ISSN- 2394-5125 VOL 06, ISSUE 05, 2019

IOT INTERACTION SYSTEM WITH GAZE TRACKING AND NATURAL HEAD MOVEMENTS

M. Pravin[1], S. Prabhu[2], P. N. Sureshkumar[3]
Department of Electronics and Communication,
CMS College of Engineering and Technology,
Coimbatore, India.

ABSTRACT

This paper presents an interaction system , where users can control an IoT device by gazing at it and doing simple gestures. The ultimate goal of the proposed system is to develop an efficient interaction method between users and IoT devices so that any user including parents, children, and the elderly can control the devices intuitively. There have been various studies to enhance the interaction between the users and devices by applying the recent machine learning- based technologies, such as voice recognition, gesture recognition, gaze estimation, etc. These have focused on improving the accuracy of recognizing complex sentences or gestures; however, people with disabilities or severe diseases, the elderly, and children cannot easily perform such tasks. The proposed system consists of object detection module, gaze estimation module, hand gesture recognition module and IoT controller module.

INTRODUCTION

The recent developments in network infrastructure and smart devices have resulted in the rapid spread of the Internet of things (IoT) applications and services. The IoT is defined as a network of inter- connected things (e.g., computers, vehicles, and sensors) that exchange data and information among themselves with and other services. The number of internet-connected devices is now dramatically growing. According to a recent study on the prediction of IoT market share, the number of IoT devices will approach 100 billion and the total amount of data generated by the users and devices will reach 35 ZB by 2020. It is therefore expected that the success of IoT will allow the users and things to be connected anytime, anywhere using any path, network, or service. These characteristics of the IoT ecosystem will improve the quality of services in various application domains such as health-care (e.g., remote patient monitoring and treatment), transportation (e.g., smart transportation systems), and home automation (e.g., smart appliances).

A typical architecture of the IoT comprises a perception layer, a network layer, and an application layer. Numerous studies have been conducted to address the challenging issues in each layer of the architecture. The perception layer first interacts with physical devices like RFID, sensors, and actuators and then connects the devices to the network of IoT. The network layer of the IoT is responsible for transmitting data between different things, applications, and services using heterogeneous networks and communication protocols. Finally the application layer exploits the data from the underlying layers to build and provide the required services.

Interoperability of IoT devices, applications, services, and users is an important factor for the successful implementation of IoT. Therefore, the networklayer and its corresponding technologies (e.g., Bluetooth,Wi-Fi, 6LoWPAN, ZigBee, Z-Wave, MQTT, and CoAP, etc.) have been usually considered the most important components since they are closely related to the connectivity, interoperability, and scalability of IoT architectures.

Similarly, the IoT standards, frameworks, and platforms in the application layer aim to maximize the interoperability by abreacting the layers of the IoT architecture and providing an efficient user interface (UI).

However, the current technologies mainly focus on improving the machine-to-machine communication/interaction, rather than the interaction between users and machines. For example, some IoT platforms designed for smart home automation provide a web-based UI and a mobile application to register, manage, and control the smart home appliances connected to them. Users must first open the website or the mobile application, explore a page to select a menu, find a room or location, and finally select the device to be manipulated from a list. After selecting the device, the users can check the status or control it by touching or clicking the buttons on the webpage or mobile page. However, this UI and procedure will become tedious and time-consuming to the users with the current rapid increase in the number of IoT ready devices. Additionally, users who are not familiar with smart devices, such as children and seniors or severely ill patients or the disabled, will encounter difficulties in using IoT applications and services.

This paper focuses on the interaction between people and IoT devices, which can affect the usability of the IoT platforms. It briefly depicts various approaches to help users interact with smart devices.It has a proposed smart IoT interaction system with gaze estimation and object detection modules.

LITERATURE REVIEW

Improving the Usability of Remote Eye Gaze Tracking for Human-Device Interaction

Author: Kang-A Choi, Chunfei Ma, and Sung-Jea Ko, Fellow, IEEE. This paper presents a novel single-point calibration-based remote eye gaze tracking (REGT) method. The proposed method consists primarily of two steps. First, a user calibration database (UCDB) that contains user-specific data associated with multiple calibration points (CPs) is constructed for a certain number of users. Second, for new users, the candidate calibrated data are retrieved from the UCDB by simply requesting the users to look at a single CP on the screen centre (SC). Experimental results show that the proposed REGT method with single- point calibration demonstrates a highly competitive accuracy compared with conventional methods employing multipoint calibration.

Hand Gesture Based Remote Control System Using Infrared Sensors and a Camera

Author: Faith Erden and A. Enis Cetin, Fellow, IEEE. A multimodal hand gesture detection and recognition system using differential Pyro-electric infrared (PIR) sensors and a regular camera is described. Any movement within the viewing range of the differential PIR sensors are first detected by the sensors and then checked if it is due to a hand gesture or not by video analysis. If the movement is due to a hand, one dimensional continuous-time signals extracted from the PIR sensors are used to classify/recognize the hand movements in real-time. Classification of different hand gestures by using the differential PIR sensors is carried out by a new winner-take-all (WTA) hash based recognition method. Jaccard distance is used to compare the WTA hash codes extracted from I-D differential infrared sensor signals. It is experimentally shown that the multimodal system achieves higher recognition rates than the system based on only the on/off decisions of the analog circuitry of the PIR sensors.

DeepGaze II: Reading fixations from deep features trained on object recognition

Author: Matthias Kummerer, Thomas S. A. Wallis and Matthias Bethge,Fellow, IEEE. Here we present DeepGaze II, a model that predicts where people look in images. The model uses the features from the VGG-19 deep neural network trained to identify objects in images. Contrary to other saliency models that use deep features, here we use the VGG features for saliency prediction with no additional fine-tuning (rather, a few readout layers are trained on top of the VGG features to predict saliency). The model is therefore a strong test of transfer learning. After conservative cross-validation, DeepGaze II explains about 87% of the explainable information gain in the patterns of fixations and achieves top performance in area under the curve metrics on the MIT300 hold-out benchmark. These results corroborate the finding from DeepGaze I (which explained 56% of the explainable information gain), that deep features trained on object recognition provide a versatile feature space for performing related visual tasks. We explore the factors that contribute to this success and present several informative image examples.

Directional Correlation Filter Bank for Robust Head Pose Estimation and Face Recognition

Author: Si Chen, Dong Yan, and Yan Yan, Xiamen University. During the past few decades, face recognition has been an active research area in pattern recognition and computer vision due to its wide range of applications. However, one of the most challenging problems encountered by face recognition is the difficulty of handling large head pose variations. Therefore, the efficient and effective head pose estimation is a critical step of face recognition. In this paper, a novel feature extraction framework, called Directional Correlation Filter Bank (DCFB), is presented for head pose estimation. Specifically, in the proposed framework, the 1-Dimentional Optimal Tradeoff Filters (ID- OTF) corresponding to different head poses are simultaneously and jointly designed in the low-dimensional linear subspace. Different from the traditional methods that heavily rely on the precise localization of the key facial feature points, our proposed framework exploits the frequency domain of the face images, which effectively captures the high-order statistics of faces. As a result, the obtained features are compact and discriminative. Experimental results on public face databases with large head pose variations show the superior performance obtained by the proposed framework on the tasks of both pose estimation and face recognition.

A personal Health Care Office Chair

Author: Sajid Mubashir Sheikh and Ibo Ngebani, University of Botswana. Over the years, Medical Doctors and Researchers have reported that prolonged sitting as well as bad sitting postures can be a danger to human health such as chronic diseases attributed to rheumatic disorders as well as pains in muscle and connective tissues of tendons, joint capsules and ligaments. To address these issues, in this work, we propose a personal health care office chair which is developed using IoT technology. The chair detects and monitors how long a person sits on the chair, as well as detects the sitting posture of the person. Warnings are signalled on the LCD as well as signed through sound from a buzzer. A Wi-Fi module is used to transmit the data to the internet for recording.

Crowdsourced Children Monitoring and Finding with Holding Up Detection Based on Internet of Things Technologies

Author: Lien-Wu Chen, Tsung-Ping Chen, Hsein-Min Chen and Ming-Fong Tsai, Senior Member IEEE. In this paper, we propose a crowdsourced children monitoring and finding (CCMF) framework to detect holding-up behaviours and find missing children using wearable devices and surrounding smartphones based on IoT technologies. In the monitoring mode, the CCMF framework can prevent young children from taking away by strangers with bad intentions. In the finding mode, the CCMF framework can cooperatively find missing children equipped with wearable devices consisting of mobile iBeacon and 3-axis accelerometer modules through crowdsourced sensing networks formed by smartphone users with outdoor GPS and indoor IoT localization. According to our review of relevant research, CCMF is the first children monitoring and finding solution that can detect holding-up postures of a target child and provide the guiding path to a lost child through crowdsourced sensing networks. An iOS-based prototype with Arduino wearable devices and mobile/static ibeacon nodes is implemented to verify the feasibility and superiority of our framework. Experimental results show that CCMF outperforms existing methods and can significantly increase recognition success rates and efficiently reduce false alarm rates of holding up detection.

EXISTING SYSTEM

There have been various studies to enhance the interaction between the users and devices. In the section, we briefly review the previous approaches based on head pose estimation, gaze estimation and eye tracking, and gesture recognition. The recent efforts to support the IoT system interaction with user behavior detection techniques are also introduced.

Head pose estimation can be used to monitor the user’s attention while interacting with devices. For example, a driver’s head pose estimation can be used to detect drowsiness or to monitor the driver’s attention. Eye tracking and gaze estimation are employed to estimate the gaze position by monitoring the user’s pupil movement. These technologies are particularly useful for severely ill patients who cannot easily perform daily life voluntary tasks.

In contrast to the model-based approaches, appearance-based approaches extract visual features from input images to learn a direct mapping function of the image features and gaze points. These

ISSN- 2394-5125 VOL 06, ISSUE 05, 2019

approaches do not require high resolution input images or expensive hardware setup and calibration; therefore can be more suitable for a practical environment. A neural-network based real-time approach with a standard webcam has also been studied. The work from introduced a way to help people draw lines and type text on the screen with their eye movements using a deep multi-layer perceptron algorithm. As these methods basically rely on feature representations to learn mapping functions, identifying useful features from images is important. The recent advances of deep learning approaches like convolutional neural networks (CNNs) have increased the efficiency of image understanding and classification tasks.

Finally, several research works have developed a unified framework with various technologies to support efficient IoT system interaction. The input modalities used in these studies can be categorized into the followings: 1) photo, 2) gesture, and 3) gaze. Even though many studies have been conducted to improve the interaction between the users and devices, the successful implementation of IoT environments and services still face some challenges. First, some of the existing works are not suitable for IoT services and applications. The head pose based approaches, target small and controlled conditions and cannot be directly applied to IoT environments such as a house or room.

The eye tracking and gaze estimation approaches generally detect the gaze positions on a small screen in front of the users. However, in IoT environment, smart devices and appliances can be installed anywhere in the room or the building; hence these techniques must be improved to work far from. Second, some previous works require the users to wear additional devices, such as a smart watch, headband or smart glasses. As the number and category of IoT devices are growing, the scalability and interoperability of IoT interaction systems need to be further addressed. However, this requirement is restrictive for those who are not familiar with or not able to use smart devices.

PROPOSED SYSTEM

The ultimate goal of the proposed system is to develop an efficient interaction method between users and IoTdevices so that any user including patients, children, and the elderly can control the devices intuitively. The hardware module called “Watch” is located at the center of a room (e.g., a living room or a patient’s room). This module estimates user’s gaze position and detects the IoT devices installed in the room. Another module called “Do” is installed around the user (e.g., near the arms) to detect hand gestures. The proposed system works as follows. First, the “Watch” module records the opposite side of the user to detect and recognize the types of IoT devices installed in the room. Second, the “Watch” module detects the user’s head region and then computes a fine- grained head pose information (i.e., pitch, yaw, and roll) to estimate the user’s gaze position. With this information, the proposed system can identify the target device. Then, the “Do” module captures the user’s hand gestures. A combination of hand gesture information and the type of selected IoT device is then translated into an IoT command and transmitted to IoT platforms. Finally, the device is manipulated according to the command. For example, the “Watch” module recognizes the position and the type of each installed device (i.e., a lamp, monitor, fan, refrigerator from left to right). If a user stares at a device at the left-most side, the "Watch" module detects the user's gaze position and predicts the type of the device to be controlled as a lamp. Finally, the user can turn the lamp on or off with a swipe gesture performed near the "Do" module installed around the user.The schematic architecture of the proposed system consists of "Watch" phase and "Do" phase. The goal of the "Watch" phase is to identify the target device by estimating the user's gaze position and classifying the device category. For this, a gaze estimation module and an object detection module based on deep neural networks have been adopted. The "Do" phase is to capture the user's hand gestures to compose an IoT device control command using the device and gesture information. Therefore, the "Do" phase exploits a gesture recognition module with an embedded sensor and an IoT controller module which interacts with the IoT platform. The rest of this section provides details on each module.

The proposed system mainly consists of

1) Object detection module: - The object detection module captures a picture of the room to detect and recognize the installed IoT devices using deep learning approaches.

2) Gaze estimation module: - The gaze estimation module records and detects the face of the user from the video stream.

3) Hand gesture recognition module: - The hand gesture recognition module is a simple hardware module equipped with proximity-based sensors to detect hand gestures.

4) IoT controller module: - The user's command to control an IoT device is recognized and then sent to the IoT controller module.

OVERVIEW

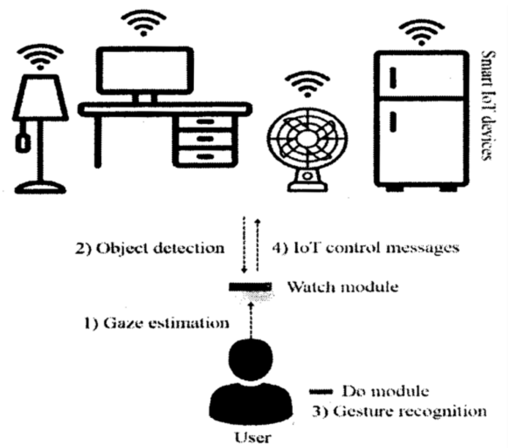


Fig 3.1 Overview of the proposed system

ISSN- 2394-5125 VOL 06, ISSUE 05, 2019

INDOOR GAZE ESTIMATION BY HEAD POSE ESTIMATION

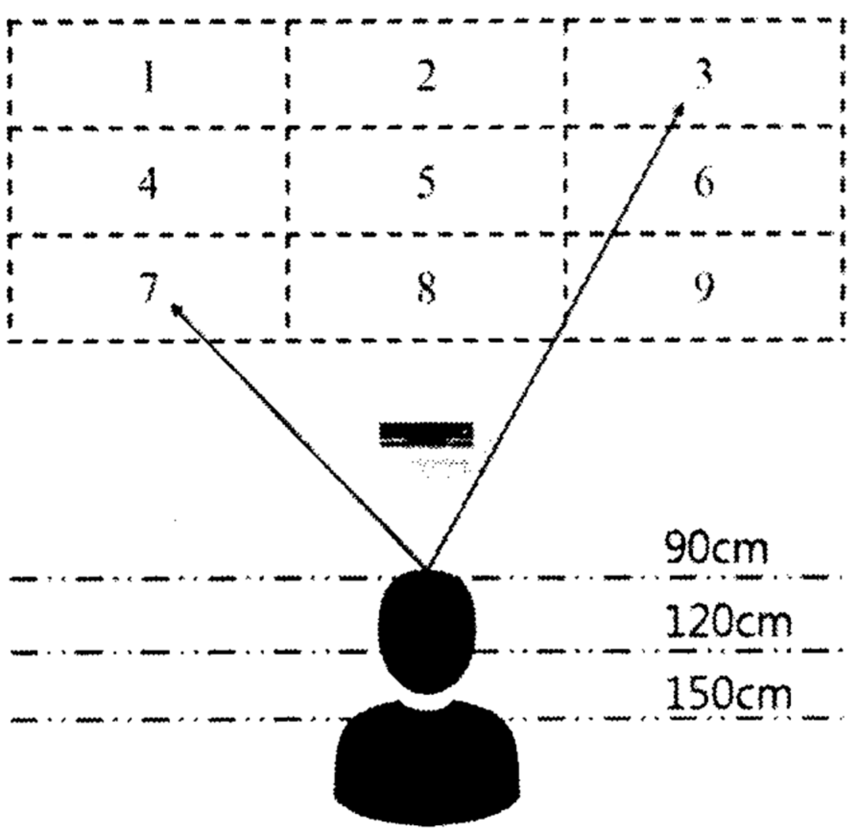


Fig 3.2 Overview of indoor gaze estimation by head pose estimation

EXPERIMENTAL PROCEDURE

To evaluate the performance and feasibility of the proposed system, we conducted quantitative experiments with a prototype. The prototypes of the "Watch" module and "Do" module were configured using commercial webcams and small single-board computers. The deep learning models used for gaze estimation and object detection were trained using a desktop PC with commercial graphic cards supporting parallel computations. The IoT platform used in our experiment is a prototype implementation from the author's previous work.

PROTOTYPE IMPLEMENTATION

First, the "Watch and the "Do" hardware modules were installed in a room. The Watch module consists of a small single-board computer equipped with a front and back cameras for gaze estimation and object detection, respectively. The "Do" module consists of a unified sensor board and a gesture sensor for hand gesture recognition.

To perform the user study, a dashboard application that can display the video frames recorded by the

"Watch" module is implemented. The dashboard application displays the results of gaze estimation and objects selection (Object form), gesture recognition (Event form), and the composed IoT command (Details form). Here, a user was asked to control a device with hand gestures. For example, an instruction, "Turn down the volume of the TV". It can be found that the user gazed at the TV in front of him and then performed the "right" hand gesture to turn down the volume of it.

Fig 9.1 Supported hand gestures. Each image represents "right", "left", "backward"and "forward" gesture left to right, respectively

Device	Gesture	Meaning
Cleaner, rice, cooker, washing machine	Forward	Turn on
c	Backward	Turn off
Air conditioner	Forward	Turn on
	Backward	Turn off
	Left	Temperature up
	Right	Temperature down
tv/monitor	Forward	Turn on
	Backward	Turn off
	Left	Volume up
	Right	Volume down
Fan	Forward	Turn on
	Backward	Turn off
	Left	Speed Up
	Right	Speed Down
Door	Forward	Open
	Backward	Close
	Left	Lock
	Right	Unlock

Table 9.1.(a) Device Type and its corresponding gestures

ISSN- 2394-5125 VOL 06, ISSUE 05, 2019

PERFORMANCE EVALUATION

The classification accuracy of gaze estimation for different classifier models and features are discussed. To train the final gaze estimator, we adopted decision tree, random forest, SVM, and k-NN algorithms. The dataset consists of 28,350 training images and 9450 testing images from nine subjects. The classifiers were trained to output the label of the grid from the input feature vectors obtained from the DeepGaze framework. For quantitative evaluation of the proposed system, we also developed two more classifiers using gaze features obtained from the head pose estimator and a model-based gaze tracking method. Without the distant feature, the decision tree algorithm with OpenFace features showed the best result (76.06%), which is slightly higher than the SVM algorithm with DeepGaze features (75.95%). On including the distance feature, the performance of every classifier, except the OpenFace-based models, improved significantly, and the maximum gain was 14.12% point for the decision tree algorithm with HopeNet features.

The model-based approaches require high-resolution eye images to compute geometric local features; therefore, the performance gain would be limited if the OpenFace-based method cannot detect the eye region from the images. As a result, the random forest model with the face orientation features from DeepGaze and the face distance feature demonstrated the highest classification accuracy (87.56%) among all the models used in the experiment.

The results show that the random forest models with DeepGaze, HopeNet, and OpenFace outperform the CNN-based approaches. The maximum accuracy was observed for the random forest model with the DeepGaze features. However, the CNN-based approaches performed competitively without any feature engineering. It is expected that adding optimization techniques will improve the overall performance of the CNN-based approaches.

The dataset used for our experiment totally consists of 3065 images of IoT devices (235 images of an air conditioner, 774 images of a rice cooker, 407 images of a washer, 625 images of a cleaner, 315 images of a fan, 530 images of a TV, and 179 images of a door). Due to the lack of a large-scale well-refined publicly available home appliance image dataset, we collected sample images of each device type from various web search engines. After downloading the images, we manually labeled them to annotate the bounding box and device category information for the YOLO detectors.

The classification model was trained for 1000 iterations. For evaluation of the classification accuracy, we performed a 10-fold cross validation and computed the average precisions. On average, the proposed system achieved a high accuracy of 92.25%. Among the devices, the highest classification accuracy was found in the fan (97.17%), while the lowest was for the air conditioner (80.69%). The experimental results on IoT device detection and classification can be interpreted as follows. First, there are no strong correlations between the number of training samples and the classification accuracy in our domain. For example, even though both the air conditioner and the door have relatively low number of samples (235 and 179, respectively) there exists a big difference between their classificationperformances (80.69% and 93.56%), respectively). However, it is difficult to find unique characteristics in the appearance of air conditioners, which results in their low classification accuracy.

ISSN- 2394-5125 VOL 06, ISSUE 05, 2019

At the 100th iteration, the system produced a poor result of 4.16% on average. However, the performance sharply increases until 300th iteration (73.70%) and then shows a gradual increase until 1000th iteration (92.25%). The details of the classification accuracy of each device type according to the number of iterations can be found in Table II.

The performance of gaze estimation, IoT device detection, and IoT device categorization of the proposed approach is acceptable (around 90%). Therefore, the overall architecture and design of the proposed system should be feasible for use in an IoT environment with further improvements.

Iter	AC	RC	Washer	cleaner	Fan	TV	Door
100	3.72	8.84	1.33	2.80	4.71	5.30	2.45
200	18.41	73.18	63.75	54.40	70.31	50.60	57.65
300	25.19	82.95	93.29	83.54	91.42	60.81	78.74
400	32.56	89.31	82.99	78.46	93.47	85.76	81.78
500	41.65	90.82	93.40	88.26	95.01	86.00	87.29
600	38.97	91.34	94.72	90.71	93.48	88.81	90.21
700	37.93	92.15	93.67	91.33	94.80	89.27	91.48
800	48.46	93.27	96.61	92.71	96.34	90.71	93.11
900	78.71	93.38	96.83	92.68	96.22	89.62	93.60
1000	80.69	93.24	95.83	92.72	97.17	92.54	93.56

Table 9.2.(a) Classification accuracy according to the number of iteration

DISCUSSIONS AND RESULTS

In this work, we proposed a novel smart IoT interaction system. The proposed method helps users easily control IoT devicws by gazing at a device and performing simple hand gestures. Deep convolution neural network based approaches were adopted to estimate user's gaze position and to detect and recognize the IoT devices. A simple module with a gesture sensor was implemented. The final IoT command to be transmitted to the IoT platform was composed based on the detected IoT devices and hand gestures. We presented the feasibility of the proposed system through various experiments.

The main contributions of the work are as follows. First, we proposed a novel approach to utilize head pose estimation and object detect technologies for IoT interaction. The previous works on gaze estimation focused on tracking the gaze position on a display. However, these approaches cannot be applied to the IoT environment that requires indoor gaze estimation. In this work, we solve this problem by a combination of object detection and implicit gaze estimation from the fine-grained head pose information. Second, without any additional devices, such as a headband or watch, the users can specify the target device. The previous works used wearable devices to detect user's head pose or gaze direction, which can be uncomfortable for many years. With this work, the users can interact with the IoT devices intuitively by installing simple modules.

Results :

Fig R1 : Screenshot of result

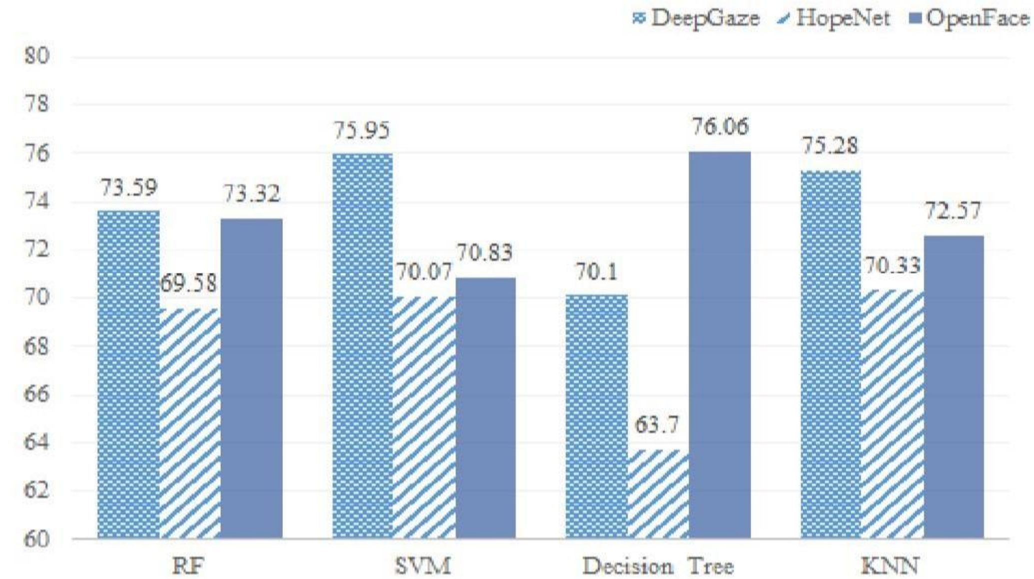


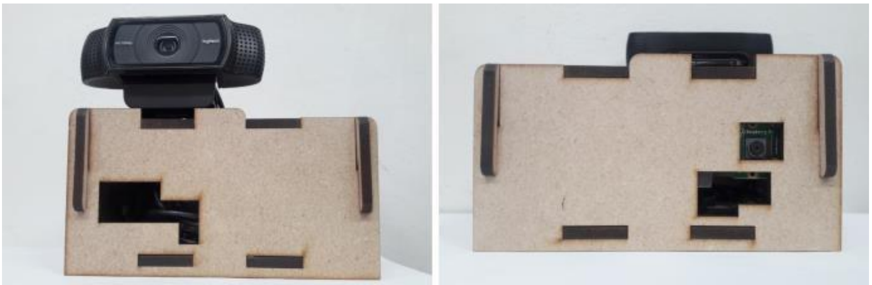
Fig. 11. Classification accuracy with face orientation feature only

Fig R2: Example images used for training the gaze estimator

ISSN- 2394-5125 VOL 06, ISSUE 05, 2019

Fig R3 : The result of IoT device detection:

Fig R4 : Prototype Implementation:



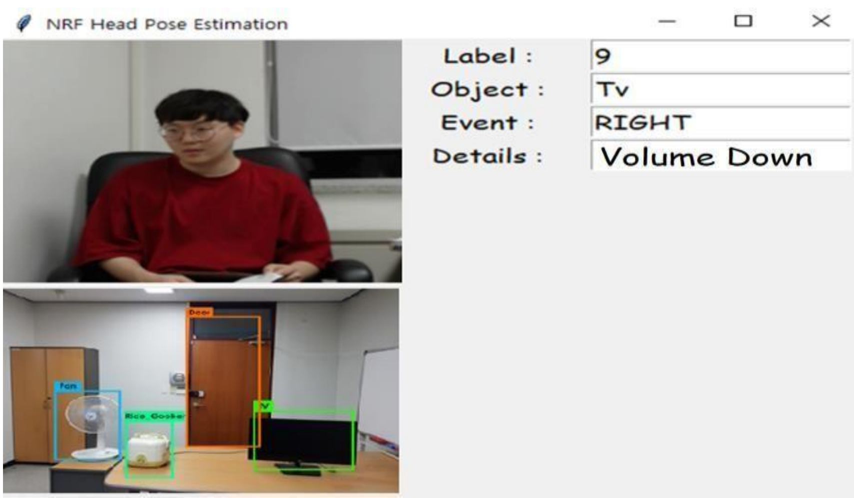


Fig R5 : Implementation of dash board application

Fig R6 : Example of training samples :

CONCLUSIONS

Finally, the IoT controller module generates a complete IoT command based on the information received from the object detection module, the gaze estimation module, and the hand gesture recognition module. The target device is identified using the gaze position obtained from the gaze estimation module and a category of the device in that location, from the object detection module. TheIoT Commands are generated based on the category of the device and a recognized user gesture. For example, “Up” gesture with the device category “air conditioner” is translated into the command “Increase the temperature”.

The proposed method helps users easily control the IoT devices by gazing at a device and performing simple hand gestures. For the Watch phase, deep convolutional neural network based approaches were adopted to estimate user’s gaze position and to detect and recognize the IoT devices. For the Do phase, a simple module with a gesture sensor was implemented. The final IoT command to be transmitted to the IoT platform was composed based on the detected IoT devices and hand gestures.

ISSN- 2394-5125 VOL 06, ISSUE 05, 2019

REFERENCES

[1] O.B. Sezer, E. Dogdu, and A. M. Ozbayoglu, "Context-aware computing, learning, and big data in internet of things: a survey," IEEE Internet Things J., vol. 5, no. 1, pp. 1-27, Nov. 2018

[2] B.J. Borges Neto, H.T. Silva, M. R. Assuncao, A. R. Mini, and A. A. Loureiro, "Sensing in the collaborative internet of things," IEEE sensors J., vol. 15, no. 3, pp. 6607-6632, Mar. 2015

[3] C. Perara, M. Compton, P. Christen, D. Georgakopoulos, and A. Zaslavsky, "Semantic-driven configuration of internet of things middleware," in Proc.SKG, Beijing, China, 2013, pp. 66-73

[4] R. Mahmoud, T. Yousuf, F. Aloul, and I. Zualkernan, " Internet of things (IoT) security: current status, challenges and prospective measures, " in Proc. ICITST, London, UK, 2015, PP. 336-341

[5] Y. Zou, J. Xiao, J. Han, K. Wu, Y. Li, and L. m. Ni, "GRfid: a device- free rfid –based gesture recognition system, " IEEE Trans. Mobile Comput., vol. 16, no.2,pp. 381-393, March 2017.

[6] K.Wu, H. Li, W. Lu, Y. Yi, Y. Liu, q. Zhang, and L, M , Ni, 'Hjam: attachment transmission in wlans, "in Proc. IEEE INFOCOM, Orlando,
FL, USA,2012, pp.1449-1457

[7] Y. Zou, W. Liu, K. Wu, and L. M. Ni, "Wi-fi radar: recognizing human behaviour with commodity wi-fi, " IEEE Commun. Mag., vol. 55, no. 10, pp. 105-111, Oct. 2017.

[8] K. Wu, J. Xiao, Y. Yi, D.Chen, X. Luo, and L. M. Ni, " Csi-based indoor localization, " IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 7, pp. 1300-1309, Jul. 2013.

[9] W. Kluge, F. Poegel, H. Roller, M. Lange, T. Ferchland, L. Dathe, and D. Eggert, " A fully integrated 2.4-GHz IEEE 802.15.4-complaint transceiver for Zigbee applications, " IEEE J. Solid-State Circuits, vol. 41, no. 12, pp.2767-2775, Dec. 2006.

[10] H.B. Pandya and T.A. Champaneria, "Internet of Things: survey and case studies," in Proc. EESCO, Visakhapatnam, India, 2015, pp. 1-6.

[11] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, 'Internet of things: a survey on enabling technologies, protocols,

and applications, " IEEE Commun. Surv. Tuts., vol 17, no.4, pp. 2347- 2376, Jun 2015.

[12]C. Bormann, A. P. Castellani, and Z. Shelby, " CoAP: an application protocol for billions of tiny internet nodes." IEEE Internet Comput.,vol. 16, no. 2, pp. 62-67, Mar.2012.

[13] OneM2M Partners (2018). OneM2M [Online]. Available: http://onem2m.org

[14] Open Connectivity Foundation. (2018). OCF [Online]. Available http://openconnectivity.org/

[15] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF), " Comput. Vis. Image Underst., vol.110, no. 3, pp. 346- 359, Jun 2008.

[16] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vis., Vol. 60, no. 2, pp. 91-110, Nov. 2004.

[17] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proc. IEEE, vol. 86, no.11, pp. 2278-2324, Nov. 1998.

[18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, " You only look once: unified,

ISSN- 2394-5125 VOL 06, ISSUE 05, 2019

real time object detection.” in Proc. CVPR, LasVegas, NV, USA, 2016, pp.779-788.

[19] M. Vidal, A. Bulling, and H. Gellersen, “Pursuits: spontaneous interaction with displays based on smooth pursuit eye movement and moving targets.” in Proc. UbiComp, Zurich, Switzerland, 2013,pp. 439- 448.

[20] E. Velloso, M. Wirth, C. Weichel, A. Esteves, and H. Gellersen, “Ambigaze; direct control of ambient devices by gaze, “ in Proc. DIS, Brisbane, QLD, Australia, 2016, pp. 812-817.

[21] S. Mayer and G. Soros, “User interface beaming-seamless interaction with smart things using personal wearable computers, “ in Proc. BSN, Zurich, Switzerland, 2-14, pp. 46-49

[22] M. Khamis, A. Hoesl, A. Klimczak, M. Reiss, F.Alt, and A. Building, “Eyescout: active eye tracking for position and movement independent gaze interaction with large public displays, “ in Proc.UIST, Quebec City, QC, Canada, 2017, pp, 155-166.

[23] M. Budde, M. Berning, C. Baumgartner, F. Kinn, T. Kopf, S. Ochs, F. Reiche, T, Riedel, and M. Beigl, “Point & Control- interaction in smart environments, “in Proc. UbiComp, Zurich, Switzerland, 2013, pp. 303- 306.

[24] G. Alce, A. Espinoza, T. Hartzell, S. Olsson, D. Samuelson, and M. Wallergard, “Ubicompass: an iot interaction concept,” Adv. Human-Computer Interact., vol.2018, pp. 1-12, Apr.2018

[25] Z.Zhu and Q. Ji, Novel eye gaze tracking techniques under natural head movement, ”IEEE Trans. Biomed. Eng., vol.54, no. 12, pp. 2246- 2260

[26] OpenIoT Consortium.(2018) OpenIoT [Online]. Available http://www.openiot.eu

[27]J. J. Cerrolaza, A. Villanueva, M. Villanueva, and R. Cabeza, “Error characterization and compensation in eye tracking systems.” in Proc. ACM ETRA, Santa Barbara, CA, 2012, PP. 205-208.

[28] J. Chen and Q. Ji, “3D gaze estimation with a single camera without ir illumination,” in Proc. ICPR,Tampa, FL, USA, 2008, pp. 1-4.

[29] B. Ahn, D-G. Choi, J. Park, and I. S. Kweon, “ Real-time head pose estimation using multi-task deep neural network, “ Rob. Auton. Syst.,vol. 103, pp.2074-2083

[30] D. Verweij, A. Esteves, V-J Khan and S.Bakker, “Wavetrace: motion matching input using wrist-worn motion sensors,” in Proc. ACM CHI, Denver, CO, USA, 2016, pp. 2319-2331.