Group 4: Forest Conservation Using Chainsaw Detection And Forest Fire Prediction

Abhinav Dudeja Arpan Kapoor Manish Kumar Piyush Gangle
21111001, 21111016, 21111037, 21111046
{abhinavd21, arpank21, manishk21, piyushg21}@iitk.ac.in
Indian Institute of Technology Kanpur (IIT Kanpur)

Abstract

Deforestation and uncontrollable fires have devastated many parts of India over the past decade, causing severe damage to forest ecosystems. This includes the release of carbon locked in the biomass. Our team has read various solutions that are proposed yet to overcome this problem. After analysing these solutions we have developed Machine learning enabled Internet of Things system which is capable to detect forest fire (based on environment's climatic conditions like humidity and temperature) and prevent illegal logging by detecting chainsaw sound.

1 Introduction

"Nature", after reading this word various things comes in mind like tree, animals, food, fresh air, etc. obviously, it is a nature which gives us food, shelter and environment to survive. In this environment there is a very crucial thing which is "Trees", we should thanks to the nature for giving us such dense forests, which are the major contributor to increase the level of oxygen as well reduce the level of carbon dioxide. A single tree absorbs 48 pounds of carbon dioxide every year which is the most effective carbon dioxide absorbing machine we have. But there are some sad part as well it is commonly seeing that the size of forest is reducing day by day. According to the article [1] around 15 million to 18 million hectares land of forest (size of Belgium land) are destroyed every year. In which large amount of land is destroyed by illegal means, so here we need to do something which would be very useful to catch those illegal loggers, some facts [2] are shown below:

Major drivers for deforestation (Approx.)			
Timber logging	Fuel wood (charcoal)	Wildfires	Livestock grazing
60%	25%	10%	5%

A tree is a true companion of human being but it is saddest part that, a football ground size forest is wiped out in every second, however in major countries there exists strong laws against cutting the tree, but it is hard to monitor the entire forest constantly. So here in this project we are going to focus on above major drivers and give a cost-effective solution by which illegal tree cutting will be detected in real time and wildfires can be predicted.

1.1 Overview of the project

In this project we have used Convolution neural network(CNN) for chainsaw detection and Support vector machine(SVM) for Fire prediction.
In case of Chainsaw detection we have use the microphone, by which we can record the sound and this sound send to model which will convert the sound into mel-spectogram, which feed to model(Convolution neural network) and model will predict about that this sound is of chainsaw or

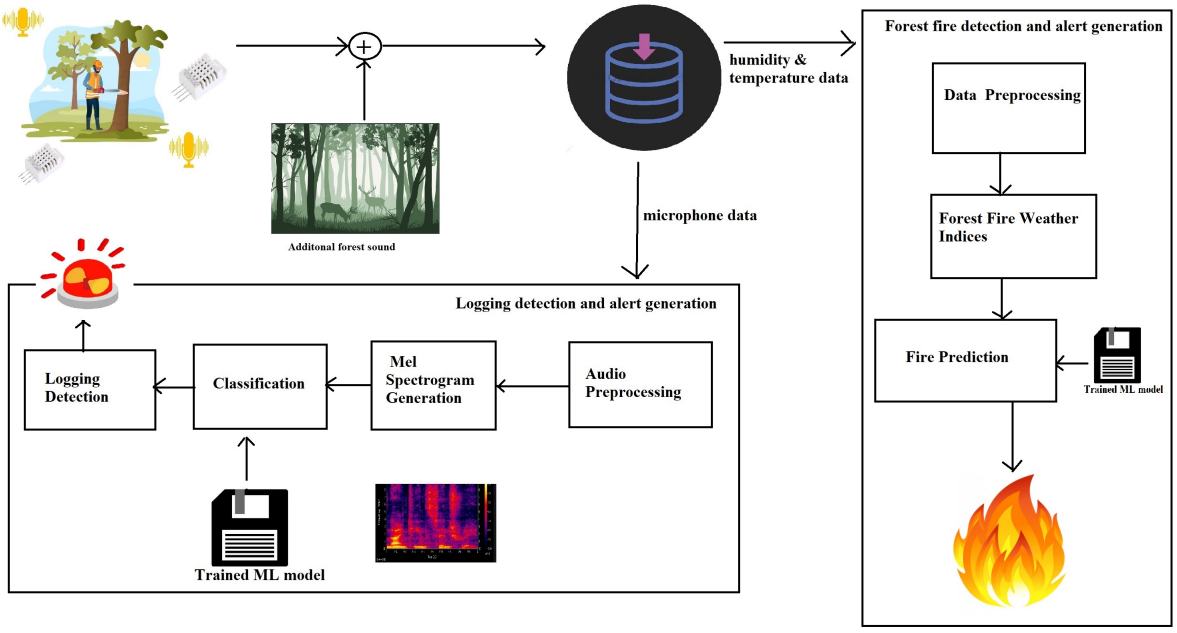

Figure 1: Diagrammatic illustration of our approach

not.
In case of Forest Fire Prediction we have uses data of temperature, humidity, wind speed and rain, which feed to model(Support Vector Machine), then initially there are some calculation of Forest Fire Indices (shown in section 4.2.1) after which model will predict about forest fire.

1.2 Motivation

Motivation behind building this project is to design intelligent system which can be used to conserve the forest from illegal logging and natural/man-made fire.

1.3 Application

Our project generate the alarm when there is a sound of chain-saw near the microphone, with that illegal loggers can be detected. so this project could be so useful for governments to address the issue of deforestation, also governments can take appropriate security measures when there is a prediction of wildfires.

1.4 Report Structure

This report contain information of two module of our project Deforestation Detection and Forest Fire Prediction, so each section contain information of those module in respective order.

The structure of this report is as follows:

- Related Work – In this section we have covered the literature survey and compared our approach to what already exists.
- Proposed Idea - This section contains an overview of what approach have been used in this project.
- Methodology - Working of the project and details about the datasets and machine learning models are contained in this section.
- Results - This section contains the results from our project.
- Discussion and Future Work - We propose ideas for further enhancement of the project

- Conclusion - Project are concluded in this section
- Individual Contributions - we mention individual contributions of all the group members with percentage

2 Related Work

2.1 Deforestation Detection

- In [3], authors have used microphone sensor to detect the environment sounds. After recording audio they performed MFCC (Mel-frequency cepstral coefficients) feature extraction and used Neural network to classify the class of audio as chainsaw sound or not. In this approach they attained accuracy of around 80%.
- In [4], authors have used microphone and gyro sensor to detect logging. In this approach they have not used any kind of Machine learning. Authors have recorded the values captured by sensor and based on some threshold values they predicted whether logging is being performed on not.
- In [5], the authors use a microphone to record sound from which they extract 5 temporal and 5 spectral features. These features are then used in a custom distance based classification model, which is compared with 3 different ML models: Gaussian Mixture Models, K means clustering and PCA. Their custom distance based model achieves higher accuracy (92%) compared to the 3 ML models.
- In [6], the authors use satellite images(10 meters spatial resolution Sentinel-1 and Sentinel-2 images), train multi-layer perceptron, long short term memory Recurrent neural network and classify land in different classes, with that find out the deforestation. their model with MLP gives 86.8% producer accuracy and in case of LSTM gives 98 % producer accuracy, however, both model user accuracy can not able to cross 68.4%.
- In [7], the authors uses microphone and record the sound after that apply Modified MFCC feature based dynamic time warping (MDTW) and spectral feature based Gauss-Bayesian classifier (SGBC). The DTWM algorithm gives average accuracy 83.27% for clean signals and 70.75% for noise mixed signals The SBGC algorithm gives average accuracy 93.47% for clean signals and 90.52% for noise mixed signals.

All the papers studied for this part have either used no ML models or have used traditional ML models, which is unlike our approach of using deep learning (CNN). Also, none of the papers discussed have used mel spectrogram images as features for training their model.

2.2 Forest Fire Prediction

- [8], proposed a system for forest fire prediction where they used the atmospheric data for the day to predict if the conditions were favourable for a forest fire. In this paper, the data points used were temperature, humidity, solar radiation, wind speed and precipitation. The system proposed in this paper does not consider the weather conditions for the past days. This system does not use any attribute to take into consideration the condition of forests due to past days which can result in wrong predictions. Our project uses indices that are calculated from the past data thus making a better prediction.
- [9], proposed a GIS (Geographic Information System) based system where weather and geographical attributes were used to find the most fire susceptible areas for better monitoring and conservation of forests. In this system satellite imaging was used to find the geographical data points such as slope, curvature, distance to settlement, distance to drainage, etc. This system works on modelling the most fire susceptible areas. Our system requires only the weather data that can be collected at much lower cost and more efficiently.

The papers discussed above either disregard the importance of past days' data or require a lot of data collection that requires a lot of investment and time. Our system works on basic data points that are responsible for forest fires i.e., Temperature, Humidity, Precipitation, etc. Also, our system does not disregard the conditions from past days by using the indices that are calculated using previous day's data.

3 Proposed Idea

We are using microphone sensor to capture the sound of forest's environment (chainsaw detection part) and Digital humidity & temperature sensor (dht22 sensor) to get humidity and temperature values of environment (Fire prediction part).

3.1 Chainsaw Detection

We have converted the audio classification problem into a binary image classification problem by generating the mel spectrogram corresponding to the recorded audio. The generated mel spectrogram is then classified into 2 classes: Chainsaw and Other sounds using a Convolutional Neural Network (CNN).

3.2 Forest Fire Prediction

The proposed system for forest fire prediction generates a fire prediction level on a 5-point scale that goes from extremely high probability to very low probability. We use SVM algorithm to predict fire probability. Our system uses atmospheric data points (Temperature, Humidity, Wind Speed, Precipitation) and forest fire weather indices (FFMC, DMC, DC, etc) to generate a prediction for wildfires.

4 Methodology

4.1 Chainsaw Detection

4.1.1 Dataset used

We have used the strongly labelled audio samples from Google Research Audioset[10]. This dataset includes labelled parts of YouTube videos and our collected data includes 431 chainsaw audio samples and 1486 audio samples of rain, thunderstorm, wind and wild animals.

4.1.2 Data Pre-processing

The audio samples have been downloaded from YouTube in mp3 format using `youtube-dl` and `ffmpeg`. The python notebook `getStronglyLabelledData.ipynb` includes the code for downloading these samples.

The collected audio samples are then loaded into python using the library `librosa` at a sample rate of 22050Hz. The loaded audio data is converted to mel spectrogram using the function `librosa.feature.melspectrogram`. This is then plotted using `matplotlib` and saved to disk as 2 classes: Chainsaw and Others.

4.1.3 Machine Learning Algorithm

We have used a Convolutional Neural Network written using `keras` and `tensorflow` to train on the 2 classes of mel spectrogram images generated from the audio samples. It uses 3 Convolution and Pooling layers and is summarised below:

The code for generating the mel spectrograms as well the training of the model is present in the python notebook `cnn.ipynb`. The code for classification/inference of a new sound is present in `cnnPredict.ipynb`.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 222, 222, 32)      896
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 111, 111, 32)      0
_________________________________________________________________
batch_normalization (BatchNo (None, 111, 111, 32)      128
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 109, 109, 32)      9248
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 55, 55, 32)        0
_________________________________________________________________
batch_normalization_1 (Batch (None, 55, 55, 32)        128
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 54, 54, 32)        4128
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 27, 27, 32)        0
_________________________________________________________________
batch_normalization_2 (Batch (None, 27, 27, 32)        128
_________________________________________________________________
flatten (Flatten)            (None, 23328)             0
_________________________________________________________________
dense (Dense)                (None, 64)                1493056
_________________________________________________________________
dropout (Dropout)            (None, 64)                0
_________________________________________________________________
dense_1 (Dense)              (None, 2)                 130
=================================================================
Total params: 1,507,842
Trainable params: 1,507,650
Non-trainable params: 192
_________________________________________________________________
```

Figure 2: CNN model summary

4.2 Forest Fire Prediction

4.2.1 Dataset used

We have used UCI dataset[11] for forest fire data. This dataset contains features of two categories:

- Weather Data – These features are basic weather data points that are temperature, humidity, wind speed and rainfall. These data points are collected on day-to-day basis.
  - Temperature in Celsius at Noon (12PM)
  - Humidity in percentage at Noon (12PM)
  - Avg. Wind Speed in MPH (Miles per hour) for past 24 hours
  - Rainfall in inches for past 24 hours
- Forest Fire Weather Indices – These indices are numeric ratings for how easily the fuel in the forest will catch fire and how fast will the fire spread through the forest. There are 6 fire weather indices that help in better understanding and prediction of forest fires.
  - FFMC (Fine Fuel Moisture Code) - Numeric rating of moisture content of small sized fuel in the forest like litter, dried grass, and leaves. This rating determines how easily fine fuels can catch fire. Higher the rating higher the risk of fire.

- DMC (Duff Moisture Code) - Numeric rating of moisture content of medium-sized fuel in the forest like small sticks and branches.
- DC (Drought Code) - Numeric rating of moisture content of large-sized fuel like logs. This rating is a useful indicator of seasonal drought effects on forest fuels.
- ISI (Initial Spread Index) - Numeric rating of expected rate of fire spread. This index is based on wind speed and FFMC values. ISI does take into consideration fuel type and spread rate may vary for different fuel types at same ISI value.
- BUI (Build Up Index) - Numeric rating of total amount of fuel available for combustion.
- FWI (Fire Weather Index) - Numeric rating of fire intensity. It is calculated using ISI and BUI values.

The code for forest fire prediction is in `SVMFirePrediction.ipynb`

4.2.2 Data Pre-processing

Dataset was divided into two sets of data, We combined the two sets. All the null values were removed and the labels were made consistent.

4.2.3 Machine Learning Algorithm

Non-Linear SVM (Support Vector Machines) Algorithm was used to train the machine learning model on this data. This model was developed using Python programming language. We used the polynomial kernel function with degree 2. The prediction probability of forest fire was divided into five classes going from extremely high to very low probability of fire with equal intervals (i.e., each class of probability has a range of 0.2).

4.3 Real world implementation

For a real world implementation, we intend to use the following hardware components:

- Raspberry Pi for executing the classification models
- DHT22 as the humidity and temperature sensor
- Microphone for capturing the sounds in the forest

The Raspberry Pi can send alerts on the smartphone of concerned forest officials.

5 Results

5.1 Chainsaw Detection

Our model was able to obtain an accuracy of 91% with a validation set accuracy of 88%. To get more insight about the module watch the demonstration video[12].

5.2 Forest Fire Prediction

Our model was able to get a validation set accuracy of above 90% and f-score of above 0.92. This result shows that our model is not generating false alarms. To get more insight about the module watch the demonstration video[13].

6 Discussion and Future Work

- Use of additional temporal and spectral features along with the mel spectrogram to improve chainsaw model
- Explore use of recurrent neural network along with CNN.
- Deployment of forest fire prediction system as an addition to existing fire detection systems to reduce the damage to forests.

7 Conclusion

This project is proposed to save the nature, we are giving here approach of deforestation detection and Forest fire prediction. As we train our chainsaw model with good amount of dataset, that's why our model gives pretty good accuracy. We develop a cost effective and scalable approach to address the issue of deforestation.

8 Individual Contributions

Individual Contributions						
Members Name	Forest Fire Prediction	Chainsaw Sound Detection	Presentation	Report	Literature Survey	Percentage Contribution
Abhinav Dudeja	✓	×	✓	✓	✓	25
Arpan Kapoor	×	✓	✓	✓	✓	25
Manish Kumar	✓	✓	✓	✓	✓	25
Piyush Gangle	×	✓	✓	✓	✓	25

References

[1] "Deforestation - wikipedia." https://en.wikipedia.org/wiki/Deforestation.

[2] "5 causes of deforestation." https://onetreeplanted.org/blogs/stories/deforestation-causes.

[3] I. Mporas, I. Perikos, V. Kelefouras, and M. Paraskevas, "Illegal logging detection based on acoustic surveillance of forest," *Applied Sciences*, vol. 10, no. 20, 2020.

[4] G. A. Mutiara, N. Suryana, and O. Mohd, "Multiple sensor on clustering wireless sensor network to tackle illegal cutting," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 10, no. 1, pp. 164–170, 2020.

[5] S. F. Ahmad and D. K. Singh, "Automatic detection of tree cutting in forests using acoustic properties," *Journal of King Saud University - Computer and Information Sciences*, 2019.

[6] L. Shumilo, B. Yailymov, M. Lavreniuk, and Y. Bilokonska, "Remote sensing approaches for deforestation identification in ukraine," in *2020 IEEE 5th International Symposium on Smart and Wireless Systems within the Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS)*, pp. 1–4, 2020.

[7] S. g, "Acoustic signal classification for deforestation monitoring: Tree cutting problem," *Hilarispublisher*, 2018.

[8] G. E. Sakr, I. H. Elhajj, G. Mitri, and U. C. Wejinya, "Artificial intelligence for forest fire prediction," in *2010 IEEE/ASME international conference on advanced intelligent mechatronics*, pp. 1311–1316, IEEE, 2010.

[9] A. Tiwari, M. Shoab, and A. Dixit, "Gis-based forest fire susceptibility modeling in pauri garhwal, india: a comparative assessment of frequency ratio, analytic hierarchy process and fuzzy modeling techniques," *Natural Hazards*, vol. 105, no. 2, pp. 1189–1230, 2021.

[10] "Google research audioset." https://research.google.com/audioset/download_strong.html, 2021. [Online; accessed 20-November-2021].

[11] "Algerian forest fires dataset data set." https://archive.ics.uci.edu/ml/datasets/Algerian+Forest+Fires+Dataset++, 2019.

[12] "Chainsaw detection demonstration." https://youtu.be/1k4Uf7K_TMU.

[13] "Forest fire prediction demonstration." https://youtu.be/wABJDTKFhm4.