Speech Recognition Driven Assistive Framework for Remote Patient Monitoring

Marc Jayson Baucas and Petros Spachos
School of Engineering, University of Guelph, Guelph, ON, Canada

Abstract—**Health care resources have started to become scarce due to their increase in demand. Hospitals have begun to run out of space, forcing them to deny admission of patients. Remote Patient Monitoring (RPM) has the potential to help citizens who suffer from chronic diseases and provide environments were easy to access healthcare is available. RPM allows people to receive the same amount of care without having to difficulties to find a spot at a hospital ward. However, some roadblocks end up preventing RPM from being implemented by more healthcare providers. Data integrity, user privacy, and high power consumption are some of these concerns. With data transmission and transaction, privacy and confidentiality have always been an issue. High power consumption is a concern due to RPM's demand for continuous data collection. This paper proposes a framework that reinforces the RPM system to address these concerns. The design not only allows better data filtering for privacy but also a more responsive system with the use of controlled surveillance and speech recognition. Overall, this framework provides an opportunity for RPMs to be a viable implementation for healthcare providers.**

I. Introduction

Due to the ageing population, the number of people admitted to hospitals has increased. With the increase of those in need of special care, human resources and space in hospitals are starting to become scarce. Over the past decade, most improvements in healthcare services have originated from the integration of technology [1], [2]. Many healthcare-related establishments have turned to this fast-evolving market as a source of solutions for most of their concerns. Technology is being used to create a system that allows patients to receive the same amount of care at low resource cost. As a result, it has given rise to many solutions that attempt to alleviate this strain.

One of the promising solutions is Remote Patient Monitoring (RPM) systems [3]. An RPM system can provide high-quality care to patients such as the elderly or the chronically ill even within their own homes [4]. This system is called telemonitoring. In this system, patients are given devices to wear that check on specific vitals based on their current illness. This information is then sent by these devices (i.e. Electrocardiography (ECG) and Heart Rate monitor) to their respective health care providers [5]. The service providers then respond accordingly based on the data they received from each patient.

RPM can potentially provide benefits to not only the health care providers but also the patients. First, RPM can be more convenient and cheaper for patients as well as health care service providers in terms of general and specific care [6]. Patients are allowed to remain in their homes while being able to receive professional care. At the same time, patients that live in far-flung areas no longer need to travel long distances to receive immediate care [6]. This benefit also gives the ability for health care providers to extend their services to the citizens who are unable to be physically present in their clinics for checkups. Lastly, healthcare providers can detect any anomalies that might be wrong in a patient's condition [4]. This benefit significantly increases the chances of early detection for emergencies, which in turn helps the patient in the long-term.

Though RPM provides convenience and benefits to its patients, it also comes with some concerns. Firstly, the integrity of the data is significant to RPMs. RPMs give healthcare providers a chance to detect anomalies in a patient's health earlier. However, there is no way to determine the integrity and accuracy of the data clinics receive from their devices in its current state [7]. Data that is normal can end up being transmitted as an emergency once it loses in integrity. Integrity is lost if data is reported differently from its original value. Mistakes in communicating or translating data could result in a loss of resources or even the patient's life.

Privacy is also a big issue for most patients [8]. It is less likely for a health care provider to implement RPM systems due to this issue [6]. Healthcare services are held responsible for keeping their patient's information private and confidential. A faulty system can easily misinterpret wirelessly transmitted data. Malicious individuals are the biggest offenders in misusing this information.

Another issue is high power consumption. RPMs make use of wireless devices to transmit vitals from the patient to the medical clinics. Continuous transmission of data over wireless networks can result in higher power consumption. Batteries power most monitoring devices that are used by health services. In situations when batteries are quickly depleted by continuous use or without a replacement, the patient cannot be monitored or diagnosed.

This paper introduces a framework that addresses the data integrity and privacy concerns among patients as well as the high power consumption issues of health monitoring devices to create a better system. The rest of this paper is as follows: An overview and in-depth discussion of the system and its different aspects in Section II. Section III reports some preliminary results of the proposal. Finally, Section IV are the conclusions to the paper.

II. System Overview and Design

This section discusses the overall design of the system, highlighting its key aspects and how each works to create the proposed framework.

A. Overview

The proposed framework aims to address the concerns mentioned in the previous section. Mainly, it addresses issues with data integrity and privacy as well as improving its overall power consumption. The framework is a system that takes the sensing abilities of devices partnered with neural networks to allow a more interactive response system for patients in times of emergencies. Most RPMs have patient data being streamed all the time to the doctors and physicians. Such a design resulted in an issue with privacy and patient confidentiality. RPMs need to monitor their patients' vitals 24/7 to catch any emergency. However, this results in high power consumption and exposes patient data.

The proposed framework reduces these concerns by integrating a system that records and streams data only when given the authorization by the system during emergencies. The idea is to have a system that will interact with the patient during an emergency. If the patient permits or does not respond for a given time frame, then a video feed revealing the patient's current status will be sent to the clinic. It is then up to the clinic to decide the best action during the presented situation.

This system provides proper controlled surveillance for the patient. Also, it maintains privacy for its users while saving power by only recording and transmitting data if prompted. The system starts by staying in low powered status or sleep mode. Most of its functionality is turned off and will only activate once triggered. The system serves as a filter and will only send the data to the clinic if any anomaly or significant change is detected. This design creates a more confidential system that only cares about the well-being of the patient without getting too much data or using up too much power. Upon detection of an anomaly, the system decides on its severity. If the system determines an emergency, a listening device placed within the household of the patient interacts with the user. It then asks the user if medical attention is needed or if the trigger is just an alarm.

Overall, this creates a framework that improves RPM systems in terms of maintaining the data integrity and privacy for each patient while keeping the system energy efficient.

B. Related Literature

The first two issues were the integrity and privacy of the patients' data. Previous works attempt to create a mobile RPM system that utilizes wireless technology to expand the scope of healthcare services. In [9], they proposed a wireless personal digital assistant (PDA) which allowed real-time transmission of a patient's vitals to authorized medical staff. However, their design only focuses on reacting based on incoming data. As a result, data integrity becomes crucial. Without a proper interface, compromised data could be misinterpreted, resulting in situations not handled appropriately. Instead of relying heavily on the transmitted data from the PDA or the transducer, our framework could potentially make the system more reactionary. With our design, clinics can now administer help efficiently and monitor patients during emergencies by providing a smarter system that can diagnose the situation before data is transmitted. The system can now pre-process data before it is forwarded to detect emergencies earlier. We can protect a patient's privacy by only collecting information if an anomaly or an emergency is detected.

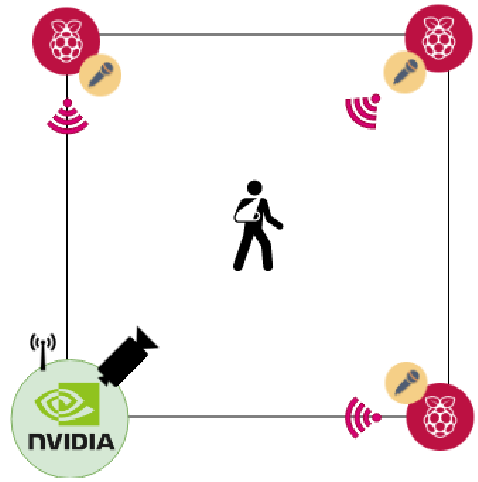

Fig. 1: Proposed RPM room setup.

Another mentioned issue with current RPM systems and health monitors is high power consumption. In [10], they proposed a low-power ECG monitoring system using Internet of Things(IoT) network. However, their design relies on the health monitor alone to provide reliable data which increases the power consumption of the device. In our introduced design, power consumption is less of a concern since most of the system is on low power unless triggered by the health monitor during emergencies. We use IoT to drive the network and wake-up logic to enable low-energy states. At low power mode, the system becomes a data filter for the health monitor before a patient's information is sent to the healthcare providers. As a result, power is conserved. Also, data can be easier to manage by reducing the transmission rate while maintaining the quality of service. Overall, our proposed design can act as an extension to previous works on wireless health monitors by providing a smart system that can manage each device efficiently.

C. Design

The proposed framework is shown in Fig. 1. It is composed of three major sections:

i. Sound recording,
ii. surveillance capture, and
iii. speech classification.

With the presence of an already existing health monitor, this framework serves as a filter between patient and healthcare service. It only reports data that is needed by the patient for treatment and maintenance.

D. System components

We used a combination of hardware and software components that build towards our framework. In terms of hardware, the system uses three Raspberry Pis and an NVIDIA Jetson TX1 developer kit. The Raspberry Pis are the sensing devices that will be placed strategically around the room. They record the patient's response based on the health monitor data. They can also filter out the data being sent by the monitor to the clinic. We loaded the Raspberry Pis with a Raspbian operating system. The Jetson board already came with a preloaded with an Ubuntu 16.04 operating system which we used in this prototype. The sound is recorded with the use of an STM32 NUCLEO-64 board with an attached X-NUCLEO-CCA02M1 expansion board to serve as a digital MEMS (Micro-Electrical-Mechanical System) microphone. This microphone is connected to the Pi and controlled using a python script. The code used in recording the sound uses the PyAudio library. This library allowed us to configure the recording stream to match the default configurations of the MEMS microphone that we used. By default, the microphone records with a sampling rate of 16 kHz and a 16 bit sound resolution. A patient can be monitored effectively by having the three Raspberry Pis placed on each corner. Each Pi will wirelessly transmit the sound file to the Jetson board for processing. The Jetson serves as the central processing hub for the system due to its computing capabilities via its built-in GPU.

Before the sound gets classified by the model, a script will merge the sound clips into one waveform. This classifier is a Convolutional Neural Network (CNN). It was the selected network due to its ability in creating acoustic models [11]. CNN is also known as a powerful toolkit for speech recognition [12]. The CNN used in this framework is coded using Python. The code makes use of a combination of neural network training libraries; Tensorflow and Keras. It also uses Scipy and Numpy for its sound processing and feature extraction. The CNN makes use of 4 types of layers to build the architecture: Convolutional Layer, Pooling layer, Normalization layer, and Full-Connected or Dense layer. The neural network was built using blocks. Each block contained; the convolutional layer, pooling layer, and the normalization layer. The input starts by going through three of the previously mentioned blocks. Then, the network flattens the output before the dense layer uses it for predictions based on the trained model.

We chose to use a CNN due to our intention to train our model to detect more complex terms for future iterations. During our system's preliminary stages, the classifier is only prepared to respond to two keywords; "yes" or "no" response while ignoring everything else. This design provides a simpler model for proof of concept. We aim to not only cover simple answers. For future work, we plan to create a fully programmable interface that can react to more situations.

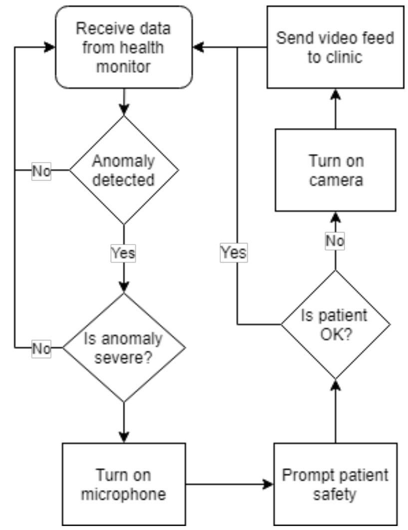

Fig. 2: General design flow.

E. Design flow

The general design flow of the system is shown in Fig. 2. The system starts with the Pi detecting an anomaly from the data sent by the worn health monitor. Based on the issue, the Pi reports it and does either one of two things; returns to data monitoring or interacts with the patient. If the anomaly is severe, the system asks if the patient needs immediate assistance. The Pi then switches to active mode by turning on its microphone peripheral to listen to the patient's response. If the patient does not need help, then the system returns to low power mode. However, if the patient does need assistance or does not respond within 15 seconds of the prompt, then the Jetson is prompted to turn on its built-in camera. This camera is used to survey the patient in the room. The video feed is sent to the clinic to show the patient's current state. The clinic is then expected to respond accordingly to how they evaluate the situation.

In terms of the system prototype, we created a recording mechanism using the Raspberry Pi and the digital microphone. The module can record sound data and send it to the NVIDIA Jetson board via wireless socket transmission. The system transmits the sound data as packets and is rebuilt as a file once it reaches the Jetson board. A detailed representation of the operation flow of the Raspberry Pi in the design is shown in Fig. 3. We also set the trained classifier within the Jetson board. Its job is to process the sound file once the Raspberry Pis have completed their transmission. Based on the results of the classifier, the board with either do one of two things; turn on the camera or return to low-power mode while waiting for the next transmission. A detailed representation of the operation flow of the Jetson in the design is shown in Fig. 4.

III. Preliminary Results

This section provides any results from the initial prototyping of the system. In terms of the classifier, the model is trained to detect either a keyword of "yes" or "no".

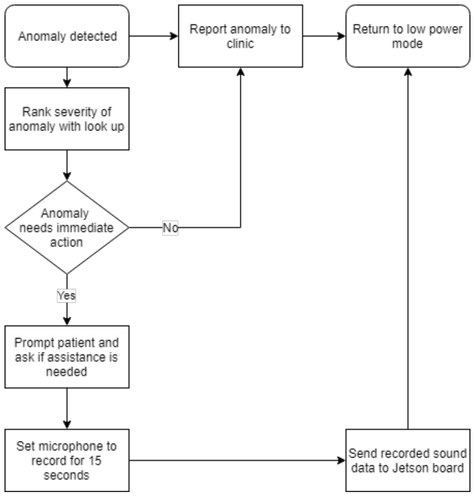

Fig. 3: Raspberry Pi operation flow.

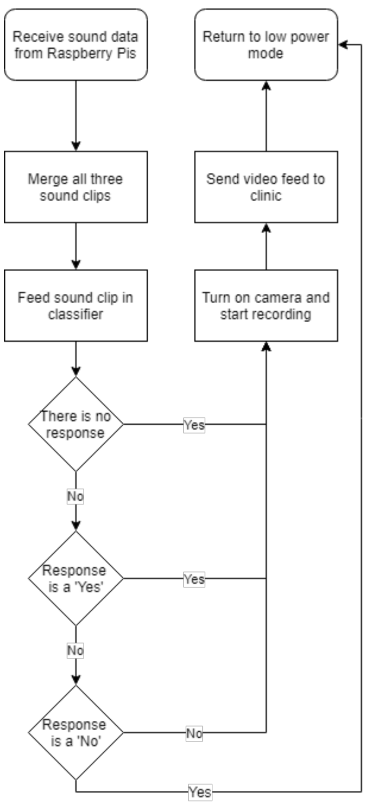

Fig. 4: Jetson operation flow.

We conducted the training of the model with the Speech Commands Dataset provided by Google.

During training, the dataset was split 70-30; for training and testing respectively. The results where the verification accuracy reaches the intended accuracy are shown in Fig. 5. The results where the verification loss approaches the accounted loss are shown in Fig. 6. As it is shown in the figures, as the number of steps approached an epoch value of 12, the accuracy and loss started to fit within their verification

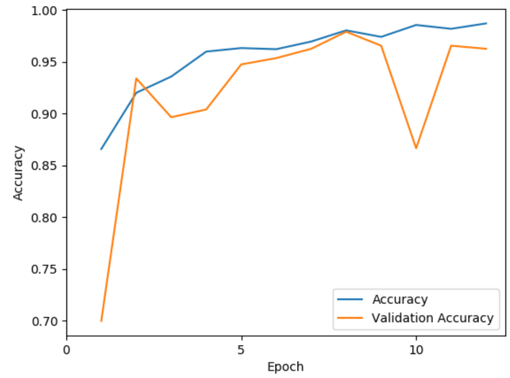

Fig. 5: Training accuracy and validation.

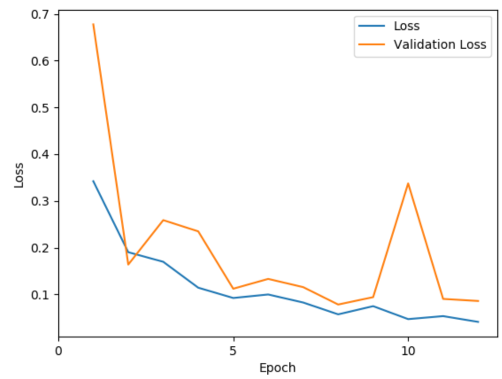

Fig. 6: Training loss and validation.

curves. This behaviour indicated the increase in the model's precision in classifying the intended keywords. The resulting classification accuracy of the trained model is 97% and further supported by a subsequent loss of 3%, which is promising. To examine the feasibility of the system further, we have identified that should conduct more experiments in different rooms and under different conditions.

IV. Conclusion

Data integrity, user privacy, and high power consumption is a concern for RPM systems. Any healthcare provider who wants to implement an RPM system should first address these issues. We proposed a framework that can be integrated into an RPM system to build towards minimizing these concerns. This framework is a sensing system that makes use of speech recognition to create an interactive design that caters to patients during emergencies. Also, this system creates a filtering mechanism which could potentially improve the privacy and confidentiality of any patient's data. The classifier used in this framework is a CNN that yielded an accuracy of 97% and supported by a training loss of 3%. Further experimentation is needed to evaluate the system under different conditions.

References

[1] L. Catarinucci, D. de Donno, L. Mainetti, L. Palano, L. Patrono, M. L. Stefanizzi, and L. Tarricone, "An iot-aware architecture for smart healthcare systems," *IEEE Internet of Things Journal*, vol. 2, no. 6, pp. 515–526, Dec. 2015.

[2] B. Xu, L. D. Xu, H. Cai, C. Xie, J. Hu, and F. Bu, "Ubiquitous data accessing method in iot-based information system for emergency medical services," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1578–1586, May 2014.

[3] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 1, pp. 156–167, Jan. 2006.

[4] T. Bratan and M. Clarke, "Optimum design of remote patient monitoring systems," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2006, pp. 6465–6468.

[5] P. Zhang, Y. Kogure, H. Matsuoka, M. Akutagawa, Y. Kinouchi, and Q. Zhang, "A remote patient monitoring system using a java-enabled 3g mobile phone," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2007, pp. 3713–3716.

[6] A. Raman, "Enforcing privacy through security in remote patient monitoring ecosystems," in *2007 6th International Special Topic Conference on Information Technology Applications in Biomedicine*, Nov 2007, pp. 298–301.

[7] A. Kara, "Protecting privacy in remote-patient monitoring," *Computer*, vol. 34, no. 5, pp. 24–27, May 2001.

[8] A. Abbas and S. U. Khan, "A review on the state-of-the-art privacy-preserving approaches in the e-health clouds," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1431–1441, July 2014.

[9] Yuan-Hsiang Lin, I-Chien Jan, P. C. . Ko, Yen-Yu Chen, Jau-Min Wong, and Gwo-Jen Jan, "A wireless pda-based physiological monitoring system for patient transport," *IEEE Transactions on Information Technology in Biomedicine*, vol. 8, no. 4, pp. 439–447, Dec 2004.

[10] E. Span, S. Di Pascoli, and G. Iannaccone, "Low-power wearable ecg monitoring system for multiple-patient remote monitoring," *IEEE Sensors Journal*, vol. 16, no. 13, pp. 5452–5462, July 2016.

[11] S. Park, Y. Jeong, M. S. Kim, and H. S. Kim, "Linear prediction-based dereverberation with very deep convolutional neural networks for reverberant speech recognition," in *2018 International Conference on Electronics, Information, and Communication (ICEIC)*, Jan. 2018, pp. 1–2.

[12] A. Incze, H. Jancs, Z. Szilgyi, A. Farkas, and C. Sulyok, "Bird sound recognition using a convolutional neural network," in *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)*, Sep. 2018, pp. 000295–000300.