PAPER • OPEN ACCESS

Smart Medical Chatbot with Integrated Contactless Vital Sign Monitor

To cite this article: Wan Muhamad Asyraf Wan Zaki *et al* 2019 *J. Phys.: Conf. Ser.* **1372** 012025

View the article online for updates and enhancements.

Recent citations

- Andrew Reyner Wibowo Tjiptomongsoguno *et al*

IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection–download the first chapter of every title for free.

Smart Medical Chatbot with Integrated Contactless Vital Sign Monitor

Wan Muhamad Asyraf Wan Zaki[1], Muhammad Faiz Md Shakhih[1], Muhammad Hanif Ramlee[1,2,] and Asnida Ab Wahab[1,2*]

[1]School of Biomedical Engineering and Health Science, Faculty of Engineering, Universiti Teknologi Malaysia, 81310, Johor, Malaysia
[2]Medical Device and Technology Centre (MEDITEC), Institute of Human Centered Engineering (iHuman), Universiti Teknologi Malaysia, 81310, Johor, Malaysia.

*asnida@biomedical.utm.my

Abstract. Artificial intelligence is one of the important fields in modern technologies to help us strive for better life. Healthcare industries nowadays spend a lot of money researching on how artificial intelligence can help improve their services and give the highest satisfaction to their customers. Most healthcare organisations have a passive relationship to their patients when it comes to communication and this situation is often worsened because of a lack of inter-operability between client and provider. Mobile applications on the other hand have become one of the effective strategies in bridging the interaction between provider and end user. In this study, an automated self-learning system is designed to provide conversational healthcare for personalised proactive experience. This system is developed along with the in cooperation of contactless monitoring device using a vision-based real-time monitoring of vital signs which allow patients to monitor their oxygen level, heart rate and respiration rate. This system is also automatically calibrated across patients, allowing precise measurement using highest probability method and natural language processing. Results obtained from the comparative analysis show a promising result with an error of 1.16 for pulse sensor and 2.917 for ECG which are below the threshold error. This allows user to accurately measure vital signs in a non-obtrusive way, and to provide them with the data required to determine to the right timing for any intervention procedure needed. The developed system would also help to bridge the gap of interoperability between client and medical provider

1. Introduction
One of the requirements for general artificial intelligence (AI) is to interact with the machine through natural language. This AI field includes dialog systems, chatbots or spoken dialog systems. The chatbot is also drawing upon ever-growing medical question range, to broaden its already significant wealth of medical expertise. Many seemingly static scenes contain subtle changes that are invisible to the naked human eye. However, it is possible to pull out these small changes from videos through the use of algorithms via motion magnification [1]. Motion magnification gives a way to visualize these small changes by amplifying them and to pull out interesting signals from these videos, such as the human pulse [2]. Motion magnification techniques can be divided into two categories which are Lagrangian and Eulerian approaches [5,4,3]. Lagrangian approach explicitly extracts the motion field (or optical flow) and uses it to move the pixels [10]. Lagrangian perspective is to analyse the angle of motion of the pixels (particles) of interest in the tracking image [6]. However, the Lagrange perspective approach has the following shortcomings where it is necessary to accurately track and estimate the trajectory of

particles, which requires more computational resources. Unlike the Lagrangian perspective, the Euler perspective does not explicitly track and estimate the motion of the particle, but instead fixes the perspective in one place, such as the entire image. In 2012, Wu et al. started from this perspective and proposed a method called Eulerian Video Magnification [7]. Eulerian techniques are useful for revealing subtle motions, but they are hand-designed, and do not take into account many issues such as occlusion. Wadhwa et al. improved the technology in 2013 and proposed a phase-based image motion processing technique [8]. In this study, the eulerian technique is used to extract the heart rate of the target person, where occlusion does not affect the end result much. Current systems utilize various sensors which require constant and secure physical contact with the patients during monitoring process and have a passive relationship to their clients when it comes to communication. This often worsens the problem because of a lack of inter-operability between client and provider [9]. The use of personal health assistants on our phone has suddenly become definite responses for certain needs, allowing precise measurement without the need for individual adjustments and contactless vital sign acquisition. However, for patient with specific conditions, a less intrusive monitoring method is critically required to avoid any unnecessary distress to the patient [10]. This study is focus on chatbot operation and process, where the text presented to them by the user (a process known as "parsing"), inferring what user mean and/or want, and determine a series of appropriate responses based on that patient information and vital sign prior to responding to a series of algorithms that interpret and identify the common disease such as common cold (viral rhinopharyngitis), tension-type headache (TTH) and migraine with exception of cancer related disease.

2. Methodology

2.1 Database and Automated Diagnosis
The system encompasses of state of art chatbot interface along with vision based vital sign solutions that offers user to accurately measure vital signs in a non-obtrusive way, and to provide them with the data required to determine the need for any intervention procedure [11]. The method of database recall for disease identification, as shown in Figure 1 which was similar in ideas and formulaa to those of the "maximum likelihood method" and (head, ear, eye, nose and throat) HEETN exam, were applied to the automated diagnosis. The likelihood was weighted with a prior probability of each disease. Clinical data include quantities that are continuous, discrete, and bivariate [12]. Each of the data is therefore considered to be bivariate in terms of quantity either a symptom occurred or not. The format item that is specified here and used directly for automated diagnosis is designed to select from the three: YES, NO, and UNKNOWN. In this way the answer can be chosen.
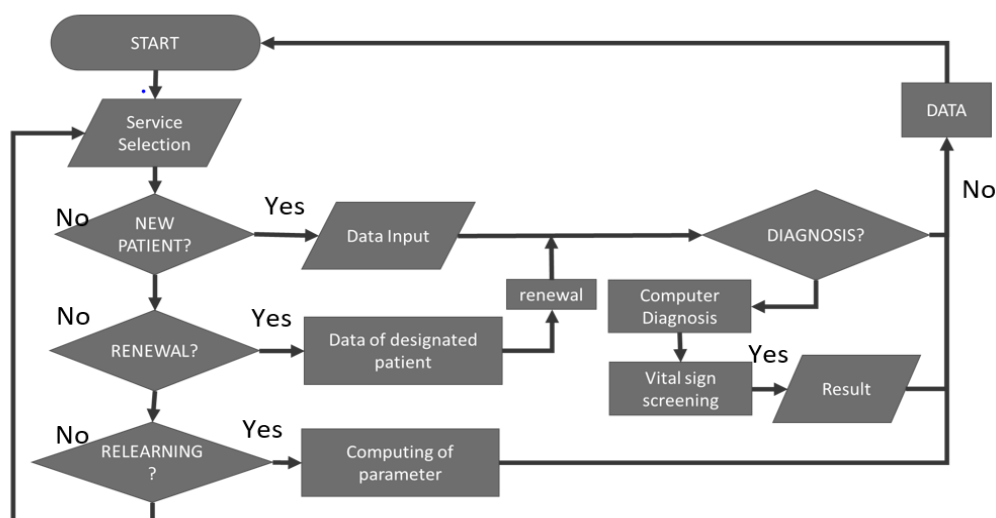

Figure 1. Flowchart of Automated Diagnosis.

2.2 Motion Magnification

Heart rate is measured from subtle change in colour and movement under skin epidermis, due to variation in volume and oxygen saturation of the blood in vessel at every heartbeat. Each cardiac cycle appears in peak within small range e.g. 50 and 90 bpm.

2.3 Proposed Algorithm Pipeline

The automated diagnosis system is integrated with a chatbot and motion magnification for vital sign sensing. The architecture, as shown in Figure 2, consists of user interface and server with different sets of parameters. The left one (machine learning) is called the encoder, while the below one (corresponding logic and rule parser tokens) is called the decoder. The encoder machine learning conceives a sequence of context tokens one at a time and updates its hidden state. After processing the whole context sequence, it produces a final hidden state, which incorporates the sense of context and is used for generating the answer. For this purpose, a SoftMax layer over vocabulary of HEETN exam is maintained in the decoder machine learning [12]. At each time step, this layer takes the decoder hidden state and outputs a probability distribution over all words in its vocabulary.
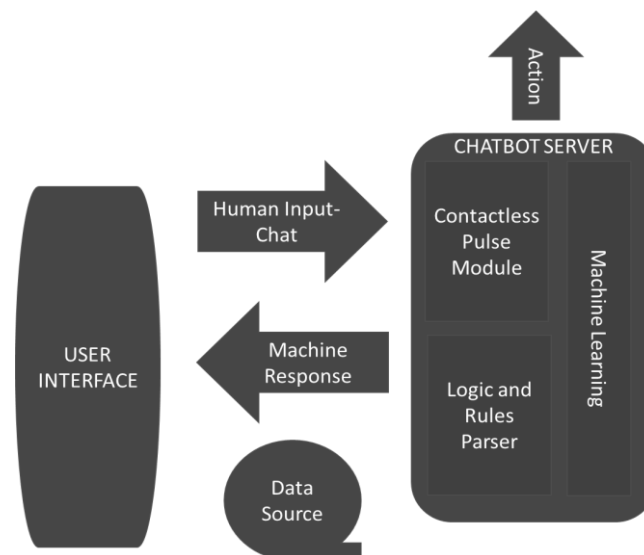

Figure 2. Overview of the system, which compromise of user interface and a chatbot server that contain principle component to realise automated diagnosis.

First, decoder hidden state is initiated with final encoder hidden state, $h_0^{dec} = h_n^{enc}$. Then, logic and rule parser token are passed as first input to the decoder and update first hidden state, $h_0^{dec} = rnn_\theta[h_{n-1}^{dec}, \omega_t]$. Next is to sample (or take one with max probability) first word from first SoftMax layer $\omega_{t+1} \sim p_{t+1} = softmax(g_\varphi(h_t^{dec}))$. Finally, pass this word as input, update hidden state and generate new word. Repeat step 4 until logic and rule parser token is generated or maximum answer length is exceeded. More simply, the network predicts the next word in the sequence by providing it with a correct prefix. Training is performed via maximum likelihood training, which leads to classical cross-entropy loss [12]:

$$L = \sum_{t=1}^{m} \sum_{i=1}^{|V|} I[y_t = i] log\hat{p}_{t,i} = \sum_{t=1}^{m} log\hat{p}_{t,y_t} \qquad (1)$$

Where, y_t is a correct word in reply at time step *t*.

2.4 Contactless vital sign monitor

As mentioned above, the first step of image amplification technique is to spatially filter the video sequence to obtain basebands of different spatial frequencies. The topmost image, that is, the image that in contrast with changes in blood volume in a portion of the peripheral microvasculature had the lowest spatial frequency and the highest signal-to-noise ratio, hence can use the maximum magnification, and the magnification of the next layer is sequentially decreased. In addition, it is easy to approximate the heart rate signal where images with higher spatial frequencies, such as the original video image, approximate with Taylor series expansion for reconstruction, as shown in Figure 3. In a nutshell, the contactless vital sign monitor can be visualised in block diagram as follows:
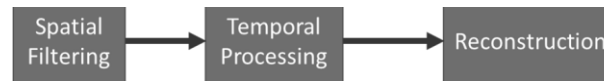

Figure 3. Step for vital sign via Eulerian video magnification.

3. Result and Discussion

Comparative analysis between our developed method and existing methods was performed. Based on Figure 4, vital sign result was based on the Eulerian magnification, specifically the heart rate measurement. All results in this section were processed with temporal filters unless otherwise noted.
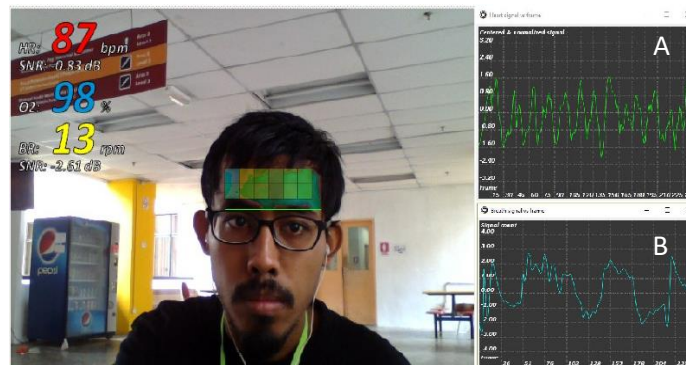

Figure 4. User Interface of the Contactless Vital sign Monitor. Graph A shows the heart rate while graph (B) shows the breath rate.

3.1 Disease identification accuracy

Disease identification metrics such as Dice indexes (DI), Jacard indexes (JI), false positive (FP) and false negative (FN) were used to validate the performance of proposed automated diagnosis of common disease. In this test, performance of the developed system on common cold (viral rhinopharyngitis), TTH and migraine as shown in Table 1 were measured

Table 1. Performance of measured result

Disease	Acc (%)	Precision	Sensitivity	Specificity	F-Measure	DI	JI
C. Cold	96.667	0.9678	0.9670	0.0126	0.9583	0.9753	0.9711
TTH	93.333	0.9357	0.9340	0.0187	0.9096	0.9640	0.9390
Migraine	89.876	0.8850	0.8825	0.0346	0.8550	0.8976	0.8314

Based on Table 1, it can be seen that migraine had lower accuracy due to a lot of subclasses for the machine to pin down exact root symptom. Yet, the system still performs better than threshold accuracy, which is over 96.667% for common flu, and 93.333 for TTH. The overall system had more than 0.85 true positive or sensitive, which provide high probability sample for disease identification.

3.2 Comparison with pulse sensor and electrocardiogram

Based on comparison with pulse sensor and ECG system in Figure 5, it is found out that the developed system had mean difference of 1.25 for pulse sensor and 2.92 for ECG for the heart rate measurement. This is because, our contactless system suffers from outdoor occlusion and changing intensity, but with closed room, the difference might be lower.
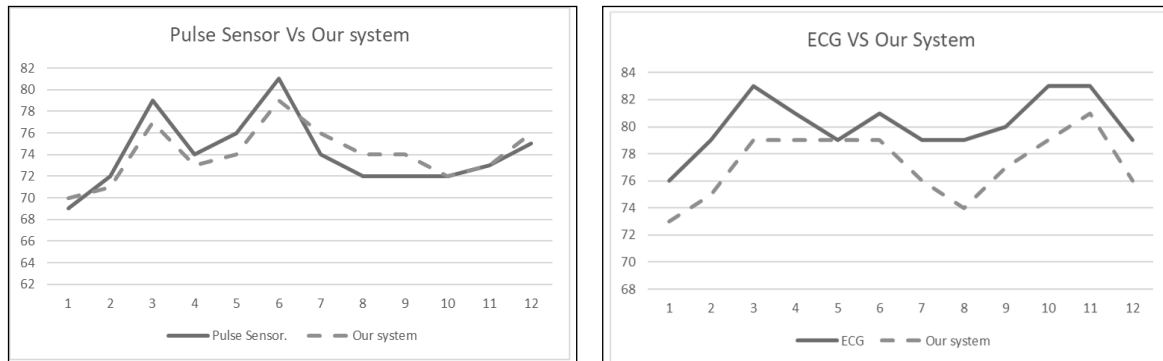

Figure 5. The results of 12 sample readings from pulse sensor, ECG and our system.

Mean average error (MAE) was calculated to determine the difference between two continuous heart rate measurements. This was performed to obtain the error contributes to MAE in proportion to the absolute value of the error. In addition, the feasibility of the system can be observed in ways that disregard the direction of which a measure that place emphasis on this is significant on same scale as the data being measured [13]. The formula to calculate MAE is as follows:

$$MAE = \frac{1}{n} \sum |Actual - Forecast| \qquad (2)$$

From the pulse sensor, the MAE attained is 1.167 and for ECG is 2.917. This is obvious because pulse sensor calculates heart rate based on constitutes another means of determining the timing of cardiac cycles via continuous monitoring of changes in blood volume in a portion of the peripheral microvasculature while ECG measure the bio-potential generated by electrical signals that control the expansion and contraction of heart chambers. ECG should have highest accuracy to benchmark for heart rate, but as per monitoring and initial diagnosis, the threshold for error is less than 3 in which the system performs above the threshold [14].

4. Conclusion

In terms of computer diagnosis or medical information system, the diagnostic system with learning capability is very useful. As regards, the automated diagnosis, bivariate in terms of quantity either a symptom occurred or not, using the highest probability method where there are many diseases need to be identified, with 99.667% for common cold, 93.333% for TTH and 89.876% for migraine, method of maximum probability is practical and powerful [15]. With the in cooperation of contactless vital sign monitor, an error of 1.16 for pulse sensor and 2.917 for ECG which are still below threshold error offers user to accurately measure vital signs in a non-obtrusive way, and to provide them with the data required to determine to the right timing for any intervention procedure. This system helps bedridden patients who usually stay at home yet require a constant health monitoring without having to wear or attach to bulky medical equipment [16] which can bridge the gap of inter-operability between client and medical provider.

Acknowledgments

References

[1]   Zhang, Y., Pintea, S.L., van Gemert, J.C. 2017. Video Acceleration Magnification. *IEEE Conf. on Computer. Vision and Pattern Recognition.* 502-510

[2]   Brandtzaeg, P. B., & Følstad, A. 2017. Why People Use Chatbots. *Internet Science Lecture Notes in Computer Science.* **10673**, 377-392.

[3]   Wu, H.Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., Freeman, W. 2012. Eulerian video magnification for revealing subtle changes in the world. *ACM SIGGRAPH 31.* **4,** 65–8

[4]   Wadhwa, N., Rubinstein, M., Durand, F., Freeman, W.T. 2014. Riesz pyramids for fast phase-based video magnification. *IEEE Int. Conf. on Computer. Photography.* 1-10

[5]   Wadhwa, N., Rubinstein, M., Durand, F., Freeman, W.T. 2013. Phase-based video motion processing. *ACM Trans. Graph.* **4**, 80

[6]   Liu, C., Torralba, A., Freeman, WT, Durand, F. and Adelson, E.H. 2005. Motion magnification. *ACM Transactions on Graphics (TOG),* 519–526

[7]   Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F. and Freeman, W. 2012. Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics (TOG).* **4**, 65.

[8]   Fujita, H., Selamat, A., & Omatu, S. 2017. New Trends in Intelligent Software Methodologies, Tools and Techniques *Proceedings of the 16th International Conference SoMeT_17.* (Amsterdam: IOS Press, Incorporated). 139 - 152

[9]   Wootton, R., & Batch, J. 2005. Telepediatrics: Telemedicine and child health. London: *Royal Society of Medicine Press.* **4**, 286–293

[10]   Reisner, A. T.; Shaltis, P. A.; McCombie, D.; Asada, H. H. 2008. Utility of the Photoplethysmogram in Circulatory Monitoring. *Anesthesiology.* **8**, 950-958.

[11]   Hibbard JH, Stockard J, Mahoney ER, Tusler M. 2004. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. *Health Serv Res.* **4**, 5 – 26

[12]   Alepis, E., and Lambrinidis, C. 2013. M-health: supporting automated diagnosis and electonic health records. *SpringerPlus*, **2**,103.

[13]   Hyndman, R. and Koehler A. 2005. Another look at measures of forecast accuracy. *International Journal of Forecasting* **4**, 679-688

[14]   Valenza, G., Citi, L., & Barbieri, R. 2017. Linear and Nonlinear Parametric Models in Heart Rate Variability Analysis. *ECG Time Series Variability Analysis*, 87-116.

[15]   Yeung, D. 2017. Keynote. *Proceedings of the 2nd Workshop on Deep Learning for Recommender Systems - DLRS 2017.*

[16]   Tasli H.E.; Gudi, A.; den Uyl, M. 2014. Integrating Remote PPG in Facial Expression Analysis Framework. *Proceedings of the 16th International Conference on Multimodal Interaction, Istanbul*, 74-75.