HUMAN DETECTION IN CATASTROPHE SCENARIOS USING UAV AND DEEP LEARNING TECHNIQUES

A PROJECT REPORT

Submitted by

V. DIVYA PRABA (820420104016)
G. NUHA (820420104038)

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

ANJALAI AMMAL MAHALINGAM ENGINEERING COLLEGE,
KOVILVENNI – 614403

ANNA UNIVERSITY: CHENNAI 600 025

MAY 2024

ANNA UNIVERSITY: CHENNAI 600 025

BONAFIDE CERTIFICATE

Certificate that this project report **“Human detection in catastrophe scenarios using UAV and Deep Learning Techniques**” is the bonafide work of **“DIVYA PRABA V, NUHA G”** who carried out the project work under my supervision.

SIGNATURE
Dr.K. VELMURUGAN M.E., MBA., Ph.D.
HEAD OF THE DEPARTMENT
Professor,
Department of Computer Science
and Engineering,
Anjalai Ammal Mahalingam
Engineering College,
Kovilvenni – 614 403,
Thiruvarur (DT).

SIGNATURE
Mrs. S. FARJANA FARVIN M.E.,
SUPERVISOR
Assistant Professor,
Department of Computer Science
and Engineering,
Anjalai Ammal Mahalingam
Engineering College,
Kovilvenni – 614 403,
Thiruvarur (DT).

INTERNAL EXAMINER

EXTERNAL EXAMINER

ABSTRACT

For effective life-saving during disasters, prompt and accurate detection of human presence is essential. Unmanned Aerial Vehicles (UAVs) equipped with computer vision technology offer a promising solution for rapid reconnaissance in disaster-stricken areas.

This abstract presents a novel approach utilizing the You Only Look Once (YOLO) object detection algorithm with deep SORT (Simple Online and Realtime Tracking) for human detection from UAV imagery in disaster scenarios. YOLO's real-time processing capabilities enable swift identification of humans amidst debris and challenging environmental conditions.

Deep SORT significant improvements in tracking accuracy and robustness compared to traditional tracking algorithms, making it a valuable tool in computer vision and artificial intelligence applications.

Through extensive experimentation and evaluation, the effectiveness and efficiency of the proposed method are demonstrated, highlighting its potential to enhance the speed and accuracy of search and rescue missions during disasters.

The integration of YOLO and Deep SORT with UAVs represents a significant advancement in disaster response technology, facilitating rapid and precise detection of human presence to expedite life-saving efforts

ACKNOWLEDGEMENT

The successful accomplishment of our efforts into reality brings us in the forethought of gratitude to the almighty God who has showered his blessings to us all through our life and this project work.

We convey our deep sense of gratitude to the Managing Trustee Justice, **Thiru. A. RAMAMURTHY**, Anjalai Ammal Mahalingam Engineering College, Kovilvenni, who has given us an opportunity to be a student of this renowned Institution.

We devote our sincere thanks to **Dr. S. N. RAMASAMY, M.E., Ph.D.,** Principal, Anjalai Ammal - Mahalingam Engineering College, Kovilvenni, who has given us an opportunity to do our project work in an intellectual environment.

We are highly beholden to **Dr.K.VELMURUGAN, M.E., M.B.A., Ph.D.,** Head of the Department, indeed he had taken interest in every student of our department. His motivation and untiring efforts make us complete this Under Graduate program successfully.

We take this opportunity to record our deep sense of gratitude to **Mrs. S. FARJANA FARVIN, M.E.,** Assistant Professor, Department of Computer Science and Engineering, and our project guide for the constant help, encouragement and guidance offered during the course of this project work.

We extend our profound thanks to our parents, family and friends for their moral support extended to us during this project work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

UAV	-	UNMANNED AERIAL VEHICLE
AI	-	ARTIFICIAL INTELLIGENCE
ML	-	MACHINE LEARNING
DL	-	DEEP LEARNING
CV	-	COMPUTER VISION
YOLO	-	YOU ONLY LOOK ONCE
DEEP SORT	-	DEEP SIMPLE ONLINE REALTIME TRACKING

CHAPTER-1

INTRODUCTION

1.1 INTRODUCTION

Countless calamities have struck the world. It has impacted the lives of humans ranging from immediate injuries to long term consequences. A lot of lives have suffered because of these calamities. Calamities such as floods, earthquakes, hurricanes, wildfire etc. can cause loss of life, lifetime injury, destruction of their living environment, lack of basic needs and many more unspoken losses. The rescue operation after the disaster requires a lot of professionally trained individuals and organizations to help the people in need. It includes the saving of lives, providing immediate medical care and ensuring the safety of the victims. The professionals include fire fighters, paramedics, search and rescue teams and volunteers are there to help the people who were struck under debris or in a risky condition due to the disaster occurred.

Traditional search and rescue efforts often face significant challenges in accessing remote or hazardous areas, leading to delays in locating survivors and providing timely assistance. However, recent advancements in unmanned aerial vehicle (UAV) technology combined with the power of deep learning techniques offer a promising solution to enhance human detection in catastrophe scenarios. By leveraging the aerial vantage point provided by UAVs and the capabilities of deep learning algorithms, it becomes possible to rapidly identify and pinpoint individuals in need of assistance across various terrains and conditions.

1.1.1. ARTIFICIAL INTELLIGENCE

Ever since the pandemic hit the world, Artificial Intelligence has been rapidly evolving and has become a significant part of the daily life of human beings. In the recent advancements in the technology field, the use of artificial intelligence in almost all industries has not only reduced work time but also increased the efficiency of work place. According to Alan Turing, one of the pioneers of Artificial

Intelligence has said: “A computer would deserve to be called intelligent if it could deceive a human into believing that it was human”. It helps us to solve the difficult problem that are unable to solve more efficiently.

1.1.2. MACHINE LEARNING

It is a branch of the Artificial Intelligence that uses algorithm to learn and make predictions. The algorithm helps the machine to from the large set of data and produces the better development model to predict. For example, if we want a computer to recognize an object, we can’t give the all the information on what the object looks like. Thousands of photos of the object can be provided, and the algorithm can identify recurring patterns in the pictures to identify the object's attributes without the need for human intervention. Over time, the algorithmic algorithms improve their ability to identify items, even in previously unseen images. Today’s technological achievements uses the machine learning in all applications like Pattern recognition, Human Detection, Recommendation systems, Self-driving cars, Automated robots for household and industrial purposes and so on. Machine learning can be categorized into 3 based on the learning. They are Supervised learning, Unsupervised Learning and Reinforcement Learning. The most widely used machine learning algorithm is Supervised Learning. It uses datasets which are labelled and train to predict and recognize patterns.

1.1.3. DEEP LEARNING

It is one of the subsets of machine learning algorithms that uses multiple layers to extract high-value features from the input data. Deep learning is modeled with a similar logic to how the human brain works. It mimics the behavior of the neurons. The algorithm creates an artificial neural network (ANN) with multiple layers that can learn and make decisions on its own. When training the deep learning model, large datasets are needed for the model to be successful. One of the most important advancements in deep learning called transfer learning, is the use of the pre-trained model. These pre-trained models can help train the model with large datasets. Deep Learning aims to optimize computers to think and act using the model based on the human brain. In order to generate precise predictions, the model can identify any complicated patterns in pictures, videos, and other comparable datasets. Large corporations such as Microsoft, Google, and Meta have started using deep learning techniques to train their large language models (LLMs). It solves the challenges in the classification and regression issues and generates the noteworthy forecasts.

1.1.4. CONVOLUTIONAL NEURAL NETWORKS (CNN)

Neural networks are a subset of artificial intelligence and machine learning, serving as the fundamental building blocks for deep learning algorithms. Inspired by the functionality of the human brain, neural networks are designed and structured to tackle complex tasks including classification, pattern recognition, matching, clustering, and optimization. These networks consist of interconnected nodes or neurons that operate in parallel, enabling them to handle intricate problems effectively.

Among the types of deep learning neural network algorithms, convolutional neural networks (CNNs) are highly adaptable for analyzing and visualizing data. Although primarily utilized for image processing, CNNs can be modified to handle audio and other forms of data as well. Due to their remarkable capacity for object processing and identification, CNNs are predominantly employed in computer vision applications such as object detection and pattern recognition. They find extensive use in real-world scenarios, including facial recognition, medical image analysis, and autonomous vehicles. The functioning of convolutional neural networks (CNNs) relies on a specialized technique known as convolution, which involves the combination of two functions to generate a modified third function that represents the shape modifications induced by the other functions. In the context of CNNs, the extraction of features is performed through the use of convolutional layers.

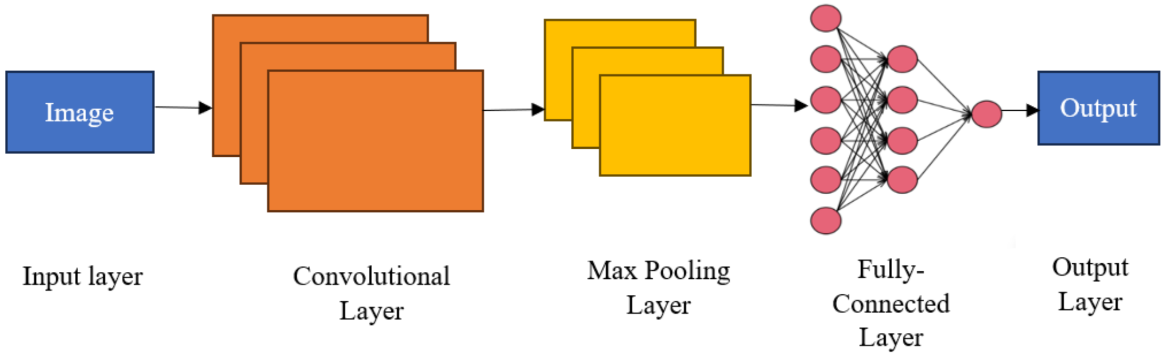

Figure 1.1. Schematic structure of Convolutional Neural Networks (CNNs) with the layers

1.1.5. YOU ONLY LOOK ONCE (YOLO)

The limitations of the convolutional neural networks have led to the development of the alternative models like YOLO, which aims to tackle and efficiently handle the complex tasks and generates the adaptable solutions for real time object detection problems. YOLO was introduced in June 2016 by Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi as an object detection model. The YOLO object detection algorithm processes the entire image during training and testing rather than using the sliding windows or region proposal techniques. It’s ability to perform real time object detection in a single pass through a neural network unlike complex multiple stage of the traditional object detection. It makes fewer background errors and its speed is higher when comparing to other object detection models such as Fast R-CNN and Mask R-CNN. It outperforms Fast R-CNN and Mask R-CNN in terms of accuracy, speed and efficiency. YOLO divides the input image into a grid and predicts bounding boxes with confidence scores and class probabilities directly from this grid.

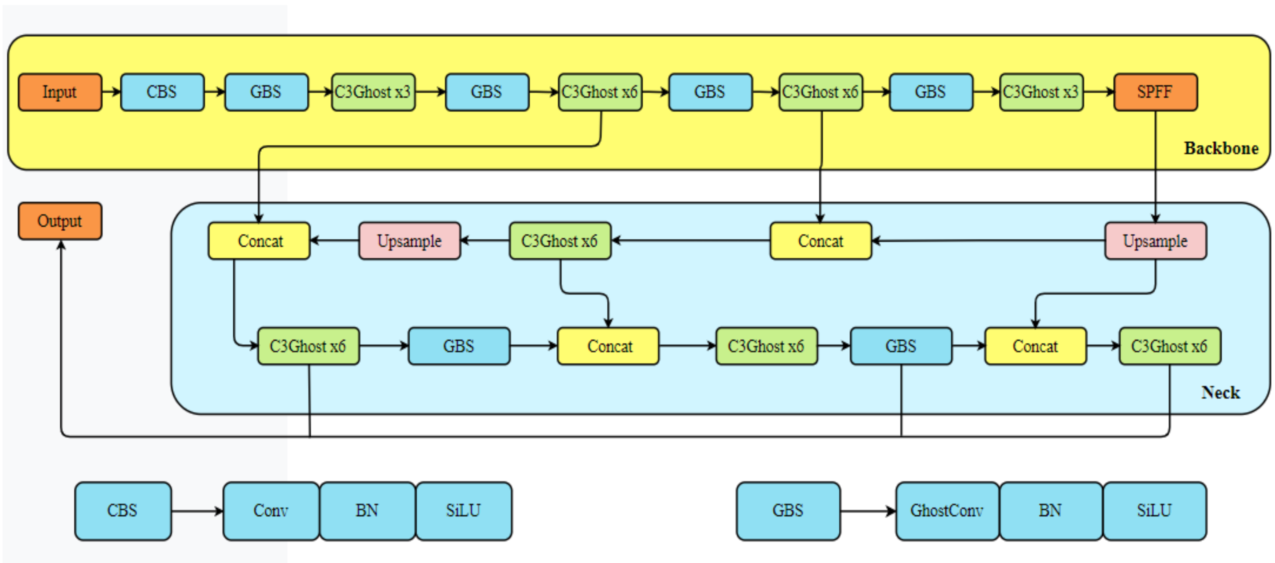

Figure 1.2. Schematic Structure diagram of YOLOv8

1.1.6. DEEPSORT

DeepSORT (Deep Simple Online Realtime Tracking) is a state-of-the-art object tracking algorithm that combines deep learning-based object tracking with appearance information to improve tracking accuracy and handle occlusions effectively. It extends the original SORT algorithm by integrating appearance information based on a deep appearance descriptor, reducing the ID switching problem and enhancing tracking performance. DeepSORT has been used in various applications, such as multi-target tracking, pedestrian detection, and re-identification systems, demonstrating superior accuracy and efficiency in real-time tracking scenarios.

1.2. PURPOSE OF THE PROJECT

The purpose of this project is to develop an efficient and reliable system for human detection in catastrophe scenarios utilizing Unmanned Aerial Vehicles (UAVs) equipped with deep learning techniques. In times of natural disasters, such as earthquakes, floods, or wildfires, conventional methods of search and rescue operations often face challenges due to the complexity of the affected environment, limited accessibility, and time sensitivity. The integration of UAVs and deep learning algorithms presents a promising solution to enhance the speed, accuracy, and safety of locating and rescuing survivors in such scenarios.

1.3. SCOPE OF THE PROJECT

The scope of this project is to integrate Unmanned Aerial Vehicles (UAVs) with advanced deep learning techniques for real-time human detection in catastrophe scenarios. It encompasses the development of software and hardware components for seamless UAV-deep learning integration, employing image processing to analyze high-resolution imagery, optimizing deep learning models for robust human detection, ensuring real-time performance, adaptability to diverse environments, and user-friendly interface design. Ethical considerations regarding privacy and data security will also be addressed, with the overarching goal of providing an efficient and reliable system to aid search and rescue operations in disaster-stricken areas.

1.4. OBJECTIVE OF THE PROJECT

The objective of this project is to develop a robust and efficient system for identifying and tracking individuals amidst disaster situations. Leveraging the YOLOv8 object detection framework and DeepSORT (Deep Simple Online Realtime Tracking), the project aims to enhance disaster response efforts by providing accurate real-time information about human presence and movement in affected areas. By employing cutting-edge deep learning techniques, the system intends to assist emergency responders in locating survivors, coordinating rescue operations, and ultimately, minimizing casualties in critical situations such as natural disasters or humanitarian crises. Through the integration of YOLOv8's rapid object detection capabilities with DeepSORT's robust tracking algorithms, the proposed system seeks to offer a comprehensive solution for human detection and tracking, contributing to more effective disaster management strategies.

CHAPTER-2

LITERATURE REVIEW

2. LITERATURE REVIEW

A literature review is a survey of scholarly sources on a specific topic. It provides an overview of current knowledge, allowing you to identify relevant theories, methods, and gaps in the existing research that you can later apply to your paper, thesis, or dissertation topic.

Demir, K. A., Cicibas, H., & Arica, N. "Unmanned Aerial Vehicle Domain: Areas of Research", 2015

Unmanned aerial vehicles (UAVs) domain has seen rapid developments in recent years. As the number of UAVs increases and as the missions involving UAVs vary, new research issues surface. An overview of the existing research areas in the UAV domain has been presented including the nature of the work categorized under different groups. These research areas are divided into two main streams: Technological and operational research areas. The research areas in technology are divided into onboard and ground technologies. The research areas in operations are divided into organization level, brigade level, user level, standards and certifications, regulations and legal, moral, and ethical issues. This overview is intended to serve as a starting point for fellow researchers new to the domain, to help researchers in positioning their research, identifying related research areas, and focusing on the right issues.

Sarker, I.H. "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions", 2021.

Deep learning (DL), a branch of machine learning (ML) and artificial intelligence (AI) is nowadays considered as a core technology of today's Fourth Industrial Revolution (4IR or Industry 4.0). Due to its learning capabilities from data, DL technology originated from artificial neural network (ANN), has become a hot topic in the context of computing, and is widely applied in various application areas like healthcare, visual recognition, text analytics, cybersecurity, and many more. However, building an appropriate DL model is a challenging task, due to the dynamic nature and variations in real-world problems and data. Moreover, the lack of core understanding turns DL methods into black-box machines that hamper development at the standard level. This article presents a structured and comprehensive view on DL techniques including a taxonomy considering various types of real-world tasks like supervised or unsupervised. In our taxonomy, we take into account deep networks for supervised or discriminative learning, unsupervised or generative learning as well as hybrid learning and relevant others.

Yamashita, R., Nishio, M., Do, R.K.G. et al. "Convolutional neural networks: an overview and application in radiology", 2018.

Convolutional neural network (CNN), a class of artificial neural networks that has become dominant in various computer vision tasks, is attracting interest across a variety of domains, including radiology. CNN is designed to automatically and adaptively learn spatial hierarchies of features through backpropagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers. This review article offers a perspective on the basic concepts of CNN and its application to various radiological tasks, and discusses its challenges and future directions in the field of radiology. Two challenges in applying CNN to radiological tasks, small dataset and overfitting, will also be covered in this article, as well as techniques to minimize them. Being familiar with the concepts and advantages, as well as limitations, of CNN is essential to leverage its potential in diagnostic radiology, with the goal of augmenting the performance of radiologists and improving patient care.

Dina Chahyati, Mohamad Ivan Fanany, Aniati Murni Arymurthy, "Tracking People by Detection Using CNN Features", 2017.

Multiple people tracking is an important task for surveillance. Recently, tracking by detection methods had emerged as immediate effect of deep learning remarkable achievements in object detection. In this paper, we use Faster-RCNN for detection and compare two methods for object association. The first method is simple Euclidean distance and the second is more complicated Siamese neural network. The experiment result show that simple Euclidean distance gives promising result as object association method, but it depends heavily on the robustness of detection process on individual frames.

B. Valarmathi, Jain Kshitij, Rajpurohit Dimple, N. Srinivasa Gupta, Y. Harold Robinson, G. Arulkumaran, Tadesse Mulu, "Human Detection and Action Recognition for Search and Rescue in Disasters Using YOLOv3 Algorithm", 2023.

Drone examination has been overall quickly embraced by NDMM (natural disaster mitigation and management) division to survey the state of impacted regions. The proposed model is used YOLOv3 (you only look once) algorithm for the detection and recognition of actions. In this study, we provide the fundamental ideas underlying an object detection model. To find the most effective model for human recognition and detection, we trained the YOLOv3 algorithm on the image dataset and evaluated its performance. The proposed work shows that existing models are inadequate for critical applications like search and rescue, which convinces us to propose a model raised by a pyramidal component extracting SSD in human localization and action recognition.

Zaman, Fadhlan & Tahir, Noorita & Yusoff, Yusnani & M Thamrin, Norashikin & Hasmi, Ahmad Hafizam " Human Detection from Drone using You Only Look Once (YOLOv5) for Search and Rescue Operation", 2023.

Drones are unmanned aerial vehicles that can be remotely operated to perform a variety of tasks. They have been used in search and rescue operations since the early 2000s and have proven to be invaluable tools for quickly locating missing persons in difficult terrain and environment. In certain cases, automated human detection on drone camera feed can help the responder to locate the victims more effectively. In this work, we propose the use of a deep learning method called You Only Look Once version 5, or YOLOv5. The YOLOv5 model is trained using data collected during a simulation of search and rescue operations, where mannequins were used to represent human victims.

Al Mudawi N, Qureshi AM, Abdelhaq M, Alshahrani A, Alazeb A, Alonazi M, Algarni A. Vehicle Detection and Classification via YOLOv8 and Deep Belief Network over Aerial Image Sequences , 2023

The paper by Al Mudawi et al. proposes a new approach for vehicle detection and classification using UAV (Unmanned Aerial Vehicle) imagery, combining the YOLOv8 object detection algorithm and a Deep Belief Network. The proposed model consists of five stages: pre-processing (defogging and semantic segmentation), feature extraction, vehicle detection using YOLOv8, vehicle classification using a Deep Belief Network, and post-processing. The study used the VEDAI (Vehicle Detection in Aerial Imagery) and VAID (Vehicle Aerial Imagery from a Drone) datasets to evaluate the performance of the proposed approach, and the results demonstrate that the combined YOLOv8 and Deep Belief Network model achieved an accuracy of 95.6% on the VEDAI dataset and 94.6% on the VAID dataset for vehicle detection and classification. The authors highlight the potential of this computer vision-based solution for various aerial imaging applications, including intelligent traffic monitoring systems, and emphasize the importance of pre-processing techniques, such as defogging and semantic segmentation, to improve the performance of the object detection and classification models.

Xiao X, Feng X. Multi-Object Pedestrian Tracking Using Improved YOLOv8 and OC-SORT, 2023

In this paper presents a novel approach for multi-object pedestrian tracking by enhancing the YOLOv8 object detection model with innovative techniques like

SoftNMS, GhostConv, and C3Ghost Modules, resulting in improved precision and efficiency. By integrating the OC-SORT tracking algorithm with a mobileNetV2-based ReID model, the system effectively handles occlusion and ensures robust identity verification for consistent tracking. The approach outperformed existing methods on benchmark datasets like MOT17 and MOT20, achieving high scores in HOTA and MOTA metrics, showcasing its effectiveness in challenging real-world scenarios. Overall, the combination of advanced detection and tracking methodologies in this study demonstrates significant advancements in multi-object pedestrian tracking capabilities.

S. J. Pan and Q. Yang, "A survey on Transfer Learning," IEEE Transactions on Knowledge and Data Engineering, Oct. 2010.

Transfer learning is a machine learning technique that aims to leverage knowledge gained from one task to improve performance on a related task. The survey by Pan and Yang provides a comprehensive overview of transfer learning, including its definition, taxonomy, applications, and key algorithms. The authors discuss the motivations behind transfer learning, such as the limited availability of labeled data and the need to adapt models to new domains. They also present a taxonomy of transfer learning based on the type of knowledge transferred, the learning settings, and the differences between the source and target tasks. The survey covers a wide range of transfer learning algorithms, including instance-based, feature-based, and parameter-based methods, and discusses their strengths and limitations. Finally, the authors highlight several real-world applications of transfer learning, such as image classification, natural language processing, and bioinformatics, and discuss future research directions in this rapidly evolving field

CHAPTER-3

SYSTEM DESIGN

3. SYSTEM DESIGN

3.1. EXISTING SYSTEM

The YOLOv5 model, which is used for human detection, has several limitations that can impact its performance in real-world applications. One of the primary issues is the speed and accuracy of the model, which can be affected by domain-specific biases. This means that the model may perform well in specific domains but may not generalize well to other domains, leading to false detections and poor performance. The model's performance can also be affected by environmental factors such as changes in lighting conditions. This can make it inconvenient to use in real-world applications where lighting conditions may vary significantly. Another issue is the training data used to train the model. If the data is biased, the model may not perform well on new test data, leading to inaccurate detections. This is particularly true in crowded scenes or when objects are far away from the camera, where the model may struggle to accurately detect objects.

In addition, YOLOv5 is a purely detection model, unlike other models such as YOLOv8 and YOLO-NAS, which also include segmentation learning. This means that YOLOv5 may not be as accurate in detecting objects in complex scenes with multiple objects or overlapping objects.

3.2 PROPOSED SYSTEM

The proposed system of our project is to detect humans using YOLOv8 and Deep SORT involves combining the YOLOv8 object detection algorithm with the Deep SORT tracking algorithm to achieve accurate human detection and tracking in real-time scenarios. This system integrates precise detection capabilities from YOLOv8 with effective multi-object tracking provided by DeepSORT.

By leveraging the strengths of both algorithms, this system aims to enhance the accuracy and robustness of human detection and tracking processes.The YOLOv8 algorithm, known for its speed and accuracy in object detection, is enhanced with modifications like spatial attention, feature fusion, and context aggregation. It utilizes a Convolutional Neural Network (CNN) with a backbone (CSPDarknet53) and head consisting of multiple CNN layers to predict bounding boxes and class probabilities.

On the other hand, DeepSORT is an improved version of the SORT algorithm, offering real-time multi-object tracking capabilities.By combining YOLOv8 and DeepSORT, the system aims to address challenges in human detection and tracking, such as occlusions, scale variations, and complex backgrounds. This integration is designed to provide a comprehensive solution that ensures accurate detection, robust tracking, and reliable counting of humans in real-time scenarios.

The fusion of YOLOv8 and DeepSORT represents a significant advancement in human recognition and tracking systems, offering a powerful solution for real-time human detection and tracking applications.Overall, the proposed system of human detection using YOLOv8 and DeepSORT aims to deliver a seamless integration of precise detection and effective tracking, enabling enhanced performance in scenarios requiring accurate human detection and tracking across video sequences.

3.3 SYSTEM ARCHITECTURE

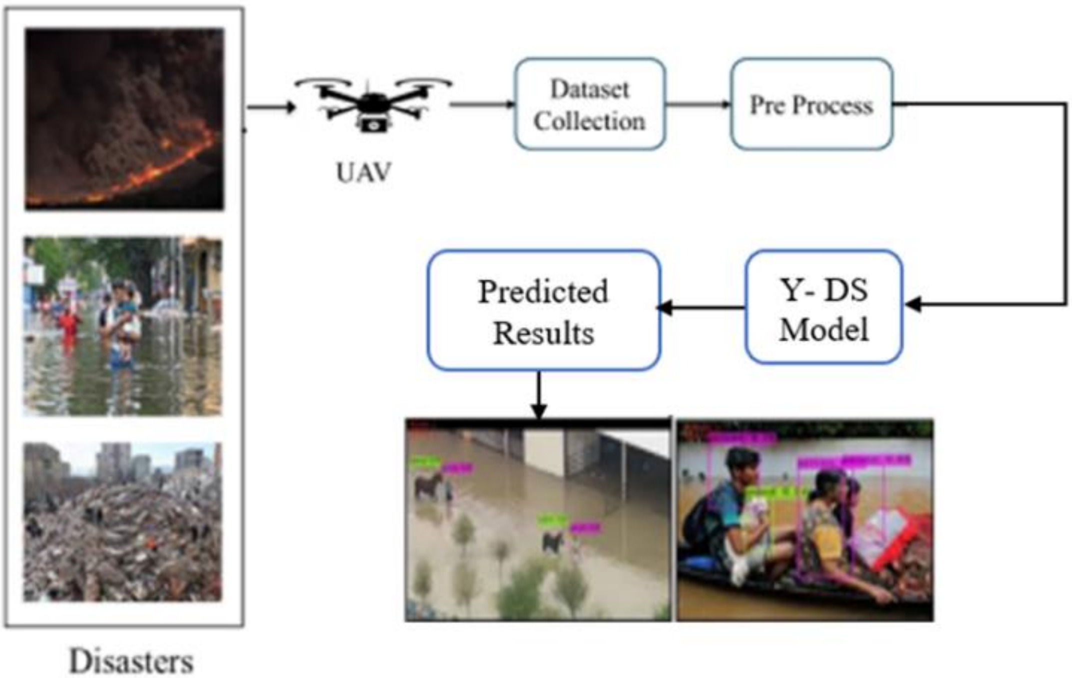

Figure 2.1. System Architecture of the proposed system

3.3.1. REAL-TIME HUMAN DETECTION AND LOCALIZATION

One of the key features of the system is real-time human detection and localization using deep learning algorithms. By processing the images captured by UAVs in real-time, the system can identify and locate individuals in disaster areas, enabling swift and targeted search and rescue operations. This capability is essential for saving lives and ensuring efficient response to catastrophic events.

3.3.2. OBJECT DETECTION WITH YOLOV8

YOLOv8 is a lightweight and efficient object detection model designed for real-time applications, suitable for systems with limited computational resources like mobile or edge devices.YOLOv8 utilizes a backbone architecture, typically a modified version of Darknet-53, to extract features from input images and a head architecture to translate these features into detectable objects and their attributes.

3.3.3 OBJECT TRACKING WITH DEEPSORT

DeepSORT is a state-of-the-art object tracking algorithm that assigns unique IDs to detected objects and tracks them across video frames.DeepSORT integrates with object detection models like YOLOv8 to provide continuous tracking of objects, ensuring accurate monitoring and tracking of humans in dynamic environments.The system features DeepSORT for assigning and maintaining unique IDs for objects, enabling effective tracking and monitoring of individuals in catastrophe scenarios.

3.4 DETAILED DESIGN

3.4.1 USE CASE DIAGRAM

A use case diagram is a visual representation within the Unified Modeling Language (UML) that illustrates the interactions between users (actors) and a system to achieve specific goals or functionalities (use cases). Actors are typically represented as stick figures, while use cases are depicted as ovals. The diagram showcases how users interact with the system to accomplish tasks, with arrows indicating the flow of communication or interaction. Use case diagrams help stakeholders understand the functionalities of a system, the roles of different actors, and the overall scope of the system's behavior. They provide a high-level view of the system's requirements and serve as a foundation for further system design and development. A use case diagram for human detection in disaster scenarios using YOLO and Deep SORT shown in the Fig 3.2 involves several entities: dataset collection which are images from UAV, preprocessing, annotation, training the human detection model and evaluation. In this project, we mainly focused on training and evaluating of the proposed model.

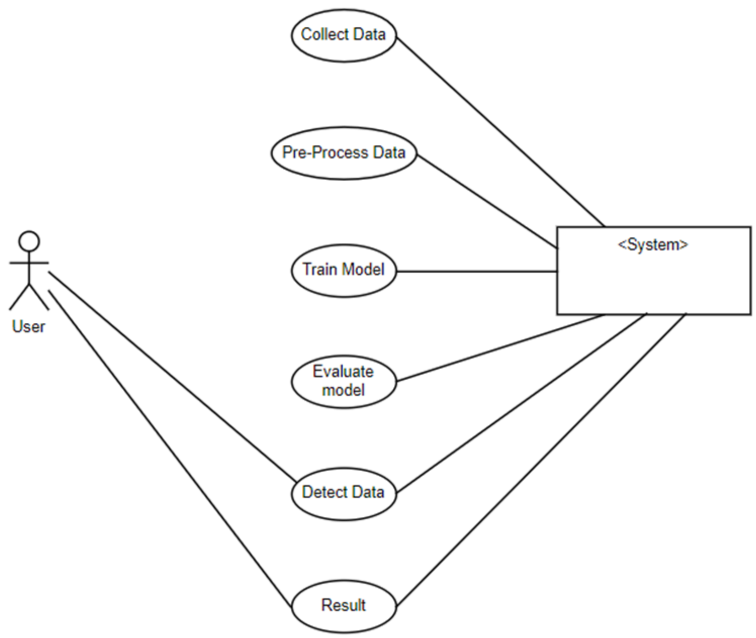

Figure 3.2. Usecase Diagram

3.4.2. CLASS DIAGRAM

A class diagram is a visual representation in the Unified Modeling Language (UML) that depicts the structure of a system by illustrating the classes, their attributes, methods, and the relationships between them. Classes are represented as rectangles, with compartments detailing their properties and behaviors. Associations between classes are depicted by lines, showing how they are related, and can include multiplicity and role information to specify the cardinality and role of each class in the relationship. Inheritance relationships show the hierarchy between classes, interfaces specify contracts that classes must implement, and dependencies indicate relationships where changes in one class may affect another. Class diagrams serve as a foundational tool for understanding the static structure of a system, aiding in communication between stakeholders, and guiding software design and implementations. In a class diagram Fig 3.3 provides a static view of the application, representing the types of objects residing in the system and their relationships. The class diagram consists of two primary classes: User and System, which represent the main entities in the system. The User class interacts with the System and can view the results of the model executed by the System. The System class encompasses attributes such as a human detection model, object track, frame, and boundary boxes. The System's operations include setting an ID, updating boundary boxes, collecting image datasets from the User, and detecting and tracking humans.

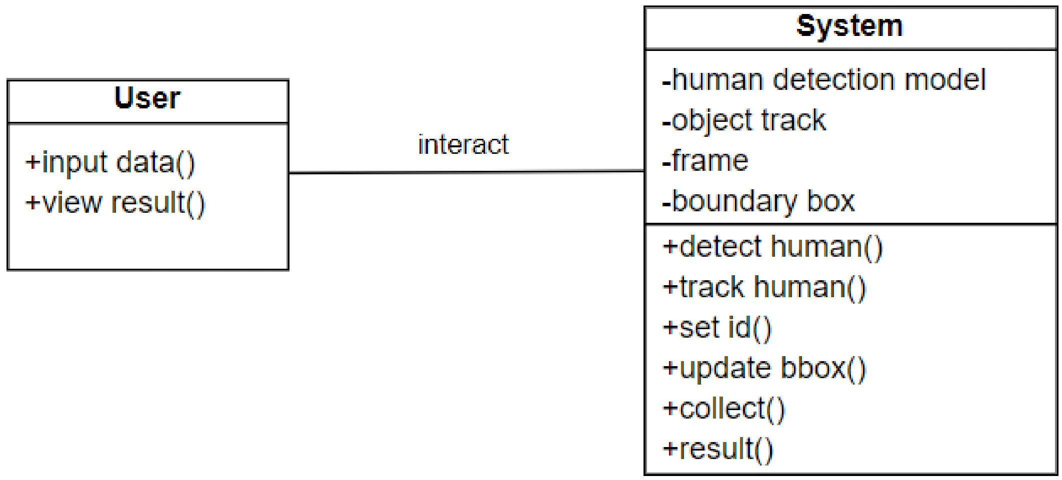

Figure 3.3. Class Diagram

3.4.3 SEQUENCE DIAGRAM

A sequence diagram is a type of interaction diagram in the Unified Modeling Language (UML) that illustrates how objects interact in a particular scenario or sequence of events within a system. It portrays the dynamic behavior of a system by showing the sequence of messages exchanged between objects over time. Objects are represented as vertical lifelines, with messages exchanged between them depicted as horizontal arrows. The sequence of messages is arranged chronologically from top to bottom, indicating the order in which interactions occur. Optional elements such as loops, branches, and parallel execution can also be represented to capture complex behaviors. Sequence diagrams provide a visual representation of the flow of control and data during the execution of a scenario, aiding in understanding system behavior, identifying potential issues, and guiding system design and implementation. The sequence diagram for the proposed system shown in the Fig 3.4 for the human detection in disaster scenarios illustrates the interaction flow between the user and the system. The system's operation involves collecting a dataset necessary for training the model, preprocessing the collected dataset, training and fine-tuning the model using the preprocessed data, testing the model, and predicting the presence of humans in an image or video stream sent by the user. The system then processes the data using the trained model and returns the result to the user.

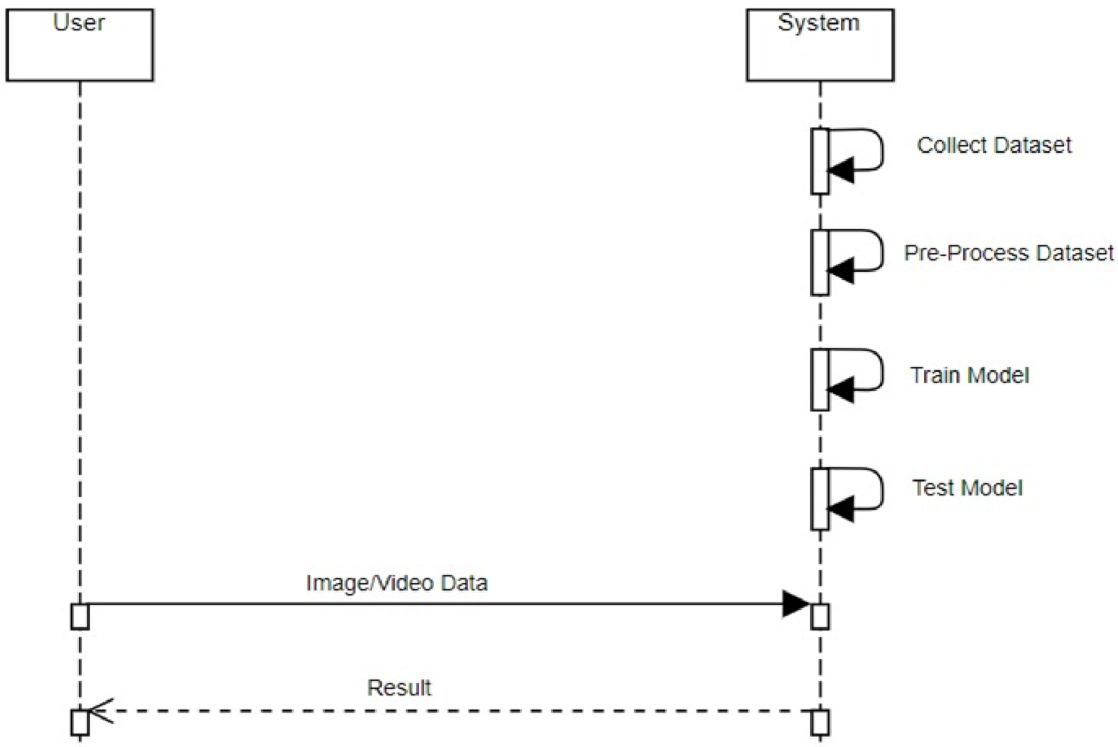

Figure 3.4. Sequence Diagram

3.4.4. ACTIVITY DIAGRAM

An activity diagram in the Unified Modeling Language (UML) is a behavioral diagram that represents the flow of activities or processes within a system. It provides a visual representation of the dynamic behavior of a system, focusing on the sequence of actions and decisions that occur to achieve a particular goal. Activities are depicted as rounded rectangles, with arrows indicating the flow of control between them. Decision points are represented by diamonds, branching the flow into alternative paths based on conditions. Forks and joins indicate parallel execution of activities and synchronization points, respectively. Activity diagrams help to model business processes, workflows, and system behaviors, aiding in understanding complex processes, identifying bottlenecks, and designing efficient systems capable of handling various scenarios. They serve as a valuable tool for communication between stakeholders and guide the development and implementation of software systems.

The activity diagram for the proposed system as shown in the Fig 3.5. for human detection in disaster scenarios using UAV and deep learning techniques illustrates a sequential flow of actions. The process begins with the system collecting a dataset, which is then preprocessed. After preprocessing, the dataset is utilized to train and fine-tune YOLO and Deep Sort model. The model's performance is evaluated, and if satisfactory, it is deployed for human detection. When a user sends an image or video for human detection, the system processes it using the trained model and returns the result.

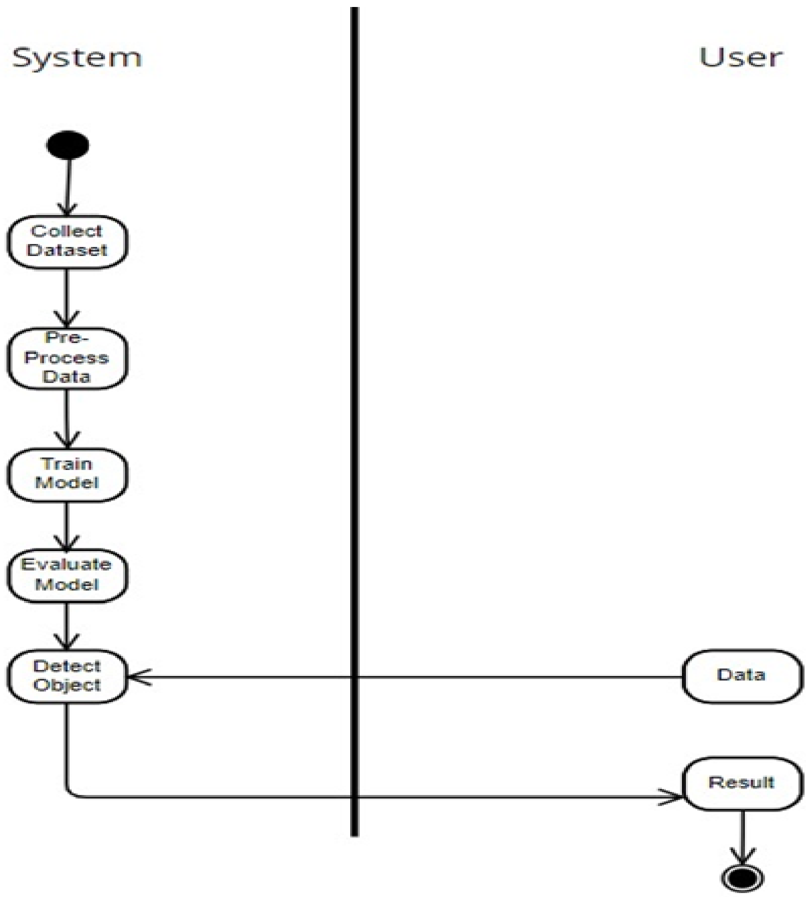

Figure 3.5. Activity Diagram

3.4.5 STATE DIAGRAM:

A state diagram, also known as a state machine diagram, is a behavioral diagram in the Unified Modeling Language (UML) that represents the dynamic behavior of a system in response to external stimuli or events. It depicts the various states that an object or system can be in and the transitions between those states triggered by events. States are represented as nodes, while transitions are depicted as arrows between the nodes, annotated with the triggering events and conditions. State diagrams also include initial and final states to represent the starting and ending points of the system. They provide a clear visualization of the possible behaviors and states of a system, helping to understand its functionality, identify potential problems, and design robust systems capable of handling various scenarios and transitions effectively. The state diagram shown in the Fig 3.6 entails 5 main states: “Dataset Collection”, “Pre-processing”, “Training Yolo model”, “Evaluating the model” and “Obtaining result”. The process begins with the system collecting a dataset, which is then preprocessed. After preprocessing, the dataset is utilized to train and fine-tune a deep learning model. The model's performance is evaluated, and it is deployed for human detection. When a user sends an image or video for human detection, the system processes it using the trained model and returns the result.

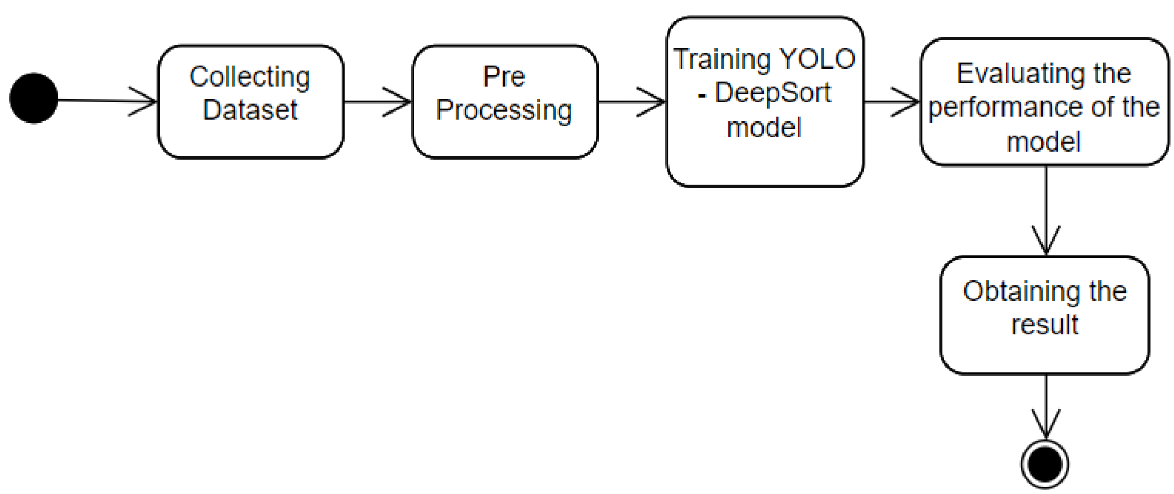

Figure 3.6. State Diagram

3.5 FUNCTIONAL REQUIREMENTS

3.5.1. LEVEL 0 DATA FLOW DIAGRAM

The level 0 DFD is shown in the figure 3.7 and it represents the process of the overall concept of the Human Detection System.

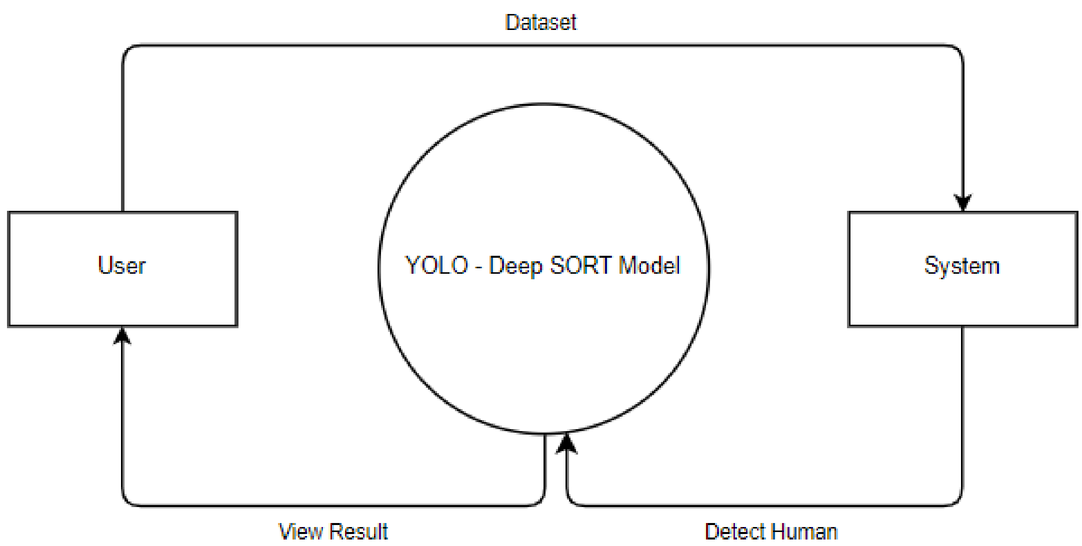

Figure 3.7. Level 0 Data Flow Diagram

3.5.2 LEVEL 1 DATA FLOW DIAGRAM

The level 1 Data Flow Diagram (DFD) shown in the figure 3.8 provides a detailed breakdown of the processes identified in the higher-level context diagram or Level 0 DFD.

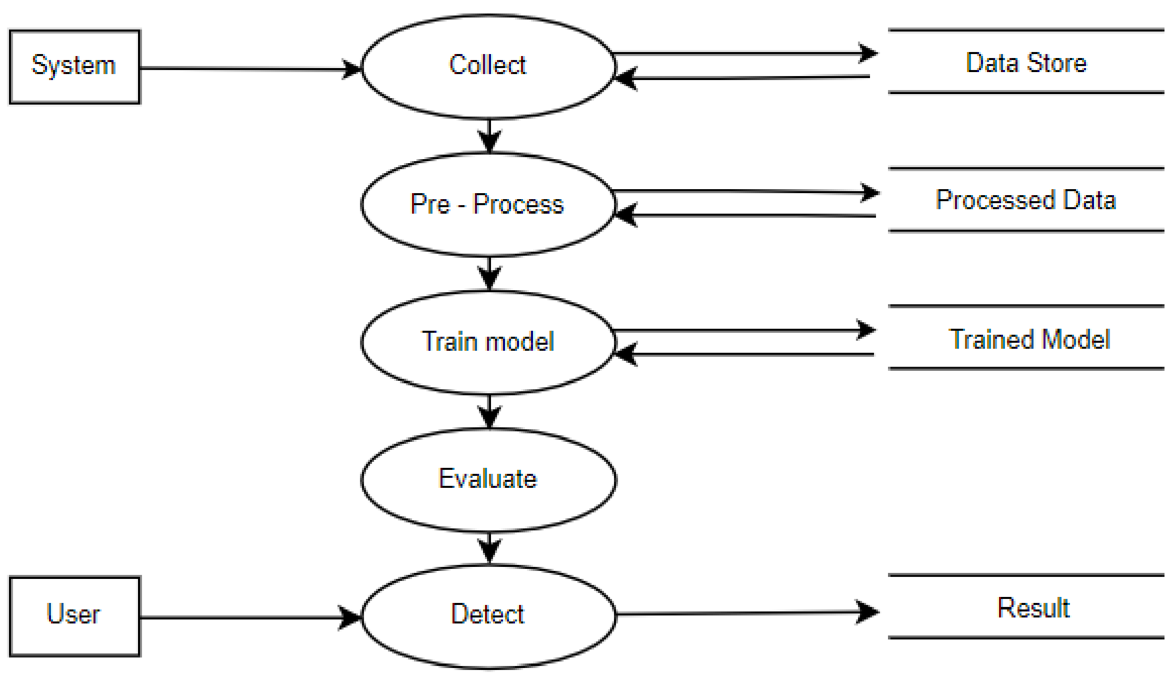

Figure 3.8. Level 1 Data Flow Diagram

CHAPTER 4

MODULE DECOMPOSITION

4. MODULE DECOMPOSITION

1. Data Collection
2. Pre-Process
3. Train Yolo-Deep SORT model
4. Evaluation of the model

4.1 DATA COLLECTION

In this module, images captured by the drone are collected and processed. The collected images are then categorized into positive and negative samples, which are used for training the deep learning model for human detection in disaster scenarios.

4.2 PRE-PROCESS

In this module, after cleaning and preprocessing the acquired dataset to enhance the model's performance, we proceeded with annotating the refined dataset. This annotation process involves the assignment of labels or annotations to the dataset. Various types of annotations, such as classification labels, bounding boxes for object detection, segmentation masks for image segmentation, and other relevant annotations, were meticulously applied during this phase.

4.3 TRAIN YOLO-DEEP SORT MODEL

In this module, the deep learning model object detection algorithms such as Deep SORT and YOLO (You Only Look Once) were used. YOLO is a popular object detection algorithm known for its high accuracy and efficient object detection mechanism, while Deep SORT is an extension of SORT (Simple Online Realtime Tracker) that incorporates deep learning techniques for data association.

4.4 EVALUATE THE MODEL

Comparing the model with other algorithms or approaches, assessing its real-world application in disaster scenarios, and testing its scalability and generalization capabilities are key factors in determining its effectiveness. In this module, we evaluate the model with performance metrics like accuracy, precision, recall, and F1 score to determine the efficiency of the proposed model.

CHAPTER 5

IMPLEMENTATION

5. IMPLEMENTATION

5.1 PYTHON

Python is a popular programming language. It was created by Guido van Rossum, and released in 1991.

It is used for:

- Web development (server-side),
- Software development,
- Mathematics,
- System scripting.

5.1.1. PYTHON LIBRARIES

Python serves as a fundamental toolset due to its extensive collection of libraries and frameworks tailored to the development and exploration of neural networks and advanced machine learning models. TensorFlow, crafted by Google, stands as a cornerstone in this field, providing an expansive platform for constructing and training various neural network architectures, including convolutional and recurrent networks, alongside transformers. PyTorch, a creation of Facebook's AI Research, offers an alternative approach with its dynamic computation graph and intuitive design, garnering widespread adoption among scholars and practitioners for its ease of use and flexibility.

5.1.2. OPENCV

OpenCV is one of the most widely used and comprehensive computer vision libraries, providing over 2,500 optimized algorithms for tasks like object detection, face recognition, and video processing. It is open-source, efficient, and supports multiple programming languages including Python, C++, and Java. OpenCV is used by major tech companies like Google, IBM, and Intel, and is suitable for a wide range of applications from surveillance to robotics.

5.1.3. TENSORFLOW

TensorFlow is a powerful open-source machine learning library developed by Google, providing a flexible architecture for deploying computation across a variety of platforms, from desktops to mobile devices. The Python API allows developers to easily build and train complex deep learning models, including capabilities like image recognition, natural language processing, and predictive analytics. TensorFlow integrates seamlessly with the broader Python scientific computing ecosystem, and offers a high-level Keras API to simplify model development. Beyond the core library, TensorFlow provides numerous extensions and add-ons for specialized machine learning domains, as well as tools for deploying trained models to production environments, including mobile and edge devices. With its comprehensive feature set, robust performance, and strong community support, TensorFlow has become a go-to choose for both research and real-world machine learning applications in Python.

5.2 NVIDIA GE FORCE

The NVIDIA GeForce drivers are essential software components that enhance the gaming experience and performance of PCs by optimizing graphics processing. These drivers can be downloaded from the official NVIDIA website, where users can access the latest versions to ensure smooth operation of their NVIDIA GPUs. It's important to choose the correct driver type, either "Standard" or "DCH," based on the system configuration to ensure compatibility and proper installation. NVIDIA regularly updates its drivers to provide improved functionality and support for various products like GeForce, TITAN, NVIDIA RTX, Data Center, and GRID. Users can also benefit from the GeForce Experience application included in the driver download, offering additional features and optimizations for gaming and graphics settings.

5.3 HARDWARE SPECIFICATION

- Device name: ASUS
- Processor: 11th Gen Intel(R) Core(TM) i5-11400H @ 2.70GHz 2.69 GHz
- Installed RAM: 8.00 GB (7.74 GB usable)
- System type 64-bit operating system, x64-based processor

5.4 SOFTWARE SPECIFICATION

- Platforms: Python IDE, Anaconda.
- Frameworks and Libraries: OpenCV, TensorFlow, Scikit-learn, PyTorch, Ultralytics.
- Data Source: Custom Dataset collected using UAV and Labelled.

CHAPTER 6

ADVANTAGES AND APPLICATIONS

6.ADVANTAGES AND APPLICATIONS

Advantages:

1. High Performance: YOLOv8 combined with DeepSORT/OC-SORT achieved high performance in tracking and time violation detection in parking areas, showcasing the effectiveness of this combination.
2. Improved Tracking Accuracy: Integrating DeepSORT with YOLOv8 resulted in enhanced tracking accuracy and robustness, demonstrating the effectiveness of this combination
3. Efficient Object Detection: YOLOv8 is particularly effective at handling challenges like occlusions, varying scales, and complex backgrounds, making it suitable for real-time car detection applications
4. Seamless Integration: YOLOv8 and DeepSORT work seamlessly together, with YOLOv8 providing precise detection and DeepSORT offering effective tracking across video sequences

Applications:

1. Vehicle Detection and Tracking: YOLOv8 and DeepSORT are utilized for car detection and tracking in intelligent traffic systems, demonstrating effectiveness in handling challenges like occlusions and complex backgrounds.
2. Object Tracking in Videos: Research proposes real-time and recorded video-based object detection and tracking using YOLOv8 and DeepSORT, highlighting their critical capabilities for computer vision systems
3. Real-Time Object Tracking: YOLOv8 and DeepSORT are combined to implement multiple object tracking in Python, showcasing their powerful combination for object tracking applications.

CHAPTER 7

RESULT

7. RESULTS

7.1. SAMPLE CODE

YOLO.py

```
import os
from ultralytics import YOLO
# Load a model
model = YOLO("yolov8n.yaml")  # build a new model from scratch
# Use the model
results = model.train(data=os.path.join(ROOT_DIR, "google_colab_config.yaml"),
epochs=200)  # train the model
```

Main.py

```
import os
from ultralytics import YOLO
import cv2
VIDEOS_DIR = os.path.join('/content/drive/MyDrive/YOLOv8 Model with
custom dataset', 'videos')
video_path = os.path.join(VIDEOS_DIR, 'penn4.mp4')
video_path_out = '{}_out.mp4'.format(video_path)
cap = cv2.VideoCapture(video_path)
ret, frame = cap.read()
H, W ,_ = frame.shape
out = cv2.VideoWriter(video_path_out, cv2.VideoWriter_fourcc(*'MP4V'),
int(cap.get(cv2.CAP_PROP_FPS)), (W, H))
model_path = os.path.join('.', 'runs', 'detect', 'train', 'weights', 'last.pt')
# Load a model
model = YOLO(model_path)  # load a custom model
threshold = 0.5
```

```
while ret:
    results = model(frame)[0]
    for result in results.boxes.data.tolist():
        x1, y1, x2, y2, score, class_id = result

        if score > threshold:
            cv2.rectangle(frame, (int(x1), int(y1)), (int(x2), int(y2)), (255, 255, 255), 4)
            cv2.putText(frame, results.names[int(class_id)].upper(), (int(x1), int(y1 -
10)),
                        cv2.FONT_HERSHEY_SIMPLEX, 1.3, (0, 255, 0), 3,
cv2.LINE_AA)
    out.write(frame)
    ret, frame = cap.read()
cap.release()
out.release()
cv2.destroyAllWindows()
```

7.2 OUTPUT

7.2.1. HUMAN DETECTION AND TRACKING

7.3 TESTING

7.3.1 TEST CASE 1

Detection with limited person

7.3.2 TEST CASE 2

Detection with unlimited persons

CHAPTER 8

CONCLUSION AND FUTURE ENCHANCEMENT

8. CONCLUSION AND FUTURE ENCHANCEMENT

Human detection in disaster scenarios highlights the critical role of advanced technologies and methodologies in enhancing search and rescue operations. The research and development in human detection techniques, such as the use of hybrid algorithms, convolutional neural networks, and thermal imaging, have shown significant promise in improving the efficiency and accuracy of locating disaster victims. These techniques, combined with multi-sensor technologies like drones and robotics, offer a comprehensive approach to detecting and rescuing trapped victims under debris in disaster situations.

The need for integrated sensor technology, machine learning, and 5G networks underscores the importance of real-time data collection and analysis for swift and precise human detection in challenging environments. Overall, the evolution of human detection systems towards multi-sensor integration and advanced technologies is crucial for enhancing the speed, accuracy, and effectiveness of search and rescue missions in disaster scenarios.

REFRENCES

[1] Korteling JEH, van de Boer-Visschedijk GC, Blankendaal RAM, Boonekamp RC, Eikelboom AR. Human- versus Artificial Intelligence. Front Artif Intell. 2021 Mar 25; 4:622364. doi: 10.3389/frai.2021.622364.

[2] M. I. Jordan, T. M. Mitchell, Machine learning: Trends, perspectives, and prospects. Science349,255-260(2015). DOI:10.1126/science.aaa8415

[3] Nasir Abdul Jalil, Ha Jin Hwang, and Norazryana Mat Dawi. 2019. Machines Learning Trends, Perspectives and Prospects in Education Sector. In Proceedings of the 3rd International Conference on Education and Multimedia Technology (ICEMT '19). Association for Computing Machinery, New York, NY, USA, 201–205. https://doi.org/10.1145/3345120.3345147

[4] Yann LeCun, Yoshua Bengio & Geoffrey Hinton. Deep learning. Nature 521, 436–444 (2015). https://doi.org/10.1038/nature14539

[5] S. J. Pan and Q. Yang, "A survey on Transfer Learning," IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345–1359, Oct. 2010. doi:10.1109/tkde.2009.191

[6] Sarker, I.H. "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions". SN COMPUT. SCI. 2, 420 (2021). https://doi.org/10.1007/s42979-021-00815-1

[7] M. Abdel-Basset, H. Hawash, R. K. Chakraborty, and M. Ryan, "Energy-Net: A Deep Learning Approach for Smart Energy Management in IoT-Based Smart Cities," IEEE Internet of Things Journal, vol. 8, no. 15, pp. 12422–12435, Aug. 2021, doi: 10.1109/jiot.2021.3063677.

[8] N. Shlezinger, Y. C. Eldar and S. P. Boyd, "Model-Based Deep Learning: On the Intersection of Deep Learning and Optimization," in IEEE Access, vol. 10, pp. 115384-115398, 2022, doi: 10.1109/ACCESS.2022.3218802

[9] M. Al-Qatf, Y. Lasheng, M. Al-Habib and K. Al-Sabahi, "Deep Learning Approach Combining Sparse Autoencoder With SVM for Network Intrusion Detection," in IEEE Access, vol. 6, pp. 52843-52856, 2018, doi: 10.1109/ACCESS.2018.2869577

[10] V. Monga, Y. Li and Y. C. Eldar, "Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing," in IEEE Signal Processing Magazine, vol. 38, no. 2, pp. 18-44, March 2021, doi: 10.1109/MSP.2020.3016905

[11] Yamashita, R., Nishio, M., Do, R.K.G. et al. "Convolutional neural networks: an overview and application in radiology". Insights Imaging 9, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9

[12] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.

[13] Introduction to convolution neural network, Geeks for Geeks. Available at: https://www.geeksforgeeks.org/introduction-convolution-neural-network/

[14] P. Yang, X. Sun, W. Li, and S. Ma, "Automatic academic paper rating based on modularized hierarchical convolutional neural network," Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2018. doi:10.18653/v1/p18-2079

[15] Dina Chahyati, Mohamad Ivan Fanany, Aniati Murni Arymurthy, "Tracking People by Detection Using CNN Features", Procedia Computer Science, Volume 124,2017, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2017.12.143

[16] Z. Shi, Y. Ye, and Y. Wu, "Rank-based pooling for deep convolutional Neural Networks," Neural Networks, vol. 83, pp. 21–31, Nov. 2016. doi: 10.1016/j.neunet.2016.07.003