REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Information Technologies

FIRE AND SMOKE DETECTION BASED ON ARTIFICIAL INTELLIGENCE TECHNIQUES

Ali Faris Mansor AL-KHAFAJI

Master's Thesis

Supervisor

Asst. Prof. Dr. Oğuz KARAN

İstanbul, 2023

FIRE AND SMOKE DETECTION BASED ON ARTIFICIAL INTELLIGENCE TECHNIQUES

Ali Faris Mansor AL-KHAFAJI

Information Technologies

Master's Thesis

ALTINBAŞ UNIVERSITY

2023

The thesis titled FIRE AND SMOKE DETECTION BASED ON ARTIFICIAL INTELLIGENCE TECHNIQUES, prepared by ALI FARIS MANSOR AL-KHAFAJI, and submitted on 27/12/2023 has been **accepted unanimously** for the degree of Master of Science in Information Technologies.

____

Asst. Prof. Dr. Oğuz KARAN

Supervisor

Thesis Defense Committee Members:

Asst. Prof. Dr. Oğuz KARAN	Department of Software Engineering, Altınbaş University	____________________
Asst. Prof. Dr. Abdullahi Abdu IBRAHIM	Department of Computer Engineering, Altınbaş University	____________________
Asst. Prof. Dr. Serdar KARGIN	Department of Biomedical Engineering, İstanbul Arel University	____________________

I hereby declare that this thesis meets all format and submission requirements of a Master's thesis.

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Ali Faris Mansor AL-KHAFAJI

Signature

DEDICATION

First and foremost, I want to express my gratitude to Allah Almighty for providing me with the mental understanding, health, and stamina necessary to finish this course of study.

I would like to thank my parents, the people who deserve all the meaning of thanks and gratitude for supporting me from the beginning of my life to what I have reached now. Words will not be enough to express how they tried for me in all my life stages. Their grace and efforts are immeasurable, and I cannot repay them.

I sincerely thank everyone who supported me and helped me in my academic career especially my supervisor Asst. Prof. Dr. Oğuz KARAN. efforts and cooperation in supervising this thesis were instrumental in its successful completion.

In recognition of their unwavering support, I dedicate this thesis to my parents. Their continuous encouragement has been my driving force throughout my academic career, bridging the gap between old traditions and new opportunities.

PREFACE

First and foremost, I want to thank my Almighty Allah; without his blessing, I would not. Have been possible to complete the master’s degree.

I dedicate my humble efforts to my lovely Family Those love, encouragement, affection, and praying day and night make me capable of getting all the success and privilege.

ABSTRACT

FIRE AND SMOKE DETECTION BASED ON ARTIFICIAL INTELLIGENCE TECHNIQUES

AL-KHAFAJI, Ali Faris Mansor

M.Sc., Information Technologies, Altınbaş University,

Supervisor: Asst. Prof. Dr. Oğuz KARAN

Date: December/2023

Pages: 64

An efficient monitoring system is required for accurate fire and smoke detection to stop the fire and guarantee the safety of the occupants' lives. This calls into question the existing sophisticated capabilities of fire alarm systems and highlights the need for a thorough smoke and fire detection system. It's critical to detect smoke and fire. Given the time and accuracy needed for fire detection, convolutional (deep) neural networks have been developed to recognize objects, even though flames usually inflict significant damage. YOLO was applied, and a suggested method called YOLOv8 was also created. Still, much research has been done on using real data with deep learning. The suggested method was to use an image-rich Smoke and Fire database. The findings show that the suggested approach performs better than others in accuracy, model size, and detection speed. Consequently, we demonstrated how to apply the YOLOv8l algorithm as quickly as possible to identify important fire and smoke properties; identification is faster and more accurate than earlier methods; the YOLOv8 algorithm achieves an average @mAP of 96.6%.

Keywords: Artificial Intelligence, Convolutional Neural Network, Deep Learning, Image Processing, Fire and Smoke Detection, YOLOv8.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

AI	:	Artificial Intelligence
YOLO	:	You Only Look Once
DNN	:	Deep Neural Network
ANN	:	Artificial Neural Network
CNN	:	Convolutional Neural Network
SSD	:	Single Shot Detection
FASTER R- CNN	:	Faster Region-Based Convolutional Neural Networks
ML	:	Machine Learning
DL	:	Deep Learning
RGB	:	Red Green Blue
SVM	:	Support Vector Machine

1. INTRODUCTION

Effective detection, early warning, and combat techniques are severely hampered by multi-dimensional fire phenomena, their unpredictable locations, and diverse causes. Traditional fire detection systems have long relied on a variety of sensors to determine the presence of fire by examining signals such as smoke, flame, and heat. By observing numerous events, such as changes in temperature and smoke concentration in the air, this analytical technique aims to accomplish the critical task of separating fire from other environmental conditions. Heat, light-sensitive, and smoke detectors are all on the list of traditional fire detectors, and they have all been proven reliable and affordable over time. However, there are many challenges in the way of accurate fire detection, including limitations that still apply to traditional sensors. Their poor response in situations where smoke distribution moves slowly, often because of the relatively slow spread of smog, is a major drawback. Furthermore, the performance and lifespan of these sensors are constantly challenged by constant exposure to external factors, especially dust collection. Their limited applicability for fire detection in high-risk locations, such as wide-open spaces and outdoor locations, exacerbates these issues. Therefore, there is an urgent need to research and create more reliable and effective fire detection technologies. Up to 50% of all recorded incidents are related to fires, resulting in catastrophic losses in terms of loss of life and damage. Efforts to improve fire detection technology therefore require limitations that still apply to traditional sensors. A critical limitation is the role that mitigation plays in reducing these negative effects. In addition, this study highlights a crucial aspect of fire dynamics, namely the fact that smoke production is often an early sign of impending tragedy [1]. Detecting smoke during this crucial stage may open the door to rapid life-saving action. Fires pose a constant threat to people's lives and property and the continued functioning of institutions and industries against the backdrop of modern urbanization manifested in rapid economic growth. Since fires can occur suddenly and have the potential to cause widespread devastation, early detection and warning systems are crucial. Three main approaches now dominate the fire detection and alarm systems market: sensor-based methods, image processing methods, and deep learning methods. This thesis explores cutting-edge fire detection technologies and techniques to go beyond the limits of traditional sensors and improve early warning and battle tactics [2]. By doing so, we want to reduce the negative impacts of fire incidents, thus protecting the lives and property of people in our increasingly fire-prone environment. Early smoke detectors could not pinpoint the exact location of smoke, and firefighting equipment could only accurately detect it when the smoke was restricted. As a result, effective smoke localization has recently become the focus of computer vision tasks associated with smoke detection. Historically, most vision-based smoke detectors have relied on simple heuristic features [3][4]. These methods display smoke features such as color, velocity, opacity, and direction graphically. However, these feature-based techniques have been insufficient to characterize smoke flow externally, especially when ambient conditions vary, leading to poor performance of standard smoke detection algorithms. Improvement is required to increase generality and suppress interference in smoke detection algorithms. Some natural objects in forested areas may mimic smoke, which can confuse. Under normal circumstances, disclosure forms may be read incorrectly. Furthermore, smoke has diverse optical properties at different stages of combustion, making it difficult for detection algorithms to adaptively collect high-dimensional information corresponding to different stages. Since there are no feature-based techniques now to characterize smoke movement externally, a successful wildfire smoke identification system must be able to visually depict aspects such as sociality, velocity, opacity, and smoke direction. Conventional smoke detection systems struggle to adapt to changing operating settings [5].

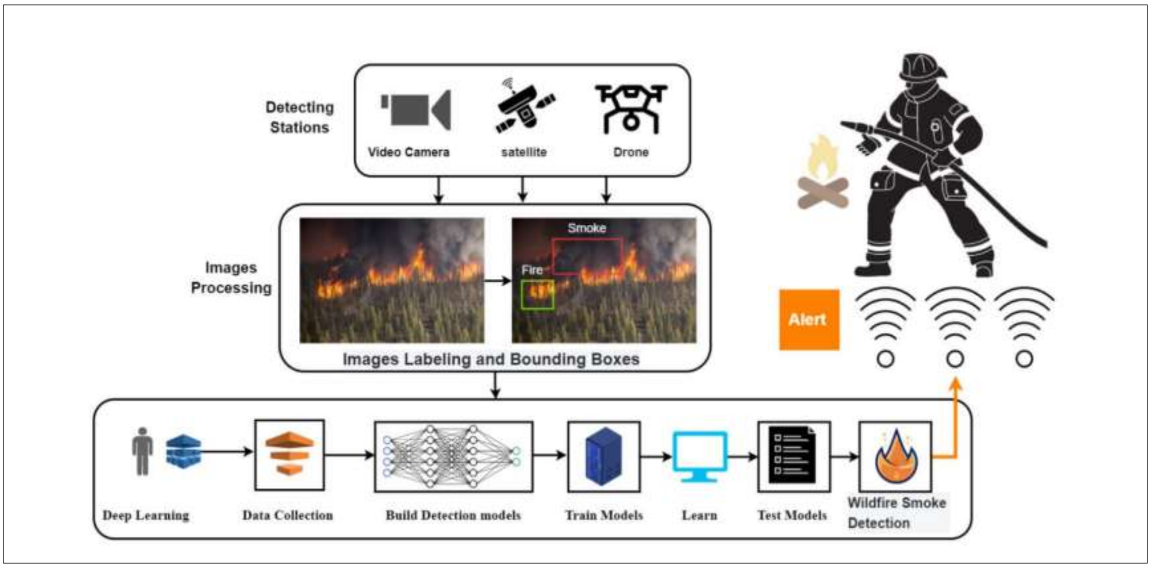


Figure 1.1: Shows a Basic Method for Detecting Wildfire Smoke Using Deep Learning.

still faces significant hurdles, particularly in crucial application areas that are gaining attention. One of the most significant challenges is image processing's poor speed and insufficient precision, where exact recognition and location accuracy are critical [6]. Traditional object identification approaches rely on the creation of features and architectures by hand. However, because of uncertainties caused by fluctuating illumination conditions and diverse views in application contexts, these approaches frequently fail to perform adequately. Fortunately, the fast improvement of computer systems has made it possible to apply object detection methods to deep learning frameworks, providing an efficient answer to these issues [7]. Achieving optimal performance and accuracy in experimental scenarios is a significant challenge when working with object detection algorithms [8]. Machine learning and various deep learning techniques, such as R-CNN and YOLO, are commonly employed for object recognition tasks within the realm of convolutional neural networks, aiding in the automation of object identification [9].

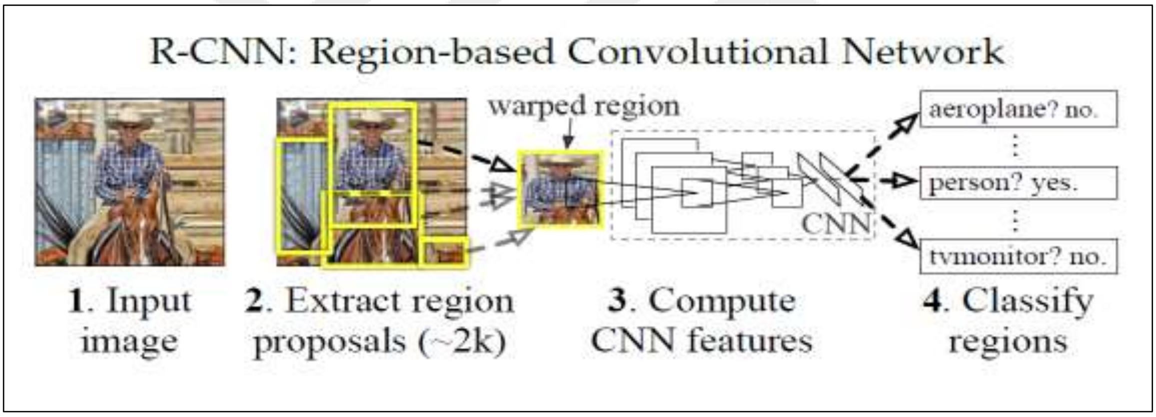


Figure 1.2: R-CNN.

According to reference, substantial datasets with Regression problems are where unsupervised deep learning is most frequently utilized [10]. The development of deep learning and machine learning gave rise to the idea of deep neural networks, or DNNs. An Artificial Neural Network (ANN) is what a Deep Neural Network (DNN) is between the input and output layers [11]. DNNs are taught to identify certain patterns in the job at hand and output the likelihood that data will be like the feed input. For instance, provided that the network has been trained on several dog breeds, predicting a dog breed from a picture. The user may examine the findings and choose from probability whether the network should show the recommended label and return it.

The deep neural network may forecast an outcome for a particular task and draw conclusions based on its prior experience, in addition to working by the algorithm. A DNN stays away from explicit programming for every encounter in this situation. It identifies that the proper input is transformed into the output by mathematical manipulation, regardless of whether the connection is linear or non-linear. Each output's probability is computed as the network moves through the levels. "Deep" networks are referred to as such since they consist of several layers, and each of these mathematical procedures is regarded as a layer.

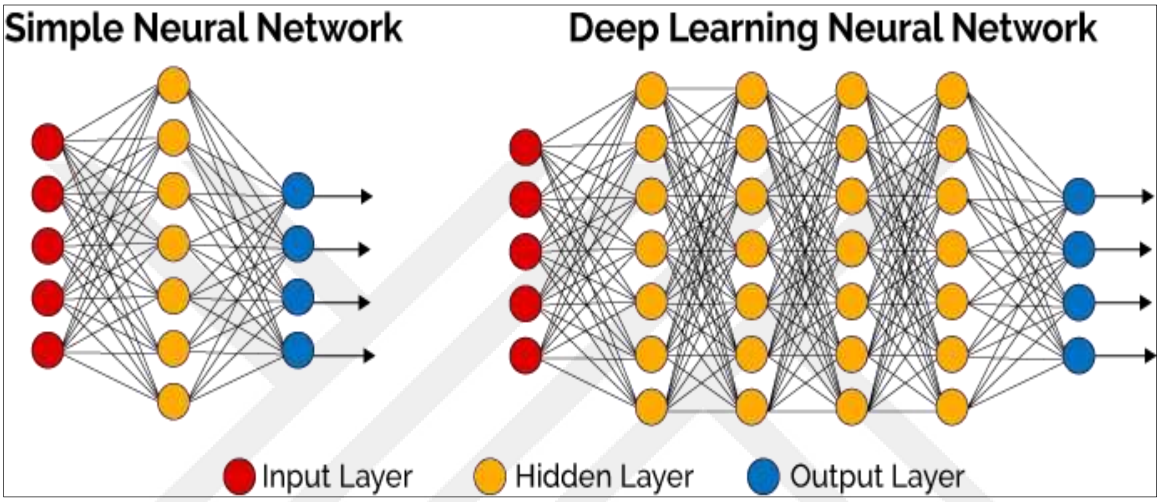


Figure 1.3: Deep Neural Network Conceptual Representation.

appropriate labels are essential for training object identification models, involving meticulous data set preparation and weighting, which can be a time-consuming and labor-intensive process [12]. Advanced object recognition approaches often employ transfer learning, which entails adapting a pre-trained network to suit the specific task at hand. Recent years have witnessed significant advancements in convolutional neural network (CNN)-based deep learning algorithms, introducing innovative approaches to fire detection. CNNs excel at extracting valuable information from images, diving deep into semantic data, and facilitating comprehensive model training, thereby eliminating the challenges and constraints associated with manual feature selection. For instance, [13]. In 2016 YOLO Redmont, and other object recognition algorithms were developed, indicating great progress in the field of intelligent object recognition. An important advantage of YOLO is its high processing speed, which enables real-time applications in augmented reality, self-driving cars, video surveillance, and road fracture detection [14]. YOLO, a method for real-time object detection that is widely applied. In various fields, including fire and smoke detection [15]. The depth-based YOLO model leverages features learned from a deep CNN to identify objects in both images and videos, making it an attractive choice for real-time applications due to its known speed and accuracy [16]. Between 2015 and 2019, YOLO was used primarily for live smoke and fire detection. However, after 2020, the focus shifted towards enhancing detection speed and reducing false positives. Some researchers have improved the previous YOLO method by using advanced optimization techniques to enhance recognition speed [17] [18]. In addition, some companies have chosen lightweight YOLO models to boost performance, while other research studies have explored optimization methods and shallow neural networks to achieve similar goals. Deep learning-based object detection techniques are now divided into two categories: single-stage methods and two-stage methods. Two-stage methods, such as R-CNN and R-FCN series, provide excellent accuracy in image detection but take longer to process and can be expensive to run [19] [20]. On the other hand, single-stage methods are more cost-effective and convenient, such as You Only Look Once (YOLO), Single Shot Detector (SSD), and Non-Convolutional Single Shot Detector (DSSD) [21]. This allows it to be deployed on embedded devices without compromising performance due to model compression. Different versions of the YOLO algorithm have been created over time, ranging from YOLO to YOLOv8. YOLOv8 was used for fire monitoring in this article. YOLOv8 by Ultralytics offers several improvements over its predecessors. YOLOV8 separates the input image into cells using a grid-based approach to solve the problem, enabling efficient object detection. Since its introduction, the Yolo algorithm has undergone continuous improvement, starting with version V1, and continuing through version V5 and the latest version, YOLOv8, released in early 2023 [22]. YOLOv5 takes advantage of newer network architecture and includes advanced layering and Swish activation techniques. And construction. Fusion engineering reduces complexity while increasing detection speed and accuracy. YOLOv8 includes five models: YOLOv8n, YOLOv8s, YOLOv8m YOLOv8l, and YOLOv8x, all of which provide powerful APIs through command line calls. There is no official document, although there is some accessible information about the Yolov8 architecture. Yolov8 is helped by its improved header hierarchy. In addition, it switches from object detection with anchors to object detection without anchors. By optimizing data and eliminating previous power outputs, YOLOv8s Up sampling PAN-FPN stage dramatically boosts performance. YOLOv8 has the best discovery rate among these options. Performance, providing the perfect combination of accuracy and computing speed. YOLOv8 is the best solution for general fire detection tasks due to its exceptional performance. As a result, this study presents a YOLOv8-based approach to fire detection as a universal algorithm for fire image analysis. The object identification method used by YOLOv8 makes use of the Darknet-53 architecture, which has a convolutional neural network as its core. A customized version of ResNet designed specifically for object detection tasks is called Darknet-53. Its importance for object detection cannot be emphasized, since it exceeds the YOLOv7 algorithm's prior backbone, which was based on the Darknet-19 architecture. Darknet-53 plays a key role in YOLOv8 by extracting features from incoming photos and passing those collected characteristics to a detection head that forecast bounding boxes and class labels for objects in the images. Notably, this detecting head functions without using anchors, therefore there is no requirement for established anchor boxes. In contrast, YOLOv8 uses pixel-level predictions to precisely establish bounding boxes. The crucial role of Darknet-53 is to improve the model's detection skills greatly in its capacity to extract complex and nuanced characteristics.

1.1 MOTIVATION

Promoting safety and preventing the terrible effects of fires are the main motivations behind and goals of this study in fire and smoke detection. More sophisticated research is needed in the field of fire detection because traditional techniques such as temperature checking and sampling have their limitations. Thanks to the success of deep learning methods, recent developments in computer vision, especially in areas such as object identification and recognition, have created new opportunities. This study's primary objective is to alter the YOLO (You Only Look Once) principle for fast and accurate fire and smoke detection. We focus on using the YOLOv8 model, which is one of the newest models in the YOLO family and known for its fast object detection. We want to enhance object detection to get the best results, especially in quickly detecting fires. Based on our study, YOLOv8 showed higher accuracy and a good balance between accuracy and computing speed. Therefore, the YOLOv8m version, which provides a high level of accuracy and excels at detecting even small flames, is the basis of our proposed model. By integrating this algorithmic model into the entire system, we can accurately predict the occurrence of fires, reducing damage to public fire resources.

1.2 PROBLEM STATEMENT

The current focus is on the urgent need to use deep learning methods to improve fire and smoke detection, with a focus on incorporating the YOLOv8 algorithm. Early and accurate identification of these risks is essential given the significant threats that fire and smoke disasters pose to human life and property. However, conventional methods are usually unable to guarantee rapid and accurate detection, especially at different sizes and in a variety of environmental conditions.

a. Increasing the accuracy of fire detection and providing the model with the ability to distinguish between fire and smoke well.

b. Enhancing the speed model's ability to quickly execute detection tasks, ensuring efficiency and timeliness.

c. Taking advantage of the inherent efficiency of the YOLOv8 method includes reducing neural layers to reduce parameters and improving generalizability.

d. This algorithm also includes an improved tracking mechanism that has great potential to enhance the recognition of small objects in long-range conditions.

1.3 OBJECTIVES OF THIS THESIS

The purpose of this thesis This initiative's main objective is to use deep learning algorithms to protect people from flames. This has a significant effect on the likelihood that individuals will survive right now. The emergency units' arrival This work makes use of deep learning and artificial intelligence in real-time self-monitoring. to locate smoke and flames. The goal of this project is to recognize smoke and fire in photos using a deep-learning algorithm. We tackle frequent issues in the fire and smoke detection industries with YOLO algorithm models. These challenges have already been imparted; the great bulk of criticism centers on the methods involved. Smoke and fire detection can be challenging, particularly in remote and huge areas. But other elements surfaced as significant barriers, including the ability of light to obfuscate the distinction between fire and itself. Therefore, the main objective of this study is to highlight the variations in. Performance and detection capabilities of several YOLO algorithm models, with an emphasis on the effectiveness of YOLOv8 in detecting smoke and fires in locations. Risks connected to fire Effectively handling these circumstances can greatly limit damage to persons and property.

1.4 CONTRIBUTION

The main goal is to improve fire and smoke detection using cutting-edge technologies that produce better and smoother results. As we move from artificial intelligence to deep learning, our goal is to achieve the best possible results under appropriate and acceptable conditions. Catley's important contributions, as detailed in this paper, are as follows:

a. We've compiled a library of images depicting typical fire scenarios, including forest scenes, highways, and high-rise buildings. This dataset is a useful tool for creating fire detectors, allowing training and evaluation of the proposed model.

b. The YOLOv8 model has proven its performance in fire and smoke detection and is currently the best fire detection method in terms of accuracy and speed. Experimental results support the importance of increasing the accuracy and speed of fire detection.

c. To compare several YOLO models - especially YOLOv8 - with other detection models, including Fast R-CNN, Faster R-CNN, and YOLO4v, we conducted a comparative study.

2. LITERATURE REVIEW

The numerous research projects that have been carried out in the field of fire and smoke detection are covered in this section. The advancement of artificial intelligence has aided in the development of fire detection models for use in fire and smoke detection systems over time. These models have shown improved speed and accuracy to minimize harm and preserve as many lives as possible. The effectiveness of fire detection techniques has significantly improved during the last few decades. This enhancement stands out considering the quick progress being made in computer vision thanks to machine learning and deep learning models. Convolutional Neural Networks (CNNs) were developed, and this has significantly improved detection performance. This chapter provides insight into the ongoing issues related to public blazes and continuing investigation. The introduction of the newest models for fire and smoke detection follows the presentation of pertinent solutions that have been developed to handle these issues.

2.1 PROBLEM DIFFICULTY

Smoke is usually the result of fire, and the two are often linked. The flame is constantly producing smoke, which may or may not start the fire. Even though these two materials are intertwined, they each have unique properties that can help us understand fire and smoke so we can build and engineer detection features. Environmental resilience systems are often used for fire and smoke detection in a variety of situations. Lighting conditions in the environment may range widely, from bright sunlight to total darkness. Smoke and fire detection systems need to work well in a variety of lighting situations, which may affect how visible smoke and flames are. Thermodynamics and Humidity How smoke and flames can be affected by ambient temperature and humidity conditions. Extreme temperatures can affect the sensitivity and reaction times of some detectors. Dust and Materials Airborne particles, dirt, and dust can interfere with sensors. Rapid Spread of Fires are likely to spread rapidly, especially under certain conditions or when accelerants are present. Limitations of sensors Optical detectors among other diverse sensing technologies Effective answers to these problems often combine sophisticated sensor technology, machine learning algorithms, and deep knowledge of the specific area in which the detection system is implemented. To improve the accuracy and reliability of fire and smoke detection systems while reducing false alarms, ongoing research and development activities are essential.

2.1.1 Features of Fire

A combustible substance that encounters an appropriate amount of heat and oxygen will ignite and produce a flame. The precise type of light it releases depends on the makeup of the burning material and intermediary processes. It typically emits light in visible, infrared, and ultraviolet ranges.

Shape Variation Techniques for detecting fire rely on an expert team's understanding of the velocity, color, and form of flames and that's not all. The flame frequently acquires a triangular shape, being wider at the base and tapering toward the top, depending on factors including wind direction, wind speed, and chemical combustion agents [23]. The geometry of the flame can also be influenced by the size and shape of the source where combustion occurs. The characteristics of a fire zone in a photograph might likewise vary across frames [24]. By examining changes in blob shape between two subsequent frames, a subject matter expert (SV) makes use of the fast changes in flame form. This approach, according, calls for the utilization of three distinct professionals who handle Viewing the same issue from several angles while accounting for factors like color, motion, and shape variety.

Color Soot, a black substance produced by incomplete combustion of organic components, contributes to the complex colors of fire. The reddish-orange firelight is produced by soot particles, which act as perfect blackbodies. With significant emissions in the visible and infrared spectrum, the gas emits a faint blue light once combustion is complete. The temperature of the blackbody radiation and the chemical composition of the emission affect the color. RGB color information is iteratively converted into mathematical spaces that partition brightness and hue to replicate human color perception [25]. The HSI (Hue-Saturation-Intensity) color model is one of the most convenient models for expressing colors in a way consistent with human perception. Shapes are represented by three colors in digital images [26]. Red, green, and blue levels (R, G, B). They correspond to color receptors in the human eye, each of which is sensitive to a wavelength. When combined, RGB levels allow color to be represented on a computer. Rule-based color models have been used by researchers such as Selek and Demirel to efficiently distinguish luminance from color, especially in classifying flame pixels. As the temperature rises, the dominant color of the flame changes from white, the lightest shade of organic matter, to yellow, orange, and red. Depending on the chemical composition of the source, the exact colors may or may not correspond to typical temperatures (red/orange/yellow/white). For example, barium nitrate flames burn bright green, a distinctive color often seen in wildfires.

Figure 2.1: Examples of Fire and Smoke in Various Environments.

2.1.2 Features of Smoke

In public fire footage, smoke is a key signal. Motion, color, and energy properties of smoke pictures are used as feature vectors and fed into a CSAdaBoost (Cost-sensitive Adaboost algorithm) classifier to perform smoke detection [27]. However, distractions like thick fog and moving objects that resemble smoke sharply reduce the accuracy of recognition. Along with the air it captures, smoke is made up of airborne materials and gases generated during combustion or pyrolysis. The kind of fuel and the conditions of combustion affect the amount and quality of smoke. Less smoke results from increased oxygen, particularly when complete combustion takes place. Despite improvements in smoke video recognition, wildfire movies are not well-suited to the present algorithms. Due to things like solar feedback and the presence of multiple moving leaves, they could set off false alerts. Additionally, stationery items like trees, Smoke, grass, and other bright materials and people's clothing can all be mistaken for one another. In order to address this problem, motion zones are identified from successive frames using an appropriate backdrop model, which lowers the likelihood of false positives caused by static distractions like blue sky and gray foliage.

Shape smoke normally has a dense form at the foot of its source and progressively disperses into the surroundings as it rises. When utilizing traditional detection techniques, this natural fluctuation in smoke size and form might lead to categorization errors. Due to buoyancy and wind effects, smoke plumes often ascend vertically or slant upward. They continue to grow in diameter, taking the form of an upside-down cone. The flow of the smoke's direction is constantly changing, which causes a dynamic development in the texture.

Color An important aspect of anticipating fire behavior is the color of the smoke that is released. Chen presented a color-based smoke detection technique in 2006 [28]. It consists of two decision rules: a static rule based on chromaticity and a dynamic rule based on diffusion. The former depends on the smoke's grey color, while the latter depends on how easily it spreads. The kind and intensity of the fuels present are revealed by the smoke's hue, which also offers hints about the fire's likely behavior. White smoke frequently signals the early stages of material consumption and is connected to the release of moisture and water vapor. It may also allude to the burning of combustible materials like grass or twigs. However, thick black smoke shows that heavy fuels have burned insufficiently. Occasionally used to denote the burning of man-made items like tires or buildings. Darker smoke often indicates a more volatile fire, but grey smoke may indicate a fire slowing down as it runs out of fuel. In further studies, the study of color and texture traits in moving areas has been used to identify smoke [29]. These areas are isolated by classifying pixels at the level of motion after segmenting motion according to frame differences, followed by background removal. Using a basic chromaticity-based static rule that relies on color thresholds and the presumption that smoke often displays grey hues, first-pixel categorization as smoke pixels is accomplished. For further improvement, more dynamic decision rules may be used.

2.2 INTRODUCTION TO AI

The goal of artificial intelligence is to create and comprehend intelligent beings. The swift advancement of technology and the widespread accessibility of digital data in several sectors have significantly transformed data patterns. This has led to the creation of novel algorithms and techniques to assess and tackle many contemporary issues. A subfield of artificial intelligence called machine learning enables robots to pick up knowledge from their experiences and surroundings. rather than designing them explicitly for specific use cases.

Deep learning is a type of machine learning that uses layers of artificial neurons to analyze massive amounts of unstructured data for tasks such as image identification and language processing. Technological discoveries have accelerated the development of new algorithms and solutions. As well as the availability of digital data. Consider current issues.

Machine learning, a subset of AI, enables computers to learn and develop on their own without explicit programming, it excels at creating sophisticated models, making them appropriate for difficult circumstances with traditional approaches, and it accelerates data analysis, discovering profitable possibilities or possible problems rapidly [30]. Combining machine learning with AI and cognitive technologies improves its ability to analyze large amounts of data efficiently. Many academics regard ML to be an essential component of AI. Learning is a fundamental aspect of human cognition, and ML models can replicate human cognitive capacities due to their plasticity [31]. ML comprises a variety of strategies for dealing with real-world problems. It allows computer systems to learn problem-solving instead of explicit programming. In identifying client demands in tweets, for example, the algorithm learns regular word patterns through training. samples. In machine learning, we distinguish between unsupervised, supervised, and boosting approaches. Unsupervised machine learning finds latent data patterns in the absence of a specified "correct" solution. Supervised ML learns from examples, allowing models to learn on their own without the need for explicit rule writing. Semi-supervised learning blends labeled and unlabeled data to provide more detailed insights.

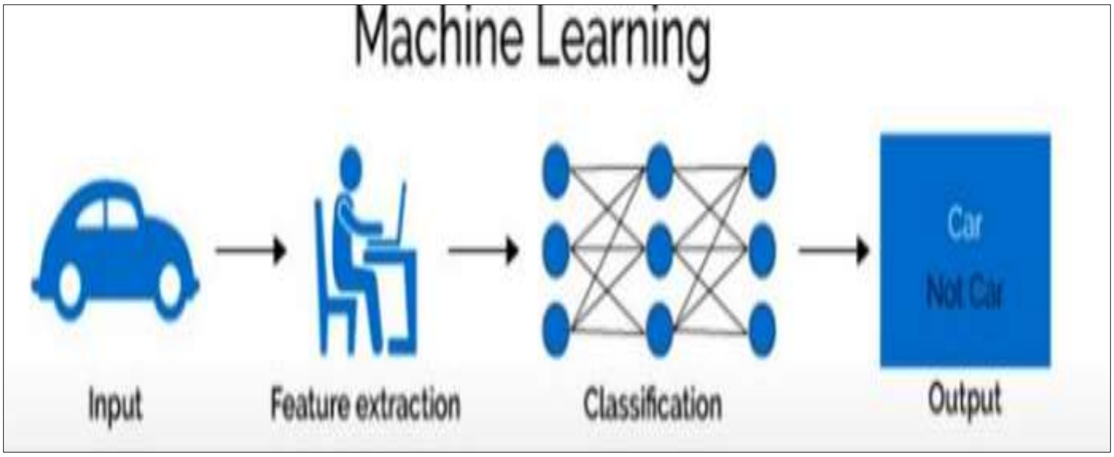


Figure 2.2: How Does Machine Learning Work.

Deep learning is a machine learning technique that combines representations and mapping functions and integrates these concepts into a hierarchical system. This can be achieved by studying computational models or graphs at different processing stages with the aim of learning representations for data with multiple abstract layers [32]. Deep learning models are capable of learning Raw data can be used to discover low-level features as well as obvious and distinct features. As the complexity of the data being processed increases. It was an intense learning experience. It has been utilized in fields like machine translation, bioinformatics, medication design, computer vision, speech recognition, natural language processing, voice recognition, social network filtering, and medical picture analysis, materials screening, and drawing game software, with similar results in some Domains cases, Preferred over human specialists [33].

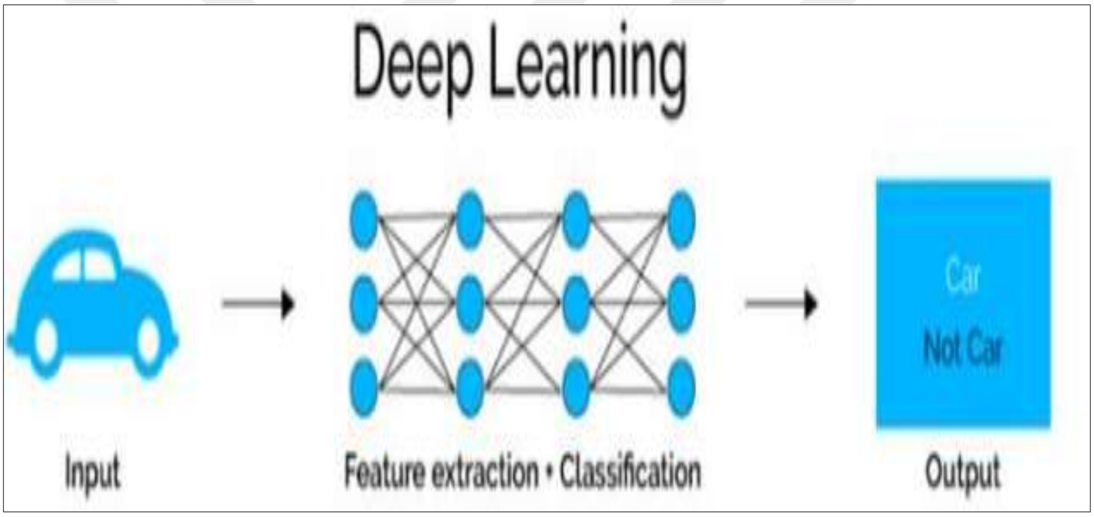


Figure 2.3: How Does Deep Learning Work.

Its latest division has also been likened to machine learning. Deep learning has already arrived. It is the result of numerous surveys and research and development in the field of machine learning. Explain how artificial intelligence, machine learning, and deep learning are related. The field of artificial intelligence is broad .It is a broad field that developed in the middle of the last century and witnessed the growth of artificial intelligence. He helped develop machine learning, which evolved into deep learning. It has many parallels in the field of object recognition and, more recently, investigations in the field of fire detection.

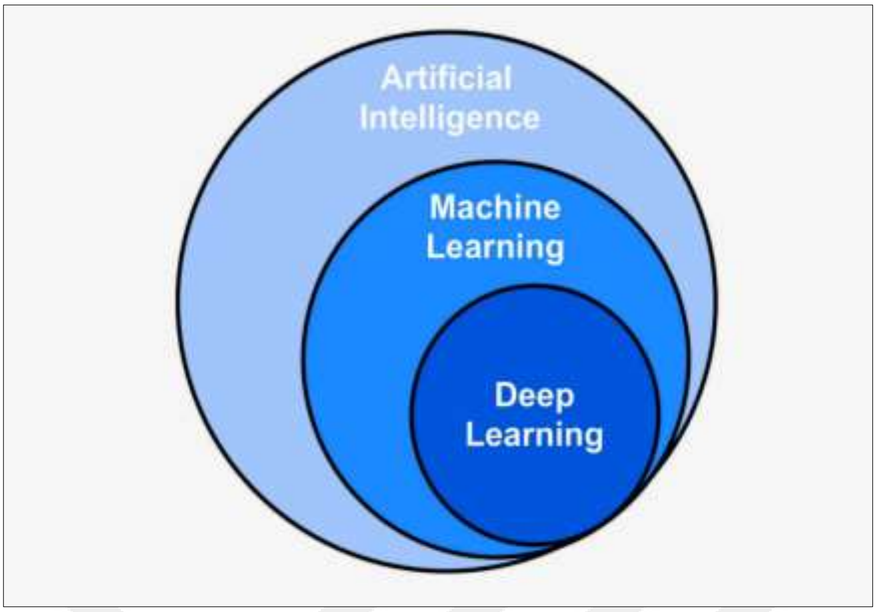


Figure 2.4: The Relationship Between AI, ML, and DL.

Computer vision One of the most essential artificial intelligence techniques is computer vision. Computer vision is an interdisciplinary scientific topic of computer science that focuses on how computers may be made to capture human visual skills at a high degree of comprehension. For digital picture or video content [34]. It is thought that this sector can only operate at a limited capacity due to the lack of large-scale computations and data. Recent advances in artificial intelligence, machine learning, and neural networks, in addition to the amount of data available, have proven successful for computer vision to capture a large amount of attention A leap into outperforming humans in some object detection and labeling tasks. This field has had a lot of success in fixing difficulties. pertaining to visual character recognition (OCR), machine inspection, the creation of 3D models, medical imaging, the safety of vehicles, motion capture (mocap), surveillance, biometrics, and fingerprint recognition. Computer vision is used to tackle a variety of difficulties. Some of these issues are depicted in the image below.

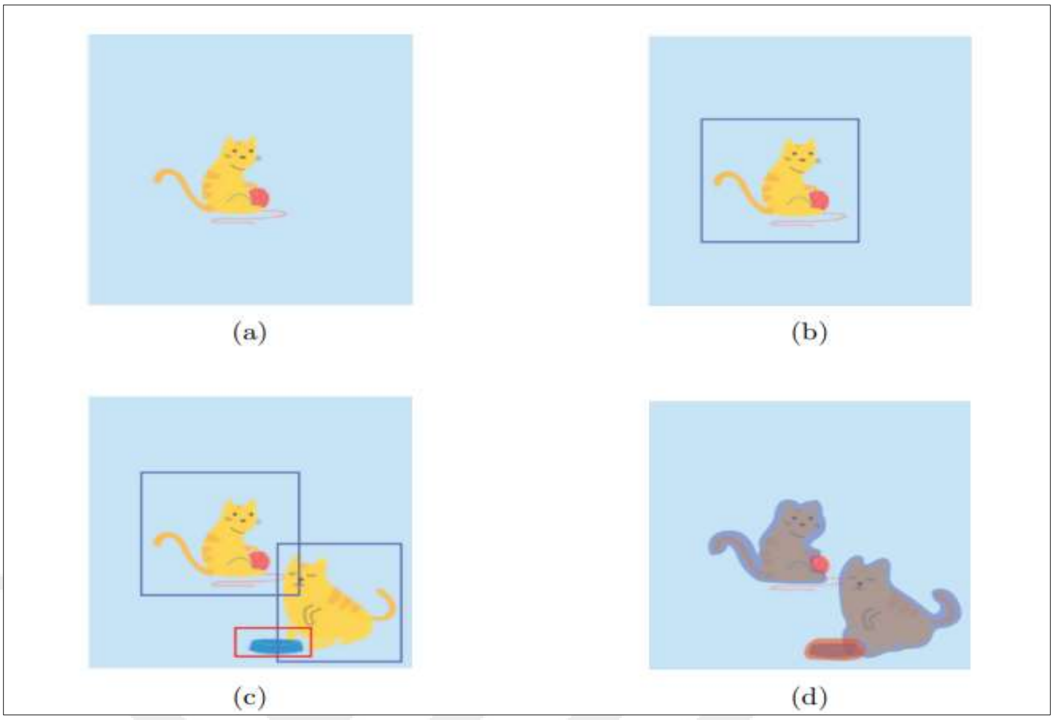


Figure 2.5: Types of Computer Vision Issues.

Figure depicts four distinct types of computer vision challenges:

a. Classification
b. Classification + Localization
c. Multi-Object Detection
d. Semantic Segmentation

2.3 CONVOLUTIONAL NEURAL NETWORKS (CNN)

One basic kind of neural network that is used for tasks like object identification, image recognition, and picture categorization is the convolutional neural network, or CNN for short. Most object recognition models in deep learning are built on convolutional neural networks, which are frequently employed in this field. Before classifying a picture, CNN image classification processes it Should there be any animals in the photograph, the result will be (dog, cat, or horse). The computer interprets the input picture as a pixel inside an array, with the image resolution having an impact on the array of pixels, h x w x d. The computer interprets these three letters as (w = width, d = dimension, and h = height). For example, a picture of size 4 by 4 by 1 denotes a grayscale image array of the matrix since 1 denotes the image's single color, which is gray; in contrast, an image of size 6 by 6 by 3 denotes a color image since RGB stands for (Green=G, Blue=B, and Red=R) A CNN model with deep learning, such as SSD or YOLO, uses a sequence of convolution layers to process each unique image as it trains or tests. This convolution layer has filters (Kernels) and a fully connected layer. Using pooling, CNN will recognize an object inside the image with values between 1 and 0 by using the SoftMax activation function [35] [36]. The entire convolution layer procedure and input picture categorization are shown in the chart below.

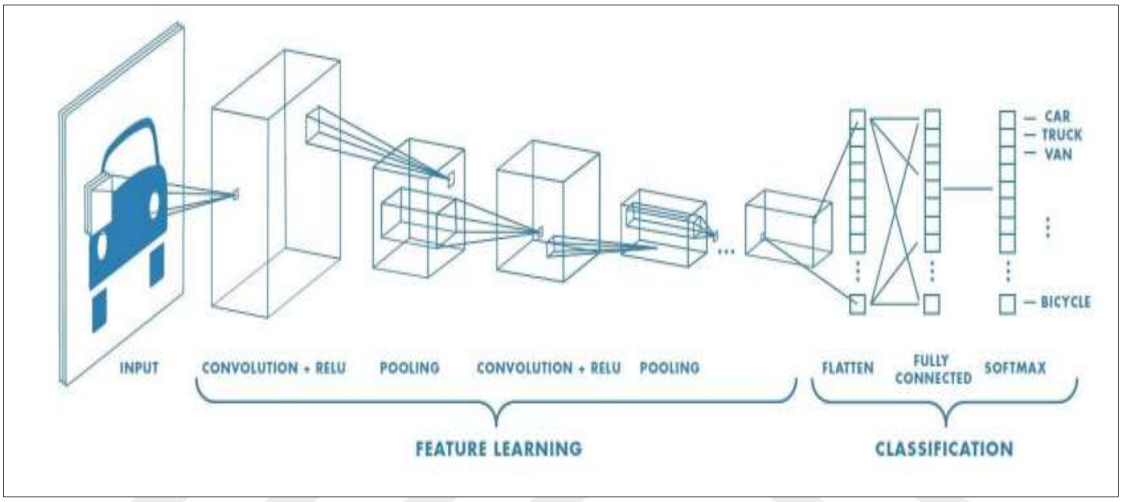


Figure 2.6: Workflow of a Neural Network.

2.3.1 Convolution Layer

A crucial layer in a convolutional neural network is the convolution layer. Its objective is to take features or weights out of pictures that are generated by both linear and nonlinear methods. Convolution is a method for optimizing image quality. The output value, not the output tensor, is what is obtained when a certain set of numbers, called a kernel, is applied to the input aggregated to be called tensors. The kernel components and input elements define the map [37].

2.3.2 Pooling Layer

In charge of decreasing the internal size of maps and features to preserve data and any potential small feature distortions or transformations, as well as lowering already-existing features that are subsequently learned [37].

2.3.3 Fully Connected Layer

It is important to note that the output features of the pooling and convolution layers are filtered so they are completely connected to the layers, turning them into one-dimensional integers and connecting them to a single number. These layers, which connect inputs to outputs with learnable weights, are also referred to as dense layers [37]. Features are now transformed after being gathered from the layers. Results of convolutional neural networks via a set of linked sublayers. The number of layers is often the same as the final layer contract that is produced [37].

The picture is mostly a matrix composed of different values for each pixel. A grayscale image, for instance, is two-dimensional and has one channel with pixels with values ranging from 0 to 255. Other images or matrices include many channels, such as RGB images, which have red, blue, and green channels. A multicolored picture, also known as an A matrix with three channels red, blue, and green represents an RGB picture. As seen in Figure 2.7, each channel, or layer, is composed of pixels with values ranging from 0 to 255.

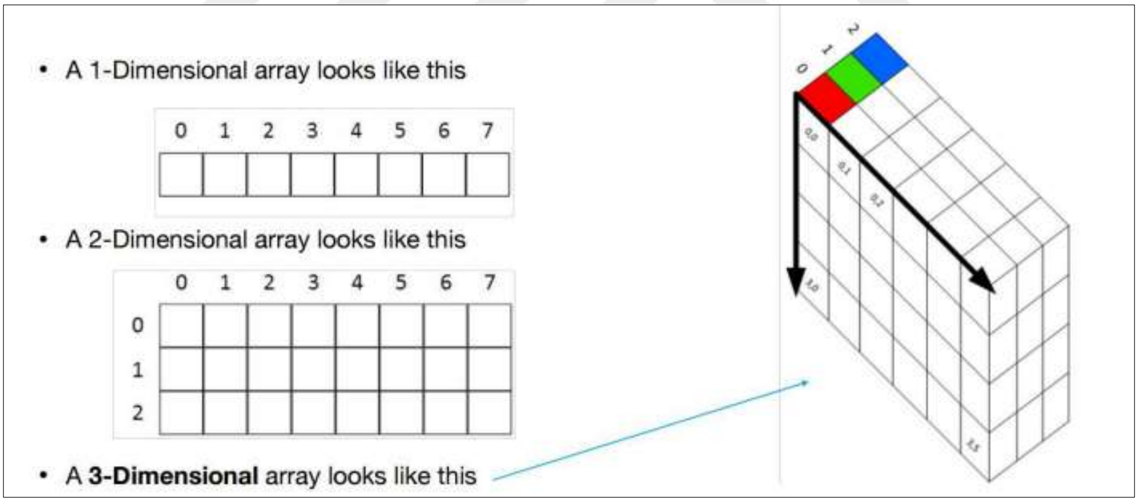


Figure 2.7: Channels for Images.

2.4 TECHNIQUES FOR DETECTION BASED ON COMPUTER VISION

For use in applications like ships, aircraft hangars, and tunnels, computer vision algorithms have been developed for automatic video fire or smoke detection. They focus mostly on cramped, congested areas. The development of trustworthy video fire detection systems in big or open spaces has received a lot of research attention. Many computer-vision-based fire detection systems operate by separating out color and motion, two crucial characteristics.

2.4.1 Detection of Color and Motion

The RGB color model extracts fire information from images by analyzing the color components of pixels, which are red (R), green (G), and blue (B). A pixel is classed as "fire" if R exceeds a predetermined threshold (RTH) and the condition R > G > B is met. This model emphasizes the red (R) channel's dominance in fiery colour detection, where RTH is critical in defining the detection threshold [38].

$$\mathrm{R} > \mathrm{RTH} \tag{2.1}$$

$$\mathrm{R} > \mathrm{G} > \mathrm{B} \tag{2.2}$$

The R component is considered the major one for fire detection; hence, it is required that R have a defined threshold. As a result, a mask may be produced and used to identify the fire pixels in a picture. For constructing the mask, this approach employs the following formulas. This approach had an 82% detection accuracy.

$$\mathrm{R1}(x, y) = (\mathrm{R}\ (x, y) > \mathrm{G}\ (x, y))\ \&\&\ (\mathrm{G}\ (x, y) > \mathrm{B}\ (x, y)) \tag{2.3}$$

$$\mathrm{R2}(x, y) = (\mathrm{R}\ (x, y) > 190)\ \&\&\ (\mathrm{G}\ (x, y) > 100)\ \&\&\ (\mathrm{B}\ (x, y) < 140) \tag{2.4}$$

$$\mathrm{M} = \mathrm{R1}\ \&\&\ \mathrm{R2} \tag{2.5}$$

The color models are RGB and HSV. The flicker of the flame is considered in this detecting method [39]. By generating a luminance time-derived matrix and applying the RGB and HSV fire color models across the color cumulative matrix, one may identify pixel locations that exhibit a high flicker frequency. The previously determined pixel areas are combined, and spatial color variation is assessed to filter out objects with comparable properties to fire. The time color variation on the indicated regions is assessed to filter out things having comparable properties to fire but a consistent time color variation. The result will then be classified as either positive or negative detection. This approach detected fire with 94% accuracy in the provided frame.

Yi (x, y) denotes the brightness of pixel (x, y) for frame 'i', and the luminance time derivate matrix is computed as follows:

$$\mathrm{Di}\ (x, y) = |\mathrm{Yi}\ (x, y) - \mathrm{Yi}{-}1(x, y)\ | \tag{2.6}$$

The cumulative time derivation is computed using:

$$\mathrm{CTi}\ (x, y) = \alpha \mathrm{CTi}{-}1(x, y) + (1 - \alpha)\ \mathrm{wi}\ (x, y)\ \mathrm{Di} \tag{2.7}$$

Fire Color Models (RGB and YCbCr): This approach identifies fire pixels by employing both the RGB and YCbCr color models. Y indicates luminance in the YCbCr color space, whereas Cb and Cr represent chrominance blue and chrominance red, respectively. In the preparation step, images are captured using a still camera, noise reduction filters are used, scaling occurs, and mathematical calculations are used to build a probable fire pixel mask in the segmentation stage. In the classification stage, this mask is used to extract observed fire zones in each frame. The strategy, which entailed discussing the color ranges and formulae used to generate the mask achieved 89% detection accuracy [40]. Traditional vision-based approaches rely on manual feature extraction, which includes elements such as color, motion, texture, and form.to recognize fire or smoke. Initially, Celik presented a color-based fire detection system that distinguished luminance and chrominance information using the YCbCr color space [41]. This approach combines the RGB and YCbCr color spaces to provide more exact fire detection by following rules to identify an image as 'fire,' resulting in a binary representation of the discovered fire region.

2.5 MACHINE LEARNING-BASED DETECTION TECHNIQUES

Through the application of non-linear functions, machine learning algorithms are employed to build a relationship between unusual data and its components. These algorithms are trained on data to internalize the features of the data, using either supervised or unsupervised learning techniques. Several fire and smoke detection systems were researched and are detailed below.

In research conducted a simple motion detection method was used, followed using the irregularity property of smoke as a feature set. Other factors, to create feature vectors, more elements were added, such as the smoke area, the border of the smoke, the number of smoke blocks, and the foreground area. The last step was using a Support Vector Machine (SVM) categorization [42]. Even though this approach has been reported to achieve good results, the issue of false alarm rates must still be addressed. developed a method for detecting fire and smoke using quaternionic wavelet characteristics. Moving objects with positive vertical velocity are detected as probable fire and smoke zones in this technique. Furthermore, quaternion Gabor wavelets are used to extract local spectral, spatial, and temporal features of fire areas.

A 1D temporal filter kernel is used to capture random flickering activity within the fire zones to reduce false alarms. Entropy, Contrast, Angular Second Moment, Inverse Difference moment, and Image Pixel Correlation are five texture descriptors based on GLCM (Gray-level Co-occurrence Matrix). Furthermore, the mean value of the potential smoke patches is included in the feature set used to train an SVM model [43]. Even though this method is computationally demanding, testing data across a variety of settings shows that it outperforms some of the earlier techniques. In Ref Timothy suggested a multi-spectral video-based technique for detecting early forest fires. They identified smoke plumes using multiple forms of imaging (RGB, mid-wave infrared, and long-wave infrared). The smoke plume's temporal and spectral properties were analyzed using multi-spectral and multi-temporal Principal Component Analysis (PCA) across a series of video frames [44].

The spatial texture of a plume in the higher-order main components was examined using statistical descriptors from the Gray Level Co-occurrence Matrix (GLCM). PCA effectively segregated smoke plumes, and the correlation texture descriptor best recognized their presence. The third stage was to categorize the smoke using a threshold correlation of spatial texture. The research did not include any quantitative or qualitative data. evaluation of the algorithm's performance. presented a real-time method for detecting fire flames and smoke in videos that relied on foreground image accumulation and optical flow." The method is divided into four steps: extracting moving pixels and regions using frame differential analysis, identifying smoke candidate regions using static color models, creating foreground accumulation images for smoke detection, and calculating smoke motion features using block-based image processing and optical flow. The approach assumes that smoke is normally grey, with about equal values for the R, G, and B components in smoke pixels.

To extract motion characteristics, optical flow methods are applied, which reduce computing complexity by treating core pixels in smoke blocks as feature points. A propagation neural Network is used to classify smoke features. However, this is not the case. When the smoke is not grey in hue, the procedure may not be performed properly [45]. suggested an outdoor smoke detection system that is resistant to changes in the environment during the day. They utilized a Gaussian Mixture Model (GMM) to estimate the background, the Integral Image approach to extract Haar features from probable smoke regions, blurring and morphological image processing techniques, and the AdaBoost algorithm to recognize smoke objects [46].

The method was said to be computationally efficient and successful in identifying outside smoke, although it was evaluated on a small number of video clips and only qualitative performance indicators were supplied. sought to build a false alarm reduction system by analyzing multi-source data, such as meteorological geography information and visual and infrared camera data. By combining Wireless Sensor Network (WSN) approaches with Backpropagation Network (BPN), Radial Base Function Network (RBFN), Dynamic learning Vector Quantization (DLVQ), and Multi-Layer Perceptron (MLP), the study achieved a detection rate of more than 98% and a false alarm rate of 1.93% [47]. used data analysis to identify forest fires and residential fires by evaluating factors such as the Canadian Fire Weather Index (FWI) fire temperature, ionization, photoelectric CO gas, and other technologies. The data was analyzed using Distributed Artificial Neural Networks (ANN) and Naive Bayes. Their investigation found that residential fire detection was 81% accurate and forest fire detection was 92% accurate [48]. Reference used accessible data to forecast forest fire behavior. Their study found that the following models had the highest accuracy: AdaBoost (91%), Random Forest (RF, 88%), Artificial Neural Network (ANN, 86%), and Support Vector Machine (SVM, 81%) [49].in Ref undertook research to forecast the occurrence and magnitude of forest fires by evaluating a worldwide dataset of burnt areas for the year 2015. Their findings suggested that the following machine learning models performed well: XGBoost (94%), logistic regression (81%), RF (89%), and Multilayer Perceptron (MLP, 90%) [50]. Ref investigated ANN and logistic regression models for forecasting human-caused fires in a comparative analysis. Their study produced an ANN using Geographical Information System data. For no-fire correct prediction, the accuracy is 85% accuracy [51]. Diyana created a technique for monitoring forests to spot fires early. employing UAVs equipped using cameras that are thermal, optical, or both. Rotary-wing drones fly at low altitudes to confirm potential fire locations whereas fixed-wing drones soar at intermediate altitudes. Their system, which was constructed on the SSD model using MobileNetv1 as the backbone and COCO dataset weights, achieved 94% accuracy, triggering warnings for authorities upon fire detection [52]. employed Sentinel satellite images and a Random Forest classifier to map fire severity, reaching an amazing 98% accuracy [53]. performed research on the geographical performance of forest fires by taking weather, vegetation, and infrastructure into account. They used the MARS-DFP classifier, and their study yielded an accuracy of 86.5% accuracy [54]. with the highest performance in predicting forest fire occurrences with an accuracy of 0.809, the frequency ratio-logistic regression 0.909, the frequency ratio-multilayer perceptron 0.858, the frequency ratio-classification and regression tree 0.847, and the frequency ratio-support vector machine 0.959 were the next highest AUCs [55].

2.6 TECHNIQUES FOR DETECTION BASED ON DEEP LEARNING

The many studies conducted to create models for smoke and fire detection systems are described in this section. Numerous studies have been conducted to determine if fire or smoke is present in photographs using machine learning and deep learning models as artificial intelligence has advanced. However, we examined CNN-based fire/smoke detection models in this work. Numerous areas of machine learning have advanced significantly since deep learning was introduced. Numerous fields, including voice identification, natural language processing, and image object detection and classification, have successfully employed deep learning approaches. Researchers have done several investigations on fire. and smoke detection based on deep learning to boost productivity. A portion of this study has been gathered and examined to enhance and recognize the problem.

The development of CNNs has resulted in considerable performance advances in a variety of computer-based vision uses, including image identification and picture categorization. LeNet-5, a CNN algorithm, obtained one of the first effective results in this field by identifying hand-written characters [56]. Researchers have recently been able to do large-scale dataset analysis thanks to the development of very potent GPUs and the accessibility of massive datasets. to produce extraordinarily deep CNNs. For example, created AlexNet, a deep CNN network that excelled in the 2012 ImageNet Challenge [57]. Furthermore, various CNN variants have demonstrated outstanding performance in picture categorization [58]. In the authors described a method for detecting forest fires by deploying drones to gather photographs of the region, which were then examined with a basic CNN and the YOLOv3 model. The accuracy and speed of our deep learning-based UAV fire detection system exceeded expectations, with an algorithm recognition accuracy close to 83% [59]. YOLOv5 was upgraded to integrate adaptive stabilization training with K-means++ technology, which improved performance and speed of detection while minimizing fire damage. Three YOLOv5 models—small, medium, and big—with different loss functions, CIoU and GIoU, were examined by the authors. Using a synthetic technique, they grew their dataset from 4,815 to 20,000 pictures. The revised model obtained an, especially when employing the CIoU loss function. The average accuracy and recall were 86.0% and 78.0%, respectively, outperforming the baseline YOLOv5 by 4.4% on average [60]. showed off a technique for utilizing an enhanced YOLOv3 model with a larger detection area and faster detection speed to detect fires 24 hours a day, even at night. The study stressed the significance of reliable data collection for successful deployment They added 9,200 photographs from diverse sources to their database and used data augmentation techniques such as image rotation. Custom cameras and the YOLOv3 real-time fire detection model were used in the technology, which achieved an excellent average accuracy of 98.9%. Nonetheless, differentiating non-fire sources of illumination, such as incandescent light or high-beam lights, remains a difficulty Deep learning-based techniques for fire detection in smart cities have recently gained popularity in several instances, hybrid techniques incorporating different deep-learning fire detection algorithms have been developed [61]. For example, Al-Turjman suggested a hybrid strategy for fire detection in smart cities that incorporated CNN with recurrent neural networks (RNN) The suggested method obtained great accuracy while having a low false alarm rate [62]. Huang suggested a YOLOv3-based technique for fire detection in outdoor situations in another investigation. This method was tested on a real-world picture dataset and obtained an accuracy of 92.8% Light-YOLOv4, a portable, real-time smoke and flame detector, was introduced [63]. For starters, they employed a lighter network backbone compared to YOLOv4. Second, cross-scale relationships that are bidirectional were included. Ultimately, they divided the convolution and computed the channel and geographic area separately, all enhancements above YOLOv4. Light-YOLOv4 detected flames with an accuracy of 86.43% and smoke with an accuracy of 84.86%, respectively. Additionally, it demonstrated a mean average precision of 85.64% for flame and smoke detection at a confidence threshold of 0.5 and a mean average of 70.88% for smoke detection at the same confidence level, respectively [64]. A framework was used in similar research to analyze several iterations of YOLO, such as YOLOv3, YOLOv4, and YOLOv5, to pinpoint the best locations for UAV landings with the goal of reducing system failures and enhancing safety The findings demonstrated the better performance of YOLOv5 with large network weights over its rivals using a database of 11,268 satellite photos with an accuracy rate of 70% In addition, YOLOv4 performed better than YOLOv3. Additionally, YOLO models have shown success in the detection of illnesses like cancer using image processing [65].

An alternative strategy was suggested to enhance the ability of YOLOv5 to detect breast cancer. Using 10,239 photos from the CBIS-DDSM, the small, medium, huge, and extra-large YOLOv5 weight models were evaluated. using drones and video cameras. The improved YOLOv5x model outperformed the others, obtaining an MCC of 93.6%, demonstrating the promise of these strategies for detecting corpus. The upgraded YOLOv5m model outperformed rivals like YOLOv3 and Faster R-CNN with an accuracy of 96.5% and mAP of 96% [66]. Using the YOLOv7 model as your foundation, suggest a better fire and smoke detection technique with an 8520-image diversified dataset. We switched out the E-ELAN module for the partial convolutional module to increase the training pace of the YOLOv7 model.

Additionally, we address the issue of excessive gradients during training brought on by poor-quality data with significant errors by using the Focal-EIoU loss function on the collected dataset from the actual world, the model is verified. According to experimental findings, the proposed technique improves FPS by 38% while achieving precision for fire and smoke detection of 85.5% and 85.7%, respectively [67]. The YOLOv7 model was improved in this study to identify smoke from forest fires in challenging wilderness environments. The innovations, such as SPPF+, BiFPN, and decoupled heads, resulted in considerable performance increases. An AP of 78.6%, AP50 of 86.4%, and AP75 of 81.7% were obtained by the optimized model, representing improvements of 2.7% in AP, 3% in AP50, and 5.5% in AP75. With an AP50 of 83.4%, CBAM consistently beat ECA and SE in the ablation investigation on attention processes. On a specially created smoke picture dataset, experimental findings showed that the improved YOLOv7 model beat cutting-edge and multi-stage object detection models, obtaining an AP50 of 86.4% and an APL of 91.5%.

With an AP50 of 82.5% and an APL of 87.3%, YOLOv7 fared better than average. Especially, conventional The YOLOv7 technology has been improved to overcome the limits of previous forest fire smoke detection devices, which could only identify one fire at a time and were restricted to tiny, confined spaces [68]. The use of YOLOv8 algorithms for the detection and identification of forest fires using drone photos is explored in this research. The performance of numerous YOLO models is examined in the study, with special attention paid to the YOLOv8n model, which showed greater performance. The training picture dataset will be improved and expanded in the future, along with the YOLOv8n model training settings and parameter optimizations for batch size and anchor value. The study's outstanding 95.7% Mean Average Precision (MAP) accuracy was attained [69].

3. METHODOLOGY

3.1 YOLO ARCHITECTURE

YOLO functions as a real-time object detection system and was first released in 2016. Regression is how the YOLO algorithm finds objects. It processes the input image to categorize the item and get its bounding box coordinates. The typical components of YOLO architecture are the Head, Neck, and Backbone. These are their roles. a network built to take information out of pictures and use it for other purposes. Located between the Head and the Backbone, the Neck is essential to feature fusion and allows for better use of characteristics that the Backbone extracts. The Head uses the characteristics that were retrieved to recognize the item. The YOLO algorithm's original Backbone network topology had two completely linked layers after every 24 CNN layers. The CNN layers serve the purpose of the image's characteristics, and the ensuing fully linked layers forecast the output coordinates and probabilities. The Backbone network design of YOLO has undergone significant changes because of the technology's developments. Furthermore, the algorithm forecasts the likelihood of an object being found within the box that borders. A grid cell oversees detecting an object if its center is inside it, as shown in the picture. There will be many bounding boxes for each grid. During training, four descriptors can be used to identify each bounding box. Sets the width and height of the element's center enclosing box to the category to which it belongs.

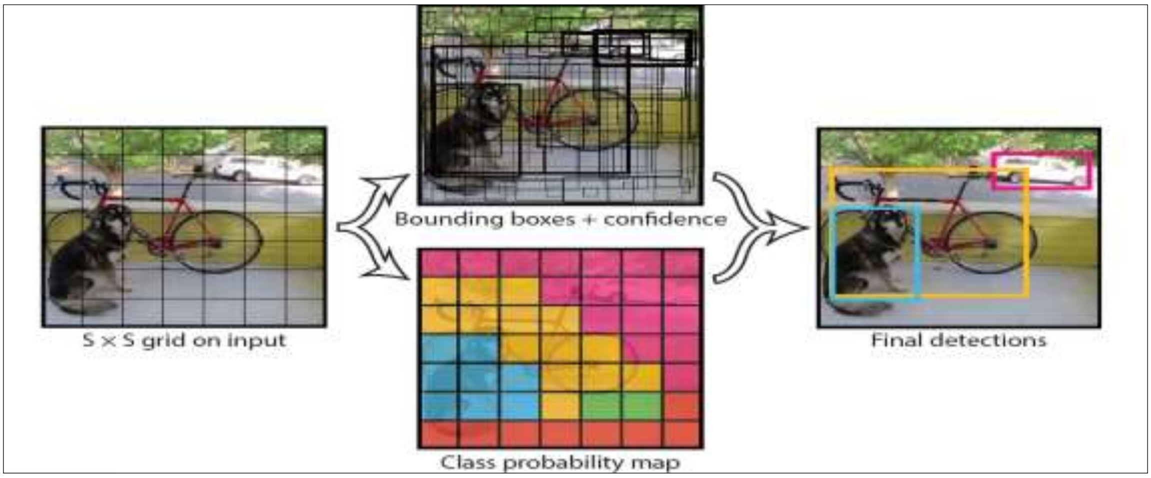


Figure 3.1: An Illustration of How YOLO Predicts on a Single Network and How the Final Categorization is Realized by Combining the Numerous Observed Segments.

Redmon unveiled Darknet-53 in 2018, a network comprised of 53 CNN layers. The goal of this setup was to extract complex characteristics to solve the problem of recognizing tiny targets while maintaining a balance between processing speed and accuracy. In contrast, the fully linked layers that follow predict the output coordinates and probabilities, to extract the image's characteristics. The design of YOLO's Backbone network has undergone significant changes because of the technology's developments Redmon presented Darknet-53 in 2018, which was made up of 53 CNN layers [70]. It was intended for this setup to extract complex characteristics, successfully solving the problem of identifying tiny objects while juggling accuracy and processing speed. With YOLOv8 being the newest member of the YOLO family [71]. The YOLO traffic timeline is shown in Figure.

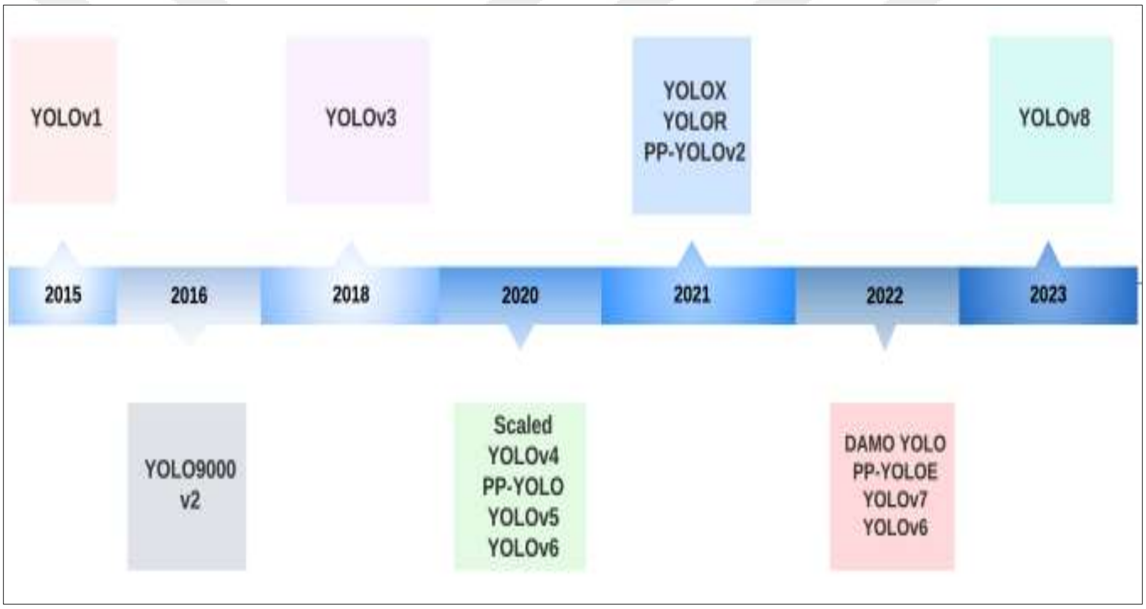


Figure 3.2: Below is a Glossary of Important YOLO Architectural Keywords.

3.2 YOLOV8 OBJECT DETECTION NETWORK

The conventional approach of feature extraction for fire and smoke detection was the subject of several studies. The primary issue with such strategies was the time required to compute these feature extractions. As a result, real-time fire and smoke detection was sluggish and performed poorly. These techniques also produced several false positives and errors in background detection. We proposed employing the YOLOv8 model for fire and smoke detection in response to the recent advancement of deep learning algorithms. Imaging processing is more quickly completed for fire and smoke detection. The ideal method to deal with the detection of these items in this situation is YOLOv8. By labeling the practice photographs, the model was then created using a neural network. YOLOv8 layers, too. To evaluate the effectiveness and accuracy of the YOLOv8 approach, we also trained, verified, and tested it.

The most recent installment of the well-known "You the Only Look Once" series, YOLOv8 was released by Ultralytics on January 10th, 2023 [71]. An improved model of the YOLO (You Only Look Once) series is the YOLOv8 model created by Ultralytics. It is intended to be quick, precise, and simple to use. To enhance performance and versatility for tasks like object detection in pictures, image segmentation, and object classification, this new version builds on earlier versions. Owing to its quicker and more accurate object identification capabilities, YOLOv8 is the favored option for a variety of computer vision applications. Developers may easily adapt its user-friendly framework to their projects or datasets. This model is adaptable, enabling adjustments to better fit various needs and circumstances, and it's made to be simple to use even for people with no prior experience.in computer vision and machine learning. We observe that the model varies amongst us in terms of speed and shape. YOLOv8 represents a significant development in the fields of object identification, picture classification, and instance segmentation by building on the foundation set by the successful YOLOv5 model. Despite the lack of a publication, the community debates and the accessible repository offer important insights into the revolutionary improvements made possible by YOLOv8.

The Backbone, Neck, and Head are the three key architectural elements that distinguish the YOLO framework from its forerunners, and YOLOv8 follows suit [72]. As shown in the illustration, these parts cooperate, with the Head carrying out the prediction of item positions and classifications and the Neck integrating this information. YOLOv8 Using an anchor-free model instead of an anchor-box model method featured in prior YOLO models, is one of the standout aspects of YOLOv8 [73].

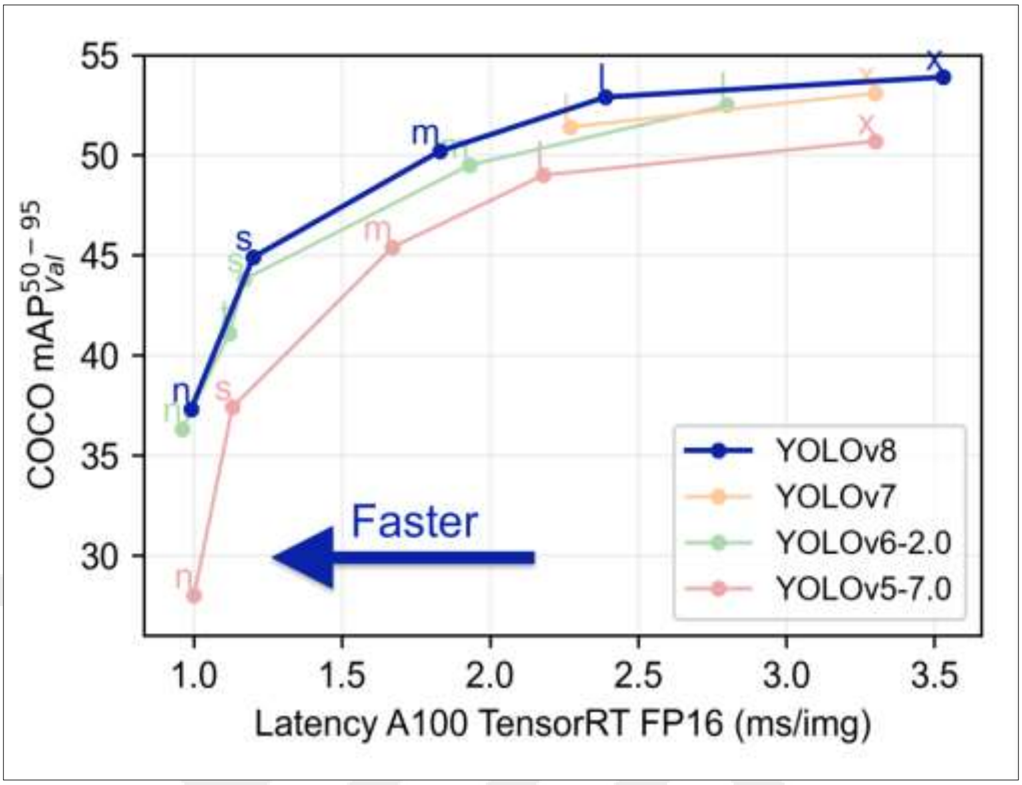


Figure 3.3: Performance Comparison on the MS COCO Benchmark.

This ground-breaking change enables the model to directly anticipate the center of an object, minimizing problems with anchor boxes including lack of generality and trouble with irregularities. YOLOv8 accelerates NMS, or Non-Maximum Suppression, is an essential post-processing procedure in charge of sorting through candidate detections following inference, by lowering how many box forecasts there are. As the essential building blocks of neural networks, convolutions also make their appearance in YOLOv8. The stem's original 66 convolution has been replaced with a 33 convolution, and C2f, which replaces C3, has been introduced.in a model structure that is more effective and adaptable [73].

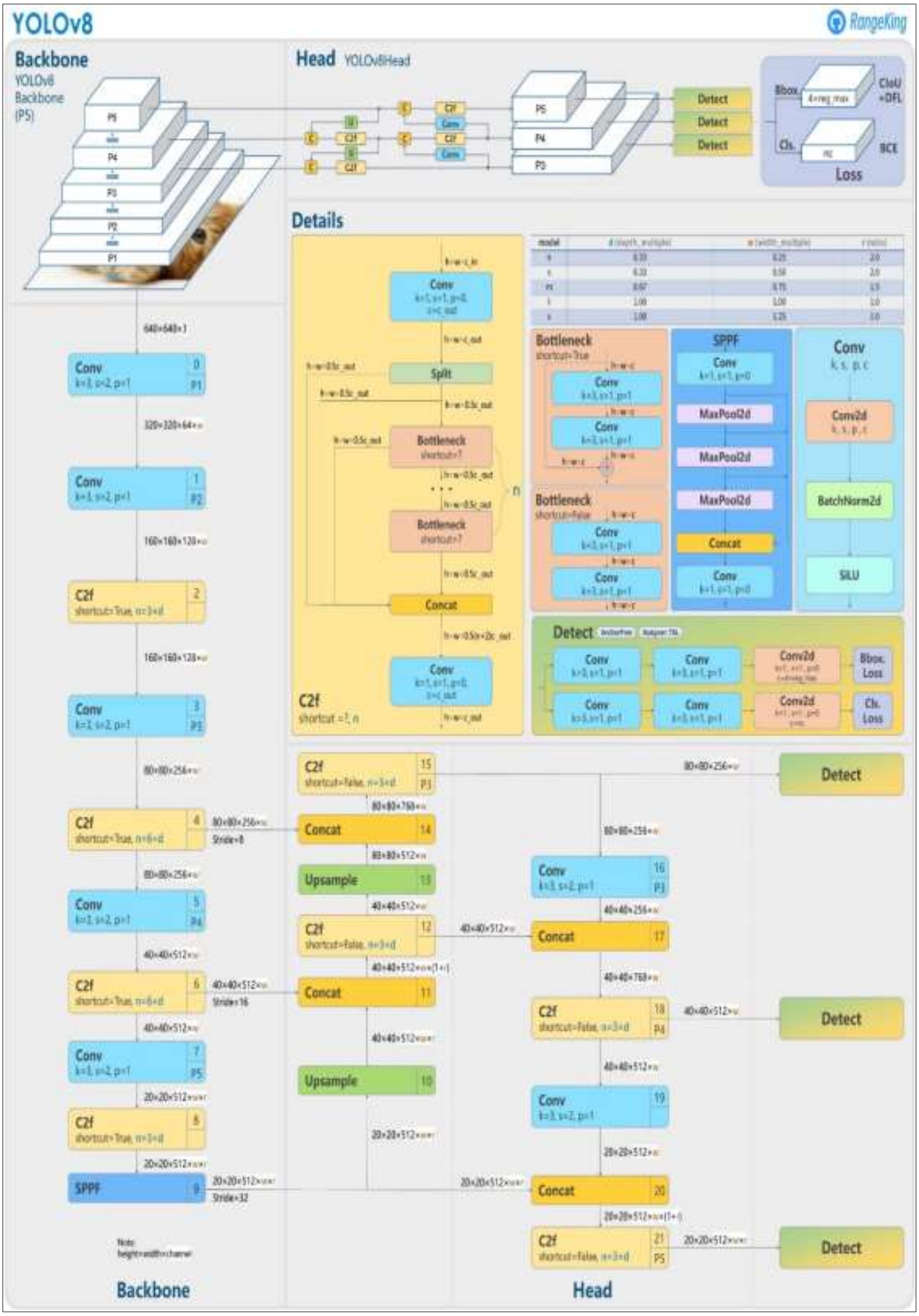


Figure 3.4: Yolov8 Architecture.

This new structure uses all the Bottleneck's outputs, which are two 3 3 convolutions utilizing leftover connections, as opposed to the C3 setup, which only used the last Bottleneck's output The stronger model structure ensured by this modification pushes the limits of the potential of YOLOv8 in computer vision applications [74].

The stronger model structure ensured by this modification pushes the limits of the potential of YOLOv8 in computer vision applications. The Pyramid Pooling Feature in Space (SPPF) is a feature that YOLOv8 implements in an ongoing attempt to accommodate objects of varying sizes. Additionally, this model employs a mosaic augmentation technique for online picture augmentation during training. By joining four photos together, this method enables the model to learn objects in novel places, under partial occlusion, and against various background pixels [75]. Figure depicts the YOLOv8 architecture created by RangeKing on GitHub. It differs from YOLOv5 in the following ways.

a. The C3 module has been replaced with the C2f module.
b. Replace the first 6 9 6 transform on the spine with a 3 9 3 transform. Conversions 10 and 14 must be removed from the YOLOv5 setup.
c. Convert the original 191 transformation of the bottleneck to a 3 9 3 transformation.
d. Use the split head to remove the object step.

The basic building block has been changed, with C2f replaced by C3, and the initial transformation of the stem 6 9 6 has been changed to 3 9 3. The letter “f” indicates the total number of features, the letter “e” indicates the growth rate, and the CBS is a component block of Conv, BatchNorm, and SiLU later in Figure 2, which is a diagram describing the unit. The core of the original convoy size has been changed from 1 9 1 to 3 9 3, however, the bottleneck remains the same as in YOLOv5.

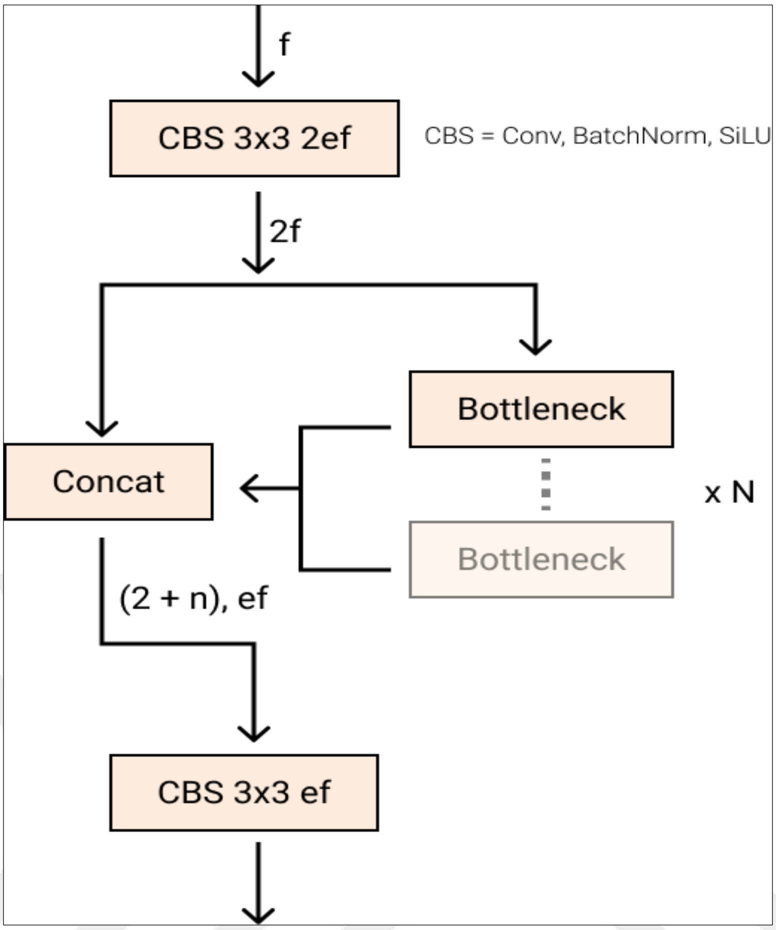


Figure 3.5: Module Yolov8 C2f.

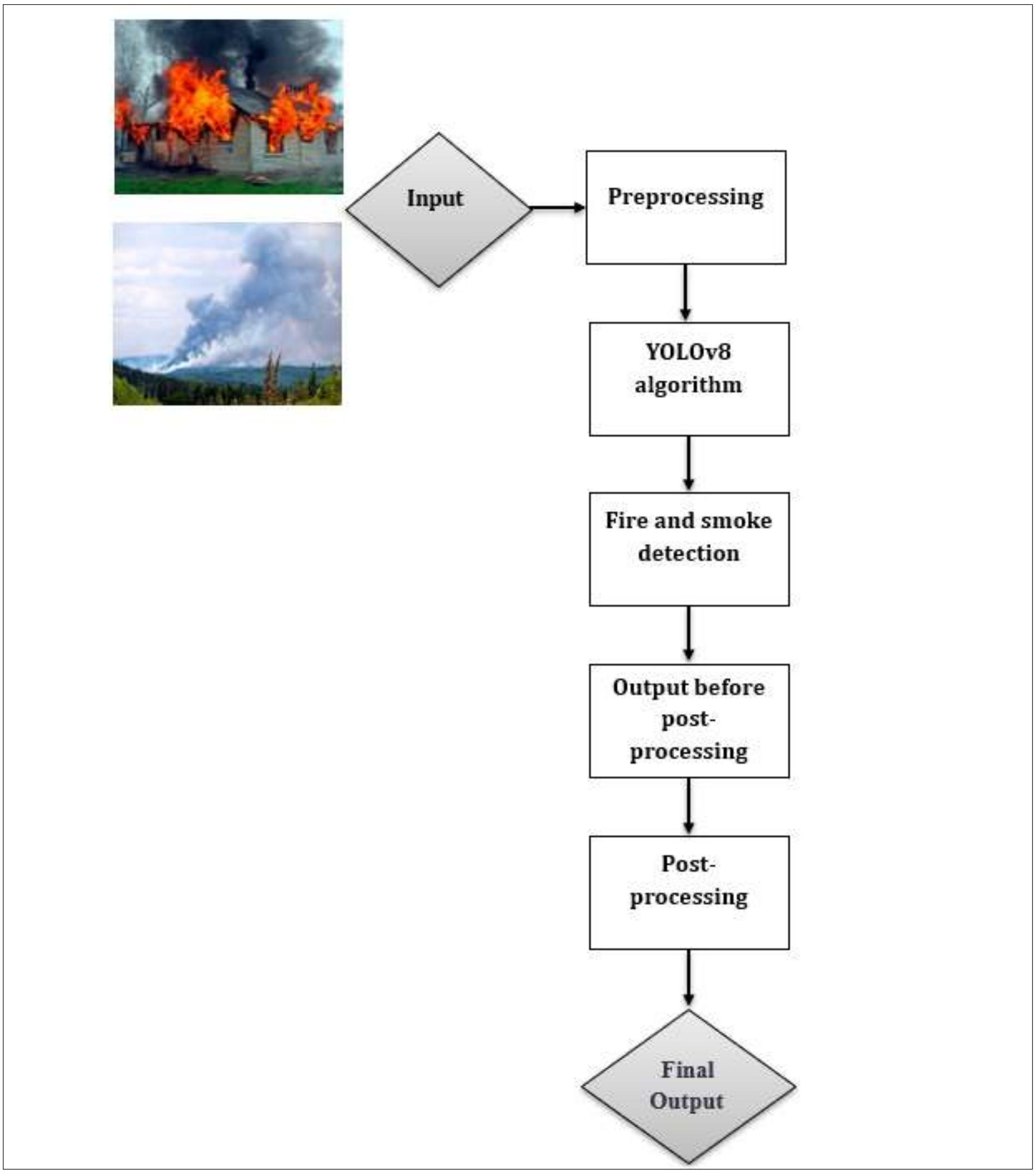
significant improvement in the YOLO series, YOLOv8 raises the bar for the next computer vision efforts. YOLOv8 keeps developing, becoming a more desirable option in the field of real-time object detection thanks to high rates of accuracy and creative architecture modifications, and greater developer features. Like its predecessors, YOLOv8 is available in a variety of sizes, including nano (n), small (s), medium (m), large (l), and extra-large (x), to accommodate the varying demands of various applications. Table 1 lists the characteristics of these five variants, which we will thoroughly investigate.

Table 3.1: Contains Information on the Yolov8 Variations Investigated.

Model	Size (Pixels)	Paramar(M)	Flops(B)
YOLOV8n	640	3.2	8.7
YOLOV8s	640	11.2	28.6
YOLOV8m	640	25.9	78.9
YOLOV8l	640	43.7	165.2
YOLOV8x	640	68.2	257.8

3.2.1 Detailed YOLOv8 Algorithm Based Fire and Smoke Detection

At various stages of our inquiry, we could find it helpful to provide observations and procedural recommendations on the method. These comments could include more background information, explanations, or noteworthy details pertaining to each phase. The graphic illustrates the addition of comments as well as the insertion processes at every stage of the algorithm. We will discuss how it operates. More picture frames that were taken with a camera, drone, or other device are shown and data is gathered for them. As part of the pre-processing stage and the subsequent stage, this stage resizes and normalizes the input photos. Preprocessing makes ensuring that the incoming data is consistent, which helps the YOLOv8 algorithm. produce precise projections After that, the YOLOv8 algorithm phase will be carried out executing. At this stage, the picture will be divided into grid segments by the algorithm, which will be executed in each of them Using the associated data, create bounding boxes and forecast the bounding box and object class probability for each grid cell. YOLOv8 uses a grid-based technique to detect objects with speed and accuracy. The model will learn to identify and localize flames and smoke following the stage, which is the most crucial phase in fire and smoke detection, using the YOLOv8 training technique conducted on a dataset including drawings of fires and smoke. The model is trained by exposing it to labeled data, which helps it identify patterns associated with smoke and fire. The output phase, which looks at the bounding boxes, is the next. Executing. At this stage, the picture will be divided into grid segments by the algorithm, which will execute in each of them. Using the associated data, create bounding boxes and forecast the bounding box and object class probability for each grid cell. YOLOv8 uses a grid-based technique to detect objects with speed and accuracy. The model will learn to identify and localize flames and smoke following the stage, which is the most crucial phase in fire and smoke detection, using the YOLOv8 training technique conducted on a dataset including drawings of fires and smoke. The model is trained by exposing it to labeled data, which helps it identify patterns associated with smoke and fire. The output phase, which looks at the bounding boxes, is the next.



Figures 3.6: Flowchart a Comprehensive Overview of Yolov8 Algorithm-Based Fire and Smoke Detection.

3.3 DATASET

3.3.1 Dataset Description

The dataset is a crucial part of the model because each object detection model needs one to be trained. Some individuals mistakenly believe that a dataset is simply a collection of images that they may obtain from any source. This will unavoidably result in a poor object detection model. The data set can be made up of either pictures or numbers depending on the model utilized; We employed an image dataset for our model, which was also used to build Roboflow, as seen in the picture. Images were manually annotated with the words "fire, smoke." There are 9899 photos in the collected dataset. We discovered that the data amount was insufficient to produce accurate detection findings. By utilizing methods like picture augmentation, the number of pictures can be expanded. output for the training example per session 15% of the photos were in a grayscale range. 16,828 photos from the "Applied To" collection are optimized and shrunk to (640x640)to hasten detection, between in mosaic scaling. The training, validation, and testing (Train, Val, Test) datasets were then split apart. The Table displays the data used following the data increase.

Figure 3.7: Images of Several Classes Taken from the Unique Dataset.

The table displays how the dataset was used and processed. The data set was split into training, validation, and testing as indicated in the table after the data was augmented and processed. This improves detection and accuracy.

Table 3.2: Dataset Description Using Data-Augmentation Methods.

Dataset	Training set (82) %	Validating set (12) %	Testing Set (6) %	Image Ratio	Target Classes
Fire Smoke-Detection Dataset	13,854	1,980	990	640 * 640	Fire = 0 Smoke = 1

3.3.2 Image Pre-Processing

The experimental findings in the preceding section demonstrate that our network's accuracy rate is only approximately 95.7%. The problem of overfitting is one of the causes of poor network performance. It does not work well on test data with only 9899 photos. To address this issue, we train the network using more datasets. We boost the quantity of photos by employing various data augmentation techniques. This concept benefits our network. To improve one's performance. Data augmentation is a popular method for increasing the variety and diversity of data collection to improve the performance and resilience of deep learning. Deep learning models for artificial intelligence are also available. During our actual data-gathering procedure, we used data augmentation techniques on a few photographs to make the dataset more realistic. We used several data augmentation techniques on photos, as seen below. In this situation, for each training example's output pictures Between -15° and +15° rotation Grayscale was used on 15% of the photos. Color temperature between -20° and +20° Mosaic was used. Mosaic zooms merge four photos into a single image. Therefore, the things in the image are linked together. They are scaled down from the original image. This kind the increase is beneficial in terms of detection YOLOv4 uses little things in photos for the first time [76]. The figure depicts the Increase mosaicism was used to combine four photos.

Figure 3.8: Case of Sample Data Augmentation.

This approach assisted us in improving the visibility and clarity of dark pictures, making them more appropriate for analysis and model training. A spin augmentation technique was used. As a data augmentation phase, this approach applies random rotation to selected photos. The rotation angle was picked at random between 15 and +15°. By rotating photos, where items might appear in multiple directions, the dataset gets more realistic. This boosting strategy allows the model to learn and generalize more well, even when confronted with pictures that have been rotated during inference or testing. The enriched data was then joined with additional primary data to form a bespoke dataset. In this unique data set the original raw data and the enriched data are then combined. As a result, it became more reflective of real-world events and encompassed a broader variety of variants, improving the dataset's quality and increasing the model's generalizability. The expanded data gave more instances with changeable features, such as photos whose brightness was altered or rotated, which aids in bias reduction. Furthermore, this improves the resilience of the trained model. A custom dataset is a significant resource for a variety of applications since it allows us to access a more complete and diversified set of data for certain needs. This allows for improved model training, validation, and assessment, resulting in more accurate and dependable outcomes in the target domain. Figures illustrate a sample enlargement of the image for each example, as well as the discrepancies between the photos. This is something that shows us how to deal with it.

Figure 3.9: Case of Data Augmentation.

3.3.3 Data Splitting

Data splitting improves the model's performance and capacity to generalize to new data while preventing overfitting. It is advised to employ data incision before beginning to use the algorithm. This method was used to split existing data into two or three subgroups. These subsets are referred to as training, validation, and test sets. It is used to create feature sets from which the model may learn. The challenge of training the model on data entities is addressed by this SpitNN approach. exactly is a massive data set We can readily use this data for training, allowing our model to learn more complex and diverse properties. There are various methods for separating data into training and testing sets. the

The most popular method is to Make use of random sampling. Completely random selection. It is a simple strategy to apply, and this method assures frequency distribution. Within the training and test sets, the results are roughly comparable [77]. A separate test suite is kept ensuring the model's correctness in real-time. Consequently, rather than separating the current dataset into training, validation, and test sets, the data was segmented into training and validation sets based on the ratio. The split percentage is very dependent on the type of model and the data collection itself. When it comes to images and videos, the model takes a lot of training, thus a bigger amount of the data is used for training. When numerous hyperparameters are established in the model, it is advised to keep a larger proportion of the data for the validation set. After dividing the data set based on the proportion of photos utilized, we proceed as follows:

a. Training Set

A training set is a subset of a larger dataset. To understand the actual relationship between these aspects, the model employs a training set for learning the features from the dataset, fitting the parameters, and changing weights. 82% of the initial dataset is divided into sections and utilized as a training set.

b. Validation Set

A validation set is used to assess the model's fit to the data set. that it is used to fine-tune and alter the algorithm's (neural network's) hyperparameters. The model does not learn anything from the validation set, but it does utilize it to modify Hyperparameters to produce a better suited overall model. Validation is enabled. It is a hybrid set that is used in training for testing reasons but not as a final exam for hiring. 1980 photos make up 12% of the data.

c. Testing Set

A test data set is a subset of data used to assess a deep learning model's performance after it has been trained on the training set and verified on the validation set. The test set of data is significant because it offers an impartial assessment of how well the model generalizes to previously unknown data and how resistant it is to various variations in data distribution. We can ensure that the model is unaffected by the training data and can generate accurate predictions on new data in this manner. When the validation set is finished and the final model is chosen, the test set should be utilized as the ultimate form of evaluation. The percentage of data in this set was 6%. The test set data also aids in the generation of more trustworthy performance estimates and the reduction of variation in model assessment.

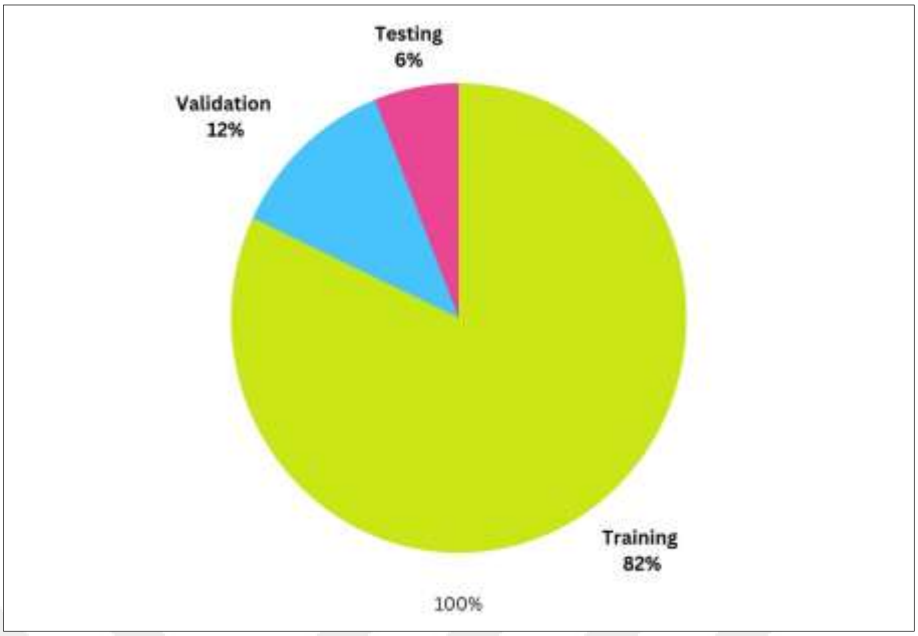


Figure 3.10: Analyze Data Set Splitting.

3.3.4 Model Training and Evaluation

A model was evaluated based on accuracy, training speed, and computing needs to determine the feasibility of an object recognition model for fire detection tasks. The investigation primarily focused on YOLOv8, an object recognition technique that uses a convolutional neural network to simultaneously evaluate bounding boxes and image class probabilities, and in particular Selection, A dataset containing training and validation data spanning over,125 epochs was used to evaluate the performance of the, YOLOv8l model in identifying public-directed fires. The comprehensive comparison technique was essential while discussing deep learning architectures for object identification. The study proposed a trustworthy instruction and assessment approach specifically designed to evaluate the effectiveness of deep learning frameworks. This approach requires a specific number of training periods, and then by evaluating the "best model" it is selected depending on how well it performs on the validation set. To comprehensively examine the models' ability to learn over time, rigorous training was conducted over 125 epochs. The performance of individual models was evaluated uniformly on the validation dataset throughout the training procedure. The "best model" for smoke and public fire detection after 125 epochs was selected based on its performance, with a strong emphasis on recall. Given its potentially disastrous implications in real-world applications, the focus of this scenario was to reduce false positives. With this update, the nominal proportions of each data set component are clearly shown on the graphic. The graphic illustrates the process of gathering data and going from growing to splitting it. Setting aside a suitable portion for testing, validation, and training is best practice. To assess the model's performance and make sure it applies well to fresh, untested data, this segmentation is essential.

Our plan successfully links the flow of the initial data to various processing stages, data augmentation, and dataset splitting for assessment and training.

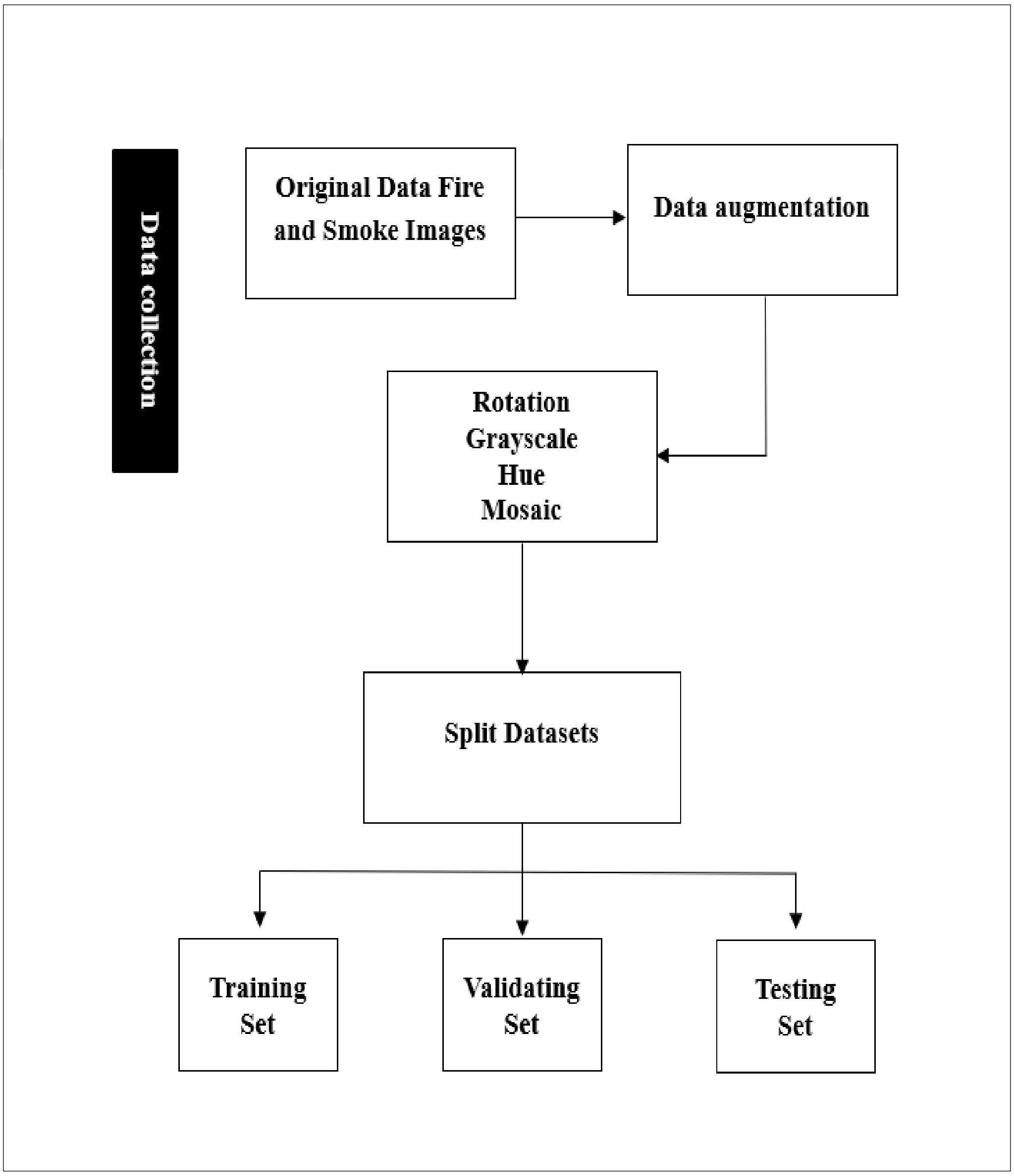


Figure 3.11: Data Splitting and Flow for Training Fire and Smoke Detection Models.

4. EXPERIMENTS AND DISCUSSIONS

4.1 PERFORMANCE METRICS

While gathering the data and training the machine learning model are crucial first steps on the deep learning journey, it's also crucial to evaluate the coach model's performance. It is crucial to understand how effectively the model generalizes unseen data and whether it can be used to address the issue. We utilize the measures of accuracy, recall, and mean average precision (mAP) to evaluate the YOLO models. In this study, the performance of the suggested YOLOv8-based technique was evaluated using a variety of assessment measures.

These measurements include Mean Average Precision (mAP), Precision (P), Recall (R) and F1 score.

Average Mean Precision (mAP): It is a statistic used to assess how well object detection models perform. The Average Precision (AP) for all classes is averaged out to determine the value of mAP. It represented the accuracy of object detection, hence a higher score in this metric indicates an improvement in the model's detection accuracy when the IoU value is reached. or more than 0.5. As a result, mAP may be calculated as.

$$mAp = \quad \frac{1}{n}\sum_{i=1}^{n} Api \tag{4.1}$$

where n is the number of courses, and APi is the AP (Average Precision) of class i. The area under the precision-recall curve is calculated in the conventional definition of AP. It measures the accuracy-versus- recall trade-off by averaging over all classes after calculating the average precision (AP) for each class. [78].

In every case, the accuracy and recall values fall between the ranges of 0 and 1. As a result, AP likewise lies between the same two numbers, 0, and 1.

Tabe 4.l: Performance Evaluation of the Yolov8 Training Step.

	Precision	**Recall**	**mAp@0.5**	**mAp50-95**
All Class	0.956	0.916	0.966	0.925
Fire	0.972	0.97	0.991	0.966
Smoke	0.941	0.864	0.942	0.883

Precision & Recall: Precision quantifies the percentage of correctly positive classifications and assesses the accuracy of your forecasts [79]. Following is the definition of Precision:

$$\text{Precision} = \frac{\textit{Ture postive}}{\text{Ture postive+Fals postive}} \tag{4.2}$$

Recall examines your ability to locate every positive, and it shows what percentage of positive cases the classifier properly detected. Recall's formula is provided by.

$$\text{Recall} = \frac{\textit{Trun postive}}{\textit{Trun postive+Fals Negative}} \tag{4.3}$$

The harmonic means of Precision and Recall, or F1 score, attempts to strike a balance between these two variables. This can be inferred as.

$$\text{F1 score} = 2 * \frac{\textit{Recall} * \textit{Precision}}{\textit{Recall} + \textit{Precision}} \tag{4.4}$$

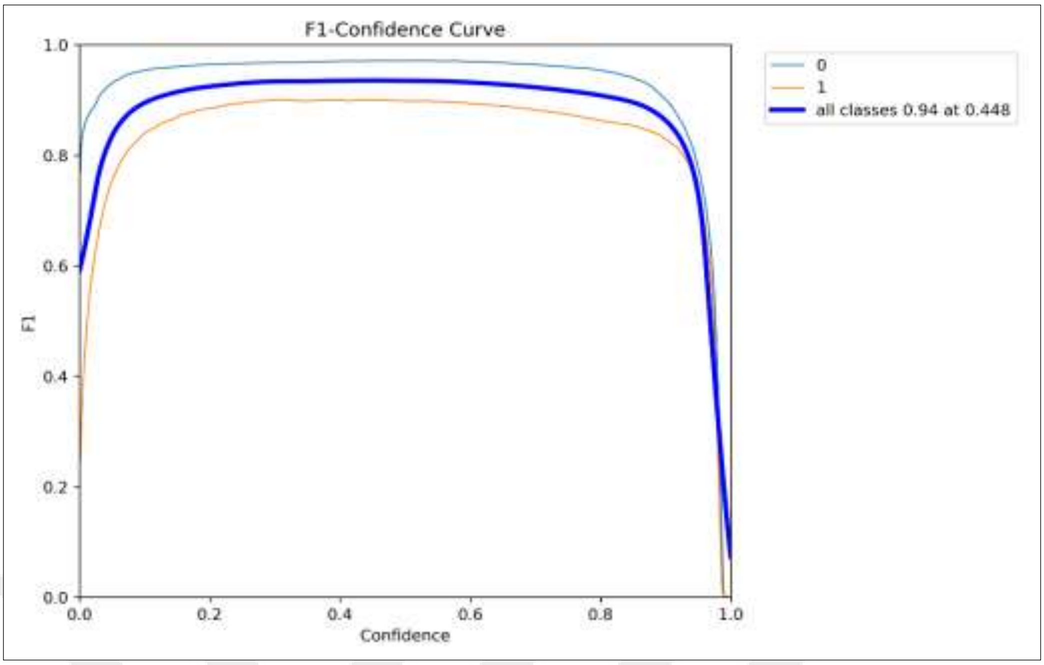


Figure 4.1: Our Yolov8 Model Received a Score of.

Based on the values in Table (4.1), our model gets an F1 score of 0.94 after the training phase. The figure below depicts the accuracy of our model at each epoch throughout both the training and validation processes.

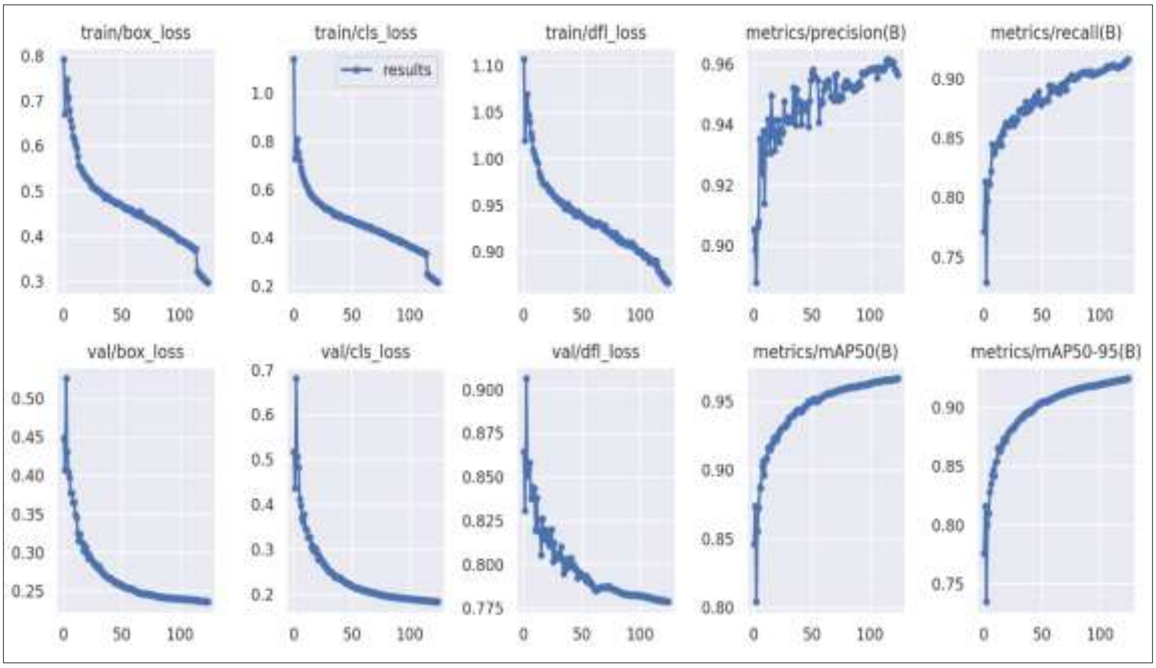


Figure 4.2: Graph of the Results and Losses.

While the losses of bounding boxes and object classes decrease with the number of epochs, Figure (4.1) demonstrates that our model's precision increases with the number of epochs suggesting that more epochs lead to greater precision.

4.2 MODEL PREFERENCE

According to the results, YOLOv8, one of the most sophisticated single-stage target detection models, has an excellent mAP@0.5 for general fire classification and recognition, including buildings, although there is still potential for improvement. The detection accuracy of YOLOv8l is 0.966. The table displays the YOLOv8l training time across 125 epochs Ultralytics new This experiment compared its performance using the same data set and parameter parameters.

Table 4.2: Training Parameter.

Parameter	Value
Batch-size	32
epochs	125
Learning rate	0.001
Image-size	640*640
Trian-time	5.383

4.3 DETECTION OF FIRES AND SMOKES

The fire dataset includes images of fires from three scenes: interior, urban road, and forest. Several methods for data augmentation training have been discovered after extensive testing. This is a more convenient case-based approach to providing information. Significantly improves accuracy. When the model is trained, the background is shown. The image also carries weight. While the model has been trained to perform at the basic level, it is preferable to train it in the city. The model only includes images of fires on urban highways. We used both ends and achieved up to 96.6% average accuracy in detection on both ends. Detection of building fires and urban road fires Detection of fires inside residences ,Forests. Database A set was obtained from 9,889 images, including images of fire and smoke, from the Internet, however, the data was enhanced using filters. Then the increment became 16828 and it was taught and published. The smoke detection data was separated. As mentioned earlier, smoke detection works. I am trying to detect smoke. This is like having a second chance to identify a fire even if it has already started. The fire filter did not detect it.

4.4 VISUALIZATION OF RESULTS

The figure depicts the identification of several fire occurrences using the YOLOv8 model. Expectations The model facilitates remarkable precision and alignment, which closely matches the original thorough fire test photographs. Amazingly, the model's performance stays consistent even when charged with finding smaller and smaller pockets of fire throughout repeated fire seasons. Despite these obstacles, the model is still capable of detecting fire and smoke detection instances with a high level of confidence, confirming its superior performance in fire detection tasks.

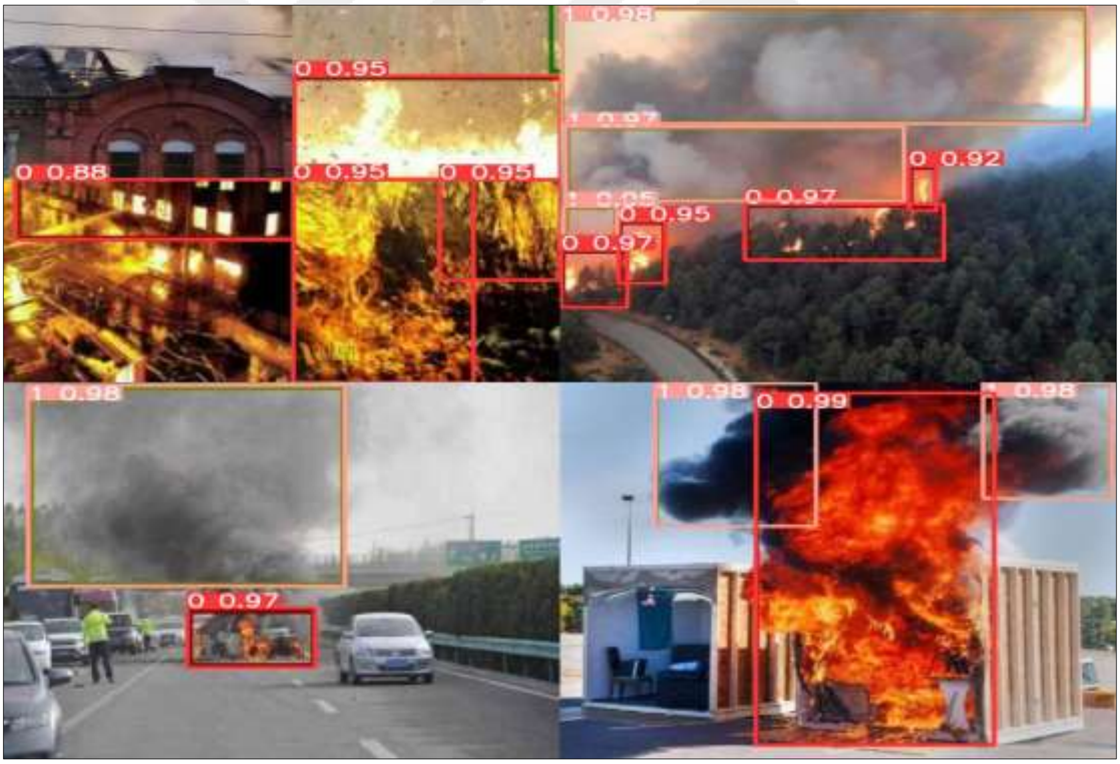


Figure 4.3: Object Detection Examples Using Yolov8.

5. CONCLUSION AND FUTURE WORK

5.1 CONCLUSION

Fire and smoke detection systems are critical in protecting lives and preventing risks from spreading. Current fire and smoke detection devices are ineffective in areas with plenty of space and space External circumstances. Deep learning, machine learning, and artificial intelligence have given rise to new techniques for resolving various issues. Using typical procedures, it is difficult to solve. This thesis is based on computer vision, and we utilize data on Fire and smoke through groups from the Internet, except whether this data is used through security cameras located in external surroundings or other. It makes use of the most recent deep learning developments by transferring learning to lightweight neural networks. This very high level was established in this article. Classification Deep CNNs can attain performance even when... little data is available. Overfitting is an issue induced by a restricted collection of visual data during training, which leads to weakness. The ability of neural network models to tackle this issue is that we use several data Reinforcement strategies to expand our training sets. Allows for notable gains in performance deep inside the Yolov8 model without requiring a major increase in the number of Learnable parameters. We want to increase our existing efforts in the future. Creating a strong algorithm for detecting flames and smoke Capable of detecting fire and smoke. This thesis makes use of a unique collection of 9,899. Each category addresses the issue of fire and smoke detection. We utilized the data after augmenting the data in this work. it grew to 16,828 photos. We worked and trained the data on the first data set, but the average accuracy was mAp of We then enhanced the data using filters and ran on the second data set, and the accuracy was at the required level mAp of 96.6% We utilized 125 data points for training on the first and second data sets.

5.2 FUTURE WORK

Future work extending this analysis to more datasets will be helpful to further our understanding of these models in various scenarios and circumstances. This would increase the body of knowledge and make the results more broadly applicable. Furthermore, there is a great chance to build on these findings and investigate fresh approaches to enhance architectural performance because deep learning is a dynamic field with ongoing research. To improve detection accuracy and resilience, future research may focus on tweaking structures, optimizing hyperparameters, or even researching ensemble approaches.

All things considered, this work establishes the foundation for further research while also advancing the field of computer vision in smoke and wildfire detection. Aimed at reducing the devastating impacts of wildfires and other calamities and increasing the effectiveness of detection systems. By building on these findings, scientists may now go deeper, make fresh suggestions for enhancements, and push the frontiers of the field to create forest fire detection systems that are more effective and efficient.

REFERENCES

[1] H. Yim, S. Oh, and W. Kim, “A study on the verification scheme for electrical circuit analysis of fire hazard analysis in nuclear power plant,” Journal of the Korean Surgical Society, vol. 30, no. 3, pp. 114–122, 2015.

[2] W. Wang, W. Mao, and H. E. Jianliang, “Smoke recognition based on deep transfer learning,” Journal of Computer Applications, 2017.

[3] Abdusalomov, A.; Baratov, N.; Kutlimuratov, A.; Whangbo, T.K. An Improvement of the Fire Detection and Classification Method Using YOLOv3 for Surveillance Systems. Sensors 2021, 21, 6519.

[4] Yazdi, A.; Qin, H.; Jordan, C.B.; Yang, L.; Yan, F. Nemo: An Open-Source Transformer-Supercharged Benchmark for Fine-Grained Wildfire Smoke Detection. Remote. Sens. 2022, 14, 3979.

[5] Xu, R.; Lin, H.; Lu, K.; Cao, L.; Liu, Y. A Forest Fire Detection System Based on Ensemble Learning. Forests 2021, 12, 217.

[6] Ghosh, Rajib, and Anupam Kumar. "A hybrid deep learning model by combining convolutional neural network and recurrent neural network to detect forest fire." Multimedia Tools and Applications 81.27 (2022): 38643-38660.

[7] Li, Zhongjin, et al. "Optimizing makespan and resource utilization for multi-DNN training in GPU cluster." Future Generation Computer Systems 125 (2021): 206-220.

[8] He K M, Zhang X Y, Ren S Q, et al. Deep Residual Learning for Image Recognition. Proc. of The IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770–778.

[9] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. Proc. Of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580–587.

[10] Wang J Q, Chen K, Yang S, et al. Region Proposal by Guided Anchoring. Proc. of The IEEE Conference on Computer Vision and Pattern Recognition, 2019: 2965–2974.

[11] Juergen Schmidhuber. Deep learning in neural networks: An overview. Neural Networks, 61, 04 2014.

[12] Anguelov D, Erhan D, Szegedy C, et al. SSD: Single Shot Multibox Detector. Proc. Of the European Conference on The Computer Vision, 2016: 21–37.

[13] Xu, Hao, Bo Li, and Fei Zhong. “Light-YOLOv5: A Lightweight Algorithm for Improved YOLOv5 in Complex Fire Scenarios”. Applied Sciences 12.23 (2022): 12312–12312.

[14] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]/Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.

[15] Mukhiddinov, Mukhriddin, Akmalbek Bobomirzaevich Abdusalomov, and Jinsoo Cho. “Automatic Fire Detection and Notification System Based on Improved YOLOv4 for the Blind and Visually Impaired”. Sensors 22.9 (2022): 3307–3307.

[16] Wu, Hufeng, et al. “Ship Fire Detection Based on an Improved YOLO Algorithm with a Lightweight Convolutional Neural Network Model”. Sensors 22.19 (2022): 7420–7420.

[17] Xue, Zhenyang, Haifeng Lin, and Fang Wang. “A Small Target Forest Fire Detection Model Based on YOLOv5 Improvement”. Forests 13.8 (2022): 1332–1332.

[18] Lin, Ji, Haifeng Lin, and Fang Wang. “A Semi-Supervised Method for Real-Time Forest Fire Detection Algorithm Based on Adaptively Spatial Feature Fusion”. Forests 14.2 (2023): 361–361.

[19] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems 28 (2015).

[20] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[21] Liu, Wei, et al. "Ssd: Single shot multibox detector." Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016.

[22] Jocher, Glenn, Ayush Chaurasia, and Jing Qiu. "YOLO by Ultralytics." URL: https://github. com/ultralytics/ultralytics (2023).

[23] Fggia, P., Saggese, A., Vento, M.: 'Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion', IEEE Trans. Circuits Syst. Video Technol., 2015, 25, (9), pp. 1545–1556.

[24] A. Rafiee, R. Tavakoli, R. Dianat, S. Abbaspour, and M. Jamshidi, "Fire and smoke detection using wavelet analysis and disorder characteristics," in Proc. 3rd Int. Conf. Comput. Res. Develop. (ICCRD), vol. 3. Mar. 2011, pp. 262–265.

[25] T. Çelik and H. Demirel, "Fire detection in video sequences using a generic color model," Fire Safety J., vol. 44, no. 2, pp. 147–158, Feb. 2009.

[26] Rafael C. Gonzalez and Richard E. Wood, Digital Image Processing, 2Ed. New Jersey: Prentice Hall, 2002.

[27] Zhao Y, Li Q, Gu Z. Early smoke detection of forest fire video using CS Adaboost algorithm[J]. OptikInternational Journal for Light and Electron Optics, 2015, 126(19): 2121-212.

[28] Thou-Ho (Chao-Ho) Chen, Yen-Hui Yin, Shi-Feng Haung and Yan-Ting Ye. "The smoke detection for early fire alarming system based on video processing". Proceeding of the 2006 international Conference in intelligent information Hiding and Multimedia Signal Processing (IIH-MSP'06).

[29] S. Calderara, P. Piccinini, and R. Cucchiara, "Smoke detection in video surveillance: A mog model in the wavelet domain," in ICVS, Santorini, Greece, May 2008, pp. 119–128.

[30] Expert System. What is machine learning? a definition. Blog, Machine Learning, 2017.

[31] Neisser, U. (1967). Cognitive psychology. Thinkingjudgement and Decision Makin. https://doi.org/10.1126/science.198.4319.816.

[32] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. Nature, 521(7553):436–444, May 2015.

[33] D.Ciresan, U. Meier, and J. Schmidhuber. Multi-column deep neural networks for image classification. In 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Jun 2012.

[34] B. Gardiner, Sonya Coleman, and Bryan Scotney. A design procedure for gradient operators on hexagonal images. pages 47 – 54, 10 2007.

[35] L. O. Chua, CNN: A vision of complexity, vol. 7, no. 10. 1997.

[36] H. C. Shin et al., "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics, and Transfer Learning," IEEE Trans. Med.

[37] J. Tan, Q. Yang, J. Hu, Q. Huang, and S. Chen, "Tropical Cyclone Intensity Estimation Using Himawari-8 Satellite Cloud Products and Deep Learning," Remote Sens., vol. 14, no. 4, pp. 812, 2022.

[38] Punam Patel and Shamik Tiwari. Flame detection using image processing techniques. International Journal of Computer Applications, 58(18):13–16, Nov 2012.

[39] Yong feng QI and Yuan lian HUO. Color-based method for face detection and location. Journal of Computer Applications, 29(3):785–788, May 2009.

[40] N.I. Zaidi, N.A.A. Lokman, Mohd Daud, M.S. Achmad, and K.A. Chia. Fire recognition using rgb and ycbcr color space. 10:9786–9790, 01 2015.

[41] Celik, T., Ozkaramanli, H., Demirel, H.: Fire Pixel Classification Using Fuzzy Logic and Statistical Color Model. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1205-1208(2007).

[42] Jing Yang, Feng Chen, and Weidong Zhang. Visual-based smoke detection using support vector machine. In 2008 Fourth International Conference on Natural Computation. IEEE, 2008.

[43] Zhou Yu, Yi Xu, and Xiaokang Yang. Fire surveillance method based on quaternionic wavelet features. In Lecture Notes in Computer Science, pages 477–488. Springer Berlin Heidelberg, 2010.

[44] Timothy M Davenport. Early Forest Fire Detection using Texture Analysis of Principal Components from Multispectral Video. PhD thesis.

[45] Chunyu Yu, Zhibin Mei, and Xi Zhang. A real-time video fire flame and smoke detection algorithm. Procedia Engineering, 62:891–898, 2013.

[46] Hyuntae Kim, Daehyun Ryu, and Jangsik Park. Smoke detection using GMM and AdaBoost. International Journal of Computer and Communication Engineering, 3(2):123–126, 20.

[47] Arrue, B.; Ollero, A.; Matinez de Dios, J. An intelligent system for false alarm reduction in infrared forest-fire detection. *IEEE Intell. Syst.* 2000, *15*, 64–73.

[48] Bahrepour, M.; van der Zwaag, B.J.; Meratnia, N.; Havinga, P. *Fire Data Analysis and Feature Reduction Using Computational Intelligence Methods*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 289–298.

[49] Rubí, J.N.; de Carvalho, P.H.; Gondim, P.R. Application of machine learning models in the behavioral study of forest fires in the Brazilian Federal District region. *Eng. Appl. Artif. Intell.* 2023, *118*, 105649.

[50] Shmuel, A.; Heifetz, E. Global Wildfire Susceptibility Mapping Based on Machine Learning Models. *Forests* 2022, *13*, 1050.

[51] Vega-Garcia, C.; Lee, B.S.; Woodard, P.M.; Titus, S.J. Applying Neural Network Technology to Human-Caused Wildfire Occurrence Prediction. *AI Appl.* 1996, *10*, 9–18.

[52] Zhang, Q.; Xu, J.; Xu, L.; Guo, H. Deep Convolutional Neural Networks for Forest Fire Detection. In *2016 International Forum on Management, Education and Information Technology Application*; Atlantis Press: Amsterdam, The Netherlands, 2016.

[53] Gibson, R.; Danaher, T.; Hehir, W.; Collins, L. A remote sensing approach to mapping fire severity in south-eastern Australia using sentinel 2 and random forest. *Remote Sens. Environ.* 2020, *240*, 111702.

[54] Tien Bui, D.; Hoang, N.-D.; Samui, P. Spatial pattern analysis and prediction of forest fire using new machine learning approach of Multivariate Adaptive Regression Splines and Differential Flower Pollination optimization: A case study at Lao Cai province (Viet Nam). *J. Environ. Manag.* 2019, *237*, 476–487.

[55] Mohajane, M.; Costache, R.; Karimi, F.; Bao Pham, Q.; Essahlaoui, A.; Nguyen, H.; Laneve, G. Oudija, F. Application of remote sensing and machine learning algorithms for forest fire mapping in a Mediterranean area. *Ecol. Indic.* 2021, *129*, 107869.

[56] LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. Gradient-based learning applied to document recognition, Proceedings of the IEEE. 86 (11): 2278–2324.

[57] Best, N., J. Ott, and E. J. Linstead. 2020. Exploring the efficacy of transfer learning in mining image-based software artifacts. Journal of Big Data 7(1): 1–10.

[58] Namozov, A., and Y. Im Cho. 2018. An efficient deep learning algorithm for fire and smoke detection with limited data. Advances in Electrical and Computer Engineering 18(4): 121–128.

[59] Jiao, Z.; Zhang, Y.; Xin, J.; Mu, L.; Yi, Y.; Liu, H.; Liu, D. A Deep Learning Based Forest Fire Detection Approach Using UAV and YOLOv3. In Proceedings of the 1st International Conference on Industrial Artificial Intelligence (IAI), Shenyang, China, 23–27 July 2019; pp. 1–5.

[60] Wang, Z.; Wu, L.; Li, T.; Shi, P. A Smoke Detection Model Based on Improved YOLOv5. Mathematics 2022, 10, 1190.

[61] Abdusalomov, A.; Baratov, N.; Kutlimuratov, A.; Whangbo, T.K. An Improvement of the Fire Detection and Classification Method Using YOLOv3 for Surveillance Systems. Sensors 2021, 21, 6519.

[62] Al-Turjman F, Al-Karaki JN, Al-Bzoor Z (2021) Hybrid deep learning-based approach for fire detection in smart cities. Sensors 21(6):2186.

[63] Huang J, Luo Q, Wang J, Guo J (2020) Fire detection in outdoor scenes using YOLOv3. IEEE Access 8:114978–114985.

[64] Wang, Y.; Hua, C.; Ding, W.; Wu, R. Real-time detection of flame and smoke using an improved YOLOv4 network. Signal Image Video Process. 2022, 16, 1109–1116.

[65] Nepal, U.; Eslamiat, H. Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs. Sensors 2022, 22, 464.

[66] Mohiyuddin, A.; Basharat, A.; Ghani, U.; Peter, V.; Abbas, S.; Bin Naeem, O.; Rizwan, M. Breast Tumor Detection and Classification in Mammogram Images Using Modified YOLOv5 Network. Comput. Math. Methods Med. 2022, 2022, 1–16.

[67] Jie Lian, Xinyu Pan, Jin Guo. An Improved Fire and Smoke Detection Method Based on YOLOv7.in IEEE Xplore. 2023.

[68] Kim, S.-Y.; Muminov, A. Forest Fire Smoke Detection Based on Deep Learning Approaches and UnmannedAerial Vehicle Images. Sensors 2023, 23, 5702. https://doi.org/10.3390/s23125702.

[69] Xuanbo Jia, Yike Wang, Taiming Chen. Forest Fire Detection and Recognition Using YOLOv8 Algorithms from UAVs Images.in IEEE 5th International Conference on Power, Intelligent Computing and Systems (ICPICS) 2023.

[70] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).

[71] G. Jocher, A. Chaurasia, and J. Qiu, ‘YOLO by ultralytics,’ Ultralytics, Los Angeles, CA, USA, Jan. 2023.

[72] J.-H. Kim, N. Kim, and C. S. Won, ‘High-speed drone detection based on YOLO-V8,’ in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Jun. 2023, pp. 1–2.

[73] A. Vats and D. C. Anastasiu, ‘‘Enhancing retail checkout through video inpainting, YOLOv8 detection, and DeepSort tracking,’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2023, pp. 5529–5536.

[74] A. Aboah, B. Wang, U. Bagci, and Y. Adu-Gyamfi, ‘‘Real-time multiclass helmet violation detection using few-shot data sampling technique and YOLOv8,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2023, pp. 5349–5357.

[75] S. Tamang, B. Sen, A. Pradhan, K. Sharma, and V. K. Singh, ‘‘Enhancing COVID-19 safety: Exploring YOLOv8 object detection for accurate face mask classification,’ Int. J. Intell. Syst. Appl. Eng., vol. 11, pp. 892–897, Feb. 2023.

[76] A. Bockokovskiy, C. Wang, H. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection, " 2020, url: https://arxiv.org/abs/2004.10934.

[77] Max Kuhn and Kjell Johnson. Feature Engineering and Selection. Chapman and Hall/CRC, jul 2019.

[78] Young, G. O., and J. Peters. "Synthetic structure of industrial plastics (Book style with paper title and editor)." (1964): 15-64.

[79] H. Zhu, H. Wei, B. Li, X. Yuan, and N. Kehtarnavaz, ‘A review of video object detection: Datasets, metrics and methods,’ Appl. Sci., vol. 10, no. 21, p. 7834, Nov. 2020.