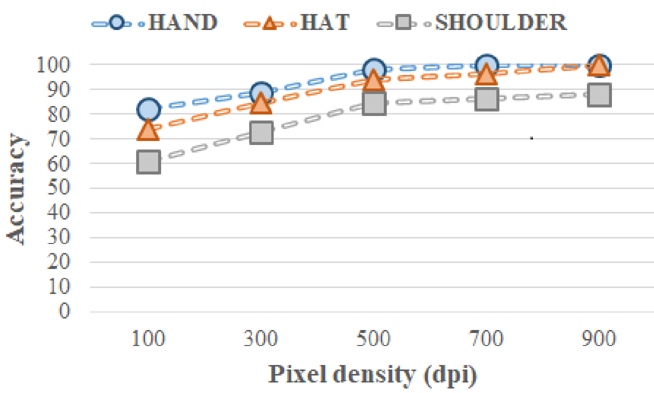


Figure 8. Comparison of accuracy for “three scenarios” in terms of pixel density

Figure 8 compares the accuracy of the three scenarios “HAND”, “HAT” and “SHOULDER” by varying the pixel density ranging from 100 dpi to 900 dpi. The accuracy is measured by the percentage of the valid classification. In comparison, the scenario HAND has the best performance than the scenario HAT. Even though the HAND yields more favorable performance, the kid cannot always carry the camera in his/her hands and this is the main reason that this paper proposes the "Smart-Hat" system.

Figure 9 further compares the accumulated time delay of seven stages ($T_1 + T_2 + \ldots + T_7$) by varying the pixel density ranging from 100 dpi to 700 dpi. It can be observed that the accumulated time delay is 1.85 seconds exactly in the case of 100 dpi and it takes 2.1 seconds to perform the classification in the case of 300 dpi. It takes 2.8 seconds for the 500 dpi and it takes 4.1 seconds to perform the classification in the case of 700 dpi. If the maximal allowed delay is 3 seconds, 500 dpi is the best policy because within 3 seconds (2.8) it can perform the classification. However, in the case of 700 dpi, it cannot perform the task because it needs more time for the classification as the pixel density is higher and it can be seen from Figure 9 that 700 dpi has crossed the boundaries of 3 seconds time delay.

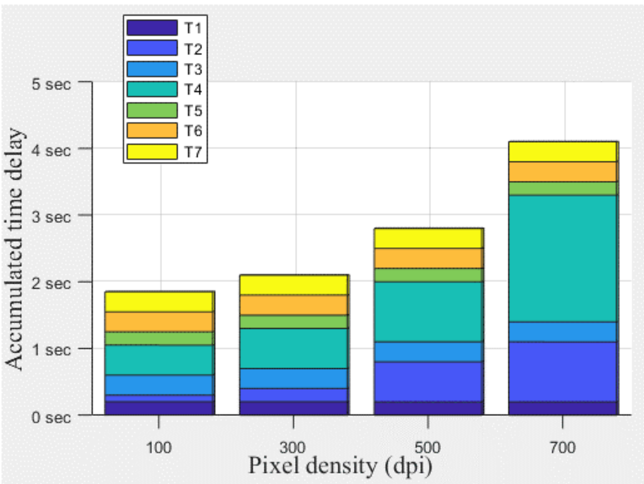


Figure 9. Comparison of "Accumulated time delay" by varying pixel density

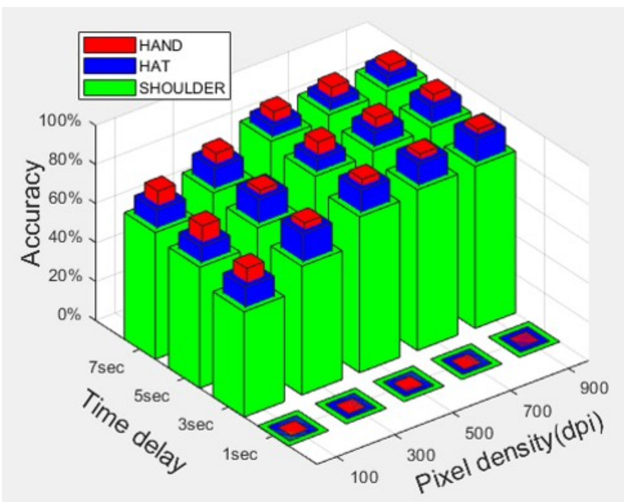


Figure 10. Comparison of "Accuracy" for three scenarios in terms of Time delay by varying pixel density

Figure 10 shows the accuracy by varying the pixel density ranging from 100 dpi to 900 dpi in three scenarios HAND, HAT and SHOULDER. The scenario HAND outperforms the other two mechanisms in terms of accuracy. The main reason is that the camera in the HAND can capture the image perfectly and thus classification result is better than the other two mechanisms. As shown in Figure 10, the performance of 900 dpi at 7 seconds allowed time delay was more favorable than that of the other pixel density. However, in the case of 100 to 900 dpi and the allowed time delay is 1 second, there is no output result. This occurs because that the allowed time delay is less and the processing time is larger than the delay bound. As shown in Figure 10, accuracy increased with the number of pixel density but in the case of 500, 700 and 900 dpi the accuracy remains almost same even though the pixel density is increasing because pixel density has a very low impact on accuracy.

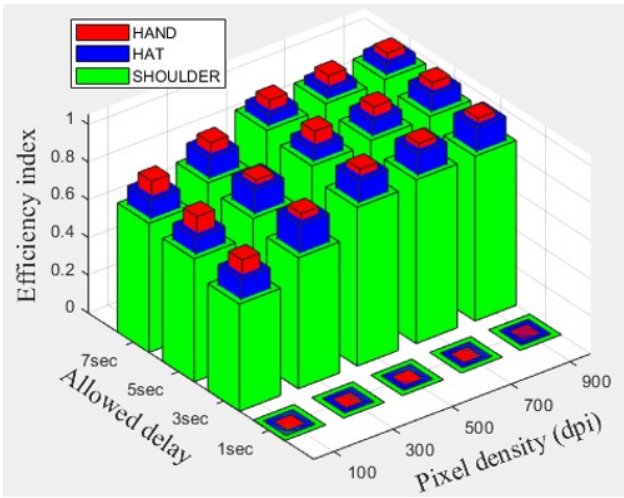


Figure 11. Comparison of "Efficiency index" 'for different pixel density by varying the allowed delay

Figure 11 shows the efficiency index for 100 to 900-pixel density by varying the allowed time delay from 1 to 7 seconds. It is obvious that the image with high pixel density requires much time for processing. However, given a constraint of an allowed delay time, it constrains that the time for processing the image cannot be larger than the allowed delay. This implies that the pixel density should have an upper bound for the given allowed delay. Hence, the maximal accuracy can be achieved.

The efficiency index considers two parameters: the maximum accuracy and current accuracy. The maximum accuracy is achieved when the image has maximal pixel density but guarantees that the image processing time does not exceed the given allowed delay. For example, assume that the image with a pixel density of 500 dpi requires a processing time of 3 seconds. Assume that image with 500 and 300 dpi can have accuracies 0.95 and 0.88, respectively. Given the allowed delay time 3 seconds, if the image pixel is only

300 dpi, then the accuracy is not good since the image with 500 dpi can be processed within 3 seconds. In this case, the efficiency index can be calculated by

$$1 - (0.95 - 0.88)/0.95 = 0.9263$$

Let $I_{efficiency}$ denote the efficiency index, which is equal to the current accuracy subtracted from the maximum accuracy, as shown in Exp. (1).

$$I_{efficiency} = 1 - [((Maximum\ Accuracy - Current\ Accuracy))/(Maximum\ Accuracy)] \quad (1)$$

A large value of current accuracy can lead to a large value of efficiency index since the maximum accuracy is fixed under a given allowed delay time.

In the experiment of Figure 11, efficiency index is shown by varying the pixel density ranging from 100 dpi to 900 dpi. It shows that the efficiency index increases as the time delay increases. This occurs because it takes more time to classify high pixel density images and for 500 dpi 3 seconds maximal allowable delay is the best time as it is shown in Figure 9. In comparison, the efficiency index of HAND is higher than those of HAT and SHOULDER. Recall the performance results of Figure 10. Using hand to control the camera can obtain a better quality of the picture than HAT and SHOULDER. Therefore, the HAND has the best accuracy. Since the calculation of the efficiency index as shown in Exp. (1) is related to the accuracy, the HAND has higher efficiency index value than HAT and SHOULDER, as shown in Figure 11.

We invited 10 kids and their parents to give feedback by wearing our Smart Hat. More than 85 percent of kids say "It's interesting", "It's comfortable", "It's funny", "I like to wear the Hat". In addition, some good comments collected from their parents are "Smart Hat is very helpful for my kid", "Playful nature of the hat has an immediate grasp on children" and "My kid is excited to wear the hat". Figure 12 compares the accuracy of the four different ages of kids in three scenarios HAND, HAT and SHOULDER by varying pixel density ranging from 100 dpi to 700 dpi. As shown in Figure 12, the kids of age 5 years old yields the best performance and outperforms the other three ages of kids. The kids with 3 – 4 years old will have the capability to handle the Smart Hat and it is the learning stage of kids. Hence, the accuracy increases with the pixel density. A kid with 2 years old doesn't know the importance of the pixel density as well as handling the Smart Hat, causing that the accuracy remains low even though the pixel density is increasing. As a result, our proposed Smart Hat will help the kids between the ages of 3 – 8 by aiding them in their primary learning task to identify objects in the real-time environment without the supervision of third- party.

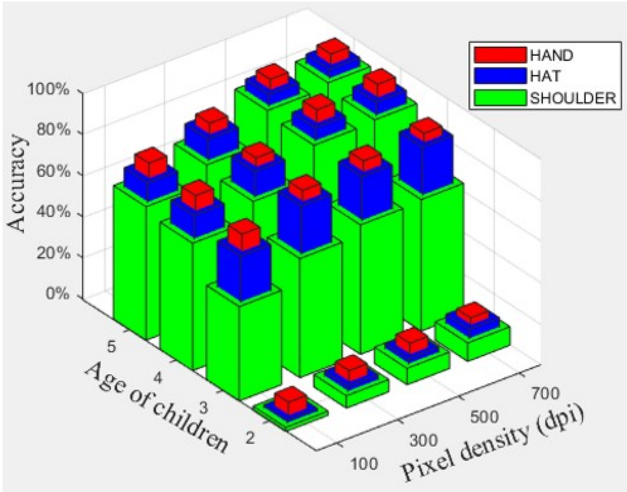


Figure 12. Comparison of "Accuracy" in terms of different ages of kids by varying pixel density for three scenarios

Figure 13 compares the accuracy with various distances from Access point (AP) to the Smart Hat. As shown in Figure 13, the response time increases as the distance from the AP to Smart Hat increases. This occurs because that when the distance is long, the signal will become weak and there will be packet loss during the transmission. As a result, the accuracy decreases with the distance from the AP to Smart Hat. The accuracy decreased to 88 % when the distance is 15 meters but when the distance is 30m and 50m accuracy decreased to 76% and 65 %.

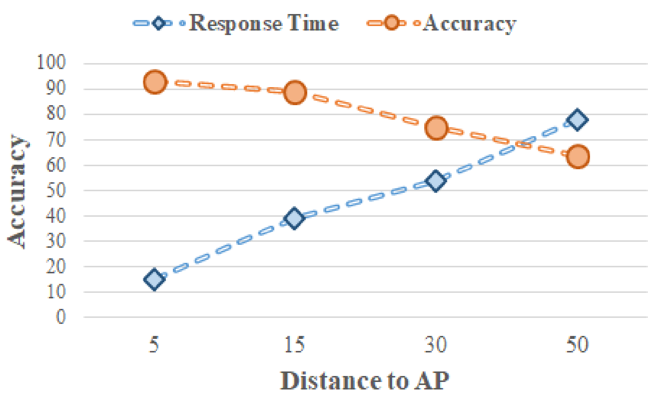


Figure 13. Comparison of "Accuracy" by varying the distances between AP and Smart Hat

Figure 14 shows that the results of object recognition are done by Inception-v3 model with the TensorFlow. The results are astonishing. On an Intel Core i7-4790 CPU @3.60GHz with 12GB of RAM, we got more than 90 % of accuracy and very confident Classification. The experiments were done on different illumination conditions. The intensity of light plays a major role in object recognition. The high intensity of light is required to have better results in object recognition. In general, white light is required to have a clear vision of an object because the spectrum

Figure 14. Examples of various objects recognition on the system

of white light has a mixture of different colors Violet, Indigo, Blue, Green, Yellow, Orange, and Red (VIBGYOR). However, our proposed system is outdoor learning and the intensity of light will be high during the daytime. Finally, all the experiments were conducted by the researchers except the sixth experiment.

7 Conclusion

Kid education is very different from adults, as the mind of the kid is very different and sensitive. Successful preschool education plays a significant role in a kid's development and it has to be done in a very careful manner. This research focused on investigating how to support self-learning of modern-day preschool kids to recognize the objects in the outside environment. The main effort is to provide these kids with a Smart Hat, in order to help kids in their primary learning task of identifying objects just by using their own Smart Hat. The proposed solution is to address the problem of creating a learning platform for the kids by combining IoT, AI and speech corpus models to provide an interactive learning experience for the particular objects in the outside environment. Compared to traditional learning, our proposed strategy eventually increases kids learning capability with great memorization skills. Since our proposed system is unique and effortless learning method, the kids will be more fascinated to learn new things by using our Smart Hat. At the end of the paper, performance results confirm the goodness of our design choices. The future work would consider that kids can interact with the content of the image. That is to say, kids can ask some questions about the content of the image and get a response from our system. Also, we will embed the GPS module in our device to provide the real-time tracking of children to their parents when they explore in the outside environment. Therefore, it will be a practical solution to parents for supporting the learning needs of their kids.

References

[1] W. Cho, Y. Ma, Y. Huang, A Smart Socket-Based Multiple Home Appliance Recognition Approach over IoT Architecture, *Journal of Internet Technology*, Vol. 16, No. 7, pp. 1227-1238, December, 2015.

[2] C. Li, C. Lee, C. Y. Weng, A Secure Three Party Node Authentication and Key Establishment Scheme for the Internet of Things Environment, *Journal of Internet Technology*, Vol. 19, No. 1, pp. 147-155, January, 2018.

[3] S. Raja, T. Sampradeepraj, Internet of Things: A Research Oriented Introductory, *International Journal of Ad Hoc and Ubiquitous Computing*, Vol. 29, No. 1-2, pp. 4-14, September, 2018.

[4] S. Y. Moon, J. H. Park, J. H. Park, Authentications for Internet of Things Security: Threats, Challenges and Studies, *Journal of Internet Technology*, Vol. 19, No. 2, pp. 349-358, March, 2018.

[5] F. Li, Y. Du, From AlphaGo to Power System AI: What Engineers Can Learn from Solving the Most Complex Board Game, *IEEE Power and Energy Magazine*, Vol. 16, No. 2, pp. 76-84, March-April, 2018.

[6] Q. Cui, Z. Zhou, C. Yuan, X. Sun, Q. M. Wu, Fast American Sign Language Image Recognition Using CNNs with Fine-tuning, *Journal of Internet Technology*, Vol. 19, No. 7, pp. 2207-2214, December, 2018.

[7] W. Cho, Y. M. Huang, Long-Term Motion and Pace Detection System with Wearable Shoe Module, *Journal of Internet Technology*, Vol. 16, No. 4, pp. 581-591, July, 2015.

[8] D. Mulfari, A. L. Minnolo, A. Puliafito, Building TensorFlow applications in smart city scenarios, *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*, Hong Kong, China, 2017, pp. 1-5.

[9] L. L. Lewis, Y. A. Kim, J. A. Bey, Teaching Practices and Strategies to Involve Inner-city Parents at Home and in the School, *Journal of Teaching and Teacher Education*, Vol. 27, No. 1, pp. 221-234, January, 2011.

[10] V. Grlaš, M. Marinović, Ecological Education in Preschool Education through Visual Communication, *The 33rd International Convention MIPRO*, Opatija, Croatia, 2010, pp. 959-963.

[11] A. S. Drigas, G. Kokkalia, A. Economou, Mobile Learning for Preschool Education, *International Journal of Interactive Mobile Technologies (iJIM)*, Vol. 10, No. 4, pp. 57-64, October, 2016.

[12] B. Beschorner, A. Hutchison, iPads as a Literacy Teaching Tool in Early Childhood, *International Journal of Education in Mathematics, Science and Technology*, Vol. 1, No. 1, pp. 16-24, January, 2013.

[13] A. Causo, P. Z. Win, P. S. Guo, I. Chen, Deploying Social Robots as Teaching Aid in Pre-school K2 Classes: A Proof-of-concept Study, *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, 2017, pp. 4264-4269.

[14] I. Culjak, D. Abram, T. Pribanic, H. Dzapo, M. Cifrek, A Brief Introduction to OpenCV, *35th International Convention*

MIPRO, Opatija, Croatia, 2012, pp. 1725-1730.

[15] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the Inception Architecture for Computer Vision, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, 2016, pp. 2818-2826.

Biographies

I-Hsiung Chang received the Ph.D. degree in education from the National Chengchi University, Taiwan in 2012. He is currently a Ph.D. student with the Department of Computer Science and Information Engineering from 2018, Tamkang University and also an Assistant Professor in the Department of Early Childhood Education, TOKO University, Taiwan. His current research interests cover the joint of Education and Information Technologies in AI-IoT, Robot and Big Data.

Huan-Chao Keh received both Master and Ph.D. degrees from the Department of Computer Science at Oregon State University, USA. He is currently a full professor in the Department of Computer Science and Information Engineering at Tamkang University, Taiwan. He served as the Dean of Student Affairs from 1997 to 2003, the Dean of Academic Affairs from 2004 to 2014 and the Vice President for Academic Affairs from 2014 to 2018. He has been the President of Tamkang University since August 1, 2018. He published more than 60 research papers and developed a leading Unique Device Identification (UDI) system which can be used to mark and identify medical devices within the healthcare supply chain. His current research interests include Data Mining, Internet of Things, Artificial Intelligence and Clinical Medical Information Systems.

Bhargavi Dande received the M.S. degree in Computer Science and Information Engineering from Tamkang University, Taiwan, in 2019. She is currently pursuing the Ph.D. degree in Computer Science and Information Engineering at Tamkang University, New Taipei, Taiwan. Her current research interests are Internet of Things, Wireless Sensor Networks and Artificial Intelligence.

Diptendu Sinha Roy received the Ph.D. degree in Engineering from Birla Institute of Technology, Mesra, India in 2010. In 2016, he joined the Department of Computer Science & Engineering, National Institute of Technology (NIT) Meghalaya, India as an Associate Professor, where he has been serving as the Chair of the Department of Computer Science & Enginering since January 2017. Prior to his stint at NIT Meghalaya, he had served the Department of Computer science & Engineering, National Institute of Science & Technology, Berhampur, India. His current research interests include software reliability, Distributed & Cloud Computing and IoT, specifically applications of artificial intelligence/ Machine Learning for smart integrated systems. Dr. Sinha Roy is a member of the IEEE Computer Society.