iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** II **Month of publication:** February 2020

DOI: http://doi.org/10.22214/ijraset.2020.2108

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: 7.177
Volume 8 Issue II Feb 2020- Available at www.ijraset.com

An Novel Approach of CNN - Machine Learning Model integrated with Android for Women's Safety (SAS)

Helly Patel[1], Parth Lathiya[2], Bhoomit Patel[3], Hirpara Nidhi[4], Divya Ebenezer[5]
[1, 2, 3, 4]*BE- 4 th Year Students,* [5] *Assistant Professor Department of Computer Science and Engineering, Babaria Institute of Technology, Varnama, Gujarat, India*

Abstract: ***The most threatening issue of today's era is related to women's safety. Despite some government laws regarding women's safety, women still feel afraid to travel alone at any hour of the day. The cases of females being harassed, molested, assaulted or violated are increasing significantly. It is estimated that 35% of the females (women, girls', babies) have experienced physical and/or sexual violence at some point in their lives whether it be at work, home or on streets. Today, though women are excelling in almost every field, they always have a concern regarding their safety at any point of hour. This paper represents our proposed system (mobile application) named as "SAS(Safe Society)".The application aims to provide a sense of safety and security amongst all the women and girls' around. The application is corely based on voice recognition which will help creating a safe and secure society by detecting the distress sound(signal) of victims and then after it will give details of victims location with the help of GPS location tracker to the emergency contacts mentioned and to nearby police station with the help of using GSM(Global System for Mobile communication). This paper includes both the design and implementation of SAS. This application has an advantage that it will play a siren sound along with a "In danger" message displaying on the contact's mobile. Keywords: Android, GPS, registered emergency contacts, voice recognition of distress sound using Machine Learning, Melspectogram***

I. INTRODUCTION

The issue of women's safety has been ongoing since ages and still inspite of being curbed, it has accelerated with the passage of years. At an unknown place, it seems unsafe for women to travel lonely at midnight. At this time the mobile phone can be the best friend of the user. There are ample amount of advantages of using the latest mobile technology that can provide a helping hand in declining the rates of violence, rapes, molestation against women. They are easily accessible, crowdsourcing and affordable scalability. Women is getting kidnapped at every 45 minutes, raped at every 50 minutes, 15 to 20 dowry deaths every day[1]. New Delhi, the capital city of our nation India in year 2012 provoked the attention of the people of the nation as well as of the entire world. A fatal assault took place on 16th December 2012 at a place named Munirka, a neighborhood situated in south Delhi. This fatal assault was a rape incident which shoke the entire nation. A 23-year-old woman was not only molested but also hit mercilessly by a gang at 9:30 PM with a male counterpart. At that time she was traveling on a private bus with her friend.[2]. Many existing security systems which includes varoius types of mobile apps, shock generating jackets, GPS and GSM based location monitoring, manually triggered SOS switch available in the market failed to find a full security solution and had many drawbacks[3].

The atrocities against the women can be now brought to an end with the help of an application called SAS. This declines the risk of violence happening to women worldwide. The application also provides assistance when there arises a requirement of it which in turn helps us to identify the location of the victim's in danger. Our application differs from the other developed and existing applications since it has some diverse key features which diverse. The working is as follows:

1) Firstly, the application will recognize whether the voice is a distressed voice (i.e SCREAM) of the victim or not. This will be done with the help of our trained datasets using the concept of machine learning.
2) Then, if it recognizes it as distress(SCREAM), it will send a message displaying "HELP" along with the live location to the nearby police station with the help of GPS.
3) The above message is also sent to the registered emergency contacts of the user.
4) Above all that, there is also integrated a siren sound with this message. The message upon its arrival will play a siren sound on the mobile phone.

This message also includes the latitude and longitude of the victim. Android, a world-wide known software that has striked the era of mobile modernization, is a comprehensive platform to run a mobile phone. It is powered and based on Linux Kernel. This is the first open platform that has acquired much popularity by it's robustness. The current system is developed on the core concept of the android comprehensive platform. Android is an open-source os for mobile device for now a days. Google team has developed this platform that permits writing the code in Java language with much ease[4].

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 8 Issue II Feb 2020- Available at www.ijraset.com

II. EXISTING SYSTEMS

Among the worst countries in crime, India has an disgusting track record in all forms of sexual exploitation. In public transports, on streets, and even offices, Indian women are in a constant state of alertness. The currently existing women safety applications prevelant in the market have a push button to send an SOS message indicating danger that sometimes does not provide an helping hand in some critical situations. Large number of applications are available for women safety that alerts the victim's family, friends or the police. Some of the existing applications are discussed as follows:

1) *SafetiPin:* Complete Safety App: This developed application has proven to be really helpful for the women. The women can beforehand know whether the unknown place they are traveling to is safe or not. To make safer decisions regarding their mobility MySafetipin may be a Personal Safety app for women that can ensure their safety and security, based on the calculated safety score of an area[5].
2) *TellTail[6]:* the application allows the user to be tracked through the GPS on their phone or the vehicle. The user can send instant alerts to a group of contacts as well. The application provides location even if no GPS on your cell phone as location can be monitored using the vehicle's GPS.
3) *Fightback:* Mahindra faction developed and proposed this application for assuring the security of people in general. This application sends a message through E-mail, GPS, SMS, and GPRS to your nearly family members, friends or the emergency contacts indicating the message that "user is in trouble". This application works on that mobile which supports Android Java Programming. With the help of this application, the victim location can also be sent through SMS using maps, which is an exciting feature of this application [7].
4) *On Watch:*This, a personal security application, in an distress situation, helps the user to send an alert through this application to their friends and family. Through this application one can send "Time Based Alert" if one doesn't reached on a given and specified time on the specified locality then within one hour it will alert to your family and friends. This app has 6 custom-made alert methods that allow you the flexibility to inform your friends, Police Station or a combination of all two if you need help[8]
5) *"VithU"[9]* is an emergency application that, send alert on click power button of your Smartphone 2 times consecutively, every 2 minutes to your contacts that you provide into the app as the designated receivers or guardians.

III. PROPOSED SYSTEM

The purpose of this is to design a system based on Machine Learning which will aid towards a better approach to defining women safety.The main aim is to develop a model that is helpful to feel women safe and more empowered.This system helps women when she is in dangerous situation such that travelling alone at night etc. In this system, here is the complete solution for voice command recognition and emergency detection based on audio signals entirely integrated in a low-consuming embedded platform. The system combines an active operation mode where distress calls are captured. The registration is required for any person who wants to use this application. User would be able to use all the facilities provided by the system for providing more safety and security to them. Facilities include sending the emergency(SOS) signal to the nearest police station and registered emergency contacts that has been provided by user. The proposed system also sends the emergency signal to the people nearby the victim in specific range along with live location. The live location that is provided by the system can help the police to reach toward the victim and there is siren also that is continuously active and can be turned off by the authority only. Audio Recognition and its transformation to image must be done to get successful result.

Flowchart

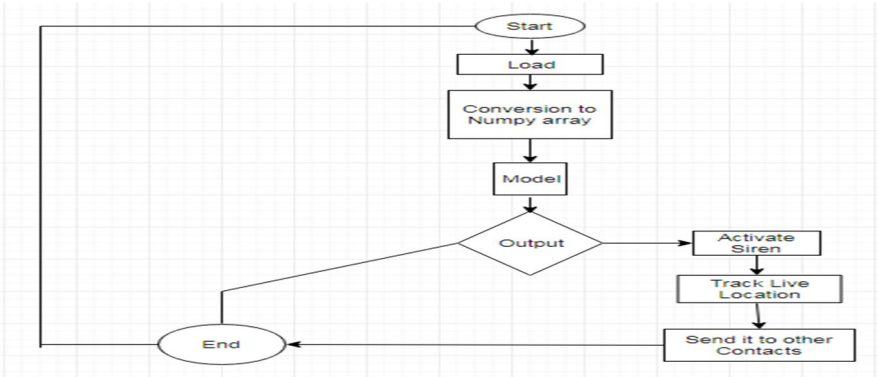


Figure.1

©IJRASET: All Rights are Reserved

1) *Input:* Firstly the user's distress voice will serve as an input to the application.
2) *Conversion to Num Pyarray:* All the inputs will be loaded and then be converted to a NumPy array for voice recognition purposes. This will also serve as an input to the neural network system.
3) *Model:* After conversion to NumPy array, this array will be given as an input to the trained model(of distress sounds/signals) for recognition of distress signal, so that the app can be triggered. Here, in this model, this input will be stored for enhancing the efficiency of the trained model. We will be using deep learning techniques for preparing the trained model. It will contain ample and various amounts of distress sound. Various matching algorithms will also be applied here.
4) *Output:* After distress sound is recognized, the siren will be activated with a "help" message on the victim's registered emergency contact number. Along with the above message, the victims live location will also be sent to the registered emergency contact number as well as to the nearby police stations for further actions.

The important modules are:

A. Database

SCREAMS are stored in the database(dataset). User have to provide the numbers of close ones from whom they aspect to acquire help. The voice have to be saved for further processing of voice recognition. Contacts and voice are saved in the database for getting more efficient results. This database is saved in memory of mobile. It is SQLite database. The range of the number of emergency contacts which has been provided by user is between 2-5.

B. Voice Recognition

This module is used to identify the keyword from the input voice from the user side.That voice spoken by the victim is compared with the registered different kinds of screams.

Usually, it is easy to classify data in simple formats like CSV files and images. But what if we need to pass our machine learning judgment on a source of information like audio signals. Well, this can be done by converting audio signals into the form of images that are called spectrograms. These are the pictorial representation of the audio sound with the help of various features like sampling rate and frequency with amplitude. So, a Mel spectrogram will convert the standard time vs amplitude to time vs frequency vs amplitude formats on the x-axis, y-axis and color axis respectively.

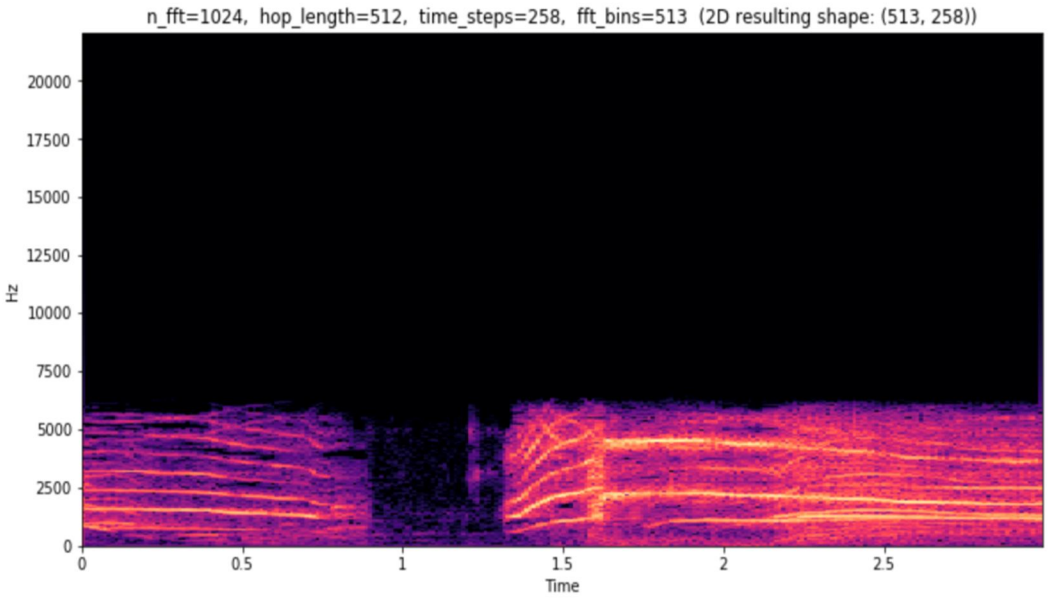


Figure.2 (Melspectrogram)

This is matched with the help of a converted image for better accuracy and efficiency. If both converted images are matched then the message will be sent. For this module, we will train the datasets with various types of distress voice(scream).

©IJRASET: All Rights are Reserved

C. Location Tracking and Address Finding module

For tracking the location and find the address the GPS of mobile device must be on.the longitude and latitude of the current location of the user can be recognize by this module. Using longitude and latitude location which has been identified by this model is given to the emergency contacts of the victims or to the nearest police station via message followed by actual address.

D. Message Sending Module

The longitude and latitude coordinates will be received by the GPS Application Program Interface (API).The emergency message is sent to registered contact numbers along with the longitude and latitude and an exact address(live location) of the user. When a message is sent then the notification is given in the form of a loud siren sound.

IV. WORKING OF PROPOSED SYSTEM

There are no applications for the person's safety if the person is in danger. When Some person who is in danger has somehow get chance to call for the help, he/she have to explain the exact location of place and generally there is not that much of time in that situation. drawbacks in the existing systems are discussed below:

1) The person who is in danger can't explain and show the position and situation.
2) The person does not know the necessary details.

In the proposed system there is no need of any actions , it will automatically identifies the situation of danger and gives alert signal to emergency contact provided by user along with the proper address of current location.with this application you will never fill alone!!!!

A. System Architecture

This work develops a women's safety system which provides the current location details of the women in danger using concepts of GPS, CNN, and Melspectogram to their family members. The voice recognition module comprises various steps that include the conversion of audio clips into spectrum images and then feeding those images into a layer of CNN to acquire the desired results.
The working of the proposed system is stated and explained below in brief:

B. Data Set

The dataset is used for training and testing purposes. The dataset either gives positive or negative results. Based on this, two types of the dataset are developed for our system:

1) *Positive Dataset:* This dataset is composed of around 2000 clips of various SCREAM voices of women. The audio clips are of equal length(i.e 3 sec).
2) *Negative Dataset:* This dataset is composed of around 2000 clips of various background voices like the chattering of people, the voice of heavy rains, the voice of swaying trees, etc. of equal length(i.e 3 sec).

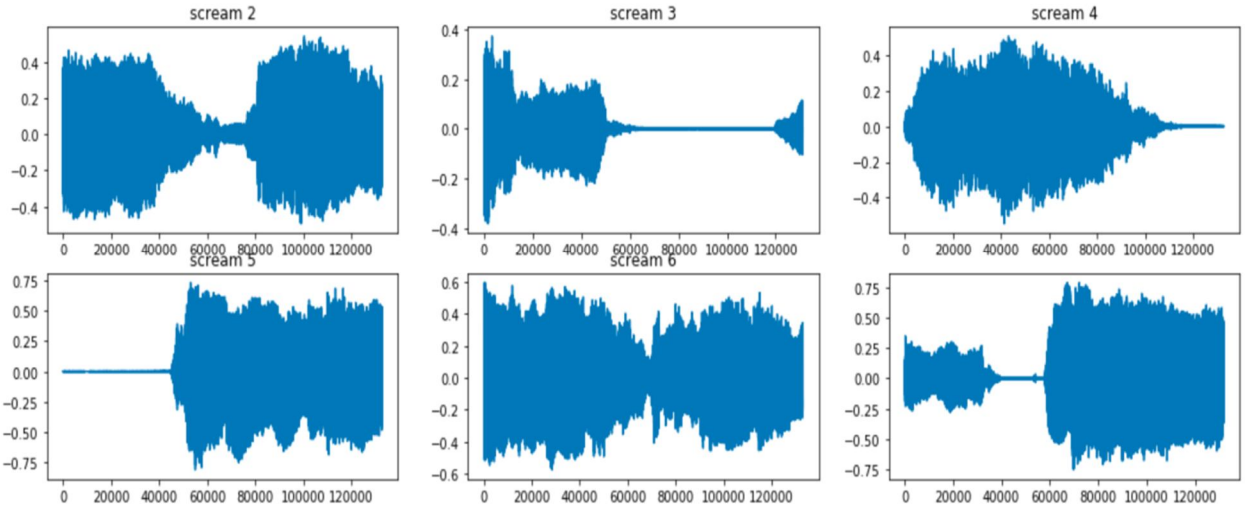


Figure .3(Audio Clips Of Datasets)

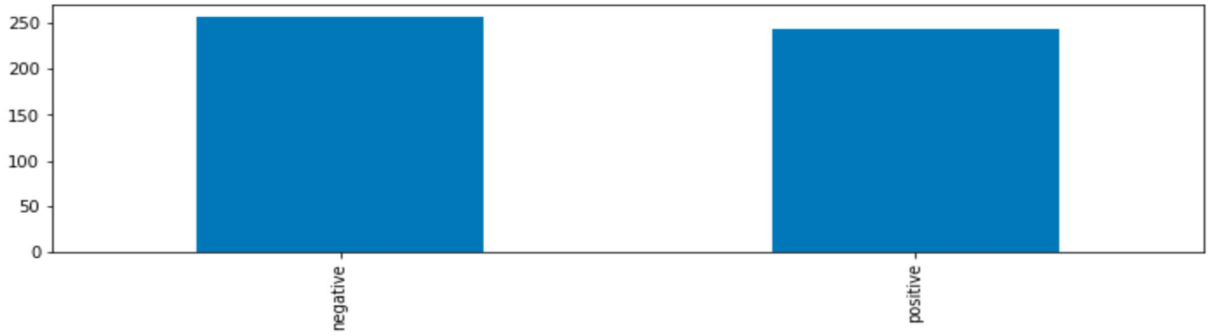
©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 8 Issue II Feb 2020- Available at www.ijraset.com

Based on the voice recognition of the victim, it will give positive or negative results as shown in figure 4.

```
train_df['category'].value_counts().plot.bar()
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x157469690>
```




```
validate_df['category'].value_counts().plot.bar()
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x157456f90>
```

Figure.4

C. Conversion to Mel-Spectograms

Each audio file was converted into a spectogram which is a visual representation of the spectrum of frequencies over time. The spectrogram is the representation of the squared magnitude of the Fourier transform (STFT) of the sound signal. This output graph alters the audio frequencies to a form that is perceived easily by humans[10].

The vital components used in these new developments are window length that implies the frame of time to carry out Fourier Transform on with hop length that is the number of instances between its next frames.The most common window size for such a transformation to occur is 2048 that results to about 10ms, the smallest and easiest acceptable period that a human ear can perceive and differentiate between. The Librosa library makes it possible to create such spectrograms easily on the flow[11].

These spectrograms are a way to represent audio across various time and frequency dimensions.

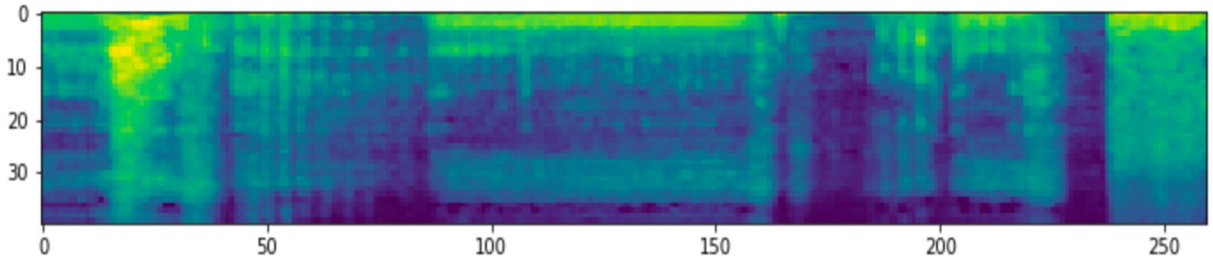


Figure.4[11](Converted Image)

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 8 Issue II Feb 2020- Available at www.ijraset.com

D. Feature Extraction Module(CNN)

CNN(Convolution Neural Network) encodes the information from the voice-features as well as the image features. This data is sent to a convolution layer with a kernel-size of 1 which also acts as a fully connected layer[12].

With the help of CNN module, the input image is analyzed and processed while in the phase of convolution and attributes a label on the later phase.

A standard most understandable architecture can be viewed in figure 3 below. Firstly, image sample is passed into the neural network, which called be called as Input Image. After which, this image passes through a large number of steps, i.e., the convolutional part of the model. Lastly,the model runs it's prediction and outputs a digit on the input image. The most critical and complicated component in the model is the convolutional layer of this model. It aids the model by compressing the image to aid faster computations and modify it to stabilize the weights and enhance its generalization of inputs. The convolutional process gets rid of all the irrelevant details that are not needed and only considers the important feature weights that actually determine the correct output of the input image[13].

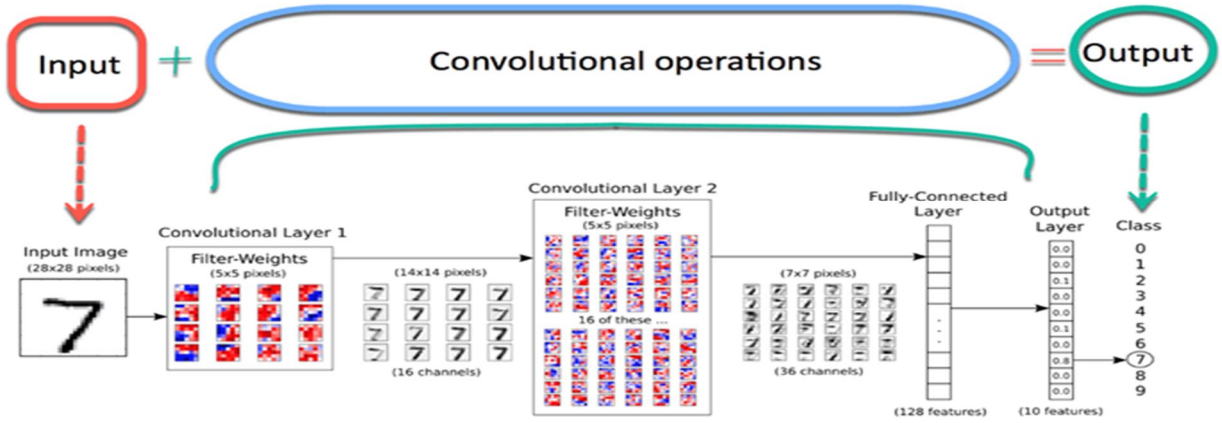


Figure.4[13]

The building block of the convolution neural network contains a pair of hidden plies as a convolution ply and a pooling plyformate. The input contains many number of localized features organized as a number of feature maps for the further use of this input. The size of feature and details maps gets smaller at upper layers as more convolution and pooling operations are try for operations. Usually one or more fully connected hidden layers are added on top of the final convolution neural network layer in order to combine the features and details across all frequency bands before feeding to the output layer of the model[14].

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 251, 38, 32)	896
batch_normalization_5 (Batch	(None, 251, 38, 32)	128
max_pooling2d_4 (MaxPooling2	(None, 125, 19, 32)	0
dropout_5 (Dropout)	(None, 125, 19, 32)	0
conv2d_5 (Conv2D)	(None, 123, 17, 64)	18496
batch_normalization_6 (Batch	(None, 123, 17, 64)	256
max_pooling2d_5 (MaxPooling2	(None, 61, 8, 64)	0
dropout_6 (Dropout)	(None, 61, 8, 64)	0
conv2d_6 (Conv2D)	(None, 59, 6, 128)	73856
batch_normalization_7 (Batch	(None, 59, 6, 128)	512
max_pooling2d_6 (MaxPooling2	(None, 29, 3, 128)	0
dropout_7 (Dropout)	(None, 29, 3, 128)	0
flatten_2 (Flatten)	(None, 11136)	0
dense_3 (Dense)	(None, 512)	5702144
batch_normalization_8 (Batch	(None, 512)	2048
dropout_8 (Dropout)	(None, 512)	0
dense_4 (Dense)	(None, 2)	1026

Total params: 5,799,362
Trainable params: 5,797,890

Figure.5 Summary Of CNN(Convolution Neural Network)

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 8 Issue II Feb 2020- Available at www.ijraset.com

The given model from the figure 05 takes in the created input spectrograms and pass it via both the RNN and CNN layers parallelly combining their outputs and then passing it from the dense layer alongside the SoftMax activation function to carry out its classification. These CNN block of the network comprises of 2D convolution layer backup up by a 2D Max pooling layer.
Finally, the Batch normalizations are performed and at last the 2D Max Pooling is carried out that reduces spatial dimensions of the image which will help prevent model from overfitting problem. This cycle of operations (2D Convolution > Batch Normalization > 2D Max Pooling)is carried out three times.

E. GPS Module

The Global Positioning System is known as a location tracker. It tracks the current location in the form of both the longitude as well as the latitude. In case if GPS is disabled then the system will only send the longitude and latitude through SMS. So, the Internet is mandatory.

V. RESULT

The main purpose of the work is to provide safety and security to women in dangerous situations.The developed neural network model will be able to distinctly recognize and classify a situation of a victim and her natural environment with the help of the voice recognition system that is specifically trained to detect women's voice and carry out a judgement accordingly, thereby helping them in a hazardous situation by calling in for help assigned to her.
In the backend, These final training images are then fed to our convolutional neural network for learning and adjusting the weights. Since our dataset only contains two categorical outputs, the last layer is to be used as a softmax layer and give either '0' or '1' as the prediction as shown in figure 6.

Figure.6(Category Stats)

Each time the model uses the 'categorical_crossentropy' loss and 'rms' optimiser. Crossentropy since the labels are one hot encoded, and not sparsh_crossentropy. The RMSprop optimizer is like gradient descent algorithm with momentum. The RMSprop optimizer will be able to supress the number of oscillations performed in a vertical fashion.

```
Epoch 18/25
32/32 [==============================] - 8s 256ms/step - loss: 0.1366 - accuracy: 0.9542 - val_loss: 2.1741e-04 - val
_accuracy: 0.9626
Epoch 19/25
32/32 [==============================] - 8s 259ms/step - loss: 0.1752 - accuracy: 0.9490 - val_loss: 0.0051 - val_acc
uracy: 0.9917
Epoch 20/25
32/32 [==============================] - 8s 259ms/step - loss: 0.2039 - accuracy: 0.9427 - val_loss: 0.6591 - val_acc
uracy: 0.9813
Epoch 21/25
32/32 [==============================] - 8s 255ms/step - loss: 0.1654 - accuracy: 0.9329 - val_loss: 0.1736 - val_acc
uracy: 0.9346
Epoch 22/25
32/32 [==============================] - 8s 258ms/step - loss: 0.1243 - accuracy: 0.9625 - val_loss: 0.0751 - val_acc
uracy: 0.9720
Epoch 23/25
32/32 [==============================] - 8s 252ms/step - loss: 0.1878 - accuracy: 0.9597 - val_loss: 0.0074 - val_acc
uracy: 0.9346
Epoch 24/25
32/32 [==============================] - 8s 261ms/step - loss: 0.1308 - accuracy: 0.9490 - val_loss: 0.0110 - val_acc
uracy: 1.0000
Epoch 25/25
32/32 [==============================] - 8s 255ms/step - loss: 0.1378 - accuracy: 0.9512 - val_loss: 0.0019 - val_acc
uracy: 0.9813
```

Figure.7(Accuracy Results)

©IJRASET: All Rights are Reserved

The accuracy of our proposed system is 93%.

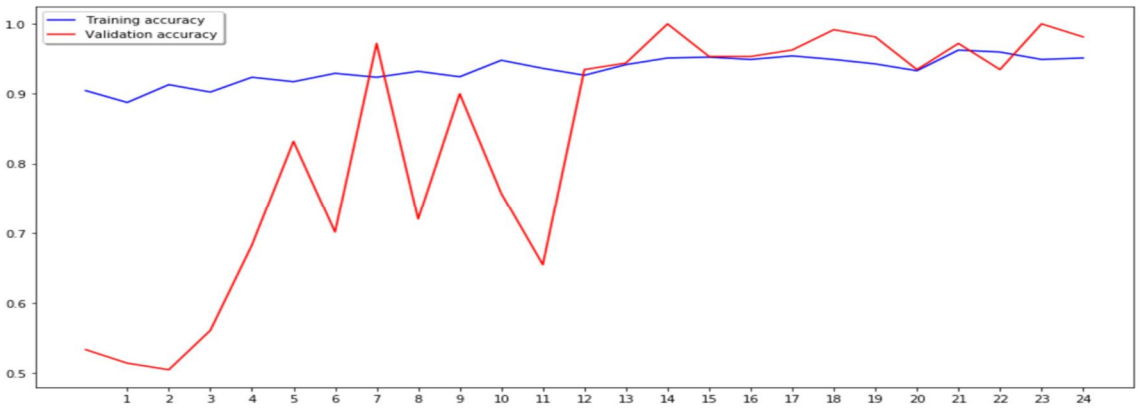


Figure.8 (Validation Loss)

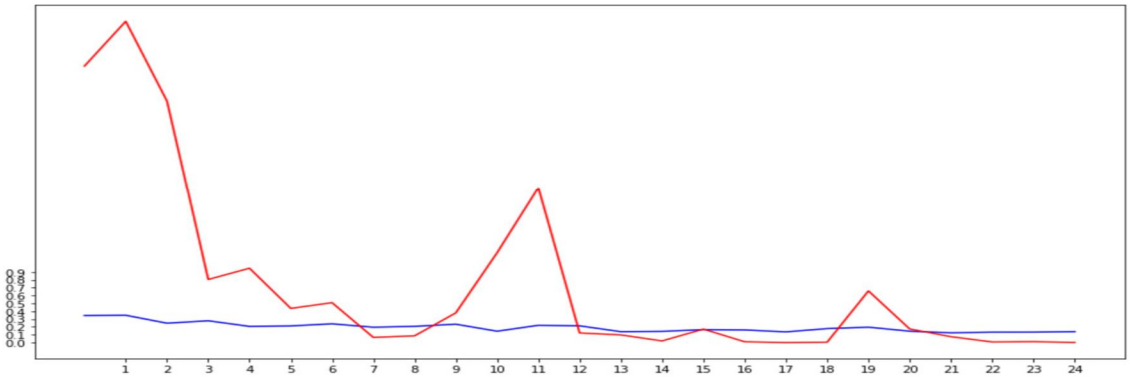


Figure.9 (Training Loss)

After identifying the danger situation through our model the live location is sent to the given emergency Contact numbers.

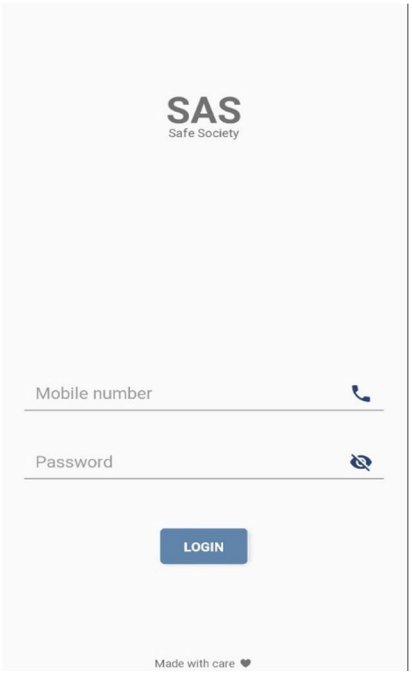


Figure. 10

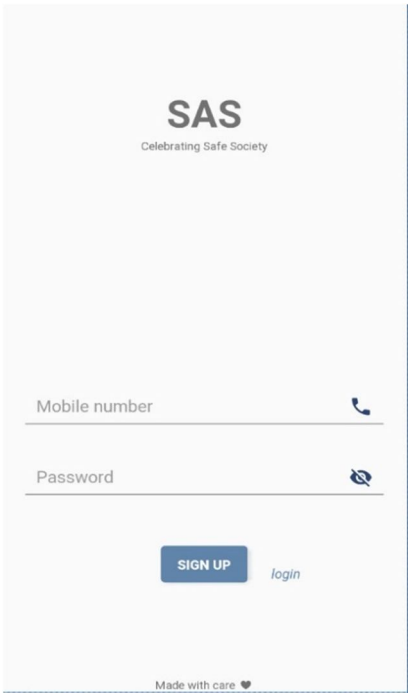


Figure. 11

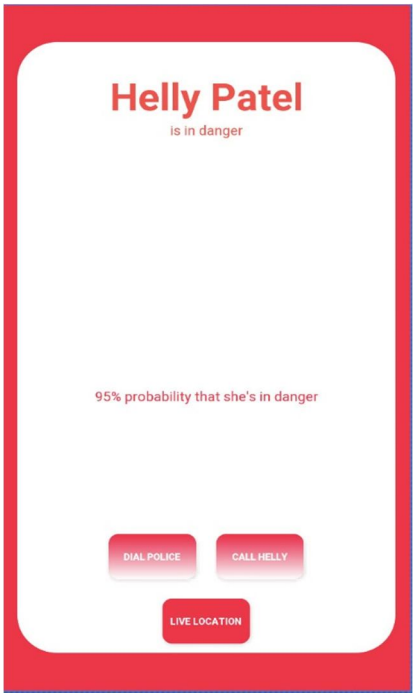


Figure. 12

©IJRASET: All Rights are Reserved

Figure.13

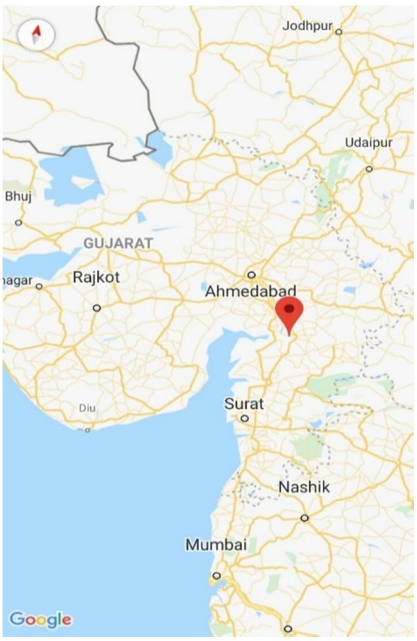


Figure.14

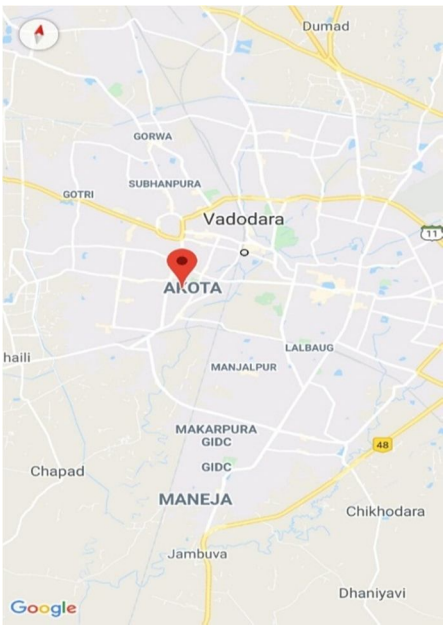


Figure.15

The above figures demonstrate the application GUI made on the android platform. Figure 10 shows the interface that pops up when the user first time opens the application after installation for login purpose followed by figure 11 that depicts the signup option. Then, the main interface is shown in figure 12that demonstrates the possibility of the victim in danger. It also shows three main options: daily police, Chelly, and live location. When the user clicks on a live location, a map is shown as depicted in figures 13,14,15. The map has the facility to zoom in and know the exact location of the dreadful situation that might be taking place.

VI. FUTURE WORK AND ENHANCEMENTS

There's a lot of possibilities and combinations that can be applied to this model so that it can become more efficient in space and time. With some more work, it could even predict the exact scenario of the situation just by analyzing the audio and trained labels. It has tremendous possibility especially with the use of the recurrent neural networks where it keeps getting intelligent and intelligent and starts picking up the best hyperparameters it needs to detect a scene.If there're any future security issues, they'll be resolved with the use of technologies evolving in the future. Furthermore, as the technology advances, the current applications can also adapt to the changes to the desired environment for user's better experience and model's performance.

VII. CONCLUSION

This mobile application is very helpful for the woman. These systems based on Machine learning can be used to better understanding the dangerous situations better than any currently existing technology. With the implementation and use of this technology, it is possible to provide the level of freedom, care and help towards the women and feel comfortable even at nights, letting them work as equals to Men and enjoy the freedom that they're supposed to have since the beginning.Not only does this help individuals possessing this technology but it'll also help collect data from users and classify threats at different locations, classifying the threat regions from the database and hopefully reduce the number of criminal activities against women to a greater extent. This system will provide a chance to live a stress free life by assuring them their security.

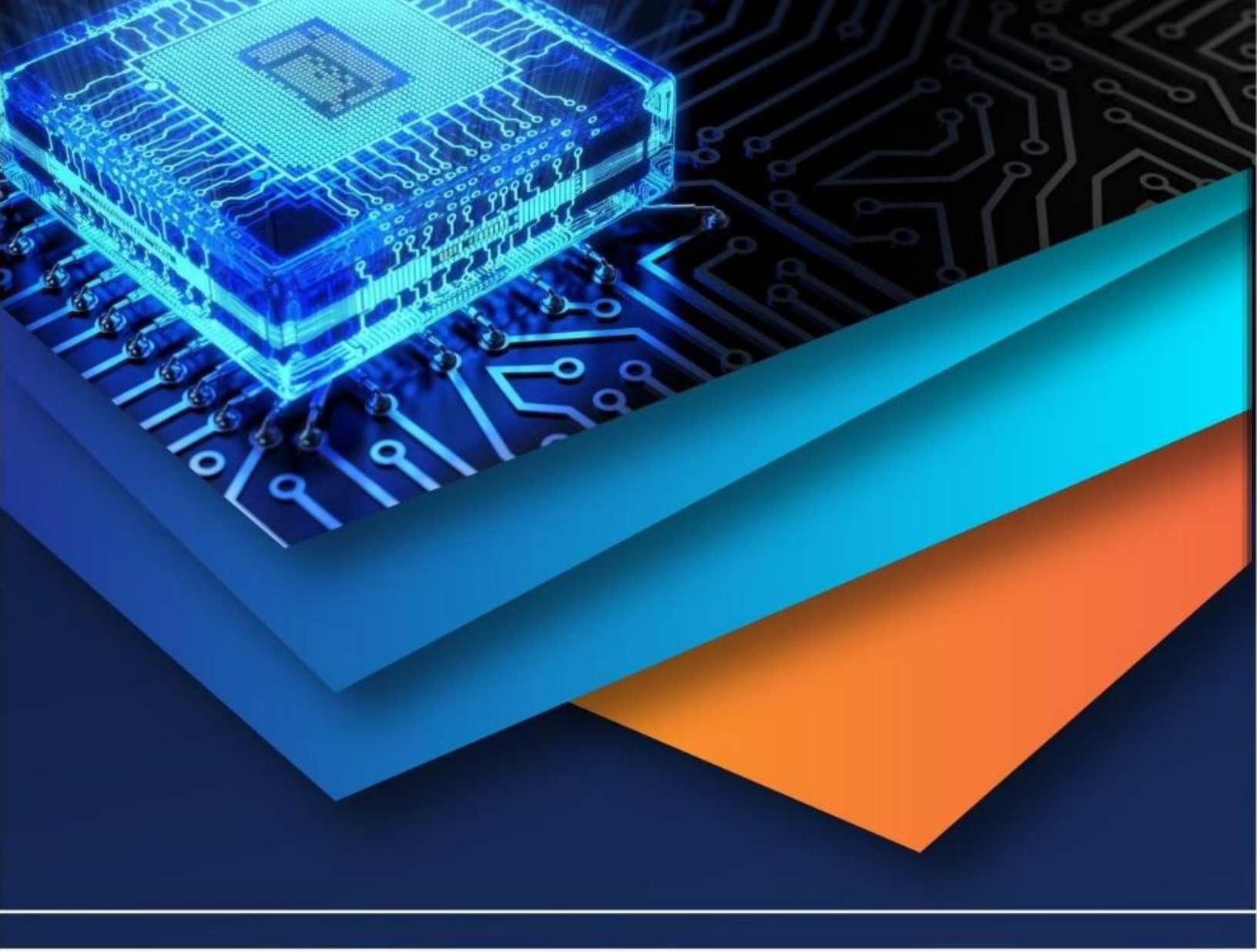
REFERENCES

[1] "Reverse Room Techniques for Irreversible Data Hiding",Journal of Chemical and Pharmaceutical Sciences 08(03): 469475, september 2015.
[2] I Safe AppDr. Sridhar Mandapati1, Sravya Pamidi2, Sriharitha Ambati3 IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 17, Issue 1, Ver. I (Jan – Feb. 2015), PP 29-34 www.iosrjournals.org
[3] Sruthi CK1, Mridul K, Zayan K, Ashin V, Sreehari K, "AI BASED WOMEN SECURITY SYSTEM",International Journal of Scientific Research and Review, ISSN NO: 2279-543X
[4] VaijayantiPawar et al Int. Journal of Engineering Research and Applications,SCIWARS Android App for Women SafetyISSN : 22489622, Vol. 4, Issue-3(version 1), March 2014, pp.823-826
[5] "Safetipin",An Android App Developed By Dr.KalpanaViswanath, https://play.google.com/store/apps/details?id=com.safetipin.mysafetipin&hl=en_IN

©IJRASET: All Rights are Reserved

[6] TELLTAIL",Delhi Integrated Multi-Modal Transit System Ltd.,
[7] Fightback" Android App Developed By CanvasM Technologies, 26 June 2013,"http://www.fightbackmobile.com/welcome"
[8] "OnWatch" Android App Developed ByOnWatch, 10 November 2012"https://play.google.com/store/apps/details?id=com.onwatch".
[9] " VITHU", Android App developed by Indian Crime Television series "GUMRAH"
[10] https://towardsdatascience.com/using-cnns-and-rnns-for-music-genre-recognition-2435fb2ed6af
[11] Design and Implementation of Women Safety System Based On Iot Technology, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7 Issue-6S3 April 2019
[12] Wav2Letter: an End-to-End ConvNet-based Speech Recognition System,RonanCollobert Facebook AI Research, Menlo Park locronan@fb.com; Christian Puhrsch Facebook AI Research, Menlo Park cpuhrsch@fb.com ;Gabriel Synnaeve Facebook AI Research, New York gab@fb.com
[13] https://www.guru99.com/convnet-tensorflow-image-classification.html

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)