University of Sindh Journal of Information and Communication Technology (USJICT)
Volume 6, Issue 2, July 2022
ISSN-E: 2523-1235, ISSN-P: 2521-5582
Website: http://sujo.usindh.edu.pk/index.php/USJICT/

Controlling Android Based Smart Phone in Sindhi

Rama Shahani, Asia Khatoon Soomro, Asadullah Burdi, Fida Hussain Khoso, Dil Nawaz Hakro, Maryam Hameed,
Dawood University of Engineering and Technology, Karachi
Jamescook University, Australia

Abstract: Due to its various domestic and industrial applications, Artificial Intelligence (AI) has given birth to many of the attractive speech recognition applications since previous decades. Various speech recognition systems are available, but the availability is rare other than the case except English language Although, speech recognition systems in many languages other than English are a rare occurrence for non-native English speakers who rely on speech to control a variety of devices or applications. The aim of this research is to create an android application, where "Sindhi Language" could be recognized and processed, and necessary actions are performed according to given commands through Android phones. The Android application has been designed to understand selected Sindhi words and sentences, spoken by the user. This system recognizes voices of people in an environment with moderate level of noise as well as without noise. The Android-based Sindhi speech recognition framework is implemented by using React Native technology. The technology aids in the development of applications by allowing them to be reloaded instantly and without the need to recompile them. Native Sindhi speakers performed better than non-native Sindhi speakers in terms of accuracy. Users who speak Sindhi will benefit from this Android application, which makes it easier to use smartphones using local language rather than having to speak English.

Keywords: Sindhi Language, Android, React Native, Speech Recognition

I. INTRODUCTION *(HEADING 1)*

A.I refers to a wide range of intelligent behaviors such as learning, reasoning, memory, feeling, understanding, judgement, and other intelligent activities that can be replicated artificially using real machines [1]. Automatic Speech Recognition (ASR) is a method of converting an acoustic waveform into text that should be similar to the information being conveyed by spoken words [2]. Speech recognition takes input through voice, in the same way as a computer receives input through keyboard and mouse [3]. With the development of the mobile and internet communications, ASR systems have been applied in different fields in the recent times. New application directions include keyword spotting, multilingual language interpretation, natural spoken dialogue, and human-machine interaction [4]. Google Speech Recognition (GSR) and Apple Siri are popular applications that enable people to interconnect with search engines by using their voice to perform tasks such as sending email and texts, calling contacts and businesses, listening to music, browsing the web, and other common tasks [5].

II. SINDHI LANGUAGE

With fifty-two letters in the alphabet, Sindhi is one of the world's oldest languages. This language is spoken, read, and written, all over the world [6]. The official language of Sindh province is Sindhi, spoken by 53 million speakers in Pakistan, and different areas of the world [7]. Sindhi is a world-renowned ancient and unique language with a complex grammar and extensive morphology. There are numerous Sindhi-language magazines, social media sites, newspapers, explicit websites, and web journals that provide an information on various Sindhi topics. Sindhi grammar differs from that of English and other languages, as do the meanings and implications of Sindhi lexicons. The use of diacritics in Sindhi text causes the meaning, number, and gender of Sindhi lexicons to change [8].

III. LITERATURE REVIEW

Android speech recognition can be used in a variety of situations, including home appliances, security systems, aids for handicapped people, robotics, artificial wearable sensor [15], visual speech recognition [16] voice to text and others [17].

A. Voice to Text

The research presented by [9], proposed an Android application that will help users locate a lost mobile phone in the vicinity, regardless of whether the phone is on mute. As a result, it saves time. This application has been used in two different modes. (1) Through the use of voice (2) Through the use of text messages. To locate the lost phone, the user must speak and submit a unique keyword. The keyword in the database can be checked using the "Text matching algorithm." Though the keyword is entered correctly, the phone will ring, even if it is set to mute. For encryption, the Advanced Encryption Standard (AES) algorithm is used. This application allows a user to locate a lost mobile phone in a specific area and has been designed to be user-friendly for people with disabilities.

B. Home Appliances

In [10], the study proposed a home-based smart system which allows user to handle and control various home appliances smart home automation system that allows users to control electrical appliances in their homes by using an

Android smartphone. The home automation system recognizes the input speech and allows to control home appliances with voice commands. Support Vector Machine (SVM) classifier that extracts the voice commands, has been used to recognize speech. General Packet Radio Service (GPRS) technology is used to implement the home automation system wirelessly. The application has been developed using the Java programming language. Electrical appliances such as fans, current sensors, light sensors, and light switches are incorporated into a device, and commands are performed given by user. The android application acts as a transmitter, sending ON/OFF commands to the receiver. The primary outcomes with a classification rate 96% are given by the sound classification by applying SVM classifier.

C. Security Systems

An android application for a security system applied on the door by using Bluetooth technology [11]. GUI based along with easiness and simple interface allows commands to be sent quickly from a mobile to the appliance such as checking the status of lock of the door, the status when door is unlocked or checking of the door security and its types. The user commands are responded by the security system by actuating various actions and then acknowledging back to the user or the user mobile. If user loses his or her phone, the typical method of lock namely key and the lock will still be in the working interface as backup. A device can be communicated up to 100 meters far enough, by the "piconet" set up, supported by the Bluetooth. This application is implemented by the Java language. The door, a keyhole, and two LEDs made up the user interface. The power indicator was one LED, and the door ajar indicator was the other. A pin could be actuated for unlocking and locking of the door by the system via small margin. The status of door could also be checked by the system if user presses a button.

D. Aids for Handicapped people

The study of [13] proposed a smart wheelchair for physically challenged persons, which uses an android phone to regulate the rotation of the wheelchair through voice and gesture movement as well as to read SMS, e-mail, and news. There are eight sensors in total implemented in this android application, two of which are Infrared Radiation (IR) sensors, and the rest are light detection, smoke detection and temperature sensors. The main three units has been applied in this application. To recognize voice and gesture by means of android and motor control by means of signal conditioning. In this application the panic button has also been added, which will sound a buzzer in the event of an emergency. Using an infrared sensor, obstacles are removed. The disabled people can rotate the wheelchair inside of the house without any assistance. This makes it easy to use.

IV. METHODOLOGY

The framework discussed in this paper is being introduced as an application on Android, running above the version of Android 5.0. The flow chart of the speech recognition system is shown in the Figure 1, where it transmits the voice towards the mobile interface. The voice was captured by the android application. The captured voice was compared for execution with the database. If the command is found against the command issued, the system will be able to perform the activity. If the given command could not match for some reason, the alert message will be appeared.

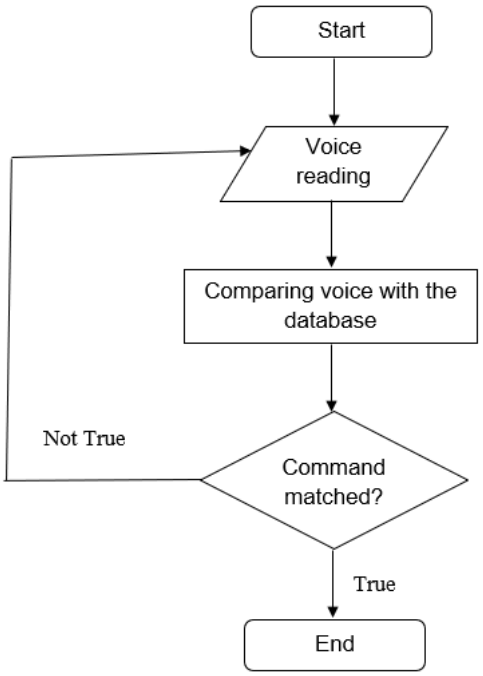

Figure 1: Flow chart of Speech recognition android application

The parameters of the speech recognition framework are as follows: The sampling frequency measured was 16 kHz. The corpus has been developed using 20 words in Sindhi language. A total of twenty-five subjects as Sindhi native speakers whereas fifteen Sindhi non-native speakers were approached and asked to repeat the entire sentence 15 times in the session of recording. Every sentence was replaying to ensure that the same word was spoken included into the captured signal.

V. EVALUATION

When native Sindhi speakers were evaluated, the results were as follows: The word "پيغام (message)" had the lowest accuracy score of 70%, while the words "تصويرون (gallery)" and "وقت (time)" had an accuracy score of 80%, and the words "موسم (weather)", "ٻڌاءِ (tell)", "لطيفا (jokes)", "نالا (contacts)", "کول (open)", "نقشو (map)", and "موڪل (send) had the highest accuracy score of 90%. The overall accuracy of native Sindhi speakers was 95%.

When non-native Sindhi speakers were evaluated, the results were as follows: The words "نالا (contacts)", and "تصويرون (Pics/Gallery)" had the lowest accuracy score of 50% and 60%, whereas the words "پيغام (message)", "لطيفا (jokes)", "نقشو (map)", and "موسم (weather)" had an accuracy

score of 70% and the words "کول (open)", "موڪل (send)", "وقت (time)", and "ٻڌاءِ (tell)" had the highest accuracy score of 80%. Non-native Sindhi speakers were 85 percent accurate overall. Non-native Sindhi speakers had the lowest precision when compared to native Sindhi speakers.

VI. WORKING

To view the application while it is being built in the react native technology, the USB Debugging must be activated on an Android phone. To activate USB debugging, first enable "Developer options" through the phone's settings and then tapping seven times the "Build number" line at the end. Then go back to settings and enable USB debugging [14] as shown in the Figure 2. The application will be permitted by the user initially to record audio as shown in the Figure 3.

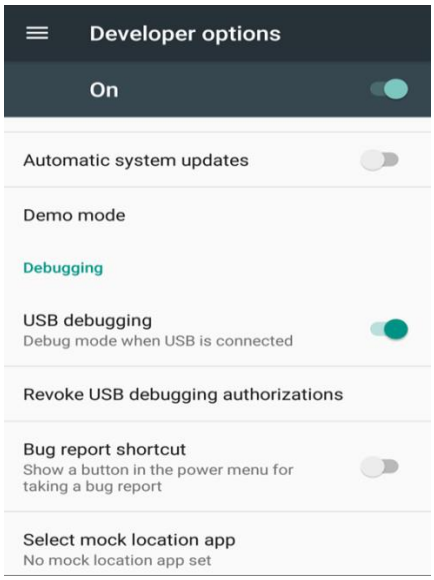

Figure 2: USB Debugging Activation

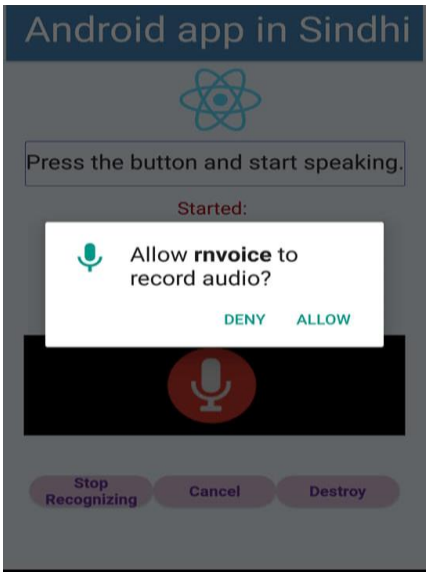

Figure 3: Recording permission by user

The application has only one point of view, the primary view. The main view is where the application starts. The user will be asked through the main view of the screen to speak by pressing the mic button. It will finally come to an end, and then the recognition will begin. If the user does not want his or her voice to be heard or recognized, the three choices are also available at the bottom of the application screen as shown in Figure 4.

Figure 4: Android application main view

A. *Pitch and Error*

Pitch is a critical feature of voiced speech. It contains the knowledge that is exclusive to the speaker. The application will generate an error if the voice of user has not been matched as shown in Figure 5.

Figure 5: Error message

VII. Results

The result of android application is as follows: If user commands "نقشو کول" through android speech recognition system, the action will be performed according to the desired command, and "google map" will be displayed on the android screen as shown in Figure 6.

Figure 6: Result of ASR in Sindhi language

When the user inputs unclearly under some condition, for example, in a noisy environment where several people are conversing, and the language input cannot be used, the "Unknown command" alert message will be shown on the android screen as shown in Figure 7 by the android application.

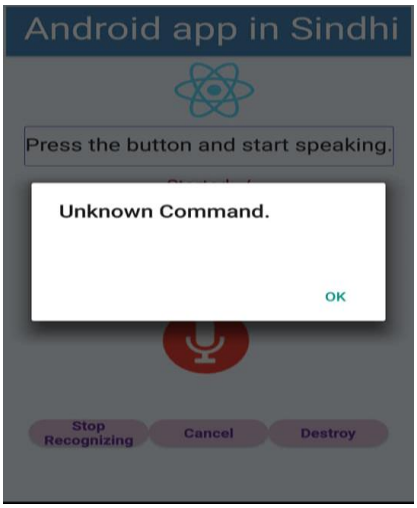

Figure 7: Alert Message

A total of 180 subjects were chosen, with 100 Sindhi native speakers and 80 non-native Sindhi speakers. A total of 800 words were chosen for native and non-native Sindhi speakers to speak. To ensure that the application works, all recordings were made in a quiet environment with some background noise. Non-native Sindhi speakers had a total accuracy of 85%, whereas native Sindhi speakers had a total accuracy of 95%.

VIII. Conclusion

Many studies on voice recognition for different languages have been conducted, but no such study has been done on Sindhi voice recognition in Android. This paper presents the study of controlling Android based smart phones by using voice in Sindhi language. A user can easily use this Android Sindhi speech recognition system where he/she does not have to type a command. Voice-activated control systems can reduce the amount of time and effort needed. The amount of time spent typing would be greatly decreased, and it would also be helpful to people with disabilities. The use of React Native as an easy and efficient technology for the current study and the resulting android application could pave the way for the creation of further ASR applications in other alternative native languages and using different devices, allowing users to use speech recognition without having to learn English language.

References

[1] Li, D., & Du, Y. (2017). Artificial intelligence with uncertainty. CRC press.

[2] Ghai, W., & Singh, N. (2012). Analysis of automatic speech recognition systems for indo-aryan languages: Punjabi a case study. Int J Soft Comput Eng, 2(1), 379-385.

[3] Gosavi, S. D., Khot, U. P., & Shah, S. (2013). Speech recognition for robotic control. Int. J. Eng. Res. Appl, 3(5), 408-413.

[4] Eljagmani, H. M. (2017). Arabic Speech Recognition Systems(Doctoral dissertation).

[5] Assefi, M., Liu, G., Wittie, M. P., & Izurieta, C. (2015). An experimental evaluation of apple siri and google speech recognition. Proccedings of the 2015 ISCA SEDE.

[6] https://sindhinlp.com/

[7] Hakro, D. N. (2015). Enhanced Segmentation and Feature Extraction for Sindhi Optical Character Recognition (Doctoral dissertation, Universiti Sains Malaysia).

[8] Dootio, Mazhar. (2018). An analysis and solution of computational linguistics problems of Sindhi text. 10. 8-24.

[9] Tekwani, B., Bhandiwad, V., & Dhuri, K. (2018, August). FIND IT An Android application to find misplaced phone of user in the vicinity. In 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2018 2nd International Conference on (pp. 769-774). IEEE.

[10] Sangeetha, S. B. (2015, March). Intelligent interface based speech recognition for home automation using android application. In Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015 International Conference on (pp. 1-11). IEEE.

[11] Potts, J., & Sukittanon, S. (2012, March). Exploiting Bluetooth on Android mobile devices for home security application. In Southeastcon, 2012 Proceedings of IEEE (pp. 1-4). IEEE.

[12] Alexenko, T., Biondo, M., Banisakher, D., & Skubic, M. (2013, March). Android-based speech processing for eldercare robotics. In Proceedings of the companion publication of the 2013 international conference on Intelligent user interfaces companion (pp. 87-88). ACM.

[13] Khadilkar, S. U., & Wagdarikar, N. (2015, January). Android phone controlled voice, gesture and touch screen operated smart wheelchair. In Pervasive Computing (ICPC), 2015 International Conference on (pp. 1-4). IEEE.

[14] https://reactnative.dev/docs/running-on-device

[15] Yang, Q., Jin, W., Zhang, Q., Wei, Y., Guo, Z., Li, X., ... & Ren, T. L. (2023). Mixed-modality speech recognition and interaction using a wearable artificial throat. Nature Machine Intelligence, 5(2), 169-180.

[16] Bhaskar, S., & Thasleema, T. M. (2023). LSTM model for visual speech recognition through facial expressions. Multimedia Tools and Applications, 82(4), 5455-5472.

[17] Li, J. (2022). Recent advances in end-to-end automatic speech recognition. APSIPA Transactions on Signal and Information Processing, 11(1).