

Voice-Activated Emergency Response: Integrating Arduino Nicla Voice with MIT App Inventor for Elderly Care in Smart Homes

Wei-Tzu Chin

New Taipei Municipal Banqiao Senior High School, Southampton University Foundation Program,
Taiwan, China

Abstract. Nowadays, the average lifespan of people is increasing steadily. A multitude of elderly people have the tendency to reside in their own preferred dwellings. However, as a result of this, it might be too late to discover some elders' sudden abnormal situations. Therefore, it is a great concern for them to instantly and effectively deliver their distress signals to the outside world under consciousness, due to the fact that unanticipated accidents may happen all at once. This proposed system integrates the embedded system "Arduino Nicla Voice" in detection with MIT App Inventor in display on the user's interface and deploys Bluetooth Low Energy(BLE) as the communication channel. First, 8 different values of window overlap were tested and the best outcome was used in the upcoming validation method. The overall system performance in terms of accuracy, precision, recall, loss, and F1-score were examined with a 10-fold stratified cross validation. A Receiving Operating Characteristic(ROC) graph and Area Under the ROC Curve(AUC) based on predicted probability were exhibited to see the performance of each fold and the micro-average as well. Then, limitations were discussed so as to improve the robustness of this PERS system for future studies.

Keywords: Personal Emergency Response Systems(PERS), Arduino Nicla Voice, MIT App Inventor, BLE Communication, Machine Learning in Healthcare, Elderly Care Technology, Voice Recognition Technology.

1. Introduction

Based on information from the World Health Organization(WHO), it is predicted that the proportion of people aged over 60 worldwide will increase from 12% to 22% from 2015 to 2050, i.e., 1.1 out of every 5 will be elderly in 2050 [1]. Due to extended life expectancy family planning policy and a falling birth rate, the over-60s are expected to rise to 20% to 38% of the total in China in 2050, one of the most rapidly-ageing rates in the world. On the other hand, in countries like Italy and Japan, the over-60s are predicted to be 43% in 2050 [2].

While the demographic transformation can be seen in every nation in the world, an increasing number of senior citizens are believed to prefer independently dwelling in their own homes, a concept known as "ageing-in-place", which is considered more cost-effective and desirable than institutional care [3].

However, for some vulnerable groups of seniors, there can be sudden occurrences such as falls, heart attacks, and strokes in need of urgent medical aid. According to the World Health Organization (WHO), it is estimated that around 684,000 people died due to falls. It is the second-leading cause of fatal accidental injuries. The majority of fatal falls happen among adults over the age of 60 [4]. Therefore, it is imperative for measures to be taken to respond to urgent situations so as to prevent them from further health hazards. In an interview with WHO in 2016, a 92-year-old man, Mr. Bjorkman, stated the alarm saved his life during the time he previously experienced acute dyspnea. He mentioned that the alarm was a life-saving tool when he required immediate medical attention [5]. In order to provide instant help to the geriatrics who are aware of their own crises, the integrated Personal Emergency Response System (PERS) was proposed.

[13] and [15] afford solutions with advanced technologies for tracking activities of daily living (ADL). Systems are capable of distinguishing abnormal situations from normal living patterns. Nonetheless, privacy has raised concerns and focuses in recent years, due to which older adults have a lower willingness to install a monitoring system (Health & physiological monitoring type B) for

lifestyle patterns compared to PERS [6]. To nip delayed emergency response in the bud, a system was proposed which would triggered passively only when detecting the keyword. In ordinary life, users can diminish their worries about personal data leakage, while a similar goal can still be achieved. As long as the user is aware of their own danger, he or she can instantly utter the keyword, and so the emergency message will be sent to connected devices promptly with minimal delay. In this way, the dignity of older individuals can be maintained as their autonomy to live on their own is enhanced. In the phase of population ageing, the proposed system gives elder groups the opportunity to not be looked after by other people, and they can also deliver their urgent need for help when in distress; hence, further medical measures can be taken.

This paper is structured as follows: Section II is a literature review. In Section III, methodology is examined. Section IV displayed and evaluated the findings and limitations. System integration, keyword selection, and possible future improvements are discussed in Section V. Lastly, Section VI concludes the paper.

2. Literature review/Related Works (1500 words)

Table 1. Comparative Study of Voice Recognition in Emergency Situations

Paper Title	Literature inforamtion(A uthor, Year Published)	Sample or Participant s1	Research Methodology2	Major Findings3	Relevance4
An Acoustic Fall Detector System that Uses Sound Height Information to Reduce the False Alarm Rate[7]	Mihail Popescu, Marjorie Skubic, Marilyn Rantz/ 2008	5 types of fall actions are mimicked by a stunt actor in 10-15 minutes. These moves are each repeated for 3-5 times.	K-nearest neighbour(KN N) algorithm was used in sound recognition in the proposed acoustic fall detection system.	False alarm rate was greatly reduced with the deployment of signal of height of the detected sound.	Fall is an emergency situation especially for elders. In this paper, the height of sound was used to improve the false alarm rate, while Stratified Cross-Validation was used in our study so as to higher accuracy. The overlap between consecutive windows is 50% in this paper, while the overlap between windows

					are approximate ly 87 in our study.
--	--	--	--	--	--

Audio Sound Event Identification for Distress Situations and Context Awareness [8]	J.E. Rougui, D. Istrate and W. Soudene, 2009	9 sound types and speech mixed with 5 noises of every life were included in the module. 11 classes were trained on 70 minutes and signals of daily life noises were trained on 18 minutes.	Gaussian Mixture Model(GMM) was exploited to classify events-homogeneous segments with the efficiency representation of Mel Frequency Cepstral Coefficient (MFCC).	When signal to noise ratio(SNR) is larger or equal to 20db, error recognition rate was about 25%. Accurate recognition was performed with the sound of the electric shaver and for hair dryer because of its spectral composition. While when there is water flushing and bathroom noises present, performances need to be improved.	With the proposed segmentation module, no data pre-processing was required, while data was pre-processed the manually. SNR were all precisely recorded in the proposed research, whereas no SNR was precisely measured in our research.
--	--	--	--	--	---

Development of an automated speech recognition interface for personal emergency response systems[9]	Melinda Hamill ¹ , Vicky Young ² , Jennifer Boger ² and Alex Mihailidis ^{*2} , 2009	4 male and 5 female ranging from 20-30 years participated and asked to speak as loud as AN4. Also, they were asked to speak the normal volume for “yes” and “no” and 12 utterances per subject.	Communication and dialog was enabled to interface with the proposed personal emergency response system(PERS) by using automatic speech recognition (ASR). Hidden Markov Model(HMM) was deployed for word recognition in the prototype using Sphinx 4 speech engine. Java Speech application programming interfaces(APIs) were used as the interface for researchers to interact with the underlying Sphinx 4.	It was found that ASR’s accuracy in identifying words can be influenced by the location of noise interference. Also, as the number of responses of either ‘yes ’or ‘no ’in dialog reduces, the correct recognition rate for the beam-former would increase (from 50% to 90%) as a result.	The research aim can be linked to our study focus, both of which were proposing the opportunities for elders to live independently after retirement, while they can contact the outside world in time when an emergency occurs. However, the prototype of this research is dialog-based, whereas our system does not include a dialog function for users.
---	---	---	---	---	---

Embedded Implementation of Distress Situation Identification Through Sound Analysis [10]	Dan Istrate, Michel Vacher*, Jean-François Serignat*, 2007	Database of 38 minutes was collected from 11 male and 10 female ranging from 20-65 years. Around 2/3 of the audio files were recorded at laboratory, while environmental sound, files from the Internet, and files from a commercial CD all account for around 10% of the total database.	For sound classification, acoustic parameters of LFCC (Linear Frequency Cepstral Coefficients) with 24 filters were used for the Gaussian Mixture Model (GMM) algorithm. While the autonomous system RAPHAEL was deployed for speech recognition.	With the SNR of 10db or higher, the proposed system demonstrated higher and reliable performance, with 0% and 5% (or less) in detection error rate and classification rate respectively. Dissimilar patterns could be found in voice and speech recognition. The error rate of the former ranged from 9% to 37%, while 22% for the latter.	Though both of the studies proposed an embedded system, in this system, the patient has the opportunity to cancel the alarm before a distress signal being sent outward, whereas in our study, there is a limitation for people to deal with a falsely detected signal. SNR was recorded in this system, however, no precise SNR was measured in our research.
---	---	--	--	---	---

Unsupervised Machine Learning for Developing Personalised Behaviour Models Using Activity Data [11]	Laura Fiorini 1,*, Filippo Cavallo 1, Paolo Dario 1, Alexandra Eavis 2 and Praminda Caleb-Solly 3 1, 2017	Data was collected from 17 independent homes in the UK for 100 days. The participants behaved as they normally do in their everyday lives. None of any other information about the participants was included since it was conducted anonymously.	Activities classification was based on multi class Support Vector Machines(SVM). To identify the users daily behaviour patterns, the robust K-Means (KM) and Self-Organizing Maps(MAP) were deployed and compared. To reduce the number of components in the dataset as well as visualise the feature space two-dimensionally, Principal Components Analysis (PCA) and Sammon's Mapping (SM) were applied.	The five groups compared had accuracies ranging from 0.76 to 1.00. When each room was considered separately and independently, the overall accuracy increased. 55 days were found sufficient to construct a learning set. Also, the "blind" approach proposed was conducive for group and distinguish clusters of similar patterns. Features derived from this research helped to differentiate participants with similar behaviours. Support could be offered to those in need specifically for the module to distinguish participants with specific patterns in behaviours.	There is a similarity in the aim to enable elders to live independently at their home and systems to make instant responses when there is an emergency. In the proposed research, after generating a confusion matrix, the F-Measure was calculated to find the best trade-off based on users' daily behaviour patterns. In our research, however, after generating a confusion matrix, this research focused on finding the final result in terms of accuracy, precision, recall, F1 score, and loss to avoid overfitting.
---	---	--	--	---	---

Distant Speech Recognition in a Smart Home: Comparison of Several Multisource ASRs in Realistic Conditions [12]	Benjamin Lecouteux, Michel Vacher, François Portet, 2011	21 persons (including 7 women) of the average age of 38.5 ± 13 years, ranging from 22 to 63 years, performed daily living activities to be recorded in 7 microphones in the smart home. For Phase 1, the total time collected was 38 minutes 46s per channel (862 sentences), and 40 minutes 27s per channel (917 sentences) for phase 2.	Annotated data could be better adjusted with the collaboration of Maximum Likelihood Linear Regression (MLLR). And Hidden Markov Model (HMM) and trigram language model was applied to the A* decoder in the speech recognition toolkit Speeral.	It is found that the accuracy of Automatic Speech Recognition (ASR) enhances when utilising 7 microphones in a multi-smart home environment. Though beam forming presented 16.8% Word Error Rate improvements, the Driven Decoding Algorithm (DDA), particularly the 2-level DDA approach, performed an 8.8% WER and 96.8% F-Measure, superior to other methods.	Relevance could be observed from a similar aim of supporting elders and disabled persons live independently in their own residence. The natural man-machine supports individuals with voice and tactile command interaction in this study, while there were only voice elements included to trigger the assistance.
--	---	---	---	---	--

Influence of time and length size feature selections for human activity sequences recognition [13]	Hongqing Fang a,n, Long Chen a, Raghavendiran Srinivasan b , 2013	Data was from 10 activities performed by 2 volunteers. where Database was collected in CASAS smart apartment testbed (3 bedrooms, 1 bathroom, 1 kitchen, 1 living/dining room inclusive) in Washington State University campus for 55 days. Activities recorded were 600 instances, and 647,485 for motion sensor events.	Viterbi algorithm based on a hidden Markov model was exploited to find and recognise activities sequences in this research. A 3-fold cross validation was applied.	Better results could be obtained with a small activity length size feature value. Relatively, there is a higher tendency that worse results be generated with a larger length size feature.	3-fold cross validation was demonstrated in this research with a larger amount of data, while 10-fold cross validation was applied in our research with a smaller amount of data. Active monitoring sensors keep track of activities of daily living(ADLs) specific groups, whereas our system acts passively, i.e. only when detecting keywords, the system gives response.
--	---	---	--	---	--

Complete Sound and Speech Recognition System for Health Smart Homes: Application to the Recognition of Activities of Daily Living [14]	Michel Vacher¹, Anthony Fleury², François Portet¹, Jean-François Serignat¹, Norbert Noury², 2009	7 male and 3 female of average age 37.2±14 years volunteered to utter 20 distress sentences, 10 regular sentences and 3 phone conversations (45 in total) from 1 to 10 meters from the microphone. Database included 1985 audio files of total duration 35 min 38s made of daily life sounds.	In the echo cancellation system, the method Minimum Mean Square Estimator Short-Time Amplitude Spectrum Estimator (MMSE-STSA) was implemented. GMM in segmentation section helped to classify the audio data into different classes of events. 24 filterbank LFCC features were included in sound classification. A 10-fold stratified cross-validation method was used to evaluate the impact of the attribute selection.	The error rate increased to 30% for distress sentences recognition due to the impact of poor acoustic in the flat. The assessment indicated that the PIR sensors provided the most valuable information.	Our study and this study both proposed a system to detect daily activities to help develop an individual's autonomy in the preferred living space. Cross validation was deployed to compare the performance of different attributes in the filter selection method, while the same method was used to avoid overfitting when looking into the model performance.
---	--	--	---	---	---

SVM-Based Multimodal Classification of Activities of Daily Living in Health Smart Homes: Sensors, Algorithms, and First Experimental Results [15]	Anthony Fleury, Member, IEEE, Michel Vacher, and Norbert Noury, Senior Member, IEEE, 2010	Participants are 6 male and 7 female of average age of 30.4 ± 5.9 years, ranging from 23 to 43 years. 51 min 40s was the mean execution time, while 11min 33s difference between the largest and the smallest, as there was no time limit for participants to perform activities.	“Leave-one-out” cross-validation method processed the validation process in terms of the classification in degree 1 and 2 polynomial and Gaussian kernel .In the case of Gaussian kernel, the optimal value for the parameter of the kernel can be determined by minimising the global error rate. Principal component analysis (PCA) was deployed in feature extraction. SVM algorithm was implemented in data classification due to the limited number of samples.	Frame that were well-classified was 86% of the global percentage, with 75.9% for polynomial kernel and 86.2% for the Gaussian kernel. The probability of misclassification was found to be higher in hygiene and toilets than any other classes.	Cross validation is also used in our research, while it was to avoid overfitting in the overall results. The aim of this research is to find the loss of autonomy of an elders in time by classifying different activities of daily living (ADL) using cross validation. Both of the proposed system and our system provide assistance to elderly and handicapped individuals to dwell solitarily.
--	--	---	---	---	---

Single activity sensor-based ensemble analysis for health monitoring of solitary elderly people [16]	Ohbyung Kwon a,†, Jae Moon Shim b, Geunchan Lim c, 2012	150 elders participated for 16 days in August 2010 for activities being monitored for 5-second cycle time. Data was transmitted every hour, which 1236 activities were in total.	The u-care system was implemented in this study. However, due to only one independent variable in case-based reasoning and Artificial Neural Networks (ANN), with the phase of optimising u-care systems, a higher-level context identification method was applied in analysis and prediction of future state.	Test efficiency could be improved with a limited size of training set, while the time required for generating estimation rules is minimised. The quality of health monitoring would be enhanced after the overall accuracy being boosted and significant outliers being found. based on the proposed methods	Our system and the system proposed in this paper provide a higher number of opportunities for elderly people to live in solitarily sustainably.
--	---	--	--	--	---

PATIENT STATUS MONITORING FOR SMART HOME HEALTHCARE [17]	M. Shamim Hossain, 2016	More than 100 people, among which were mostly male and ranging from 20-45 years, participated to perform under conditions of normal, tensed, and feeling pain.	Local feature (LF) was deployed in feature extraction for speech recognition. Interlaced derivative patterns (IDP) also extract features from the key frames of the patient's face selected. SVM functioned to determine the optimal separator between two classes with maximum margin.	Highest accuracy was found in identifying normal conditions, while the lowest accuracy was found in identifying tensed conditions. Results also depicted accuracy can be greatly increased by combining modality.	In this system, an alert could be delivered to corresponding stakeholders and ambulance or smart transportation service when an abnormal activity was detected, which could be referred to our system so that the distress signal would be delivered to connected devices (while sending alerting messages to stakeholders is included in our future scope.)
---	--------------------------------	---	--	--	---

Health and emergency-care platform for the elderly and disabled people in the Smart City [18]	Aamir Hussain , Rao Wenbi , Aristides Lopes da Silva , Muhammad Nadher , Muhammad Mudhish, 2015	Participant s include remote emergency service providers, doctors and insurance providers, as well as local informatio n collected from patients' sensors connected.	Decision was made by the sensor's node itself. Fuzzy rules were to monitor and determine the nature of the events.	Users would only have to turn on and do measurements on the device, dataset collected from sensor nodes and actuators could be delivered to the repository.	Both the proposed system and our system aims to provide emergency response when sensors detect abnormal health conditions (the proposed system) or keywords (our system). Emergency messages would be sent to remote servers in this study, while the distance is limited between the sensor and the server.
---	---	--	--	---	--

Intelligent decision support system for dementia care through smart home [19]	K.S.Gayathria, *, K.S.Easwarakumar, 2016	Samples were from smart home dataset, most of which included uncertain data and domain knowledge.	Markov Logic Network (MLN), Artificial Neural Network (ANN) and rule based systems were compared in terms of decision support.	The F-measure of MLN in decision making was discovered to be the best in comparison with ANN, for the latter not being capable of generating output when incomplete inputs were given.	Notwithstanding different target users, both our system and the proposed system used common algorithms to support the decision making process. precision, recall, F-measure (i.e. F1 score in our research) were recorded down and optimised in both studies.
---	--	---	--	--	---

Activity Recognition and Dementia Care in Smart Home [20]	K. S. Gayathri, K. S. Easwarakumar, and Susan Elias, 2018	N/A	Markov Logic Network (MLN) structure was used in Event Pattern Activity Modelling Framework (EPAM) to construct a relationship between abnormal factors and recognise irregular occupant behaviour.	Data-driven approaches provide solutions to challenges in activity dynamics and data uncertainty. Knowledge-driven approaches effectively dealt with challenges in activity granularity, contextual knowledge, and activity diversity.	There is a similar aim in diminishing the reliance on caretakers and excessive care for specific groups in society.
---	---	-----	---	--	---

Improving Human Activity Recognition Performance by Data Fusion and Feature Engineering [21]	Jingcheng Chen 1,2, Yining Sun 1,2 and Shaoming Sun 1,2,3,* , 2018	13 male and 4 female participated and completed tasks in three adjacent places at the same laboratory building. Participants randomly performed 9 activities and carried out each activity at least 10 times. Prior to the start of each activity, participants stood still for 30s to undertake research requests.	A genetic algorithm-based feature selection algorithm with fixed activation number (GFSFAN) was used to select appropriate feature subsets for human activity recognition(HAR). Six classifiers, Centre-Nearest Neighbours, K-Nearest Neighbours(KNN), LDA, NB, Random Forests(RF) and Linear Support Vector Machine (SVM), were applied to evaluate the classification performance.	KNN, Random Forest and SVM have the optimal results in classification for high F-measure (FM). Among all, KNN requires a longer time period under the circumstance of a large amount of data. Comparable classification performance can be maintained with the GFSFAN algorithm for a large number of features being discarded. Higher accuracy for HAR could be obtained with RF and SVM classifiers.	It was proved in this research that higher window overlap performs better in training, nevertheless , 85% to 50% (5% reduced each time) overlap in our research, the optimal overall performance was found with 70%. As the value of our results were close, the overlap theory proposed in this research is acceptable. Also in this research, wearable sensors actively recognise abnormal, if any, human activities by adapting six scales of algorithm methods. In our research, however, the sensor would be triggered
--	--	---	--	--	---

					passively and react.
--	--	--	--	--	----------------------

3. Materials and methods (750 words)

3.1 Overall Description

In order to provide timely aid to individuals facing health emergencies while alone, this research proposes a comprehensive system. This system utilises a well-trained model capable of recognising specific keywords. When the keyword is detected, the model will promptly deliver a signal via a BLE network to the connected mobile phone, alerting others to the urgent situation that requires immediate assistance. This system integrates the embedded Arduino Nicla Voice and App Inventor for connection and display. A flow diagram of the system is exhibited in Figure1.

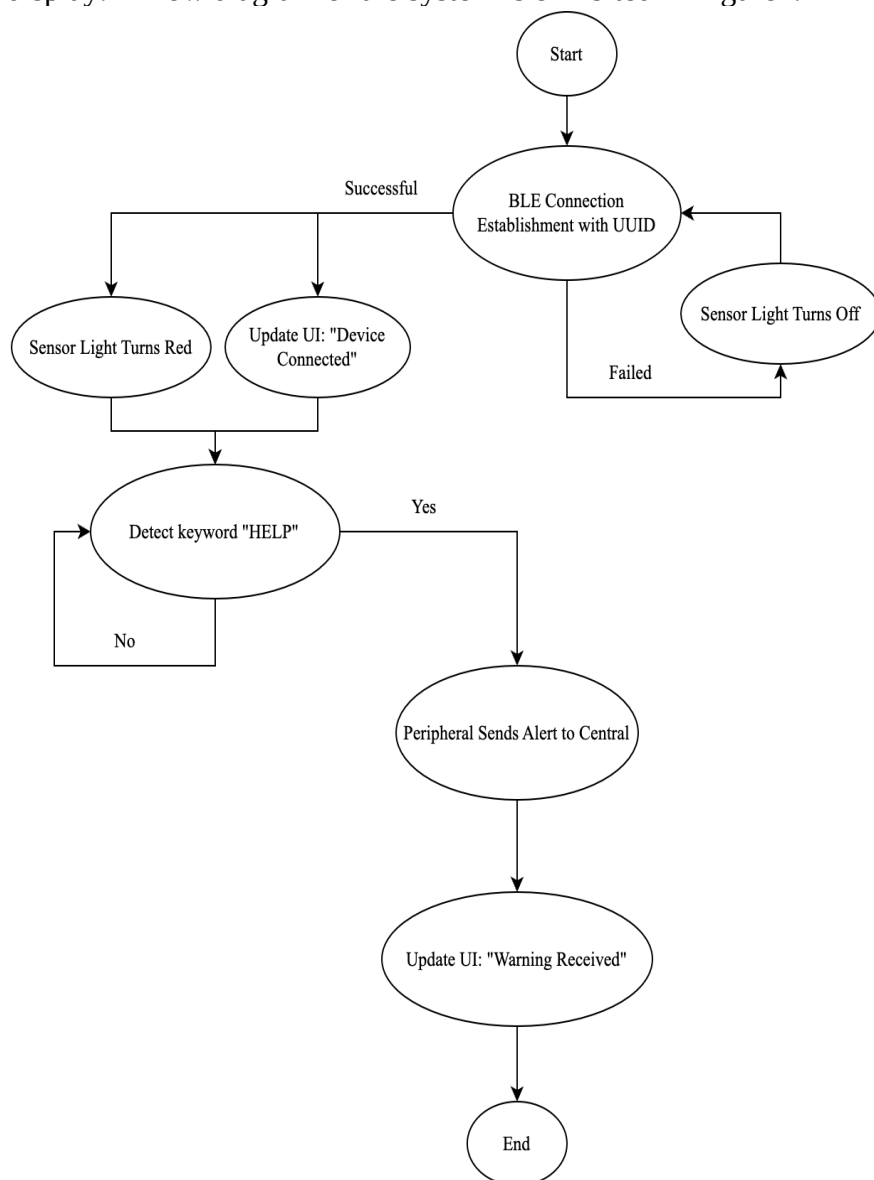


Fig. 1. System Flow Diagram

3.2 Hardware Configuration

The hardware involved in our system is the Arduino Nicla Voice. The primary processor is nRF52832, which includes a 64 MHz Arm® Cortex®-M4F microcontroller, 64 KB SRAM and 512 KB of Flash, The microprocessor Syntiant® NDP120 Neural Decision Processor™ processes data, while the built-in Syntiant® Core 2™ in the microprocessor handles deep neural network inference tasks [22][23]. The ANNA-B112 Bluetooth® module facilitates low-energy communication. Sensors, including a high-quality MEMS microphone, along with a 6-axis IMU and a 3-axis magnetometer, are capable of detecting ambient sounds in the environment comprehensively.

3.3 Data Acquisition

The database consists of 2005 keyword recordings and 20 noise recordings. The recordings were evenly distributed into 2 classes: the keyword “help” and noise. duration of the keyword contains 38 m 45 s. A similar audio duration of 38 m 29 s was collected for noise class. 10 males and 14 females participated in the simulation to utter the keyword “help” for one time in three different situations. The situations were set as “weak health condition”, “imperative and critical health condition”, “when falling down”. The age of participants was between 17 to 29 years. Though there were volunteers participating at the beginning stage of the experiment, due to limited resources, the rest data collection process of the experiment was conducted by the author. Therefore, the diversity of the data was limited in around 3.5% of the total recordings; around 96.5% of the total keyword files were recorded from the author. Noise collection was primarily from discourses of participants in the background. Participants’ sounds were recorded in a quiet room in New Taipei City, Taiwan, for no other violation in ethical and privacy concerns. The sound was recorded with the use of the microphone on Arduino Nicla Voice. Distances between signal and the microphone ranged from 0.1 to 3 metre. Signal-to-noise ratio (SNR), however, was not applicable in this research due to limitations of apparatus and resources.

3.4 Data pre-processing

Prior to the machine learning stage, audio data was preprocessed to improve the model’s accuracy. In the preliminary stage, audio files were randomly split into 80% in the training set and 20% in the testing set.

The default setting of the platform for feature extraction used Log Mel filterbank energy features, an approach for audio processing in machine learning. Parameters such as frame length was set to 0.032 s, and frame stride was set to 0.024 s. Predefined parameters include a filter number of 40, Fast Fourier Transform (FFT) length fixed at 512 points, and both Low Frequency and High Frequency set to 0 Hz, indicating that appropriate values were automatically selected by the system. The preemphasis coefficient was set at 0.96875 for filtering. Through splitting audio signals into individual windows, signals are transformed from time-domain to frequency-domain; therefore, it is easier to process sound recognition, classification, and spectral analysis. For time series data in the Impulse Design section of Arduino Machine Learning Tools, the immutable built-in window size is 968ms. The sampling frequency was set to be 16000 Hz. Since window size cannot be modified, different window increases and overlaps were therefore tested so as to find the optimal model performance. Table 1 shows the results of the eight tested values. The best value for overall performance was achieved when the window overlap fixed at 70%, i.e. the window increase set to 290 ms. (In fact, the precise value was 290.4 ms, but the system automatically ignored the 0.4 ms.)

Table 2. System Performance for Different Window Settings

window increase / window overlap	Accuracy	Precision	Recall	F1 Score	Loss
145ms / 85%	0.985	0.968	0.948	0.958	0.0148
194ms / 80%	0.986	0.986	0.955	0.963	0.014
242ms / 75%	0.979	0.94	0.957	0.948	0.0213
290ms / 70%	0.988	0.984	0.964	0.974	0.0125
339ms / 65%	0.976	0.954	0.952	0.953	0.0242
387ms / 60%	0.977	0.969	0.947	0.958	0.0229
436ms / 55%	0.976	0.967	0.947	0.957	0.0244

484ms / 50%	0.97	0.959	0.944	0.951	0.0299
-------------	------	-------	-------	-------	--------

3.5 ML model on Nicla Voice

In the proposed system, the Syntiant® NDP120 Neural Decision Processor™ was applied to the embedded chip to facilitate machine learning operations. The Syntiant® NDP120 Neural Decision Processor™ is designed for audio and sensor processing purposes in terms of power-constrained systems. Its neural processing capability allows multiple operations while maintaining long-term low power consumption simultaneously. It is built with Syntiant Core 2™ programmable deep learning architecture, which allows the chip to natively run multiple Deep Neural Networks (DNN) on other structures such as Convolutional Neural Network(CNN), Recurrent Neural Network(RNN), and Fully Connected Network (FCN)[24].

Additionally, compared to Syntiant Core 1™ embedded in Syntiant's NDP100 and NDP101 devices, the NDP120 processes complex model learning with higher efficiency. Apart from DNN, a programmable HiFi 3DSP is also involved, which is applicable to conventional audio processing. The processing capability and memory configuration of nRF52832 also provide high efficiency in machine learning operations and data processing.

3.6 Data Transmission and Communication: BLE Network

ANNA-B112 module Bluetooth Low Energy(BLE) network is integrated to transmit signals from Arduino Nicla Voice to the designated mobile device, which in this case refers to the MIT App Inventor interface. The embedded BLE module supports a variety of Bluetooth protocols, including Bluetooth Low Energy Serial Port Service and GATT client and server protocols [25]. Data organisation and communication methods between BLE devices are specified in the GATT protocol. This module is supported by u-connectXpress software, which is approachable for beginners who do not require in-depth knowledge when developing BLE applications. In [25], based on the RF performance, it can be known that a high receiver sensitivity ensures data can be received in weak signal situations, while a higher output power extends the boundary for effective communication. [25] also states that throughput decreases when multiple devices are connected, meaning a longer time is required to transmit the data.

There are several advantages to using BLE in communication. As the name implies, it consumes eminently little electricity in data transmission, extending its longevity and providing the convenience of no need for additional power supply. Moreover, the construction of a network using the BLE protocol is relatively faster, reducing the complexity and waiting time. For each BLE device, there is a unique UUID (Universally Unique Identifier) assigned to each service and characteristic, respectively, ensuring specificity in communication. Each character in the UUID represents one hexadecimal digit. In our custom profile, the service UUID is d75fdea4-d846-4613-9fbc-b781fa353d75, while characteristic UUID is af15640-ec27-407e-97da-0f2e829e189, both of which were written in codes on Arduino Integrated Development Environment (IDE) as well as MIT App Inventor.

Figure 2 explains the principle of the BLE network involved in our system. At the connection stage, there are two significant roles: the central role and the peripheral role. The central role initiates a connection and transitions from initiating state (the state that requests a connection by responding to advertising packets from a specific device) to connection state(the state of a connection with another device). On the other hand, a device that accepts a connection request and shifts from the advertising state (the state that sends out advertising packets and addresses incoming packets in response to advertising packets from other devices) to the connection state is regarded as the peripheral role[26]. Though ANNA-B12 module supports simultaneous roles as a peripheral and a central device, here it is the peripheral role that accepts the connection request. Hence, the embedded module on Arduino Nicla Voice and the designated mobile device are defined as the peripheral role and the central role,

respectively, in this system. The effective communication range of the BLE network typically ranges from a few metres to tens of metres. The strength of connection is relatively high and stable within a short distance, and vice versa. In addition to distance, the strength of the signal can also be affected by barriers (doors, walls et cetera) between the connection and other wireless networks. When the aforementioned factors exist in the connection stage, the BLE signal consequently diminishes. During the connection stage, the display of Received Signal Strength Indication (RSSI) on the peripheral device helps users to understand the strength of the signal and actions can therefore be taken when aiming to amplify the strength of the connection. In our system, the embedded chip was adopted, which is pre-equipped with support for BLE communication. Our study aims to achieve high accuracy in keyword detection, and the default BLE network provides support in two significant parts:

Voice Recognition:

Audio data could be transmitted from an embedded microphone using BLE communication, enabling the system to detect and identify specific keywords.

Wireless Data Transmission:

Collected data was able to be transferred to the MIT App Inventor interface on the mobile device via BLE communication, which facilitates the opportunities for remote monitoring as well as data sharing.

The pre-configured BLE communication function in the embedded module, Arduino Nicla Voice, allows the author to concentrate on the development of other segments in our system and meet the eventual objectives.

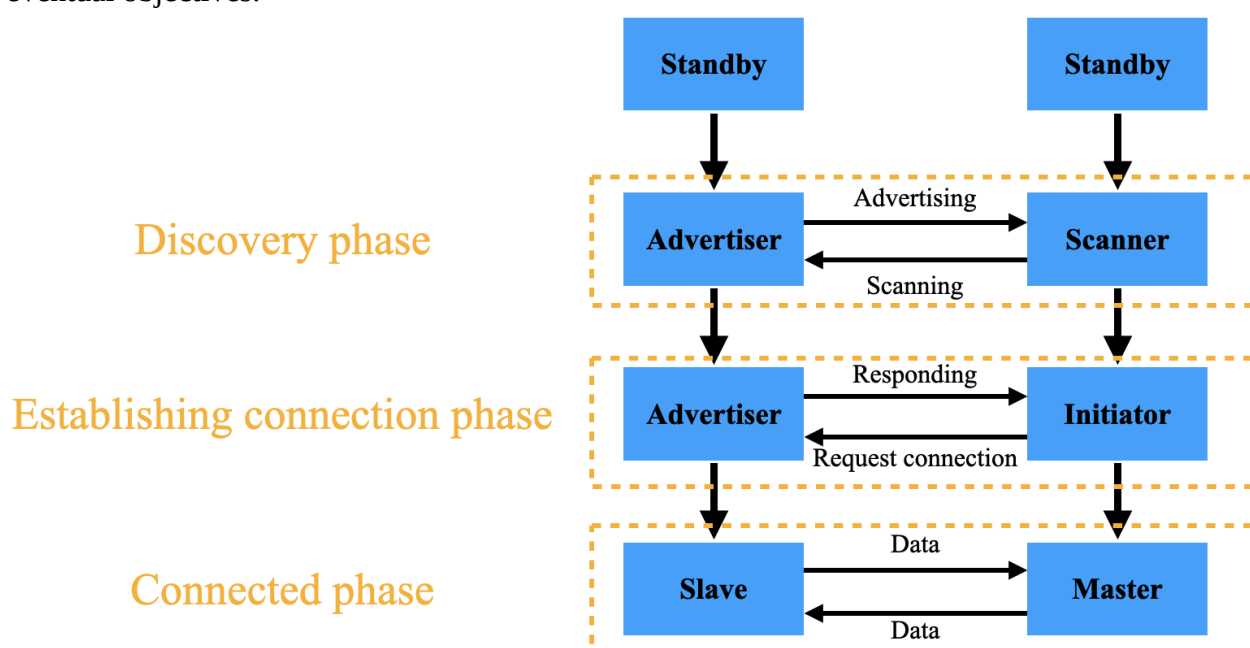


Fig. 2. Principle of BLE network

4. Findings & Evaluation

The optimal value of window overlap(70%) was found after tailoring the window increase. Then, 10-fold stratified cross-validation was employed so overfitting issues can be mitigated under the limited diversity of the database. This method provides a more comprehensive evaluation of performance since the model had to assess the data in terms of different training sets and testing sets. For the validation process, the data was split into 90% for the training set and 10% for the testing set. For each fold, a confusion matrix exhibited the performance in different different categories, with True Positive (TP) on the top-left corner, False Positive (FP) on the bottom-left corner, False

Negative(FN) on the top-right corner, and True Negative(TN) on the bottom-right corner (Figure 3).

$$\begin{bmatrix} TP & FN \\ FP & TN \end{bmatrix}$$

Fig. 3. Confusion Matrix for Model Evaluation

5. Results

Based on the confusion matrix, the model performance was accessed by looking into several metrics, including accuracy, precision, recall, loss, and F1 score, which were examined in model training in Table 2. In our model evaluation, higher accuracy and a lower loss value represent better overall performance. Additionally, the commonly used tools to examine binary classification problems are the Receiver Operating Characteristic(ROC) Curve and Area Under the Curve (AUC), whose results are shown in the graph below.

Table 3. 10-Fold Cross-Validation Performance Metrics

	Accuracy	precision	Recall	Loss	F1 score
Fold 1	0.987	0.979	0.975	0.0130	0.977
Fold 2	0.968	0.941	0.935	0.0319	0.938
Fold 3	0.972	0.947	0.949	0.0276	0.948
Fold 4	0.976	0.972	0.942	0.0242	0.957
Fold 5	0.978	0.965	0.957	0.0223	0.961
Fold 6	0.966	0.931	0.952	0.0340	0.941
Fold 7	0.975	0.976	0.980	0.0251	0.978
Fold 8	0.976	0.974	0.940	0.0241	0.957
Fold 9	0.976	0.959	0.957	0.0242	0.958
Fold 10	0.977	0.968	0.950	0.0229	0.959
Overall performance	0.9751	0.9612	0.9537	0.02735	0.9574

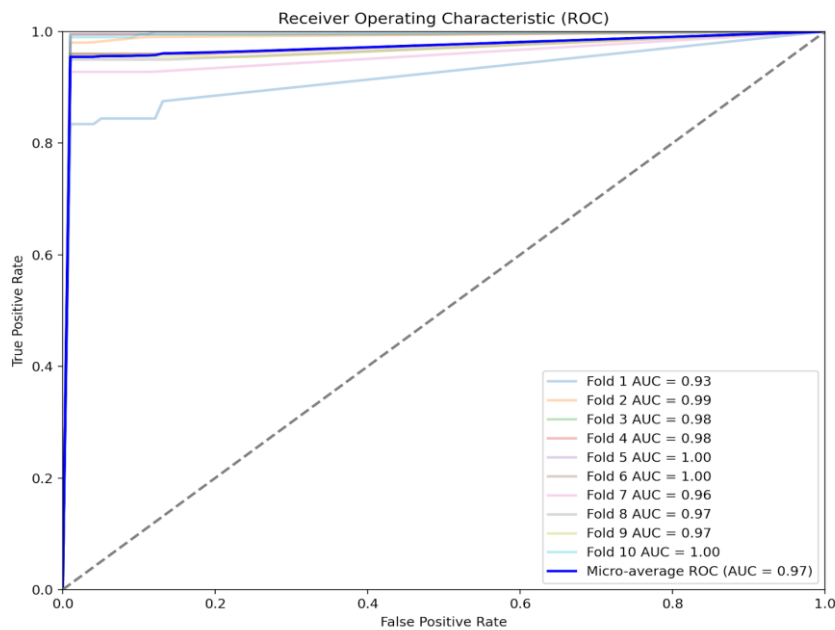


Fig. 4 ROC and AUC graph

In the system evaluation, the fairness and robustness of the model were established by using a 10-fold stratified cross-validation. Table 2 shows the accuracy, precision, recall, loss, and F1 score of each fold. ‘Overall performance’ row exhibited the mean value of the 10 folds and provided evidence for the consistency of the dataset. The value is around 0.975 for mean accuracy, 0.961 for precision, 0.954 for recall, 0.0274 for recall, and 0.9587 for F1-score.

Moreover, the performance of each fold was examined by looking at the predicted probability. Therefore, the Receiving Operating Characteristic (ROC) graph and Area Under the ROC curve (AUC) showed the capability of each fold in distinguishing the correct classes at different thresholds. The timestamp was set at 290 ms for each file. Also, micro-average ROC (AUC is 0.97) is included in this graph to show the comprehensive overall performance of every fold. Performance in correct prediction has outstanding results in fold 5 and fold 10 with an AUC equal to 1. The comparatively inferior result was the predicted probability in fold 1 with 0.93 of AUC.

5.1 Code and Block Display

The lines and the blocks below display the code in Arduino Integrated Development Environment (IDE) in the system and behind the interface of MIT App Inventor, respectively.

```
#include "Nicla_System.h"
#include "NDP.h"
#include <ArduinoBLE.h>

BLEService warningService("d75fdea4-d84b-4b13-9fbc-b781fa353d75");
BLEUnsignedCharCharacteristic isWarning("e903a02f-c7a8-468a-b533-7c0a263943ee", BLERead | BLENotify);

void onWarningLabelMatch(char* label) {
    // if SmartHome_NiclaVoice receives message, send signal via
    // bluetooth to central
    // notify that we recognized a keyword
    Serial.println("Help");
    BLEDevice central = BLE.central();
```



```
// If a central device is connected
if (central) {
    isWarning.writeValue(0);
}
}

void setup() {
    Serial.begin(9600);
    nicla::begin();
    nicla::disableLDO();
    nicla::leds.begin();

    NDP.onMatch(onWarningLabelMatch);
    Serial.println("Loading synpackages");
    NDP.begin("mcu_fw_120_v91.synpkg");
    NDP.load("dsp_firmware_v91.synpkg");
    NDP.load("alexa_334_NDP120_B0_v11_v91.synpkg");
    Serial.println("packages loaded");
    NDP.getInfo();
    Serial.println("Configure mic");
    NDP.turnOnMicrophone();
    NDP.interrupts();

    // Wait for the serial connection to be established
    while (!Serial)
        ;

    // Initialize the BLE module
    while (!BLE.begin()) {
        Serial.println("starting BLE failed!");
    }

    BLE.setLocalName("SmartHome_NiclaVoice");
    BLE.setAdvertisedService(warningService);
    warningService.addCharacteristic(isWarning);
    BLE.addService(warningService);
    // Start advertising the BLE service
    BLE.advertise();
    Serial.println("- Bluetooth device active, waiting for
        connections...");
    String address = BLE.address();

    Serial.print("Local address is: ");
    Serial.println(address);
}

void loop() {
    // Here is the main code to run repeatedly:
```

```
// Check for incoming BLE connections
BLEDevice central = BLE.central();

// If a central device is connected
if (central) {
    Serial.print("- Connected to central: ");
    Serial.println(central.address());

    // Set the LED color to red when connected
    nicla::leds.setColor(red);

    // while the central is still connected to peripheral:
    while (central.connected()) {
        //Serial.println(BluetoothStatus.written());
        delay(1000);
    }
    Serial.print("- Disconnected from central: ");
    nicla::leds.setColor(off);
    Serial.println(central.address());
} else {
    delay(1000);
}
}
```

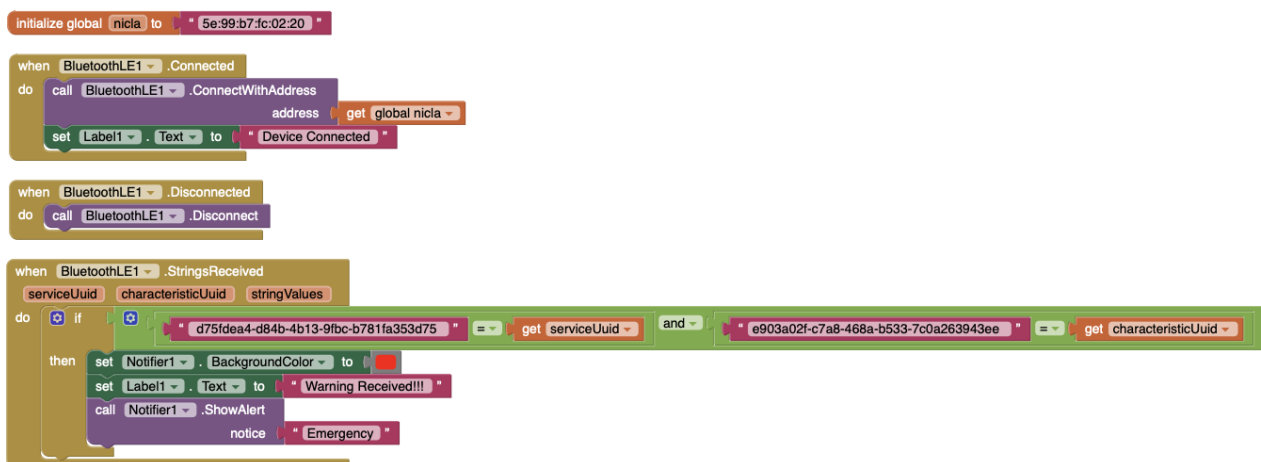


Figure 4. BLE Communication Flow in Arduino IDE and App Inventor Integration

In this system, Arduino Nicla Voice processes sound detection and transmits signals via BLE communication. When the keyword “help” is detected, the system delivers an emergency signal via BLE service. This function is shown in the Arduino IDE in terms of the configuration of `BLEService` and `BLEUnsignedCharCharacteristic` in lines 4 and 5. The attribute ‘`BLERead`’ in line 6 allows the connected central role to read the current value of this characteristic. The peripheral role with attribute ‘`BLENotify`’ in line 6 notifies the central role only when the value changes, and so the central device does not have to read the value actively, which is an efficient way to save the battery. On the other hand, App Inventor was deployed as the interface for users. When devices are successfully connected, the connected message will be updated; notifications will also pop up when alarms are received.

The multi-layer emergency notification system was realised through the efficient signal processing and communication code on the Arduino IDE, along with the approachable interface for users in App Inventor. Works are clearly divided; Arduino code is in charge of data processing and detecting, and when the keyword is recognised, it delivers the signal via BLE. Then, App Inventor is applied to the display after receiving such signals. Overall, this integration not only enhances practicality but also the users' experiences.

6. Limitation

6.1 Insufficient Diversity in Database

Exceptionally excellent metrics were calculated from the machine learning of the model. The possibility of outstanding capability of the embedded chip is not excluded; however, the oddly impressive results might be due to the restricted diversity of samples. Though distress sounds in various scenes were mimicked, a high proportion of 96.5% of total samples were recorded by the author, which could potentially result in the constrained learning within the author's acoustic tone, volume, speed, accent, pronunciation clarity, and emotional expressions, etc. Results show the possibility that the model is used to the features of the author's voice, therefore causing a potential inability to adapt to dynamically changing voices. Moreover, though few of the recordings were from volunteers, the age of volunteers were relatively young compared to the groups the system was designed for. There can be existing voice differences in terms of tones and ways of uttering the keyword between different age groups.

6.2 Lack of Practical Testing

Due to the constraints of the author residing in an ensuite room in student accommodation, while this system is inappropriate to be tested in other environments except for places resembling actual living scenes, it was not feasible to conduct real-world testing of the system. Therefore, in this study, the evaluation of the model was confined to the machine learning stage. Further study could focus on deploying and validating the system's performance in the real world.

6.3 Recording Keywords Without Real-World Context

Keywords were all recorded in a quiet room without ambient noise to avoid ethical and privacy concerns. Nevertheless, the lack of real-world context included can lead to incomprehensible model training. As a result, the model may not perform perfectly when it is applied to the actual scenes while being surrounded by daily life noises. A future study can be conducted in the context of daily living sounds.

6.4 Prohibitive Cost

The cost of the embedded chip Arduino Nicla Voice in this system is relatively higher than that of other commercial sensors due to its top-notch quality in data processing and ambient sound detection, as well as its ability to deliver signals with minimal delay. Components, including the nRF52832 SoC, Syntiant® NDP120 Neural Decision Processor™, and sensors increase the overall price. Hence, the high cost can limit the accessibility for elderly people. In the future, a substitute microcontroller and processor can be used to assemble and reproduce similar systems and maintain marginally inferior yet still high quality.

6.5 Challenges in System Integration

MIT App Inventor limits the compatibility on the versions of devices it supports. The interface may not be able to run successfully on iOS and some Android devices, and so the emulator was used in the experiment. However, with the aim of placing the system in the real-world, other powerful platforms will be used to explore compatibility.

Distant wireless communication was not included in this module. Therefore, the number of devices that can be informed is limited due to the peripheral (Arduino Nicla Voice) only connecting to one central device, so messages will only be sent to the sole device. Future improvements are discussed in the discussion section.

7. Discussion

7.1 System Integration challenge and achievement

7.2 Device Compatibility and Functionality Limits

MIT App Inventor does not support iOS or Web development compatibility. Also, devices with Android versions lower than 7 and some Android devices of version 7 or higher but not listed in the compatible devices list cannot be connected and show the messages on the screen for display. Therefore, the testing strategy was to find another compatible device to connect to the display. Also, it is feasible to test out the system's performance on an emulator.

7.3 Challenges with Remote Data Transmission

MIT App Inventor was deployed as a display interface because includes user-friendly BLE components for quick implementation. Due to the built-in BLE system in the embedded Arduino Nicla Voice, data transmission is limited within a short distance (around 1 to 10m). When the distance between the sensor and connected device increases, the signal becomes weaker and unstable. Therefore, remote data transmission between two distant places is not feasible in the current system with BLE communication.

7.4 Scalability and Modularity

For future development, the system can be scalable by including more remote users to whom emergency messages are sent. On this scale, the central device will have to incorporate distant wireless communication such as LoRa (Long Range). Therefore, the system can be enhanced and be comprehensive when the distress signal can be received by multiple users, such as families, neighbourhoods, and the local medical centre. In this way, the modularity of the system is structured so that modules are clearly divided into two parts: BLE communication between the peripheral Arduino Nicla Voice and the central device, and distant wireless communication between the central device and multiple users' devices. Hence, due to the modular design, it is easy to maintain and update the system if needed.

7.5 Optimising Power Consumption

The attribute 'BLENotify' allows the central device to not have to actively read the value on the peripheral role regularly, because the peripheral will notify or advertise when there is a change in value. The central role will only receive the data when it is necessary through this notifying method, so power consumption can be optimised.

7.6 Security and Privacy Considerations

For the connection in our system, the same UUID was written in the code for the hardware and blocks for the App Inventor interface. Only the devices that know the specific UUID can connect and interact with them, which provides a certain level of data protection because the UUID cannot be randomly guessed in general terms. Thus, unauthorised access can be prevented.

7.7 Keyword Selection

The ultimate target of this research is that Arduino Nicla Voice can successfully differentiate specific keywords from noises. In this research, noises can be easily obtained, while the keyword chosen is "help". In this segment, the pros and cons are examined in keyword selection.

Advantages

“Help” was chosen as the specific keyword because it is intuitive and a single syllable, which requires less effort and time to speak out when vulnerable individuals encounter health emergencies. Its simplicity allows users to easily memorise and utter. The applicability for users includes, but not limited to, older adults, for it is a widely used word among different age groups and the multicultural world. Also, in system design and development, a single keyword like “help” comparatively reduces the complexity because no additional computation is required to distinguish the context and process complicated sentences.

Disadvantages

The potential disadvantage can be for an individual who is incapable of uttering the keyword (e.g. choked, with post-stroke conditions/symptoms etc) Under these circumstances, it can be considered a challenge for specific groups of people to trigger the sensor so as to deliver distress signals. In addition, false positives might result in false alarms when the word “help” appears in daily discourse. As the research focusing on the simplicity of using the system, the sole keyword “help” only indicates an emergency situation. However, the type of emergency cannot be known at the beginning stage, which does not make it convenient for hospitals to prepare corresponding measures and medical equipment as there are diverse possibilities in emergencies.

7.8 System Deployment in Real-World

The proposed system is mainly designed for scenes of solitary elderly people in their own dwellings. They can speak out the keyword to trigger the system in order to get help when encountering emergency conditions such as falling, a stroke, or a heart attack. This system can also be applied to the homes of individuals who may experience the abrupt onset of diseases. In the event that the individual is in distress yet being solitary or not being noticed by other people instantly, they can utter the keyword to deliver a distress signal to make other people aware of their emergency situations.

Due to the self-powered design of the embedded system, in which BLE communication was exploited, no external power supply is needed to compute a vast amount of data. This provides convenience for more places to install the system.

With the aim of deploying the system in real life, a device was needed to connect with Arduino Nicla Voice at the preliminary stage. This designated device involves both BLE and WiFi communication functions, allowing it to send emergency messages to remote devices (e.g. families, neighbourhood, and community healthcare systems) via WiFi connection after receiving BLE signals. In this way, the burden of domiciliary care can be effectively reduced while the autonomy of elders is enhanced.

7.9 Future Scope

In the future, the diversity of the dataset should be enhanced in terms of including various background sounds and increasing participants among ages among the system's potential users. In this way, the model can adapt to varied noises as well as different keyword utterances. In addition, SNR can be recorded so that advanced signal quality and accuracy can be obtained. Also, under the circumstances of inferior model performances, problems can be discovered through the analysis of SNR.

In [27], participants suggest such a system could be used outside their home. Therefore, it is believed to be a potential possibility to integrate this system into wearable devices, thereby enhances their safety when being outside alone and their overall autonomy

A dialog function is created to allow system users to cancel the erroneous alerts in [9]. Inspired by this idea, further research would be conducted by adding another keyword for cancelling the erroneous alerts, whereby people can respond to the wrong signal detected by uttering the cancelling keyword to turn off the alarm. Consequently, medical resources would not be wasted on false alarms.

Though the convenience and ease of learning offered by MIT App Inventor were exploited in our system for certain advantages, limitations in compatible devices and functions still exist and can constrain its prevalence. In the future, other development frameworks such as Flutter, React Native, and Progressive Web Apps(PWAs) are considered to achieve true cross-platform compatibility. Additionally, the scalability of the PERS system can be ameliorated by involving remote wireless communication methods such as LoRa; thereby, multiple peripherals from different users can be included to boost the availability and flexibility of this PERS system.

8. Conclusion

This research proposed a PERS system leveraging Arduino Nicla Voice and MIT App Inventor. It operates simply with the utterance of the single keyword - “help”, primarily targeting solitary elderly groups in society. Our findings highlight the outstanding performance of the embedded system in sound detection processing and the effectiveness of alerts delivered via the local BLE communication. The high accuracy and F1 score in 10-fold stratified cross validation showcase the robustness of the machine learning module in keyword identification.

Nonetheless, the exceptional results might be due to the lack of diversity in the database, in which the model is accustomed to the author’s voice characteristics. Real-world testing and including real-world context are limited because of the environmental challenges and privacy concerns respectively, while system integration with compatible devices and multiple users is still yet to be ameliorated. Additionally, the prohibitive cost can potentially restrict adoption and prevalence.

Looking forward, the scalability of this PERS system is promising. With the opportunity to integrate communication networks for broader distances such as by deploying LoRa technology, the alerts can be delivered to a greater variety of users, including families, neighbourhoods, and local health services. Furthermore, the involvement of diverse dataset, real-world testing, cost-effective design, and dialog incorporation for cancelling the alarm can be evolved in future developments.

In summary, this study addresses the potential problems elderly individuals can encounter when being alone in their accommodation. It offers an innovative approach to developing the autonomy of older adults. Thus, the well-being of vulnerable populations can gradually be achieved.

Acknowledgements

The author would like to appreciate great assistance from Ciao An Lee, who provided professional guidance throughout the research. Implementation had been conducted smoothly with the support of Mr. Lee in the machine learning and operation session. His aid facilitated the execution of the study and the author's insights in this field.

References

- [1] World Health Organization: WHO. Ageing and Health. 1 Oct. 2022, www.who.int/news-room/fact-sheets/detail/ageing-and-health.
- [2] “Indicators.” WHO Data, platform.who.int/data/maternal-newborn-child-adolescent-ageing/indicator-explorer-new/mca/percentage-of-total-population-aged-60-years-or-over.
- [3] Wiles, Janine, et al. “Older People and Their Social Spaces: A Study of Well-being and Attachment to Place in Aotearoa New Zealand.” *Social Science & Medicine*, vol. 68, no. 4, Feb. 2009, pp. 664–71. <https://doi.org/10.1016/j.socscimed.2008.11.030>.
- [4] ---. Falls. 26 Apr. 2021, www.who.int/news-room/fact-sheets/detail/falls#:~:text=Each%20year%20an%20estimated%20684,medical%20attention%20occur%20each%20year.
- [5] ---. From Walking Canes to Wrist Alarms: WHO’s Global Survey on Assistive Technologies. 11 Feb. 2016, www.who.int/news-room/feature-stories/detail/from-walking-can-es-to-wrist-alarms-who-s-global-survey-on-assistive-technologies.

- [6] Mihailidis, Alex, et al. "The Acceptability of Home Monitoring Technology Among Community-Dwelling Older Adults and Baby Boomers." *Assistive Technology*, vol. 20, no. 1, Mar. 2008, pp. 1–12. <https://doi.org/10.1080/10400435.2008.10131927>.
- [7] Popescu, Mihail, et al. "An acoustic fall detector system that uses sound height information to reduce the false alarm rate." 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008, <https://doi.org/10.1109/iembs.2008.4650244>.
- [8] Rougui, J.E., et al. "Audio sound event identification for distress situations and context awareness." 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009, <https://doi.org/10.1109/iembs.2009.5334581>.
- [9] Hamill, Melinda, et al. "Development of an Automated Speech Recognition Interface for Personal Emergency Response Systems." *Journal of Neuroengineering and Rehabilitation*, vol. 6, no. 1, July 2009, <https://doi.org/10.1186/1743-0003-6-26>.
- [10] Istrate, Dan, et al. "Embedded Implementation of Distress Situation Identification Through Sound Analysis." HAL (Le Centre Pour La Communication Scientifique Directe), July 2007, hal.science/hal-00957417.
- [11] Fiorini, Laura, et al. "Unsupervised Machine Learning for Developing Personalised Behaviour Models Using Activity Data." *Sensors*, vol. 17, no. 5, May 2017, p. 1034. <https://doi.org/10.3390/s17051034>.
- [12] Lecouteux, Benjamin, et al. "Distant speech recognition in a smart home: Comparison of several multisource ASRS in realistic conditions." *Interspeech* 2011, 2011, <https://doi.org/10.21437/interspeech.2011-604>.
- [13] Fang, Hongqing, et al. "Influence of Time and Length Size Feature Selections for Human Activity Sequences Recognition." *ISA Transactions*, vol. 53, no. 1, Jan. 2014, pp. 134–40. <https://doi.org/10.1016/j.isatra.2013.09.001>.
- [14] Vacher, Michel, et al. "Complete Sound and Speech Recognition System for Health Smart Homes: Application to the Recognition of Activities of Daily Living." *InTech eBooks*, 2010, <https://doi.org/10.5772/7596>.
- [15] Fleury, Anthony, et al. "SVM-Based Multimodal Classification of Activities of Daily Living in Health Smart Homes: Sensors, Algorithms, and First Experimental Results." *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, Mar. 2010, pp. 274–83. <https://doi.org/10.1109/titb.2009.2037317>.
- [16] Kwon, Ohbyung, et al. "Single Activity Sensor-based Ensemble Analysis for Health Monitoring of Solitary Elderly People." *Expert Systems With Applications*, vol. 39, no. 5, Apr. 2012, pp. 5774–83. <https://doi.org/10.1016/j.eswa.2011.11.090>.
- [17] Hossain, M. Shamim. "Patient status monitoring for smart home healthcare." 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2016, <https://doi.org/10.1109/icmew.2016.7574719>.
- [18] Hussain, Akhtar, et al. "Health and Emergency-care Platform for the Elderly and Disabled People in the Smart City." *Journal of Systems and Software*, vol. 110, Dec. 2015, pp. 253–63. <https://doi.org/10.1016/j.jss.2015.08.041>.
- [19] Gayathri, K., and K. S. Easwarakumar. "Intelligent Decision Support System for Dementia Care Through Smart Home." *Procedia Computer Science*, vol. 93, Jan. 2016, pp. 947–55. <https://doi.org/10.1016/j.procs.2016.07.281>.
- [20] Gayathri, K. S., et al. "Activity recognition and dementia care in Smart Home." *Advances in Theory and Practice of Emerging Markets*, 2018, pp. 33–47, https://doi.org/10.1007/978-3-319-78378-9_2.
- [21] Chen, Jingcheng, et al. "Improving Human Activity Recognition Performance by Data Fusion and Feature Engineering." *Sensors*, vol. 21, no. 3, Jan. 2021, p. 692. <https://doi.org/10.3390/s21030692>.
- [22] Arduino. Arduino® Nicla Voice. Arduino, n.d. Web. Accessed 31 Oct. 2023. <https://docs.arduino.cc/resources/datasheets/ABX00061-datasheet.pdf>.
- [23] Nordic Semiconductor. nRF52832 Product Specification v1.4. 10 Oct. 2017, https://infocenter.nordicsemi.com/pdf/nRF52832_PS_v1.4.pdf.

- [24] Syntiant. Syntiant Product Brief NDP120. n.d., https://static1.squarespace.com/static/6488b0b8150a045d2d112999/t/650b29219d8c3f0e27f4651e/1695230242406/Syntiant-Product_Brief_NDP120.pdf.
- [25] u-blox. ANNA-B112 DataSheet UBX-18011707. 14 Mar. 2023, https://content.u-blox.com/sites/default/files/ANNA-B112_DataSheet_UBX-18011707.pdf.
- [26] Bluetooth SIG. The Bluetooth LE Primer V1.1.0. 17 Jan. 2023, https://www.bluetooth.com/wp-content/uploads/2022/05/The-Bluetooth-LE-Primer-V1.1.0.pdf?__hstc=44473531.3ccd99af25786010176b1818f9f3a3d2.1702609437500.1702609437500.1702609437500.1&__hssc=44473531.5.1702609437500&__hsfp=1214847747&hsCtaTracking=8e3cb9ce-2e7b-471a-b5cc-2343a4915b6a%7C090f705a-f0df-4f4b-8a54-c8f97c73eb69.
- [27] Singh, Deepika, et al. "Ambient Assisted Living Technologies From the Perspectives of Older People and Professionals." *Lecture Notes in Computer Science*, 2017, pp. 255–66. https://doi.org/10.1007/978-3-319-66808-6_17.