PROCEEDINGS

FPTU STUDENT RESEARCH CONFERENCE 2019

Can Tho, August 2019

QS STARS RATED FOR EXCELLENCE 2012

FPT Education

FPT UNIVERSITY

FPT UNIVERSITY

THE STUDENT RESEARCH CONFERENCE PROCEEDINGS

Can Tho, August 2019

CONTENTS

ABOUT THE CONFERENCE

The Research Festival is an annual activity of FPT University. Its goals were to improve research and presentation skills and to open opportunities for students to study, exchange and develop practical products. The Research Festival 2019 was held at FPT University, Can Tho campus. The preliminaries were organized in 25th July 2019 and the final round was on August 15th-16th, 2019.
In 2019, the Organizing Committee selected 9 groups to compete in the Final Round of IoT Showcase.

The Scientific Council include the following members:

1. Associate Professor **Mohd Fadzil Hassan** - Computer & Information Sciences Department, Universiti Teknologi PETRONAS, Malaysia.
2. Professor. **Hsiang Cheng Wang** - Vice Dean of Engineering College, National Chung Cheng University, Taiwan.
3. Dr. **Tran Nguyen Bao Tran**, Lecturer, Head of Mechatronics Department, Ly Tu Trong College.
4. Associate Professor. **Pham Hung Quy**, Office of Science Management and International Affairs, FPT University.
5. Dr. **Tran Duc Chung**, Computing Fundamental (CF), FPT University.

ABOUT THE PROCEEDINGS

Research Festival Proceedings is an annual publication of FPT University. The Proceedings gathers all research papers of students and lecturers, all of which had been defended in front of the scientific council.

At present, Students' Conference on Scientific Research Proceedings was published in English. This is for internal use only.

Readers can find it in the Library of FPT University in all four campuses, as well as online at http://ds.libol.fpt.edu.vn

Implementing CCTV-Based Attendance Taking Support System with Deep Face Recognition

Duong Xuan Hoa
ICT Department
FPT University
Hanoi, Vietnam
hoadxse04730@fpt.edu.vn

Le Van Thanh
ICT Department
FPT University
Hanoi, Vietnam
thanhlvse04819@fpt.edu.vn

Tran Quang Huy
ICT Department
FPT University
Hanoi, Vietnam
huytqse04731@fpt.edu.vn

Nguyen Tien Tuan Anh
ICT Department
FPT University
Hanoi, Vietnam
anhntt96@fpt.edu.vn

Le Dinh Duy
ICT Department
FPT University
Hanoi, Vietnam
duyldhe130655@fpt.edu.vn

Abstract

Face recognition (FR) has received considerable attention to the field of security, especially the use of closed-circuit television (CCTV) cameras in security monitoring. Although significant advances in the field of computer vision have been made, advanced face recognition systems provide satisfactory performance only in controlled conditions, and they deteriorate significantly in the face of Real-world scenarios such as lighting conditions, motion blur, camera resolution, etc. This article shows how we design, implement, and conduct the empirical comparisons of several machine learning open libraries in building attendance taking (AT) support systems using indoor security cameras called ATSS. Our trial system deployed to record the appearances of 120 students in 5 classes who study at the 3-th floor of FPT Polytechnic College building. Through measurement of technical indicators, we also calculate costs as well as benefits that the system brings in economic terms.

Keywords: Face recognition, CCTV, and Attendance Taking System.

I. Introduction

1.1. Problem and Motivation

Every day, the CCTV system operates to monitoring inside the building for security. The system's resources allow developers to build computer vision based applications to integrate with CCTV. FR is an excellent biometric technique for identity authentication. It is possible to apply FR technology for automatic attendance taking at schools. There are several benefits from attendance considering using the existing camera system such as: save time and effort, provide striking evidence for quality assurance and human resource management tasks, avoid intermediary of infectious diseases [1]. The existing attendance taking system using fingerprint recognition facing several challenging due to large intra-class variability and substantial inter-class similarity mentioned by Dyre and Sumathi [2]. Ngo et al. have combined the data from the academic portal with different FR techniques for the task of taking attendance in the classroom [1]. The result shows that their system works smoothly. However, the investment costs for procurement, camera installation at the classroom, and a large number of videos processing are expensive. This research describes the solution to apply deep FR technology to perform AT via the existing CCTV system, which takes advantage of available resources better and more suitable for different situations.

Figure 1. Real installation of the attendance taking system ATSS at 3-th floor, FPT Polytechnic School.

1.2. Related Works

Recently, deep learning techniques have made many significant achievements in FR, such as deep convolutional neural networks [3], use a cascade of multiple layers of processing units for feature extraction. They learn various levels of representations that correspond to different levels of abstraction. Those techniques called deep FR. The evolution of the FR is around network architectures and loss functions. Deep face model trained on the large dataset. We often lack resources to learn a complex model with minimal training samples for a specific face recognition task. Therefore using the pre-trained model as transfer learning usually applied [4]. Wang and Deng have reviewed many of model techniques since 2014 [5]. Most of the review models

provide more than 98% to almost 100% of accuracy on the tested datasets. Ranjan et al. summarized FR's component consist of three modules are usually needed for an automatic face recognition system [6]: face detector, landmarks extraction, and feature description as shown in figure 2.

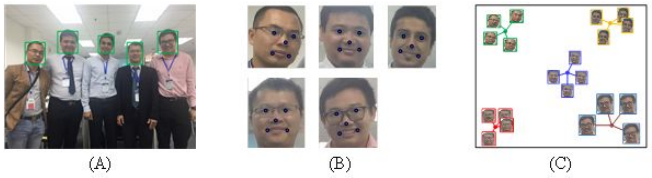


Figure 2. Three main components of Deep Face System. (A) demonstrates the face detection from an image frame, (B) shows the landmarks points of the cropped faces and (C) plot featured data in feature space archived by feature descriptor.

1.3. Problems of Face Recognition in Attendance Taking System using CCTV

The picture seems simple when we think that we only need to use FR to determine if students are present. However, when we go into our investigation, we have faced some problems:

Required almost 100% accuracy: Attendance usually affects students directly. Many schools also require attendance as part of the assessment process. See an example of the course syllabus for EBIO 6300 of the University of Colorado in semester fall-2013 [16]. At FPT polytechnic, attendance required minimum is 80% (over 30 slots of studying). There are several strategies to solve the problem, not only special technical efforts such as additional policy, system support…etc.

Constrained of the environment: installed equipment mainly used for security purposes instead of attendance taking [17]. The cameras hung at the intersection in the corridor such as elevator hall, corridor corner. The AT must not generate any affects to existing CCTV system.

Performance of the existing methods in a real environment: Even if the accuracy of ArcFace [13], the highest archive algorithm mentioned in [5] is up to 99.83% on MS-Celeb-1M test set. Algorithms almost work well in an ideal environment, which may not be satisfied in the real settings due to the effect of motion, camera resolution [18], light conditions. The attendance taking task may not require to response in runtime; however, the delay should be as short as possible. Most of the libraries asked for high processing time.

1.4. Contribution of this Paper

The main contribution of this paper is to develop a complete algorithmic process that, at each step, has been studied and evaluated to find the appropriate processing method for automatic attendance system using CCTV problems. The system consists of 4 major parts: the Job Master, Job Workers, a central database, and user interface applications. The Job Master plays the role of a navigator, and controls AI processing units, which relates to the issues of performance of the system are focused parts of this research.

The existing solution cannot solve mention problems with AT using CCTV without adding some calibrations. In this paper, we describe the way to provide experiments, combine techniques, and to prepare some environment settings to tune our attendance taking system. Performance indicators, along with business requirements, are considered to indicate detailed recommendations on investment costs as well as the benefits for particular system scale.

Several surveys summarize, and literature reviews the FR libraries ([5], [6], and [19]). In this study, we provide an empirical comparison on several latest FR libs as well as classification algorithms through our dataset in our real environment project. Ngo et al. also conducted a review on libraries, but the reviewed facial embedding libs was limited [1].

The remainder of the paper organized as follows: section 2 describes the architecture design of the system. The details of how a single module is tuned analyzed in section 3. The tested results are shown in Section 4, and brief conclusions are finally discussed in section 5.

II. SYSTEM DESIGN

The attendance system, called ATSS connects to the CCTV system. ATSS operates dependently on the CCTV system but does not leave any effect on the existing system. Figure 3 depicts the general picture of the system. There are several components in the system: the media recorder, job master, job workers, central database, and user interface applications. The media recorder plays the role to record frames captured from CCTV system for further processes. The job master constructs schedule and arranges tasks to job workers by using data leveraged from the academic portal. The processed data then stored in the central database and accessed by the user for reporting and manipulating data via the web application. The attendance data also submitted to academic portal. System admin can configure all parameters.

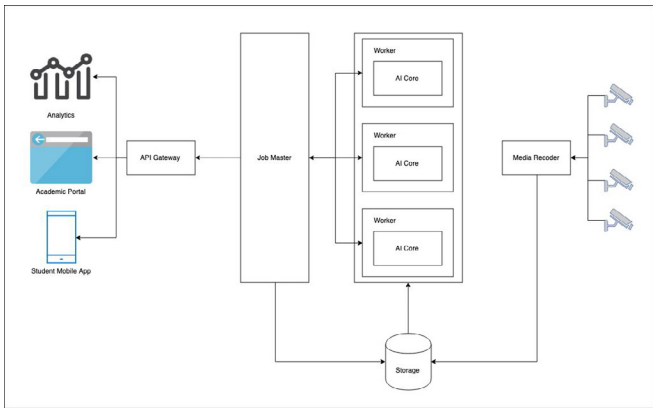


Figure 3: ATSS system architecture.

2.1. Job Master

As described in the previous part of this section, the job master is responsible as the navigator of the data streams

and process — the system run according to the schedule data. The corresponding FR model is also loaded based on the list of students instead of constructing a significant model for identifying the whole students. The exchange data structure is between ATSS, and the academic portal shows in UML in figure 4.

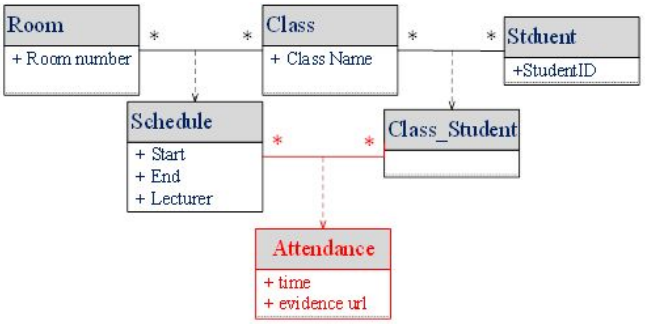


Figure 4: UML represents the data structure of exchanged data between the ATSS and the academic portal. Classes and relationships are described in black to represent the data that ATSS receives from the academic portal. The classes and relationships highlighted in red represent the attendance data structure that FAS will send to the Academic port. Actual data is formatted in Json.

According to the performance of the models reviewed by [1] and [5], due to our limited computation resources that let deep FR's performance cannot satisfied to perform run-time response. We group archived frames and corresponding schedule into several jobs and push to a queue. Each of them then is dequeued to process by the job workers. Figure 5 demonstrates the process of dividing the archived frames into jobs.

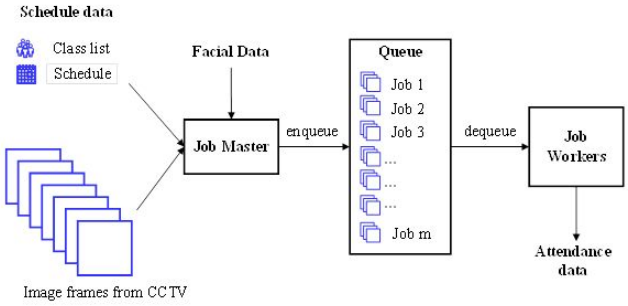


Figure 5: Archived image frames from CCTV are divided into jobs and stores in the queue for next process.

2.2. Data Collection Application

To collect training dataset, we have built a Mobile application to record the user's face video. The user uses the front camera of the phone to record different angles of the face in the lowest resolution of 720p, 30 fps. The average length of the trained videos is about 30 seconds. Videos recorded by data collector then uploaded to the database server for further processing via a web service.

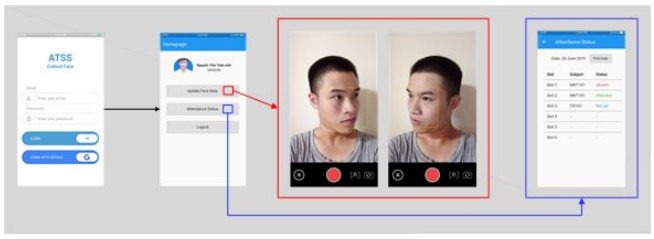


Figure 6: screen captures of mobile data collector application. Students / Academic staff use mobile application to records the video-containing-faces of the student. Video then uploaded to server for feature processing.

2.3. Job workers

The job workers perform the simple task of an FR problem. All descriptions of FR building block are available in this module. The video frames are processed directly in this module. All Processes are in parallel through the arrangement of the job master. They are constructed based on master-slave architecture [20]. The details of the job workers shown in Figure 7. There are two tasks handled by the workers:

Face identification: Photos of batches taken from the queue passed to the identification module. The system does not have to process the whole pixels of an image. The region of interest cropped for the subsequent processing. Face and landmark detections are executed to retrieve bounding boxes of the faces as well as facial landmarks points. All cropped faces passed to vectorization for the classification task. The summarization algorithm performed to make the final decision.

Data collection: Features are extracted from video faces collected by the mobile application, and then processing in frame processing component and stored as feature vectors.

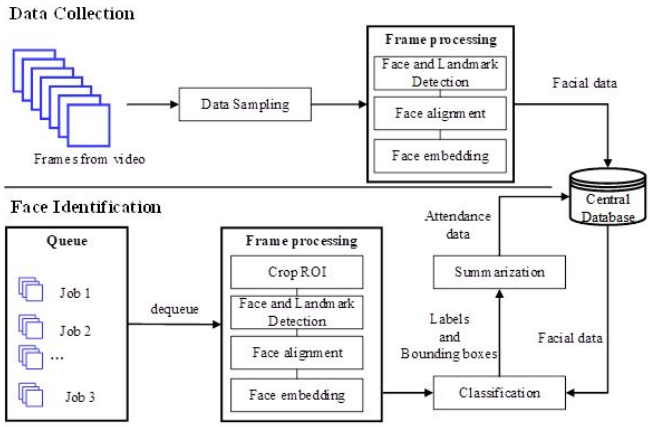


Figure 7: Components of an Ai processing unit.

There are several techniques for face data augmentation reviewed by Wang et al. [21], such as face rotation, transformation… etc. However, we applied a compelling face embedding technique for feature description and based on our observation during the system development. We notice that face alignment brings efficient.

III. DETAILED DESIGN

In this section, we describe in detail how we configure for each system component. The description mainly lies around the modules that generate the most effective, including data sampling, batch processing, the defining region of interest (ROI), frame processing, and summarization algorithm.

3.1. Data Sampling

We do not take all retrieved frames from uploaded video by the data collector application to build training dataset. Instead of doing so, on each video, we perform head pose detection, which using HopeNet[28], to detecting exactly 3 face turning angles. As our experiential, the best value of 3 face turning angles are -0.3 rad, 0 rad and 0.3 rad. After that, we perform face detection to extract 3 faces each angle. Totally, we have 9 faces evenly spread from the input video that is the number of training samples for each class. Feature vectors then extracted from the cropped faces to be stored in the facial database.

3.2. Region of Interest

We found that FR only works well within a certain distance; the problem may be related to the camera resolution. So we calculated the area to take attendance, which is a rectangle called attendance area and asked for a rule to guide students to go through this area to record their presence.

Denote: $M_i \in R^{n\times d} \ \forall\ i = 1..7$ are the training set of 7 members in project team. We calculate the center μ_i of M_i:

$$\mu_i = \left[\frac{1}{n}\sum_{j=1}^{n} M_{i_{j,1}} \ \cdots \ \frac{1}{n}\sum_{j=1}^{n} M_{i_{j,d}} \right] \ \forall\ i = 1..7$$

$s_j^{(i)}$ is the feature vector capture member i^{th} when he stands j meters away from the camera, his face turn directly to camera. We compute the distance $d_i\left(s_j^{(i)}, \mu_i\right) = \sqrt{\sum_{z=1}^{d}\left(s_{j\ z}^{(i)} - \mu_{iz}\right)^2}$ from $s_j^{(i)}$ to μ_i. Set $g_j = \frac{1}{m}\sum_{i=1}^{3} d_i(s_j^{(i)}, \mu_i)$ is the mean of particular distance to camera of 3 members. Figure 8 illustrates the best area for FR.

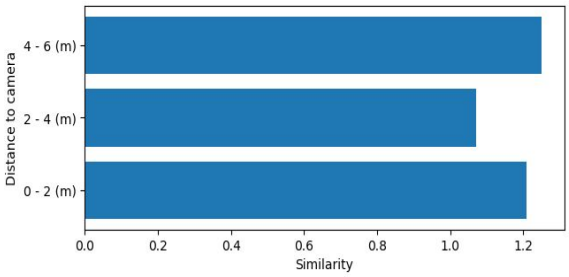


Figure 8: illustrates the similarity between the test faces and the mean of training faces (Euclidean distance, the closer is better) corresponding to a particular range to camera.

The camera is hanging at the height of 2.8m. The distance from the face to the camera is about 2 to 4 m, which is the ideal distance for FR. In practice, we choose 2.5m to 3.5m. To speed up the processing, we set up so that three people can stand in horizontal line and still be able to check. We set the height of the attendance area is 100 cm (equivalent to 3-floor tiles). We then cropped the image to fit the size of 5 humans step in the attendance area. The area called the region of interest ROI [22].

3.3. Frame Processing

The system only considers faces that appear in the ROI. Ngo et al. had a review of Face detection libraries. According to the results obtained from them, we use MTCNN [7] as a face and landmark detector. The faces are aligned based on five returned landmark points (left eye, right eye, nose, left corner mouth, and right corner mouth). We have tested some different FR libraries and test their output feature vectors with several machine learning algorithms include parametric and non-parametric, generative, and discriminative [23].

3.4. Responding time

At FPT polytechnic, if the student is late for the first 15 minutes, or leave early before the last 15 minutes, he/she will be counted as absent. Attendance usually takes place at this time. To increase flexibility in attendance as well as to avoid affecting the class, we define the AT process consist of 2 phases: check-in and check-out for this process. The check-in time starts 15 minutes earlier than the lesson and ends after the first 15 minutes. Likewise for check-out at the end of class. Students present in front of the attendance taking area at the right time considered as present. The AT may not strictly require runtime responding. However, it should get a response as fast as possible. In this situation, we set 2 minutes as the size of a job. We tried many different sizes; however, 2 minutes is the best-observed value due to the aspects of io waiting time, processing time, and memory consumption.

3.4. Summarization algorithm

Using pure FR to classify input faces generates a mess, due to motion blur, the deflection angle of the faces to the camera. Each human goes through the ROI, creating dozens of frames. Some of the samples are classified correctly, but some others were not. Our idea is to tracks every face in ROI as object tracking [24], the class assigned to a particular sample most times will be the final decision. However, this number of identifiers must be more than one threshold, which we have set is ten frames. The Summarize algorithm consists 2 main steps as follow:

Step 1- connected bounding box detection: Each bounding box have 4 attributes: $x_{min}, x_{max}, y_{min}, y_{max}$. There are any 2 of bounding boxes considered connected if they are in at most two frames away from each other and their ratio of intersection is greater than $p = 0.4$ as we found in the experiment, or they connected with a common bounding box. A connected component of bounding boxes contain all

bounding boxes pairwise connected. They are the movement of a face in a video.

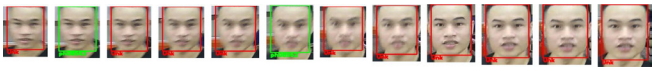


Figure 9: Example of a connected bounding boxes.

Step 2 – final decision: For each connected component, we use a sliding window algorithm with window size is $s = 10$ and slide to the right to determine the range for labeling. For each window, we label by using the time that label appear t_{appear} and the predicted ratio $\frac{t_{appear}}{s}$ is used to determine whether that label can label all the boxes in that window or not. As in the experiment, the best value for this ratio is 0.6. If no label satisfies the condition, all boxes in that window will be labeled "Unknown".

Figure 10: Example of a sliding window.

As shown in Figure 10, all wrong labels are labeled again, and in the next window, we use the old (not updated) label and continue labeling. This function will reduce the rate of the wrong detection in our system.

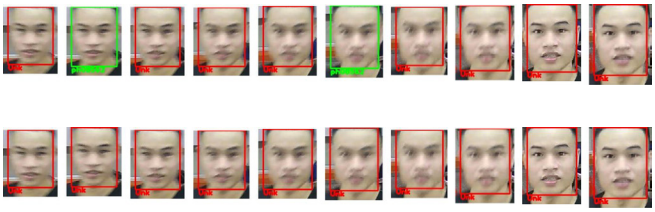


Figure 9: Example of the output of the summarization algorithm, (A) illustrates the pure classification results, (B) The most frequent label (class) assigned to the whole boxes appear in the window..

IV. Experiment and Result

4.1. Experiment

We used the pre-trained models to build the system. Therefore, in order to carry out the evaluation of appropriate models, we use the facial data of 120 students as training data (as described in sections 2.2 and 3.1, there are 9 facial images for each student). Test data were recorded in 2 sessions, with 7490 labeled images in attendant time. We manually labeled each face that appears in the ROI in videos. Figure 11 shows the frequencies of appearance of each label in the test set. All experiments are conducted using system of Intel Core i5 3470 3.20GHz, GPU NVidia GTX 1050ti 4GB, RAM 8GB.

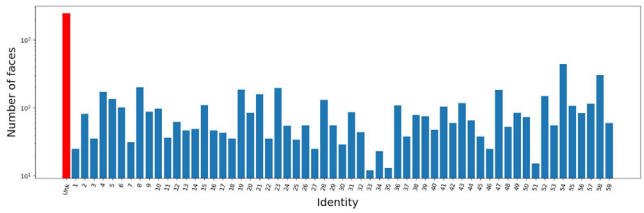
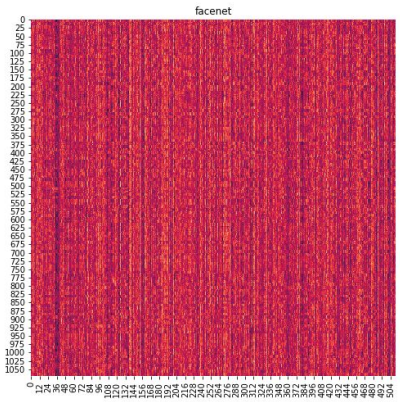


Figure 11: the frequencies of appearance of 60 labels (including unknown) in the test set. The first chart represents the number of unknown samples in the test set.

It is easy to see that the number of samples of "unknown" class is very prominent. The reason for this is because we used the camera facing the staircase in the building for installing our system. Many students who were walking to their classroom at different floors go through this area.
We used MTCNN as a facial detector [7], according to Ngo et al [1] and our observation, the results of detected faces are complete. The extracted faces then automatically be aligned vertically so that the next parts can operate smoothly. There are many famous and latest methods be selected as candidates for feature descriptor. In this research we have tested our system with FaceNet introduced by Schroff et al [12] that dispensing with the alignment step and simply applying end-to-end training for an embedding space. Vectors in this space represent faces and the Euclidean distance between vectors represent how different the faces are. Another method is Arcface by introduced by Jiankang Deng et al [13]. They use a new loss function to learn highly discriminative features for robust face recognition. Feature data then passed to train the classifier.

4.1. System Accuracy

Figures 12 and 13 illustrate the face embedding data for FaceNet and ArcFace respectively. Part A of each figure shows the heat map of the data on the feature space. Meanwhile, part B visualizes featured data in 2-dimensional space using technique t-SNE [25]. Each face is described by FaceNet with a 512-dimensional vector meanwhile ArcFace uses a 512-dimensional vector to embed the same faces. We can observe that extracted facial data by both descriptors visualized by t-SNE is linearly separable, which is fit to the classification task.



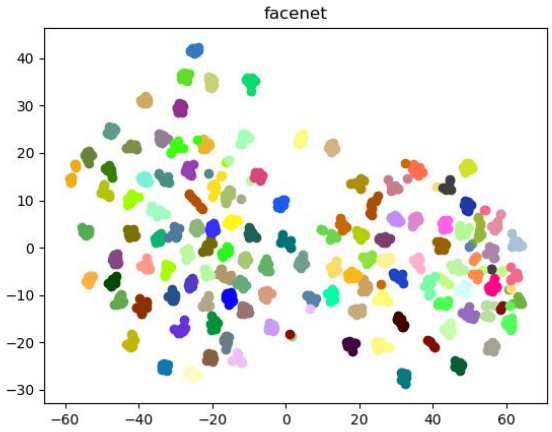
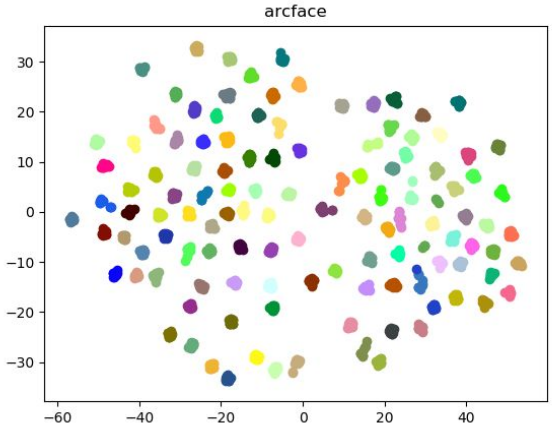


Figure 12: (A) the heat map to represent training set in feature space, learned by Arcface, (B) The extracted vectors by FaceNet visualized in 2 dimensions space by t-SNE.



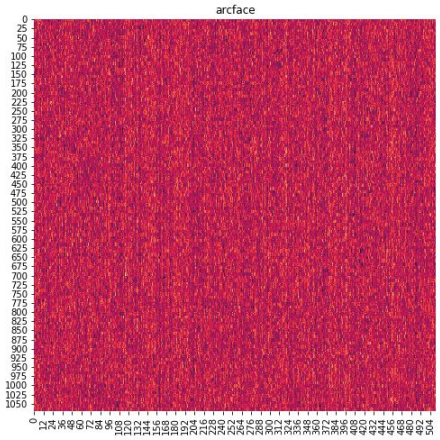


Figure 13: (A) the heat map to represent training set in feature space, learned by Arcface, (B) The extracted vectors by Arcface visualized in 2 dimensions space by t-SNE.

We build 4 classifiers: SVM Linear Kernel (Linear SVM) [23], SVM RBF Kernel (RBF-SVM) [26], Gaussian Naive Bayes (NB) [23] and Weighted- KNN [27]. In this report, we only present the best results we obtained for each classifier. Table 2 shows the number of errors for each combination. In order to measure the performance of each method, we build a confusion matrix for each pair, denoted byM(f,c)R500×500 ∀ f∈FaceNet, ArcFace, c∈RBF-SVM,NB,KNN,Linear-SVM. We then calculate total error E(f,c) corresponding to each M(f,c) by the sum of all cells not belonging to the diagonal of the matrix, such that: E(f,c)= i=1500j=1500Mi,j(f,c)*(i≠j)

Table 1. The number of errors in the test set corresponds to the combination of the methods.

Classifiers Descriptor	**Linear SVM**	**RBF SVM**	**WKNN**	**NB**	**Mean**
FaceNet	0.32	0.33	0.29	0.32	**0.31**
ArcFace	0.886	0.803	0.913	0.75	0.83

Through the numerical results obtained from Table 1, we can see the best results obtained from the combination of Feature extractor ArcFace and the classifier KNN. In Figure 14 we shows the fusion matrix of the combination.

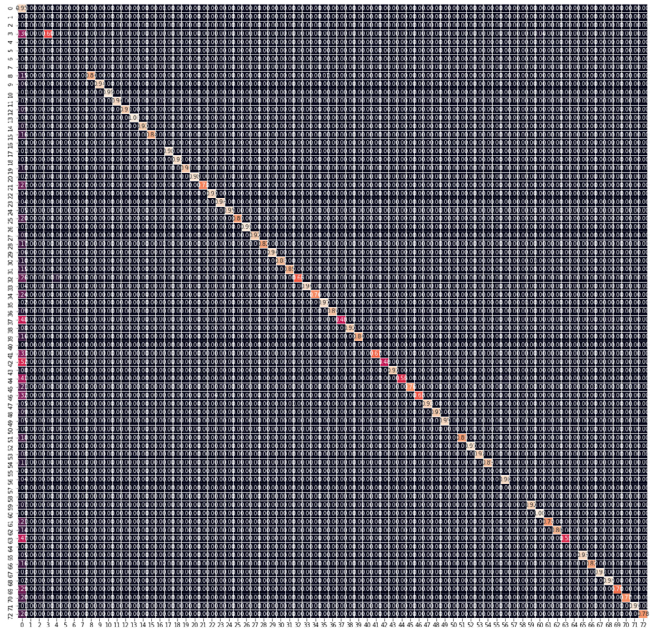


Figure 14: Confusion matrix archived from the combination of Arcface (pre-trained model LResNet100E-IR, ArcFace@ms1m-refine-v2) and KNN (K=5 and confident distance threshold = 1.07). The first column and row represent the label of "unknown", others columns and rows show the result of corresponding class. Displayed data is normalized due to the large number of "unknown" samples.

We can see that when using Arcface as feature extraction the result was really good compared to some algorithms using similarity based on distance like KNN or Linear SVM and Generative classifier provided poor results. This is predictable because we don't have definition about unknown class, so the result when predict P(y | x) is low.

There are many samples in the scope of research but misclassified into the "unknown" class. This is significantly improved when using the summarization process. Figure 15 illustrates the fusion matrix after executing summarization. Our setting archives accuracy of **91.3%** on the test dataset, which was measured before applying our summarization algorithm. As explained before (Section 3), the results can be improved by using our summarization algorithm. In Figure 15, we show the confusion matrix of the system and the accuracy is **92.7%**. Although the accuracy only increase 1.4%, the ratio of False Positive has decreased significantly to 1% as show in figure 16.

If we remove unknown sample from our data, the accuracy can be up to 98.5%. Our results almost reach the results of Arcface even we don't have standard conditions and we get many difficulties as listed before in school environment.

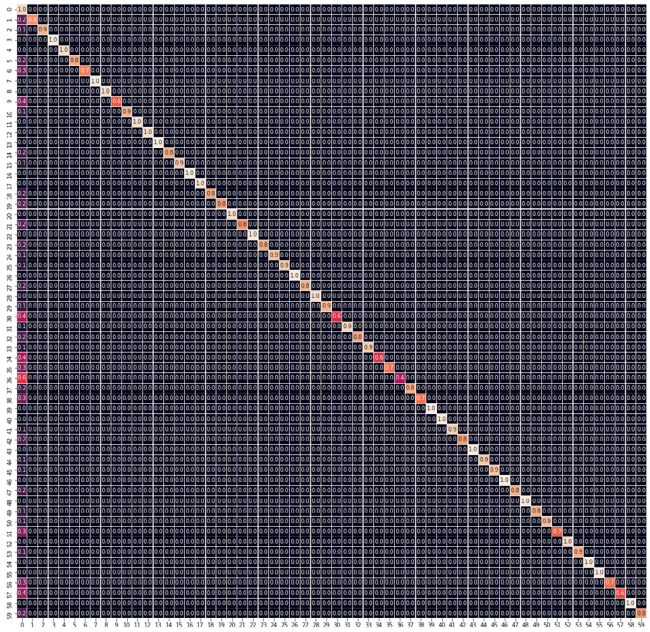


Figure 15: The obtained Confusion matrix after executing summarization process. The first column and row represent the label of "unknown", others columns and rows show the result of corresponding class. Displayed data is normalized due to the large number of "unknown" samples.

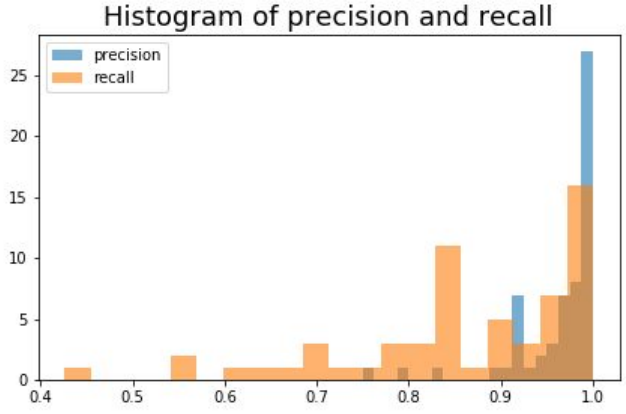


Figure 16: Histogram of precision after using summarization

4.2. System processing time

The processing time of the system is not suitable for the purpose of runtime responding. While detection and alignment are not too affected by the number of faces on the same frame, this greatly affects the feature extractions. This lets the system performance to be decreased linearly. The predicted speeds of the classification models are very fast, so we do not present in this section. Figure 17 shows the statistical numbers for each processing stage.

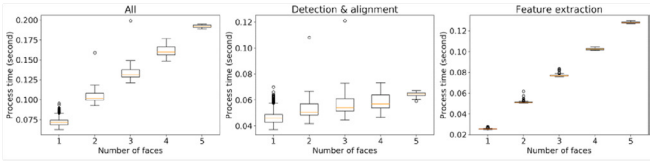


Figure 17: (A) shows the processing time of the FR processing, (B) illustrate processing time of face detection (including landmarks points) and face alignment, (C) visualizes the processing time of face description, corresponding to number of faces in the same frame.

Computers that we use to operate the system can only run up to 4 job workers (4 processes) at the same time. Figure 18 shows the ability to utilize system resources to parallel executing multiple jobs at the same time.

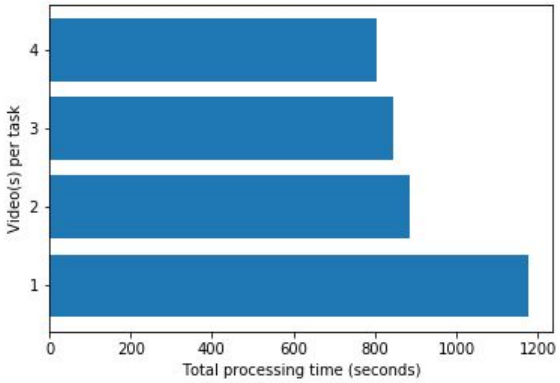


Figure 18: (A) shows the processing time of the FR processing, (B) illustrate processing time of face detection (including landmarks points) and face alignment, (C) visualizes the processing time of face description, corresponding to number of faces in the same frame.

As mentioned in section 3.4, each attendance session lasts 25 minutes, we choose each job length to be 2 minutes. So if there are 4 cameras, it needs to be processed in ~ 802 seconds, so before check-out, students can receive attendance records at the time of their check-in. The system depends on the number of cameras connected.

V. CONCLUSIONS

In this paper, our immediate goal is to build a system, which can support lectures in attendance taking in FPT Polytechnic. We showed the result of our real experiment

and the feasibility of attendance taking support system in university environment using CCTV Camera.

In our environment, our system had to work with 200 students in 5 classes. By experimentation, we found out that it's impossible to achieve state-of-the-art accuracy in university environment, but we had improved accuracy using our algorithm using movement of the face to remove some wrong results. It seemed promising and showed that we still can increase accuracy of our system.

In this thesis, we propose a full system solution powered by state-of-the-art facial recognition model, from hardware, to procedures for handling many streaming videos with unknown faces recorded by CCTV. By taking the advantage of multi process and job scheduling, we also leverage the hardware efficiency of face recognition to minimize system cost but still meet the required response time..

Since attendance taking system required high precision, our system is still struggling with low recall, this will be a huge disadvantage if the test environments are not optimal. Concerning the results of our single classifiers, we can also expect to improve recall by not using a single classifier technique, but a combination of parametric (SVM, DNN) and non-parametric classifier (KNN) to better define "Unknown" class, which can increase the recall performance. Our system is also depended on preprocessing . In the future, we can use some other models to improve our results. Currently, our system's target can handle up to 500 students. However, because of some difficulty that we have mentioned above, we have evaluated our system only on 120 students. In the near future, we will increase the size of evaluation dataset to verify the comprehensiveness of our system.

References

[1] Ngo Tung Son, Le Phuong Chi, Phan Truong Lam, and Tran Van Dinh. 2019. Combination of Facial Recognition and Interaction with Academic Portal in Automatic Attendance System. In Proceedings of the 2019 8th International Conference on Software and Computer Applications (ICSCA '19). ACM, New York, NY, USA, 299-305. DOI: https://doi.org/10.1145/3316615.3316638.

[2] Dyre, Shoba, & Sumathi, C.P. (2016). A SURVEY ON VARIOUS APPROACHES TO FINGERPRINT MATCHING FOR PERSONAL VERIFICATION AND IDENTIFICATION. http://doi.org/10.5281/zenodo.232968

[3] Stephen Balaban .Deep learning and face recognition: the state of the art .[J] arXiv preprint arXiv:1902.03524

[4] Yu, H., Luo, Z., & Tang, Y. (2016). Transfer Learning for Face Identification with Deep Face Model. 2016 7th International Conference on Cloud Computing and Big Data (CCBD). doi:10.1109/ccbd.2016.014

[5] M. Wang, W. Deng, Deep face recognition: a survey, arXiv arXiv:1804.06655v8 (2019).

[6] R. Ranjan, S. Sankaranarayanan, A. Bansal, N. Bodla, J. C. Chen, V. M. Patel, C. D. Castillo, and R. Chellappa. Deep learning for understanding faces: Machines may be just as good, or better, than humans. IEEE Signal Processing Magazine, 35(1):66–83, 2018.

[7] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. IEEE Signal Processing Letters, 23(10), 1499–1503.doi:10.1109/lsp.2016.2603342.

[8] Hao Wang, Zhifeng Li, Xing Ji, Yitong Wang .Face R-CNN .arXiv preprint arXiv:1706.01061.

[9] Sachin Sudhakar Farfade, Mohammad Saberian, Li-Jia Li .Multi-view Face Detection Using Deep Convolutional Neural Networks. arXiv preprint arXiv:1502.02766.

[10] Feng, Z.-H., Kittler, J., Awais, M., Huber, P., & Wu, X.-J. (2017). Face Detection, Bounding Box Aggregation and Pose Estimation for Robust Facial Landmark Localisation in the Wild. 2017 IEEE Conference on Computer Vision and pattern Recognition Workshops, (CVPRW).doi:10.1109/cvprw.2017.262.

[11] C. Sagonas, E. Antonakos, G, Tzimiropoulos, S. Zafeiriou, M. Pantic. 300 faces In-the-wild challenge: Database and results. Image and Vision Computing (IMAVIS), Special Issue on Facial Landmark Localisation "In-The-Wild". 2016

[12] Florian Schrof, Dmitry Kalenichenko and James Philbin "FaceNet: A Unified Embedding for Face Recognition and Clustering" arXiv: 1503.03832V3 [cs.CV], 17 June 2015.

[13] J. Deng, J. Guo, and S. Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. arXiv preprint arXiv:1801.07698, 2018.

[14] Jiawei Han, Micheline Kamber, and Jian Pei. 2011. Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[15] Feng, Z.-H., Kittler, J., Awais, M., Huber, P., & Wu, X.-J. (2017). Face Detection, Bounding Box Aggregation and Pose Estimation for Robust Facial Landmark Localisation in the Wild. 2017 IEEE Conference on Computer Vision and pattern Recognition Workshops, (CVPRW).doi:10.1109/cvprw.2017.262.

[16] Course Syllabus EBIO 6300: Phylogenetic Comparative Methods - Fall 2013: online: https://www.colorado.edu/smithlab/sites/default/files/attached-files/EBIO6300_PCMsSyllabus.pdf. Last visited 1 June, 2019.

[17] CCTV Technology Handbook, System Assessment and Validation for Emergency Responders (SAVER), July 2013.

[18] Zhou, Y., Liu, D., & Huang, T. (2018). Survey of Face Detection on Low-Quality Images. 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). doi:10.1109/fg.2018.00121.

[19] Masi, I., Wu, Y., Hassner, T., & Natarajan, P. (2018). Deep Face Recognition: A Survey. 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI).doi:10.1109/sibgrapi.2018.00067

[20] Li Jianglin, Zhang Daojie, & Zhang Siyuan. (2016). The application and research on master-slave station distributed integration business architecture. 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). doi:10.1109/imcec.2016.7867396

[21] Xiang Wang, Kai Wang, Shiguo Lian .A Survey on Face Data Augmentation .arXiv preprint arXiv:1904.11685.

[22] Ron Brinkmann (1999). The Art and Science of Digital Compositing. Morgan Kaufmann. p. 184. ISBN 978-0-12-133960-9.

[23] book: Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) 1st Edition.

[24] Alper Yilmaz, Omar Javed, and Mubarak Shah. 2006. Object tracking: A survey. ACM Comput. Surv. 38, 4, Article 13 (December 2006). DOI=http://dx.doi.org/10.1145/1177352.1177355.

[25] van der Maaten, Laurens and Hinton, Geoffrey. "Visualizing Data using t-SNE." Journal of Machine Learning Research 9 (2008): 2579--2605.

[26] Wang J., Chen Q., Chen Y. (2004) RBF Kernel Based Support Vector Machine with Universal Approximation and Its Application. In: Yin FL., Wang J., Guo C. (eds) Advances in Neural Networks – ISNN 2004. ISNN 2004. Lecture Notes in Computer Science, vol 3173. Springer, Berlin, Heidelberg.

[27] Hechenbichler and Schliep, 2004 Hechenbichler, K., and Schliep, K.P. (2004). Weighted k-nearest-neighbor techniques and ordinal classification. Discussion paper 399, SFB 386, Ludwig-Maximilians University, Munich. http://www.stat.uni-muenchen.de/sfb386/papers/dsp/paper399.ps.

[28] N. Ruiz, E. Chong, J. M. Rehg, "Fine-grained head pose estimation without keypoints", Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, pp. 2074-2083, Jun. 2018.

Automated Medical Dispensing Machine

Nisha Angeline C V[(1)], Rahul Ganesh E[(2)], Siva Pratheep S[(3)], Nishanthan V[(4)]

Department of Information Technology,

Thiagarajar College of Engineering,

Madurai, India.

{nishaangeline[(1)],pranavchithiran[(2)],sivapratheep484[(3)],vnishanthan2[(4)]}@gmail.com

Abstract*— Automatic Medical Dispensing Machine(AMDM) is a machine that supply medicines to people as per the medication prescribed by the Doctor electronically. The Automatic Medical Dispenser is loaded with medicines and their unique QR code. Patient information is retrieved from the cloud. The Doctor can prescribe medicines as an E-Prescription which is sent as a QR code to the patient. The QR code scanner in the dispenser machine scans the QR code of the patient and scans through the medicines available in the machine. It finds the right match for the prescription and dispenses it. The bill amount for the medicine will be detected via E-Payment.

Keywords—Prescription, QR Code, Dispenser, Android App, Medicines

I. INTRODUCTION

Today's multi-speciality hospitals are buzzing with inpatients and outpatients. According to the WHO Statistics, millions are harmed due to improper medication and nearly 47600 people die every day due to over dosage. There is a growing need for automation in order to avoid human errors and save time in the health sector. Hence it is proposed to develop an automated medicine dispensing machine that avoids human errors while searching for the tablets thus providing prescribed drugs to the concerned patient without any time delay.

II. CHALLENGES FACED DURING DRUG DISTRIBUTION

A. Maintaining the Integrity of the Specifications

The following problems are identified to pose a huge challenge to the drug distribution industry.

- Scarcity of talented pharmacists:

About 85% of WHO Member States report to have less than 1 pharmaceutical personnel per 1000 population [1]. This makes it necessary to automate drug distribution.

- Errors in dispensing:

Even Skilled labor is prone to error. A study says that 0%-45% of errors are made during drug distribution [2]. The errors are majorly due to mislabeled storage containers, lack of attention and lethargy during drug distribution [3].

III. NEED FOR AUTOMATIC MEDICAL DISPENSING MACHINE

A. Objectives

The main objectives of this product are as follows

- Is to reduce in-correct medicine intake due to oversight of the pharmacist.
- Avoid Self-medication of the Patient.
- Avoid delay in distribution of medication.
- Avoid Drug Duplication using Patient History via the Mobile app.
- To prompt the patient especially Senior Citizens to take their medication at the right time via the Mobile App developed

B. Making use of Barcode and QR code

The use of bar code instead of the handwritten transcript has proved to be efficient in dispensing drugs, this can be proved by a study that states the errors have been dropped from 0.19% to 0.07% when a barcode is used [4]. The errors in transcription are higher in number when it comes to strength and dosage. 40 Danish communities came together to identify four major reasons for the errors that occur during the transcription stage are [5].

- Handwritten prescriptions.
- Similarities in packaging and naming.
- A lack of effective control of prescription label and medicine.
- Lack of attention caused by interruptions.

A QR is more efficient in data storage and in retrieving data compared to a barcode. The information for dispensing medicines can be encoded in a QR code and can be decrypted using a QR code reader. This will prevent the dispensing of a higher dosage of medicines than recommended and avoids misinterpretation of the transcript by the physician.

C. Mechanised Drug Distribution

Humans are susceptible to error. They are inefficient in a process involving drudgeries over a long period. This is due to being busy, time constraints, fatigue, lack of attentiveness and similarities in drugs [2]. It is impossible to overcome these disadvantages as they are part of our Human nature. But human interventions in such delicate positions can be eliminated. Machines, however, are much better at repeating the same work over and over again. Where humans get bored and distracted and accustomed to what is in front of them, machine thrives when able to repeat the same thing over and over again.

D. Tackling Medication Duplication

Therapeutic duplication is the practice of prescribing multiple medications for the same indication or purpose without a clear distinction of when one agent should be

administered over another. For example, prescribing both ibuprofen and acetaminophen for PRN(Pro Re Nata) mild pain, or prescribing both Zofran and Compazine for PRN nausea and vomiting [11]. It is possible that people can have multiple illnesses at the same time, the medications they take for the individual can affect the nature of the other. This can lead to side effects and can also reduce the efficiency of both drugs. So, it is preferred to use combinations of drugs that are mutually exclusive and thereby preventing interacting with each other. The more the number of physicians gets involved, the higher the probability of potentially inappropriate drug combinations [12]. This due to that one physician is unaware of the drug prescribed by the other. Digitalization of transcription of drugs can help to overcome as it may allow the physicians to know the drugs prescribed by the other.

IV. Existing Solutions

A. *Semi-Automated Dispensing Machine using Barcode.*

A Barcode-assisted system was developed for medical administration and dispensing systems that fully integrate the information flow among the ordering, dispensing and administrating the system [15]. The Objectives of this system were to manage physician's orders, retrieve historical information regarding the patient's prior medications, obtaining the patient's identity and channelizing the skilled labor for métier works. However, the system was semiautomated. The study hospital-employed nurses to scan the barcode on the wristbands of the patients or the badges of the patients. The nurses had to take note of the medications that were prescribed to the patient and they have to manually fetch it from the pharmacy. It also was not considered cost-efficient, as each nurse allocated for the prior are provided with a laptop with seamless wireless connectivity to fetch the medication data from the hospital server.

B. *Drug Data Transfer System*

At BWH(Brigham and Women's Hospital), the Physician enters the medication details into a Computerized Physician Order Entry Program [16]. Medication orders are sent electronically to the pharmacy information system, approved by pharmacists, and then executed throughout the hospital. Medicines that are to be commonly used are stocked in semiautomated medication-dispensing devices in the nursing units, whereas less commonly used medications are filled from the pharmacy, dispensed in unit doses, and stored in patient-specific drawers before administration. This system does not record the medications given to the patient and without this historical data of medication, this system is prone to prior mentioned Drug Duplication problem.

C. *Computerized Physician Order Entry(CPOE) for Neonatal Ward.*

The clinical information system of Sayan-HIS(Hospital Information System) includes functionalities for order entry [17]. When the physician's order is entered into the computer, the prescription system delivers the requested order for medications, lab tests, and imaging to their relevant target hospital sections at the appropriate time. The user interface of the prescription system remained intact during the study period. The system constrains the selection of drugs and their possible pharmaceutical forms (vial, ampoule, tablet, etc.) through drop-down lists and pre-constructed orders. A minimum and maximum dose reminder are also available. This alert is to some extent useful for adults but not for neonates whose dose may vary up to tenfold for the same drug because of the weight-based dose calculation.

V. Proposed Solution

A. *Modules in our Solution*

a. Automated Medical Dispenser Machine(AMDM)
b. DocHelp App
c. MediHelp App
d. AMDM-Admin App

B. *Automated Medical Dispenser Machine*

1) *Raspberry Pi*

It is a mini computer capable of doing activities like browsing the internet, playing high-definition video, making spreadsheets, word-processing, and playing games, cameras etc.

In this work, AMDM is used to scan the QR code and extract the Doctor's prescription from the Patient Mobile app. It is then used to dispense the prescribed dosage.

2) *ZBar Barcode Reader*

ZBar is an open-source software suite for reading bar codes from various sources, such as video streams, image files, and raw intensity sensors.This converts the image into string given by the Raspberry pi.

3) *WebCam*

The Web Cam provides a clear 5MP resolution image or 1080p HD video recording. Here it is used to scan the QR code and feed the image to the ZBar Barcode Reader.

C. *MediHelp, DocHelp,AMDM-Admin - Android App*

Today's society is said to belong to the information age where information is considered to be a valuable asset. Different technologies are made so as to ensure the proper handling of information. Mobile devices use such technologies to provide information access and storage.

The DocHelp App is used to give the patient information to the Doctor on scanning the Patients Unique QR Code instead of the manual way of going through paper medical reports for the patient's history. It also saves the time for the Doctor's need for conversation with the patient on his previous history. Using the app, the Doctor can prescribe medicines to the patient with references to the patients medical history. The prescription is generated as a encrypted QR code in MediHelp app.

The MediHelp app allows the patient to go through his own medical history and also his prescription

The AMDM admin app allows the admin to keep track of the stock of the medicine present in the AMDM.

1) Firebase

Firebase is a Backend-as-a-Service (BaaS) provider allows users to build applications for the web and mobile interfaces without the need for server-side programming languages. Firebase is now coined as "a next-generation app-development platform on the Google Cloud Platform (GCP.)".Once connected, applications connect to Firebase via what is known as a "WebSocket," instead of using generic HTTP. Firebase eradicates the need for applications to make HTTP calls in order to get (and sync) user data since everything is handled with the help of a single socket connection.

Our apps make use of Firebase as a BaaS for database and user authentication. It allows them to be secure and seamless.

2) ZXing

ZXing ("zebra crossing") is an open-source, multi-format 1D/2D barcode image processing library implemented in Java, with ports to other languages. This library is made use to generate and Read QR Code within the app.

3) Material Design

Material Design is a unified system that combines theory, resources, and tools for crafting digital experiences. Google's Material Design considers "material" as a homogeneous digital fabric in which the material responds according to user interaction. It provides certain design standards for developing applications across android, web and iOS devices. The Material Design is implemented the app to make the User Interface a lot easier to interact and to make the most of the User Experience.

D. Work Flow of the System

The work flow of the system is as follows

1. The Patient needs to create an account in the app.
2. The Doctor scans the Unique QR code given the patient.
3. The DocHelp App has the following options
    - View Profile
    - View Medical History
    - View Drug History
    - Prescribe Medicines
4. View Profile allows the Doctor to view general information about the patient.
5. View Medical History and View Drug History displays the conditions and the drug intake of the respective patient
6. Prescribe Medicine allows the doctor to prescribe medicine to the respective patient.
7. The MediHelp app has the following options
    - View QRCode
    - View Patient Details
    - View Prescribed Medicines
8. View Prescribed Medicines allows the patient to view the diagnosis and the drug prescription
9. The patient gets his prescription in the form of the QR-CODE from the physician and takes it to the AMDM.
10. ZBar reader is used to decoding the QR-code. The algorithm behind the Zbar reader is successfully initiated through a tiny little camera that processes the captured QR-code image through highly developed digital refinement solutions to penetrate or perhaps read the exact QR-code.
11. This information is displayed in the system monitor along with the bill of the medicine
12. The user is next prompted to pay the bill through online payment and the transaction is completed.
13. Once payment is successful the raspberry-pi send a signal to the servo motor. The servo motor then dispenses the appropriate medicine from the rack.

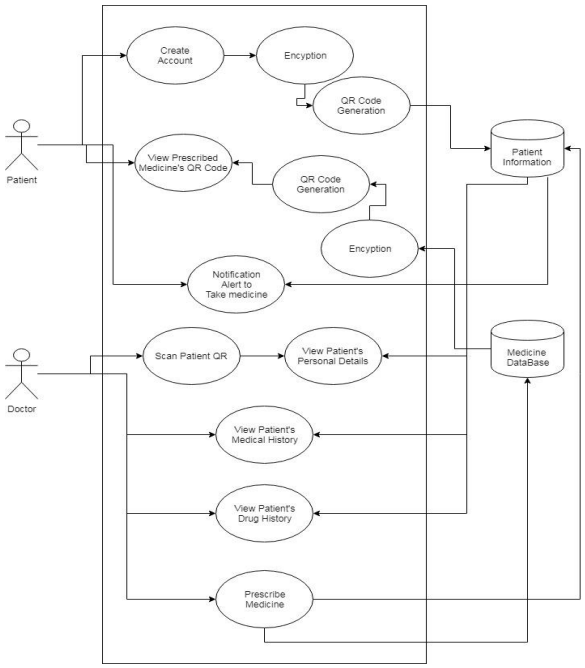


Figure 1 - Use Case Diagram for MediHelp & DocHelp.

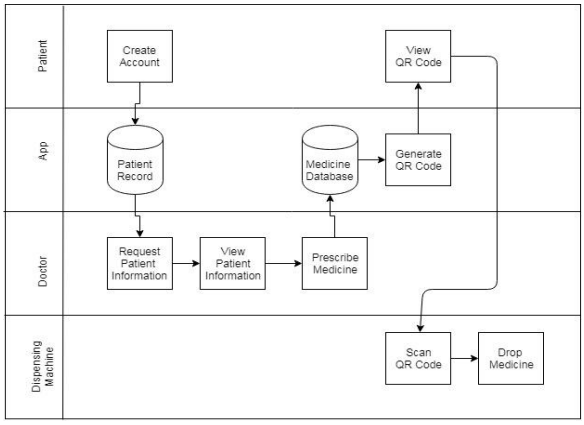


Figure 2- Swim lane diagram depicting the working of AMDM

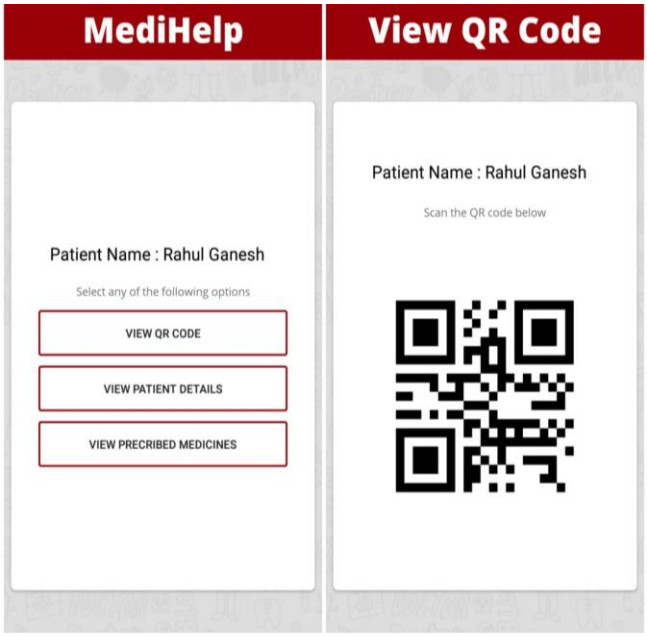


Figure 3 a – Screenshots of MediHelp app (Patient-side app)

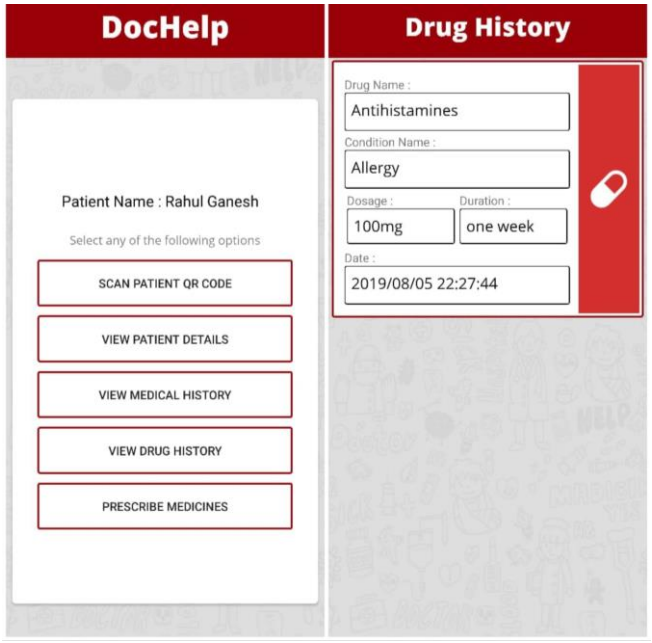


Figure 4 a – Screenshots of DocHelp app(Doctor-side app)

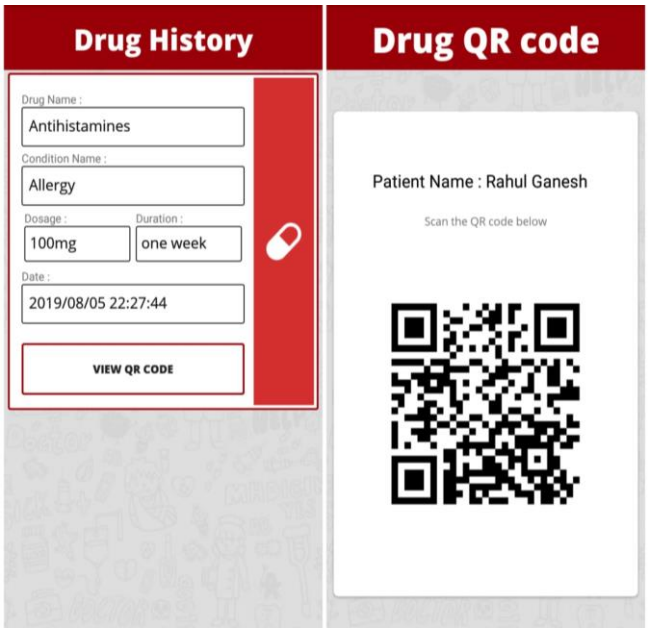


Figure 3 b - Screenshots of MediHelp app (Patient-side app)

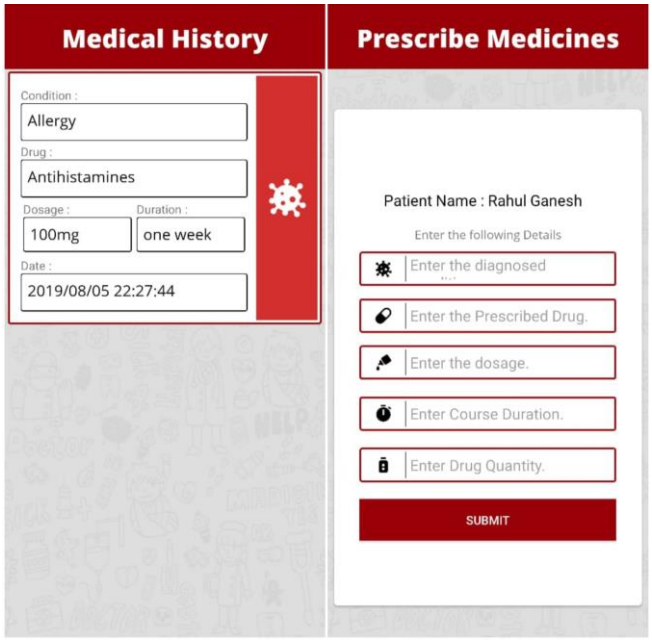


Figure 4 b - Screenshots of DocHelp app(Doctor-side app)

VI. Conclusion & Future Scope

The proposed project AMDM overcomes the human intervention in terms of cost and time. When compared to traditional manual systems, the proposed AMDM may reduce the dispensing and medication errors to a greater extent in hospital. It is planned to extend the project to identify the expiry date of the medicine in the AMDM and produce an alert to the concerned authorities to prevent intake of expired medicines. It is also planned to provide hospital management system via the app from reception to the billing section.

VII. Acknowledgment

We would like to thank Thiagarajar College of Engineering and the Department of Information Technology for providing us this great opportunity. We would like to thank Dr. Deisy, Dr. Sridevi Ms.Sylaja, Ms.Shibani, Ms.Suruthimai for supporting us in this project.

References

[1] who.int, 'Global Health Observatory (GHO) data, Density of pharmaceutical personnel'.
Available: : www.who.int/gho/health_workforce/pharmaceutical_density_text/en/

[2] Cheung, K. C., Bouvy, M. L., & De Smet, P. A. (2009). Medication errors: the importance of safe dispensing. British journal of clinical pharmacology, 67(6), 676–680. DOI:10.1111/j.1365-2125.2009.03428.

[3] Beso A, Franklin BD, Barber N. The frequency and potential causes of dispensing errors in a hospital pharmacy. Pharm World Sci. 2005;27:182–90.

[4] Maviglia SM, Yoo JY, Franz C, Featherstone E, Churchill W, Bates DW, Gandhi TK, Poon EG. Cost-benefit analysis of a hospital pharmacy bar code solution. Arch Intern Med. 2007;167:788–94.

[5] Knudsen P, Herborg H, Mortensen AR, Knudsen M, Hellebek A. Preventing medication errors in community pharmacy: root-cause analysis of transcription errors. Qual Saf Health Care. 2007;16:285–90.

[6] who.int, "General information on counterfeit medicines", available:
www.who.int/medicines/services/counterfeit/overview/en/index.html

[7] CMPI (2009), "Counterfeit drugs and China NEW", available at: http://www.cmpi.org/in-the-news/testimony/counterfeit-drugs-and-china-new(accessed on 16 July 2014)

[8] Newton et al. (2010), "Impact of poor quality medicines in the 'developing' world" Trends Pharmacol Sci. 1. 31(3), S. 99-101

[9] Sarkar U, López A, Maselli JH, Gonzales R. Adverse drug events in US adult ambulatory medical care. Health Serv Res. 2011;46(5):1517–1533.

[10] Gandhi TK, Weingart SN, Borus J, et al. Adverse drug events in ambulatory care. N Eng J Med. 2003;348(16):1556–1564.

[11] Standford Health Care, 'Therapeutic Duplication', october 2016.

[12] Tamblyn RM, McLeod PJ, Abrahamowicz M, Laprise R. Do too many cooks spoil the broth? Multiple physician involvement in medical management of elderly patients and potentially inappropriate drug combinations. CMAJ. 1996;154(8):1177–1184.

[13] Maviglia SM, Yoo JY, Franz C, Featherstone E, Churchill W, Bates DW, Gandhi TK, Poon EG. Cost–benefit analysis of a hospital pharmacy bar code solution. Arch Intern Med. 2007;167:788–94

[14] oon EG, Cina JL, Churchill W, Patel N, Featherstone E, Rothschild JM, Keohane CA, Whittemore AD, Bates DW, Gandhi TK. Medication dispensing errors and potential adverse drug events before and after implementing bar code technology in the pharmacy. Ann Intern Med. 2006;145:426–34

[15] Poon EG, Cina JL, Churchill W, Patel N, Featherstone E, Rothschild JM, Keohane CA, Whittemore AD, Bates DW, Gandhi TK. Medication dispensing errors and potential adverse drug events before and after implementing bar code technology in the pharmacy. Ann Intern Med. 2006;145:426–34

[16] Maviglia SM, Yoo JY, Franz C, Featherstone E, Churchill W, Bates DW, Gandhi TK, Poon EG. Cost–benefit analysis of a hospital pharmacy bar code solution. Arch Intern Med. 2007;167:788–94.

[17] Kazemi A, Ellenius J, Pourasghar F, Tofighi S, Salehi A, Amanati A, Fors UG. The Effect of Computerized Physician Order Entry and Decision Support System on Medication Errors in the Neonatal Ward: Experiences from an Iranian Teaching Hospital.

Design and Simulation of a Prototype to Minimize High Latency in Healthcare Internet-of-Things using Fog Computing

Saurabh Shukla
Computer and Information Sciences Department
Universiti Teknologi PETRONAS(UTP)
32610, Seri Iskandar, Perak, Malaysia
saurabh_16005867@utp.edu.my

Mohd Fadzil Hassan
Center for Research in Data Science(CeRDaS), Computer and Information Sciences Department
Universiti Teknologi PETRONAS(UTP)
32610, Seri Iskandar, Perak, Malaysia
mfadzil_hassan@utp.edu..my

Low Tang Jung
Computer and Information Sciences Department
Universiti Teknologi PETRONAS(UTP)
32610, Seri Iskandar, Perak, Malaysia
lowtanjung@utp.edu.my

***Abstract*—The number of IoT devices has just exploded in large numbers. The worldwide number of IoT devices is expected to reach 50.1 billion by 2020. Healthcare sector is going to have 30.7 % of IoT devices. This much number of IoT devices will increase the service time of cloud computing. High service time will increase the round-trip time delay between cloud servers and healthcare IoT devices. This delay in service will increase the latency for healthcare data transmission to end-users. IoT requires the data in low latency. Cloud cannot meet this requirement. The large volume of data will also lead to high network latency in analysing and acting upon the data. This project report presents the result of the wearable computing ECG device supporting the prototype simulation. Through experiments, healthcare IoT data is collected, visualized and correlated with information on the patient's physiological state. The results can be used to improve the latency between healthcare IoTs and end-users. The proposed work includes the design and development of a prototype which consists of a smart ECG product device to integrate with fog computing gateway to minimize high latency using a machine learning approach. This approach integrates healthcare IoT devices with the cloud and uses fog services.**

Keywords—internet of things, healthcare internet-of-things, cloud computing, fog computing, machine learning, reinforcement learning, fuzzy inference system, artificial intelligence

I. INTRODUCTION

By 2020, there will be 50.1 billion IoT devices. These devices will be generating 507 zettabytes of data [1]. 30.7 % of data will be made available in healthcare. Time-sensitive healthcare applications requires the data in real-time mode. Large data transmission requires high bandwidth consumption and RAM consumption. Transferring the data to the cloud for processing adds the round-trip time delay for healthcare applications[2]. The incerase in volume of latency can be from milliseconds to seconds to minutes. Few healthcare applications requires a latency of less than 60ms. Patient physical sate and health sate change continously with time.Therefore continous monitoring is required which cloud is unable to do. Medical agencies need to manage a large amount of data that may require many servers for analysis of large data [3] The servers also need to collect data from various areas and domains. This large amount of data generated by healthcare IoT is for processing, computing, and storage in the cloud. In turn, it affects the services of the medical team due to high latency.

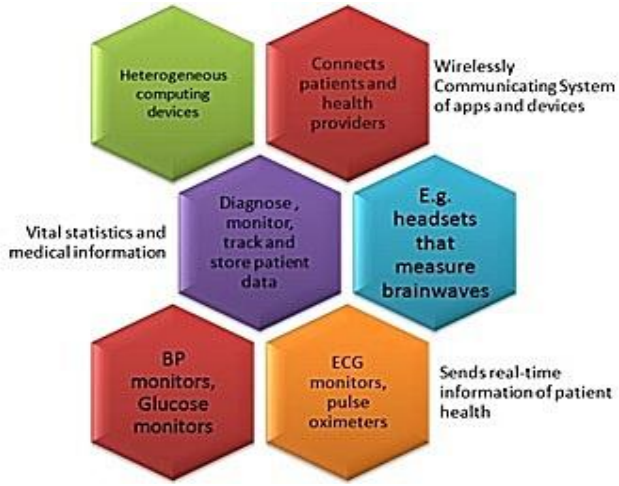


Figure 1: Healthcare IoT device applications and examples

Figure 1 shows the healthcare IoT applications and its examples that consist of Blood Pressure(BP) monitors, Glucose Monitors, Pulse Monitors, Oximeters, Electrocardiography(ECG) monitors etc. To overcome all the issue related to high latency Cisco in 2014 introduced its own concept of minimizing latency called fog computing[4]. Fog computing is an intelligent device that can acts as a middlware gateway between IoT devices and cloud servers. It acts at the edge of IoT devices. It's a kind of small cloud which performs processing and filtering at the edge[5]. Fog computing serevsr uses LAN connections. It send the end-user data in sing hop count as compariosn to multiple hop count of cloud computing. Fog comouting has various features that supports healthcare IoT devices like close proximity to end -users and distibuting and parallel processing[6]. The main objective of this research project is to design an intelligent prototype which transfers the healthcare IoT device data in real-time to end-users by minimizinh high latency. Here we are measuirng the ECG

values of patient as a live data . We have use an ECG wearable device to measures the patient ECG data. Fog computing is used in this project as a gateway to transfer the health data to doctors and patienst in real-time. We have use machine learning approach to make the fog computing gateway an intelligent middleware. The paper is organized as follows: Section 2 discusses the related work, section 3 discusses the methodology and data collection, Section 4 presents the proposed prototype. In Section 5, the conclusions are presented.

II. Related Work

Several works have been done to improve latency, response time and storage in IoT with fog computing and cloud computing in healthcare. The most relevant published works have been included in this section.

In [7] , a Smart e-Health Gateway is proposed to solve the issue of high network latency in IoT by exploiting the concept of fog computing and forming a Geo- distributed layer of intelligence between sensor nodes and cloud. The main limitation of this work is that the smart fog gateway lacks the decision making capability. However, it can cope with many challenges like mobility, energy efficiency, and scalability and reliability issues. In[8] , a Smart Gateway with fog computing is proposed to solve the issue of handling a large amount of data generated by IoT which further leads to high service latency. Here the fog nodes are not able to store the heterogeneous data.

Table 1: Features of Cloud computing and Fog Computing[9]

Feature	Cloud Computing	Fog Computing
Latency	High service latency	Low service latency and provides real-time communication
Connectivity	Cannot operate in the absence of internet	Distribution of the networking computation
Bandwidth	Higher performance issues due to the distance	Bandwidth is conserved by processing local data
Geo-distribution	Centralized intelligence causing congestion	Distributed localized intelligence improves agility
Dynamic analytics	Not suitable for real-time data due to high latency	Suitable for time-sensitive data at the edge

In [6], the fog computing as a smart gateway is proposed to provide the techniques and services as embedded data mining, distributed storage, and notification service and to reduce high latency. The main limitation of this work is that it cannot perform data aggregation and filter at the network edge. In [10], the concept of iFogstor is introduced. iFogstor helps in minimizing service latency, network latency of cloud and reduces network traffic. It helps in reducing the burden of the cloud. Solutions proved very good performance as they reduced the latency by more than 86% as compared to a Cloud-based solution and by 60% as compared to a naive Fog solution. It has few challenges like storing heterogeneous data, tracking previous records and to make decisions in real-time mode. In [11], a reinforcement learning code offloading mechanism is proposed using mobile fog to provide low latency services to mobile service consumers. They have used distributed reinforcement learning algorithm to offload basic blocks in a decentralized fashion to deploy mobile codes on geographically distributed mobile Fogs. In [12], a Hipster, is introduced that combines heuristics and reinforcement learning to manage latency-critical workloads in the cloud. Hipster's goal is to improve resource efficiency in data centres while respecting the QoS of the latency-critical workloads. The proposed technique has limitation for decision making in the real-time environment.

Table 2: The requirement of Quality-of-service (QoS) for Electronic-Health services [4].

Services for E-health	Health service examples	Medium	Response Delay
Real-time conversation	Audio conferencing	Audio	< 150 ms one-way
Real-time video conversation	Video Conferencing	Video	< 250 ms one-way
Real-time service for robotics	Tele-surgery	Audio, video, control of robotic data	< 300 ms round-trip-time
Monitoring patient in real-time mode	Exchange and transfer of patient information	Biomedical data collected via sensors	< 300 ms one-way for real-time ECG

All the mentioned techniques for latency minimization have used conventional gateways for data transmission and latency minimization. The recent works lacks the implementation of intelligent fog computing gateway which acts as a middleware between IoT device and cloud servers, Hence the recent advances for latency minimization in healthcare IoTs requires a novel approach.

III. METHODOLOGY

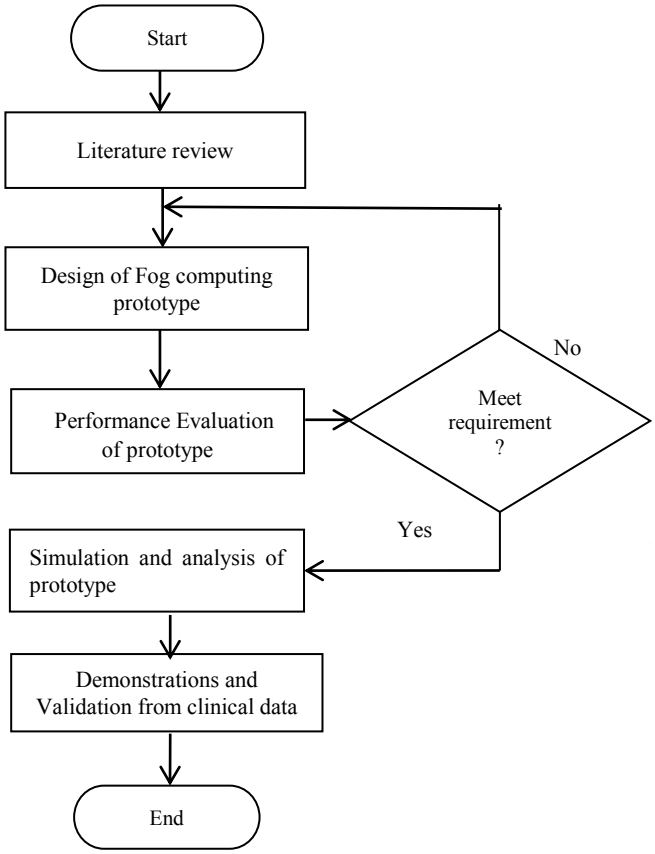


Figure 3: Proposed work methodology

A. Product Description

Figure 4: ECG wearable device

- 1 mobile android device with a dual-core processor and 4GB RAM is used. This device will act as a local gateway between healthcare IoT devices and fog device.
- 1 Laptop i-3 core processor machine with 4 GB RAM and 500 GB Hard Disk. This device will act as a fog gateway between ECG wearable sensor device and cloud service.
- One ECG Wearable Z03 Information Heart Pressure Blood Test Smart device. The device evaluates the ECG offline and flips the wrist to automatically switch the screen. The device is a wearable Hardware with Smart APP and can connect with Big Data Cloud Service.
- The real-time synchronization to the cloud is done through APP control.
- The device connected with gateways can monitor body health changes, find health data, show risk status and evaluate results.
- Other Functions: Step, Calories, Distance, sleep monitor, sport, Call reminder, Alarm Reminder, and social share, etc.
- The product is compatible with Bluetooth, heart rate, blood pressure, step counting, and other functions.
- Master chip: nRF52832, Sensor: BST-BMA253 and ECG chip: TI 129X Photoelectric chip: HRS3300DS
- Vibration: Silence / Vibration Waterproof level: IP68, Compatible systems: for Android 4.4 and above, for iOS 8.2 and above Features: Blood pressure, heart rate, ECG, multi-sports mode.

IV. RESULTS AND DISCUSSION

A. Design Description

The prototype model consists of an ECG sensor device. The model is having a display for user interaction. The designed prototype can process and send the sensor data from one to other nodes. The wearable device, which is a smart ECG monitor watch provides physiological sensor data. The data is further transmitted through (Bluetooth), WIFI or internet to a low energy network to the stable device. The connected devices processes visualize data, interact with the end-users and connects to other localized nodes and servers. The project report presents the result of the wearable computing ECG device supporting the prototype simulation. The prototype of a wearable device presented in this report is a part of a feasibility study for recording and interpreting sensor data. Sensor data is retrieved from heart rate. Sensor data are stored both in raw and processed form.

B. Initial Work

In the first phase of the project, the aim is to create a prototype for a wearable ECG device to minimize the high latency for healthcare IoTs using fog computing. An additional aim is to acquire the patients and doctor's feedback regarding real-time healthcare data transmission. The prototype consists of hardware architecture and software architecture. The hardware architecture consists of an i-5 core processor with 4GB RAM machine. This machine acts as a gateway between IoT devices and end servers. A 4GB RAM android based mobile device acts here as a fog device for local data processing. An ECG device to record the healthcare data. The software architecture consists of an

iFogSim Simulator to simulate the prototype with different parameters. The response time between the healthcare IoTs and the cloud server is affected by the physical distance between them. The fog nodes are acting as a gateway between them can help to resolve the issue of high latency due to multiple hop counts. The total latency is a combination of communications, computations and network latency. Sometimes, selecting the closest node for processing leads to high computing latency. The communication latency depends upon the number of workloads. Whereas here the computing latency is the time spent on the data while being processing on the fog server.

The service latency is the time taken by the data travel from healthcare IoT to server and back from server to healthcare IoT. High traffic and workload due to large data generated by IoT lead to high network latency. See Fig.5 is the fog computing model. It consists of three layers which are as healthcare IoT layer, fog computing layer and cloud computing layer. The first layer is the healthcare IoT layer which consists of an ECG sensor device. These devices generate Patient health data. The Data is then classified into a) High-Risk b) Low-Risk c) Regular and 4) Historical data using fuzzy inference system. Next, the classified data is sent to the fog computing layer i.e. the second layer. Where the selection of time-sensitive data is done using reinforcement learning. Virtualization of fog server is done for allocation of PHD. Then the PHD is sent in real-time mode to end-users via fog computing layer. The patient historical data is sent to the third layer i.e. cloud computing layer for future use.

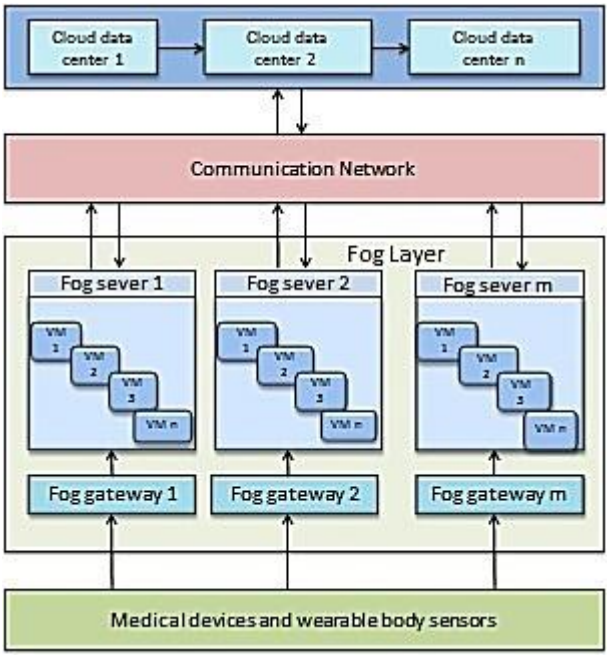


Figure 5: Fog computing model representing fog layer, fog server and fog gateway for real-time data transmission between healthcare IoT devices and cloud data center.

C. Simulation Process and Tool

Disease /Data Classification and categorization using Fuzzy Inference System are:

1) High Risk (Time-sensitive data)

2) Low Risk

3) Regular

4) Historical data

Selection from the above category and classification is done through an agent in the reinforcement learning environment. Fog nodes using reinforcement learning find the best suitable option from the classified data to be sent in real-time mode. Different request from different users for different applications inspired the application of reinforcement learning to select the time-critical data. To simulate the fog computing-based architecture and the analytical model we used iFogsim as an open source software tool and a python based spyder editor tool. This approach can satisfy the varied demands of users. There is the GUI allows the user to draw physical elements such as Fog devices, sensors, actuator, and connected links. The proposed method significantly reduces the network latency of accessing mobile services/fog. To set up parameters in the tool and patient health data values taken from the healthcare IoT. The result is classified as regular, low and risk. These classified data are like streams which are then sent to Real-Time Analyzer like SPARK in a fog computing environment. Here SPARK is open-source software used for handling big data streams in a real-time environment. In the proposed system, we are going to collect real-time data from various wearable body sensors. Cloud data centers contain the historical data of patient which are sent by the fog nodes.

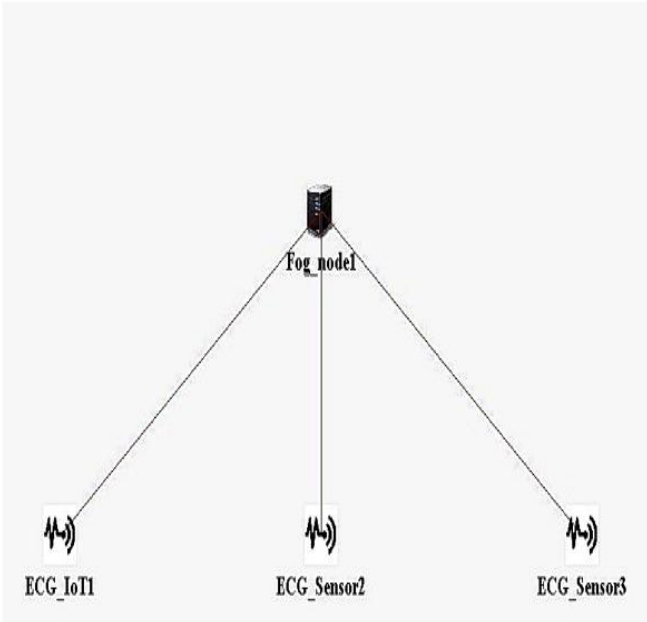


Figure 6: Graphical User Interface to build physical topology arrangements for fog computing using iFogSim simulator.

Table 3: Measurement of patient live health data by wearable device

Heart Rate(HR)	84
Blood Pressure(BP)	106/68
ECG	ECG no abnormalities

HRV Health Index	80
Fatigue Index	52
Physical and mental load	55
Physical Fitness	79
Cardiac Function	79

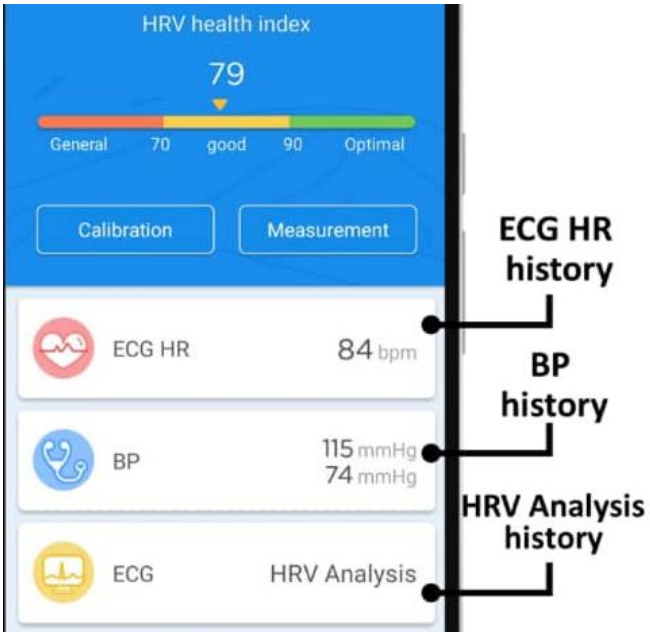


Figure 7: Heart rate value index

Figure 8: Heart rate and blood pressure

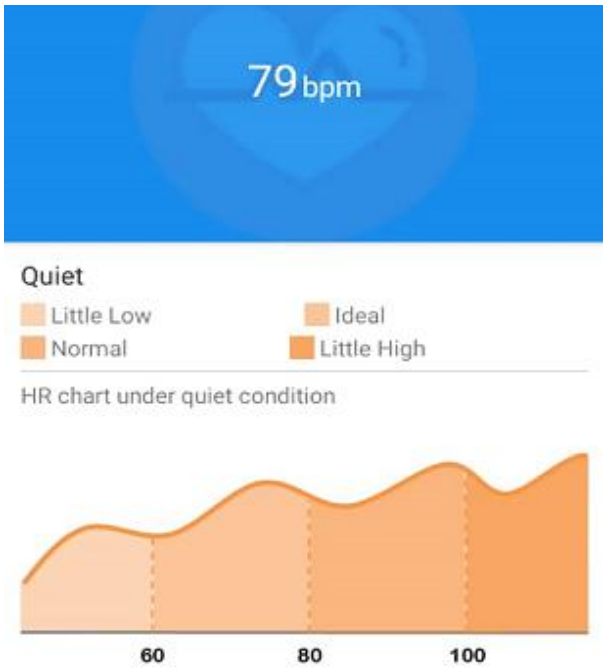


Figure 9: Heart rate chart under quiet condition

Figure 10: Average, highest and lowest heart rate values

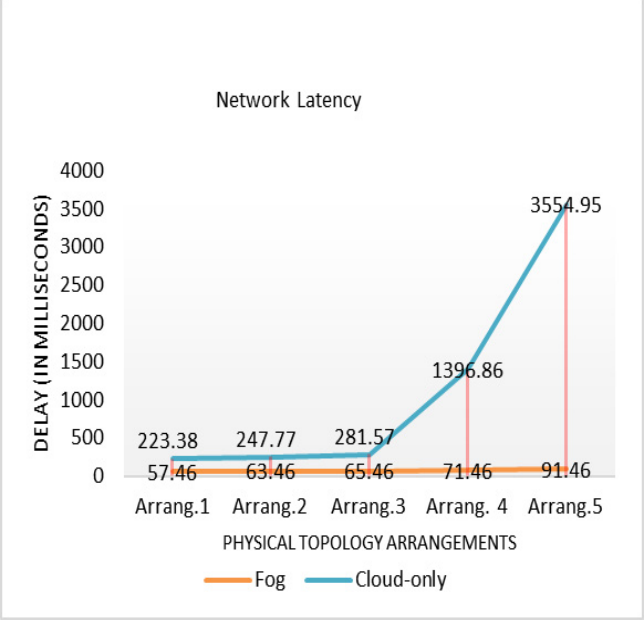


Figure 11: Network latency between fog and cloud of different physical topology arrangements**.**

VI. CONCLUSION

To overcome this limitation, an enhanced fog computing prototype has been proposed, where cloud services are extended to the edge of the network to decrease the latency and network congestion. To realize the full potential of fog and IoT paradigms for real-time analytics, several challenges need to be addressed. Here fog computing covers all the latency minimization parameters which makes it more efficient for real-time data transmission in IoT devices and cloud. The proposed work is expected to allocate the data packets to various processors of virtual machines available in fog gateway by using the technique of virtualization. Using virtualization, fog gateway can tentatively reduce the volume of data to be sent to into data centers of the cloud. Biomedical data analysis of a patient is done with the early medical warning using FIS to transfer the patient's health data in real-time mode by reducing high latency to medical agencies and doctors. The simulation is expected to show the improved performance of the proposed method with respect to network and service latency. There are some challenging issues in the successful development of fog including mobility we continue our works on virtualization, mobility, privacy and end-user issues of fog computing. The future work includes minimizing the service latency along with computational latency.

ACKNOWLEDGMENT

The authors would like to thank the Universiti Teknologi PETRONAS and Universiti Sains Malaysia for supporting this project, and to FPT University Research Festival (ResFest) 2019.

REFERENCE

[1] F. Wortmann and K. Flüchter, "Internet of things," *Business & Information Systems Engineering,* vol. 57, pp. 221-224, 2015.

[2] S.-C. Hung, D. Liau, S.-Y. Lien, and K.-C. Chen, "Low latency communication for Internet of Things," in *2015 IEEE/CIC International Conference on Communications in China (ICCC)*, 2015, pp. 1-6.

[3] B. Hammi, R. Khatoun, S. Zeadally, A. Fayad, and L. Khoukhi, "IoT technologies for smart cities," *IET Networks,* vol. 7, pp. 1-13, 2017.

[4] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*, 2012, pp. 13-16.

[5] S. Cirani, G. Ferrari, N. Iotti, and M. Picone, "The IoT hub: A fog node for seamless management of heterogeneous connected smart objects," in *2015 12th Annual IEEE International Conference on Sensing, Communication, and Networking-Workshops (SECON Workshops)*, 2015, pp. 1-6.

[6] C. S. Nandyala and H.-K. Kim, "From cloud to fog and IoT-based real-time U-healthcare monitoring for smart homes and hospitals," *International Journal of Smart Home,* vol. 10, pp. 187-196, 2016.

[7] T. N. Gia, M. Jiang, A.-M. Rahmani, T. Westerlund, P. Liljeberg, and H. Tenhunen, "Fog computing in healthcare internet of things: A case study on ecg feature extraction," in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, 2015, pp. 356-363.

[8] Y. Liu, J. E. Fieldsend, and G. Min, "A framework of fog computing: Architecture, challenges, and optimization," *IEEE Access,* vol. 5, pp. 25445-25454, 2017.

[9] M. Mukherjee, L. Shu, and D. Wang, "Survey of fog computing: Fundamental, network applications, and research challenges," *IEEE Communications Surveys & Tutorials,* vol. 20, pp. 1826-1857, 2018.

[10] M. I. Naas, P. R. Parvedy, J. Boukhobza, and L. Lemarchand, "iFogStor: an IoT data placement strategy for fog infrastructure," in *2017 IEEE 1st International Conference on Fog and Edge Computing (ICFEC)*, 2017, pp. 97-104.

[11] M. G. R. Alam, Y. K. Tun, and C. S. Hong, "Multi-agent and reinforcement learning based code offloading in mobile fog," in *2016 International Conference on Information Networking (ICOIN)*, 2016, pp. 285-290.

[12] R. Nishtala, P. Carpenter, V. Petrucci, and X. Martorell, "Hipster: Hybrid task manager for latency-critical cloud workloads," in *2017 IEEE International Symposium on High Performance Computer Architecture (HPCA)*, 2017, pp. 409-420.

DustScan Project

Viet Nguyen The
FPT Polytechnic
Hanoi, Vietnam
vietntph06711@fpt.edu.vn

Trung Nguyen Tuan
FPT Polytechnic
Hanoi, Vietnam
trungntph06712@fpt.edu.vn

***Abstract*—Air quality directly affects human health. However, in Asia, air pollution has become a dire situation. The lack of awareness and information prevents people from taking appropriate actions. DustScan provides a way to watch and get notified about air quality in a wide area, both indoor and outdoor. This project will show how severe air pollution in people's everyday life is, and help to make a better living and working environment.**

Keywords—dust, air, quality, temperature, humidity, map

I. INTRODUCTION

Air pollution is the greatest environmental risk to health today, estimated to contribute to 7 million premature deaths every year [1]. Polluted air presents the world's 4th leading contributing cause of early deaths, and burdens the global economy with an estimated annual cost of $225 billion (USD) [2].

According to 2018 World Air Quality report by IQAir [3], among 3000 cities included, 64% exceeded the WHO's annual exposure guideline for fine particular matter, also known as PM2.5. Furthermore, out of all regions, Asian show a dominating number of locations that have the highest average PH2.5 levels during 2018. Among South East Asian countries, Hanoi is ranked 2nd, only below Jakarta of Indonesia. However, as stated by IQAir, in spite of the fact that a large amount of people are affected by pollution, there is a lack of real-time monitoring devices.

In Hanoi, only under 20 stations for measuring dust quality recorded by Air Visual (IQAir), along with 3 stations recorded by The World Air Quality Project[1], and some other personal and private stations. All these stations can only detect the air in the small radius around them so there are still a lot of uncovered areas.

Furthermore, most people have no ideas about air quality inside their own home or office. Without detailed records, they don't know if their living and working environment, as well as their commuting routes are safe or not.

DustScan aims to monitor air quality in real time. For household and office, it can help record and notify if air quality is not optimal. In addition, people can also check the whole city using Dust Map to prepare themselves when heading out..

II. BACKGROUND INFORMATION

Dust is made of small particles of solid matter. It generally consists of particles in the atmosphere that come from various sources such as soil, dust lifted by wind, volcanic eruptions, and pollution. Dust in homes, offices, and other human environments contains small amounts of plant pollen, human and animal hairs, textile fibers, and many other materials which can be found in the local environment.

Fine dust, which consists of tiny particles with diameters of less than 10 micron, is a potential cause for respiratory diseases and increases risk of cancer. The smaller the size of dust particle, the deeper it can penetrate the lungs. Since the nasopharynx can't filter particles smaller than 10 micron, fine dust reaches the lungs to a certain extent.

Based on the diameter of particle, fine dust can be categorized into PM10 or PM2.5. PM10 is particulate matter 10 micrometers or less in diameter, whereas PM2.5 is one with diameter of 2.5 micrometers or less. PM2.5 is especially hazardous to human health and also the main concern of air pollution.

III. MATERIALS AND PROCEDURES

A. Materials

1) Dust sensor PLANTOWER Laser PM2.5 7003

PMS7003 is a kind of digital and universal particle concentration sensor, which can be used to obtain the number of suspended particles in the air, i.e. the concentration of particles, and output them in the form of digital interface. This sensor can be inserted into variable instruments related to the concentration of suspended particles in the air or other environmental improvement equipment to provide correct concentration data in time.

Fig. 1. Dust sensor PlanTower PMS7003

[1] https://aqicn.org

TABLE I. DUST SENSOR TECHNICAL DETAILS

Parameter	Index	Unit
Range of measurement	0.3~1.0; 1.0~2.5; 2.5~10	Micrometer (μm)
Counting Efficiency	50% at 0.3μm 98% at >=0.5μm	Micrometer (μm)
Effective Range[a]	0~500	Microgram per cubic meter (μg/m³)
Maximum Range[a]	≥1000	Microgram per cubic meter (μg/m³)
Resolution	1	Microgram per cubic meter (μg/m³)
Maximum Consistency Error[a]	±10% at 100~500μg/m³ ±10μg/m³ at 0~100μg/m³	
Standard Volume	0.1	Liter (L)
Single Response Time	<1	Second (s)
Total Response Time	≤10	Second (s)
DC Power Supply	Typ:5.0 Min:4.5 Max: 5.5	Volt (V)
Active Current	≤100	Milliampere (mA)
Standby Current	≤200	Microampere (μA)
Interface Level	L <0.8 at 3.3 H >2.7 at 3.3	Volt (V)
Working Temperature Range	-10~+60	Celsius (℃)
Working Humidity Range	0~99%	
Storage Temperature Range	-40~+80	Celsius (℃)
MTTF	≥3	Year (Y)
Physical Size	48× 37× 12	Millimeter (mm)

[a] PM2.5 standard

2) Temperature and humidity sensor DHT22

DHT22 output calibrated digital signal. It utilizes exclusive digital-signal-collecting-technique and humidity sensing technology, assuring its reliability and stability. Its sensing elements is connected with 8-bit single-chip computer.

Fig. 2. Temperature and humidity sensor DHT22

TABLE II. TEMPERATURE & HUMIDITY SENSOR TECHNICAL DETAILS

Parameter	Index	Unit
Power supply	3.3-6 DC	Volt (V)
Output signal	digital signal via single-bus	
Sensing element	Polymer capacitor	
Operating range	humidity 0-100%RH; temperature -40~80Celsius	
Accuracy	humidity +-2%RH; temperature <+-0.5Celsius	
Resolution or sensitivity	humidity 0.1%RH; temperature 0.1Celsius	
Repeatability	humidity +-1%RH; temperature +-0.2Celsius	
Humidity hysteresis	+-0.3%RH	
Long-term Stability	+-0.5%RH/year	
Sensing period	Average: 2	Second (s)
Interchangeability	fully interchangeable	
Dimensions	22*28*5	Millimeter (mm)

3) Module SIM 808 GPS/GSM/GPRS/Bluetooth

SIM808 module is a GSM and GPS two-in-one function module. It is based on the latest GSM/GPS module SIM808 from SIMCOM, supports GSM/GPRS Quad-Band network and combines GPS technology for satellite navigation.

Fig. 3. Module SIM808

TABLE III. SIM808 TECHNICAL DETAILS

Parameter	Index	Unit
Power supply	3.3-4.4 DC	Volt (V)
Bandwidth	850/900/1800/1900	Megahertz (MHz)
Bluetooth	compliant with 3.0+EDR	
GPRS multi-slot	class 12/10	
GPRS mobile station	class B	
GSM	Compliant to GSM phase 2/2+	
Operation temperature	-40 ~ 85	Celsius (℃)
GPS receiver channels	22 tracking / 66 acquisition	
Tracking sensitivity	-165	Decibel-milliwatt (dBm)
Update rate	5	Hertz (Hz)
Weight	3.3	Gram (g)
Dimensions	24*24*2.6	Millimeter (mm)

4) Arduino WeMos D1

WEMOS D1 is a WIFI development board based on ESP8266 12E. The functioning is similar to that of NODEMCU, except that the hardware is built resembling Arduino UNO.

Fig. 4. Arduino WeMos D1

TABLE IV. WEMOS D1 TECHNICAL DETAILS

Parameter	Index	Unit
Operating Voltage	3.3	Volt (V)
Digital I/O Pins	11	
Analog Input Pins	1(Max input: 3.2V)	
Clock Speed	80/160	Megahertz (MHz)
Flash	4	Megabyte (Mb)
Length	68.6	Millimeter (mm)
Width	53.4	Millimeter (mm)
Weight	25	Gram (g)

5) Other materials

- Wires
- Power sources
- Sim card
- 3D printed containers

B. Software and development setup

- Arduino IDE
- Server – Ubuntu with PHP development environment
- Firebase realtime database
- React Native for mobile apps

C. Procedures

1) Block diagram

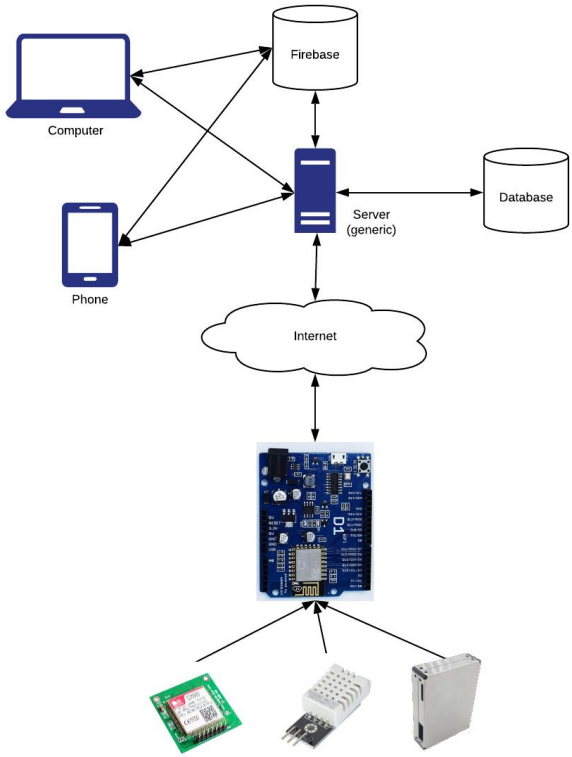


Fig. 5. Block diagram

2) Measuring and sending data

Indoor devices

Sensors detect and measure every 2 seconds then send request to the server using household or office WIFI network.

Outdoor devices

Each set of device will be attached to a moving vehicle to maximize the area. Because the dust sensor has to be placed at a certain height to prevent inaccurate readings, buses are preferred.

Module GPS produces output every second. Because the results will not be correct if there is any movement during the measuring process, an extra step has to be taken before sending the request.

By using the average velocity calculated by GPS module and setting a threshold, below which the devices are considered to be not moving, the data recorded by other sensors (temperature, humidity, dust) are sent to the server.

3) Receiving and storing data

Both data sent by indoor and outdoor devices are processed in the same way.

When the server receives data, the coordinate of the device at the time of measurement will be checked in the MySQL database. If the coordinate is not present, a new firebase will be created. Then this key, along with all the data, will be inserted into MySQL database and firebase database.

If the coordinate is already recorded, its key will be used to update firebase database. The new record will also be inserted into MySQL database.

4) Retreiving and displaying data

Both the indoor app and the map will display the latest measurements on their initial view page. These measurements are taken directly from Firebase realtime database, and they automatically update every time there is a new change.

For the map, if a user click on one marker (defined by one coordinate), they can see the details and several charts showing average value for each index of current date and month.

Similarly, for the indoor app, when clicking on one index, information and charts will also appear.

IV. Results

Both apps and map display results recorded in real time. When user click on one marker or one parameter, detailed information is displayed.

Dust Map can be accessed via http://52.221.226.126/map

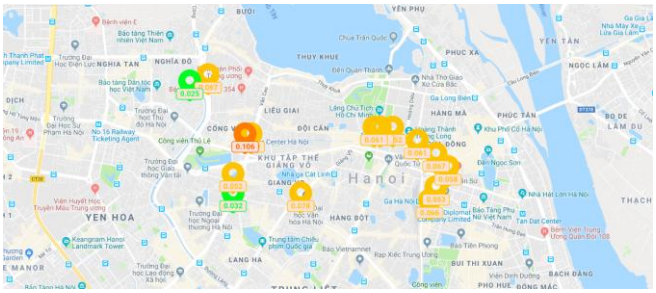


Fig. 6. Dust Map

V. Discussions

The project aims to provide a way to monitor and display air quality in real time both indoor and outdoor.

All 3 indices (dust density, temperature, and humidity) are recorded quite correctly with acceptable margin of error.

The map shows new markers and updates existing markers when there's any change detected. The markers can also change color based on dust density level.

All the charts are generated at the time of clicking button to see more, so they can update the latest calculated average values.

VI. Conclusion

DustScan has achieved the first step to create a system to measure and record air quality in real time and over a great area. Using multiple devices indoor and attached to moving vehicles give a realistic and detailed overview of what people breathe in every day.

The indoor device works as intended. However, the outdoor device depends on some factors. The first factor is the weather. Weather condition has a large impact on the data collected, especially when it rains. The second one is

traffic. Heavy traffic or traffic jam can also makes data incorrect.

Moreover, in this project demo, all the sensors used are more suitable for using indoor rather than outdoor. Therefore, in order to get better and more accurate results, future implementation must use better equipment.

In the future, the scope of the project may extend to other indices of air quality, e.g. CO2, NO2 concentration. Each index can also be predicted for better application in everyday life.

ACKNOWLEDGMENT

We would like to thank FPT Education for giving us an opportunity to learn and make this project come true. We would also like to thank our advisors, Mr. Thien Tran Huu and Mr. Long Tran Thanh, for their supports.

REFERENCES

[1] "Air Pollution Deaths Cost Global Economy US$225 Billion," The World Bank, 8 September 2016. [Online]. Available: https://www.worldbank.org/en/news/press-release/2016/09/08/air-pollution-deaths-cost-global-economy-225-billion.

[2] "2018 World Air Quality Report," IQAir, 5 March 2019. [Online]. Available: https://www.airvisual.com/world-most-polluted-cities/world-air-quality-report-2018-en.pdf.

[3] "9 out of 10 people worldwide breathe polluted air, but more countries are taking action," World Health Organization, 2 May 2018. [Online].

Helmet Detection and Extraction of License Plate Characters using the YOLO Algorithm

V. Surya, S. Sanjana
Department of Information Technology,
Thiagarajar College of Engineering (TCE),
Madurai, Tamil Nadu, India.
{zursolomon, sanjanasekarpandian67}@gmail.com

Abstract*—Two wheelers like bikes and scooters are the most desirable modes of transport in India owing to the country's economy, lifestyle choices and traffic conditions. But because of carelessness of the people and negligence of the traffic police, detection of helmet and non-helmet wearing motorcyclists is mandatory these days in order to ensure the safety of riders on the road. However, due to many constraints such as poor video quality, occlusion, illumination, and other varying factors it becomes very difficult to detect them accurately. During the past years, the advancements in deep learning models have drastically improved the performance of object detection. One such model[1] is You Only Look Once Version 3(YOLOv3) which combines both classification and object detection in a single architecture. In this paper, we introduce an approach that aims to identify violators riding bikes by ignoring the helmet rule, and scan their license plate characters using a state-of-the-art Optical Character Recognition system, and send the extracted characters to the nearest traffic police department. This system will make use of Internet of Things(IoT),Machine Learning(ML) and Artificial Intelligence(AI) concepts. One model benefitting said growth is the YOLO algorithm We tested the potential of our approach on different helmet and non-helmet wearing images. Experimental results show that the proposed method performs better when compared to other existing approaches.

Keywords —helmet detection, YOLOv3, license plate extraction, COCO

I. Introduction

Motorcycles being one of the most convenient modes of transport has led to increasing usage and thus accounts for the highest rate of total road accidents. As per one survey carried out in the year of 2014, 30% of all road accidents related to deaths were of riders on two-wheelers [2]. Also, as per the record of Chennai, India, between the period of January 1, 2013 to June 28, 2015, total 1,453 two-wheeler riders who died in road accidents were not wearing helmet [3]. Also, according to the World Health Organization (WHO), wearing a helmet can reduce the risk of severe injury by 72% and the risk of death by 39% [2]. In India, road safety has been a neglected area. Due to this wearing a helmet has been made mandatory and violation of the same attracts a hefty fine. Currently, all major cities employ surveillance system on roads to identify violators which needs a system that can automatically detect whether a motorcyclist is wearing a helmet or not. With the advancements in technology over the past few years, one could say that we are currently living the future we had once dreamed of living. With this rise in technology and with the birth of even more powerful CPUs and most importantly, GPUs, the fields of Machine Learning and Artificial Intelligence have seen a rapid growth recently. This paper aims to an automatic helmet and non-helmet detection of motorcyclists.

II. Related Work

Since the past several years, many researchers have solved the problem of helmet detection. Wen *et al.* [4] proposed a circular arc detection method based on Hough transform for helmet detection. Chiu *et al.* [5] use a Canny edge detector [6] to detect helmets. Similarly, authors in [7] use a background subtraction to detect moving vehicles. Here, authors calculate an image based statistical information by isolating the approximate region of the head of the riders. In order to extract features, Histogram of Gradient (HOG) descriptors is used from the isolated regions and then results are passed to a Support Vector Machine (SVM) classifier for classification purpose. Furthermore, authors in [8] proposed a model for vehicle detection, tracking and classification from roadside CCTV cameras. In order to make their model invariant against illumination changes and camera vibrations, they use a new background Gaussian mixture model. Silva *et al.* in [9] propose a hybrid descriptor model based on geometric shape and texture features for automatic detection of motorcyclists without helmet. They use Hough transform with SVM to detect the head of the motorcyclist. Additionally, they extend their work in [10] by multi-layer perception model for classification. In [11], Waranusat *et al.* [11] propose a system to detect moving objects with K-Nearest Neighbour(K-NN) classifier over the head region of motorcyclist to classify helmet. Similarly, authors in use background subtraction and object segmentation with k-NN classifier to detect bike-riders. All these methods which are employed with engineered features have limitations in accurate detection of helmets.

Alex *et al.* [13] introduced a Convolutional Neural Network (CNN) based method for object classification and detection. Since then CNNs became the standard model for classification as it outperformed all the traditional models such as HOG , scale-invariant feature transform (SIFT) and Local Binary Patterns (LBP). Recently, A. Hirota *et al.* use a CNN for classification of helmeted and non-helmeted riders. Although they use CNN, their helmet detection accuracy is poor with limitations of helmet color and multiple riders on a single motorcyclist.

Literature shows that the CNN can learn engineered features from raw data effectively and it outperforms over it's hand-crafted counterparts. Therefore, we use CNN to solve helmet detection problem.

III. Proposed Approach

The overall contributions of our paper are as follows:

- Use of person detection instead of motorcyclist in order to increase the helmet detection accuracy in the test image.
- Training a model using our database of helmet wearing motorist images.
- Extraction of license number plate for non-helmet wearing motorcyclist.

a) Raspberry pi

The Raspberry pi is considered as the edge device which will be placed along with the camera module in the traffic signal. The camera is triggered to take photo graphs and short videos whenever there is a stop signal. The photos and the videos are uploaded to the cloud via the raspberry pi. The server uses the photo graphs and the videos as test data and finds out if the person is not wearing a helmet and extracts the number plate of the vehicle and sends it as SMS to the concerned traffic police authority.

b) Work Flow

The block schematic of the proposed approach is depicted in Fig. 1. Here, we use two YOLOv3 models one after another in order to detect helmet and non-helmet motorcyclist. First YOLOv3 model is trained with COCO dataset which detects number of classes in an image. COCO is a large scale object detection, segmentation and captioning dataset. The cropped images of person class from the first YOLOv3 are used as input to the second YOLOv3 model which was trained on helmet dataset. This helmet dataset is prepared by us and it consists of helmeted images only. Also, the number of filters is modified to 30 in the final convolutional layer in order to detect single helmet class. At last, a license plate detection is performed using ocr space's API.

c) *Person Detection*

The first step in the proposed approach is to detect a person from the input image. In order to perform this, the input test image is passed through first YOLOv3 model which is trained on COCO dataset. YOLOv3 is the state of the art object detector and it has the capability to detect all classes of COCO dataset. It can also detect classes such as person, car, motorbike, etc., in addition to other classes in the dataset. We use a person detection class out of all other classes in order to detect helmeted motorcyclists. All other detected classes from the first YOLOv3 model have been discarded.

Here, it is worth to note the reasons for selecting the bounding box of a person i.e. person detection instead of motorcycle in the first step which is based on our empirical tests. Since the input image consists of motorcycles, cars and persons we found out that when the camera is either front or back facing, then the detector does not always detect the motorcycle or it detects motorcycle with very less confidence score. In such cases, if motorcycle detection criteria is used then it leads to many helmets as undetected in the test image. However, in the criteria of person detection, if a person is sitting on a motorcycle, then the head region covers the helmet, and the foot region contains the license number plate area in motorcycle as shown in the Fig 1. For this reason, we use person detection criteria instead of detecting motorcycle or person and motorbike together. This also increases the detection score of helmet in the input test image.

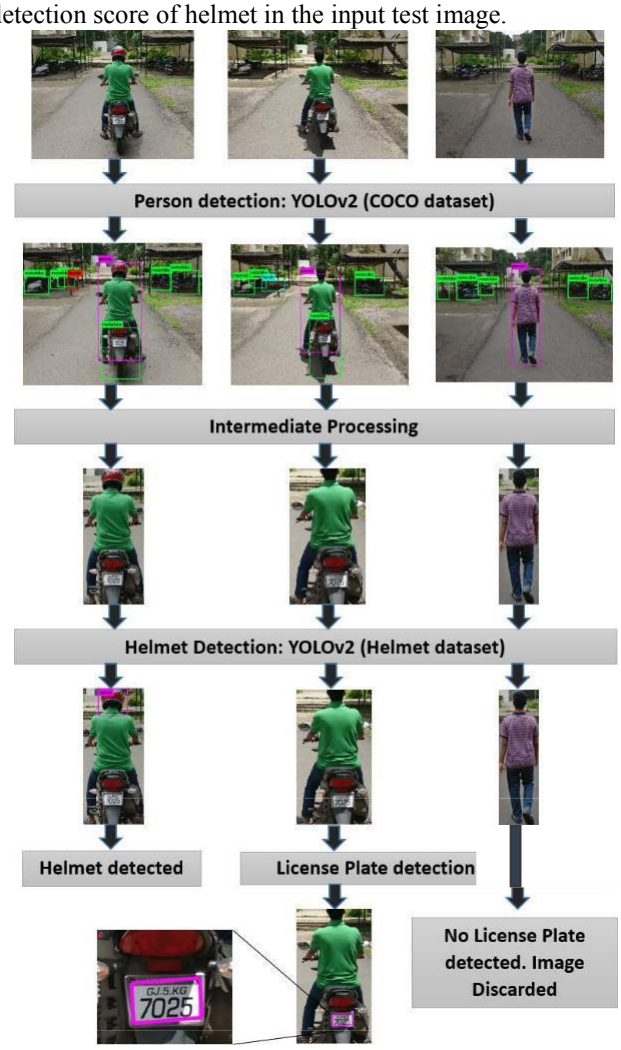


Fig. 1: Block schematic representation of the proposed approach.

d) *Intermediate processing:* As discussed earlier, YOLOv3 detects all the classes from the COCO dataset and in the pro-posed approach we use person detection only in order to detect helmeted and non-helmeted motorcyclist. All the other classes except person are discarded by intermediate processing. Also, the detected person's bounding box is cropped automatically and that image is used for further processing.

e) *Helmet Detection:* The main step of the proposed approach is to detect helmeted and non-helmeted motorcyclists in an image. For this step, we again use

YOLOv3 model which is trained on dataset of helmeted image. The cropped images of detected person are used as input here to YOLOv3 model. Since YOLOv3 model is trained on helmeted image dataset, whenever a test image consists of helmets, same is classified and detected in that cropped image.

f) *License Plate Detection:* If helmet is detected from the output of the second YOLOv3 stage, then the process is stopped for that cropped image. However, if helmet is not detected from detected person's image, then the cropped image is required to pass through the license plate detector, Optical Character Recognition(OCR), which detects the license plate with coordinates. These coordinates are used further in order to extract the license plate. Note that if the number plate is not detected, then it means that instead of a person on the motorbike, the first YOLOv3 had detected a pedestrian.

IV. Algorithms Used

A. You Only Look Once-YOLOv3

During the past several years, the advancements in deep learning models have drastically improved the performance of object detection. One such model is YOLO which combines both classification and object detection into a single architecture. It's upgraded version i.e., YOLOv3 with 23 convolutional layers and 5 max pooling layers is very fast which can detect the helmets in real time with high confidence score. Also, it has a capability to learn general representations of objects, which is very helpful in helmet detection as helmets can be of various shapes and sizes. This motivates us to use this model for our approach. Moreover, YOLOv3 predicts bounding boxes and class probabilities directly from full images in one evaluation. It also learns about the contextual information about classes in addition to their appearance. Due to this, during training, when we give images of person wearing a helmet, it learns that helmet is generally worn in the head region, which enables it to detect helmet more accurately.

B. Optical Character Recognition*(OCR) API*

"The OCR.space Online OCR service converts and scans images of text documents into editable files by using Optical Character Recognition (OCR) technologies. It uses state-of-the-art modern OCR software. The recognition quality is comparable to commercial OCR SDK software (e. g. Abbyy)" [From https://www.ocr.space]. Due to the robust support and vast features we use the same to detect license plate on the motorcyclist. The default detector has been trained for Indian number plates but we can add a country by training the detector on large example dataset of that country's license plates. Instead of adding a new layout of license plate, which would have been cumbersome, we relied on the region extraction of license plate. The region extraction feature works better and it results the coordinates of the number plate in the image.

V. Experimental Results

The experiments were conducted on a machine with Intel i5 6200U CPU, 8GB RAM and 2 GB NVIDIA GeForce GT920 dedicated GPU. The programs for helmet detection are written in Python 3.6 with the help of the various libraries such as OpenCV 3.0 and Pytorch which is a deep learning research platform that provides maximum flexibility and speed.

A. Database

In the proposed approach, we use two datasets in order to detect helmeted and non-helmeted motorcyclists. At the first stage of YOLOv3 model, we use COCO database which consists number of images of different classes. The second database which consists of helmeted images is used at the second YOLOv3 model.

a) *Training:* For person detection at first YOLOv3 stage, we use COCO dataset in order to train that network. For detecting helmet at the second stage, we trained YOLOv3 model with helmeted image dataset. This dataset is prepared by us and it has a total of 800 helmeted images. Here, we trained the YOLOv3 model for one class detection only i.e., helmet. Since our dataset contained around 1000 images only, we chose to train the model using weights pre-trained on ImageNet [13]. Network pre-trained on a huge and manifold dataset like the ImageNet, captures features like curves and edges in its early stages which are useful to most of the classification problems. Also, this speeds up the training process as model has already learned the elementary features. For both training and testing the batch size is set to 64 with the learning rate of 0.001. The momentum and decay are set to 0.9 and 0.0005, respectively.

b) *Testing:* For testing purpose, the helmeted and non-helmeted images have been downloaded from the ImageNet dataset . It is worth mentioning that these images were not used in the training dataset at the second YOLOv3 model i.e. helmet dataset. This testing dataset consists of a total of 409 helmeted and 403 non-helmeted images (i.e., person, car, motorcycle, etc.,). According to the norms, one can use 700 helmeted and 100 non-helmeted images, i.e. approximately 20% of training data for testing. However, we use equal amount (almost 50%) of helmeted and non-helmeted data since our model was already aware of faces (transfer learning).

B. Result and Discussion

In this section, experimental results obtained using proposed approach is presented. To check the robustness of the algorithm, we tested it on images having both helmet and non-helmet wearing motorcyclists. Results obtained using proposed method for different scenarios are displayed in Fig. 2.

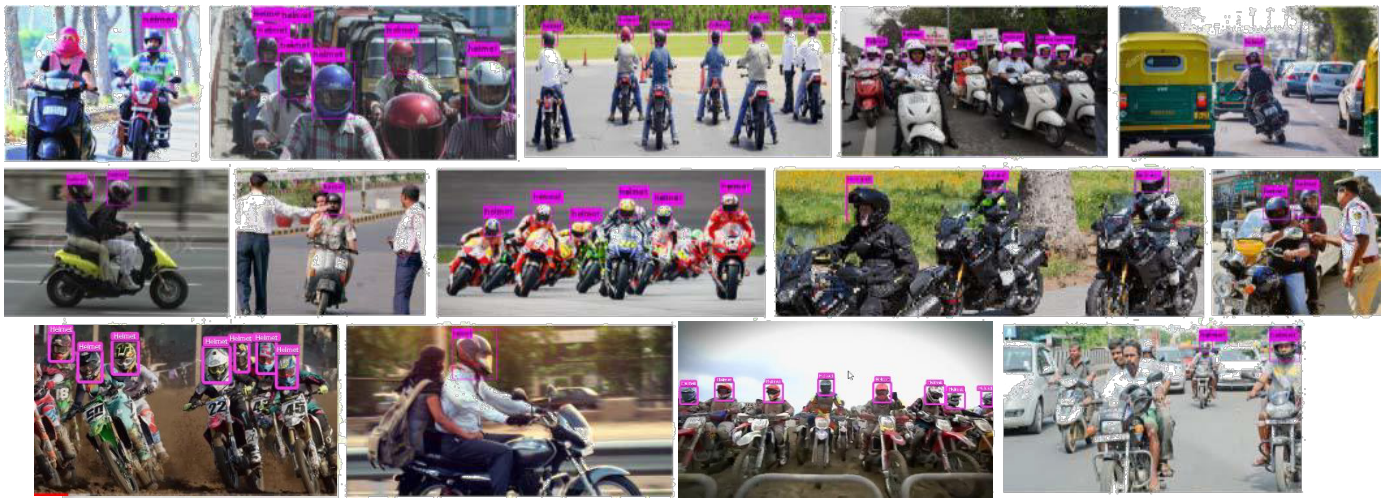


Fig. 2: Helmet detection in different scenarios.

As one can see from figure that the helmets are detected accurately in crowded areas and also in the images with single motorcyclist. Additionally, our approach can distinguish between cap and helmet despite both of them having similar features. Also, it can differentiate scarves from helmets which may pass off as helmets in some cases (see first image in Fig. 2). It is important to note that the proposed approach also detects the helmets accurately in the motorcyclists captured from side-view camera. Results for the side-view motorcyclists are displayed in the third row in Fig. 2. Furthermore, the helmet detection approach proposed by A. Hirota *et al.* fails when color of helmet is black, same as hair color. However, this limitation is overcome in the proposed method and our method can detect helmets of all colors and hapes. From the results displayed in Fig. 2, one can note that the helmets are detected and the license plates are extracted for non-helmeted images by the proposed method. Thus, our approach is robust and reliable in different scenarios.

In Fig. 3, the graph of average error during training iterations i.e., number of batches completed. Here, one can see that at the beginning of the training, the average error started very high, i.e., approximately 160, and then it reduced exponentially and finally it becomes stagnant, with a very little but significant change, after 500 iterations. We stop our training around 2500 iterations or epochs where the error was not changing much. Here, we use a learning rate of 0.001.

Finally, Fig. 4 shows the results of helmet detection at various phases of training. Fig. 4(a) displays the result obtained when 50 Epochs of training for helmet detection are completed. Here, no helmet is detected.

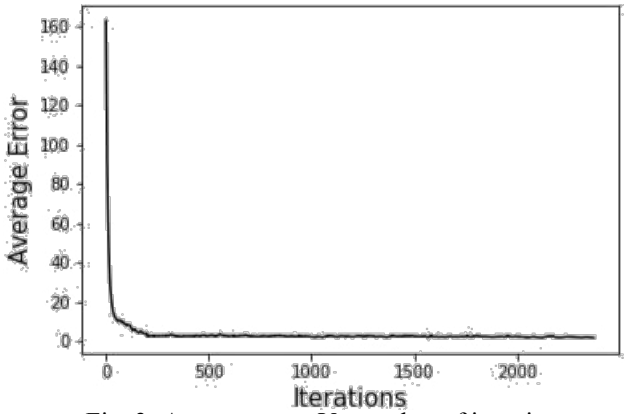


Fig. 3: Average error Vs number of iterations.

4(a) 4(b)

4©

Fig. 4: Detection of multiple helmets for different iterations:

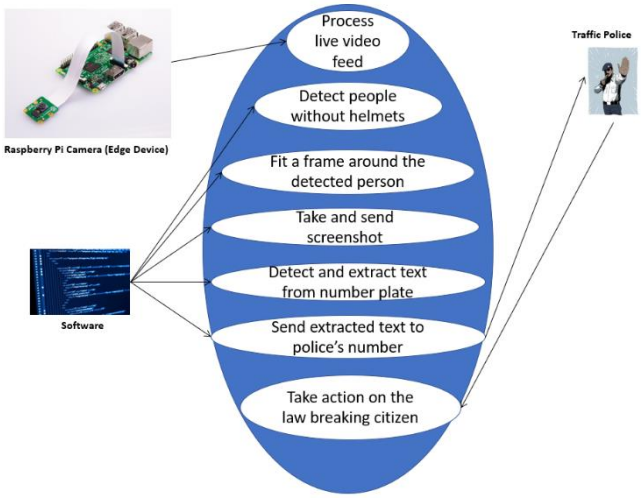


Fig. 5: Use Case Diagram of the Proposed System

in Fig. 4. However, the objects which are not helmet are also detected as class helmet. Finally, Fig.4 depicts the helmet detection results obtained when 2500 iterations are over which shows that the final weights detect all the helmets correctly without any false positive. Fig 5: shows the use case diagram of the proposed system.

V. Conclusion

An approach for automatic helmet detection using Convolutional Neural Network is proposed. In order to increase the helmet detection accuracy, we use two stages of YOLOv3 models. This YOLOv3 model ensures the person detection which is trained on COCO dataset and a custom dataset obtained by bulk downloading google search results and filtering them. This step decreases the number of helmets being undetected. The cropped images of detected person are used as input to the second YOLOv3 model which was trained on our helmeted image dataset. The proposed method has been tested on different helmeted image scenarios. Also, we evaluated the quantitative measures on the test images and same is compared with other state of the art methods. Experimental results and quantitative measures on different scenarios show that our approach is robust and reliable with a high helmet detection accuracy.

VI. Acknowledgement

We would like to thank Thiagarajar College of Engineering and the Department of Information Technology for providing us this great opportunity. We would like to thank Mrs.C.V.Nisha Angeline for guiding us in this project and Dr. Deisy and Dr. Sridevi for supporting us in this project.

VII. References

[1] J. Redmon and A. Farhadi, ―YOLOv3: An Incremental Improvement, C*oRR*, vol. abs/1804.02767, 2016. [Online]. Available: http://arxiv.org/abs/1804.02767

[2] [Online]. Available: https://scroll.in/article/757365/road-accidents-kill-382-in-india-every-day-1682-times-more-than-terrorism

[3] [Online]. Available: http://timesofindia.indiatimes.com/city/chennai/98-6-of-bikers-who-died-didnt-wear-a-helmet/articleshow/47904790.cms

[4] C.-Y. Wen, S.-H. Chiu, J.-J. Liaw, and C.-P. Lu, ―The safety helmet detection for atm‘s surveillance system via the modified hough trans-form,‖ in *IEEE 37th Annual 2003 International Carnahan Conference onSecurity Technology, 2003. Proceedings.*, Oct 2003, pp. 364–369.

[5] C.-C. Chiu, C.-Y. Wang, M.-Y. Ku, and Y. B. Lu, ―Real time recogniton and tracking system of multiple vehicles, in *2006 IEEE Intelligent Vehicles Symposium*, 2006, pp. 478–483.

[6] J. Canny, ―A computational approach to edge detection, *IEEE Transac-tions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6,
pp. 679–698, Nov 1986.

[7] J. Chiverton, ―Helmet presence classification with motorcycle detection and tracking,*IET Intelligent Transport Systems*, vol. 6, no. 3, pp. 259– 269, September 2012.

[8] Z. Chen, T. J. Ellis, and S. A. Velastin, ―Vehicle detection, tracking and classification in urban traffic,‖ *2012 15th International IEEE Conference on Intelligent Transportation Systems*, pp. 951–956, 2012.

[9] R. Silva, K. Aires, T. Santos, K. Abdala, R. Veras, and A. Soares, ―Automatic detection of motorcyclists without helmet, in *2013 XXXIX Latin American Computing Conference (CLEI)*, Oct 2013, pp. 1–7.

[10] R. R. V. e. Silva, K. R. T. Aires, and R. d. M. S. Veras, ―Helmet detection on motorcyclists using image descriptors and classifiers, in *2014 27th SIBGRAPI Conference on Graphics, Patterns and Images*, Aug 2014,
pp. 141–148.

[11] R. Waranusast, N. Bundon, V. Timtong, C. Tangnoi, and P. Pat-tanathaburt, ―Machine vision techniques for motorcycle safety helmet detection,‖ in *2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013)*, Nov 2013, pp. 35–40.

[12] K. Dahiya, D. Singh, and C. K. Mohan, ―Automatic detection of bike-riders without helmet using surveillance videos in real-time,‖ in *2016 International Joint Conference on Neural Networks (IJCNN)*, July 2016,
pp. 3046–3051.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ―Imagenet classification with deep convolutional neural networks,‖ in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012,pp.1097-1105.

Home Assistant

Nguyễn Việt Hoàng
FPT School Danang
Danang,Vietnam
vietcutun@gmail.com

Lê Thanh Tùng
FPT School Danang
Danang,Vietnam
tonylethanh2003@gmail.com

Our product makes life simpler and better with innovative technology, products and design. We provide many useful features to make your house to be the smart house such as smart Alarm system, Voice Control, AI chatbot notification…

I. Introduction (*Heading 1*)

Smart house is no longer a very over rate system but essentially needed for a convenient and modern lifestyle. To complete demand , we bring to you : Home Assistant .This is an open source home automation platform that runs on Python 3.x, designed to be easily deployed on any computer from Raspberry Pi to network storage devices (NAS) and even a Docker container to deploy on other systems in an easy-to-follow manner. We aim to develop Home Assistant to be able to used by everyone with minimal botheration and minimal lag.

II. System requiremnets

A. Hardware requirements

*1) **Raspberry pi 3 b+** : The code processing center, the brain of the house, LAN SERVER*

*2) **Google Home Mini**: Voice-controlled smart speaker, supports Google Assistant features, links home assistant with IFTTT*

*3) **Sonoff T1 US:** Smart switch has been flashed the custom rom to connect to pi via MQTT*

*4) **Xiaomi Mijia hub:** Including (door sensors, motion sensors, temperature sensors, light sensors, zigbee sockets) linked to pi via LAN*

B. Software Requirments

1) Home Assistant:

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

2) Google Assistant:

Google's Assistant is considered as an upgraded or extended version of Google Now, which allows to expand voice control capabilities, while expanding the voice controls with existing OK Google. If you've ever used Google Now, it attracts information relevant to you.

3) Tasker:

Tasker is an application for Android which performs tasks (sets of actions) based on contexts (application, time, date, location, event, gesture) in user-defined profiles or in clickable or timer home screen widgets. Tasker for Android lets you automate almost anything on your phone. Tasker is an automation app. You create a list of actions, called a **task**, which can then be executed according to a variety of contexts. You tell Tasker you want it to do A, B, and C when X and/or Y happens. You might be thinking that Tasker's like IFTTT (**If This Then That**), and you'd be half right. Tasker tasks are similar to IFTTT applets, and both services have a wide variety of plugins. IFTTT has wider service support and cross-platform usability, but Tasker can do more on the Android devices it touches with more complex actions.

4) Chatfuel:

Chatfuel is the leading bot platform for creating AI chatbots for Facebook

5) Pushbullet

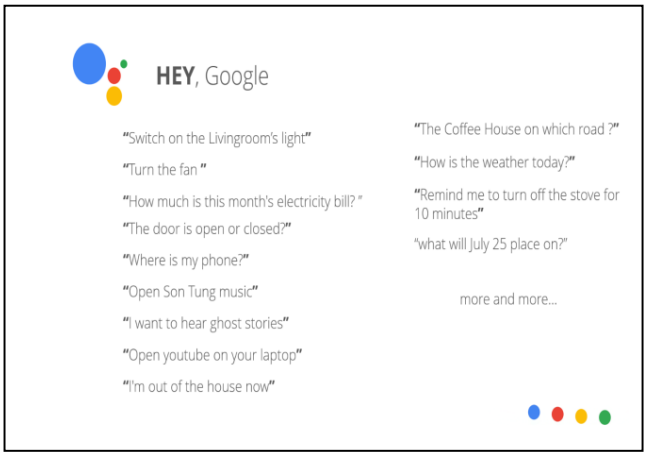
Pushbullet is a cross-platform application which bridges the gap between your Phone, Computer, Tablet so they all work together better. With Pushbullet you can see your phone's notification to your PC, share links and files easily between your devices instantaneously. This app is free with added paid features. Currently available for Android, iPhone, Windowsapplications and Add-ons for Chrome, Firefox, Opera and Safari browsers.

III. System features

A. *Voice control*

First, consumers have an ability to directly control devices by saying via Google Assistant, an well-known groundwork of voice recognitions . Besides some of already known Google's features, that lots of persons had used frequently, we improve more practical command for the voice control. People are, now on, no need for taking exility actions like going out of their warm bed just for turning every light on and off or electronic devices. When going out for dinner, for example, instead of turn off devices manually, how 'bout say "I'm leaving home" and your houses will turn off straight away and keep your security alarm awake.

Your assistant could directly answer your question about the electronic bill you have to pay for that month and give you advices to save energy by some programmed calculation.



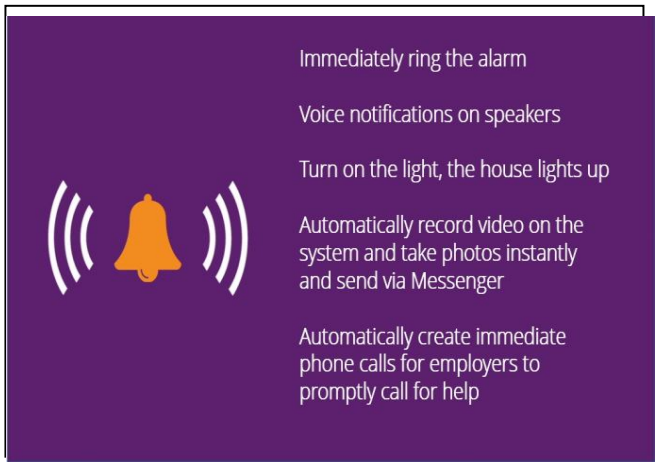
B. *Automation tasks*

A smart home could not be smart enough if it doesn't automatically respond or make the tenant comfortable.

By the help of temperate sensor, while you're inside or outside the house, your air conditional will proactive adjust the amount of energy waste but still keep your house fresh and in an ideal temperature. There are also options which you will receive notification on your phone when you have visitors. Remember the day when you have a phone call but your TV still work loudly? Now you don't need to turn it off anymore because your assistant had done it for you!... Home Assistant is the definition real automatic features where you feel the comfortable follow every action

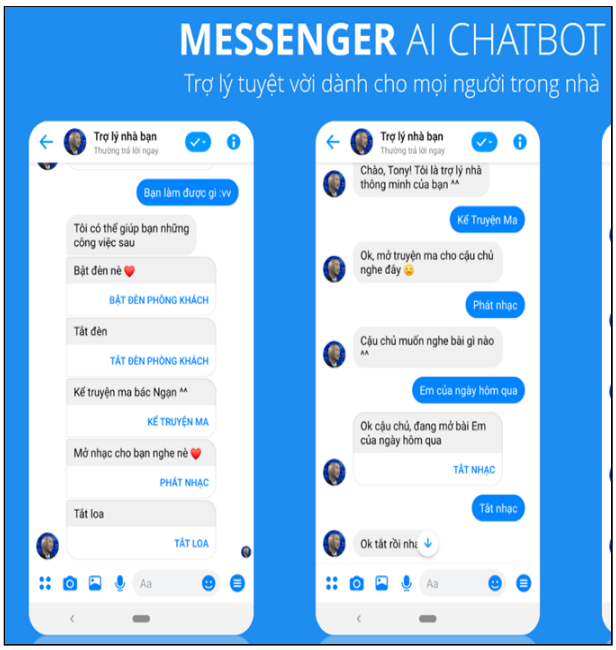
C. *Security and notification*

Safetiness and privacy always placed on top. While false alarm or wrong people recognition is happening ,which disturb you,and AI face recognition were not complete nor suitable for some people's wallet ;we attempt for more considerable solution. By the connection of your smartphone,which almost stay with you ,with devices trough wifi , Home Assistant are now can discriminate between you,your family members and your "uninvited" guess.Hence,the security alarm will work effectively without annoying your family.

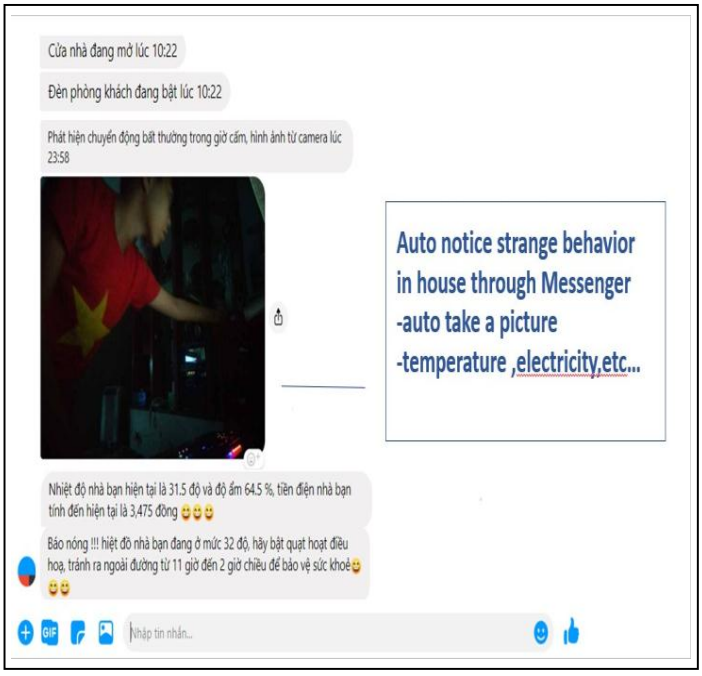


D. *AI chatbox with Facebook messenger and SMS*

Lastly,to achieve the most convinient and flexible of the system,you are able to communicate with your assistant through Messenger and SMS (incase there is no wifi).the process is simple : chatbox will give you options for the very first time you text her/him for controlling your house's devices (light,air conditioner, fridge,etc..).Through time,the system will save your activity, analysis them,then automatically adjust most used devices congruously.

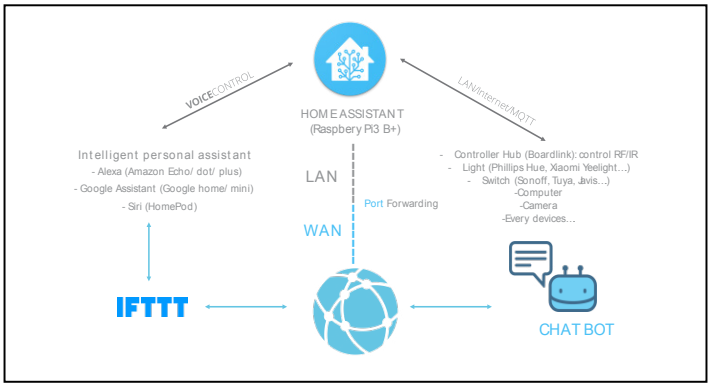


"Easy to use for every family member"

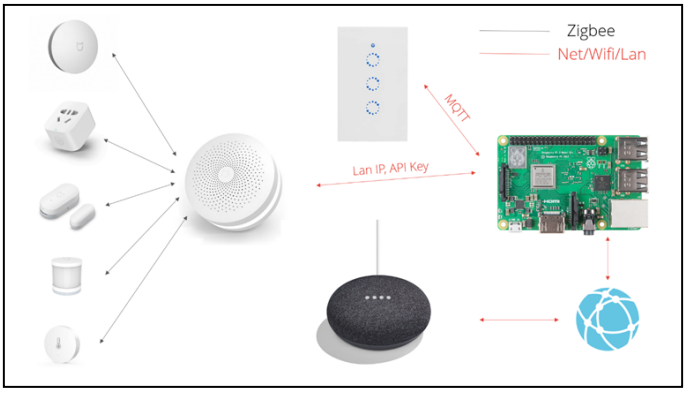


IV. How it works

Regarding the connection protocol, all IoT devices are linked to Raspberry Pi 3b+ running on the Home Assistant platform, Pi will play the roll as the brain of the house to control IoT devices, IoT devices like WI-FI switch, sensor, TV, Google Home all connect to Pi in LAN to gain the feedback speed among 100ms. For automatic devices on the phone (for example: Phone to temporarily stop watching TV programme, announcing phone contacts via speaker, SMS ……), we use phone help application to transmit HTTP Request commands to IP of HOME ASSISTANT to control indoor devices when the phone changes status, sensors, etc. Similarly, to customize the voice, google answers we use IFTTT to implement webhook, HTTP Request to control, just like that, to control the home with chatbot, we use chatfuel, to execute the HTTP Request command. The notification by IFTTT bot connects to HOME ASSISTANT using API that HOME ASISSTANT supports.

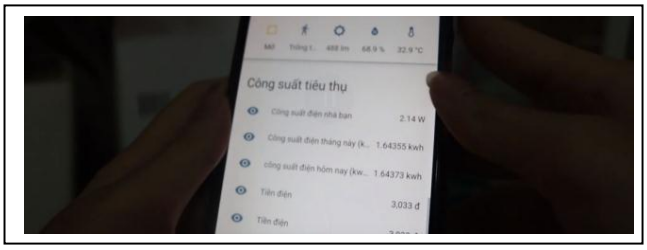


Home Assistant platform is installled on Raspberry Pi 3b+ and connect to all devices via Wifi/Lan . The controls will be send to IFTTT app services



V. Demo

Please check out our demo on this link:
https://youtu.be/Ovm2wuiXPy8



THANK YOU!

Home Automation – IoT Showcase Contest

Tran Quoc Tuan
Programming Mobile Devices
The college of FPT Polytechnic in Da Nang
193 Nguyen Thi Thap – Hoa Minh – Lien Chieu – Da Nang
tuantqpd02578@fpt.edu.vn

Nguyen Ngoc Duy
Programming Mobile Devices
The college of FPT Polytechnic in Da Nang
193 Nguyen Thi Thap – Hoa Minh – Lien Chieu – Da Nang
duynnpd02638@fpt.edu.vn

----------***----------

Abstract *- Home Automation is a smart home development model to modernize your home. Based on the actual needs of everyone. Home Automation is easily developed because it is low cost and user-friendly. Home Automation uses user-friendly technologies to send results to different devices such as mobile phones, tablets, or computers connecting WIFI, or SMS messaging via mobile wave.*

Keyword *- Home Automation, IoT Smart Home, Android IoT Home Automation.*

I. INTRODUCTION

Nowadays, with the continuous development of technology 4.0. Along with the development of mobile applications and the trend of automating electronic devices, replacing people with everyday tasks. Since then we have applied and developed the Home Automation product to meet those needs. This product will help us create automated homes that match our current needs. For examples, it can automatically turn off the light according to temperature, turn on the fan when the room temperature is too high, notice of fire detection, anti-theft, gas leak,.. All things are communicating with people through mobile application. Hopefully, these models will help future houses become smarter.

II. LITERATURE SURVEY

In current situation, there are many IoT projects developed. Through many practical surveys, we find that Home Automation is still one of the best and always expected projects.

A. Survey via website infoq.vn[1]

Most of the answers on the site give results that 80% of the users want Home Automation to completely replace everything. 15% said they don't have enough money to choose Home Automation. And only 5% do not agree with that.

B. Survey via FPT Polytechnic Da Nang[2].

Most of the students in the school want to have Home Automation System to solve their needs. In addition, they want to develop for their own homes.

III. MOTIVATION

In some cases, there may be older or handicapped people who are not able to move frequently for controlling appliances in the house, so home automation system can help these people to easily control all the appliances because it is simple and easy to use.

Home Automation also improves the standard of living. It provides an easy, flexible and less interactive interface with all low-cost functions and in flexible environment.

IV. PROPOSED SYSTEM

In this system, the android application used to send signals to the MQTT [3] facility is connected to the NodeMCU V3[4] module that provides this signal to the system to control indoor devices.

A. System Architecture

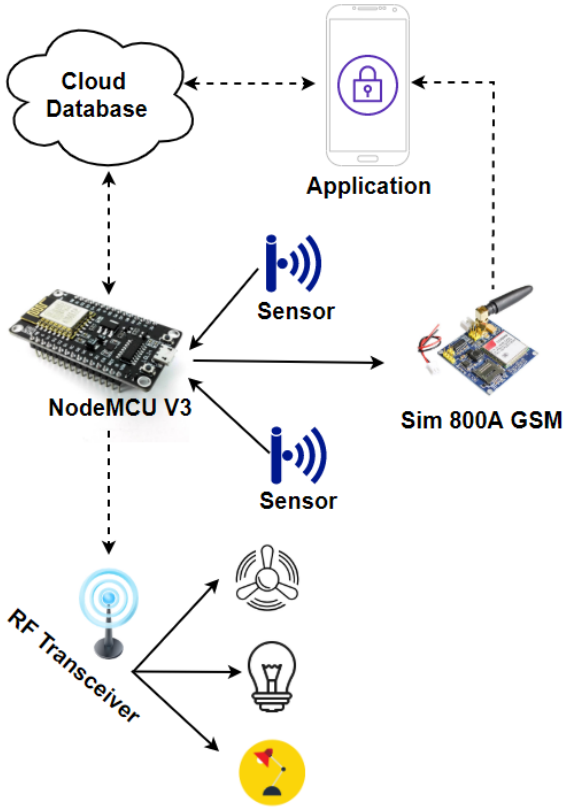


Fig- 1: Operating model

1) Android application in mobile sends the signal to the module NodeMCU V3 which is connected to module ESP8266[5] Wifi via MQTT protocol in the same network. Android application has all the GUI[6] buttons for each appliances.

2) Wi-Fi module receives the signal from the mobile application and gives this signal to NodeMCU V3 for processing.

3) We use NodeMCU V3 as a controller to control all the appliances. Each command is processed by

NodeMCU board and control RF receiver and transmitter[7] for switching on/off the appliances.

4) RF receiver and transmitter use as electrical switches, for performing on/off operation.

5) The sensors will send signals to the NodeMCU module, then when the sensors change, the signal will be sent to the SIM 800A module[8] and it will send a message to the user by SMS. As well as sending a notification to Android apps.

6) Finally, users can completely access Android apps on mobile devices and control your home in the most convenient and easiest way.

B. Mathematical Model

We use set theory for representing our system

Let [H] be the System

H = *{Start, End, Input, Output, Send (), Receive ()}*

Where,

Start = Start state of the system.

Input = Input command for the system. *{Sensor, GUI Buttons, ...}*

Output = Output command for the system. *{Output 1, Output 2, Output 3, Output 4,}*

Where,

Output 1 = Operation for Fan. *{Fan1, Fan2...}*

Output 2 =Operation for Lamp. *{Lamp1, Lamp2..}*

Send () = is a function for sending the commands to Android Apps and send the result for the android app when the sensor changing.

Receive () = is a function for receiving the commands from the android application and receive a result of the sensor.

End = End State of system.

V. IMPLEMENTATION DETAILS

A. Software Implementation

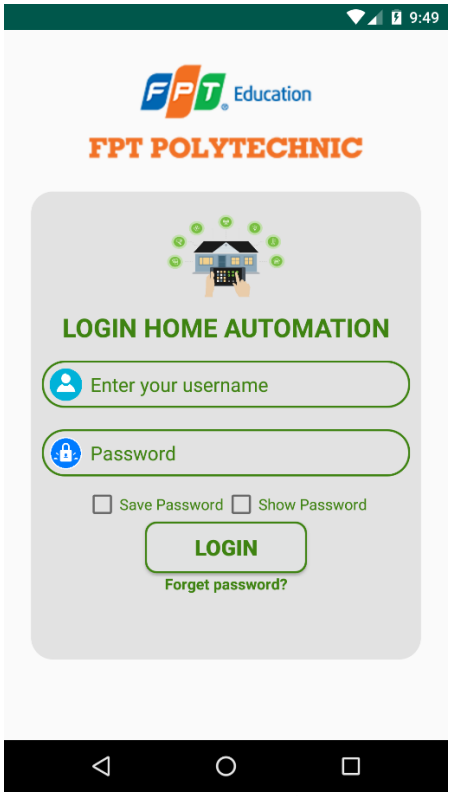


Fig- 2: Login Application

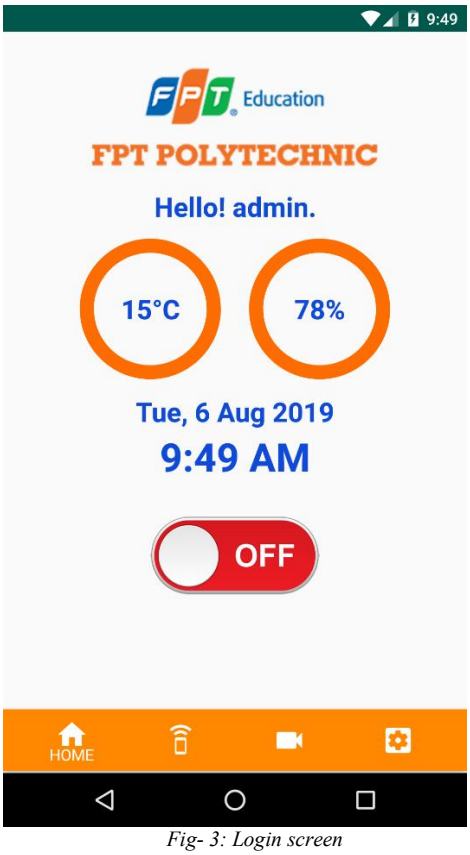


Fig- 3: Login screen

Friendly screen displays temperature and humidity in the house. Show date and time. And there is a button to on/off applications.

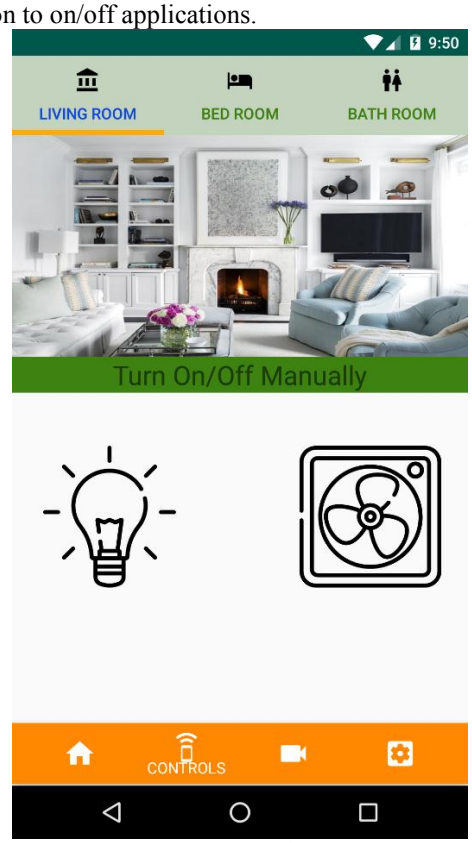


Fig- 3: Controls screen

Monitor controls the devices in the house. Display in an intuitive way easy to use.

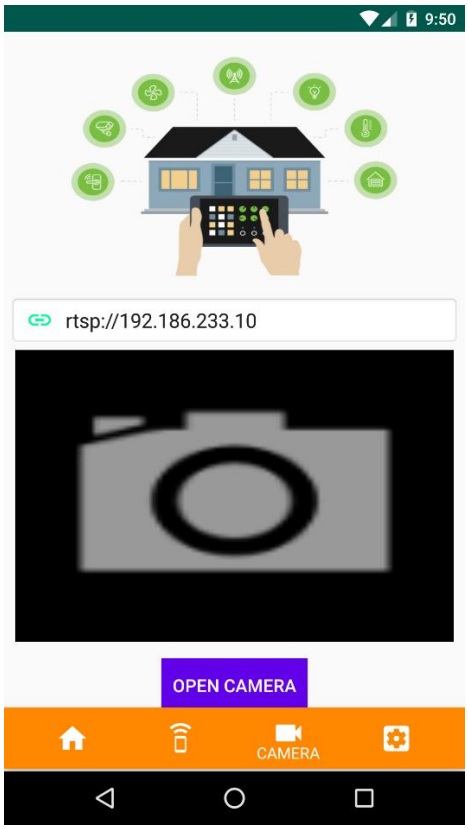


Fig- 4: Camera screen

This mode helps us to observe the whole house when needed, as well as when there is an alarm.

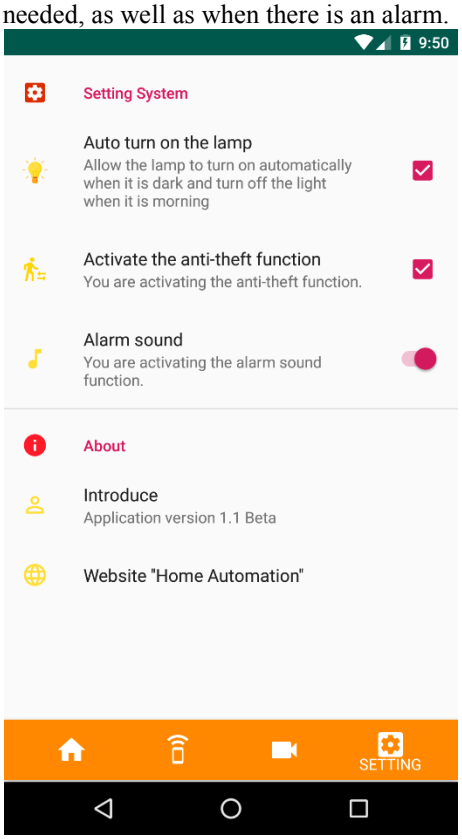


Fig- 4: Settings screen

The entire application settings screen such as automatically turn on and off the light, or turn off the sound...

B. Hardware Implementation.

In the hardware implementation, we are using NodeMCU V3 as a controller. We use Arduino IDE [9] software to write programs for this module to work. It adopts a wifi module ESP8266 directly attached to the board to work. ESP8266 Wi-Fi module is used for communication between the android mobile app and NodeMCU V3 via MQTT protocol. This through protocol we connect between Android Application and NodeMCU V3 to handle the given requests as well as receive information from sensors to send data back and forth to the application as well as the NodeMCU V3 processor. If having request from the application, it will be sent to the cloud and ask the processor NodeMCU V3 to make that request. When the sensors have an impact, it sends the results to the processor and requests execution.

VI. FUTURE SCOPE.

In our system, we have provided Android software but have not had IOS application yet. As well as the software interface is also not optimized UI / UX. Therefore, we will continue to upgrade the software to expand more functions and will expand many platforms such as Website, IOS App in the future.

VII. CONCLUSION

Home Automation is not a new IoT product but it's never outdated. Home Automation is indispensable in every home. Home Automation System provides interface between various types of home and electrical appliances like lamp and fans etc. Users can control and use easily the appliances they need. Moreover, users can save energy and manage their home flexibly when using Home automation system

VIII. ACKNOWLEDGEMENT

In addition to the efforts of the group to develop this product, we would like to thank our internal guide teacher. Teachers in the IT department Mr. Nguyen Van Nam and Mrs. Vu Thi Thanh Huyen for their valuable and helpful suggestions.

Special thanks also to Mr. Nguyen Khanh - Deputy Director of FPT Polytechnic College for his indispensable support and suggestions.

Thank FPT Education Organization for organizing this useful playground.

REFERENCES

[1] InfoQ.vn "This is one of Vietnam's leading community of online market survey" Available in: http://www.infoq.vn

[2] FPT Polytechnic Da Nang "The College Of FPT Polytechnic in Da Nang" Website in: https://caodang.fpt.edu.vn

[3] MQTT Cloud "Message Queuing Telemetry Transport" Wikipedia: https://en.wikipedia.org/wiki/MQTT

[4] NodeMCU V3 "This is an open source IoT platform." Wikipedia: https://en.wikipedia.org/wiki/NodeMCU

[5] Module WIFI ESP8266 "The ESP8266 is a low-cost Wi-Fi microchip with full TCP/IP stack and microcontroller capability" Availabe in: https://github.com/esp8266/Arduino

[6] GUI "Graphical User Interface" Available in: https://developer.android.com/guide/topics/ui

[7] RF Receiver and Transmitter "Radio Frequency Receiver and Transmitter Module" Available in: https://en.wikipedia.org/wiki/RF_module

[8] SIM 800A module "SIM800 module is a GSM quad band module." Available in: https://wiki.eprolabs.com/index.php?title=SIM_800A_GSM_GPRS_Module

[9] Arduino IDE "Arduino integrated development environment" Website: https://www.arduino.cc/en/Main/software

KIDOOS: Kid Obesity Observation System

Muhamad Hariz Adnan
Computer and Information Sciences
Universiti Teknologi PETRONAS
32610, Seri Iskandar, Perak, Malaysia
m.hariz_g03017@utp.edu.my

Mohd Fadzil Hassan
Center for Research in Data Science(CeRDaS),
Computer and Information Sciences
Universiti Teknologi PETRONAS
32610, Seri Iskandar, Perak, Malaysia
mfadzil_hassan@utp.edu..my

Izzatdin Abdul Aziz
Center for Research in Data Science(CeRDaS),
Computer and Information Sciences
Universiti Teknologi PETRONAS
32610, Seri Iskandar, Perak, Malaysia
izzatdin@utp.edu..my

***Abstract*—Many ways and techniques such as assessment of body composition and data mining techniques have been applied to predict childhood obesity, but only a few managed to produce accurate results due to lack of reliable data. IoT devices can be used to collect lifestyle data from the kids. The data will be processed and combined with survey data to make predictions. A system of kid obesity observation system consisting of children lifestyle data collections, predictions of obesity and overweight risks using artificial intelligence and lifestyle monitoring and preventive plans is proposed.**

Keywords—internet of things, childhood obesity, prediction, data mining, artificial intelligence, obesity risk factors

I. Introduction

According to the United Nation (UN), The number of overweight children under five rose 38% between 2000 and 2016 in the region. It has become a worrying global epidemic. Obesity can be defined as an excessive accumulation of body fat [1]. Evidence show that childhood obesity persists into adulthood. With advances in technology, people are becoming less concerned about living a healthy life.

The calculation of Body Mass Index (BMI) is a method that is used to determine child obesity. A child with BMI at or over the 97th percentile is considered obese [2]. Prediction means an indication in advance based on observation, experience, or scientific reason [3]. A childhood obesity prediction system classifies young children who are at the risk of obesity. The system indicates in advanced based on their early growth records and other factors.

The risk factors of childhood obesity are genetic, parental obesity, sedentary lifestyle, and high sugar intakes [4]. These risk factors could be obtained from the children, their parents, other family members, medical experts, articles, and health organization. Other than the risk factors, congruent techniques for prediction should be chosen wisely. Current achievements on childhood obesity prediction can be labeled as unsatisfactory and far from achieving the objective [4]. There is a lot of weaknesses that could be improved in the current obesity prediction system. One of the weaknesses in childhood obesity prediction systems is in the data collection.

The objective of this project is to develop a system targeted to kids where the system can collect lifestyle and health-related data using IoT devices and survey; identifies significant risk factors and protective factors of childhood obesity, and predict the risk of the kids to be obese and overweight. The system is Kid Obesity Observation System (KIDOOS).

The paper is organized as follows: Section 2 discusses the related work, section 3 discusses the lifestyle factor and Section 4 presents the proposed KIDOOS. In Section 5, the conclusions are presented.

II. Related Work

A. Utilization of Data Mining for Kid Obesity Prediction

Data mining is the process of analyzing and summarizing data from different perspectives and converting it into useful information [5]. In a large database, data mining is employed to detect patterns to extract hidden pieces of information [6]. Many data mining techniques have been applied to childhood obesity prediction, including decision trees, association rules, Artificial Neural Network (ANN), Bayesian networks, Naïve Bayes and Support Vector Machines (SVMs) [1, 7, 8]. The ANN, Naïve Bayes and decision tree were identified as suitable methods from related work [9].

ANN, Naïve Bayes classifier, and decision tree are data mining techniques that are applicable for childhood obesity prediction. For ANN, environment is considered as an important factor for childhood obesity prediction. The prediction result can be more promising because environment undeniably plays major role in a child's growth. The Naïve Bayes classifier focuses more on the child's personal attributes rather than the environment. However, the result of Naïve Bayes classifier has been proven acceptable for childhood obesity prediction [10].

For the decision tree, we see that the technique is easier to build than the other two methods since various attributes can be used for childhood obesity prediction. The three data mining techniques are compared in Table 1.

In the next section, the parameters identification and selection for KIDOOS is discussed.

TABLE I. Comparison of Data Mining Techniques [9]

Technique	*Strengths*	*Weaknesses*
Artificial Neural	1. Can handle a large amount of input data 2. Fault tolerance and low energy consumption	1. Efficiencies (error vs. Network size) are problem dependent 2. The run time speed may be slow when many problems are involved
Naïve Bayes	1. Accurate when sample attributes are independent 2. Simple and computationally efficient	1. High bias 2. Less accurate and performance degrades when attribute is dependent
Decision Tree	1. Fast learning and classification speed 2. Results in human understandable rules	1. Prone to error if there are many classes and a relatively small

Technique	*Strengths*	*Weaknesses*
		number of training examples 2. Can be computationally expensive to train

B. Kid Obesity Factors Identification and Selection for KIDOOS

The factors can be divided into three groups: children factors, lifestyle factors, and family/environment factors [10]. Children factors are obesity risk factors and protective factors that relate to the child himself, such as characteristics, behavior, and health [10]. Lifestyle factors are related to the child`s lifestyle, such as watching television, playing computer or video games, engaging in physical activities, daily nutrition, and the way the child is raised [10]. Family/environment factors relate to the child's family and environments, such as parental obesity, parents' education, family income, number of siblings, and the living environment [10]. The selected factors are summarized in Table 2.

TABLE II. SELECTED FACTORS FOR CHILDHOOD OBESITY PREDICTION [10].

***Category* Parameters**	***Category Parameters***
Children	Catch-up growth, adiposity rebound, premature birth, and gender
Lifestyle	Duration of breastfeeding, duration of sleep, eating junk food, eating fried food, eating fruit, eating snacks in front of TV, duration of watching TV, eating warm meals for supper, physical activity, eating soup and sandwich bought outside home, and eating snacks and chocolates bought outside home.
Family/environment	Mother BMI, father BMI, parents' weight and the number of children.

The KIDOOS IoT device will be focusing on the children lifestyle factors. The next section discusses the lifestyle factor data collections for KIDOOS.

III. LIFESTYLE FACTOR DATA COLLECTIONS FOR KIDOOS

Sedentary behaviors are factors of childhood obesity and overweight [11]. Sedentary behaviors such as watching television (TV), playing computer, and playing video games are risk factors of childhood obesity because of their negative effect such as reducing physical activities and increasing junk food consumptions [11]. It was also stated that children who watch television for more than 4 hours per day have a significantly higher BMI than those who watch television lesser than 2 hours per day. It was found that obesity odds ratio for 7 years old children that watched television more than eight hours per day is 1.55 and between four to eight hours is 1.37 [12].

It was identified that excessive time spent in watching television can demote physical activities, decrease resting metabolism in children, and bad advertisement can influence the children to eat unhealthy food [13]. Playing computer and video games indicated a lower risk of childhood obesity but possibly as much risk if regular physical activities are demoted [14].

Short sleep duration might increase the risk of childhood obesity [10]. 3 years old children who slept less than 10.5 hours or between 10.5 to 11 hours daily have 1.45 and 1.35 odds ratios of obesity at age 7 respectively [12]. The children need at least 9 hours or more than 11 hours of sleep daily [12, 15]. The effect of short sleep on children such as causing tiredness, increases child hunger and appetites, and accelerate their weight gain [15].

Physical activity is a protective factor of childhood obesity and overweight. One hour of moderate and vigorous physical activities a day could reduce childhood obesity risk by 10% [16]. Lack of physical activities such as swimming, cycling, playing outdoor, jogging, running and joining sports activities at school could result in unbalanced energy intake and energy consumption in children [10].

The sedentary behaviors, sleep duration and physical activity factors are inter-related. The data can be collected using IoT devices. The device used for this project is shown in Fig. 1 and Fig. 2.

Fig. 1. The device used for lifestyle factor data collection.

Fig. 2. The devices are attractive and suitable for kids.

The device is Mi Band 3. The device is selected because it has the functionality needed to collect the data and the price is lower (RM 90) compared to other fitness tracking devices in the market. Fig. 3. shows the data collected from the tracking device.

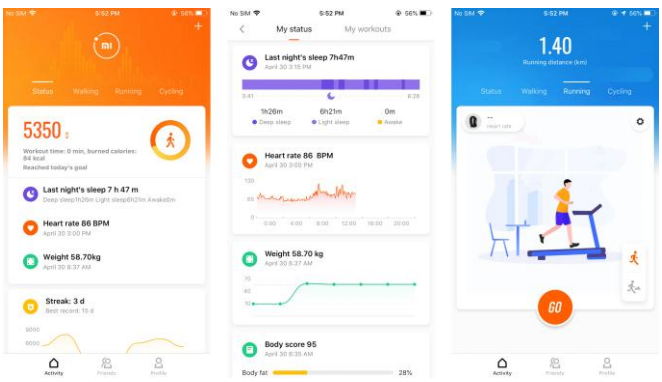


Fig. 3. The data collected from the device.

IV. KIDOOS: KID OBESITY OBSERVATION SYSTEM

KIDOOS is designed for three main functions, that are to collect the data, make prediction using data mining and artificial intelligence techniques, and provide monitoring and preventive plans. The proposed architecture of KIDOOS is shown in Fig. 4.

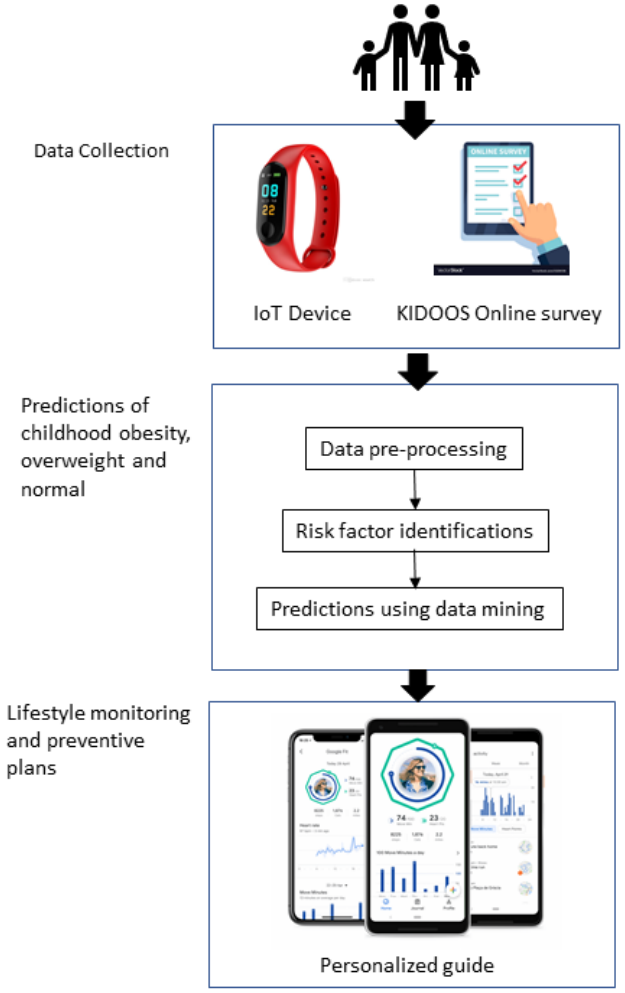


Fig. 4. The architecture of KIDOOS.

A. Data collection

Initial data have been collected from primary school children age 9 to 11 years old in Malaysia using questionnaires. The schools that were chosen as participants are located in a suburban area and an urban area. A total of 140 questionnaires were given to randomly selected children for the children and their parents to answer. Most of the data collected consisted of historical data. The target group was primary school children that were old enough to properly answer the questionnaires with the help of their parents. The data obtained from the questionnaires were processed and cleaned, including the removal of missing values and distortions, to identify suitable candidate risk factors and protective factors [17].

The system is designed to collect more data from the children and parents through online questionnaires. In addition, the data are collected from the IoT devices, as shown in Fig. 5.

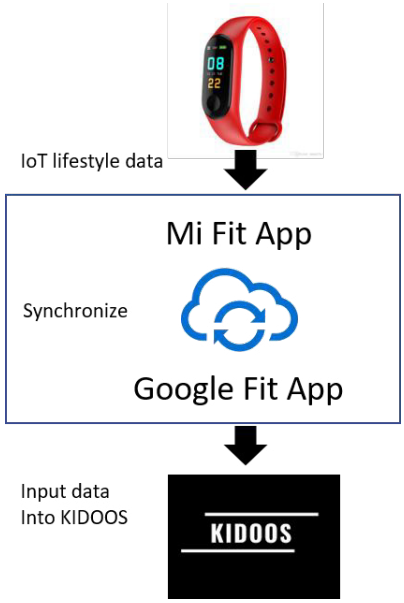


Fig. 5. The architecture of KIDOOS.

B. Predictions of Childhood Obesity, Overweight and Normal

The prediction is made using hybrid data mining techniques. Currently, the system uses Classification and Regression Tree (CART) to select important variables based on its relative importance [18]. The variables selected by CART are shown in Fig. 6 [18].

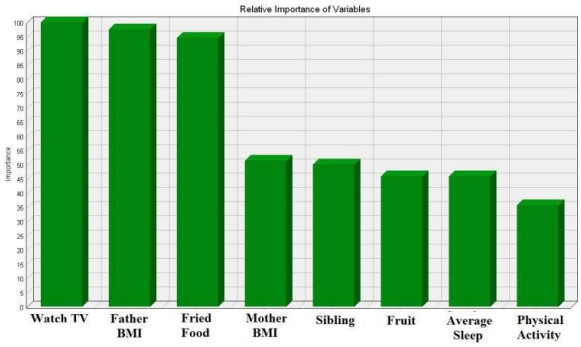


Fig. 6. Variables selection using CART.

Then, predictions will be made based on the identified variables using Naïve Bayes, means value identification and Euclidean Distance classifications. The architecture of this approach is shown in Fig. 7. The outputs of the Naïve Bayes were clustered into the positive (true) and negative groups (false) for each prediction class.

C. Lifestyle Monitoring and Preventive Plans

The system will provide a lifestyle monitoring platform and preventive plans to reduce the risks and avoid obesity and overweight. This feature is future work.

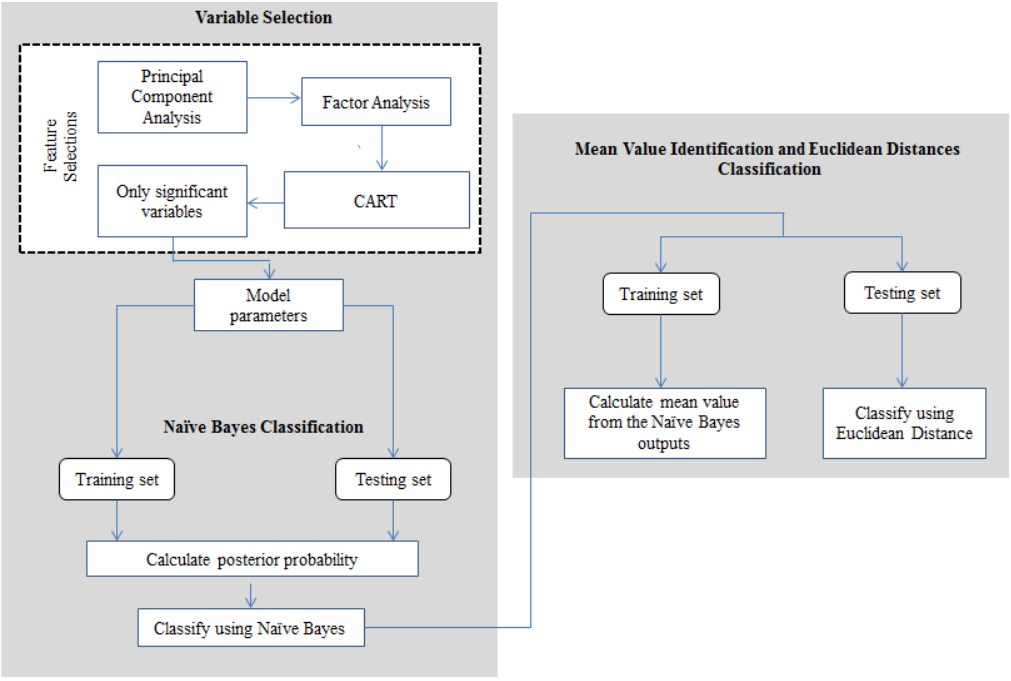


Fig. 7. The architecture of KIDOOS prediction algorithm [18]

V. Results

The results of normal, obesity, and overweight predictions using the proposed algorithm are shown in Table 3.

TABLE III. The Results of Predictions [18]

Prediction	Sensitivity (%)	Specificity (%)	Accuracy *(%)*
Normal	60	100	83
Obesity	60	78.6	70.8
Overweight	95	50	58.3

Based on the results shown in Table 3, the sensitivity of KIDOOS algorithm for normal, obesity and overweight predictions are 60, 60 and 95 percent respectively. For childhood obesity prediction sensitivity is more important than the specificity [19]. By collecting more data, enhancing the algorithm and incorporating machine learning with KIDOOS, the sensitivity, specificity and accuracy of KIDOOS can be improved.

VI. Conclusion

KIDOOS is a system that aims to collect lifestyle and health-related data using IoT devices and survey and use Artificial Intelligence techniques to predict the risk of the children to be obese and overweight. The system will provide a monitoring platform of the children lifestyle obesity factors and suggests preventive plans to the parents. In the future, the system will be developed into downloadable android and IOS mobile applications for commercialization.

Acknowledgment

The authors would like to thank the Universiti Teknologi PETRONAS and Universiti Sains Malaysia for supporting this project, and to FPT University Research Festival (ResFest) 2019.

.References

[1] S. Soni and J. Pillai, "Usage of nearest neighborhood, decision tree and Bayesian classification techniques in development of weight management counseling system," in *2008 First International Conference on Emerging Trends in Engineering and Technology*, 2008, pp. 691-694: IEEE.

[2] A. Körner *et al.*, *Obesity in Childhood and Adolescence: a review in the interface between adipocyte physiology and clinical challenges*. 2005, pp. 189-199.

[3] C. Park, *A dictionary of environment and conversation",* (in A Dictionary of the Internet.). Oxford University Press, 2007.

[4] M. H. M. Adnan and W. Husain, "A framework for childhood obesity classifications and predictions using NBtree," in *2011 7th International Conference on Information Technology in Asia*, 2011, pp. 1-6: IEEE.

[5] Q. Luo, "Advancing Knowledge Discovery and Data Mining," in *First International Workshop on Knowledge Discovery and Data Mining (WKDD 2008)*, 2008, pp. 3-5.

[6] A. Linos, I. Kotsioni, and E. Papageorgiou, "Data mining: a new technique in medical research," (in eng), *Hormones (Athens),* vol. 4, no. 4, pp. 210-212, Oct-Dec 2005.

[7] B. Novak and M. Bigec, "Application of artificial neural networks for childhood obesity prediction," in *Proceedings 1995 Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems*, 1995, pp. 377-380.

[8] S. Zhang, C. Tjortjis, X. Zeng, H. Qiao, I. Buchan, and J. J. I. S. F. Keane, "Comparing data mining methods with logistic regression in childhood obesity prediction," journal article vol. 11, no. 4, pp. 449-460, September 01 2009.

[9] B. M. A. Muhamad Hariz, H. Wahidah, and D. Faten, "A survey on utilization of data mining for childhood obesity prediction," in *8th Asia-Pacific Symposium on Information and Telecommunication Technologies*, 2010, pp. 1-6.

[10] M. Adnan, W. Husain, and N. Rashid, "Parameter identification and selection for childhood obesity prediction using data

mining," in *2nd International Conference on Management and Artificial Intelligence*, 2012, vol. 35, pp. 75-80.

[11] N. F. Krebs and M. S. J. P. Jacobson, "Prevention of pediatric overweight and obesity," vol. 112, no. 2, pp. 424-430, 2003.

[12] J. J. Reilly *et al.*, "Early life risk factors for obesity in childhood: cohort study," vol. 330, no. 7504, p. 1357, 2005.

[13] G. P. August *et al.*, "Prevention and treatment of pediatric obesity: an endocrine society clinical practice guideline based on expert opinion," vol. 93, no. 12, pp. 4576-4599, 2008.

[14] J. P. Rey-Lopez, G. Vicente-Rodríguez, M. Biosca, L. A. J. N. Moreno, Metabolism, and C. Diseases, "Sedentary behaviour and obesity development in children and adolescents," vol. 18, no. 3, pp. 242-251, 2008.

[15] R. Kuriyan, S. Bhat, T. Thomas, M. Vaz, and A. V. J. N. J. Kurpad, "Television viewing and sleep are associated with overweight among urban and semi-urban South Indian children," vol. 6, no. 1, p. 25, 2007.

[16] C. B. Ebbeling, D. B. Pawlak, and D. S. J. T. l. Ludwig, "Childhood obesity: public-health crisis, common sense cure," vol. 360, no. 9331, pp. 473-482, 2002.

[17] M. Adnan, W. Husain, and N. Rashid, "Parameter Identification and Selection for Childhood Obesity Prediction Using Data Mining," in *2nd International Conference on Management and Artificial Intelligence*, 2012.

[18] M. H. Adnan, W. Husain, and N. A. A. Rashid, "Hybrid approaches using decision tree, naive Bayes, means and euclidean distances for childhood obesity prediction," *International Journal of Software Engineering and Its Applications,* vol. 6, no. 3, pp. 99-106, 2012.

[19] M. H. Adnan, "Childhood Obesity and overweight predictions using hybrid data mining techniques," Master of Science, School of Computer Sciences, University Sains Malaysia, Penang, Malaysia, 2013.

Magic Bin – smart bin for school and office – Automatically classify trash using machine learning

Duy Tran Le
FPT University
An Giang, Vietnam
duytlse140829@fpt.edu.vn

Hieu Nguyen Mau
FPT University
Daknong, Vietnam
hieunmse140832@fpt.edu.vn

Hoa Luu Dieu
FPT University
Tien Giang, Vietnam
hoaldse140830@fpt.edu.vn

***Abstract*— With the rapid development of society, the amount of garbage is more and more increasing, which means there are more consumption and waste of resources. Classifying and recycling are essential way to reduce waste sources. However, in order to have high effectiveness in cycling, the management and classification must done well at the beginning, at the source of waste generation. We have researched and built a device that helps sorting waste to be done automatically - smart bin Magic Bin**

Keywords—

IoT = Internet of Things
ML = Machine learning
CNC = Computer Numerical Control
IDE = Integrated development environment
BLE = Bluetooth low energy

I. Introduction

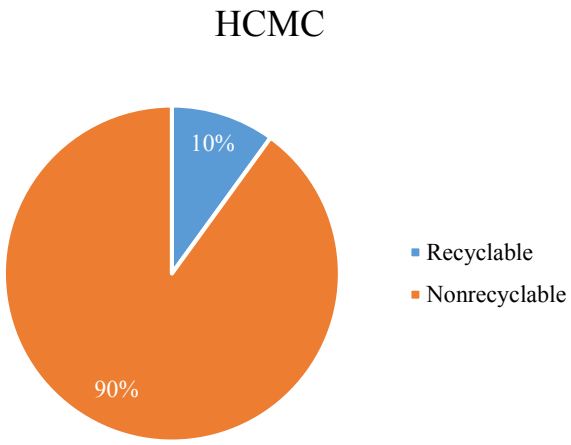
A. Purposes

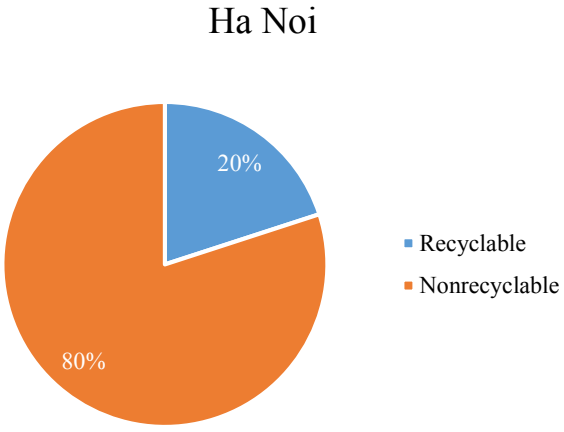
In Vietnam, the rate of waste generation depends on each urban type and ranges from 0.35 to 0.8 kg per person per day. Waste is an inevitable product of life discharged from production, business, service, living activities or other activities such as medical examination and treatment, entertainment of people. People's living standards have been increasingly improved and the industrialization process has been growing more and more deeply, waste has been created more and more with increasingly complex and diverse components. Process rubbish has become a hot issue in countries around the world, including Vietnam.

Major difficulties in inefficient waste management and treatment:

- Funding for investment in waste treatment, especially for hazardous waste is huge.
- Awareness of waste collection and treatment for officials and employees who directly do this work is not high, which greatly affects the quality and efficiency of classification, collection, transportation, handling and waste disposal.
- Law enforcement environment is not favorable although the Law on Environmental Protection has already enacted.
- Uncontrolled waste treatment solutions, interdisciplinary coordination is also ineffective in all stages of waste management.

According to statistics of environmental company *SACOTEC*. The proportion of waste that is undressed and not to be reused in Vietnam accounts for a large amount:





Recognizing the difficulties and practical needs, we have focused on researching and offering a smart trash solution to classify waste at source, making it easier to treat waste and save resources, labor power, time and money.

With the advancement of current technology, machine learning is one of the fields that are very developed to perform tasks automatically. We decided to integrate this technology into the trash system to automate the classification. At the

same time, developing the ability to create, evoke passion for learning and access new technologies of the times.

B. Research subject

In recent years, AI - Artificial Intelligence, specifically Machine Learning, is emerged as evidence of the fourth industrial revolution.

Machine Learning is a subset of AI. Machine Learning is a small area of Computer Science, it is capable of self-learning based on inputted data without having to be programmed specifically.

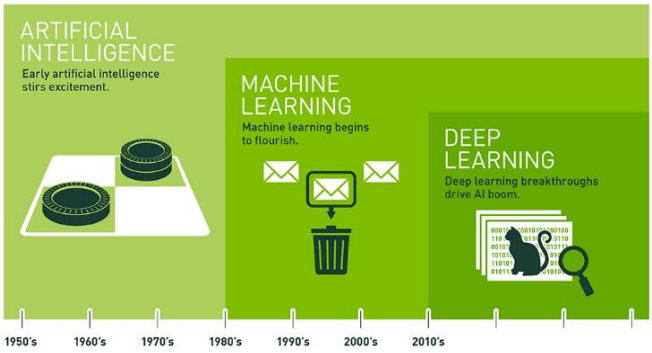


Figure 1. Machine learning

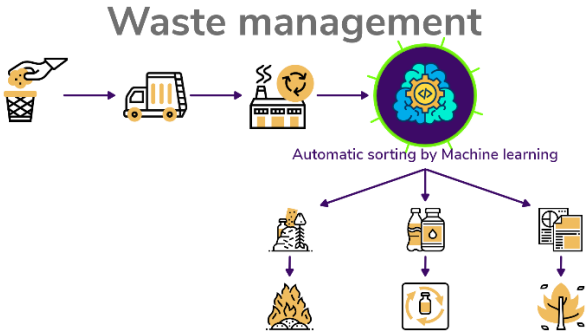
Recognizing the usefulness of this technology, we conducted research and programmed a system that is able to identify different types of garbage by using a large amount of garbage data. Through the results returned, we can consider to classify into recyclable and non-recyclable garbage.

Besides, IoT devices are an integral part of the system. Hardware includes sensor types and embedded system. We have studied, investigated and applied a number of sensors such as distance measuring sensors to detect and automate trash opening/closing door. CNC (Computer Numerical Control) is a high-precision motion axle system, so we have researched and built a similar system for trash bins to ensure accuracy in putting garbage in a suitable compartment.

C. Waste management in Vietnam

Currently, the waste management system in Vietnam is still simple and using labor power. The image below show that after the garbage collection vehicles collect garbage and move to factory, the trash must be classified by hand. However, due to lack of workforce in this field, the trash factory is overload. Therefore, it cannot classify all garbage.

Our solution is using robotic to classify trash automatically. Therefore, this problem can be solved by apply machine learning technology to classify trash quickly.

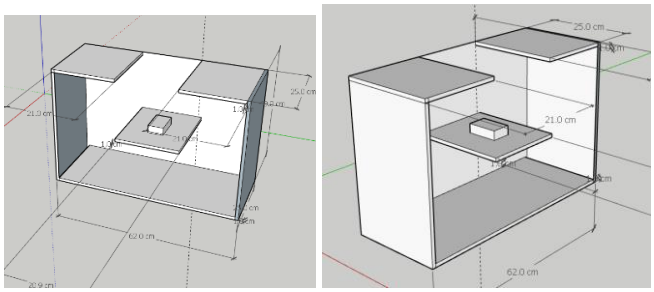


To demonstrate for our solution, we create a prototype – a bin can classify automatically trash at the source – i.e. at the point where people generate garbage.

D. Expected result

Expected prototype has some following main functions:

- Automatically open and close the door when detecting obstacles and no obstacles
- Distinguishing recyclable and non-recyclable waste into 2 compartments
- Locate the position of each bin - easy to collect and recycle garbage and for analysis purpose.
- Management application.
- Using solar energy helps to save energy in maintaining trash operation



II. IMPLEMETATION PROCESS

A. Hardware

1) Ideas

Firstly, we create a blueprint to calculate wooden frame sizes to match the size of other parts:

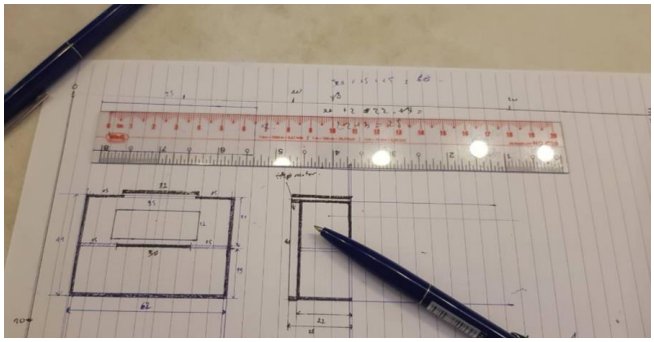
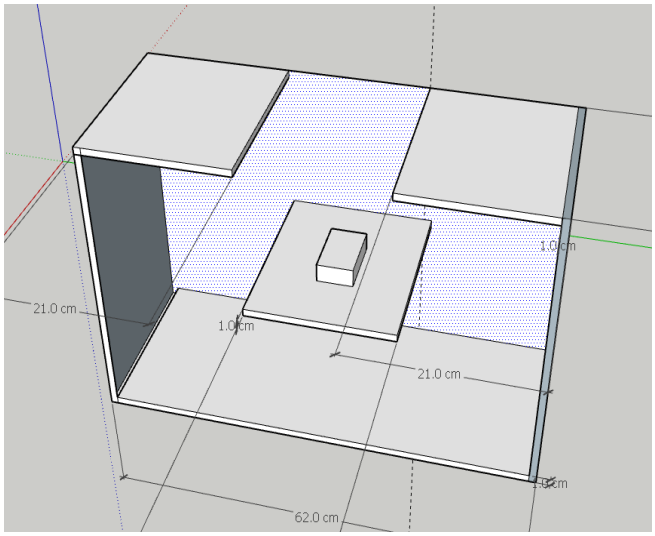


Figure 2. Blueprint

After that, we visualize blueprint in a computer with Sketchup software to get an overview of the product:



2) CNC system and stepper motor

Collect old printers that was broken but the CNC system still working. Our purpose is to control the door of the bin. Therefore, we reuses the CNC system of the old printer to make sure that our CNC system will run smoothly and correctly.

CNC: Numerical control is the automated control of machining tools

For the compartment movement system, we design a longer CNC system with maximum movement distance is 60 cm.

Figure 3. Step motor, shaft and ball bearing

Figure 4. Assembled movenment system

3) Frame

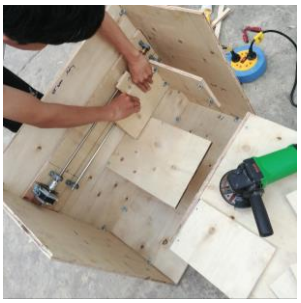


Figure 5. Measure and install wooden frame

Figure 6. Connect door controller and frame

Figure 7. Connect the door and controller; comparment and controller

4) Embeded system

We use module A4988 driver to control step motor. Microcontroller is *Arduino*.

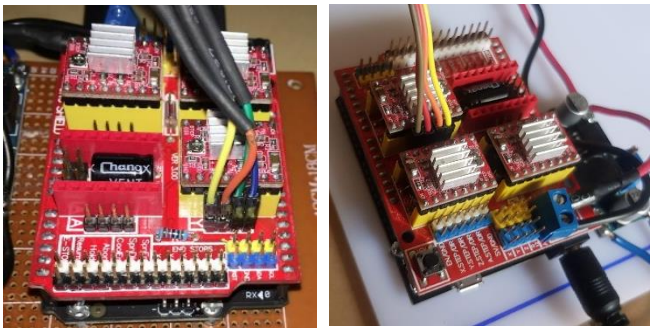


Figure 8. Stepper motor control circuit

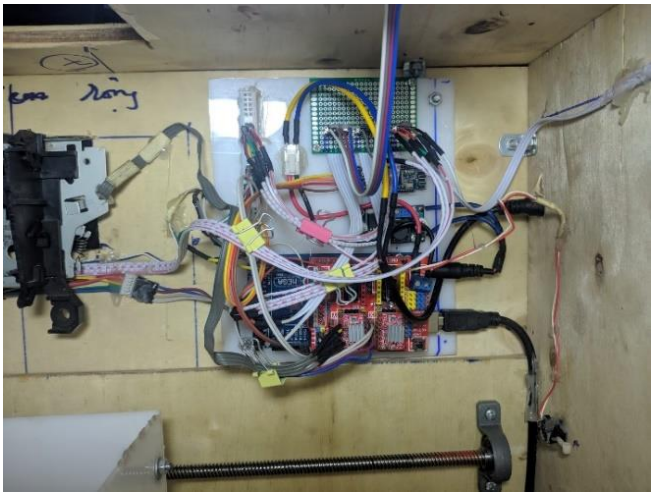


Figure 9. Assembling control circuit and wire into wood frame

5) Installing wheels and decorations

Figure 10 Lazer cutted label – Magic Bin

Figure 11. Install wheels

6) Upgrade CNC system

After running, we notice that our CNC system is not fast enough to move garbage to both side. Because the original system is use screw to pull and push the box, so it is slow. To solve this problem, we change to use pulley and belt.

Figure 12. Screw

Figure 13. Upgraded CNC - Pulley and belt

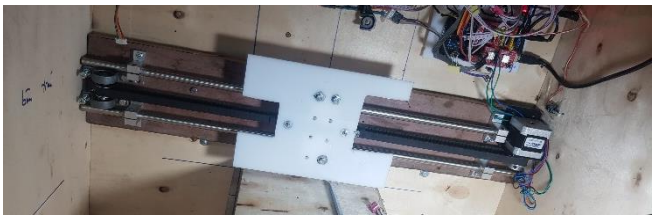


Figure 14. Assemble new CNC system

7) Decorate bin

Figure 15. Front side

Arduino: embedded system

Figure 16. Top side

B. *Software*

1) *Hardware control system*

Platform: *Arduino* and Android

Arduino is an open source hardware and software platform. Arduino hardware (microprocessor circuit boards) was born in the town of Ivrea in Italy, to build applications that interact with each other or with a more favorable environment. Hardware includes an open source circuit board designed on the Atmel 8-bits AVR processor platform, or 32-bit ARM Atmel.

Android is a mobile operating system developed by Google. It is based on a modified version of the Linux kernel and other open source software, and is designed primarily for touchscreen mobile devices such as smartphones and tablets.

In this system, *Arduino* plays as the hardware controller. *Android* phone is the core processor. After research, we notice that in this case, Android phone is the best device for controlling this system. We can use display, camera, GPS, Wi-Fi, Bluetooth and other hardware components. Moreover, in most case, Android phone is more stable than other embedded system like *Raspberry Pi*, *Arduino*. Thus, we decide to use *Android* as the main processor but it is program more complex than *Arduino*.

The local connection between Android and *Arduino* is Bluetooth 4.0 BLE, it is more security than other Bluetooth technology.

Mechanism of hardware system:

a) *Garbage classification*

Infrared sensor send data to *Arduino Mega* central processor. The microcontroller will check and get data. Then it calculates and controls the stepper motor to open or close door based on the signal, which it received from sensor.

After detect processing signal, it means that the bin has garbage.

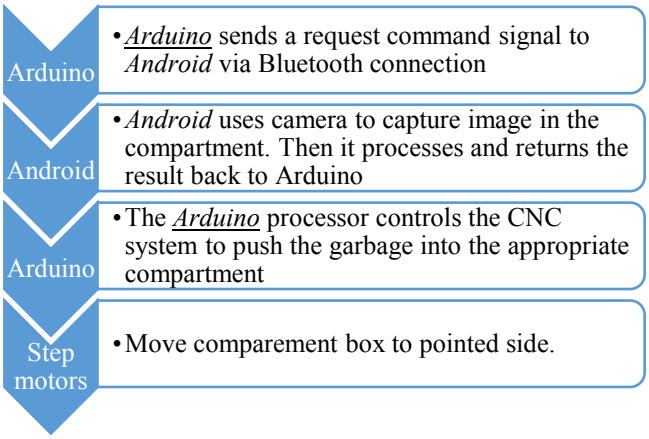


Figure 17 Garbage classification control flow

b) *Garbage monitoring*

The *Arduino* processor continuously reads the data of the ultrasonic sensor to measure the amount of garbage. Then it sends data to Android via Bluetooth BLE. Android will process and push the data to server. The server will store data send the message to the clients (management applications). Clients update the data, which displayed on the map.

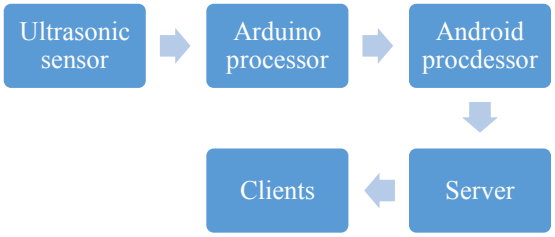


Figure 18 Garbage monitoring control flow

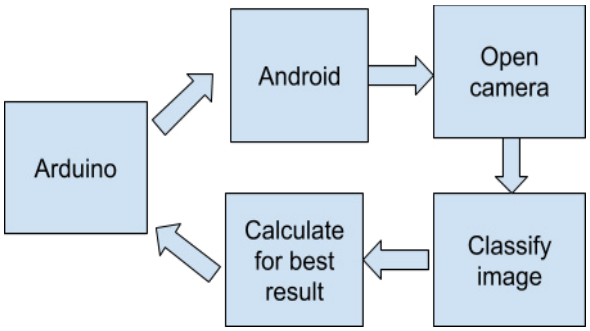
2) *Machine learning*

a) *Problem type: Image classificaiton*

Machine Learning is now one of the hottest topics around the world. However, to precise what Machine Learning is, well it's just one way of teaching the machine by feeding the large amount of data.

We create an Image Classifier, which can distinguish what type of garbage. We use one of the famous machine learning algorithms out there, which is used for Image Classification i.e. Convolutional Neural Network (or CNN).

There is control flow of classification system:



Infrared sensor: the sensor can measure distance accuracy
Ultrasonic sensor: the sensor can measure distance approximately

b) Data Preparation

For the dataset, we collect amount of garbage and capture it. After getting the trash datasets, we need to pre-process the data a bit and provide labels for each image.

In this case, we put these images in nine categories: bottle, can, plastic, glass, lunchbox, cigarette, paper, vegetable and nothing.

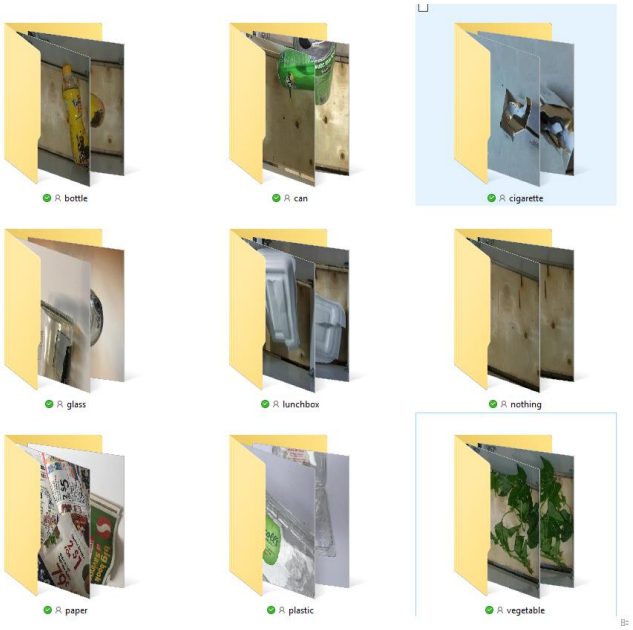


Figure 19. Trash categories

Figure 20. Bottles

Figure 21. Cans

We split dataset into three separate sets:

- 80% of images are used for training.
- 10% of images are used for validating
- 10% of images are used for evaluating the model. These images are not used in training

c) TensorFlow

Machine learning library – TensorFlow

TensorFlow is an open source software library for numerical computation using data flow graphs. The graph nodes represent mathematical operations, while the graph edges represent the multidimensional data arrays (tensors) that flow between them. This flexible architecture enables you to deploy computation to one or more CPUs or GPUs in a desktop, server, or mobile device without rewriting code. TensorFlow also includes TensorBoard, a data visualization toolkit.

d) Training the Model

There are some dataset versions, which we collected and trained:

Version 1

- Number of images of garbage types: 2412
- Data train time: 3 hours
- Objective: able to classify 6 types of garbage: plastic, glass, cardboard, paper, metal, trash.
- Disadvantages: accuracy is low, it confuses in determining between plastic and glass, metal, trash; between glass and metal, paper and metal; metal and glass.

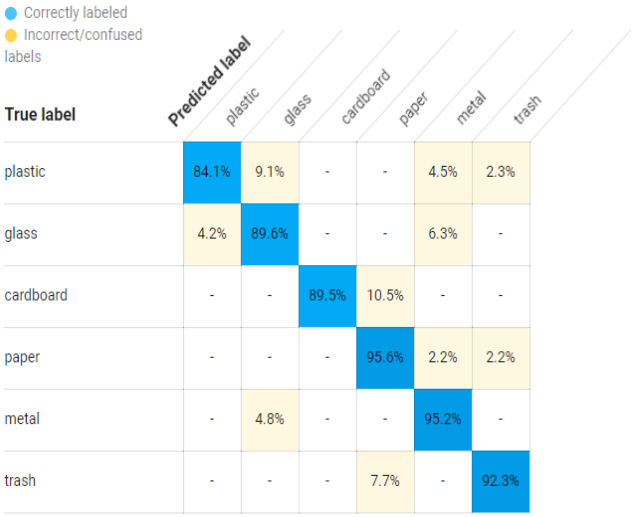


Figure 22. Dataset version 1

Version 2: Recognizing the inaccuracy of version 1, we have reduced some types of garbage to increase accuracy and increase the number of images of each type of waste.

- Number of images of garbage types: 2292
- Data train time: 3 hours
- Objective: able to classify 4 types of garbage: can, bottle, plastic, paper.
- Disadvantages: accuracy is low, it confuses in determining between can and bottle; between plastic and bottle; between paper and can, bottle.

Figure 23 Dataset version 2

Version 3: We increase a large number of images and remove some images.

- Number of images of garbage types: 5401
- Data train time: 6 hours
- Objective: able to classify 7 types of garbage: can, bottle, plastic, cigarette, lunchbox, nothing, paper.
- Advantage: high accuracy, usable, capable of detecting trash nothing.
- Disadvantages: it is confusing in determining between can and bottle, between plastic and bottle.

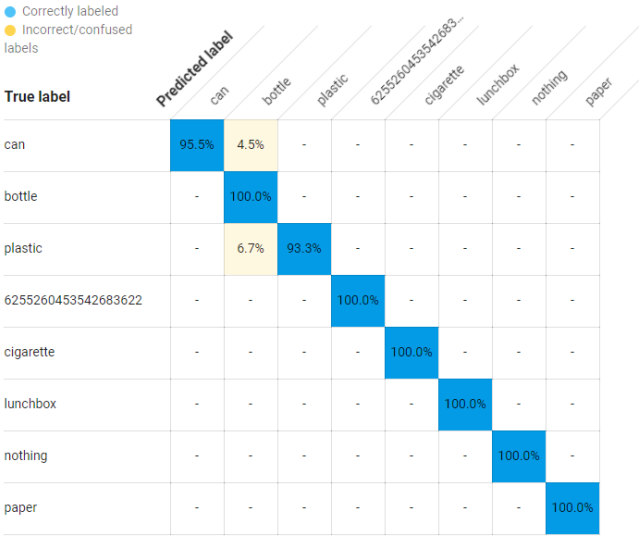


Figure 24. Dataset version 3

Version 4: Recognizing the incorrect data set in distinguishing the type of garbage bottle, we continue to update some bottle data.

- Number of images of garbage types: 5622
- Data train time: 6 hours.
- Objective: able to classify 7 types of garbage: can, bottle, plastic, cigarette, lunchbox, nothing, paper.
- Already able to distinguish 7 types of garbage.

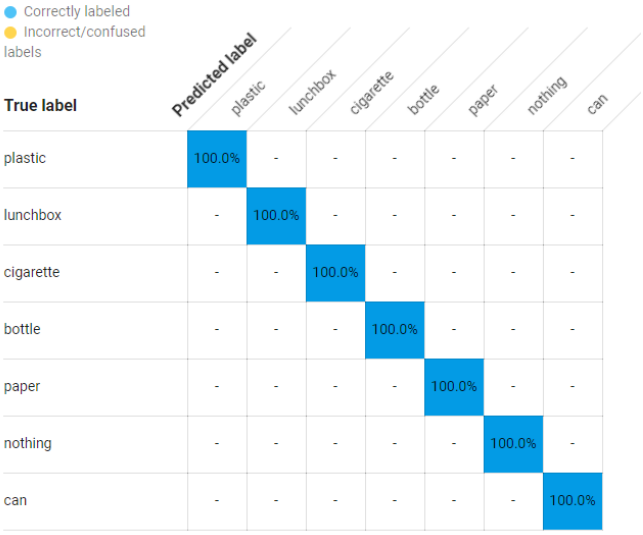


Figure 25. Dataset version 4

e) Model Evaluation - Testing

The machine learning system was able to identify most of the designated garbage, in many different conditions (distortion), the software still identifies correctly.

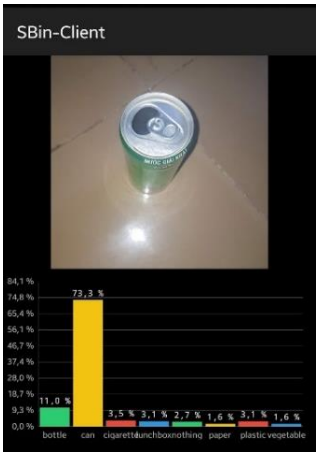


Figure 26. Result: can

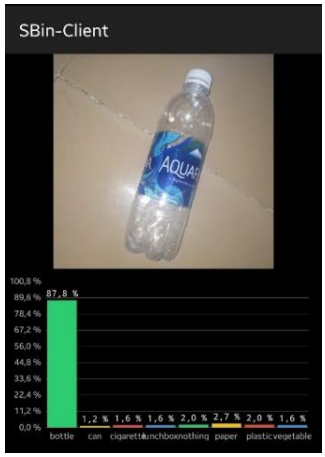


Figure 27. Result: bottle

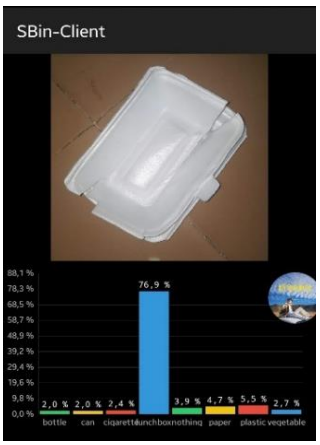


Figure 28. Result: Plastic

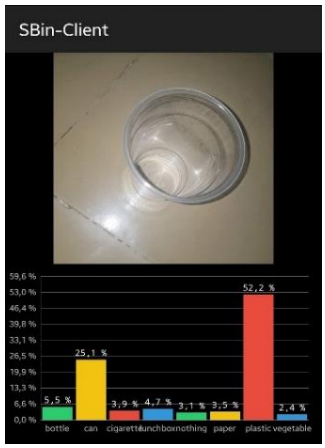


Figure 29. Result: lunchbox

3) Realtime database

a) Database management

We collect data of bin system in real time and we need update instantly in the management applications. After research, we choose Firebase as Realtime database for this project.

Firebase Realtime Database is a cloud storage database. Data is stored and synchronized in real time with each connected client, so that the management system can receive data of the trash almost immediately.

b) Database structure

There are three table in our database:

- Bin table contains some field: id (identifier), name (optional, name of the bin), module status, last active time, recyclable compartment usage, non-recyclable compartment usage, GPS data

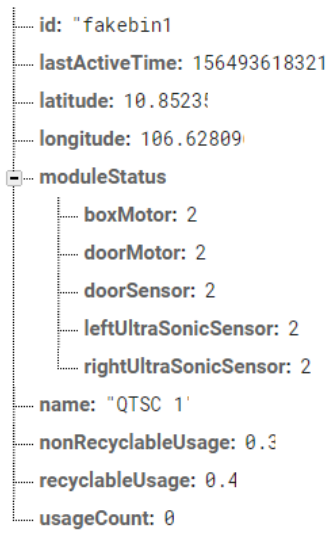


Figure 30. Bin record

- Bin usage history table: for analysis purpose

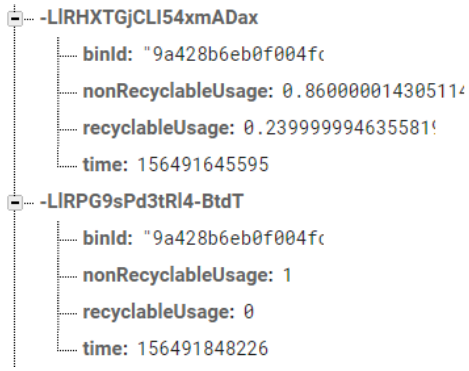


Figure 31. Bin usage record

- Classification history table: for analysis purpose

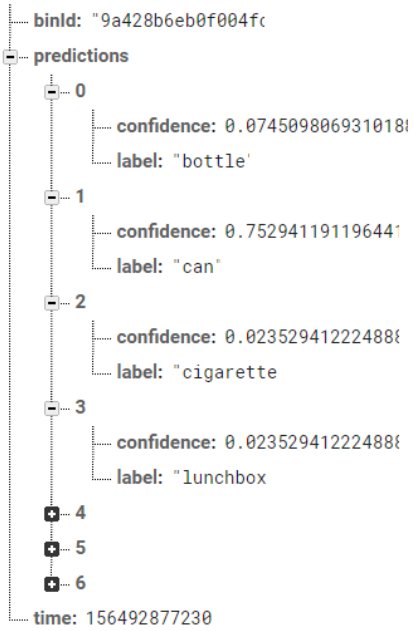


Figure 32. Classification result record

4) Management application

Platform: Android

IDE: Android Studio

Programming language: Java

a) Map of bins

Base on the GPS data of bin system and the data in database, we develop and application run on Android platform to monitor bins. Our application contains a map. It display all bins, which registered on the database.

We use Google Map API to download and display map, then we fetch data from Firebase. After that, we display the location of bins based on the corresponding data on the database.

Our application user interface is designed base on Google Material design principle.

The application supported some features:

- Map supports desktop mode and mobile mode, support different screen size, flexible user interface
- Tap a bin to see its information, include name of bin, its location, identifier, compartment usage
- Support measure approximately distance from current location to each bin, we calculate it by using Android Location API
- Search bins base on its name or location
- Notification feature shows information about bin problems, such as internet connection problem or sensor failure, or notify that what bin is full.

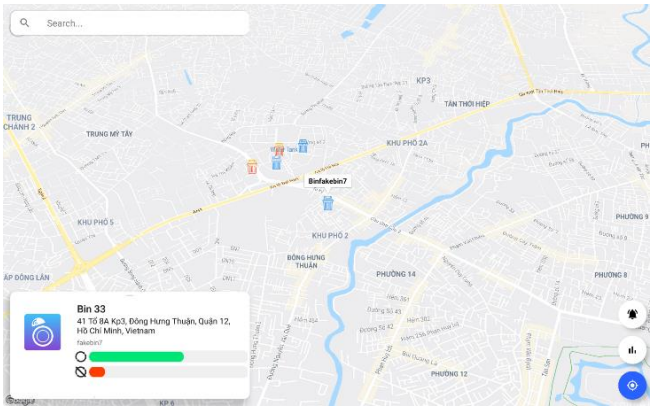


Figure 33. Desktop mode

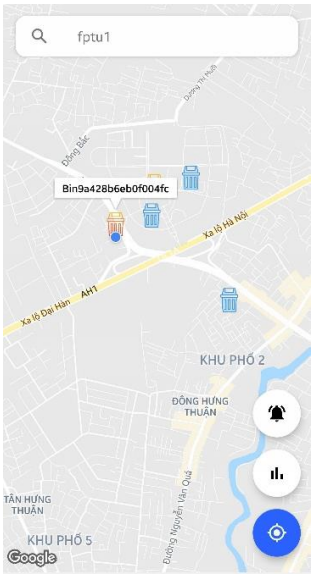


Figure 34 Mobile mode

Figure 35. Mobile mode

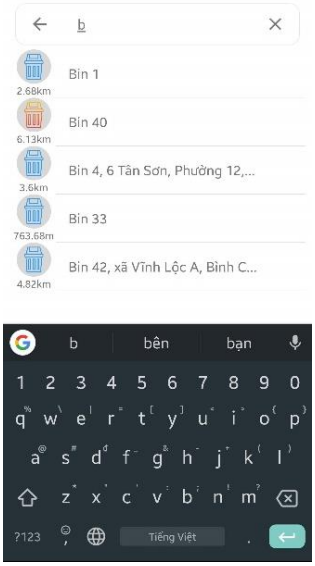
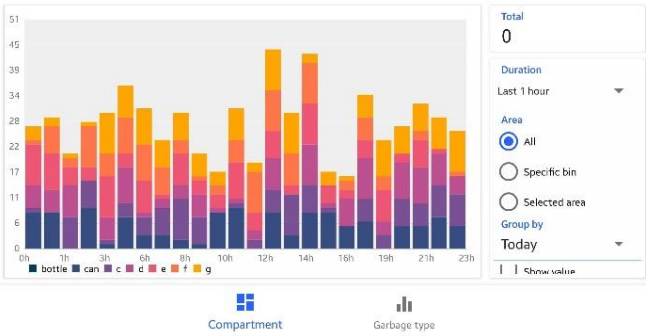


Figure 36. Bin searcher

b) Analysis application

Base on the data we collected from the bin, we can analyses and detect which location is use Magic Bin more than other location. So that we can adjust number of bin accordingly.

We also develop an application which can query and display chart. It help the manager have a general look about bin system.



III. Conclusion

Our solution - Magic Bin - cannot completely replace the traditional trash, but it is a powerful assistant in managing waste not only in Vietnam but also in the world. In Vietnam, our solution is unique. The prototype can develop as a real product and it can apply in many situations and variety environment. It helps classify trash effectively.

Moreover, the completed software can service millions of bin. In current version (1.0.4), it is stability for using. We use technology which provided by Google. So that, it is flexible, scalable and it is suitable for future development.

Our machine learning system can improve day by day. The more trash it processes the more accuracy it is.

Furthermore, our products help reduce the amount of waste significantly, making a significant contribution to environmental protection. We also raising awareness for everyone about the important of environmental protection.

IV. Future Development

Due to the limitation of time for this project, we cannot implement all expected features. In the future development, we will update the bin with some useful features:

- Using solar energy
- Replace the CNC system with the robotic arm system.
- Developed in parks, schools, hospitals, households, etc… by training new machine learning model.

V. References

[1] Arduino
http://arduino.vn/
https://www.arduino.cc/

[2] UltraSonic Sensor
http://arduino.vn/bai-viet/233-su-dung-cam-bien-khoang-cach-hc-sr04

[3] IR Sensor
https://create.arduino.cc/projecthub/adammansou/distance-measuring-sensor-assignment-faeeb4

[4] CNC
https://vi.wikipedia.org/wiki/CNC
http://arduino.vn/bai-viet/1328-gioi-thieu-ve-cnc-plotter-mini

[5] Step motor
https://en.wikipedia.org/wiki/Stepper_motor
https://www.instructables.com/id/Controlling-a-Stepper-Motor-with-an-Arduino/

How To Control a Stepper Motor with A4988 Driver and Arduino https://howtomechatronics.com/tutorials/arduino/how-to-control-stepper-motor-with-a4988-driver-and-arduino/

[6] Firebase Realtime Database
https://firebase.google.com/

[7] TensorFlow, Machine learning
https://www.tensorflow.org/
https://github.com/tensorflow/

[8] Convolutional neural network – CNN – Basic Knownlegde

[9] A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

Understanding of Convolutional Neural Network (CNN) — Deep Learning https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148

Image Recognition in Python with TensorFlow and Keras https://stackabuse.com/image-recognition-in-python-with-tensorflow-and-keras/

Building a Convolutional Neural Network (CNN) in Keras https://towardsdatascience.com/building-a-convolutional-neural-network-cnn-in-keras-329fbbadc5f5

[10] CLion - A Smart Cross-Platform IDE for C and C++ - JetBrains
https://www.jetbrains.com/clion/

[11] Android Studio and SDK tools - Android Developers
https://developer.android.com/studio/

[12] Google Map API
https://developers.google.com/maps/documentation/android-sdk/intro

Smart Hospitals

Kien Ha Ngoc Tan
Cantho FPT University
FPT University
Can Tho, Viet Nam
tankhnce130013@fpt.edu.vn

Le Nguyen Huu Tai
Cantho FPT University
FPT University
Can Tho, Viet Nam
tailnhce130024@fpt.edu.vn

Le Quoc Khoi
Cantho FPT University
FPT University
Can Tho, Viet Nam
khoilqce130023@fpt.edu.vn

Quach Luyl-Da
Cantho FPT University
FPT University
Can Tho, Viet Nam
dalq@fe.edu.vn

Nguyen Dinh Khoi
Cantho FPT University
FPT University
Can Tho, Viet Nam
khoindce13009@fpt.edu.vn

Nguyen Huynh Tuan Khoi
Cantho FPT University
FPT University
Can Tho, Viet Nam
khoinhtce140133@fpt.edu.vn

***Abstract* – IoT (Internet of Things) is an outstanding technology application in the 4.0 industrial revolution. It can be explained as a system of technological devices interconnected via internet connection. IoT is widely applied in many fields, especially in modernizing the health care sector. Our team is working on an IoT system that can contribute to health care by giving hospitals the ability to monitor patients' health in real-time. Our system collects analog signals from sensors owned by patients and then processes that data. From this, doctors and hospitals can be kept up to date about patients' latest health condition. Monitoring health status with the system is not restricted to hospitals only, as even patients and patients' relatives can do this via the website interface and applications. Our system is designed carefully so that users can use it with no remarkable difficulty.**

Keywords – IoT, smart hospitals, monitoring, heart rate, health care, heart disease

I. INTRODUCTION

In a modern busy society, people's health care need from all over the world has been vastly increasing, including in Vietnam. There are various reasons for this tendency: Population growth, awareness, psychological factors, economic development [1]. According to a survey of Q&Me (2017), 95% of males and 88% of females consider themselves as health-conscious and have good health condition [2].

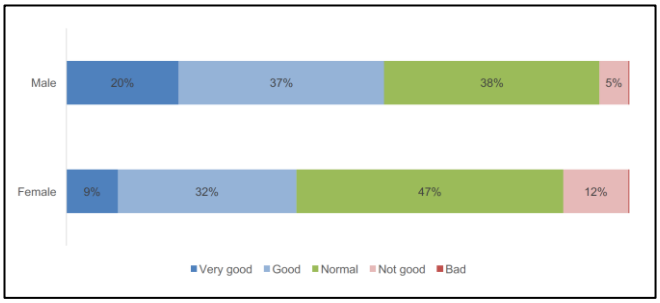


Figure 1 - Health conditions

However, our health system quality is reduced due to a variety of shortcomings. According to one study, 50% of respondents said that the main reason preventing them from dealing with health issues is "no time", the rest blamed on long waiting or the doctors' not spending a lot of time consulting, ... The financial problem is also one of the reasons why we avoid health solutions [3]

To address this phenomenon, in this study, we have developed a system to remotely monitor health conditions, especially heart-related diseases. The system uses IoT devices, integrated with MAX30100 sensor and ESP8266 circuit to perform the task of collecting, receiving and transmitting data. With this system, we can easily monitor our health status anytime, anywhere and potentially save money..

Previous related research topics:

In 2010, there was a research on "Design and development of a heart rate measuring device using fingertip" [7] conducted to improve the estimation of heart rate. The device used optical technology to detect blood flow through the fingertips for pulse detection, pulse detection, and pulse amplification. Research results showed that the inclusion rate was not significant. However, the study only stopped at receiving signals from optical sensors but could not transmit data. For MAX30100 sensors, optical signal acquisition is like the [8] study, but the algorithm is used for two completely different devices.

Another recent study in May 2019, "Machine learning and IoT-based arrhythmia diagnosis using statistical and dynamic features of ECG" [9] had IoT-enabled applications for signal recognition and transmission which could follow Remote monitoring of metronome indicators on Arduino IDE environment. The system has not integrated into user-friendly tools such as browsers, mobile applications, ... and has not designed the complete system for this study.

In the market today there are many large corporations such as Apple, Samsung, Xiaomi, ... specializing in manufacturing and trading smartwatches that integrate heart rate measurement on the device [10]. The exact ratio of these devices is relatively high. For example, if we use Samsung Galaxy S8 smartphone of Samsung and Apple Watch Series 3 smart watch to measure the heart rate of a 20 year- old (weighing 60kg) person doing an intense activity – lifting 20kg weights for 12 times with 2 minutes break after each – 3 times, the turn-on index for each device is:

- Samsung Galaxy S8: 136 BPM, 155 BPM and 164 BPM (avg: 151 BPM)

- Apple Watch Series 3: 142 BPM, 159 BPM and 170 BPM (avg: 157 BPM)

However, most devices merely warn users about their health status and do not have any connection with the hospitals' system.

Recently, the market has a Philips Monitor Mounting [11] medical device used in hospitals manufactured by Philips. The equipment can collect all relevant indicators of the heart, oxygen, ... of the patients. Philips device has a compact design which helps a physician to monitor patients easily. However, this device is restricted to hospital use, especially to emergency departments. Moreover, the compact devices that Philips produces are developed for their own devices. Thus, they have not integrated the main device's features onto mobile devices like smartphones.

From the above researches, we can conclude that our research is new and feasible in today's modern society. This does not only help us to easily monitor the health status of ourselves and our family but also timely connect patients to hospitals in case of emergency.

II. SOLUTION PLAN

For the technical process, the MAX30100 sensor will collect data from the patient and be transmitted to the central system thanks to WIFI ESP8266 circuit and data will be saved to the database. After obtaining data from patients, the system continues to distribute that data to the doctor, the patient's family, who have been decentralized before.

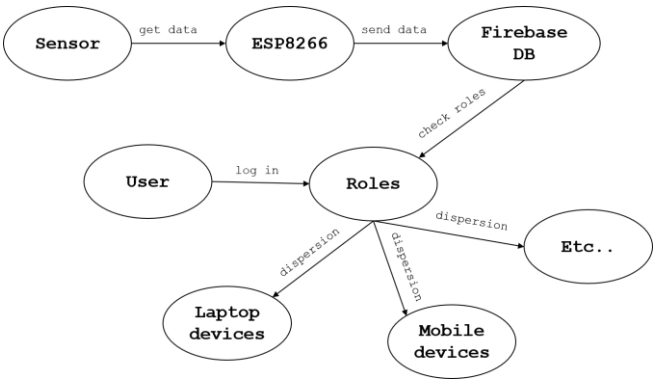


Figure 2 - Technical overview process

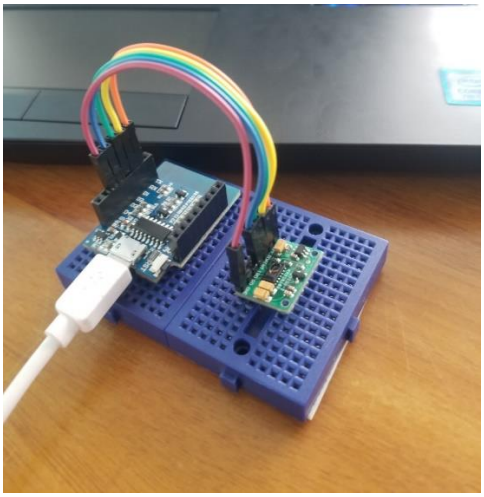


Figure 3 - Testing device of the system

1. MAX30100 [4] and ESP8266 Node MCU Lua D1 Mini [5]

Figure 4 - MAX30100 [4]

1.a. General description

The MAX30100 is an integrated pulse oximetry and heartrate monitor sensor solution. It combines two LEDs, a photodetector, optimized optics, and low-noise analog signal processing to detect pulse oximetry and heart-rate signals. [4]

The MAX30100 operates from 1.8V and 3.3V power supplies and can be powered down through software with negligible standby current, permitting the power supply to remain connected at all times. [4]

1.b. System of Max30100

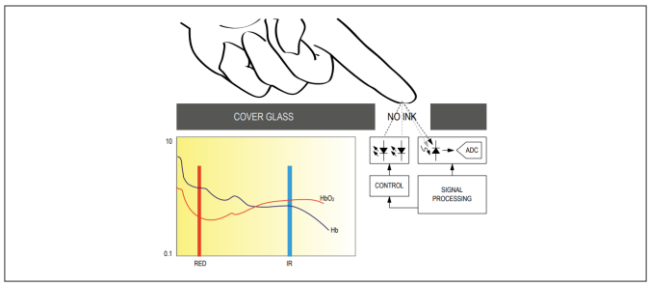


Figure 5 - System block diagram [4]

The MAX30100 is a complete pulse oximetry and heartrate sensor system solution designed for the demanding requirements of wearable devices. The MAX30100 provides very small total solution size without sacrificing optical or electrical performance. Minimal external hardware components are needed for integration into a wearable device. [6]

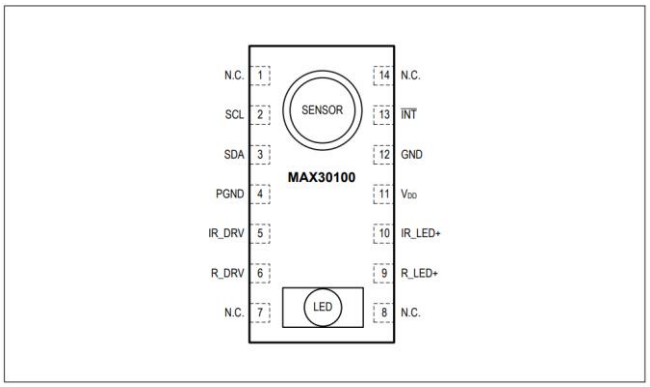


Figure 6 - Pin Configuration [4]

The MAX30100 is fully configurable through software registers, and the digital output data is stored in a 16-deep FIFO within the device. The FIFO allows the MAX30100 to

be connected to a microcontroller or microprocessor on a shared bus, where the data is not being read continuously from the device's registers. [6]

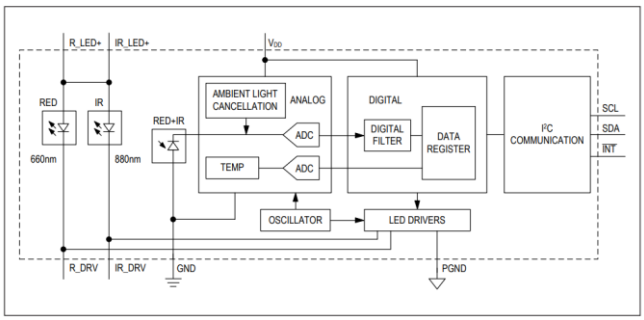


Figure 7 - Function diagram [4]

The SpO2 subsystem in the MAX30100 is composed of ambient light cancellation (ALC), 16-bit sigma delta ADC, and discrete time filter. The SpO2 ADC is a continuous time oversampling sigma delta converter with up to 16-bit resolution. The ADC output data rate can be programmed from 50Hz to 1kHz. The MAX30100 includes a proprietary discrete time filter to reject 50Hz/60Hz interference and low-frequency residual ambient noise. [6]

1.c. WIFI ESP8266 Node MCU Lua D1 Mini

Besides, the special combination of MAX30100 sensor and the WIFI ESP8266 Node MCU Lua D1 Mini kit makes the device smaller and easier to use.

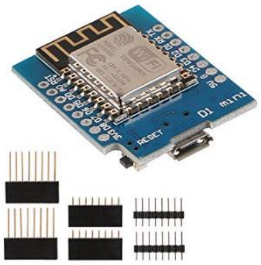


Figure 8 - ESP8266 Node MCU Lua D1 Mini

Kit WIFI ESP8266 Node MCU Lua D1 Mini is designed to be easy to use and especially can be used directly by the Arduino compiler to program and load the code. This makes the use and programming of the above applications ESP8266 become very simple. Kit WIFI ESP8266 Node MCU Lua D1 Mini has a compact design, built-in program loader, and UART CH340 interface, used for applications that need to connect, collect data and control via WIFI, especially are IoT-related applications.

Pin		
Pin	**Function**	**ESP-8266 Pin**
TX	TXD	TXD
RX	RXD	RXD
A0	Analog input, max 3.3V input	A0
D0	IO	GPIO16
D1	IO, SCL	GPIO5
D2	IO, SDA	GPIO4
D3	IO, 10k Pull-up	GPIO0
D4	IO, 10k Pull-up, BUILTIN_LED	GPIO2
D5	IO, SCK	GPIO14
D6	IO, MISO	GPIO12
D7	IO, MOSI	GPIO13
D8	IO, 10k Pull-down, SS	GPIO15
G	Ground	GND
5V	5V	-
3V3	3.3V	3.3V
RST	Reset	RST

Figure 9 - WIFI ESP8266 Node MCU Lua D1 Mini datasheet [5]

2. Smart hospitals system analysis

2.a Process overview

Patients can register to monitor their own health with smart hospitals. The hospital will issue a device (which is the system under research) that can monitor heart rate, blood oxygen levels, etc. remotely. The gathered data will be stored at the hospital. The patient's doctors, patients, and family members can view the information in real-time via the website or mobile application with the account provided by the hospital. When figures like heartrate are at an alarming level, the system will alert patients and patients' families. When the figures are at a critical level, the system will notify continuously to the nearest hospital, so the hospital can promptly prepare necessary tools to help patients.

2.b Objects using the system

The system includes the following objects: *hospital, dean, doctor, patient, patients' relatives*

2.c Database analysis

Figure 10 - Use database mySQL and Firebase

In this system, we combine the use of MySQL database to manage the accounts and information of hospitals, doctors and patients; and Firebase's database for integrating patient real-time features, processing, and data transmission.

Figure 11 - Firebase real-time database

Database for accounts, allowing users to log into the system with the basic rights granted earlier. Each object will have different rights and roles

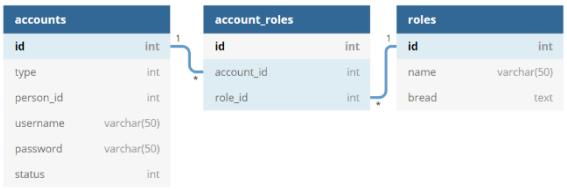


Figure 12 - Accounts and roles of system

Database for the whole system:

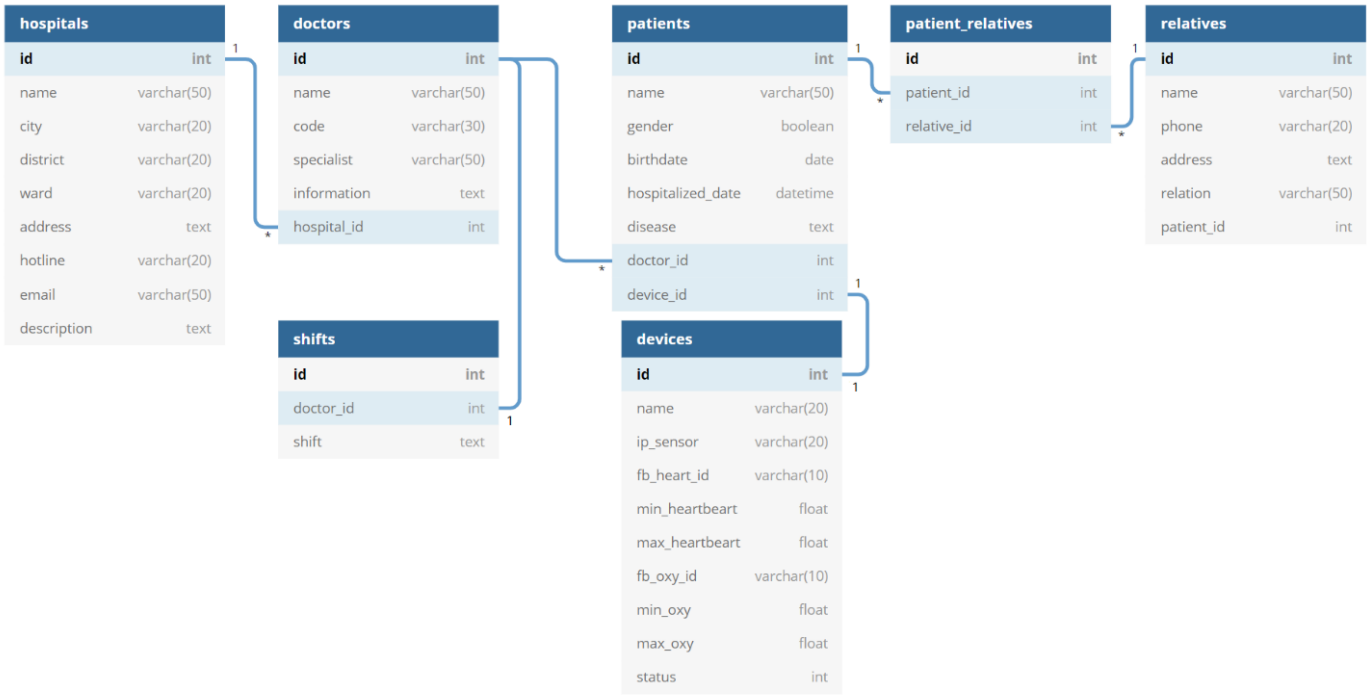


Figure 13 - Database of the system

III. ROLES:

No.	Roles name	Description
1	Hospitals' role	- View, edit, delete all tables (doctor, patient, shift, patient's family member, device manager, ...) - See all the data collected from the patient's sensor - Be entitled to send notices to doctors and assign duty schedule to doctors
2	Doctors' role	- Manage equipment and patients - Be entitled to send notices to patients, remind patients - Right to view, add, edit, delete data that sensors collect from patients
3	Patients' role	- Have the right to view your health status on devices - Have the right to ask doctors about their own health - Receive a doctor's notice in emergencies
4	Patients' relatives' role	- Be entitled to view the health status of people on devices - Receive a doctor's notice in emergencies

Table 1 - Roles of system

We have researched and analyzed the system, designed the system so that users have basic rights to view their own indicators, still assured of their health status. The signals will be automatically calculated by the system, compared and given a warning level on the user's device.

Figure 14 - Login interface of the system (illustrated image)

IV. ANALYSIS

1. Accuracy of the signal

In the first study, we asked 20-year-old (weighing 60kg) people to exercise heavily by lifting a 20kg object (lifting 12 times, each time taking a 2-minute break) and perform heart rate measurement 3 times. The results are as follows:

#	Heartrate
1	140 BPM
2	155 BPM
3	169 BPM
Average	154 BPM

Table 2 - Heart rate index for 3 measurements

Compared to the results of the 2 devices mentioned in the introduction, we have the following data sheet

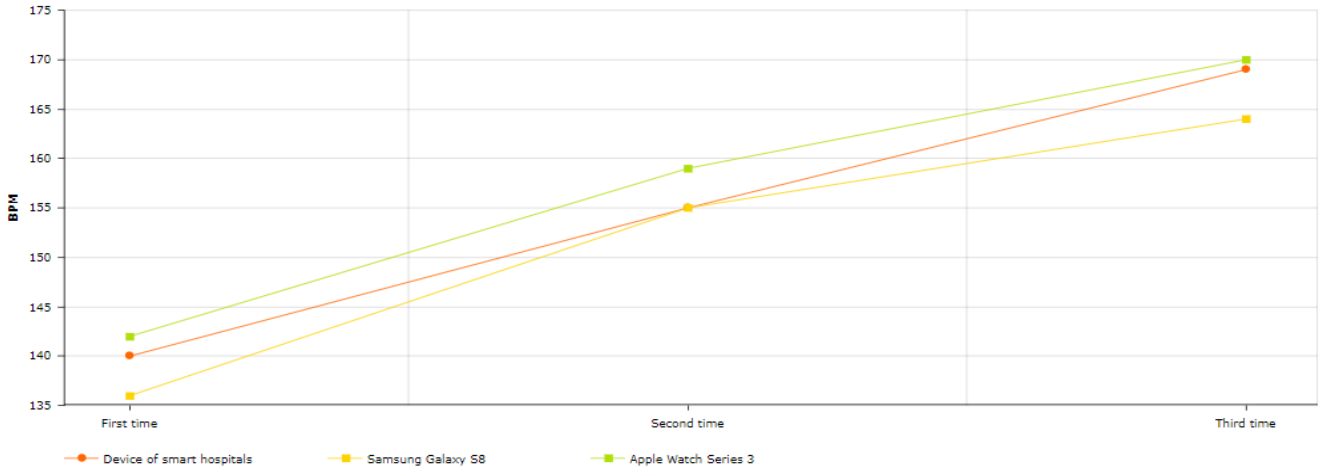


Figure 15 - Heart rate statistics table after three measurements

2. Analysis results on many people:

#	First time	Second time	Third time	Avg
1	73	72	71	72
2	80	76	77	77
3	86	85	82	84
4	75	71	70	72
5	77	72	73	74
6	86	85	82	84
7	72	67	71	70
8	88	85	85	86
9	83	78	81	80
10	81	76	79	78
11	82	79	78	79
12	70	67	65	67
13	83	81	81	81
14	75	74	71	73
15	81	77	77	78
16	81	79	80	80
17	71	69	70	70
18	74	73	72	73
19	81	77	77	78
20	88	83	86	85

Table 3 - Table measures the heart rate of 20 FPT students

V. EXPERIMENTAL RESULTS AND CONCLUSION

From the above analysis shows that the error in patient data collection is negligible. At the same time, the network latency while transmitting and receiving signals from patients to the system is also very small. Therefore, the system can be deployed in the future.

Furthermore, the system also saves time, cost. It provides an easy way to check your health anywhere.

We hope that our research will improve the quality of primary health care.

DEVELOPMENT

In the future, we will continue to research and develop this system. To be specific, we will work on how to store patients' data in the absence of internet or disconnection. When a stable connection is made to the central system, all data stored offline will be sent to the center. In addition, to enhance system flexibility, we will solve the problem of integrating the available devices. From this, we hope that our system can serve people with high precision and efficiency.

ACKNOWLEDGMENT

First, we would like to acknowledge MSc. Quach Luyl-Da and MSc. Vo Hong Khanh for their whole-hearted guidance and support throughout the process. Had it not been for their help, our research would not have been possible.

We also want to show appreciation to all of our friends in the Cantho FPT University who have offered valuable support and encouragement.

REFERENCE

[1] "Khảo sát nhu cầu chăm sóc sức khỏe ngoài giờ và tại nhà của người đến khám, điều trị bệnh tại bệnh viện phổi trung ương năm 2014 -: XEMTAILIEU," Xemtailieu.com. [Online]. Available: https://text.xemtailieu.com/tai-lieu/khao-sat-nhu-cau-cham-soc-suc-khoe-ngoai-gio-va-tai-nha-cua-nguoi-den-kham-dieu-tri-benh-tai-benh-vien-phoi-trung-uong-nam-2014-298071.html.

[2] Vietnam Market Research, "[Survey] Vietnamese health consciousness and fitness demand," LinkedIn SlideShare, 30-Dec-2017. [Online]. Available: https://www.slideshare.net/asiaplus_inc/survey-vietnamese-health-consciousness-and-fitness-demand-85392567/2.

[3] "Mọi người đều lo lắng về sức khoẻ » VieVie Healthcare - Bác sĩ bên bạn," VieVie Healthcare - Bác sĩ bên bạn. [Online]. Available: http://vievie.vn/lo-lang-ve-suc-khoe-noi-niem-khong-cua-rieng-ai/.

[4] "Pulse Oximeter and Heart-Rate Sensor IC for Wearable Health," Max30100, 01-Jan-2014. [Online]. Available: https://datasheets.maximintegrated.com/en/ds/MAX30100.pdf.

[5] "Kit RF thu phát Wifi ESP8266 NodeMCU Lua D1 Mini," Hshop.vn. [Online]. Available: https://hshop.vn/products/kit-rf-thu-phat-wifi-esp8266-nodemcu-mini-d1.

[6] "MAX30100 Pulse Oximeter and Heart-Rate Sensor IC for Wearable Health General Description," manualzz.com. [Online]. Available: https://manualzz.com/doc/9334866/max30100-pulse-oximeter-and-heart-rate-sensor-ic-for-wear.

[7] M. M. A. Hashem, R. Shams, M. A. Kader, and M. A. Sayed, "Design and development of a heart rate measuring device using fingertip," International Conference on Computer and Communication Engineering (ICCCE10), 2010.

[8] R. Strogonovs, "Implementing pulse oximeter using MAX30100 - MORF - Coding And Engineering," Morf. [Online]. Available: https://morf.lv/implementing-pulse-oximeter-using-max30100.

[9] R. L. Devi and V. Kalaivani, "Machine learning and IoT-based cardiac arrhythmia diagnosis using statistical and dynamic features of ECG," The Journal of Supercomputing, 2019.

[10] B. Pi, "Smart watch," 05-Oct-2015. [Online]. Available: https://patents.google.com/patent/US20160004224A1/en.

[11] "Patient monitoring systems: Philips Healthcare," Philips. [Online]. Available: https://www.usa.philips.com/healthcare/solutions/patient-monitoring.

Research Confest 2019

Project Smart School

Tran Do Minh An
FPT polytechnic Ho Chi Minh
391A Nam Ky Khoi Nghia
minhan2111@subee.dev

Nguyen Quoc Anh
FPT polytechnic Ho Chi Minh
391A Nam Ky Khoi Nghia
anh@subee.dev

Luu Minh Thuan
FPT polytechnic Ho Chi Minh
391A Nam Ky Khoi Nghia
thuan@subee.dev

***Abstract*— *Smart School is a cross-platform application that connects all Electrical equipment in your classroom becomes a network. It does not cost much and easy to use with many useful features, you can control it directly the Smart School application in many ways.*(*Abstract*)**

Keywords*— *Smart School, IoT Smart Home, cross-platform application.

I. INTRODUCTION (*HEADING 1*)

Based on basic needs of FPT polytechnic Ho Chi Minh students and problems in announcing news, school holidays to students ... Smart School, a cross-platform application that connects all Electrical equipment in your classroom becomes a network, so you can control them according to smart scenarios: including lighting, air conditioning, sound, and fan. You can control it directly via the Smart School application in many ways: direct editing as a remote or voice control.

II. PRODUCT DESCRIPTION

A. Software Implementation

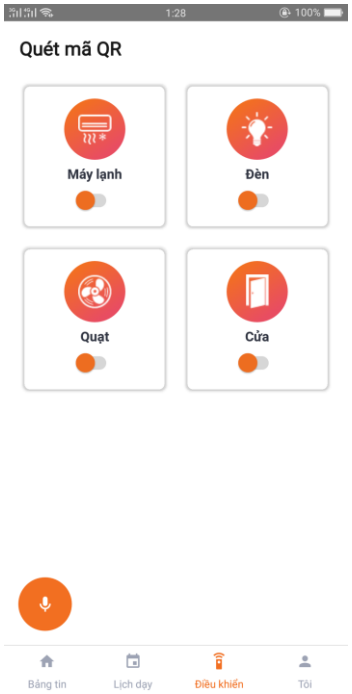


Fig 1: Control Screen

Friendly screen displays all device in the classroom, user can control them directly through the Smart School application in many ways: direct editing as a remote or voice control.

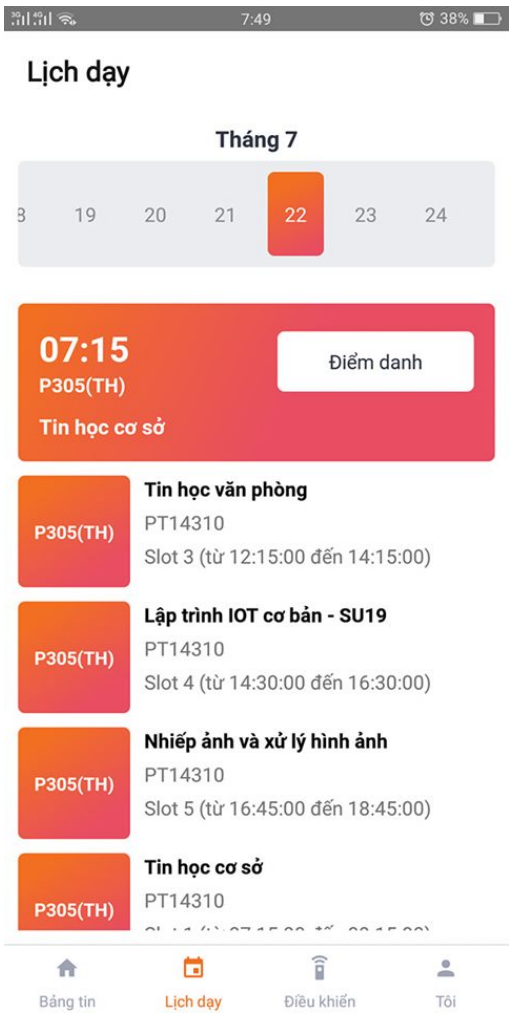


Fig 2: Attendance, Schedule Screen.

Easy to follow attendance for student and teacher, all the information you need has been encapsulated in this application. Especially the function of sending notifications to students.

B. Hardware Implemntation

In the hardware implementation, we are using Wemos-d1 as a controller. We use Andruino IDE software to write programs for this module. ESP8266 Wi-fi module is used for communication between the Smart School app and Wemos-d1. With this protocol we connect between Smart School application and Wemos-d1 to handle the given requests as well as receive information from sensors to send data back to the application.

C. The highlight of Smart School

The highlight of Smart School system is cheaper than Smart Home project, as well as improves technology, such as custom command, users can easily control the device for each action.

In addition to the highlights of electrical equipment, the application also offers convenience for students, when all the information you need has been encapsulated in this application. Especially the function of sending notifications to students. We have developed this function so that the management board, student affairs staff can send notice of the number of vacations, remind the schedule quickly, and at no cost compared to using the SMS system.

In addition to the benefits for students and the management board , the app offers teachers with a more convenient and faster means to check student attendance compared to the previous use of the school system. At the same time, teachers can follow their teaching schedule faster.

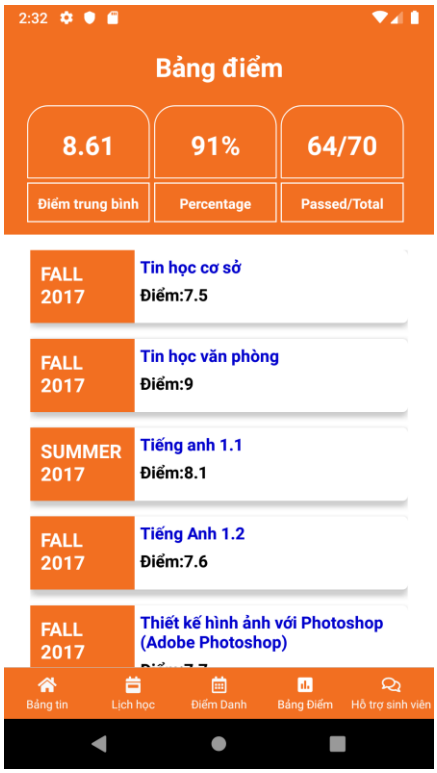


Fig 3: Score board Screen

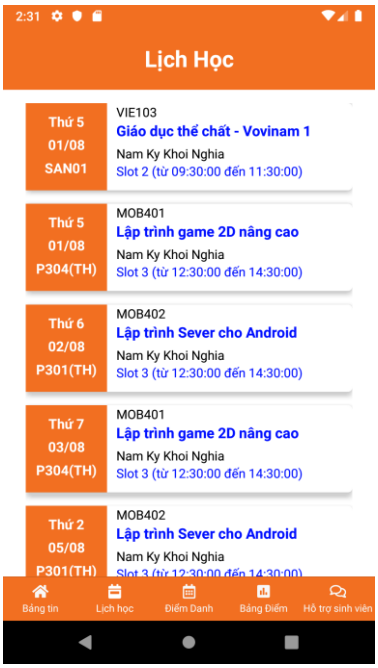


Fig 4: Schedule Screen

III. FEASIBILITY OF THE PRODUCT

In the market today, the school management software has a high price such as school management software MISA and SISAP.

MISA	SISAP
Highlight Features	**Highlight Features**
Integrated service electronic invoices Application of artificial intelligence in management of cadres, civil servants and officials Enrollment online Price: 10-12.000.000 VNĐ.	View student information(*(Attendance, Schedule, Transcript)* Parents receive immediate notice of course grades, change timetable, generate school fees, medical notification incidents ... Price:8-10.000.000 VNĐ.

These software only focus on management, and with Smart School, it combines management of professional activities and management of electrical equipment in the classroom, saving electricity and management costs, friendly.

IV. PROPOSED SYSTEM

In this system, the Smart School application used to send signals to the … facility is connected to the … module that provides this signal to the system to control indoor devices.

1. Smart School application sends the signal to the module Wemos-d1 which is connected to module ESP8266 Wifi Protocol .
2. Wi-fi module receives the signal from the Smart School application and give this signal to Wemos-d1 for processing.

3. We use Wemos-d1 as a controller to control all the appliances. Each command is processed by Wemos-d1 board to control and transmitter for switching on/off the appliances.
4. Relay module receiver and transmitter use as electrical witches, for performing on/off operation.
5. Users use Smart School app for another action.

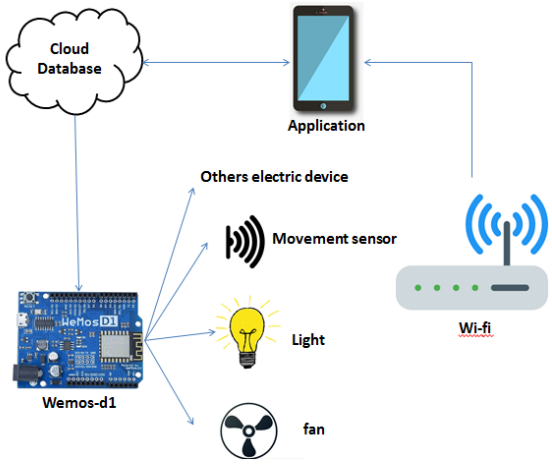


Fig 5: Operating model

V. FUTURE SCOPE.

In our system, we provided a cross-platform application that can use in Android and IOS platform. We will create and upgrade more useful features for student and teacher at school.

VI. CONCLUSION

Smart School is a cross-platform application that connects all Electrical equipment in your classroom becomes a network. Providing many features for student and teacher, it can save energy and manage the school easily when using Smart School application.

[1] Anruino VN Website: https://www.arduino.cc/

[2] React Native Website :https://facebook.github.io/react-native/docs/getting-started.html

[3] Module WIFI ESP8266 "The ESP8266 is a low-cost Wi-Fi microchip with full TCP/IP stack and microcontroller capability" Availabe in: https://github.com/esp8266/Arduino

[4] GUI "Graphical User Interface" Available in: https://developer.android.com/guide/topics/ui

[5] Arduino IDE "Arduino integrated development environment" Website: https://www.arduino.cc/en/Main/software

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.

Smart Tank for Preventing Children from Suffocation

Student 1: *Bui Trung Thanh*
Class: *Web design 1*
FPT POLY TECHNIC
Can Tho city, Vietnam
Email: *thanhbtpc00329@fpt.edu.vn*

Student 2: *Nguyen Hoang Nha*
Class: *Web design 1.*
FPT POLY TECHNIC
Can Tho city, Vietnam
Email: *nhanhpc00404@fpt.edu.vn*

Student 3: *Huynh Huu Tran Phuc Loi*
Class: *Graphic design 1*
FPT POLY TECHNIC
Can Tho city, Vietnam
Email: *nhanhpc00404@fpt.edu.vn*

Instructor 1 : Tran Van Then
Information Technology Department
FPT POLY TECHNIC
Can Tho city, Vietnam
Email: *Thentv@fe.edu.vn*

Instructor 2: Le Thanh Nhuan
Information Technology Department
FPT POLY TECHNIC
Can Tho city, Vietnam
Email: NhuanLT@fe.edu.vn

***Abstract*—Smart water tank is an IOT solution combining multi-layer warning model, which help users to prevent children from suffocation.**

Keywords—IOT, smart water tank, preventing sufocation, FPT tech.

I. DESCRIPTION OF THE PRODUCT

In fact, children really like to play with water, especially water tanks in families and schools. Therefore, children are likely to fall into water tanks, which could easily lead to suffocation. The POLY-CT group proposes a device called IOT-POLY001 that detects children falling into a water container to immediately alarm and rescue them.

II. MARKETABILITY AND HUMANITY

A. Human life, especially young children must be protected at all times.

According to statistics of the Department of Preventive of Medicine (The Ministry of Health), there are about 3,500 children drowned on average in our country. Remarkably, the number of children from 0 to 4 years old drowned was the highest, with an average of 22 cases/ 100,000 children/ year.

One of the most dangerous places for children is water tanks in every family bathroom. A child may fall into water tank and may suffocate after only 2 seconds because of adult's ignorance.

From this fact, the POLY-CT group proposes a device called IOT-POLY 001 that immediately detects and supports when a child falls into a water container.

B. The commercialization of the product in the future

The implementation of the product is very easy due to:

- Abundant equipment source which is easy to find
- Affordable price of the device
- Exceptional features
- High and fast profitable device

II. EQUIPMENT AND SOFTWARE USED FOR IOT-POLY001.

The following figures are the model, hardware and software devices that make up the IOT-POLY 001 product.

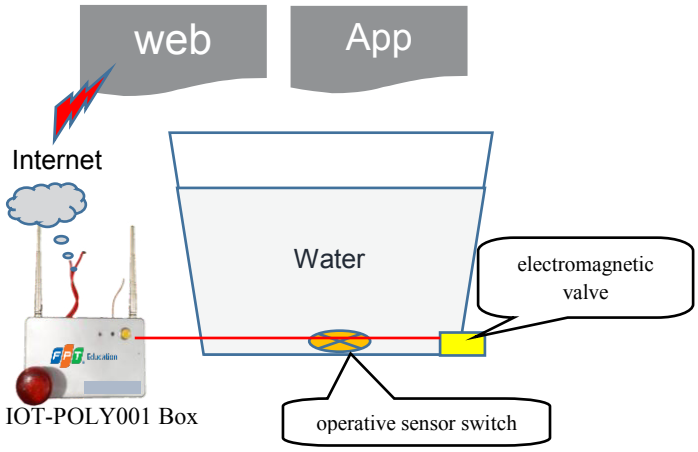


Figure 1: IOT-POLY001 operation model

A. Hardware devices

- Module Arduino UNO R3

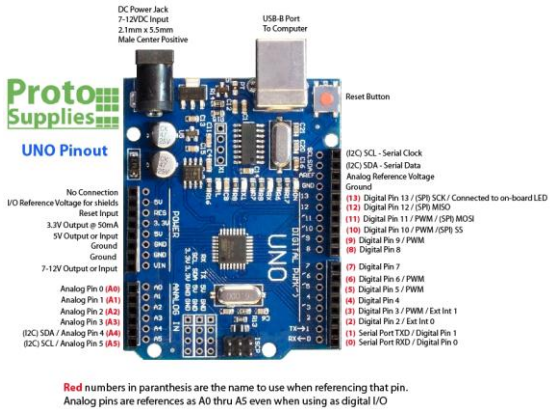


Figure 2: Module Arduino UNO R3

Arduino UNO R3 is a third Arduino UNO generation kit, with programmable capabilities for complex control applications due to the powerful configuration for ROM, RAM and Flash memory type, digital I/O inputs and outputs including many possible signal outputs PWM, analog signal readings and various communication standards such as UART, SPI, TWI (I2C). Arduino UNO R3 is used as a board to control signals received from the sensor switch and transmit the signals to the ESP 8266 module.

- Modul ESP8266

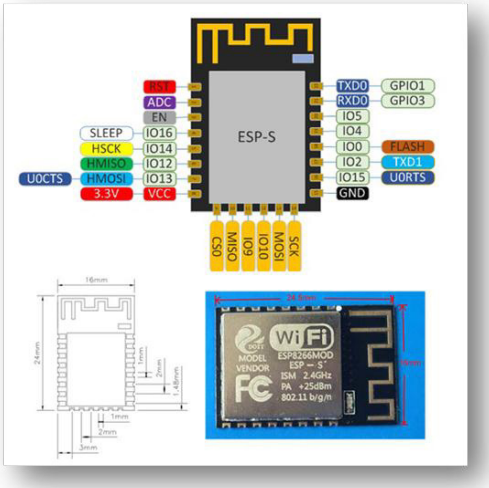


Figure 3: Module ESP8266

The *ESP8266* is a low-cost Wi-Fi microchip with full TCP/IP stack and microcontroller capability produced by manufacturer Espressif Systems in Shanghai, China

- Module USB to UART converter module

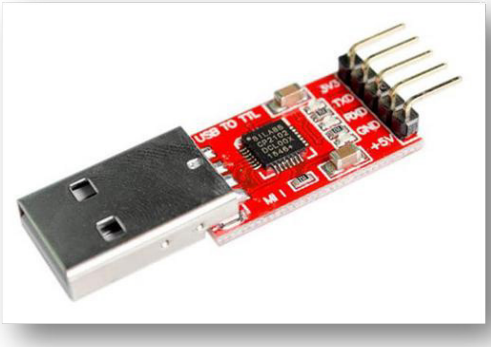


Figure 4: USB to UART

Lets your computer talk UART, you need a device that converts computer bytes to UART signals; a USB to UART converter module. This is a small device that plugs into your USB port and has at least ground, Rx and Tx outputs. It pretends to be a serial port to your computer. The computer sends data to this serial port and the module converts it to UART signals.

Is used to load program for the ESP 8266 module.

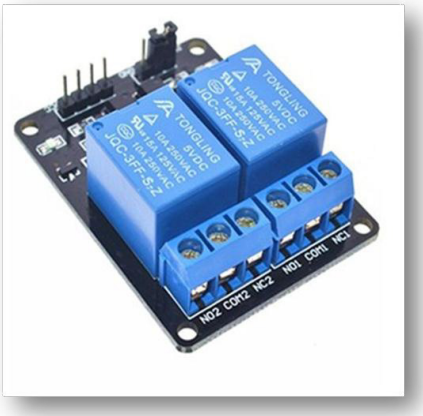


Figure 5: Relay 5VDC

Module 2 Relay activates H/L (5VDC) to use 5DVC power to feed the circuit. The trigger signal can choose high level (High – 5DVC) or low level (Low – 0VDC) via Jumper on each relay. One channel on a relay is used to close / open electromagnetic shifting-valve which discharges water and the other channel is used to close/ open the alarm.

- Limit switch

Figure 6: Limit Switch

Limit switch is placed in a water tank to have sensor whenever there is such a heavy thing (>= 1 kilogram). The sensor then activates to close the switch.

- Electromagnetic valve

Has the responsibility of instant discharging water when a problem occurs. The valve is closed/ opened with 24VDC power, which is very safe.

B. Software

- Uses MQTT protocol to communicate through CLOUDMQTT.COM.
- MQTT Dash app on smart phone

II. OPERATION MECHANISM

The following describes the process of IOT-POLY 001.

A. Normal operation state

In normal operation state, when there is no problem, the MQTT Dash sends the message "A good day" to smart phone.

Figure 7: Good message

B. Unusal occurence

When there is a thing that is heavy enough (e.g. children) falling into a water tank, the sensor switch will close. The IOT-POLY001 immediately activitvates the valve to drain off water. At the same time, it alerts the alarm and flashing lights and sends information which has the incident to the smart phone.

Figure 8: Bad message

C. Stopping the system

When we want to turn off the siren, stop discharging water, and notify the remote users via smart phone which has been tackled, we only need to press the yellow button.

Figure 9: Stop the alarm

The message being processed will be sent to smart phone.

Figure 10 :Resolved

D. Resetting the system

When we want to reset this system to be ready, reset the system to the ready state. We turn off the power supply for IOT-POLY 001 and then turn on the power again.

THE ORIENTATION OF PRODUCT DEVELOPMENT

In the future, the product will have many variations for larger applications such as family swimming pools and fish tanks.

REFERENCES

[1] C. t. C. c. n. T. V. Nam, "http://thbvn.com," Công ty CP công nghệ THB Việt Nam, [Online]. Available: http://thbvn.com/thiet-bi-giam-sat/chi-tiet/thiet-bi-chong-duoi-nuoc-cho-be-boi. [Accessed 4 8 2019].

[2] ProtoSupplies.com, "https://protosupplies.com," Portland OR 97223, USA, [Online]. Available: https://protosupplies.com/product/uno-r3-smd/. [Accessed 4 8 2019].

[3] Internet, "google.com," [Online]. Available: google/images. [Accessed 4 8 2019].

Smarthome control unit using Vietnamese speech command

Phan Duy Hung
FPT University
Hanoi, Vietnam
hungpd2@fe.edu.vn

Truong Minh Giang
FPT University
Hanoi, Vietnam
giangtmse04802@fpt.edu.vn

Le Hoang Nam
FPT University
Hanoi, Vietnam
namlhse04875@fpt.edu.vn

Phan Minh Duong
FPT University
Hanoi, Vietnam
duongpmse04921@fpt.edu.vn

Hoang Van Thang
FPT University
Hanoi, Vietnam
thanghvse04806@fpt.edu.vn

Vu Thu Diep
Hanoi University of Science and Technology
Hanoi, Vietnam
diep.vuthu@hust.edu.vn

***Abstract*—**Smart home is a very hot development area in which voice-based control devices are receiving special attention from major technology companies and researchers. Despite many studies on this problem in the world, there has not been a formal study for the Vietnamese language. In addition, many studies did not offer a solution that can be expanded easily in the future. This paper provides a speech collection and processing software and shares a dataset of speech commands is labeled and organized to the language research community. This study also designs and evaluates Recurrent Neural Networks to apply it to the data collected. The average recognition accuracy on the set of 15 commands for controlling smart home devices is 98.19%. Finally, the paper presents the implementation and performance evaluation of machine learning model on a Raspberry PI-based intelligent home control unit

***Keywords*—** Vietnamese speech command, command recognition, Recurrent Neural Networks, Raspberry PI, MQTT.

I. INTRODUCTION

Interaction and control of household devices is a fast trend, evident in the exponentially growing number of smart-homes. According to Statista, the number of active households worldwide is 67.4 million in 2019. And that number is expected to amount to 111.2 million by 2023 [1]. The goal of research in this field is to improve the interaction so that it is faster, more convenient and more flexible. Therefore, speech recognition and natural language processing with the support of Artificial Intelligence seems to be the inevitable route.

In [2], the authors create a dataset of 150 people (men and women). All the voice samples are captured in Brazilian Portuguese, with the digits "0" through "9" and the words "Ok" and "Cancel". The results show that a throat microphone is robust in noisy environment, achieving a 95.4% hit rate in a speech recognition system with multiple Neural Networks using the one-against-all approach, while a simple Neural Network could only reach 91.88%.

In [3], feature extraction methods used are the Mel frequency cepstral coefficient (MFCC). Early stages of MFCC split the input signal amplitude values into frames which are then processed by the mel-filter bank. The results of feature extraction are made into a codebook, which is then used as an input symbol on a Hidden Markov Model (HMM) to form a model for every word. The final system can recognize spoken words with an average accuracy of 93.89% in a noiseless environment, and 58.1% in noisy environments.

In [4], Du Guiming et al. propose to use the CNN principles in frequency domain to normalize acoustic variations for speech recognition. Here, the researchers use a 5- layer CNN. It can achieve isolated word recognition by training the CNN.

In [5], Jaesung Bae, Dae-Shik Kim realize that CNNs are capable of capturing the local features effectively. They can be used for tasks which have relatively short-term dependencies, such as keyword spotting or phoneme-level sequence recognition. However, one limitation of CNNs is that, with maxpooling, they do not consider the pose relationship between low-level features. Motivated by this, the researchers use a capsule network to capture the spatial relationship and pose information of speech spectrogram features in both the frequency and time axes. Compared to CNN models, the capsule-network-based systems achieved much better results from both clean and noisy data.

The above studies have shown that deep learning is the most effective solution at the present time to improve accuracy. It also seems that the end result can minimize the dependence of the problem on a specific language when the learning data is big enough. The gap here is to evaluate and customize a network architecture that matches a particular language database. In addition, the solution should be easily expanded on in the future.

The main contribution of this paper is to develop a method for recognizing Vietnamese speech commands based on deep learning technologies. A Vietnamese command dataset that includes 15 commonly used commands for smart homes (Table 1) has been labeled and is publicly available for the research

community on GitHub [6]. In addition, the source code for the data collection software for both Android and iOS is also made available on Github. Users can easily contribute data via software, and it can also be easily modified for other languages. New commands can also be added in the future. Finally this work implements and evaluates the performance of machine learning model on a Raspberry PI-based intelligent home control unit.

Table 1. Table captions should be placed above the tables

No	In Vietnamese	Equivalent English meaning
1	Đô rê mon	Doraemon (Trigger word)
2	Bật đèn	Turn on light
3	Tắt đèn	Turn off light
4	Bật điều hòa	Turn on air condition
5	Tắt điều hòa	Turn off air condition
6	Bật quạt	Turn on fan
7	Tắt quạt	Turn off fan
8	Bật tivi	Turn on television
9	Tắt tivi	Turn off television
10	Mở cửa	Open door
11	Đóng cửa	Close door
12	Khóa cửa	Lock door
13	Mở cổng	Open gate
14	Đóng cổng	Close gate
15	Khóa cổng	Lock gate

The remainder of the paper is organized as follows. Section 2 describes the data collection and processing process. Section 3 provides the selection and evaluation of deep machine learning architectures based on RNN. Next, Section 4 analyzes the application results of the selected architecture for the Vietnamese command dataset. Then, the evaluation of machine learning model on a Raspberry PI-based intelligent home control unit is presented in Section 5. Finally, conclusions are made in Section 6.

II. DATA COLLECTION AND PROCESSING

A. Data collection

In order to ensure the robustness and high accuracy of the identification process, the collected data needs to meet a number of requirements such as diversity in age, gender and region. The "SpeechCollection" application is written for both Android and iOS to be accessible to all users. Volunteers who wish to participate in data collection can download the software via Google's Play Store or Apple's App Store. The results is shared through Google cloud. Data will then be reviewed by the research team for quality and information, and the results will be stored in the final data directory.

The data collected after 1 month of the research contains voices of 293 people with fairly balanced ratios between men and women, age groups (younger than 18, between 18 and 30, between 30 and 40, older than 40) and regions (Northern, Central and Southern).

B. Data processing

1. **Data filtering and trimming :** The raw daa contributed by user might contains a lot of silences and noises. So, the first step to filte and trim the silences at the beginning and the end [7]. Because the frequency range of speech signals is from 300Hz to 3400 Hz, a simple linear bandpass filter is used to eliminate out-of-band noise. Filtered data then is trimmed to remove silences. The data is then divided into continuous frames with a length of 0.05 seconds each. The Short Time Energy (STE) on each frame is 4 calculated and compared to the average STE value. Frames with STE greater than the average value are retained whiel the rest is treated as the silence and removed.

2. **Data augmentation:** Data after trimming has different lengths, and the maximum length is less than 1.5 seconds. However, the data needs to be standardized to the same duration to make it easier to extract features used for machine learning. In addition, the data should be similar despite being collected in different environments. So, in this step the data needs to be augmented, and lengthen to a standard duration.

For background noises, audio recordings were conducted in 10 different environments (library, school, kitchen, room, road, etc.). Each recording has the same length of 2 seconds. These data are used then as background sounds, overlaid over the trimmed audio in the previous step at random. After data augmentation, we obtain the final dataset of approximately 3200 data samples for each speech command.

3. **Feature extraction:** Each 2-second sample of data is a time-series signal, from which features are extracted to provide a deep learning network input.

The first step is to apply a pre-emphasis filter on the signal to amplify the high frequencies.

The second step, a STFT transform is used because spectral analysis show that different timbres in speech signals corresponds to different energy distribution over the frequencies. The speech signal is segmented into frames of 25ms with an overlap of 15ms for each of the frame. The winstep is 10ms (25ms -15ms) and NFFT = 512, therefore each 2-second audio will be split into 200 frames, each FFT frame has NFFT / 2 + 1 = 257 frequency bins. The spectrogram has shape = (257, 200).

The third step utilizes the mel scale, a psychoacoustic scale of pitches of sounds. It is a scale that more closely represent

what the human ears capture. Each spectrum frame is multiplied with the corresponding filter, then the results are added to get a filter bank response. So, with M filters, this results in M filter bank energy vectors on a frame. In this study, the value of M selected is 13.

Finally, the logarithmic spectrum of the mel scale is converted into the time scale by using the DCT. A cepstrum is the result of taking the inverse transform of the logarithm of the estimated spectrum of a signal. Apply DCT on the 13 Log Filterbank Energies x(n) to have 13 Mel-scale cepstral coefficients. For each frame of spectrogram, there are 13 Mel-scale cepstral coefficients, so MFCC features of a 2-second sample of speech signal is a two-dimensional array with a shape of (13, 200).

III. NEURAL NETWORK ARCHITECTURE

A. *Proposed architecture*

For Deep Networks, the focus is mostly on the two major architectures: CNNs for image modeling and Long Short-Term Memory (LSTM) Networks (Recurrent Networks) for sequence modeling.

The goal of a CNN is to learn higher-order features in the data via convolutions. They are well suited to object recognition of faces, individuals, street signs, platypus- es, and many other aspects of visual data. However, in [4,5], the authors have stated that one limitation of CNNs is that, with maxpooling, they do not consider the pose relationship between low-level features [8].

Recurrent Neural Networks are in the family of feed-forward neural networks. They take each vector from a sequence of input vectors and model them one at a time. This allows the network to retain state while modeling each input vector across the window of input vectors. Recurrent Neural Networks can have loops in the connections. This allows them to model temporal behavior and gain accuracy in domains such as time-series, language, audio, and text [8].

Long Short-Term Memory (LSTM) is a type of RNN architecture that addresses the vanishing/exploding gradient problem and allows learning of long-term dependencies. A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell. LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. Therefore, we use this LSTM architecture for training models in Automatic speech command recognition [8].

When using unidirectional RNNs as generative models, it is straightforward to draw samples from the model in sequential order. However, inference is not trivial in smoothing tasks, where we want to evaluate probabilities for missing values in the middle of a time series. In bidirectional RNNs, data processed in both directions processed with two separate hidden layers, which are then fed forward into the same output layer. Therefore, this can better exploit context in both directions. Hence bidirectional LSTMs usually perform better than unidirectional ones in speech recognition.

In this study, we implement and evaluate the performance of two models, unidirectional RNNs and bidirectional RNN for automatic speech command recognition.

B. *Implementation of the Neural Network*

To make sure the model works, it is first trained with a available dataset, the Google Speech Dataset V1 [9]. The Google Speech dataset V1 has 35 commands with audio files of 1 second in length. So, for each audio file, MFCC features is a two-dimensional array with a shape of (13, 100).

In the first model, unidirectional RNNs, speech command recognition depends on the data series over time, so on the top level of model, two LSTM layers are used to extract special features with long-term dependent of audio data. Then, the weighted average of the LSTM output is fed into 3 fully connected layers in the end for classification.

In the second model, bidirectional RNNs, two Bidirectional LSTM (BiLSTM) are used. BiLSTM contains two single LSTM networks that are used simultaneously and independently to model the input chain in two directions: from left to right (forward LSTM) and from right to left (backward LSTM). Finally, the weighted average of the LSTM output is fed into 3 fully connected layers for classification.'

In both model, activation of LSTM is a tanh function, and recurrent activation is hard_sigmoid function. All of these parameters are updated during the training process on the data sets labeled via the back-propagation algorithm with an Adam optimizer with learning rate of 0.001. The batch size used was 64. The LSTM model have 38,115 parameters and the BiLSTM model have 89,315 parameters.

C. *Experiments and model analyzing*

Each model was trained for a maximum of 10 epochs. The recognition results of both proposed models are compared with the results of Douglas Coimbra de Andrade et al. [10] as shown in Table 2 and Fig. 1. That comparison proves that both proposed models have very good results.

Table 2. Accuracy results on the Google speech command dataset V1

Model	Accuracy (%)	Trainable Parameters	Epochs
Douglas Coimbra de Andrade	94.3	202K	40
Douglas Coimbra de Andrade (V2)	93.9	202K	40
Unidirectional LSTM (ours)	92.1	38k	10
Bidirectional LSTM (ours)	94.6	89k	10

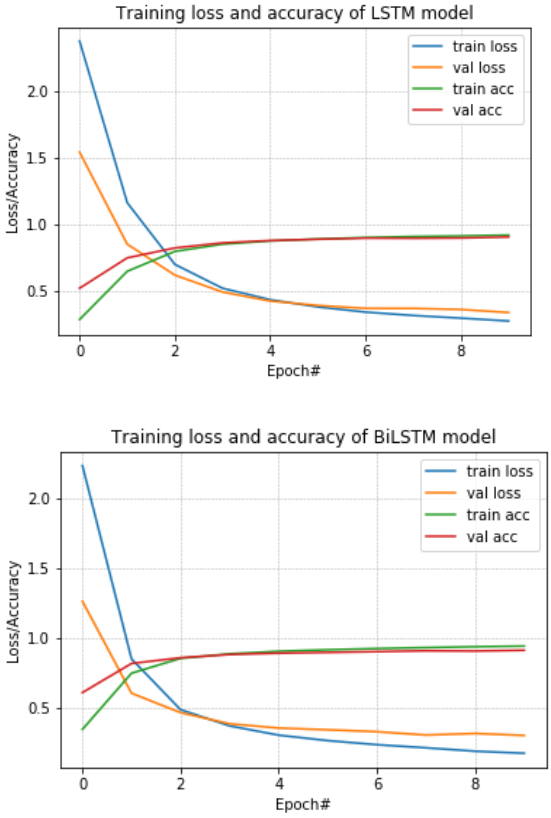


Fig. 1. Accuracy of LSTM (above) and accuracy of BiLSTM (below)

IV. APPLY TO RECOGNITION OF VIETNAMESE SPEECH COMMANDS

Using both of the above models for the collected Vietnamese command data set, the accuracy is shown in Fig. 2.

After 20 epochs, we can see that the BiLSTM model (98.19%) gives better results than the LSTM model (97.09%). However, there is overfitting in both models. Next, dropout regularization is used for reducing overfitting and improving the generalization of deep neural networks technology. Dropout is a technique where randomly selected neurons are ignored during training, i.e. "dropped-out". This means that their contribution to the activation of downstream neurons is temporally removed on the forward pass and any weight updates are not applied to the neuron on the backward pass. Dropout in Keras is implemented by randomly selecting nodes with a given probability (e.g. 20%) after each update cycle. This creates a small random amount of noise during learning and makes the architecture more flexible when processing speech data. The results of both models with dropout are shown in Fig. 3 which have significantly reduced overfitting.

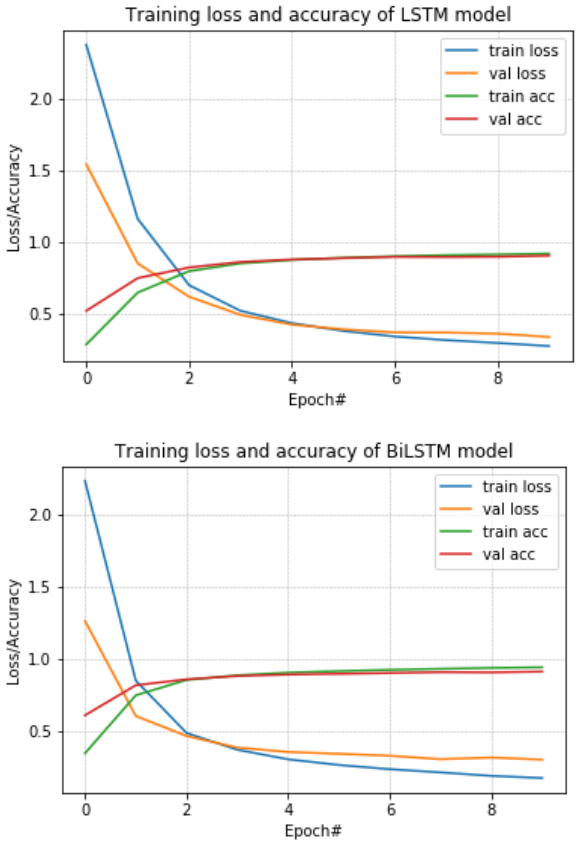


Fig. 2. Training loss and accuracy of LSTM, 97.09% (above) and of BiLSTM, 98.19% (below)

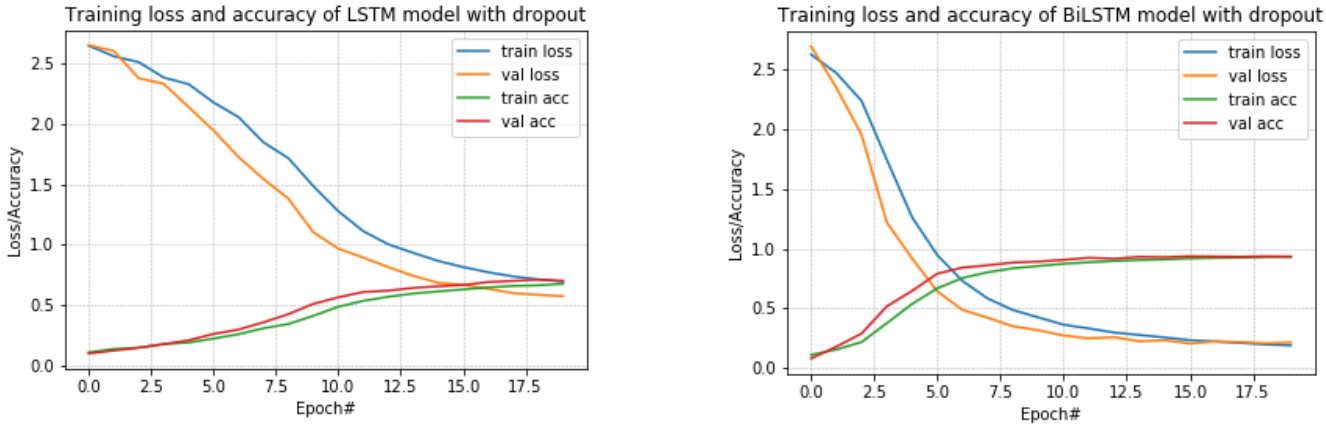


Fig. 3. Training loss and accuracy with dropout of LSTM (left) and of BiLSTM (right)

V. IMPLEMENTATION AND EVALUTION OF MACHINE LEARNING MODEL ON AN INTELLIGENT HOME CONTROL UNIT

Implementation Model for the Smart Home is described in Fig. 4, where a Raspberry PI computer V3 supports Wifi used as the home control unit. A microphone is connected via the USB port of Raspberry PI. This computer is packaged as Doraemon character. MQTT message Broker on PI will publish 15 topics corresponding to 15 commands [11]. The software on PI performs recording, identifying and publishing the topics shown in detail in Fig. 5. Distributed nodes in the home will subscribe to these topics to execute the corresponding command. The cheap and popular ESP8266 module can be used as a controller at a node [12].

The study also conducted performance tests with the following configuration:

- Raspberry Pi Model 3 Plus Rev 1.3 8
- Kernel: Linux 4.19.46-v7+
- Operation System: Raspbian GNU/Linux 9 (stretch)

The model after the train is of 1.1MB with 88688 parameters. Model when running accounts for 34-37% CPU and RSS (Resident Set Size [of Memory]) used in RAM is 218.7MB. Measure the main tasks time (record samples -> feature extract -> enqueue and dequeue-> predict-> publish) with the time library in python, Raspberry PI will complete command processing in less than 1 second. In fact, such processing time meets the needs of users.

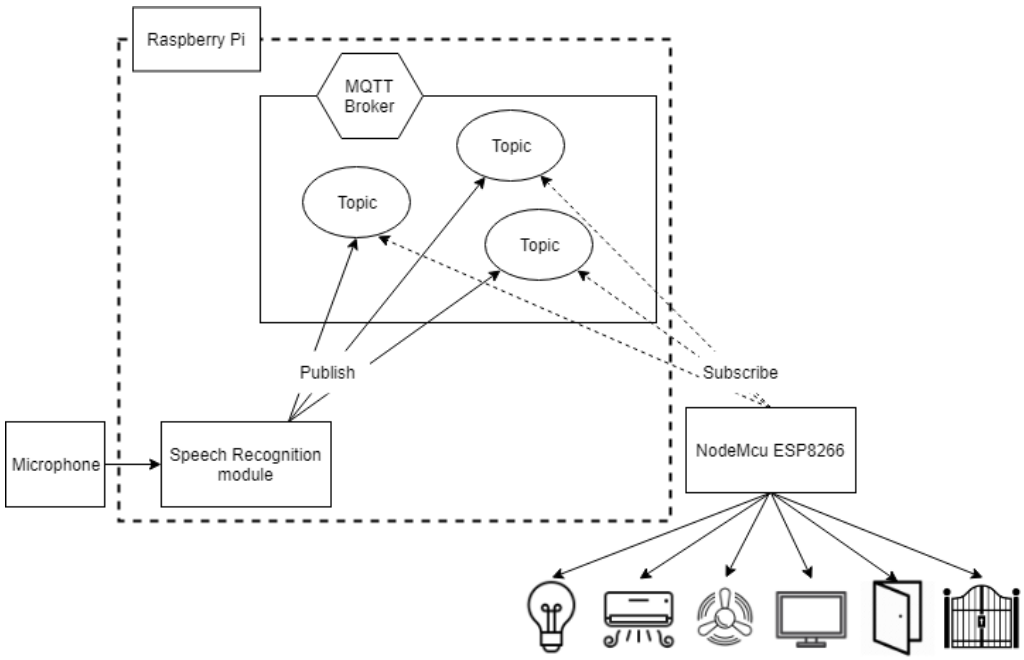


Fig. 4. Implementation Model for the Smart Home

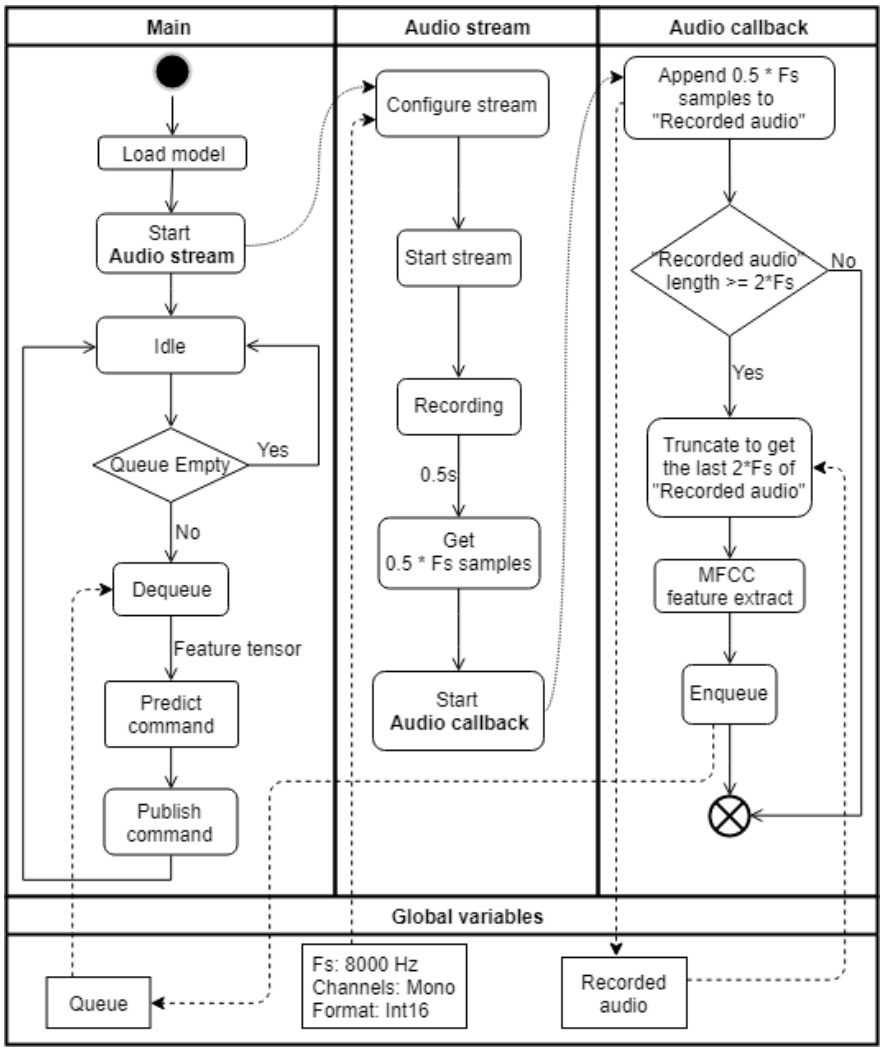


Fig. 5. Software on PI.

VI. CONCLUSION AND PERSEPECTIVES

After analysis and evaluation, we suggest using the BiLSTM model for recognizing Vietnamese speech commands. This work also contributes a Vietnamese command dataset including 15 commonly used commands for smart homes. In addition, research also give a software for collecting data on Android and iOS. The identification result of the BiLSTM model on this dataset is very good, with accuracy averaging 98,19%. The trained model is also implemented and evaluated on a Raspberry PI computer as an intelligent home control unit. The results show that the model is a good response to processing time and memory usage. The solution can be easily expanded on, for example adding commands, adding data. So future results can be improved to better meet actual problems. The solution can also be transferred to other languages. This work can serve as a good reference for many fields in Deep learning, for example, CNN [13], Pattern Recognition [14, 15,16], Optimization Methods and Regularization in Deep learning [17], etc.

REFERENCES

[1] https://www.statista.com/outlook/389/100/smart-appliances/worldwide#market-users (Accessed on 10 April 2019)

[2] Ribeiro, F. C., Carvalho, R. T. S., Cortez, P. C., Albuquerque, V. H. C. D, and Filho, P. P. R, "Binary Neural Networks for Classification of Voice Commands from Throat Microphone," in IEEE Access, vol. 6, pp. 70130-70144, 2018. doi: 10.1109/ACCESS.2018.2881199.

[3] Sidiq, M., Agung Budi, W. T., and Sa'adah, S., "Design and implementation of voice command using MFCC and HMMs method," 2015 3rd International Conference on Information and Communication Technology (ICoICT),

Nusa Dua, 2015, pp. 375-380. doi: 10.1109/ICoICT.2015.7231454.

[4] Guiming, D., Xia, W., Guangyan, W., Yan, Z., and Dan, L., "Speech recognition based on convolutional neural networks," 2016 IEEE International Conference on Signal and Image Processing (ICSIP), Beijing, 2016, pp. 708-711. doi: 10.1109/SIPROCESS.2016.7888355.

[5] Bae, J., Kim, D., "End-to-End Speech Command Recognition with Capsule Network". Proc. Interspeech 2018, 776-780, DOI: 10.21437/Interspeech.2018-1888.

[6] https://github.com/VSC-FU2019/VSC_FU

[7] Smith III, J. O., "Spectral audio signal processing", https://www.dsprelated.com/freebooks/sasp/. (Accessed on 10 Avril 2019).

[8] Josh, P. and Adam, G., "Deep Learning, A Practitioner's Approach", Chapter 4, 2017, O'Reilly Media, Inc.

[9] https://www.tensorflow.org/tutorials/sequences/audio_recognition (Accessed on 10 Avril 2019).

[10] Andradea, D. C. D., Leob, S., Da Silva Vianac, M. L., Bernkopf, C., "A neural attention model for speech command recognition", 2018, https://arxiv.org/pdf/1808.08929.pdf

[11] Tutorial, https://appcodelabs.com/introduction-to-iot-build-an-mqtt-server-usingraspberry-pi (Accessed on 10 Avril 2019).

[12] Tutorial,https://techtutorialsx.com/2017/04/09/esp8266-onnecting-to-mqtt-broker/ (Accessed on 10 Avril 2019).

[13] Nam, N. T., Hung, P. D., "Pest detection on Traps using Deep Convolutional Neural Networks," in Proceedings of the 2018 International Conference on Control and Computer Vision (ICCCV '18). ACM, New York, NY, USA, 33-38. DOI: https://doi.org/10.1145/3232651.3232661.

[14] Hung, P. D., Linh, D. Q., "Implementing an Android Application for Automatic Vietnamese Business Card Recognition", Pattern Recognit. Image Anal. (2019) 29: 156. https://doi.org/10.1134/S1054661819010188.

[15] Hung, P. D., "Detection of Central Sleep Apnea Based on a Single-Lead ECG," in Proceedings of the 2018 5th International Conference on Bioinformatics Research and Applications (ICBRA '18). ACM, New York, NY, USA, 78-83. DOI: https://doi.org/10.1145/3309129.3309132.

[16] Hung, P. D., "Central Sleep Apnea Detection Using an Accelerometer," in Proceedings of the 2018 International Conference on Control and Computer Vision (ICCCV '18). ACM, New York, NY, USA, 106-111. DOI: https://doi.org/10.1145/3232651.3232660.

[17] Nam, N. T., Hung, P. D., "Padding Methods in Convolutional Sequence Model: An Application in Japanese Handwriting Recognition," in Proceedings of the 3rd International Conference on Machine Learning and Soft Computing (ICMLSC 2019). ACM, New York, NY, USA, 138-142. DOI: https://doi.org/10.1145/3310986.3310998.