Running Head: FIRE DETECTION USING OPENCV AND G-STREAMER

FIRE DETECTION USING OPENCV AND G- STREAMER

MENTOR:-AYAN PANJA

SUBROTO RAKSHIT

PRATYUSH RANJAN

PAYEL CHATTERJEE

PRABHAT KUMAR

AJAY KUMAR MANDAL

SUBHA CHAKRABORTY

INSTITUTE OF ENGINEERING AND MANAGEMENT,
SECTOR - V , KOLKATA , WEST BENGAL

Running Head: FIRE DETECTION USING OPENCV AND G-STREAMER

ACKNOWLEDGEMENT

We would like to express our deep sense of gratitude to our professors, HOD
Prof. Amartya Sir, Prof. Ayan Sir, and all other professors, for giving us this project
Work and encourage us to do this project successfully. We have completed
this project under their guidance and under their support.

They have supported us in every step and provided all the necessary
Conditions and environment to complete this project.

Without their support, this project would have been difficult for us to complete.

Last but not the least, we would like to thank our parents and all other people,
who have supported us to do this project. Thank you very much.

Date: 29/11/2023

Running Head: FIRE DETECTION USING OPENCV AND G-STREAMER

INDEX

Chapters	Page No.
1.Abstract	*4*
2. Introduction *2.1. Multimedia over the internet* *2.2. RTSP (Real Time Streaming Protocol)*	*5-7*
3. Gstreamer *3.1. Why to use Gstreamer?* *3.2. How can we use Gstreamer in our IT Project?* *3.3. Future plans*	*8-10*
4. Open-CV *4.1. Applications of OpenCV* *4.2. Image-Processing* *4.3.Digital-Image* *4.4.Computer Vision* *4.5.Computer Vision Vs Image Processing* *4.6.Applications of Computer Vision* *4.7.Features of OpenCV Library*	*10-13*
5.Related Works	*13*
6.Application Domain	*13-14*
7.Motivation	14
8.Proposed Approach	14
9. Experimental Results	15
10. Conclusion	16

1. ABSTRACT:-

In this modern world, AI and ML play a very important role. The market of AI and ML is increasing. All the domains are now based on AI and ML or they adapting them to AI and ML. Most of the things that we utilize in this generation are based on artificial intelligence and machine learning. We are allowing computers to think and take decisions. Now computers are self-working. They have the ability to control themselves and detect anything.

So our project is also based on AI and ML. We are going the build a model which will detect smoke and fire at any distinct place and it will automatically generate an alarm for it. All the work will be computer-based. It can be utilized by any county for the fire and smoke detection system. The modules that will be used are G-streamer for the port transfer and open CV for detecting purposes. There will be cameras for detecting the fire and smoke. Signals will be shared via the internet and if the machine algorithm identifies any smoke and fire then it will send a report to the server, and it will also generate an alarm in such situations.

The purpose of the G-streamer will be to break the packets and generate the report and a pipeline will be created by it and accordingly the system will work. The main purpose of G-streamer is to create a pipeline for the project. The main purpose of Open-CV is to handle the media file and to generate a report in array format and which will transferred by the pipeline that will be created using the G-streamer.

KEYWORDS: project, pipeline, G-streamer, protocols, report, open-CV, packets

2. INTRODUCTION:-

2.1. Multimedia over the internet:-

The internet was originally designed to support data communications. Most of the traffic initially includes data such as e-mails and files. Voice traffic was served exclusively by telephone networks. As the Internet grew in terms of the number of nodes, applications, and users, the need for multimedia communication over the internet emerged. Animation, voice, and video clips are now common on the internet. Multimedia networking products like internet telephone, internet T.V, and video conferencing are available in the market. Shortly, people will enjoy other multimedia products in distance learning, distribution simulation, distributed workgroups, and other areas.

2.2 RTSP (Real Time Streaming Protocol):-

This project is based on Real-Time Streaming, so Real-Time Streaming Protocol(RTPS) will be used in this project. RTSP is an application-level networking protocol for a communication system that transfers real-time data from multimedia to an endpoint device by communicating directly with the server streaming the data.

This protocol controls the flow of data between the media devices and also establishes and controls the media stream between client devices and serves by serving as a network remote control for continuous media, such as audio and video. It does not stream the multimedia itself but communicates with the server that streams the multimedia data.

It also provides a protocol for downloading the media file in the device and also gives the control to user-related playing the file, i.e., play, pause, fast, slow, etc. The quality of the file can be controlled.

RTSP is a stateful protocol. A stateless protocol does not require the server to retain session information or status about each communicating partner for the duration of multiple requests. In contrast, a protocol that requires the keeping of the internal state on the server is known as a stateful protocol.

In a stateful protocol, there is a tight dependency between client and server. If the client sends a request to the server then it expects some kind of response, if it does not get any response then resend the request. A stateful server keeps the state of connections.

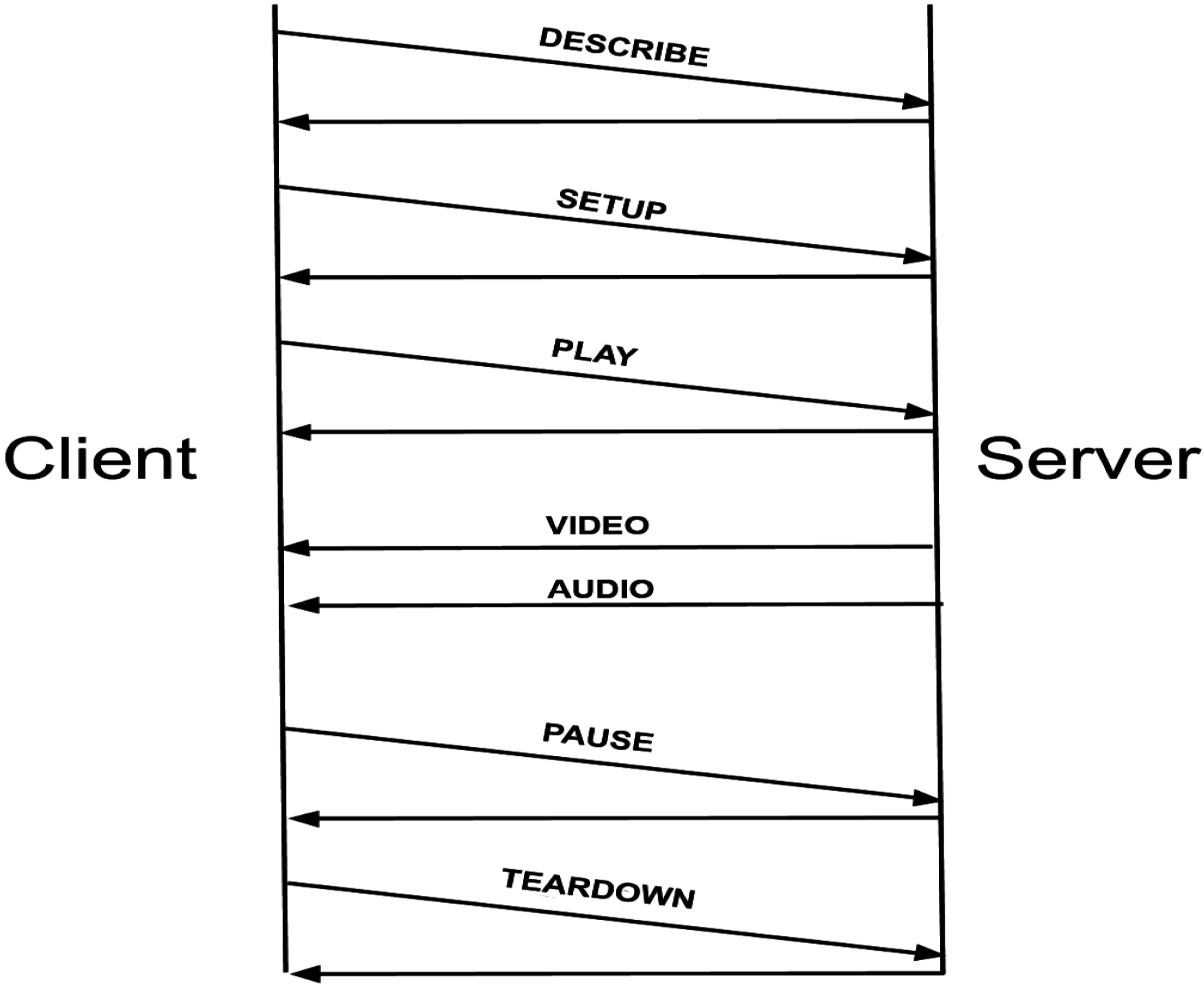


Fig 2.2.1

Fig 2.2.1 shows the connection setup between the server and the clients. Packets are transferred between the server and the clients and a communication channel is a setup between them. After the channel is set up, flies are shared in packets. There are different types of layers in the TCP/UDP and other types of layers.

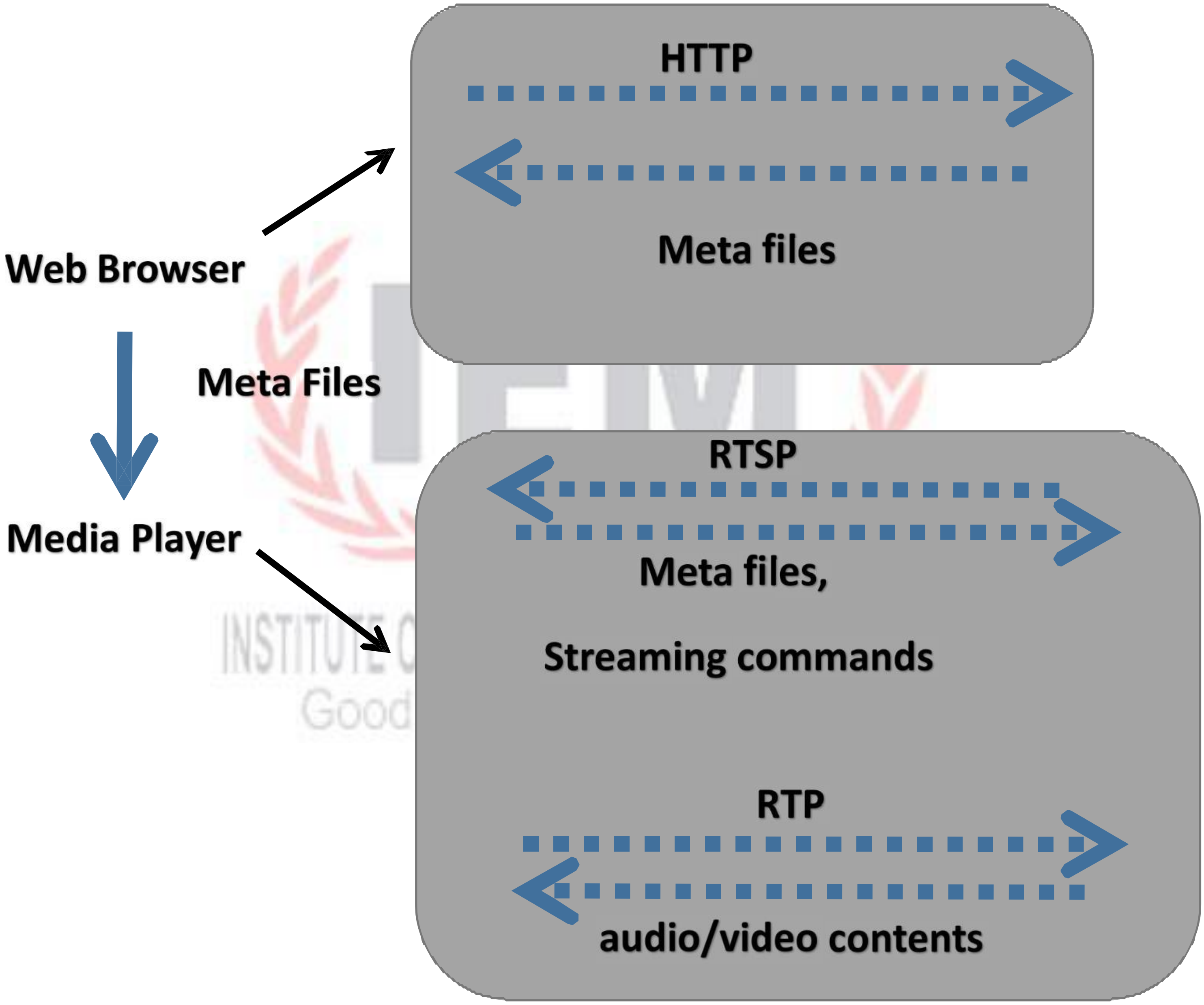


Fig 2.2.2

3. Gstreamer:-

GStreamer is an open-source multimedia framework that is mainly used to create media applications. Such as for streaming, media playback, non-linear editing, etc.
GStreamer can be used to build a system that reads files in one format, processes them, and exports them in another.
The GStreamer framework is designed to make it easy to write applications that handle audio or video or both.
It uses plugins that will provide various codecs and other functionalities.

Its most common uses are as follows:-
to extract audio track from some media file container,
to transcode video and/or audio track (change encoding format) so that I can play it in the web browser,
to change the resolution (width/height) and FPS (frames per second) of some video files,
to convert (transmit) one media file container to another (e.g. from MP4 to MPEG TS)
to stream my movie to another PC in a local network.

3.1. Why to use Gstreamer?

Dealing with multimedia is not an easy task. Therefore, selecting the proper framework to process audio and video streams becomes a key point to ensure a successful project. Developers facing this challenge should consider the following factors before making their choice for the right multimedia framework:-

coding complexity (architecture),
cross-platform support,
multimedia technologies coverage (codecs, filters, etc.),
documentation and support are available.

GStreamer excels in the above-mentioned criteria:

its intelligence plugin architecture and a comprehensive core library make your application development easy.

It provides you with well-tested elements for many of your needs.

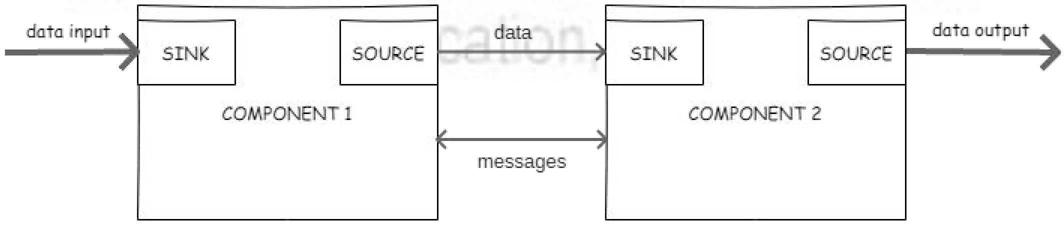
It works on all major operating systems (ie. Linux, Android, Windows, Mac OS X, iOS) and runs on all major hardware architectures including x86, ARM, MIPS, SPARC.
It counts with an extensive list of multimedia plugins (encoders, decoders, content filters...).
It allows easy integration of 3rd party ones.

Lastly, GStreamer counts with wide and well-structured documentation available for developers and a big community whose contributions ensure a continuous improvement of this framework.

3.2. How can we use Gstreamer in our IT Project?

So, we can use Gstreamer to
Extract audio from media files recorded by a Surveillance camera
Extract video or image from media files too.

Gstreamer can handle and process media files using a variety of components (elements) that can be found in Gstreamer plugins. Components can be interconnected forming a pipeline where data usually flows upstream (from source to sink, from "left to right").
In general, elements are connected like this:-



And are capable to exchange messages to inform each other about some events happening during data processing (e.g. sync messages).

For extracting the media files:-
we will need to learn more about how Gstreamer forms a pipeline for media processing.

The pipeline will be very similar to the audio one with the difference that we need to use a video decoder and video encoder.

4. Open-CV:-

OpenCV is the huge open-source library for computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today’s systems. By using it, one can process images and videos to identify objects, faces, or even the handwriting of a human. When it is integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image patterns and its various features we use vector space and perform mathematical operations on these features.

4.1. Applications of OpenCV:-

There are lots of applications that are solved using OpenCV, some of them are listed below
face recognition
Automated inspection and surveillance
number of people – count (foot traffic in a mall, etc)
Vehicle counting on highways along with their speeds
Interactive art installations.

4.2. Image-Processing:-

Image processing is a method to perform some operations on an image, to get an enhanced image, and or to extract some useful information from it.

If we talk about the basic definition of image processing then “Image processing is the analysis and manipulation of a digitized image, especially to improve its quality”.

4.3. Digital-Image :-

An image may be defined as a two-dimensional function f(x, y), where x and y are spatial(plane) coordinates, and the amplitude of fat any pair of coordinates (x, y) is called the intensity or grey level of the image at that point.

In another word An image is nothing more than a two-dimensional matrix (3-D in the case of colored images) which is defined by the mathematical function f(x, y) at any point is giving the pixel value at that point of an image, the pixel value describes how bright that pixel is, and what color it should be.

Image processing is signal processing in which input is an image and output is an image or characteristics according to the requirement associated with that image.
Image processing includes the following three steps:

Importing the image
Analyzing and manipulating the image
Output in which result can be altered image or report that is based on image analysis
Expectation-maximization algorithm
k-nearest neighbor algorithm
Naive Bayes classifier
Artificial neural networks
Random forest
Support vector machine (SVM)
Deep neural networks (DNN)
OpenCV is written in C++ and its primary interface is in C++, but it still retains a less comprehensive though extensive older C interface. All of the new developments and

algorithms appear in the C++ interface. There are bindings in Python, Java, and MATLAB/OCTAVE. The API for these interfaces can be found in the online documentation. Wrappers in several programming laIf the library find Intel's Integrated Performance Primitives on the system, it will use these proprietary optimized routines to accelerate itself.

A CUDA-based GPU interface has been in progress since September 2010.

An OpenCL-based GPU interface has been in progress since October 2012,

OpenCV is a cross-platform library using which we can develop real-time computer vision applications. It mainly focuses on image processing, video capture, and analysis including features like face detection and object detection.

Let’s start the chapter by defining the term "Computer Vision".

4.4. Computer Vision

Computer Vision can be defined as a discipline that explains how to reconstruct, interrupt, and understand a 3D scene from its 2D images, in terms of the properties of the structure present in the scene. It deals with modeling and replicating human vision using computer software and hardware.

Computer Vision overlaps significantly with the following fields −

Image Processing − It focuses on image manipulation.

Pattern Recognition − It explains various techniques to classify patterns.

Photogrammetry − It is concerned with obtaining accurate measurements from images.

4.5. Computer Vision Vs Image Processing

Image processing deals with image-to-image transformation. The input and output of image processing are both images.

Computer vision is the construction of explicit, meaningful descriptions of physical objects from their image. The output of computer vision is a description or an interpretation of structures in a 3D scene.

4.6. Applications of Computer Vision

Here we have listed down some of the major domains where Computer Vision is heavily used.
Robotics Application
Localization − Determine robot location automatically
Navigation
Obstacles avoidance

4.7. Features of OpenCV Library

Using OpenCV library, you can −
Read and write images
Capture and save videos
Process images (filter, transform)
Perform feature detection

5. Related Works:-

In 2016, a project which dealt with Real-Time Multisensor Fire Detection and Notification System using Web-Based Communication was developed. The hardware arrangement consisted of an Arduino UNO, an MQ2 smoke sensor, a TMP102 temperature sensor, and a DFRobot flame sensor, all working by the implementation of the desired outcome-based logic, putting it all together. The detection and warning about the smoke or flame are done in real-time.
In 2020, a project dealing with Fire and Smoke Detection using DeepStream and Edge Computing Approach was implemented.
In 2021, a real-time smoke and fire detection project was developed using Convolutional Neural Network (CNN) in antifire surveillance systems.

6. Application Domain:-

Fire detection:- Through camera sources, the project can detect fire in a forest area in real-time, which can be used to inform required authorities.

Smoke detection:- It can be used in the forest region where we want to detect any fire from smoke. It detects the level of smoke and accordingly predicts the presence of fire .

7. Motivation

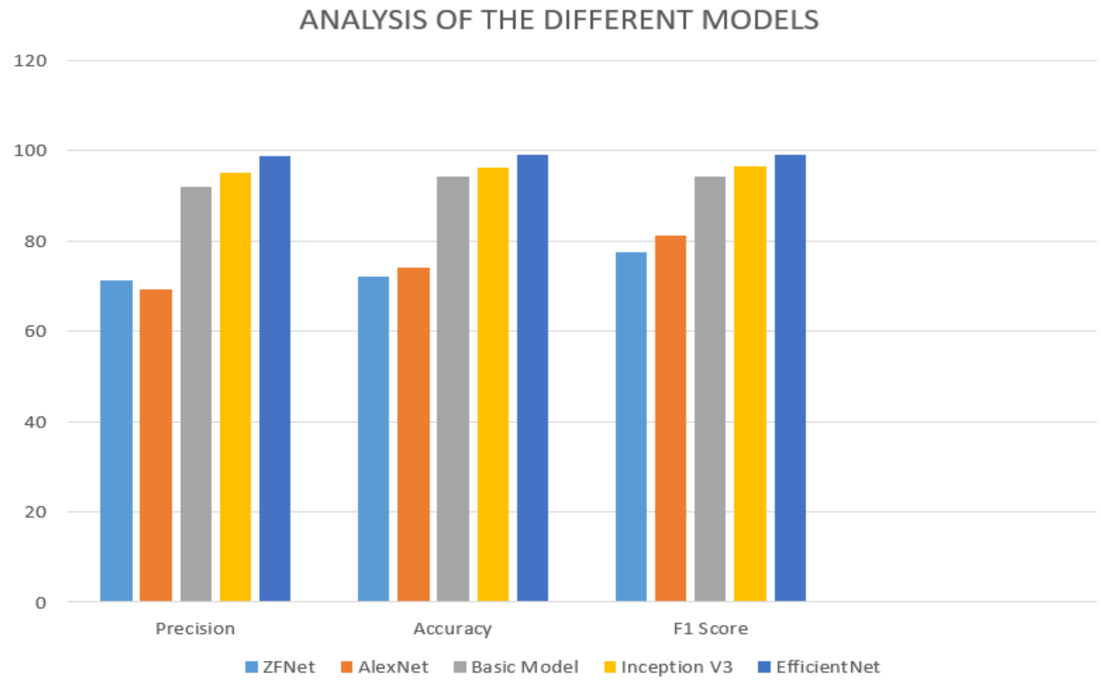
Forest Fires have deadly immediate as well as everlasting consequences, as not only it causes the loss of animal life and their habitats, large-scale burning of trees can disrupt the climatic conditions (aiding in global warming), and leading to unprecedented ecological disaster. If we could detect fire automatically and in real-time, we can possible aid authorities in taking action sooner, thereby avoiding an ecological disaster.

This is what motivated us to pursue this project.

8. Proposed Approach

In a seamless integration of cutting-edge technologies, the process begins by capturing video information from a camera source, utilizing the RTSP protocol within the powerful GStreamer framework. This video stream is then transmitted to a local host, where the OpenCV framework takes the reins to handle and process the incoming frames. The prowess of OpenCV is harnessed to facilitate efficient manipulation and analysis of the video feed.At the heart of this intelligent system lies the EfficientNet model, a groundbreaking convolutional neural network architecture introduced by Google researchers in 2019. Renowned for its ability to achieve unparalleled performance while requiring fewer parameters and computational resources, EfficientNet stands as a beacon of innovation in the realm of deep learning. Tasked with the critical role of fire prediction, this sophisticated model processes the frames received from the camera source, utilizing its robust neural network to discern patterns indicative of potential fire incidents.In essence, this integrated system seamlessly weaves together the real-time video capture capabilities of GStreamer, the versatile frame handling capabilities of OpenCV, and the predictive prowess of the EfficientNet model. Together, they form a formidable synergy, empowering the system to analyze video data with efficiency and accuracy, ultimately contributing to the timely and precise prediction of fire events.

9. Experimental Result



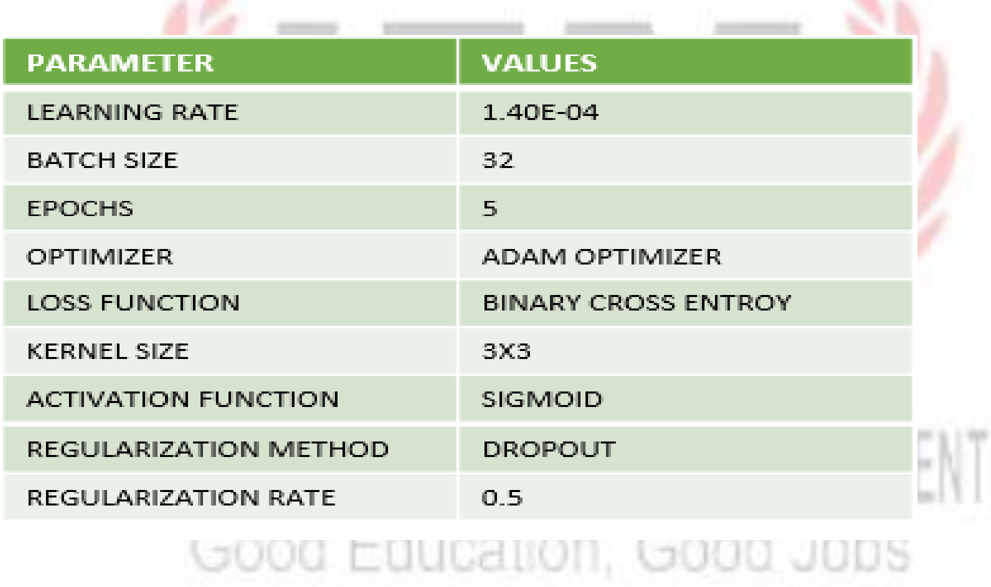
For our project, we tried models suited for our project, whose results are shown as in the figure above. The models tried are:

ZFNet, short for "Zeiler and Fergus Net," is a convolutional neural network (CNN) architecture designed for image classification tasks. Developed by Matthew D. Zeiler and Rob Fergus in 2013.

AlexNet is a pioneering convolutional neural network (CNN) architecture that significantly contributed to the advancement of deep learning in computer vision. Developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton.

Inception v3 is a convolutional neural network (CNN) architecture designed for image classification and object recognition. Developed by Google's research team, it is a successor to the original Inception model. Introduced in 2015.

EfficientNet is a groundbreaking convolutional neural network (CNN) architecture designed to achieve superior performance with fewer parameters and computational resources. Introduced by Google researchers in 2019.Among them Efficient-Net is chosen as the predicting model, after trying multiple models, it has the highest precision, accuracy and F1 Score.



PARAMETER	VALUES
LEARNING RATE	1.40E-04
BATCH SIZE	32
EPOCHS	5
OPTIMIZER	ADAM OPTIMIZER
LOSS FUNCTION	BINARY CROSS ENTROY
KERNEL SIZE	3X3
ACTIVATION FUNCTION	SIGMOID
REGULARIZATION METHOD	DROPOUT
REGULARIZATION RATE	0.5

Figure: Parameter values used for all models for experiment

10. Conclusion

The project deals with real time detection of forest fire and smoke. The video information from local CCTVs and other camera sources is taken, on which predictions is made using a CNN model, to detect forest fire. Efficient-Net is chosen as the predicting model, after trying multiple models, it has the highest precision, accuracy and F1Score as compared to other models. This system can be integrated with pre-existing warning systems(ex:- SMS warning, etc), and will aid authorities to know about forest fire faster, thereby ensuring the fire will be dealt with as soon as possible.

References:-

Authors of the said project: Robert A. Sowah, Abdul R. Ofoli, Selase N. Krakani, Seth Yayra Fiawoo
https://ieeexplore.ieee.org/abstract/document/75756

85 Authors: Yi-Chun Chen, Halim Fathoni, Chao-Tung Yang https://ieeexplore.ieee.org/abstract/document/9302759/authors#authors Authors: Sergio Saponara, Abdussalam Elhanashi, Alessio

Gagliardi Link:https: //link.springer.com/article/10.1007/s11554-020-01044-0

http://4youngpadawans.com/gstreamer-real-life-examples/ https://gstreamer.freedesktop.org/documentation/installing/on-windows.html?gi-language=c https://fluendo.com/en/what-is-gstreamer/