Identifying and Removing Advertisements from Nepali Video

(M.Sc. Project Mid Defense)

Presenter
Sarthak Pokharel
THA078MSISE18

Supervisor
Er. Bibek Ropakheti
Assistant Professor

M.Sc. in Informatics and Intelligent Systems Engineering
Department of Electronics and Computer Engineering
Institute of Engineering, Thapathali Campus

Presentation Outline

Motivation


Start the year with right foot Skin
Himalaya
Now save when you shop with your Visa Card
BHAT-BHATENI
Use your Visa Debit or Credit card to get 5% discount on your purchase at
Wheel NOW
Rin SHAKTI
सधैं चम्किरह
Ncell
कम्बो प्याकहरु
प्रस्तुत छ नारीहरुका लागि
GGO SIM
मात्र @ Rs. 99/-
17104
ggo
ADSL Topup
ADSL
Phone Number*:
016
Amount*:
Clear


B-DESI CHANNEL
क्लिन फिड मार्फत विदेशी विज्ञापनहरु
Slice गरौं
नेपालमा विज्ञापन रहित विदेशी च्यानल प्रशारण गर्ने
नेपाल सरकारको क्लिन फिड योजनाको समर्थन गरौं।
क्लिन फिड

Background

- Video advertisement : a major marketing strategy
- Companies in FMCG, financial institutions, body products and many other industries all invest heavily
- Advertisements does not always reach target audience
- Identifying advertisements makes it easier to reach target group

Problem Statement

- Nepalese media have a different broadcasting pattern
- Previous researches fail in classifying Nepali advertisement
- Media-houses/Broadcasters are not able to insert advertisement of choice
- Nepali media dataset is inadequate for proper research

Objectives of Project

- To identify advertisement sections in Nepali videos
- To crop the advertisement and prepare a clean feed video

Scope of Project

- Project Capabilities:
  - Is able to classify advertisements that are in Nepali videos
  - Is able to provide the medium for inserting new contents

- Project Limitations
  - Is not able to detect advertisements that are part of the movies
  - Is not able to classify advertisements into different categories
  - Is not able to identify musical advertisement in between movies

Originality of Project

- The major contributions of this project work are as follows:
  - Automatic advertisement detection in Nepali videos is explored
  - MFCC and LSTM are used in detecting advertisements
  - ANepali video dataset having uniform format is prepared

Potential Applications

- Broadcasters:
  - Previously contained ads can be eliminated to insert new ones

- Creators:
  - Own contents can be implemented by creators

- Audience:
  - Irritating ads can be removed by audiences
  - Contents could be made more engaging

- Government Bodies:
  - Manual work can be eliminated while applying clean feed
  - Potential researches can be performed on the domain

Literature Review - [1]

Paper	Year	Author	Methodology	Results	Weaknesses	Strengths
Robust Learning Based TV commercial detection	2005	Xian-Sheng Hua, Lie Lua, et. al.	Context based feature derived from set of 6 features were used	The technique worked well on identifying black frames of ads	Exclusive to US adverts.	Accuracy with post processing was 95.84% and without was 94.42%
Identification of new commercials using repeated video sequence	2005	J.M. Gauch, A. Shivadas	Features extracted from commercials were matched and embedded codes were detected	Worked well on unseen ads and detected the previous occurrence of ads by video transition marking	Dataset confined to US and used traditional ML algorithms	KNN had an F measure of 95% for ad- detection

Literature Review - [2]

Paper	Year	Author(s)	Methodology	Results	Weaknesses	Strengths
Detecting Ads in Video Stream using Acoustic and Visual Cues	2006	Michele Covell, Shumeet Baluja, et. al.	Identifying/ Interpreting the visual cues with dataset from 4 days	The use of visual cues were able to increase efficiency than other traditional techniques by 40%	Only worked on rebroadcast e d programs	Precision was above 99% and recall rate was 95%
Audio and Video Processing for Automatic TV Advertisement Detection	2008	Sean Marlow, David A. Sadlier, et. al.	Ad-breaks were detected by finding patterns of silent black frames	Were tested in 3 different videos and provided satisfactory result	Only worked in mpeg video stream and used 3 videos as testing datas	Precision figure s over 99.7 % were achieved

Literature Review - [3]

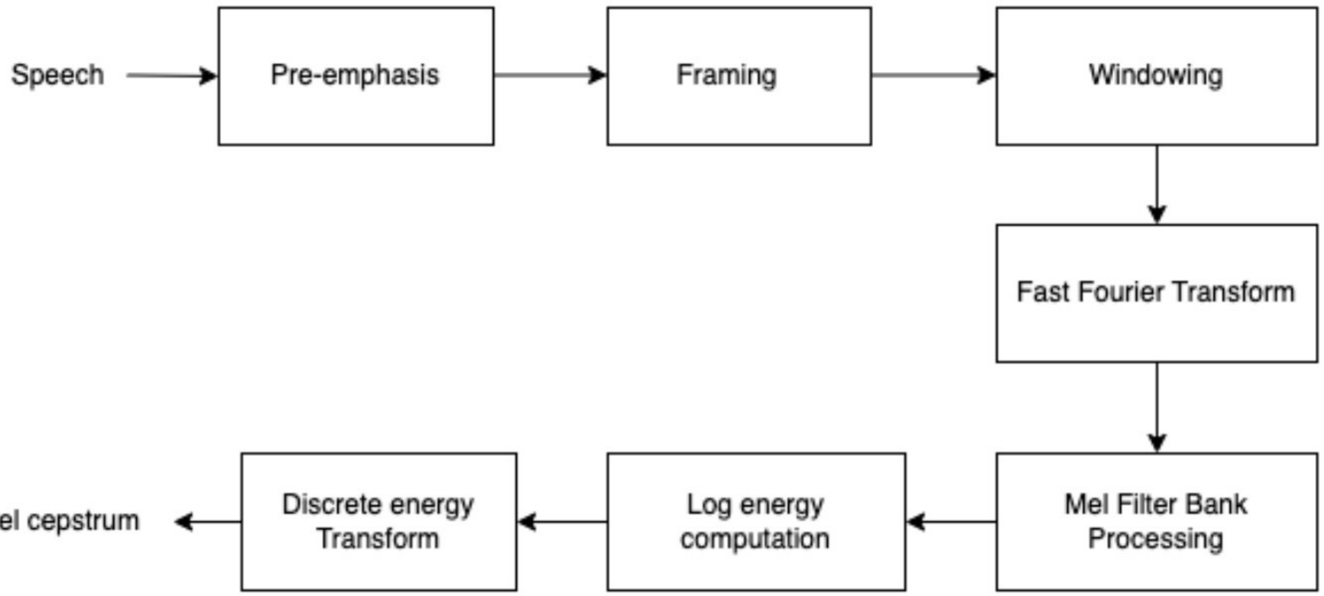
Paper	Year	Author(s)	Methodology	Results	Weaknesses	Strengths
AdNet: A Deep Network for Detecting Adverts	2018	Murhaf Hossari, Soumyabr ata Dev, et. al.	Adding 4 more layers in the VGG19 architecture	The model compared far well than the then state of the art Inception-V3 model which had accuracy of 56%	Didn't work in outdoor scenario ads and other than billboard ads	The model had an accuracy of 94% on detecting billboar d ads
An Automated Framework for Advertisement Detection and Removal from Sports Videos	2021	Abeer Toheed, Ali Javed, et. Al	Audio-visual cues of sports videos	Performance were evaluated among different SOTA techniques and performed well.	Only worked on sports videos	Provided an average accurac y of 98.75% over 5

Methodology

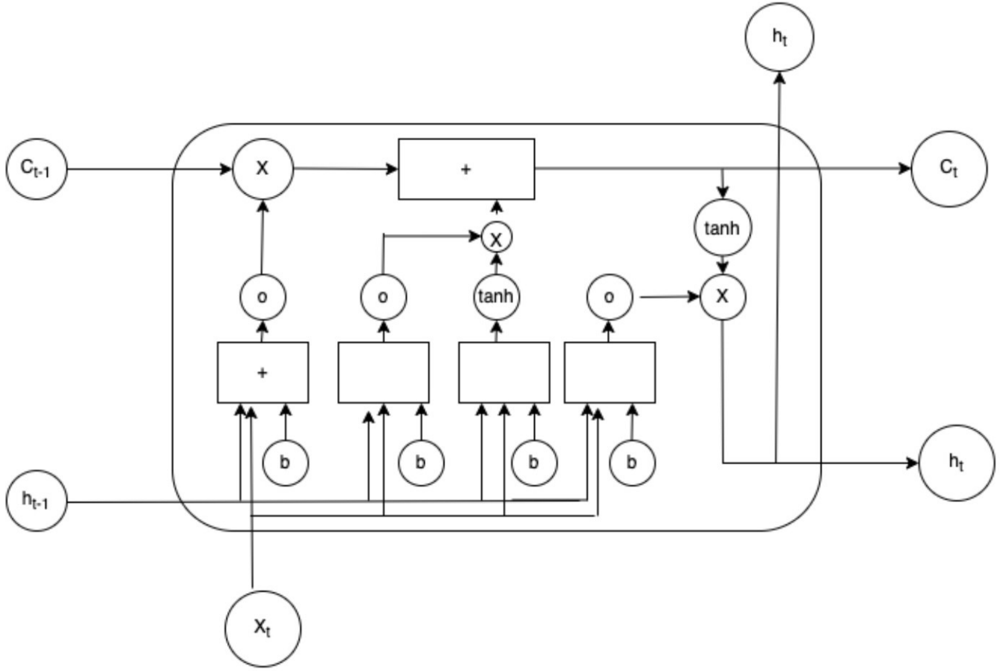
(Theoretical Formulations)

- MFCC with LSTM will be used as a deep learning model
- MFCC will extract different features from audio file
- LSTM then learns the features and classifies input into 3 different categories
- The classified advertisements are removed and clean feed is generated

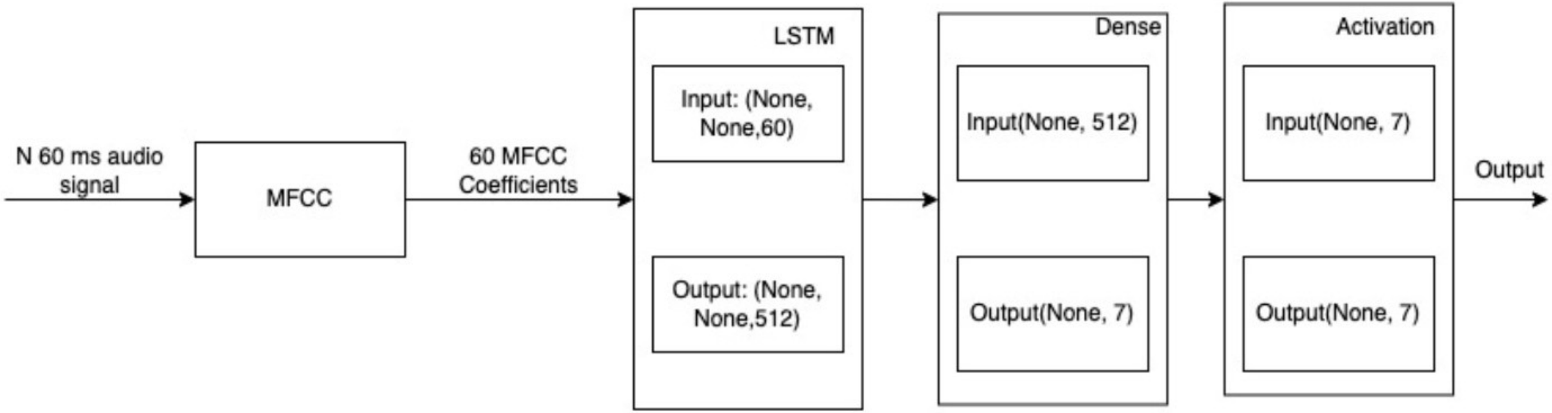
Methodology
(Mathematical Modeling - [1])



Methodology
(Mathematical Modeling - [2])



Methodology
(Mathematical Modeling - [3])



Methodology

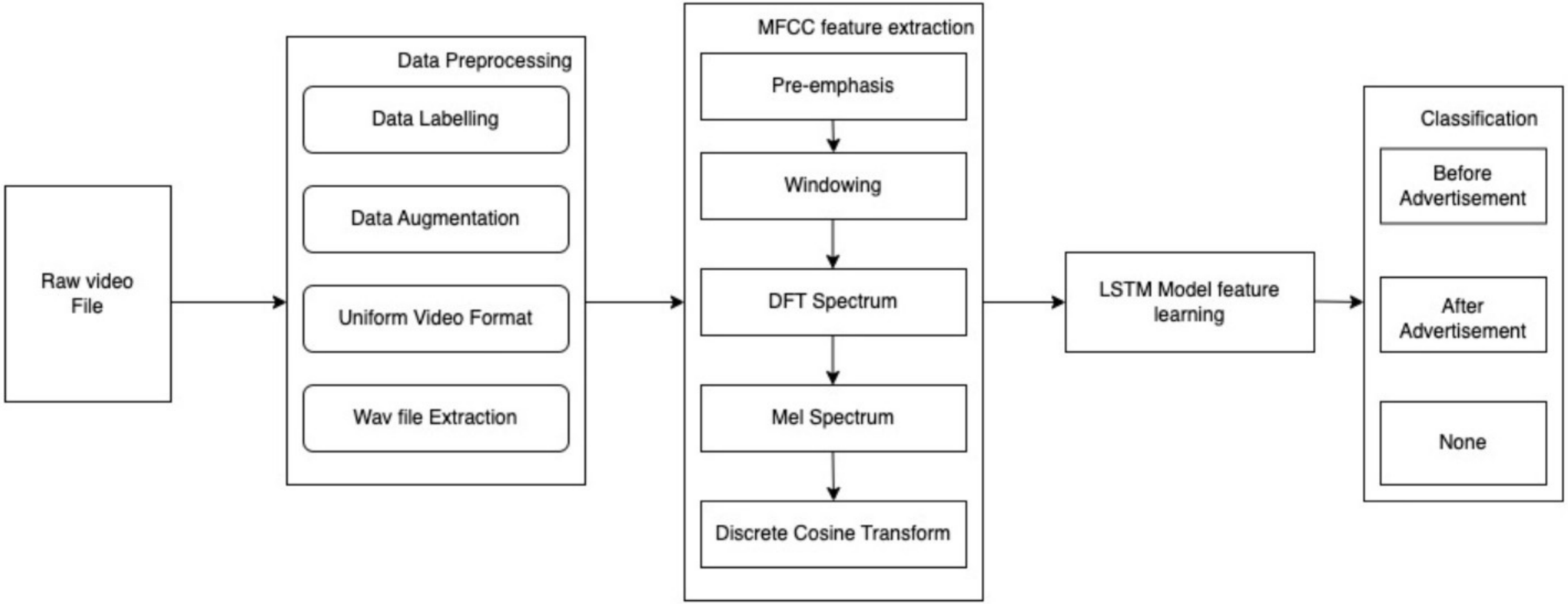
(Combining MFCC with LSTM)

- MFCC enables the ability to analyze the audio signal

- LSTM learn features generated from MFCCs to learn long-term dependencies

- LSTM has four neural network layer which interacts in a very special way

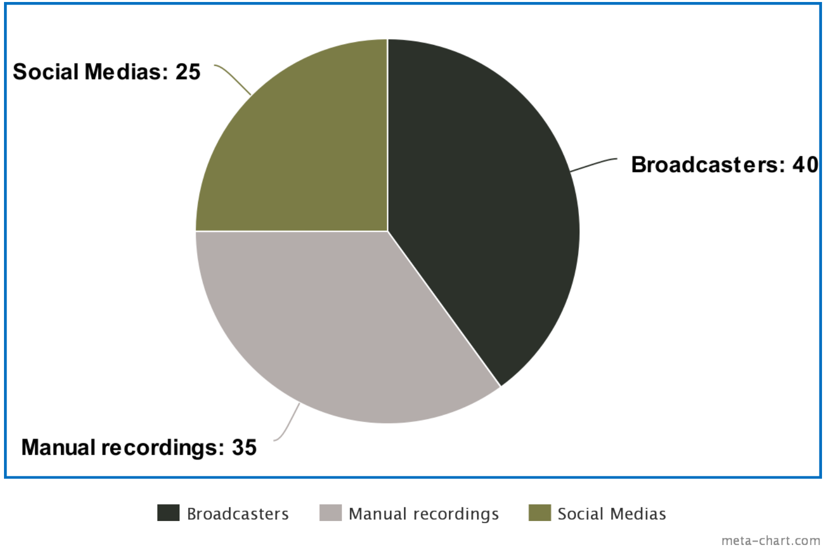
Methodology
(System Block Diagram)

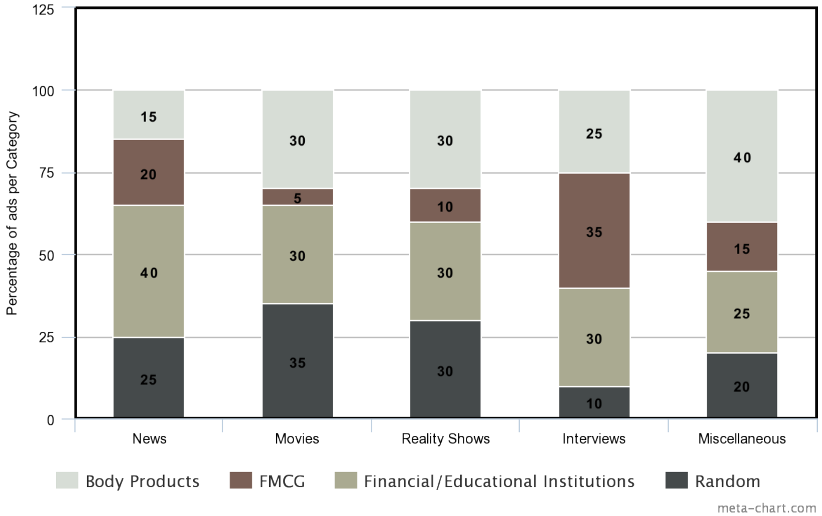


Methodology
(Instrumentation)

Hardware Requirements	Software Requirements
TPU: • Needed to execute computation intensive tasks • Obtained from TPU Research Cloud	**Python:** • Code for the project will be written in Python programming language • Python version 3.x will be used
GPU: • Needed to facilitate processing • Obtained from local computer(M1 MacBook Pro)	**IDE:** • Jupyter Notebook will be used for local development • Google collar will be used for web development
High Core CPU: • High performance RAM will be used to make execution faster • Obtained from local computer(M1 Macbook Pro)	**Tensorflow:** • Tensorflow 2.0 will be used as a Machine Learning framework

Methodology
(Sources and Analytics of Dataset)





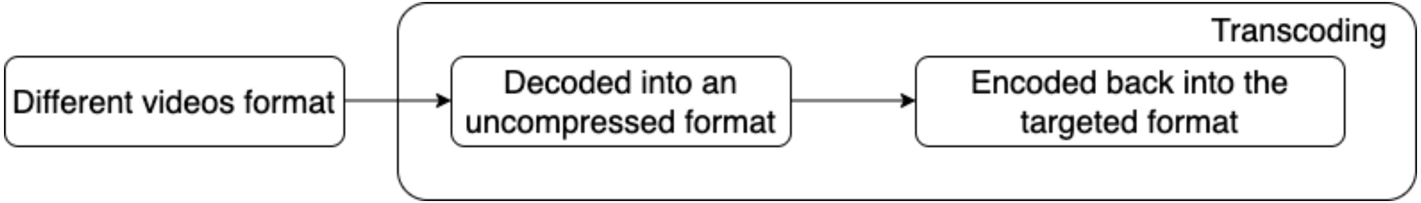
Sources of Dataset

Category of Ads

Methodology
(PreprocessingAlgorithms - [1])

- Movavi video converter will be used

- Videos are first decoded into uncompressed to produce higher quality frames

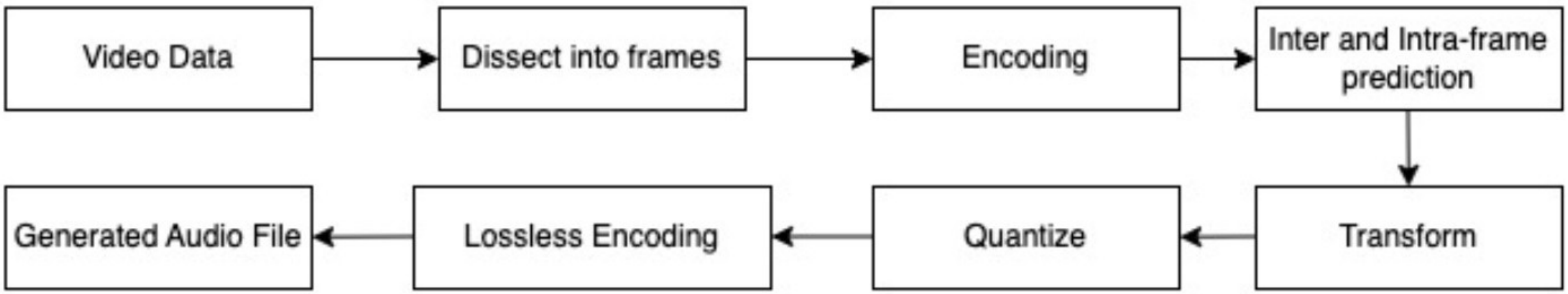


- Videos are encoded back into targeted format

Methodology

(PreprocessingAlgorithms - [2])

- FFMPEG will be used to extract audio from raw video

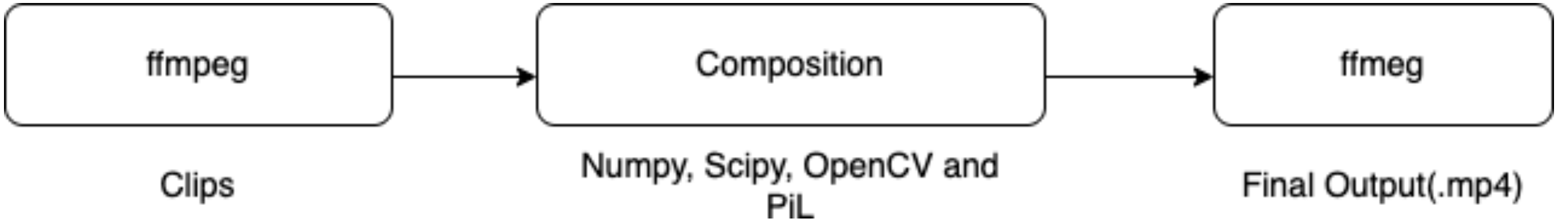


Methodology

(Post-processingAlgorithm)

- After classifying advertisements, MoviePy will be used to merge the frames

- The ffmpeg clips are composited using Python libraries and final output(.mp4) is generated



Methodology

(Elaboration of Working Principle - [1])

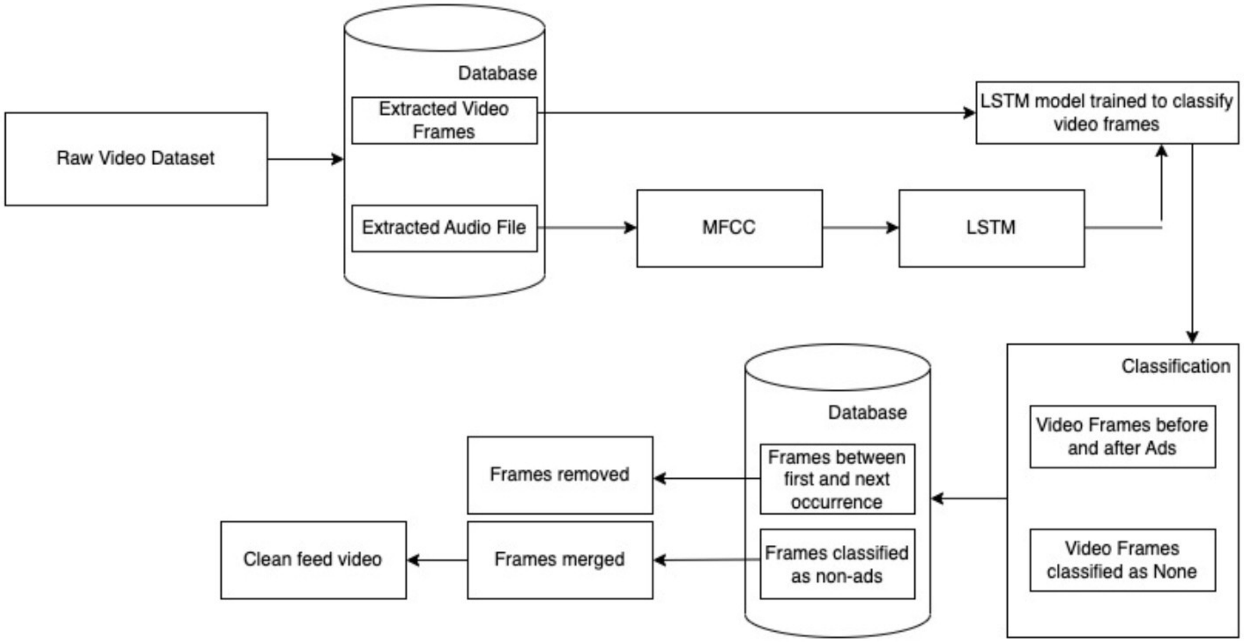
- The raw video is passed and stored in temporary database
- The video is converted into a audio(.wav) file
- MFCC features are extracted and is followed by LSTM model
- The LSTM model learns the feature and classifies it into before ad, after ad and none category

Methodology

(Elaboration of Working Principle - [2])

- The frames classified as none gets tired in database with non-ad category column
- The frames after first occurrence of break emphasis are stored as advertisement category
- The frames get stored until the next occurrence of word
- The frames with none category are combined to produce a clean feed video

Methodology
(Elaboration of Working Principle - [3])



Methodology

(Model Evaluation - [1])

- Categorical cross entropy loss function is used in the classification block

- The model terminates when the epochs is equal to 100

- Precision/Recall and F1 score is used to evaluate model performance

Methodology
(Model Evaluation - [2])

- Some manually developed metrics will also be used

$$CCR = \frac{CorrectlyClassifiedFrames}{TotalFrames} * 100$$

Correctly Classified Rate

$$NCR = \frac{FalselyClassifiedFrames}{TotalFrames} * 100$$

Negatively Classified Rate

$$PPCR = \frac{FramesIdentifiedasAdvertisement}{TotalAdvertisementFrames} * 100$$

Partially Positive Classified Rate

$$PNCR = \frac{FramesIdentifiedasAdvertisement}{TotalNon - Advertisement} * 100$$

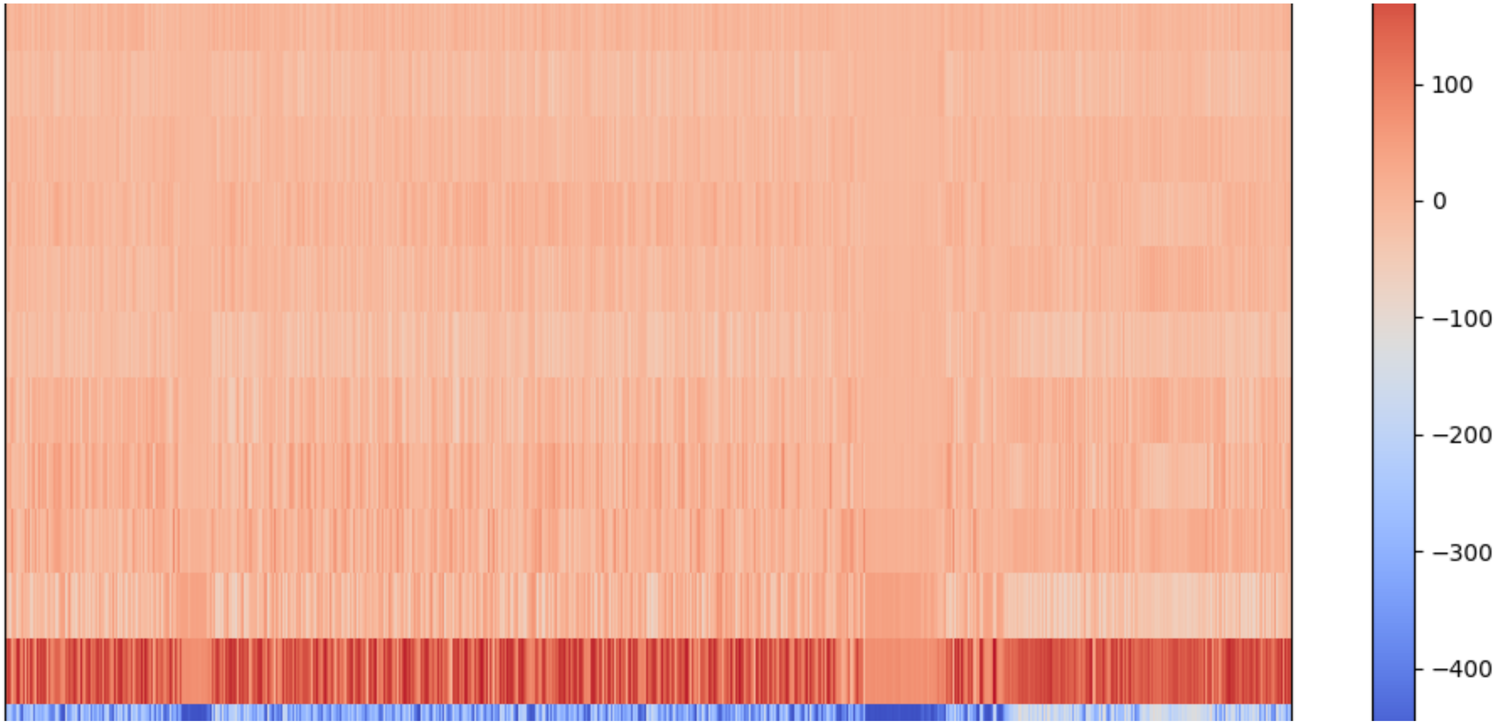
Partially Negative Classified Rate

Results - [1]

- MFCC of advertisements and non-advertisements are generated

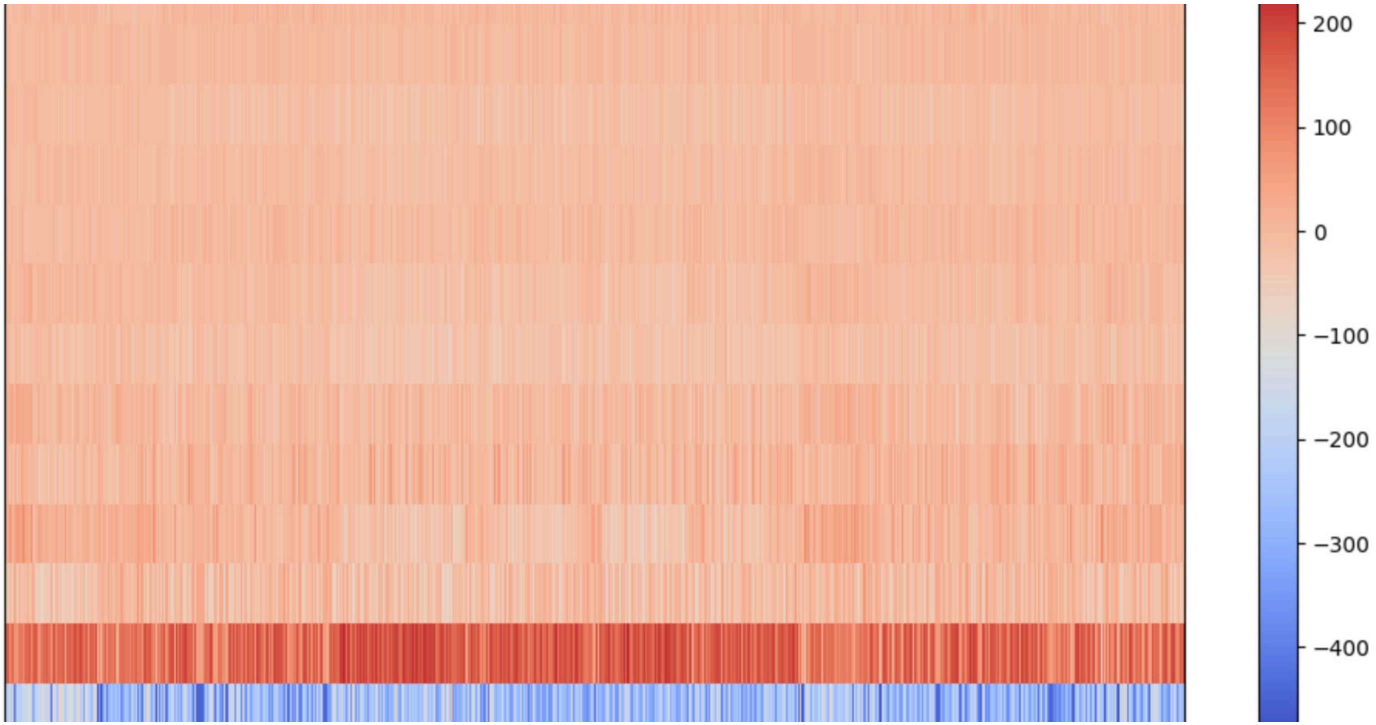
- The MFCC plots of these sections are visibly different.

Results - [2]



MFCC plot of advertisement

Results - [3]



MFCC of news section

Results - [4]

```
youtube = build('youtube', 'v3', developerKey=api_key)

def download_youtube_video(video_url, output_path='downloaded_video.mp4'):
    yt = YouTube(video_url)
    stream = yt.streams.filter(only_audio=True).first()
    stream.download(filename=output_path)
    return output_path

# Convert video file to .wav
def convert_to_wav(video_path, audio_path='output_audio.wav'):
    # Extract audio from video
    audio_clip = AudioFileClip(video_path)
    audio_clip.write_audiofile(audio_path, codec='pcm_s16le')
    audio_clip.close()

def search_videos(query, max_results=10):
    # Make an API call to search for videos
    request = youtube.search().list(
        q=query,
        part='id,snippet',
```

Python code to download YouTube video and convert it to audio

Results - [5]

```
Response', 'etag': 'QV1cpO69YAk3ES1Tg9PEIle6iwM', 'nextPageToken': 'CAEQAA', 'regionCode': 'US', 'pageInfo': {'to
नियाँका दिनभरका मुख्य समाचार | असार १८, मंगलबार, बिहान ११ बजे | AP1HD
 NIRAJAN BHANDARI HEADLINES: १.प्रधानमन्त्रीलाई भेटन एमाले ...

/watch?v=Ad930CNE78o
utput_audio.wav
                                        MoviePy - Done.
```

Response of YouTube downloading code

Discussion And Analysis

- Google project creation
- Youtube download and conversion to audio
- MFCC plot of advertisement and news section manually
- MFCC is good for audio signals.

Remaining Tasks

- Dataset finalization
- LSTM Model creation
- Model building, training and evaluation.
- Result interpretation

References - [1]

1Xian-Sheng Hua, Lie Lu, and Hong-Jiang Zhang. Robust learning-based tv commercial detection. In 2005 *IEEE International Conference on Multimedia and Expo,* pages 4 pp.–, 2005.

2J.M. Gauch and A. Shivadas. Identification of new commercials using repeated video sequence detection. In *IEEE International Conference on Image Processing 2005*, volume 3, pages II–1252, 2005.

3Michele Covell, Shumeet Baluja, and Michael Fink. *Detecting ads in video streams using acoustic and visual cues.* Computer, 39:135 – 137, 01 2007.

4Shervin Minaee, Imed Bouazizi, Prakash Kolan, and Hossein Najafzadeh. *Ad-net: audio-visual convolutional neural network for advertisement detection in videos.* arXiv preprint arXiv:1806.08612, 2018.

References - [2]

5Abeer Toheed, Ali Javed, Aun Irtaza, Hassan Dawood, Hussain Dawood, and Ahmed S Alfakeeh. *An automated framework for advertisement detection and removal from sports videos using audio- visual cues.* Frontiers Comput. Sci., 15(2):152313, 2021.

6Zihang Dai, Hanxiao Liu, Quoc V Le, and Mingxing Tan. *Coatnet: Marrying convolution and attention for all data sizes.* Advances in Neural Information Processing Systems, 34:3965–3977, 2021.

7SeÅLan Marlow, David A Sadlier, Karen McGeough, Noel E O'Connor, and Noel Murphy. *Audio and video processing for automatic tv advertisement detection.* 2001.

8Ramin Zabih, Justin Miller, and Kevin Mai. A feature-based algorithm for detecting and classifying scene breaks. In *Proceedings of the third ACM international conference on Multimedia*, pages 189– 200, 1995.

Thank You !