Voice-Based Login System

Vedant Baviskar[1], Shubham Bhagwat[2], Deepjyoti Barman[3], Gokul Viswanath[4], Mrs. R. T. Waghmode [5]

B.Tech Students, Department of Computer Science Engineering[1,2,3,4,5]

Sinhgad Institute of Technology and Science, Pune, Maharashtra, India

Abstract: *Off late, biometric modes of authentication have become pretty famous. People use their fingerprints to unlock their devices. They use their faces to unlock their devices. However, using voice recognition for unlocking has not being accepted for wide use. We were curious to see why that is and if there are downsides to that, would we be able to figure out workarounds for that.*

Keywords: Audio Content Analysis, Moving Pictures Expert Group, Voice Recognition.

I. INTRODUCTION

The project can be considered a crossover of two major industry standards. Those are authentication and voice recognition. Authentication has been in use since pretty long now and voice recognition has just now gotten attention. So, these methods together will be used together to firstly, accept a voice clip and separate the contents into the voice and noise and secondly, compare the uploaded voice with the ones stored in the cloud after registering. If the voice exists in the database, we log them in. Else, we don't. This paper will delve into how we set our methods to authenticate them.

II. MOTIVATION

Analysing an audio clip as well as getting and understanding its characteristics has been a point of interest recently. Many individuals as well as small-to-large organizations have been recently focusing on developing technology that can read and apply properties of audio and especially voice.

Such analyses and implementations can provide useful information to us about audio, and how they can be used in a wide range to applications that we use every day as part of our lifestyle. The voice is something you are, and uniquely defines you and your characteristics. This can help people and organizations introduce more such platforms that use Voice Authentication, and thus make the users' life easier while maintaining (or even strengthening) the level of security.

Digital devices are nowadays an essential part of our everyday life, as they are used for a variety of applications. Performing biometric authentication through devices can provide a stronger mechanism for identity verification as the two authentication factors, "something you have" and "something you are," are combined.

III. RELATED WORK

1. IBM uses Voice recognition and feature extraction for their services in IBM Watson (for example- Text to speech)
2. Apple, Google, Microsoft and Samsung use Voice Recognition and to some extent- Authentication too in their virtual assistants, named Siri, Google Assistant, Cortana and Bixby respectively.
3. Amazon Echo is a small smart speaker which performs actions of the user by listening to voice commands.
4. Nuance has developed an Automatic Speech Recognition (ASR) system that uses Voice Recognition.

IV. LITERATURE REVIEW

SR NO.	Paper Title	Paper Theme/Idea	Conclusion
1	Unsupervised feature learning for audio classification using convolutional deep belief networks	Application of convolutional deep belief net- works to audio data and evaluate based on the results in order to get things working	A single CDBN feature representation can achieve high performance on multiple audio recognition tasks.
2	Deep Learning for Audio Signal Processing	Review of the latest deep learning techniques for audio signal processing	Compared to conventional approaches, this deep neural networks usually require

			more computation power and more training data.
3	An Introduction to Audio ContentAnalysis Applications in Signal Processing and Music Informatics	Overview of ACA techniques and applications	Lack of large-scale training data will continue to impact the progress and methods applied to audio content analysis.
4	librosa: Audio and Music Signal Analysis in Python	Design of library along with techniques	There is room for improvement of efficiency and accuracy
5	Biometric Audio Security	Introduced us to a technique called additive synthesis, along with a predecided value of harmonics.	Provided us with a new way to think and rate the audio based on the additive synthesis and allowed us to give a value to the waves
6	Audio Data Analysis Using Deep Learning with Python (Part 1)	How to process audio and analyse it using Python and find out various characteristics of the audio samples	Audio can be analysed with Python by leveraging various libraries like Librosa and it's unique features can be found

V. CONCLUSION

The use of voice recognition in our project will be that it'll play the role of the user's credential. Typically, people use passwords, fingerprint sample as credentials. Our aim is to use voice samples as credentials and then let users login to systems using just their voice, an identity that nobody can steal from them and that they don't need to worry about forgetting after a few months.

REFERENCES

[1]. J. Foote, "Content-based retrieval of music and audio," Proc. SPIE, vol. 3229, pp. 138–147, 1997.

[2]. E. Wold, T. Blum, and J. Wheaton, "Content-based classification, search and retrieval of audio," IEEE Multimedia, vol. 3, no. 3, pp. 27–36, 1996.

[3]. S. Pfeiffer, S. Fischer, and W. Effelsberg, "Automatic audio content anal- ysis," in Proc. 4th ACM Int. Conf. Multimedia, 1996, pp. 21–30.

[4]. J. Saunders, "Real-time discrimination of broadcast speech/music," in Proc. ICASSP'96, vol. II, Atlanta, GA, May 1996, pp. 993–996.

[5]. E. Scheirer and M. Slaney, "Construction and evaluation of a robust mul- tifeature music/speech discriminator," in Proc. ICASSP' 97, Apr. 1997, vol. II, pp. 1331–1334.

[6]. D. Kimber and L. Wilcox, "Acoustic segmentation for audio browsers," in Proc. Interface Conf., Sydney, Australia, July 1996.

[7]. T. Zhang and C.-C. J. Kuo, "Video content parsing based on combined audio and visual information," Proc. SPIE, vol. IV, pp. 78–89, 1992.

[8]. J. P. Campbell, Jr., "Speaker recognition: A tutorial," Proc. IEEE, vol. 85, no. 9, pp. 1437–1462, 1997.

[9]. A. V. McCree and T. P. Barnwell, "Mixed excitation LPC Vocoder model for low bit rate speech coding," in IEEE Trans. Speech Audio Processing, July 1995, vol. 3, pp. 242–250.

[10]. K. El-Maleh, M. Klein, G. Petrucci, and P. Kabal, "Speech/music dis- crimination for multimedia application," in Proc. ICASSP'00, 2000.

[11]. Atakan Kantarci in Top 10 Voice Recognition Applications in 2021

[12]. Niko Laskaris in How to apply machine learning and deep learning methods to audio analysis

[13]. Nagesh Singh Chauhan in Audio Data Analysis Using Deep Learning with Python (Part 1)

[14]. S. Vetrivel et. al. / Indian Journal of Computer Science and Engineering Vol. 1 No. 4 240-250 "AN OVERVIEW OF MPEG FAMILY AND ITS APPLICATIONS"

[15]. Thomas Sikora, "MPEG Digital Video Coding Standards", 1997

[16]. Davis Yen Pen, "Digital Audio Compression" , Digital Technical Journal Vol. 5 No. 2, Spring 1993