iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VIII **Month of publication:** August 2021

DOI: https://doi.org/10.22214/ijraset.2021.37714

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

Digital Assistant for Sound Classification Using Spectral Fingerprinting

Ria Sinha[1], Rishi Sinha[2]
[1, 2]*BASIS Independent Silicon Valley*

***Abstract:** This paper describes a digital assistant designed to help hearing-impaired people sense ambient sounds. The assistant relies on obtaining audio signals from the ambient environment of a hearing-impaired person. The audio signals are analysed by a machine learning model that uses spectral signatures as features to classify audio signals into audio categories (e.g., emergency, animal sounds, etc.) and specific audio types within the categories (e.g., ambulance siren, dog barking, etc.) and notify the user leveraging a mobile or wearable device. The user can configure active notification preferences and view historical logs. The machine learning classifier is periodically trained externally based on labeled audio sound samples. Additional system features include an audio amplification option and a speech to text option for transcribing human speech to text output.*
***Keywords:** assistive technology, sound classification, machine learning, audio processing, spectral fingerprinting*

I. INTRODUCTION

According to the World Health Organization (WHO), 466 million people worldwide, about 6% of the population, including 34 million children, suffer from disabling hearing loss [1]. It is estimated that by 2050 over 900 million people will have disabling hearing loss. Disabling hearing loss refers to hearing loss greater than 40 decibels.

A. Causes and Effects of Hearing Loss

Several causes have been identified for hearing loss. The Hearing Loss Association of America (HLAA) has categorized hearing disabilities into two classes: (i) conductive, which include problems associated with the ear drum, ear canal, and middle ear function, and (ii) sensorineural, which include issues that affect inner ear functions as well as the brain and nervous system that interpret auditory inputs [2]. Conductive hearing loss is prompted by ear infections, benign tumors, excessive ear fluid and/or ear wax, or poor function of the ear tubes. As for sensorineural issues, researchers have labeled traumatic events, aging, hereditary, and virus or immune diseases as primary causes Many individuals are affected and require additional assistance in their daily lives, because of the multitude and diverse range of issues that lead to this disability.

Individuals in the deaf and hearing loss community have faced discrimination and oppression for centuries. This has caused challenges for them in terms of employment, higher education, and other privileges that their hearing counterparts take for granted. They are often stereotyped and marginalized in society, and their communication barriers have led to a strained relationship with the rest of the community making it difficult for them to live normal daily lives. A study published by the British Department of Health suggests that hearing-impaired individuals are 60% more susceptible to mental health and social anxiety issues than their counterparts with normal hearing abilities [3].

B. Solutions for Hearing Loss

The most popular solution for hearing loss is the hearing aid [4]. Hearing aids are electronic devices generally worn behind or inside the ear. The device is usually battery powered. It receives sound through an embedded microphone, which converts the sound waves to electrical signals that are processed, amplified, and played back using a speaker. The amplifier increases the power of the sound signal that would normally reach the ear, allowing the hearing-impaired user to listen. Hearing aids are primarily useful for people suffering from sensorineural hearing loss which occurs when some of the small sensory cells in the inner ear, called hair cells, are damaged due to injury, disease, aging, or other causes. Surviving hair cells leverage the amplified sound signal generated by the hearing aid to compensate for the loss of perception that would occur otherwise and convert the received sound signal into impulses sent to the brain via the auditory nerve. However, if the inner ear is too damaged, or the auditory nerve has problems, a hearing aid would be ineffective. The cost of a hearing aid can range from $1,000 to $6,000 [5].

For people suffering from profound hearing loss, cochlear implants may be an option. Cochlear implants are surgically implanted neuro-prosthetic devices that bypass sensory hair cells used in the ear for normal hearing and attempt to directly stimulate the auditory nerve with electrical signals.

With prolonged therapy and training a hearing-impaired person may learn to interpret the signals directly sent to the auditory nerve as sounds and speech. Cochlear implants can cost approximately $100,000, and for pre-lingually deaf children the risk of not acquiring spoken language even with an implant can be as high as 30%. A variety of assistive technologies have emerged over the years to the help the hearing-impaired. These include FM radio-based systems that transmit radio signals from a speaker to a listener and audio induction loop systems that pick up electromagnetic signals using a telecoil in a hearing aid, cochlear implant, or headset [6]. Mobile devices with touch screens and real-time text to speech transcription capabilities are starting to get used as well. Closed captioning is becoming standard in streaming media as well as television programs. Apple recently launched a feature called "Live Listen", on iOS mobile devices (e.g., iPhone, iPad) where the device becomes a remote microphone placed close to a speaker and a Bluetooth headset replays the sound live [7]. This can be useful when you are trying to hear a conversation in a noisy room or for a hearing-impaired student trying to listen to a teacher across the classroom.

C. Limitations of Current Solutions

These approaches to assisting the hearing-impaired attempt to create a real-time listening experience like a normal person by using technology to work around the defects of the ear. They can be expensive and sometimes aesthetically undesirable. Hearing aids are battery powered and not something a user would like to wear all the time. There are several situations where a hearing-impaired user may just want to be notified about interesting ambient sounds without having to wear a hearing aid. For example, a user may be sleeping and may want to get alerted if there is a knock on the door, a baby crying, a smoke alarm going off or other similar audio events that warrant some action. A digital assistant that can actively listen, process, and notify the user via a vibration alert on a smart watch can be very useful in these circumstances. Hearing-impaired people often get anxious when they are in new surroundings because systems they have in their house (e.g., a visual doorbell or telephone) may not be available in a hotel. Having a digital assistant that can intelligently process ambient sounds and notify them via their mobile device can be very useful in these circumstances and allow the hearing-impaired user to operate more confidently. The digital assistant should be customizable such that the user can specify notification preferences based on audio categories, audio types, time of day, location, etc., and allow the user the view historical alerts. The assistant can run as an app on a mobile device or integrate with smart listening devices such as Amazon Alexa and Google Home that have omnidirectional microphone arrays that can pick up sounds coming from any direction.

D. Focus of this Work

This work studies the spectral characteristics of ambient sounds and illustrates how they can be used as features to a machine learning algorithm that can classify them with proper training. The model is then incorporated into a digital assistant mobile app for the hearing-impaired. The app obtains audio signals from the microphone and first runs signal processing steps to reduce background noise and interference. Subsequently, it runs the machine learning based classifier to analyze the audio signal and classify it into an audio category and audio type. The user is then notified, based on their preferences, with a summary of the classified audio, and, for the specific type of audio, the user is presented with a meaningful description of what the machine learning classifier identified the signal as. Notifications can be stored in the system for historical viewing. Optionally, the system may include an amplifier and filter to output the received audio signal to an audio output of the user's choice or store it as an audio file for future playback. The system can also include a speech to text module that can decipher human speech and provide a text transcript of the speech in real-time on the user's notification screen. The apps machine learning classifier is periodically trained externally based on labeled audio data and updated automatically or manually.

II. AUDIO PROCESSING

Audio signals are a representation of sound that humans perceive. Sound is a vibration that propagates as a wave in a medium such as air or water. Waves are generated when a vibrating source, such as a guitar string or a human vocal cord, creates oscillating pressure changes in a medium which then propagates as a wave, much like a stone dropped into a pond would generate waves that spread out. Sound waves can be captures using a microphone that converts oscillations in the medium to electrical signals. These electrical signals can be digitized using an analog-to-digital converter and represented as well as stored as binary digits.

A. Spectral Analysis

All waves are characterized by repetitive patterns over time. Repeated patterns are quantified by their frequency, which measures the number of occurrences of the repeating event per unit of time, measured in Hertz (Hz). Higher frequency implies faster repetitions over time and vice versa. Humans can typically perceive audio signals from 20 Hz to 20 kHz range. Digitizing analog audio signals requires sampling them in time and converting it into a discrete sequence of quantized numerical values.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

The Nyquist sampling theorem states that a bandwidth limited continuous-time signal can be perfectly reconstructed from its samples if the waveform is sampled over twice as fast as its highest frequency component [8]. Since humans cannot generally hear signals above 20 kHz, sampling over 40 kHz is sufficient. The most common audio sample rate is 44.1 kHz. This is the standard for most consumer audio, used for formats like CDs [9].

A fundamental concept in signal processing is the Fourier Transform (FT). It stems from a fundamental mathematical fact that any signal can be represented as a sum of an infinite series of sinusoids. FT decomposes a signal into its constituent frequency components. In practical implementations, signals are digitized into discrete samples. Discrete Fourier Transform (DFT) converts a finite sequence of equally spaced samples of a signal into its frequency components.

In the equation below, x_n represents the n^{th} sample of a signal with a total of N samples. X_k represents the amplitude and phase of k^{th} frequency component of the signal. The power spectrum of a signal shows the intensity of all frequency components of the signal, measured by the square of the amplitude, $\|X_k\|^2/N$. The DFT is typically computed over a short window of samples using an efficient Fast Fourier Transform (FFT) algorithm. A typical window could have 1,024 samples. A signal sampled at 44.1 kHz, would have 44,100 samples per second of time which implies that a window of 1,024 samples represents 23 ms of the time domain signal. Overlapping smooth windows are used to remove spurious frequencies that can arise due to sudden truncation of the signal at the end of the window.

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{i2\pi}{N}kn}$$

Using a moving window and FFT, a spectrogram can be generated. A spectrogram is a 3D representation of a signal showings its constituent frequencies over time. Fig. 1 shows a time domain "chirp" signal where frequency is linearly increasing with time. The spectrogram clearly shows the frequency increasing linearly over time. The X-axis represents time, the Y-axis represents frequency and the colormap shows the relative intensity in any time-frequency bin. The spectrogram is generated by using a moving window over the full duration of the signal. Each window of samples is transformed using the FFT to generate frequency components and their relative power. Power is represented on a decibel (dB) logarithmic scale.

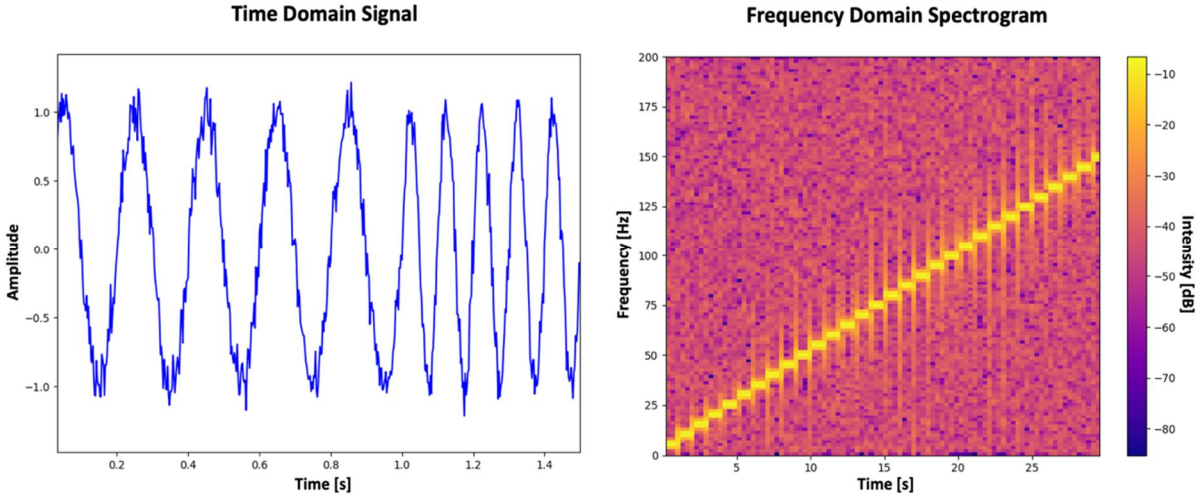


Fig. 1 Time domain "chirp" signal and its spectrogram

B. Mel Spectrograms

Humans do not perceive frequencies on a linear scale. For example, most people can easily tell the difference between 400 Hz versus 800 Hz sound but will struggle to distinguish a 10,400 Hz signal from a 10,800 Hz signal, although the frequency difference between the two sounds in either case is the same. The Mel (derived from melody) scale was developed to map a linear frequency scale in Hertz to a perceptual scale of pitches judged by human listeners to be equal in distance from one another. The reference point of 1,000 Hz is equivalent to 1,000 Mels. Based on this scale, humans perceive 3,120 Hz as 2,000 Mels and 9,000 Hz as 3,000 Mels. One can see the logarithmic compression of human dynamic range of audio frequency perception. A popular formula to convert frequency in Hz to Mels is as follows.

$$Mel = 2595 \log_{10}\left(1 + \frac{Hz}{700}\right)$$

The Mel spectrogram is a spectrogram representation where frequency is on a Mel scale. Once the power spectrum has been generated for a window of samples, a set of Mel filters is applied to gather the spectral energy in each of the Mel scale frequency bands. Each Mel filter is typically a triangular filter with a value of 1 at the center frequency and decreasing linearly to 0 till it reaches the center frequency on each adjacent side. Typically, a set of 40 such filters are used to extract spectral energy in 40 Mel frequency bins. A Discrete Cosine Transform (DCT) is then applied to the output of the Mel filter bank to remove spurious side-effects of the Mel filters. The final outputs are called Mel Frequency Cepstral Coefficients (MFCC). MFCCs are excellent features to use for audio analysis. Software packages such librosa [13], a popular python library, are available for audio processing and spectral analysis. Librosa can be used to read/write audio files, extract spectral features such as MFCC, general image plots, and perform other signal processing functions such as filtering, segmentation, and decomposition.

III.MACHINE LEARNING CLASSIFIER

The core engine to classify sounds is based on Machine Learning (ML). The ML model is trained using labeled audio data. A variety of free as well as commercial audio datasets are available online. For example, Google AudioSet [10] is collection of roughly 2.1 million audio clips, each 10 seconds long, extracted and labeled from YouTube. Similarly, the UrbanSound8K dataset [11] contains over 8,000 labeled sound files. The FSD project uses crowdsourcing of annotations of audio samples from Freesound organised using the AudioSet framework [12]. Data sets can also be generated manually by recording sounds and labelling them.

A. Training Data

We used the UrbanSound8K dataset for this work. This dataset contains over 8,000 labeled sound files each approximately 4 seconds long and of sounds encountered in a typical urban environment and labeled into 10 classes: air_conditioner, car_horn, children_playing, dog_bark, drilling, engine_idling, gun_shot, jackhammer, siren, and street_music.

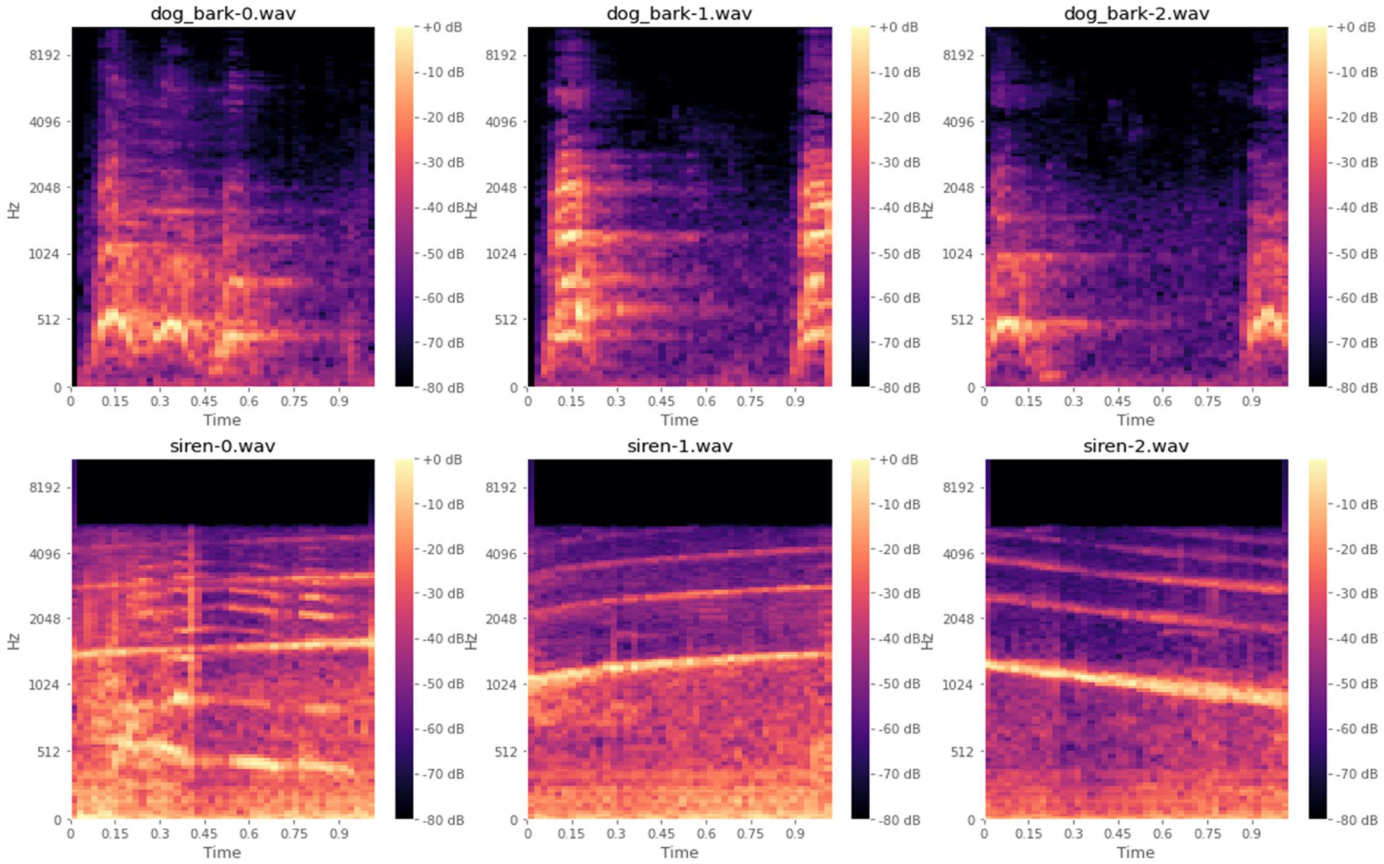


Fig. 2 Mel spectrograms of three different samples each of two different sounds (siren, dog_bark)

Fig. 2 illustrates the Mel spectrograms of three samples of two different sounds (siren, dog_bark). In the case of the siren, one can easily see the dominant frequency and some of its harmonics gradually increasing/decreasing over time. As such, the Mel spectrograms gives a powerful visual representation to the audio signal.

B. Machine Learning Model

Powerful ML models based on Convolutional Neural Networks (CNNs) have been developed for computer vision and image classification. CNNs have been used with spectrograms to successfully extract and classify sound with high accuracy and performance in [14]-[17]. Long Short-Term Memory (LSTM) is another type of neural network architecture that has been leveraged for sound classification [18]. LSTM networks are Recurrent Neural Networks (RNN) that use inputs over a period that may be related to map them to outputs. Deep neural networks that combine LSTM and CNN models have also been studied [19][20]. LSTM networks are efficient at learning temporal dependencies. They have been used in natural language processing applications where a sentence has a sequence of spoken words that are related and therefore must be considered when translating speech to text. LSTMs have been used for as phoneme classification [21], speech recognition [22] and speech synthesis [23]. LSTM network combined with CNN was also successfully used for video classification [24]. LSTM networks have been shown to be effective in classifying urban sounds in [25].

Unlike a standard feed forward neural network, LSTM have feedback connections. This allows it to process not only single data snapshots (e.g., images), but also entire sequences of snapshots of data (e.g., speech and video) [27]. This makes LSTM very applicable to problem such as handwriting recognition, natural language processing, and time series anomaly detection. At the heart of an LSTM network is an LSTM cell as shown in Fig. 3. It has an input gate, i_t, an output gate, o_t, and a forget gate, f_t. The subscript t indicates a time step. At any time step t, the cell processes the input vector, x_t, and computes various activation vectors as illustrated in Fig. 3. Matrices W and U are weights and biases that are learned during training, while σ refers to the standard sigmoid activation function. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell.

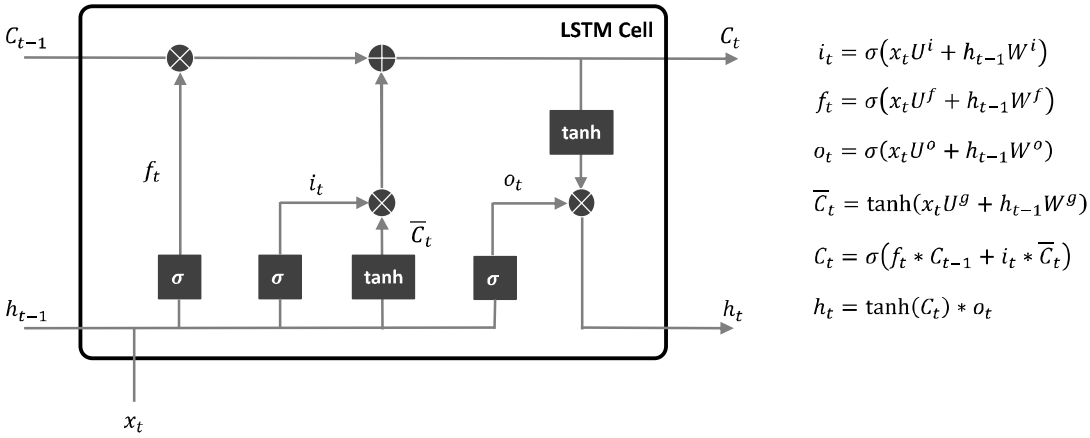


Fig. 3 LSTM cell showing the mapping of inputs to outputs

Our model for sound classification uses the LSTM architecture proposed in [25]. It is composed of two LSTM layers followed by dense layer with softmax activation function. While the LSTM produces a sequence, only the last value is propagated to the output layer. The first two layers contain 128 and 64 units, the last layer has 10 units, one per sound class. Dropout with a rate of 0.25 is applied to the output of the LSTM layers to reduce overfitting. The loss function used is categorical cross-entropy, minimized during model training using the Adam optimizer [26]. The input to the model was the magnitude of the Mel spectrogram with 128 bands, that covers a frequency range from 0 Hz to 22,050 Hz. The Mel spectrogram is computed at sample rate 44,100 Hz using a 1,024 sample window and a hop size of the same width. The LSTM based classifier was able to achieve an 83% prediction accuracy on the UrbanSound8K dataset.

IV. DIGITAL ASSISTANT DESIGN

The ML sound classification model can be incorporated into a mobile app such that it can detect and process ambient sounds, and notify a hearing-impaired user based on their preferences. The notification system could be a push notification on a mobile device

accompanied with a vibration alert and on-screen details. Vibration alerts could also be sent to a smart watch, or other visual forms of notification using LEDs or displays.

The user configures the system based on their preferences for what, when, and how they would like to be notified. Users can configure their preferences based on what sounds or categories of sounds they would like to be notified of (e.g., animal sounds, emergency, devices, vehicles, speech, music, etc.), and they can also choose how they would like to be notified (e.g., through a text message, vibration alert, or another methods). Furthermore, they can decide when they want to be notified (e.g., at work, at home, outdoors) and they can adjust when they want the system to be active (e.g., some users may want to use the system when they don't have a hearing aid on them).

Fig. 4 illustrates a flowchart of the core processing and analysis steps in the mobile app. The audio signal is captured from the device's microphone and filtered to remove background noise. Generally, this would imply that a valid audio signal above a configurable noise threshold has been received. The received signal is run through digital signal processing filters that improve fidelity and quality of the received sound, to assist downstream processing and detection. Once the sound has been isolated and filtered, the app checks to see if the user wants the specific sound to be amplified and sent to them. This may be useful if the user wants to hear the sound using a headset. If so, the received sound is amplified and sent to the audio device output (speaker or headset that is configured as the primary sound output device). A copy of the audio signal may also be stored digitally for future playback. The sound is then processed by the ML classifier which maps it to a most likely audio category (e.g., animal sounds, emergency, devices, vehicles, speech, music, etc.) and specific audio type (e.g., dog barking, ambulance siren, telephone ring, garbage truck, conversation, piano, etc.) that matches a pre-determined set of audio categories and types that the model has been trained to identify. Once audio category and type are determined, the app checks whether the user cares to be notified about the detected audio category and type based on preferences set before. If not, the system goes back to listening mode for the next audio event. If the user does want to be notified of the sound, the system first checks if the determined category was human speech. If so, it proceeds to runs a speech to text module which extracts text from the human voice signal and sends it to the notification system. If it is not human speech, the audio category and type determined by the ML classifier is summarized and sent to the notification system. For example, the system may have detected audio type "ambulance siren" of category "emergency". That information, along with the date, time duration and other relevant information may be sent to the notification system.

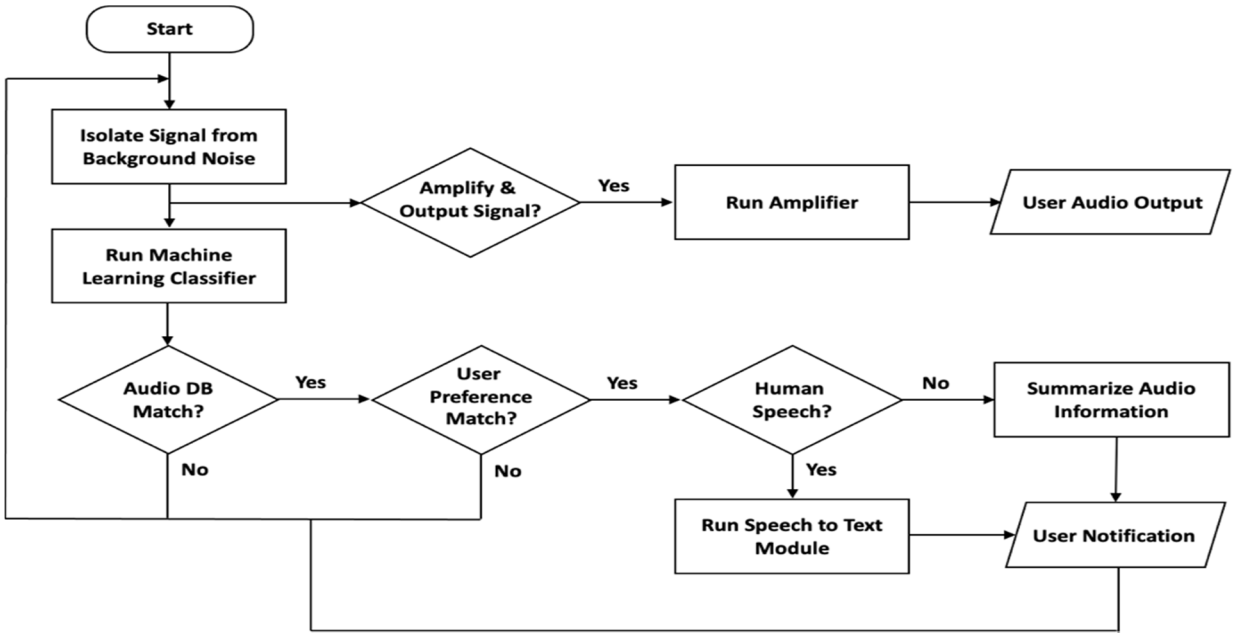


Fig. 4 Flowchart illustrating the core processing steps of the digital assistant app for the hearing-impaired

Fig. 5 illustrates the User Interface (UI) for the app. It shows a home screen, an audio amplification screen, a notification setting screen and notifications screen. The home screen shows a summary of the user's profile and provides four main functions of the application – speech to text, amplifier, notifications, and settings. The user can click to view or update their account information.

This may include their profile, account settings, as well as other important information that pertains to their condition so the system can cater to those needs.

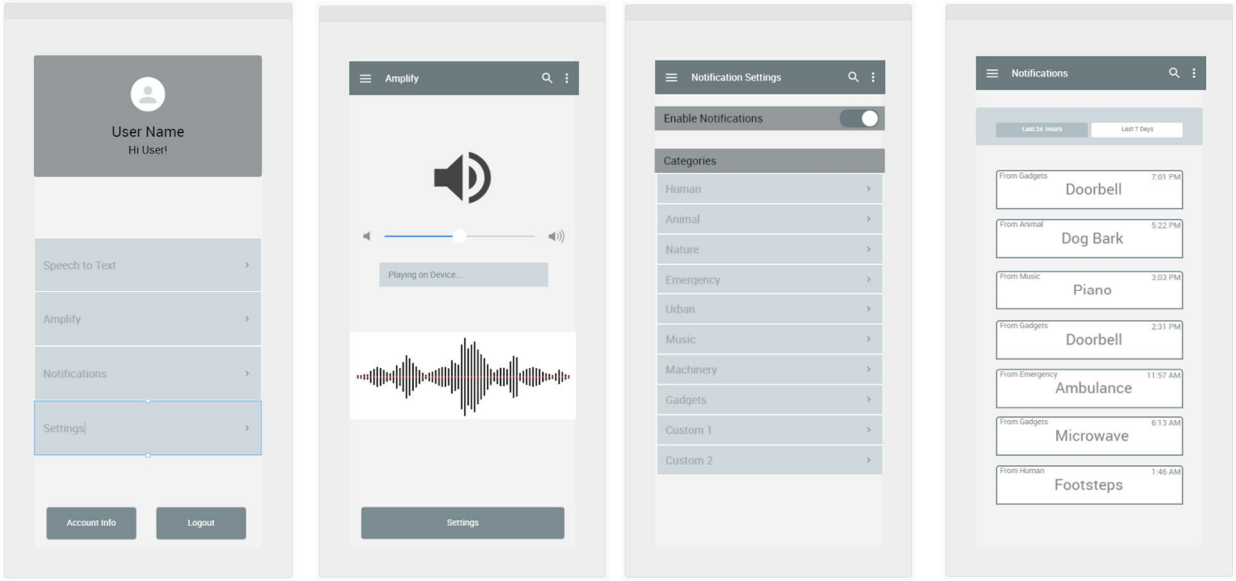


Fig. 5 User interface screens for the digital assistant app for the hearing-impaired

The amplifier feature allows the app to amplify the audio signal received over the microphone and play it back on the user's hearing device such as a paired Bluetooth headset. Notification settings allows the user to choose whether they want to allow notifications of detected sounds to be sent. If enabled, notifications will be sent to the user each time one of their preferred sounds is detected. If notifications are disabled, no alerts are sent to the user, although data may continue to be collected and stored in the background. The notification settings screen shows a sample of categories the user may prefer to be notified of. For example, if they wish to be notified of any animal sounds while taking a hike in the woods, they choose the "animal sounds" category for notifications. Finally, the notifications screen allows the user to view past notifications. It includes the category of the notification, specific audio type detected and time when the alert was generated.

V. CONCLUSIONS

In this paper we have described a digital assistant for the hearing-impaired. The assistant detects audio signals from the ambient environment of a hearing-impaired person. The audio signals are analyzed by a machine learning model that can classify audio signals into audio categories and audio types and notify the user leveraging a mobile phone notification. Machine learning models to detect and classify sounds can be effectively implemented using spectral features and deep learning neural networks. On the UrbanSounds8K dataset, an LSTM based network was shown to achieve 83% classification accuracy. The model can be implemented in the digital assistant app and is periodically re-trained and updated. The user can configure notification preferences that control what type of sounds should generate alerts. Optional system features include audio amplification as well as a speech to text module for human voice.

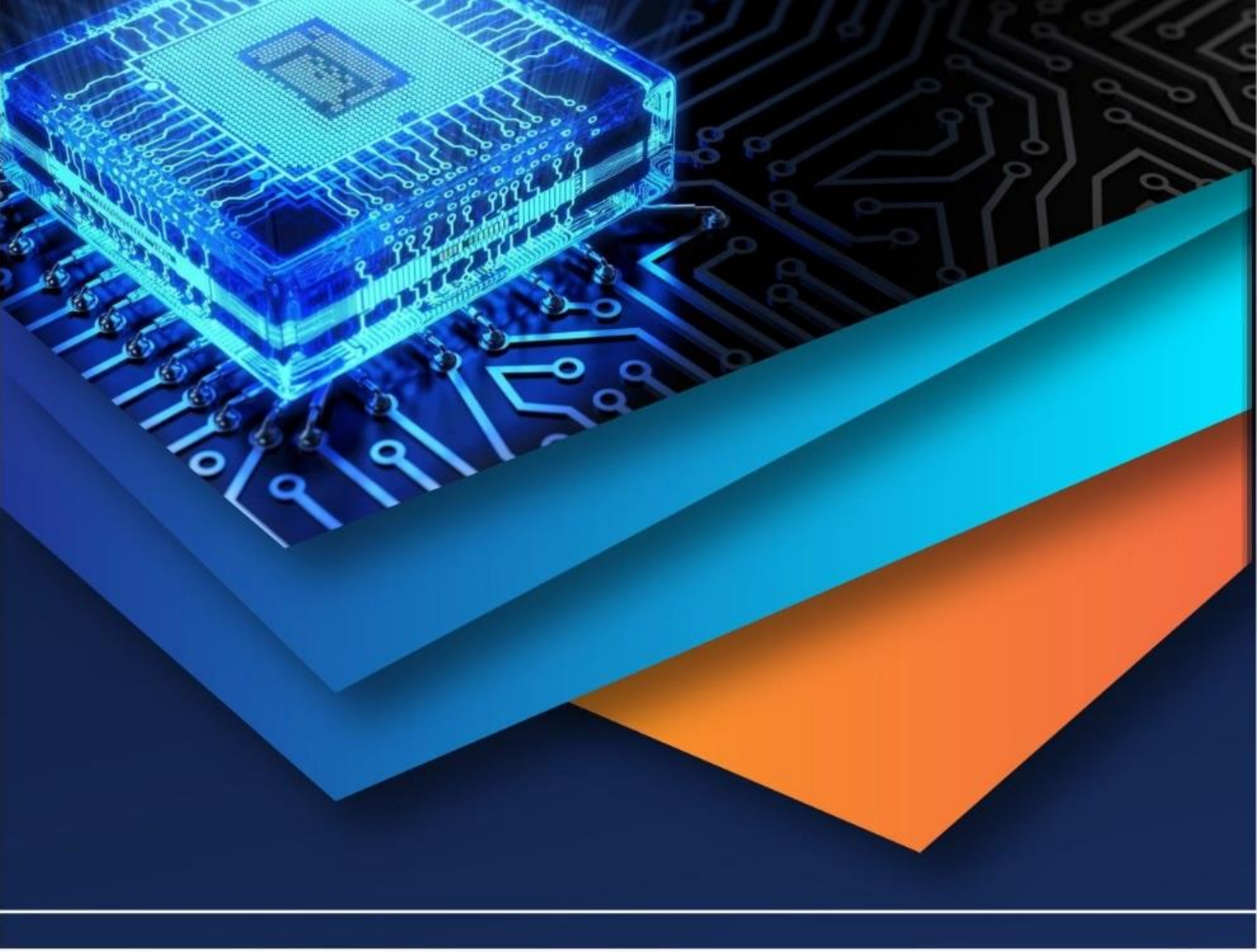
VI. ACKNOWLEDGMENT

The authors would like to thank Massachusetts Institute of Technology, Beaver Works program. We would like to acknowledge the help, inspiration and guidance from the faculty teaching the *CogWorks: Build Your Own Cognitive Assistant* and *Designing for Assistive Technology* courses.

REFERENCES

[1] World Health Organization, "Deafness and hearing loss," 2019.
[2] Hearing Loss Association of America (HLAA), "Types, Causes and Treatments," 2019.
[3] Department of Health, London, UK, "Mental health and deafness - Towards equity and access: Best practice guidance," 2005.
[4] National Institute on Deafness and Other Communication Disorders, "Assistive Devices for People with Hearing, Voice, Speech, or Language," 2018.
[5] T. Rains, "How much do hearing aids cost?", 2019.
[6] Gallaudet University and Clerc Center, "Assistive Technologies for Individuals Who are Deaf or Hard of Hearing," 2019.
[7] Apple, "Use Live Listen with Made for iPhone hearing aids," 2019.

[8] R. Oppenheim, A. Schafer, Discrete Time Signal Processing, 2014.
[9] K. Steiglitz, A Digital Signal Processing Primer: with Applications to Digital Audio and Computer Music, 2020.
[10] J. Gemmeke, "Audio Set: An ontology and human-labeled dataset for audio events," 2017.
[11] J. Salamon, "A Dataset and Taxonomy for Urban Sound Research," 2014.
[12] E. Fonseca, "Freesound Datasets: A Platform for the Creation of Open Audio Datasets," 2019.
[13] B. McFee, C. Raffel, D. Liang, D. Ellis, M. McVicar, E. Battenberg, and O. Nieto, "librosa: Audio and music signal analysis in python." In Proceedings of the 14th Python in Science Conference, p. 18-25, 2015.
[14] K. J. Piczak, "Environmental sound classification with convolutional neural networks," 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP). IEEE, p. 1–6, 2015.
[15] J. Salamon and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification," IEEE Signal Processing Letters, vol. 24, no. 3, p. 279–283, 2017.
[16] V. Boddapati, A. Petef, J. Rasmusson, and L. Lundberg, "Classifying environmental sounds using image recognition networks," Procedia computer science, vol. 112, 2017, p. 2048–2056, 2017.
[17] B. Zhu, K. Xu, D. Wang, L. Zhang, B. Li, and Y. Peng, "Environmental sound classification based on multi-temporal resolution convolutional neural network combining with multi-level features," in Pacific Rim Conference on Multimedia. Springer, 2018, p. 528–537, 2018.
[18] Y. Wang, L. Neves, and F. Metze, "Audio-based multimedia event detection using deep recurrent neural networks," in 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), p. 2742–2746, 2016.
[19] S. H. Bae, I. Choi, and N. S. Kim, "Acoustic scene classification using parallel combination of lstm and cnn," Proceedings of the Detection and Classification of Acoustic Scenes and Events 2016 Workshop (DCASE2016), p. 11–15, 2016.
[20] J. Sang, S. Park, and J. Lee, "Convolutional recurrent neural networks for urban sound classification using raw waveforms," in 2018 26th European Signal Processing Conference (EUSIPCO). IEEE, p. 2444–2448, 2018.
[21] A. Graves, S. Fernandez, and J. Schmidhuber, "Bidirectional LSTM networks for improved phoneme classification and recognition," in International Conference on Artificial Neural Networks. Springer, p. 799–804, 2005.
[22] A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in 2013 IEEE international conference on acoustics, speech and signal processing. IEEE, p. 6645–6649, 2013.
[23] Y. Fan, Y. Qian, F.-L. Xie, and F. K. Soong, "TTS synthesis with bidirectional LSTM based recurrent neural networks," in Fifteenth Annual Conference of the International Speech Communication Association, 2014.
[24] J. Yue-Hei Ng, M. Hausknecht, S. Vijayanarasimhan, O. Vinyals, R. Monga, and G. Toderici, "Beyond short snippets: Deep networks for video classification," in Proceedings of the IEEE conference on computer vision and pattern recognition, p. 4694–4702, 2015.
[25] I. Lezhenin, N. Bogach, and E. Pyshkin, "Urban Sound Classification using Long Short-Term Memory Neural Network," in Proceedings of the Federated Conference on Computer Science and Information Systems, p. 57–60, 2019.
[26] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," 3rd International Conference for Learning Representations, San Diego, 2015.
[27] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, p. 1735–1780, 1997.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY