Voice Based Form Filling System For Visually Challenged People

S. Usharani[1], P. Manju Bala[2], R. Balamurugan[3]
Associate Professor[1,2] , Department of CSE[1,2], IFET College of Engineering, Villupuram, TamilNadu, India
ushasanchu@gmail.com[1], pkmanju26@gmail.com[2]
Final year Student[3], Department of CSE[3], IFET College of Engineering, Villupuram, TamilNadu, India
balamuruganbrr@gmail.com

***Abstract*— Now a days, people are facing so many difficulties to fill the form in various places like Bank, Hospital etc., some people don't know to read and write (illiterate) and some of them are blind peoples. By using the recognition of voice, we can develop the automatic form filling system for the handicapped and blind peoples. These people can fill the form without touching anything and it is fully operated by voice. Voice Recognition play a major role in the world, to control anything using voice. It gives easy way to do the operation of the system. This recognition is a computer software program with the ability to decrypt the human speech. Voice recognition is commonly used to operate the device and write without having to use a console, mouse or press any keys. So, this form filling system is very useful for the society and reduce the wastage of time with the paper and pen, if there is any mistake while filling the form there may also cause wastage of paper but in our system the wastage of time and paper is never occurs. The people comfortable to use the app is developed to fill the form by using voice. After filling the details, the mobile is connected with the printer by using the OTG adaptor and take print of the form. It is the process of the required system that should be handled by the user. So, it is useful for the society.**

Keywords*— *Google voice recognition, Text to Speech, Speech to Text, Voice access, Button mapping technique, OTG adaptor, printer, Android studio application system.

I. INTRODUCTION

In our day to day life, peoples are filling forms in so many places like Bank, Hospital, Schools, college events etc., in the present scenario data has more value when compare to money. In future it become a tremendous development for getting data from various person and sell it out to other organization companies. If you make any mistake while filling the form some peoples just throw out that paper, they don't know what happen next if any third-party person get that partially filled details, they can do anything because it is digital world. Traditionally, repeated criminals are identified by their biometrics such as thumbprint. But criminals are smart enough to not to leave their biometrics in crime scene. In developed countries, the government create dataset which is helpful for recognize the human face which compares the suspicious act with trained dataset and information stored in database.

While filling the form in many places, there is some illiterate peoples are here. They doesn't know how to fill the form, even they can't able to read and write so there is a first problem occurs in that place then these people call some person for help to filling the form, if suppose the person is a third party criminal it will become more serious. And the next one is handicapped and blind peoples they cannot able to fill the form even it is a simple digital payment like Paytm, Phonepe, Google pay etc., so how to help these kinds of people. Let's move to innovation idea to help the handicapped and blind peoples. By using the voice recognition-based form filling system the blind people's problem is solved.

II. METHODOLOGY

Peoples are filling so many forms, but in our system the development of form filling is helpful for Banks. In this place every person must secure the confidential data, because each and every second you are watched by any other third person. In this methodology there is one app which can be used for illiterate people, blind people. These are the three different view for that only one application. For illiterate people, they know to speak and hear in our regional language but cannot able to write and some person have struggling to read. In that place we developed the app in our own regional language, for the people comfortable use it is designed as a mobile application. The person open the app and click the button its will as questions to fill the form, then the person give the answer in our regional language it fill that form in that language with the help of speech to text. For blind people, they have a personal mobile phone but how that people can open that particular app in the mobile phone. In that place the app specialization help's that people, when a blind people have a headset just plugin that, then the app will be open but how the people can accurately click the button the mobile. As same with the help of headset button these people can control that app button. If you click the headset button at the first time it asks the question next if you click the second button the app will get the input from your voice based on that process the app will work. For handicapped people, if suppose the person may be handless, they may also use that app. Just call the name of the app then the app will be open, by using the voice access the person may control the button of the application, they just call the button name then the work of particular button working action is performed. These are the three different view of the app that will be helpful for our society.

After filling the all details, you just connect the mobile with the printer with the help of OTG adaptor by using your

 ISBN 978-1-7281-6202-7

voice access the printing process may also occurs. These are the methodology we are used.

For opening the app, the headset plug action process is done, Google text to speech is used for English language, then recorded voice should be played for the regional language. To get the input from the user the speech to text method is to be used. To control all the activity of the app using the headset button the button mapping technology is used. Controlling all the process using voice with the help of voice access technology. So this app will be helpful for the society and reduce the unauthorized person can get the details form others. It is a secured process and difficulties for filling the form is totally reduced and it is a user-friendly app, these are the providence of that application. This system is not like a big embedded system you can carry out it any place and reduces the wastage of time.

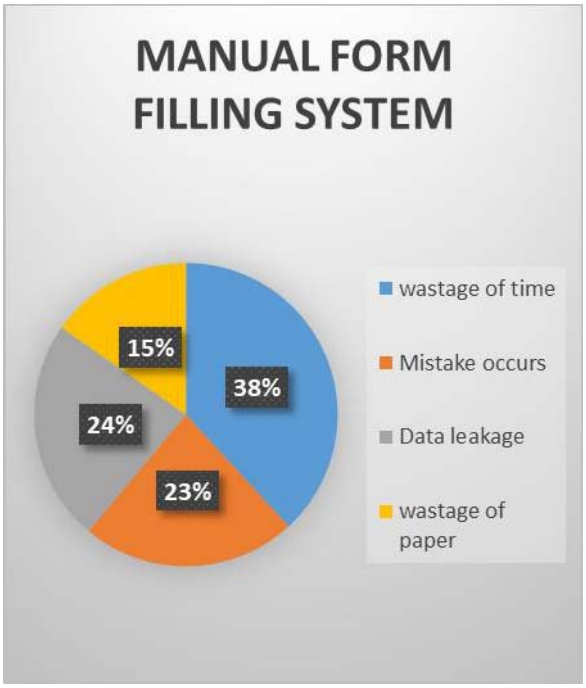

Fig-1: Statistics

In the manual form filling system, there is so many disadvantages. These problems are wastage of time, Mistake occurs, Data leakage and wastage of paper. This may occur while filling the form manually. The data flow of voice Synthesis is done with the help of text to speech and speech to text. Then finally the output is printed in the pdf formed sheet. The system is done by the android application system. The app will open when the headset is plug in. Buttons of the application will be controlled by the headset button. That is the way of filling the form in the voice-based method. After filling the all details, you just connect the mobile with the printer with the help of OTG adaptor by using your voice access the printing process may also occurs.

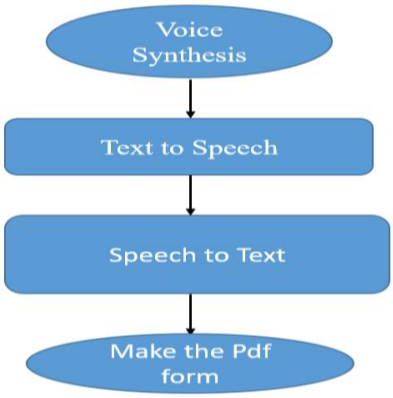

Fig 2: Data Flow Diagram

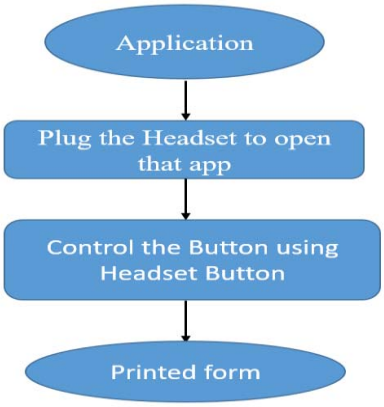

Fig-3: System flow diagram

III. IMPLEMENTATION

System is implemented with the help of Android studio IDE by the Intellij platform, Gradle script and supported plugin. In this application we are provide three modules of view. First one is for Illiterate peoples, second one for the handicapped peoples and the next module view for the blind peoples. All the three-module view is supported in one application. It is the proposed system that is developed. In this application it contains Google voice recognition, Text to Speech, Speech to Text, Voice access, Button mapping technology, OTG adaptor, printer, and Android studio application system

A. Voice Recognition

The voice Recognition is used to get the voice input from the user and type the text with the help of speech to text. It can also get the input in our regional language. These recognitions are a computer software program with the ability to decrypt the human speech.

B. Text to speech

These text to speech is only used for the English language. For our regional language the recorded voice should be used and it is directly bind with the application in the raw resource directory. The main usage of the text to speech method is to reduce the work of development process and the text can be modified, reused easily.

C. Voice access

The voice access is commonly used to operate the device without having to use a console, mouse or press any keys. The user can also access the device in the regional language with the help of the voice access.

D. Button mapping

To enhance the facility of the app the button mapping is used. Is there is any sound disturbance around the surrounding the voice access makes wrong to reduce that effect the button mapping is to be used. Here, the user may control all the operations of the app with the help of headset button with the help of mapping technique.

E. OTG adaptor

This adaptor is generally used to connect the mobile with the printer and take the print of the form in that application. It is the easy way that can be helpful to all the user and it is more comfortable. The cost of the OTG adaptor is very low cost in present it is available at rupees 30. It is also called as an interface between the mobile phone and the printer. While printing the app will hide all the button in the activity and print only the form.

F. Android application

For illiterate people, they know to speak and hear in our regional language but cannot able to write and some person have struggling to read. In that place we developed the app in our own regional language, for the people comfortable use it is designed as a mobile application. The person open the app and click the button its will as questions to fill the form, then the person give the answer in our regional language it fill that form in that language with the help of speech to text.

For blind people, they have a personal mobile phone but how that people can open that particular app in the mobile phone. In that place the app specialization help's that people, when a blind people have a headset just plugin that, then the app will be open but how the people can accurately click the button the mobile. As same with the help of headset button these people can control that app button. If you click the headset button at the first time it asks the question next if you click the second button the app will get the input from your voice based on that process the app will work.

For handicapped people, if suppose the person may be handless, they may also use that app. Just call the name of the app then the app will be open, by using the voice access the person may control the button of the application, they just call the button name then the work of particular button working action is to be performed.

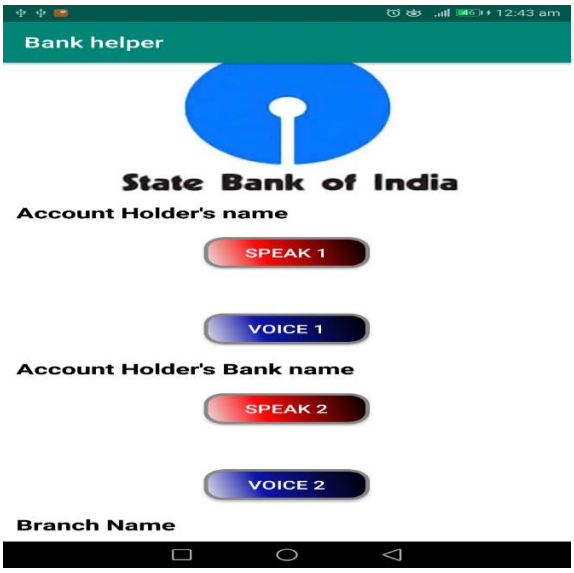

Fig 4: App View Initially

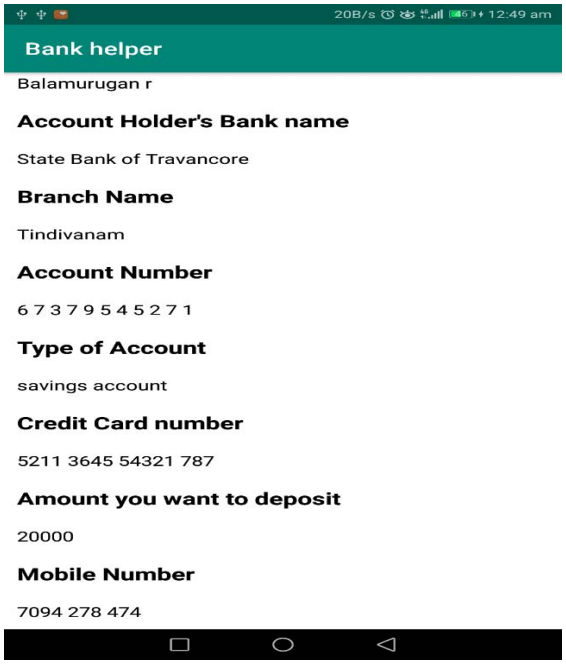

Fig 5: Final Printed form

IV. Results and Discussion

This app will produce the printed form as a output. It can be used for further process in banking. It will print the application only when the mobile is connected with the OTG adaptor and the printer, then only it produces the output of the form filling process. The recognition of voice does not produce any mistakes while filling the form. This app will be useful for the society and very easy to handle by the user. The statistic details of the current form filling system give more problems to the people and the working flow of the application that is represented in the form of flowchart. The final document after filling all the details and the button in the activity is hided and then print format.

CONCLUSION

This app will be helpful for the society and reduce the unauthorized person can get the details form others. It is a secured process and difficulties for filling the form is totally reduced and it is a user-friendly app, these are the providence of that application.

This system is not like a big embedded system you can carry out it any place and reduces the wastage of time and it is mainly useful for our society.

REFERENCES

[1] R. Serizel and D. Giuliani, "Deep neural network adaptation for children's and adult's speech recognition," in Proceedings of Italian Computational Linguistics Conference, 2014.

[2] P. Shivakumar, A. Potamianos, S. Lee, and S. Narayan, "Improving children's speech recognition using acoustic adaptation and pronunciation modeling," in Proceedings of the Workshop on Child Computer Interaction, 2014.

[3] H. Liao, G. Pundak, O. Siohan, M. Carroll, N. Coccaro, Q. Jiang, T. Sainath, A. Senior, F. Beaufays, and M. Bacchiani, "Large vocabulary automatic speech recognition for children," in Proceedings of Interspeech, 2015.

[4] J. Fainberg, P. Bell, M. Lincoln, and S. Renals, "Improving children's speech recognition through out-of-domain data augmentation," in Proceedings of Interspeech, 2016.

[5] D. Palaz, R. Collobert, and M. Magimai.-Doss, "Estimating phoneme class conditional probabilities from raw speech signal using convolutional neural networks," in Proceedings of Interspeech, 2013.

[6] T. N. Sainath, R. J. Weiss, A. Senior, K. W. Wilson, and O. Vinyals, "Learning the Speech Front-end With Raw Waveform CLDNNs," in Proceedings of Interspeech, 2015.

[7] P. Golik, Z. Tuske, R. Schl ¨ uter, and H. Ney, "Convolutional ¨ Neural Networks for Acoustic Modeling of Raw Time Signal in LVCSR," in Proceedings of Interspeech, 2015.

[8] D. Palaz, M. Magimai-Doss, and R. Collobert, "Endto-End Acoustic Modeling using Convolutional Neural Networks for HMM-based Automatic Speech Recognition," Speech Communication, 2019. [Online]. Available: https: //doi.org/10.1016/j.specom.2019.01.004

[9] H. Muckenhirn, V. Abrol, M. Magimai.-Doss, and S. Marcel, "Gradient-based spectral visualization of CNNs using raw waveforms," Idiap Research Institute, Tech. Rep. Idiap-RR-11-2018, Jul 2018. [Online]. Available: http://publications.idiap.ch/downloads/reports/2018/ Muckenhirn Idiap-RR-11-2018.pdf

[10] A. Batliner, M. Blomberg, S. D'Arcy, D. Elenius, D. Giuliani, M. Gerosa, C. Hacker, M. Russell, S. Steidl, and M. Wong, "The PF STAR children's speech corpus," in Proceedings of Ninth European Conf. Speech Communication and Technology, 2005.

[11] T. Robinson, J. Fransen, D. Pye, J. Foote, and S. Renals, "WSJCAMO: a british english speech corpus for large vocabulary continuous speech recognition," in Proceedings of ICASSP, 1995.

[12] "BEEP dictionary," http://svr-www.eng.cam.ac.uk/comp. speech/Section1/Lexical/beep.html, accessed: 01-07-2018.

[13] "MGB challenge lexicon," http://data.cstr.ed.ac.uk/asru/ MGB3/data/lm/mgb.normalized.lm, accessed: 01-07-2018.

[14] D. Povey et al., "The Kaldi speech recognition toolkit," in IEEE workshop on Automatic Speech Recognition and Understanding, 2011.

[15] F. Chollet et al., "Keras," https://github.com/fchollet/keras, 2015.

[16] M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," http://tensorflow.org/, 2015.

[17] M. J. Russell, S. D'Arcy, and L. P. Wong, "Recognition of read and spontaneous children's speech using two new corpora," in Proceedings of Interspeech, 2004.

[18] M. J. Russell and S. D'Arcy, "Challenges for computer recognition of children's speech," in Proceedings of the Workshop on Speech and Language Technology in Education (SLaTE), 2007, pp. 108–111.

[19] H. Muckenhirn, M. Magimai.-Doss, and S. Marcel, "Towards directly modeling raw speech signal for speaker verification using CNNs," in Proceedings of ICASSP, 2018.

[20] S. Kabil, H. Muckenhirn, and M. Magimai.-Doss, "On learning to identify genders from raw speech signal using cnns," in Proceedings of Interspeech, 2018. [25] "American vowels database," https://homepages.wmich.edu/ ∼hillenbr/voweldata.html, accessed: 15-07-2018.

IEEE ICSCAN 2020 ISBN 978-1-7281-6202-7