Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 3, Issue 11, 2025

WEAPON SURVEILLANCE AND DETECTION SYSTEM USING YOLO-NAS: A NOVEL APPROACH TO PUBLIC SAFETY

Muhammad Waqas Khaliq[*1], Muddasar Ahmed[2], Dr. Muhammad Adnan[3], Dr. M. Irfan Uddin[4]

[3,4]*Institute of Computing, KUST Kohat*

DOI: https://doi.org/10.5281/zenodo.17732760

Keywords

Article History
Received: 12 October 2025
Accepted: 10 November 2025
Published: 21 November 2025

Corresponding Author: *
Muhammad Waqas Khaliq

Abstract

A paradigm shift in the globe is the deepening apprehension of public safety in crowds demands the invention of modern weapon surveillance and detection methods. This thesis offers a new weapon monitor and detection system named YOLO-NAS based Weapon Surveillance and Detection System (WSDS). The system is powered with advanced deep neural network technology and contemporary computer vision algorithms that allow for faster, more accurate weapon detection, overcoming issues like partial or full occlusion and spatial resolution limitations. The study aims to provide a theoretical foundation by presenting the object detection concept in computer vision, explaining the YOLO-NAS architecture, and investigating the methods utilized by real-time monitoring systems. The methodology involves: data gathering from multiple sources, dataset preparation, the training of the YOLO-NAS model, and the performance evaluation by the use of metrics such as precision, recall, and mean average precision. The YOLO-NAS structure embeds existing surveillance cameras, utilizing the central processor to execute dynamic detection thereby triggering the alarm system to activate the response routine immediately. The study measures the capacity, customizing, and the fitness of system to the public and utilities such as schools, shopping areas, and transit stations. The results and findings part presents the effectiveness of the method and makes a comparison with those that have been mentioned before. The thesis describes how this solution reshaped current practices in the area of public security. The thesis concludes with the identification of the project's limitations providing recommendations for improving these and stressing the fact that the WEAPON SURVEILLANCE & DETECTION SYSTEM based on YOLO-NAS let us move to a new level of modern weapon surveillance systems.

INTRODUCTION

Over the last decade, the single task of guaranteeing the security and safety of an average citizen in public spaces has greatly evolved in reputation and status. The upcoming dangers arising from terrorism, mass shootings, and other violent crimes highlight the highly urgent necessity for modern technologies and creative security approaches to thwart public safety. As opposed to the traditional neural models, this research takes a bold step toward the cutting-edge YOLO-NAS (Neural Architecture Search) WSDS (Weapon Surveillance and Detection System), an unprecedented application of AI (Artificial Intelligence) enablers whose sole purpose is to detect and identify weapons in real-time.

However, The YOLO-NAS system has a huge potential to improve the existing military surveillance and gives an acceptable part in privacy and in time. Design for the operation within high risk settings like schools, shopping centers, and public transport hubs and the technology saves law enforcement and security staff from dealing with many threats. Furthermore, the emergency teams can be able to able to identify the different firearms and do a timely response to risks as speedily and effectively as possible.

Despite the impressive results of the last YOLO versions in detecting the weapons, there are some restrictions still essential. An issue is the degrading performance due to using data that is insufficient and unbalanced classes. Also, small objects cannot be detected due to the constraints of objects. Though these systems operate at an object level, they are unable to identify specific objects in the scene. For instance, some existing AI detection systems have not produced accuracies in image blurring, reduced contrast, and color distortion, dependency on high quality training data, and sensitivity to variations of different angles which are the current need to overcome these challenges. Hence, this study aims to overcome this challenge by adopting an advanced object identification system based on a deep neural network to accurately pinpoint objects in either images or videos. The YOLO-NAS model, instantiated with pre-trained weights, carries on the distinguished quality of meticulously detecting the tiniest objects with an exceptional degree of detail. YOLO-NAS achieves this level due to its ability to do away with the process of breaking an image into smaller segments, thus performing a fast and accurate recognition of objects in complex scenarios where there are diverse objects and frames involved. Its adaptability is an essential advantage; it can readily let a vast variety of these and other things be attached and adjusted to suit the different types of weapon systems and environmental conditions. In this thesis, the development milestones, implementation of the YOLO-NAS-based WSDS, and evaluation of the system will be illustrated. As a result, there is not only a concentration on the technicalities of the system but also the impact it will have on the real-world safety of the public. The adoption of computer vision high tech by law enforcement departments may be revolutionized in the sense that surveillance systems will be endowed with the effective tools necessary for the identification of dangerous elements and the prevention of their possible attacks on the public.

This chapter follows the aftermath of the explanation of the theoretical background, the methodology, and the architecture of the system is discussed. Different results are obtained and the methodological implications end up. This study is a part of the ongoing dialogue going around the issue of personal safety assistance with the complete examination of the innovative method of weapon detection in public places of mass with crowds.

Problem Statement

The current security situation demands that weapon monitoring and detection systems should be efficient and reliable to students as a primary research question. The incidents of shootings in crowded areas like schools, shopping malls, or transportation hubs is increasing significantly, which is only testifying to the inefficacy of current surveillance technologies. The traditional way of surveillance, comprising of manual examination, or Close Circuit Televisions (in short CCTV), results in unmanageable problems mostly at the time responding to serious incidents.

Time to reach the right information and creativity with accuracy in spotting and replying to arms in real-time is the main problem. Traditional video surveillance systems are going to suffer from this kind of missed detection in which law enforcement officers will not be able to quickly interact with the public and prevent potentially dangerous situations. Existing additional effective systems restrictions require a novel and advanced style that can overcome these limitations and will have a huge contribution to public safety.

The goal of this research consists of making improvements in the existing systems of weapon surveillance by presenting and evaluating the YOLO-NAS based Weapon Detection and Surveillance System. The stated issues include insufficient target accuracy stemming from data deficiency and class imbalance as well as the system's imperfect response when detecting hidden weapons of small size due to space resolution limit. The research aims to address these difficulties, providing law officers with a more dependable and speedy method to check for weapon presence in quick and dense public spots.

Consequently, the issue statement addresses the necessity to create a system that can avail the tactical reaction to the development of cases in public spaces. This YOLO-NAS architecture looks to fill the existing gaps and, through a technologically advanced model, will provide the needed solution that can substantially contribute to the elimination of gun violence and the protection of the general public.

Research Objectives

The proposed work's main objectives are:

The intended objective for this study is the creation and the deployment of a validated YOLO-NAS based Weapon Surveillance and Detection System that is able to counter the shortcomings of the existing surveillance solutions and support safer streets. The specific objectives guiding this study are as follows:

1. **Increase Weapon Recognition Accuracy:** By incorporating deep neural networks and state of the art computer vision technology into the process for recognizing weapons this increased the accuracy of the system, despite the data limitations and the solution to the class imbalances problem.
2. **Reduce Occlusion and Improve Detection of Tiny Objects:** Overcome the issues prompted by cloud outs or restricted spatial resolution especially where there are tiny objects like concealed firearms. Take up this and a more comprehensive surveillance capability is achieved.
3. **Provide Guidance to Law Enforcement Agencies:** Give a point of viewpoints and suggestions for the application of the research to police forces that would allow them to make informed decisions and to construct public areas that are free of crime and violence.

These goals are to be accomplished only by which, a comprehensive and self sustaining surveillance system of weapons emerges capable of having an accurate alert and response on time. Through addressing issues arising from both technology development and practical use, this research aims to give a picture of the real world solution for the shortcomings of the existing systems used for public safety in circumstances where the space is congested and dynamic.

Scope of the Study

The last issue I want to address is the width of our research, which is of particular importance. This study, which is specially meant for "YOLO-NAS based weapons surveillance and detection system making the application of the system to public safety in crowded places", is the main subject of investigation and a subsequent process of implementation. The scope encompasses several key dimensions:

4. **The scope encompasses several key dimensions:** Application in Various Public Settings: The article discusses the ways and possibilities of the system working under the circumstances of different public places, including the clubs, malls, and stations of transport. Having the system be able to adapt to any environment becomes crucial for its practical fitting.
5. **Identification of Multiple Weapon Types:** As the research examines the YOLO-NAS system for weapon recognition of various types which may involve firearms, knives, and explosives. This sets out to offer an all-around surveillance system that can identify risks of many kinds.
6. **Real-time Detection and Response:** The goal is materializing fast weapon recognition that may be used immediately to search for and react upon the place where law enforcement personnel and security guards are present. Quicken real-time notification and suggestions, which will largely determine how quickly and how effectively counter threats will be mitigated.

Scalability and Integration: Scalability of the YOLO-NAS is seen to be an important aspect that can enable the system to meet surveillance needs that have a lot of cameras in those situations. Moreover, the question of interoperability with current security communication systems is also examined to guarantee smooth operation.

Data Sources and Diversity: A unique data source is used and many sources such as the managed database COCO and VOC are gotten through the public domain to build and analyze the YOLO-NAS model. This means that the goal is to make the system as a whole capable of handling any condition or situation it finds itself.

Performance Metrics: The system is to be assessed based on the basic metric parameters such as precision, recall and mean average precision (MAP). These metrics enable one to know the system performance in terms of error and success in a weapons detection standpoint.

Significance of the Study

The significance of this study lies in its potential to revolutionize the landscape of weapon surveillance and detection, addressing critical challenges in public safety. The research is of paramount importance for several reasons:

Advancing Public Safety Technology: The proposed YOLO-NAS based Weapon Surveillance and Detection System represent a cutting edge application of computer vision technology. Its potential to accurately and swiftly identify weapons in real-time has profound implications for public safety, providing a proactive approach to threat mitigation.

Mitigating Gun Violence: In the face of rising incidents of gun violence in public spaces, the study's findings and the implemented system aim to contribute significantly to the mitigation of such incidents. The real-time detection and immediate response capabilities have the potential to save lives and prevent tragic consequences.

Informing Law Enforcement Strategies: By providing insights and recommendations to law enforcement agencies, the study contributes to the formulation of informed strategies for public safety. The system's data-driven approach and adaptability offer a valuable resource for optimizing security measures in diverse environments.

Technological Innovation in Surveillance: The research adds to the body of knowledge surrounding object detection and surveillance technologies. The innovative use of YOLO-NAS in weapon detection sets a precedent for the integration of advanced computer vision techniques into security systems.

Applicability in Various Public Settings: The scope of the study, encompassing different public settings, ensures that the developed system is versatile and applicable across a range of scenarios. This adaptability increases its utility for law enforcement in addressing security concerns in diverse environments.

Contributing to Academic Research: The study contributes to academic research by addressing existing gaps in the literature related to weapon surveillance. The findings and methodologies employed can serve as a foundation for further exploration and refinement of weapon detection systems.

Ethical and Responsible Technology Use: The study emphasizes the ethical implications of weapon surveillance and detection technologies. By focusing on public safety, it promotes responsible and accountable use of advanced technologies in safeguarding communities.

Research Questions

To guide the investigation and exploration of the YOLO-NAS-based Weapon Surveillance and Detection System, the following research questions form the core inquiry framework of this study:

How can deep neural networks and contemporary computer vision technology be effectively utilized to increase the accuracy of weapon recognition in crowded public spaces?

In what ways can the challenges of occlusion and spatial resolution limitations be mitigated to improve the detection of tiny objects, particularly small firearms, within a surveillance system?

How can the YOLO-NAS-based Weapon Surveillance and Detection System offer valuable insights and recommendations to law enforcement agencies for enhancing public safety in various public settings?

These research questions encapsulate the multifaceted aspects of the study, ranging from the technical aspects of the YOLO-NAS model's performance to its practical implications for law enforcement and public safety. By addressing these questions, the study aims to contribute nuanced insights into the efficacy and potential challenges of the proposed weapon surveillance system.

Chapter 2 Literature Review and Background Study

Research on the historical context and a survey of the literature make up this chapter. The background study covers the key topics required to understand this research approach. The review of the literature talks about many approaches to dealing with the accuracy problem.

Description of the Surveillance Weaponry Technologies

Various types of surveillance equipment have taken strides in the technologies to meet the purpose of security amidst the growing threats to the safety of the public. Outdated surveillance tools, exploiting human observation and closed circuit TV networks, have impeded the process of detection of hidden weapons and neutralization of possible threats [1]. The third section of the article offers a general outline of the weapon surveillance technology, focusing on the breakthroughs and the transition to the new AI computer vision technologies.

Traditional Surveillance Methods

Developing a large number of different types of technologies in the past, surveillance of weaponry was done mainly by humans who had a bit of a chance to look at some old cameras. Such ways, though they can meet some deterrence, are still short of the situation because of the complexity of modern security challenges. Monitoring manually has its setbacks with the problem of having to attend to everything continuously which can lead eventually to missing some threats [2]. To be more exact CCTV systems had such a foreseeable prospect of progress, providing remote monitoring. Tipsily, old CCTV systems could be facing the challenge of officially intercepting and following up weapons, especially in crowded or dynamic conditions.

From classical to modern computer vision approaches

The invention of computer vision technologies not only brought about a new approach to combat surveillance but also imposed newer challenges. Object detection algorithms capable of determining and location of objects inside camera images or video frames have become a vanguard in expanding the surveillance application level.

Region-based CNNs (R-CNN)

R-CNNs, a region-based CNN, a breakthrough, has introduced the object detection concept of region proposal. R-CNN proposes removing the shortcomings of classic methods by segmenting the image into regions and performing CNNs on each region. Unlike R-CNN performance was excellent but unfortunately, it had hardware time complexity limiting real-time applications [3].

Faster (R-CNN)

Faster R-CNN is another model of region-based convolutional neural networks. It can be decomposed into four parts. The first part is the feature extractor, which takes as an input an image and outputs the most distinctive information out of the raw input, represented as a convolutional feature map. The feature extractors used were pre-trained on ImageNet for classification tasks. To use them with Faster R-CNN, the last layers that are related to the classification task have been removed. Different CNN architectures are tested. These architectures are ResNet50 [10], Inception ResNetV2, VGG16 and finally MobileNetV2 [4].

As a result, The results obtained show that the model with the highest mAp among all the tested models was Faster R-CNN with Inception-ResNetV2 as a feature extractor. It achieved a mAP of 81% with the lowest log-average miss rate. However, the training and test time is unsuitable for real-time applications, as it has the longest average testing time among other models [4]. In future work, we must improve the system detection by enabling the detection of other types of weapons such as bombs. In addition, we propose to address the problem of gun classes, where the system identifies the gun type, such as Rifle, Revolver, etc.

You just look at once (YOLO)

YOLO (You Only Look Once) did not only capture the zeitgeist of object detection by suggesting the unified model that can simultaneously recognize bounding boxes and class probabilities without overlapping. YOLO accomplished real-time processing with high precision levels that its predecessor did not.

Comparison with different models

A big comparison has been taken by using different advanced algorithms. The flying bird dataset for surveillance films addresses the creation and performance assessment of flying bird recognition algorithms in surveillance films. There are 483 video clips in this dataset, totaling 28,694 frames. Whereas 28,366 occurrences of flying birds are present in 23,833 frames. The suggested collection of flying birds in surveillance movies is gathered from actual surveillance settings, and the birds display attributes including small sizes, form diversity during flight, and inconspicuous features in single frames (in certain cases). These characteristics present difficulties that must be considered when creating flying bird detection techniques for security footage. Ultimately, sophisticated (video) object identification algorithms were chosen for testing on the suggested dataset, and the outcomes showed that the dataset is still difficult for the aforementioned algorithms. as a result, SELSA (Sequence Level Semantics Aggregation) has shown better results as com- pared to all other object detection alogithms like YOLO family series, SSD, FGFA etc. [5]

Challenges and Advancements
It still had some problems, though, shown in earlier fire monitoring systems. Problems involving a lack of data to guarantee precision of model and the imbalance in the data set, alongside detection of objects of small size, appear to entice a new search for the solutions. YOLO-NAS, which is based on Neural Architecture Search (NAS) technology, is the emblem of how the sector is fighting these hurdles and optimizing the use of weapon surveillance technology as well as their effectiveness [6].

Evolution of object detection algorithms One of the most exciting developments in data driven systems
Object detection algorithms are considered to be the basis of advanced computer vision schemes providing all for image categories extraction and object localization [7]. The algorithms' evolution has seen many different versions which all tackle the lacks and inconveniences of the earlier versions to improve the precision, speed, and flexibility of the approach. It is this section that gives the reader an insight into the significant changes in object detection from yesteryears to the present day, this marks the introduction of the YOLO-NAS modeler in weapon surveillance.

Early Methods: Normalize Haar Cascades
Initial object detection algorithms, for example, Haar cascades, where features were cascaded became possible due to the introduction of feature cascades theory. Using these techniques objects were recognized based on the two dimensions of features (features were determined in/defined prior and a classifier was sequential). Unlike the magic solutions, they are best suited for narrow purposes, and the lack of effectiveness for handling multiple object classes, different scale covers, and direction variety (horizontal and vertical views) constituted the main problem.

Region-based CNNs (R-CNN)
The R-CNN (Region-based Convolutional Neural Networks) has been a new automated feature that has changed the way of object detection [8]. Similarly with R-CNN, a two-stage method has been presented which first extracts bounding boxes in the initial step and next, passes one or more boxes of the CNN to each proposal It overcame the flaw in accuracy but had an issue with being computationally unfriendly thus, this limited use in a real-time environment.

Two things many say (YOLO)
The YOLO algorithm brought a many-in-one model into object detection pipelines with the capacity to produce the same time bounding boxes and class probabilities on one object. YOLO broke the input image into a grid and labeled it as an object attribute within each grid cell [9]. Such a single shot employed the instant algorithms that allowed it to have real-time performance, making it ideal for applications, where speed is the prior factor, such as video surveillance.

Faster R-CNN and Single Shot Multi Box Detector(SSD)
The Faster R-CNN stands on the foundation of R-CNN and YOLO blended RPNs as a way to

decrease the speed of region proposal generation. All in all, this approach not only boosted the effectiveness accuracy but also enabled the model to run fast. Another significant advancement in object detection was when Single Shot Multi Box Detector (SSD) was introduced. It proposed detecting bounding boxes for objects present at different scales within a single network. SSD rather follows the rule of "fast but still accurate", an integrated solution that can be applied for a wide range of applications [10].

YOLO-NAS: Neuromorphic Computing

Several exceptional technologies are emerging that have the potential to play a critical role in the next wave of technological evolution. One of these technologies is Neural Architecture Search (NAS), which is widely used in the realm of neuromorphic computing [11]. The evolution of object detecting algorithms did not stop until a You Only Look Once Neural Architecture Search (YOLO-NAS) was developed. Through a Neural Architecture Search, YOLO-NAS seeks out the best network configurations by searching the space of all possible network architectures, thus providing higher accuracy with less computational effort. This method is not just about applying but it also aims at identifying new architectures that address specific tasks in a creative, more advanced way such as weapon surveillance [12].

YOLO-NAS: Weapon Detection

A Short but In-depth Explanation of the 'Neural Architecture Search Algorithm'. The YOLO Neural Architecture Search (NAS) is one of the breakthrough approaches to object detection and object detection, it is built specifically for small firearms detection and recognition. The purpose of the section is to provide a detailed analysis of YOLO-NAS, which will focus on its unique features and the practical use of it applicable in weapons detection.

Overview of YOLO-NAS

Neural Architecture Search (NAS)

Yolo-NAS neural architecture is based on Neural Architecture Search, a technique that allows the neural network structures to be optimized automatically. With the help of NAS, the boundaries of complex manually made designs can be extended to include scanning and immersion across the space of possible models and testing an optimal one by supplying the algorithm with the objective, such as accuracy and speed.

YOLO-NAS for sense processing through Object Detection

Through nasal integration, we derive YOLO-NAS which is an application within a YOLO architecture. This model uses the automation component of search to extract existing networks which increases the object detection accuracy and efficiency. YOLO-NAS features a great deal of applications regarding real-time operations, making them a perfect detection system for the weapon's outlook [13].

Key features of YOLO-NAS

Automatic Architecture Exploration

The main thing of YOLO-NAS is its capability to produce the optimum set of network architecture by automatically looking for results that are valid for the particular task. In this case, the model is allowed to perform this task on its own, obviating the need of hand designing it, with the ability to adapt, depending on particular requirements such as the complication in detecting weapons in a crowded and dynamic environment [9].

Precision and Efficacy are at all time high

Through that, YOLO-NAS will be looking for a compromise between precision and memory. The model strives for topologies that show good performance in terms of target detection accuracy together with a high frame rate processing [14]. The realities of the adoption of the system on site and the situation where immediate response as a prime necessity is no different than the parallel emphasis.

Application in Weapon Detection

The use of the system in weapon detection will ie The system will be used in the detection of weapons. In the surveillance of weapons, one issue that should be addressed is the space the agency will use to store the weapons in or the security and handling of such driven weapons will be a primary concern. YOLO-NAS aims to tackle the major drawbacks of weapon observation systems such as occlusion that interferes with detection and spatial resolution that is limited to seeing from a long distance [11]. Tackling this, the automatic architecture search design process makes the identification of structures that are good at detecting weapons and can successfully work even in complex scenarios.

Real-Time Weapon Detection

The real-time processing abilities of YOLO-NAS which enable it to be a valuable detection tool for use in dynamic environments to look out for weapons [15]. Developing a system that adjusts its architecture to the unique requirements of the weapons observation, YOLO-NAS will be able to do selection fast and accurately, thus providing a detection ahead the risk management.

Object Detection landscape

Plays an integral part and the identified contributions will therefore play a major role in the imaging technology. YOLO-NAS not only highlights such a possibility, but it becomes a core contributor to the broader autonomous vehicle technologies landscape by showcasing the addictiveness of the automated architecture search in improving real-time object detection. This way, we can observe gun detection (where NAS exhibits the ability to adapt to niche tasks) as an illustration of the advances in technology. In summation, YOLO-NAS comes across as an avant-garde and robust Object detection algorithm, primarily meant for weapons surveillance [16]. Neural Network Architecture Search (ReNAS) is a part of that AI which is automated optimization and goes through many empirical investigations in the proceeding chapter.

Recent Improvements in Computer Vision for Managing Public Security

The progress of computer vision technologies has been key to the evolution of the public safety domain. The adoption of modern technologies, like deep learning and CNNs, has impeached the functionality of surveillance systems with their inbuilt predictive and analytical features. It is dedicated to investigating the most essential breakthroughs implementing the crime prevention tools in this phase.

Deep learning for Surveillance

Convolution Neural Network (CNN)

With the advent Of Convolution Neural Networks, computer vision has seen a pivotal role that it plays in computer vision applications. Generative Adversarial Networks (GANs), in the scope of public safety, allow object detection to be highly sophisticated allowing a surveillance system to properly recognize and track the object(s) of interest. The major reasons why CNNs work on complex scenarios of face recognition are owing to the hierarchy and the learning feature capabilities [17].

Deep Feature Extraction

High-level feature extraction approaches help to emphasize the picture's understanding of visual information, which provides a possibility to acknowledge significantly complex patterns and unusual things. As a result of automatically discovering the hierarchical features, these techniques can assess the objects and events typical of public security.

Real-time Surveillance Systems

You One Look Once (YOLO)

The animals in the solar system experience the gravity of the planet they are orbiting or landing on. YOLO's synchronized detection and classification enables smart video stream analyses, thus, streamlining the process of immediate reaction is the strength of this career in applications where time is a critical issue, including public safety and security.

Table 2.1: Comparison of different approaches in tabular form

Paper Name	Author Name	Model used	Results	Limitations

Weapon Detection Using YOLO V3 for Smart Surveil- lance System (2021)	Sanam Narejo	YOLOv3	The achievement is accuracy and speed. YOLOv3 can detect and classify various weapons in real- time with high precision.	One primary limitation The system is not effective in detecting other types of weapons.
Internet of Things: Analysis of suspicious behavior in a surveillance camera network (2021)	Ala- Eddine Benrazek	Bron- karboch	The main achievement is its robustness in diverse environmental conditions. .	Bron-Karboch is its relatively high rate of false positives and slower processing speed and limited dataset.
YOLOv5s-Fog: Object Detection in Foggy Weather Scenarios (2023)	Xianglin Meng	YOLOv5	The main achievement is its accuracy and real-time processing capabilities.	YOLOv5 include potential inaccuracies in image blurring, color distortion, different angles and dependency on high quality training data.
Weapon Detection for Smart Surveillance System using YOLOV6 (2022)	Majeed Hussain	YOLOv6	Its streamlined architecture enhances detection capabilities.	To handle complex backgrounds and worked on only two labels.
Application for threat surveillance using python and yolo-v7 (2023)	R. Manuel, B. Ioana, and A. Gavrilas	YOLOv7	This paper improved performance in obscured objects.	Slower processing speed and did not work better on occlusions objects.
Weapon Detection in Surveillance Videos Using YOLOV8 (2023)	Raman Dugyala	YOLOv8	In this Paper, the method YOLO-v8 enhances real-time detection capabilities and accurately identifying weapon objects.	The main limitation of this model is its sensitivity to variations in lighting and background complexity, difficult to detect dummy object.

Having a single shot multi detector box (SSD)

Another step towards real-time object detection is the Single Shot Multi Box Detector which is similar to the generalized joint method to detect objects. The SSD technique predicts multiple bounding boxes at different scales within a network and simultaneously produces accurate results. This optimization facilitates fast processing speed. Through the use of these methods, public authority can be enhanced considerably.

Application in Safety Operations

To give a clear comparison of approaches, results, and contributions across different research, this thesis has given several studies in a concise tabular manner. This method summarizes pertinent trends and gaps in the area to support the context of my research while highlighting important parts of each study for simple analysis and reference as shown in the table 2.1.

Threat Detection

Computer vision technologies, which play a crucial role in threat identification in public spaces, are no longer of value but provide outstanding surveillance outcomes that are highly efficient in detecting threats. The automatically processed footage coming from the CCTV cameras helps the security officers detect potential threats like those of weapons, actions considered to be warnings, or objects that were abandoned. It is these tools, which allow the securities men to deploy a preventive approach in guarding against the emerging risks.

Crowd Monitoring

The processing of different crowd behaviors is a key activity for maintaining public safety is evident. The fine-tuned computer vision algorithms can monitor crowds spot deviations, and recognize possible dangers and warning signals when the crowds become large. By establishing this, officers can detect emergencies or security breaches and will intervene earlier.

Ethical Considerations

As computer vision technology is built into the fabric of the country's public safety, the importance of privacy and responsible use is now getting more attention. The equal caution of which is as much necessary for the operatively of these technologies as it is for the security of personal data.

Chapter 3 Research Methodology

An overview of the suggested procedures for this investigation is provided in this chapter. The proposed technique is described in detail and the dataset that will be used to measure performance is in Section s.1. The data analysis portion is covered in section 3.2. Before training the model, the phases of data preparation are covered in section 3.3, along with the methods or techniques used to gather balanced and qualitative data. The scenario for data retrieval, pre-processing, and training process for the main dataset is described in section 3.4. The efficiency evaluation used to measure the model's performance is described in section 3.5 and brief discussed the system architecture in section 3.6.

Data Collection

Data gathering is the first step in research that attempts to assemble, a population regardless of the type and size, for the reason of the deployment and evaluation of the YOLO-NAS-based Weapon Surveillance and Detection system. The approach employed herein takes into account the image-based and video-based sources from varied sources to ensure that the model is highly efficient when deployed on real time issues.

Dataset Selection and Configuration

This research has taken a dataset from Roboflow, which has configured the dataset based on classes like knives, guns, pistols, etc. Roboflow's annotation tool has been used to label the images manually. After this process, generate a dataset version by applying data augmentations like rotation or flipping to enhance model robustness. Set the train, validation, and test split ratios like 70% train, 20% validation, and 10% test, then download the dataset and integrate it into this research, updating the configuration to point to the correct data directories. Once set up, train the YOLO-NAS model by using the configured dataset and monitor the training progress, adjusting parameters or data as needed for optimal performance.

Moreover, the collection of data is structured in such a way that diverse sources are utilized to guarantee that the dataset is rich in quality. During the data collection processes, detecting knives, guns, pistols, rifles, and rockets was the task for which this research used a dataset from Roboflow. The automated annotation translation and augmentation features of Roboflow's platform made dataset preparation easier while expediting the data handling procedure. Moreover, many other datasets are available on the internet, academic publications, JSTOR, Kaggle, IEEE Explore, COCO (Common Objects in Context), and VOC (Visual Object Classes), and books provide a wealth of information's troves. Open datasets like COCO (Common Objects in Context) and VOC (Visual Object Classes) that utilize general object categories. The augmentation features of Roboflow, including brightness modifications, scaling, and horizontal flips, were employed to improve the dataset's diversity. The objective of these enhancements was to enhance the model's capacity to generalize across many contexts and circumstances.

Data Preprocessing and Augmentation

To meet the YOLO-NAS input size, all photos were reduced to 640x640 pixels. Additionally, Roboflow's augmentation tools were used in the augmentation process to apply random horizontal flips, scaling (0.8–1.2), and brightness modifications of ±10%. This assisted the model in learning how weapons appear differently from various viewpoints and in various lighting scenarios.

Model Training Configuration

This model used the YOLO-NAS Medium architecture for training since it pro- vided a nice balance between accuracy and real-time performance. A 640x640 image was used as the input image size. where it used a batch size of 32 to train the model, and for better convergence, the model started the learning rate at 0.001 using a cosine decay scheduler. The training process lasted 100 epochs, and Adam W, an optimizer with a momentum of 0.9, was employed. The bounding box regression and classification were all optimized for using the default YOLO-NAS loss function during training. Configuration includes

Model Selection: Use YOLO-NAS Medium for a balance of speed and accuracy. **Learning Rate:** The learning rate controls how fast the model updates its weights during training. A lower learning rate might result in better accuracy but can take more time to converge.

Batch Size: The batch size determines how many samples are processed before the model updates its weight. If the batch sizes are larger it will help in speeding up the training process but requiring more memory. This research has used 32 batch sizes for GPU.

Epochs: The number of epochs refers to how many times the model passes through the entire dataset during training. More epochs can improve accuracy, but too many might lead to overfitting.

Input image size: It refers to the resolution of input images, such as 640x640 or 416x416. Larger image sizes can capture more details but increase computational time, so this research make it a balancing size that can optimize between accuracy and speed.

Confidence Threshold: It is the minimum score for a detected object that can be considered valid, with typical values between 0.1 and 0.5. Lowering this threshold increases detections but can lead to more false positives in results.

Non Maximum Suppression (NMS): This threshold determines the IoU (Intersection over Union) threshold for removing overlapping bounding boxes. **Anchor Boxes:** These are predefined shapes and sizes that help predict object boundaries, which can be fine-tuned for objects like rifles, pistols, knives, etc. Their sizes are usually determined during configuration or by using clustering techniques.

Optimizer: The optimizer is used for updating model weights, with options such as Adam. It is commonly used because it adapts the learning rate dynamically. **IoU loss Function:** This function helps with bounding box regression and comes in different variants like GIoU, CIoU, or DIoU. In the detection task, if it chooses the right variant, it can lead to improved performance.

Data Augmentation : This includes the techniques, such as flipping, scaling, and rotation, to improve generalization, enabling YOLO-NAS to detect weapons under various conditions.

Number Of Classes : The number of classes defines how many object categories the model can detect, which affects the model's output structure. For this thesis, these classes have included knives, guns, pistols, rifles, and rockets.

Dropout : It is used as a regularization method to prevent overfitting. It randomly turns off units during training, helping the model generalize better.

Hardware and Setup

For the hardware and setup, if this model is utilized at a high level for a high performance computing environment to ensure efficient training, the following requirements must be made for working the system properly.

GPU: The training and evaluation processes were conducted on an NVIDIA RTX 3090 GPU. This GPU offers 24 GB of VRAM, which is crucial for handling large datasets and models like YOLO-NAS, particularly when using higher-resolution images.

CPU: An AMD Ryzen 9 5950X, a 16-core processor, was used for data preprocessing and other computational tasks that didn't require GPU acceleration.

RAM: The system had 64 GB of DDR4 RAM, allowing efficient handling of large datasets and avoiding memory bottlenecks during model training and validation. **Storage:** The system was equipped with a 2 TB NVMe SSD, ensuring fast data access and storage of model checkpoints, datasets, and logs.

Dataset Relevance

To start off, the relevant datasets and resources is deducted on the basis of the situational comprehensibility of public places weapon surveillance. The pictures and videos are sought from different surroundings such as schools, shopping centers, buses, and stations that represent the areas of prime attention and where powerful weapons detection is a must.

Inclusion Criteria

The inclusion criteria are created to make sure the dataset correctly applies to the research. To provide a visual representation of the preventable circumstances, images and videos need to be relevant to the public safety problems. The dataset has actual weapons like guns, knives, and explosions as instances and is the complete set of challenges to be answered by the proposed YOLO-NAS model.

Ethical Considerations

Ethical factors are something you need to take into account seriously when data is being collected. Privacy obstacles are at the forefront; thus ensuring that the dataset cannot jeopardize the privacy of individuals who are in public places is of utmost importance. Only information that doesn't represent any personal data is disclosed or anonymized zed on ethical Data protection standards.

Data Quality Assurance

Accurateness of the obtained information stands out as the most significant issue. Thorough controls are put in place to ensure the confirmation, truthfulness, and pertinence of the ones towards each picture or video.

Diversity and Representation

The diversity of the dataset is the main aspect of building a robust and well- trained model with the capability to give accurate results in that specific domain. Images and videos are chosen to represent different scenarios; subject to varied lighting/weather conditions and possible scenario complexities encountered in real gun control missions.

Data Documentation

The metadata of every data being collected is organized in an organized manner, including information regarding the context, source, and any additional details. Such documentation ensures openness, and re-creatability, and is a part of the dataset commissioning process for the community of researchers.

Continuous Monitoring

The process of data gathering is reflexive, and cycleicios, with imminent monitoring and revisions to generate a dataset that is reflective of the present situation. As the new challenges are shaped in the field of public safety and weapon surveillance, other datasets may be integrated to enhance the algorithm's adoption.

Data Analysis

Interpreting and analyzing data make up the vital stage in the research procedure, where collected data gets processed. First, analysis of the information is the main task, namely, data mining and pattern recognition, as well as their further technical use to create the YOLO-NAS-powered Weapon Surveillance and Detection System.

Data processing leads to interpretation

The analysis ought to begin with a fundamental survey and deciphering of the collected data. Among the tasks will be images, video and other component information examination aiming to identify the dataset properties and nature. The purpose here is to learn about the factions/scenarios/weapons that are featured in all these photos.

Pattern Identification

Data Analysis, per se, is meant for the identification of the patterns and trends that are related to the weapon monitoring and detection systems. This refers to not only the objective of looking at how weapons along with other objects have been portrayed in the dataset, but also includes a level of evaluating the credence, authenticity, objectivity, and reliability of the data sets. The interpretation of these patterns better enables the adjustment of model structure and training to suit the competition with the ones that are applied in real challenges.

Technology as an Analytical Tool

For the convenience of productive data analysis, tools that are in use are proper for the job. Software toolkits, data visualization facilities, and custom scripts may be employed to identify unforeseen patterns, correlations, and glitches in the data. Visualizations play a big role in the presentation of complicated information in an easygoing way.

Cost Considerations

Besides the costs of web access, academic publications, and data analysis instru- ments the gaining of information needs costs. Part of the budget is devoted to the operation of the research maintaining the investigation at the highest level and thorough.

Analytical Process is Dope-Mixing

Data analytics is an ongoing process of repetition, where the collected data is con- stantly being refined and can be pivoted if need be. At these stages, the researcher will notice that some patterns and insights will be discovered as observations con- tinue. This will promote necessary changes to be made to the research approach, dataset, or model design. This iterative agile method allows my study to be ap- plicable and flexible enough to solidify it for the unusual and complex operational environment surrounding IP monitoring.

Environmental impact modeling

The results from data analysis not only direct but also guide each stage of the research, thus, designing the YOLO-NAS network model becomes possible. The model builds understanding about the traits of the objects, the conditions in their environment, and the challenges they might encounter which make them competent and successful in all the different conditions.

Dataset Preparation

Dataset preparation is the most important step in the approach of applying the YOLO-NAS based system to Weapon Surveillance and Detection because it de- termines the result later. This phase involves rigorous organization, annotation, and setting of the right data structure that holds all annotated data, in a manner that the real-world problem scenarios are well-represented.

Choosing Images and Videos

The first phase of data set preprocessing is precisely the image and video col- lection from the collected data sets. Generally, this exchange is influenced by how available a particular information is to the advertisement goals. Pictures and videos will emphasize the coverage of various environments, occasions, and problem types.

Additions are presented for an Atmospheric Participation in Weapon Instances

The dataset considered follows fire arms, knives, and explosives. In this case, the YOLO-NAS model becomes aware of different weapon types, which enables it to conduct valid inferences for a different set of threat situations. Every image or video defined for the dataset shall accurately depict the weapons with details.

Annotation Process

Annotation of the dataset is a task that must be performed for preparing the dataset and includes the addition of bounding box annotations and class labels that are used to pinpoint the presence and location of weapons in each image or video frame. The text files from formats like PASCAL VOC XML, COCO JSON, and even YOLO TXT can be used, depending on the needs of the YOLO-NAS network.

Annotation Tools

Mapping out the annotation procedure, annotation utilities will be implemented. These instruments allow the objects to be cut to sharp boundaries with the classes indicated. It is pivotal that steps are taken to ensure annotations are to standard, and the background shall be corrected for the model to be able to effectively detect and localize weapons by recognition.

Dataset Organization

Data structure is an equally vital factor, especially when it comes to training algorithms. Images are systematically arranged and their annotations files, so

that each file represents an image and the entire arrangement stays coherent and ordered. Coordinating the collection of data from various sources in a structured manner ensures the smooth incorporation of the datasets into this training process.

Scenarios Representation for Different Occasions

Data structure is an equally vital factor, especially when it comes to training algorithms. Images are systematically arranged and their annotations files, so that each file represents an image and the entire arrangement stays coherent and ordered. Coordinating the collection of data from various sources in a structured manner ensures the smooth incorporation of the datasets into this training process.

Anonymization and Privacy

To comply with ethical principles, we ensure privacy and anonymity by properly concealing personal data or delinking it from the dataset. Privacy concerns are the key issue, with the assurance that the dataset protects individual privacy in the public space as the number one priority.

Documentation of Dataset

The data collection includes manual tagging, transcription as well as any additional metadata associated with the annotations. By doing so, this documentation is a necessity for good faith, repeatable research, and hence, the sharing of the dataset with the scientific community.

Quality Assurance

Proper quality assurance was observed to determine the authenticity (and validity) of the dataset. Severe monitoring will be conducted to find out and remove the rhythm, inaccuracies, or faults less any inconsistencies. The features of a precise dataset are the main features of the training of a fault resistant and efficient YOLO-NAS model.

Training Process

The training procedure is the key part of the algorithm in which the Weapon Surveillance and Detection System based on the YOLO-NAS is being built. The various stages in this part include the preparation of the dataset, loading the pre-trained model weights, and iterative fine-tuning of the model to achieve high accuracy in weapon detection even in different scenarios as shown in figure 3.1.

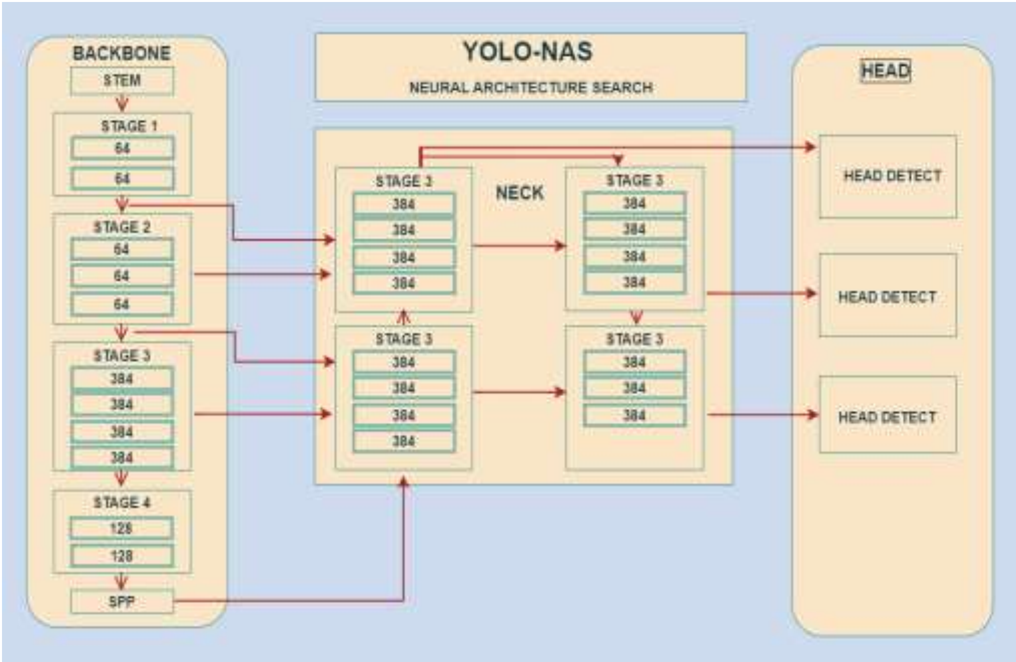

Figure 3.1: Working of YOLO-NAS

Data Pre-processing

Before the training process, the data pre-processing will be done to fit the dataset in line with the requirements of the YOLO-NAS model for training. This refers to the procedures that are being utilized, to perform actions, like ensuring the images have the right size and normalizing pixel values, in addition to creating annotations in a uniform and orderly process. The aim is to generate a uniform and logical input provided to the model.

Weights Pre-saving with Pre-trained Weights

YOLO-NAS is launched with pre-trained weights that are commonly collected from colossal image resources such as ImageNet. Pre-training the model with a diverse dataset of pictures brings all the needed knowledge to the model that will thereby bring assistance to the generalization of those weapon related features pertinent to detection.

Training Configuration

These training configuration attributes consist of specific metrics like learning rate, batch size, optimization algorithms, etcetera. However, the process of training a neural network is done by using minibatch training and stochastic gradient descent to update the model's weights iteratively as well. Although fine-tuning of the YOLO-NAS network is applied for the sake of adapting it to such features as the characteristics of weapon detection.

Model Iteration

Training can be carried out with the activation of any number of epochs where the model learns the recognition and localization of weapons in the dataset. In this process a test is triggered on the model's performance and decisions for tweaking weights are made based upon improved accuracy as well as efficiency.

Fine-tuning and Adaptability

Fine-tuning the YOLO-NAS model proves its capability of adaption here; we discuss the way of adaptation of the model to reflect unique issues of weapon detection in places when people are a critical obstruction, as well as variations in the size of the target objects.

The effectiveness of video cameras

One of the best means through which the surveillance system can provide security to both shoppers and shop staff. Our automated YOLO-NAS model is implemented through a set of surveillance cameras networked together at consistent public sites. This integration allows us to see the real-time monitoring and detection of which weapon will appear in different types of video and then identify it. The model involving facial recognition and surveillance systems is of great significance which further makes it applicable within live circumstances.

Scalability

We devised the YOLO-NAS model to be able to operate at a large scale parallelly with many cameras available. This scalability notably means that the model can be successfully launched and utilized in every way: from small scale settings to immense spaces, and this helps make the model more practically useful.

Evaluation Metrics

The evaluation procedure is dependent on critical metrics such as precision, recall, and how much a class is above an average image (MAP). These metrics presented in the table show the statistics in which the model demonstrates the accuracy, sensitivity to weapon instances, and general weapons-detection precision.

Optimization Techniques

Correction techniques such as regularization or dropout can be applied during training to prevent overfitting and to increase the generalization capacity of the proposed model, YOLO-NAS. The tools mentioned above help the model bring more resistance to variable situations.

Pseudo Code

Initialize: Load necessary libraries (YOLO-NAS, OpenCV, etc.)
Set up YOLO-NAS model for object detection
Define weapon classes (e.g 'gun', 'knife')
Load YOLO-NAS Model: Load pre-trained YOLO-NAS model weights
Capture Input Data:
Capture video input (e.g from webcam or CCTV)
While Video stream is active **do**
Capture frame from video stream
Resize frame to required input size
Normalize the frame
Perform Object Detection: Pass preprocessed frame to YOLO-NAS model
Get the list of detected objects (bounding boxes, confidence, class)
for each detected object in the list **do**
if object class is 'weapon' **AND** confidence score > threshold
Draw a bounding box around detected weapon
Label the bounding box with the weapon class (e.g'gun')
Store detection information (class, location, confidence score)
end if
if any weapon is detected **then**
Trigger alert (sound alarm, display warning, etc.)
Save detection frame with bounding boxes
end if
Display Results: Display frame with detection results in real-time
end while
Terminate System: Release video stream resources, close system

Rating and analysis of performance

The evaluation and hardware and software performance analysis are essential in the stage where the system's output is being analyzed to comment on the performance, reliability, and effectiveness of the YOLO-NAS-based Weapon Surveillance and Detection System. Herein, this part identifies the parameters, methods, and other aspects used in assessing the effectiveness of the project. The evaluation criteria for the model's performance are precision, recall, and mean average precision (MAP). These metrics deliver the numeric calculation of the detector's accuracy and their recognition of all weapons.

Precision

Precision measures the model's exactness in terms of the positive cases that it predicts. It is defined by the formula of (true positive predictions)/(true false positive predictions). A higher precision level means that overall there will be fewer false positive cases and the prediction of the model will be more accurate with a higher confidence rate.

Recall
The memory assessment tests the system's deep understanding of the weapons in general represented by all instances in the dataset. The formula says true
positive points are equal to true positive prediction divided by the addition of true positive and false negative. The higher the mean average precision for a given class indicates that the model generates more or less fitting instances of a weapon.

MAP (Mean Average Precision)
Mean Average Precision is a holistic measurement that takes precision for different stages of recall into account. It works through the computational (average precision at various levels of recall) method. This higher MAP indicates that the model is a good fit for all scenarios.

Comparative Analysis
On YOLOs performance, the existing methods and models that are utilized for weapon surveillance and detection are also compared. This comparison of the proposed approach with other benchmark models helps in identifying the strengths and weaknesses of the YOLO-NAS approach and then points out required areas so that the approach can still be improved.

System Robustness
The resilience of the system is assessed by evaluating its performance under harsh conditions including low visibility, night-time, and days with bad weather as well as highly populated conditions. An important thing that has to be appropriately analyzed to use the model in practice is how well it works under different conditions.

Real-time Performance
The real-time performance of the YOLO-NAS network will be especially valuable in the situation of apprehending crime. Performance analysis consists of checking whether the model is capable of fast and comprehensible online processing of the surveillance camera streaming and accuracy in alerting the law enforcement authorities on time.

Continuous Improvement
Assessment is also not a single-time situation. Assessment is a continuing process. The continuous monitoring of the system's performance always gives room for the improvement of all the previously established targets. Regular updates to the model based on feedback and new data contribute to its long-term effectiveness.

Areas for Development
The assessment finds out which characteristics of our strategy need to be developed or clarified. The feature presentation of hardly improving the precision, facing a difficulty, or enlarging the potential of what the YOLO-NAS model can do is the turnkey toward developing the future iterations of the system.

Visualization of Results
The results from the evaluation of gain and performance analysis are represented visually by graphs, charts, and other visualization methods. Visualization provides a way for the viewer to take in even the most complicated and incomprehensible information while still making it clear for interpretation and decision-making.

Iterative Refinement
The outcomes from the examination phase feed into the model YOLO-NAS tuning process continuously. Accomplishments can be achieved by modifying parameters, noting training data, or applying algorithmic improvements to resolve distinctive issues.

System Architecture
The architecture of the system includes PTQ (Post-Training Quantization) and QAT (Quantization-Aware Training) processes. PTQ is a technique applied after neural networks have been completely trained. This process comprises converting the weights and activations of the trained model from higher precision i-e (32-bit floating point, FP32) to lower precision (e.g. 8-bit integers, INT8). This reduces model size, and inference time and makes the model faster.
Moreover, QAT is a more sophisticated approach that introduces quantization during training. The

model is trained with quantization, simulating the effects of reduced precision like 8-bit integers during the training phase. This process has been proven less accuracy loss as compared to PTQ as shown in figure 3.2. All these make the system architecture of a YOLO-NAS-based Weapon Surveillance and Detection System unique, is the special side of the system that shows the parts of the system, the interaction between components, and their functionality. The design of the architecture involves the alignment of the cameras to perform an analysis of the data, raising an alarm, and making sure that it is scalable and easy to adapt.

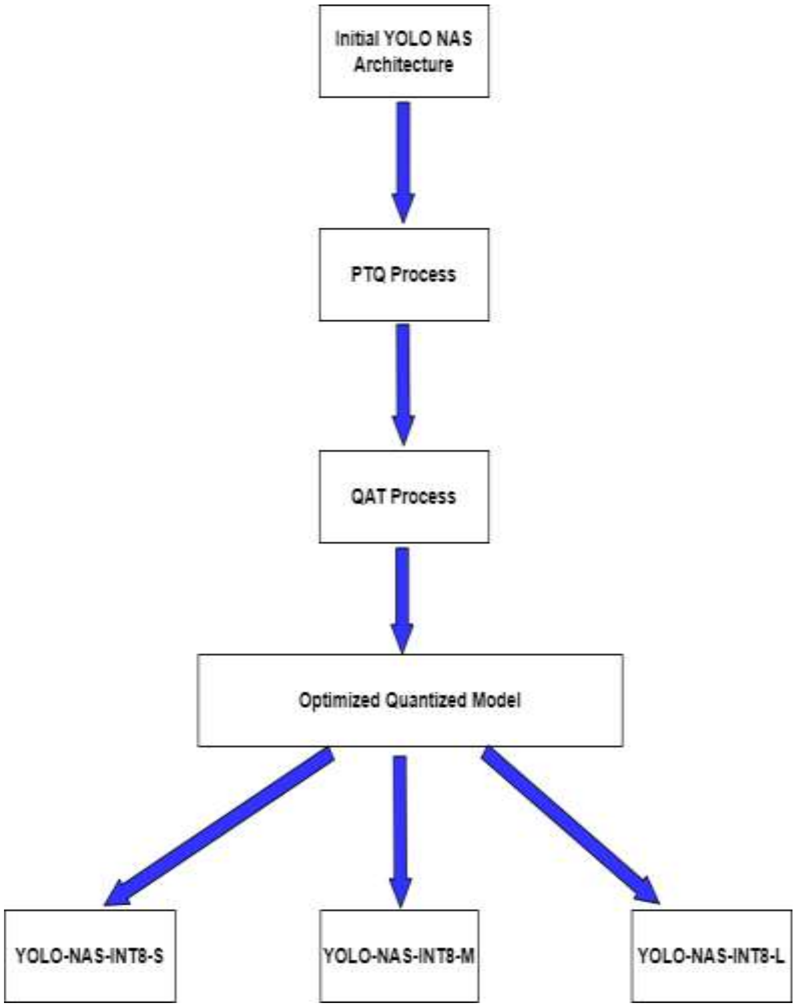

Figure 3.2: YOLO-NAS Basic Architecture

Weapon surveillance system utilizing YOLO-NAS

The heart of the system is YOLO-NAS-based module for weapons observation reaction system that 24 hours a day recognizes and identifies the weapons. This module can be built based on the YOLO-NAS architecture, as it can combine a high detection accuracy with a run time which is efficient in identifying multiple weapons types such as firearms, knives, and explosives. YOLO-NAS is a model that has been specially trained in the process of identifying and locating weapons in images or video frames.

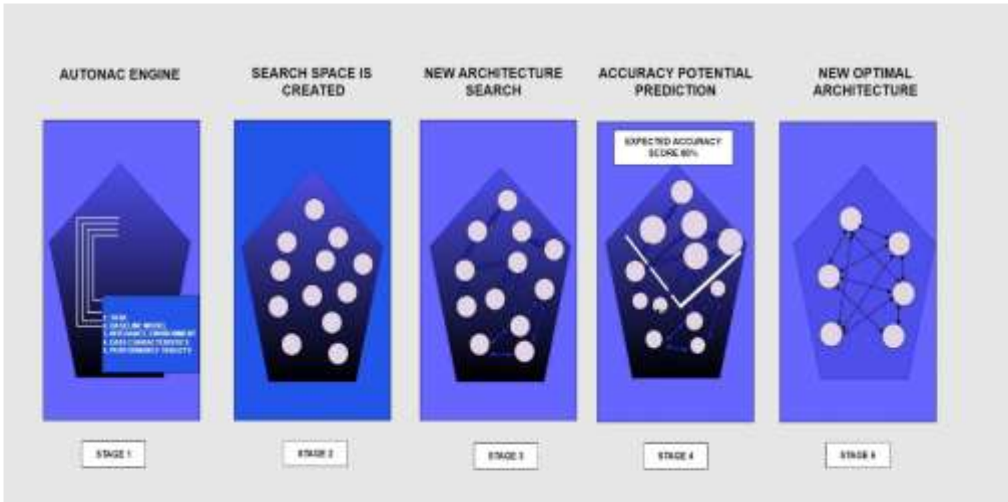

Figure 3.3: YOLO-NAS architecture stages

The above figure 3.3, particularly focuses on neural architecture search (NAS).

There are main five stages in this architecture:

Stage 1: AutoNAC Engine It defines the problem and setting up the AutoML engine by defining the specific task, making a baseline model for comparison, understanding the target surmise environment, characterizing the data, and setting performance targets.

Stage 2: Creating the Search Space This stage explores architectural options and generates a vast search space of all possible neural network architectures; encircle various combinations of layers, activation functions, and other hyper parameters.

Stage 3: New Architecture Search It defines efficient exploration and systematically searches through the vast search space, leveraging techniques. For instance, reinforcement learning, and gradient-based optimization to efficiently discover promising architectures. Moreover, this involves building and evaluating architectures to identify the most best ones.

Stage 4: Accuracy Potential Prediction This stage understanding the search space and previous evaluations and predicts the expected accuracy of the potential architectures. This helps to avoid wasting time on architectures that are unlikely to perform well.

Stage 5: New Optimal Architecture This stage shows the new optimal architecture that was found through the process of searching the architecture space. It shows the neural network with a 3x latency acceleration, 3x throughput improvement which increases accuracy. This new optimal architecture provides better performance and efficiency compared to the previous ones.

Moreover, pooling layers are a main component of convolutional neural networks (CNNs). Although they provide several benefits like reducing spatial dimensions, improving robustness, and reducing the number of parameters. Despite all these, they can also cause information loss. These layers reduce the spatial dimensions of the feature maps while processing. This process discards some of the spatial information and lead to loss of text. For instance, by tuning hyperparameters like filtering and stride, an image has a 4 x 4 matrix with a stride 2. It will extract only important features and show a new matrix 2 x 2. This causes an information loss which is needed in future processes to improve precision and accuracy. To curb this issue, skip connections have been used in the layers.

Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 11, 2025

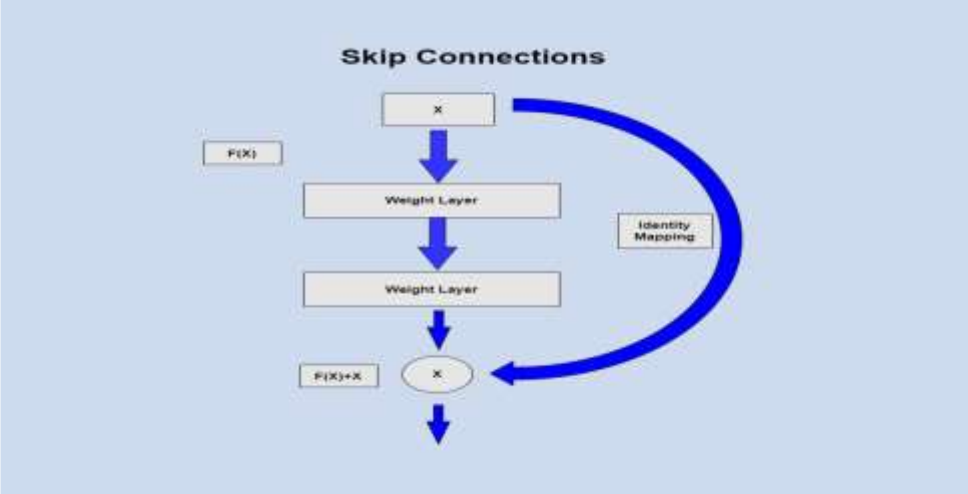

Figure 3.4: Skip connections to avoid information loss

This above figure 3.4 illustrates a process known as skip connections in deep learning. All components are in detailed appended below:
Input: The process begins with an input denoted as ”x”.
Weight Layers: These layers represent the core of a neural network. They are responsible for applying transformations to the input data using learned weights. In this case, there are two weight layers, suggesting a multi-layered neural network. **F(x):** This represents the output after the input ”x” has gone through the first two weight layers. It symbolizes the complex transformations performed by these layers.
Identity Mapping: This is a crucial part of ResNet. It simply passes the original input ”x” directly through without any modification. This preserves the original information.
Addition (+): The output of the weight layers (F(x)) is added to the original input ”x” (which went through the identity mapping). This additional operation is a key aspect of ResNet.
Output: The final output is F(x)+x.

In addition, skip connections are an essential component in ResNet (Residual Network) architecture that helps to avoid information loss. In ResNet, skip connections are used to add the input of a layer to its output, allowing the network to learn only residual functions instead of learning the complete function. This process is done by introducing an identity mapping between the input and output layers that allows the network to learn further complex features.

Surveillance Camera Assimilation

It is surveillance cameras that act as detectors situated in places open to the public that constitute the input system of the project. The cameras permanently shoot videos from the monitoring zones they are stationed in as shown in figure 3.5. The integration is a process of YOLO-NAS establishing a link to the Weapon Surveillance System for real-time feed analysis, which serves as input to the decision-making.

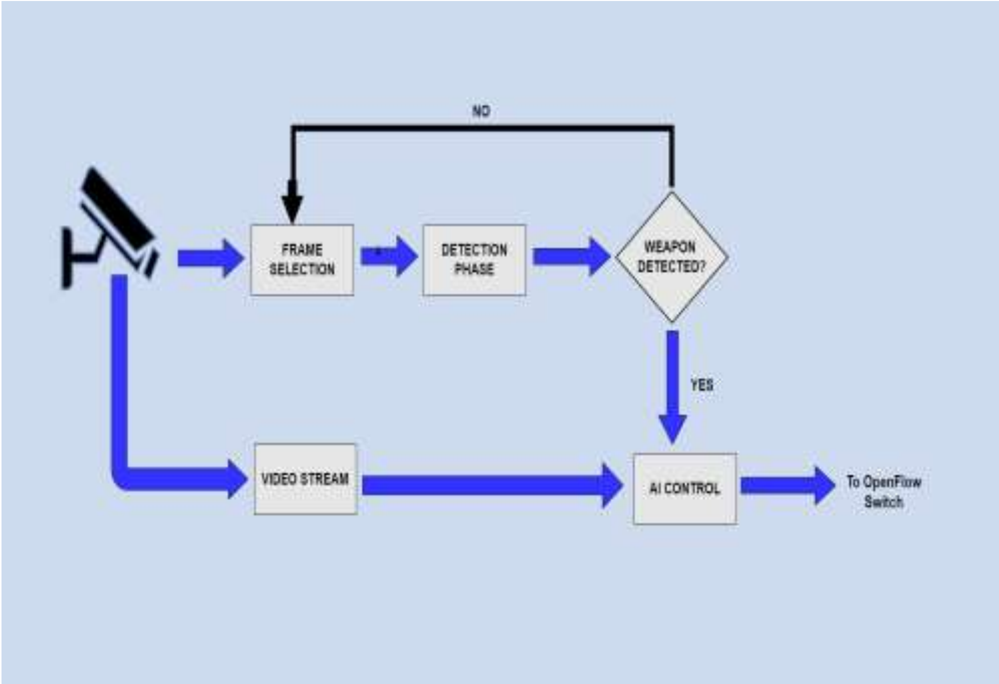

Figure 3.5: Integration of Surveillance Cameras

The apparatus of Alarm System and Communication with Authorities

The moment the system triggered when the weapon showed up in the footage. Concurrently, the system notifies the concerned authority to manage the security matters, otherwise, law enforcement. The monitoring system comes to the rescue by being the immediate alert mechanism and detecting suspected danger, acting swiftly to secure the area to avert possible threats.

Main Features of the System

A high degree of scalability and adaptability is defined by the underlying architectural design, thus, the system can support a wide spectrum of environments and scenarios. The main key features are as follows:

Scalability

The system is perfectly fitting for realizing the goals of surveillance programs in large-sized public spaces or any other securitized points that require additional coverage.

Adaptability

The adaptability of this YOLO-NAS model is astounding. Through this fine- tuning, the model performs very well in complicated environments comprising
different scenarios. The software can periodically change and be adjusted to handle various weapons, different threat conditions, and emergent threats.

Flexibility

The pick-and-deploy-as-you-need feature of the system makes it fully adaptable to situations such as the addition of sensors, introduction of technologies, or updates. This feature guarantees a progressing system, where innovations in the surveillance technology of combat weapons can be accommodated.

Data Capturing and Statistics

The construction of the architecture comprises a data logging and analytics section for a thorough evaluation to be conducted after the event and for the system optimization process. At the center of all this, a data store will keep track of critical parameters such as detected threats, system errors, and atmospheric conditions. With the help of analytical

ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 11, 2025

tools, we can get here deal with information, pattern identification, and the system enhancement guidelines setting.

Privacy Safeguards

Along with ethical aspects, the given building blocks would respectively have components of privacy protections. Identity protection of the individuals can be achieved by using anonymization methods that come with capturing the surveillance footage. The system adheres to the ethical code of privacy regulations and standards and contributes to the attainment of the right balance of safety and privacy.

The Monitoring User Interface

Consists of panels that capture sensor readings, control systems, and power distribution sources, enabling monitoring and adjustment from the monitoring stations. A user interface is a given for being certain of the monitoring and management of the system. Security officers and administrators can view important indicators from a dashboard to observe the system's status, review alerts, and do system configuration.

The software interface (UI) simplifies the monitoring and supervising of the intelligence gathering system provided by the digital surveillance tools. In conclusion, the system design incorporates a YOLO-NAS some CCTV cameras, and real-time analysis technology to create a capable Weapon Surveillance and Detection System as shown in figure 3.6. The architecture aims to improve public safety in different surrounding areas by aiming at specific attributes such as precision, efficiency, and scalability.

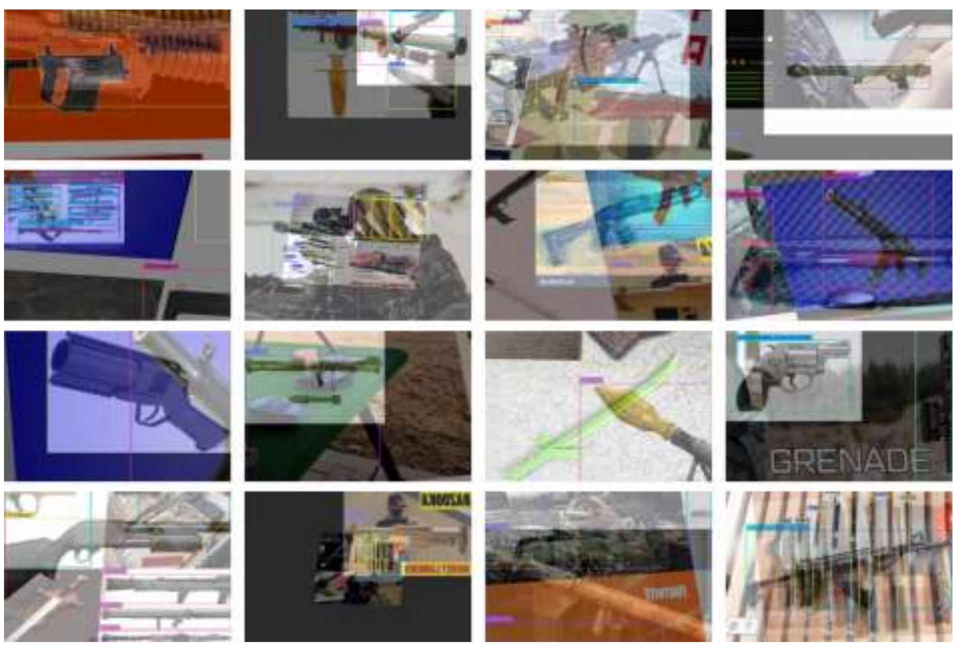

Figure 3.6: Monitoring user interface of the system

Summary

This chapter explains how the dataset is collected and how it is prepared to perform the forecasting on this type of dataset. All the preprocessing on the dataset is also discussed in the methodology along with the model architecture. YOLO-NAS working with PTQ (Pre-trained weights) and complete interface has been discussed. The problems that lead to information loss and their solutions have also been addressed. Furthermore, skip connections are used to extract all information and pass it from the first layer to the last layer which evaluates the model performance.

Chapter 4 Results

Model Design and Data Tests

Data gathering and pretreatments

This dataset is collected from Roboflow. Moreover, this dataset was produced using a variety of resources, including CCTV footage. To examine the model comprehensively, we gather pictures of weapons with varying colors, backgrounds, sizes, and shapes. The pre-processing that was done on data incorporated resizing of images, normalizing pixel values, and annotating weapons in the images. Hence, the pre-processing steps as part of model development were considered to be very important features for the scenario applicable to the real world.

To experiment, we first divide the weaponry dataset into training and test sets as shown in Table 4.1. The training set consists of 1409 instances of photos, of which 420 are classified as pistols, 340 as guns, 245 as rifles, 215 as knives, and 189 as rockets. The test set consists of 280 pictures of weapon instances, of which 84 are from the pistol class, 69 are from the gun class, 48 are from the rifle class, 43 are from the knife and 36 are from the rocket class.

Dataset	Total Images	Pistols	Guns	Rifles	Knives	Rockets
Training	1409	420	340	245	215	189
Testing	280	84	69	48	43	36

Table 4.1: Statistics About The Weapons Dataset

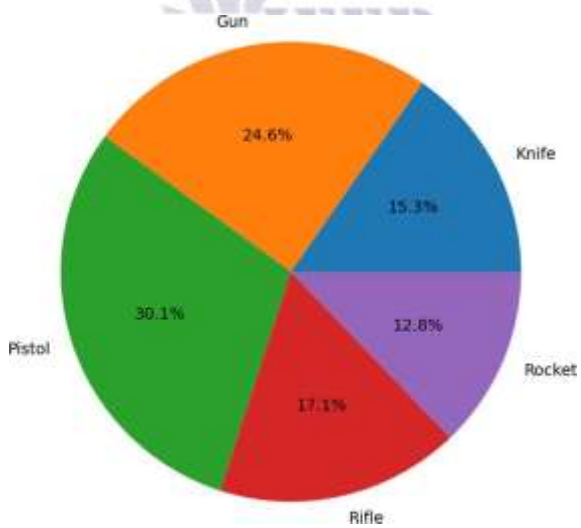

Figure 4.1: All dataset pictures distribution

The dataset used for this research shown in the following figure 4.1 has many entries with missing and corrupted images around 15% which were removed, and the resulting dataset is shown in figure 4.2 and 4.3.

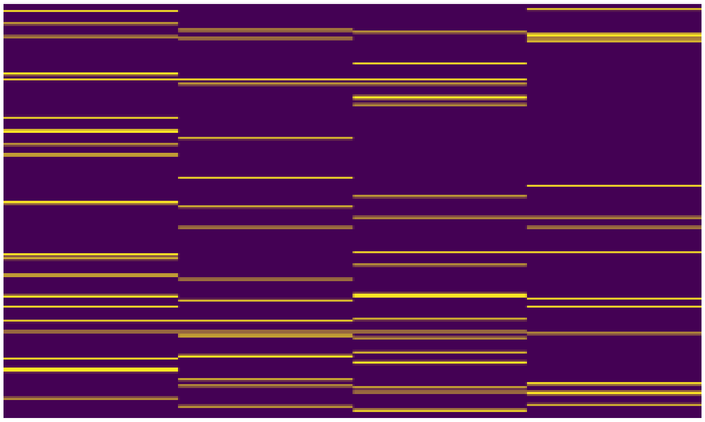

Figure 4.2: All dataset pictures

Figure 4.3: After removing corrupted pictures

Prediction on testing data

With this purpose, the YOLO-NAS model was successfully initialized using pre-trained weights that were obtained from a larger dataset similar to Image Net. This initialization has taken the central position in the model which transforms the gained knowledge from diverse images into general features that are necessary for the aim of non-firearm detection. To analyze all these things we used many algorithms.

A confusion matrix gives information on the correctness of the model by comparing the expected and actual classifications. An analysis was done on a confusion matrix class-wise as shown in figure 4.4 including the values for True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN). This matrix yields metrics such as recall, accuracy, precision, and F1 score. The percentage of accurate predictions—both true positives and true negatives—out of all forecasts is used to determine the model's accuracy, and the result is an accuracy of roughly 83. The percentage of true positive forecasts among all positive predictions, known as precision, is 90. Out of all actual positive situations, 210 of the predictions were true positive. 90 is the F1 score which is the harmonic mean of recall and precision. Moreover, recall is 0.89. These metrics demonstrate the model's overall performance and dependability in classification tasks by showing how well it can distinguish between the positive and negative classes and how resilient it is.

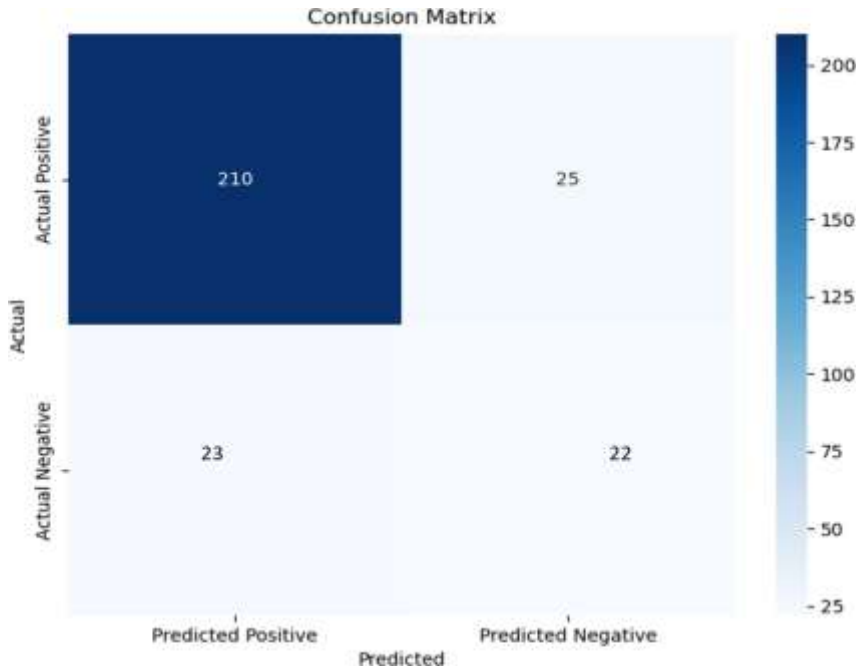

Figure 4.4: Confusion matrix of all classes

In order to analyze class-wise, figure 4.5 shows the actual and predicted counts of different types of weapons. The actual counts are represented by blue bars, while the predicted counts are represented by red bars. It seems the model does a relatively good job predicting the counts for most weapons, although there are some discrepancies. For instance, in this model, the total number of pistols is 78 but it detected 74. Overall, the model's predictions are very close to the

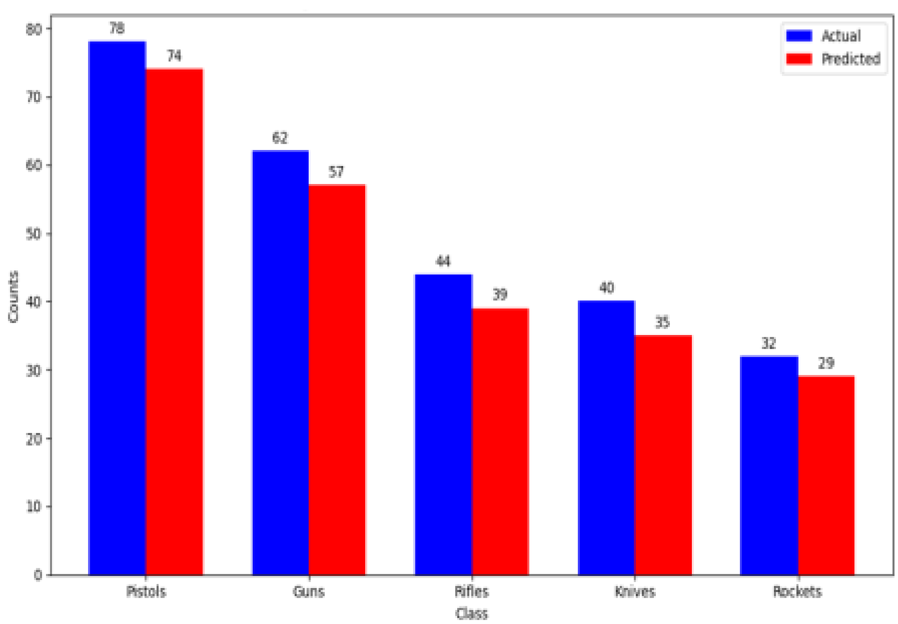

Figure 4.5: Histogram class wise actual and predicted data

Evaluate Performance

Best performers

The best model weights were saved after a model was trained on the training dataset. The below graph 4.6 has clearly illustrates that YOLO-NAS has the best precision-recall curve, followed by YOLOV8, YOLOV7, and YOLOV6. This suggests that YOLO-NAS is the best performing model of the four, followed by YOLOV8, YOLOV7, and YOLOV6. Moreover, the model is capable of detecting objects with high accuracy, even when the threshold for detection is set low. This is a critical aspect of object detection, as it enables the model to detect objects in a wide range of scenarios and environments.

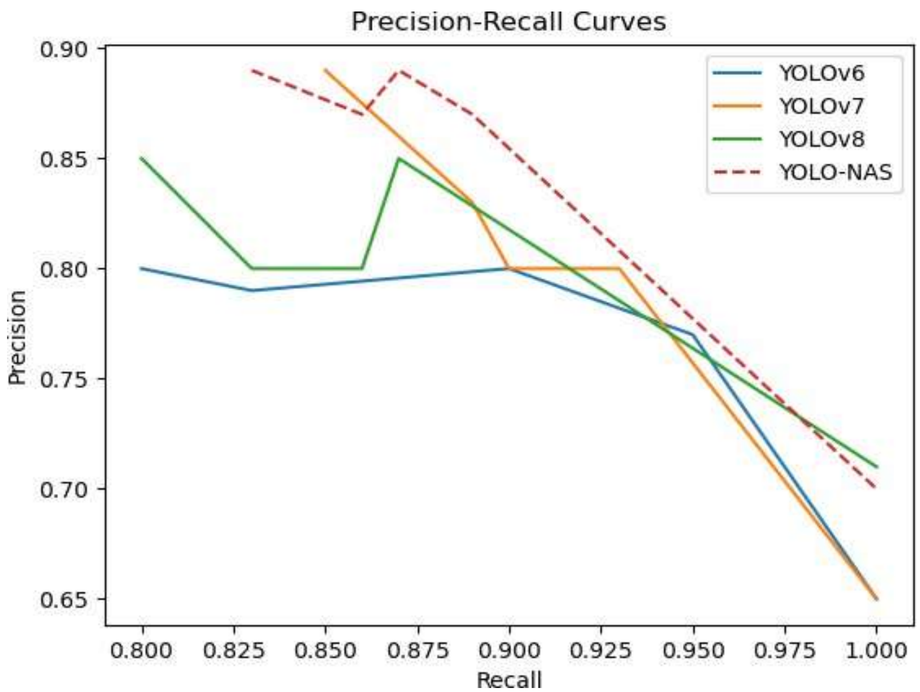

Figure 4.6: Best performance after normalization the dataset

Worst Performers

The figure 4.7 shows the precision-recall curves for four different object detection models: YOLOv6, YOLOv7, YOLOv8, and YOLO-NAS. The precision-recall curve is a plot of the precision of the model against the recall of the model. Precision is the proportion of true positives among all positive predictions, while recall is the proportion of true positives in all actual positives.

The curves show that all four models perform relatively well, with precision and recall both around 0.80. However, YOLOv6, YOLOv7 and YOLOv8 appear to perform slightly better than YOLO-NAS. This is because their precision-recall curves are higher than the other two models due to weak hyperparameter tuning. Poor hyperparameter tuning can lead to suboptimal performance. Moreover, another reason is limited training. This is because Insufficient or biased training data can lead to poor performance on unseen data which makes the model prediction unreliable.

Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 3, Issue 11, 2025

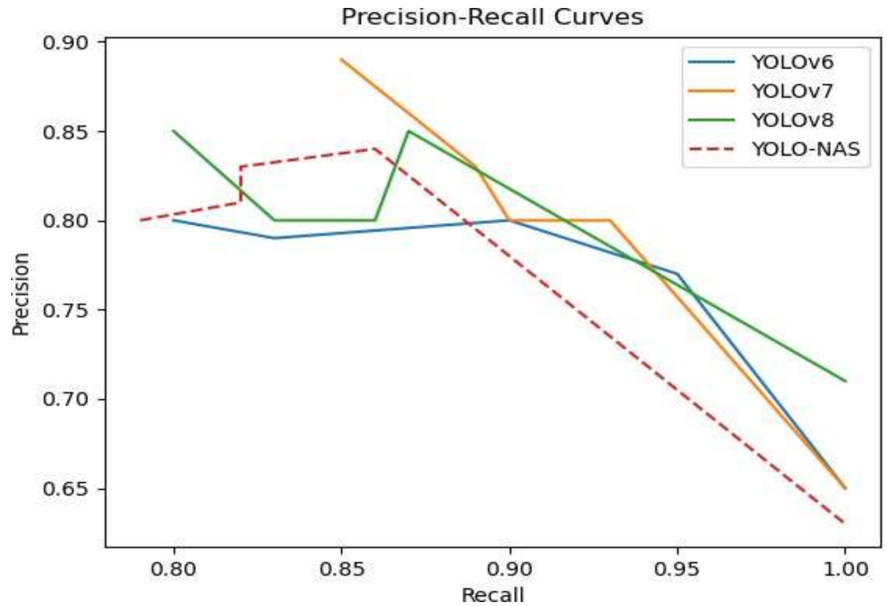

Figure 4.7: Worst performance before normalization the dataset

Interpretable Results

A Receiver Operating Characteristic (ROC) curve as shown in the figure 4.8, is commonly used to assess the effectiveness of a binary classification model. The dashed line in the picture represents the performance of a random classifier, which would have an AUC of 0.5. The orange curve illustrates the trade-off between the true positive rate and the false positive rate as the model's decision threshold changes. A higher curve indicates better performance.

In a nutshell, the ROC curve illustrates that the detection of weapons by the YOLO NAS model is satisfactory With an accuracy of 0.82 and can fairly accurately distinguish between weapons and non-weapons.

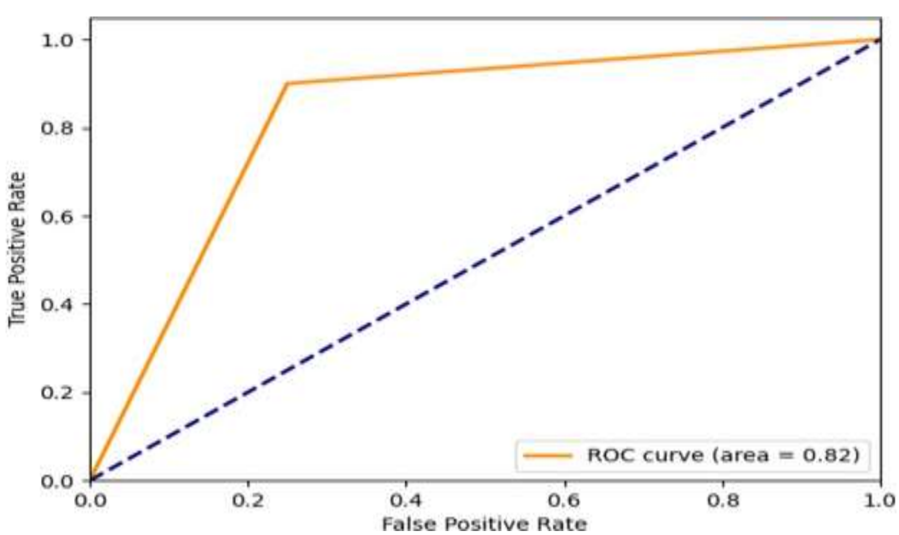

Figure 4.8: ROC Curve of the model

The below graph 4.9 shows the F1 score for weapon detection using different versions of YOLO family models, which are YOLOv6, YOLOv7, YOLOv8, and YOLO-NAS.

The X-axis represents different confidence thresholds used for classifying objects as weapons. The Y-axis represents the corresponding F1 scores for each threshold. If there is a higher threshold, it means the model is more confident in its predictions, leading to lower false positives but potentially missing some actual weapons. The F1 score measures the accuracy of the model, considering both precision where it checks how many detected weapons are actually weapons, and recall where it checks how many actual weapons are detected. As a result, YOLO-NAS emerged as the best performer in terms of F1 score, followed by YOLOv8.

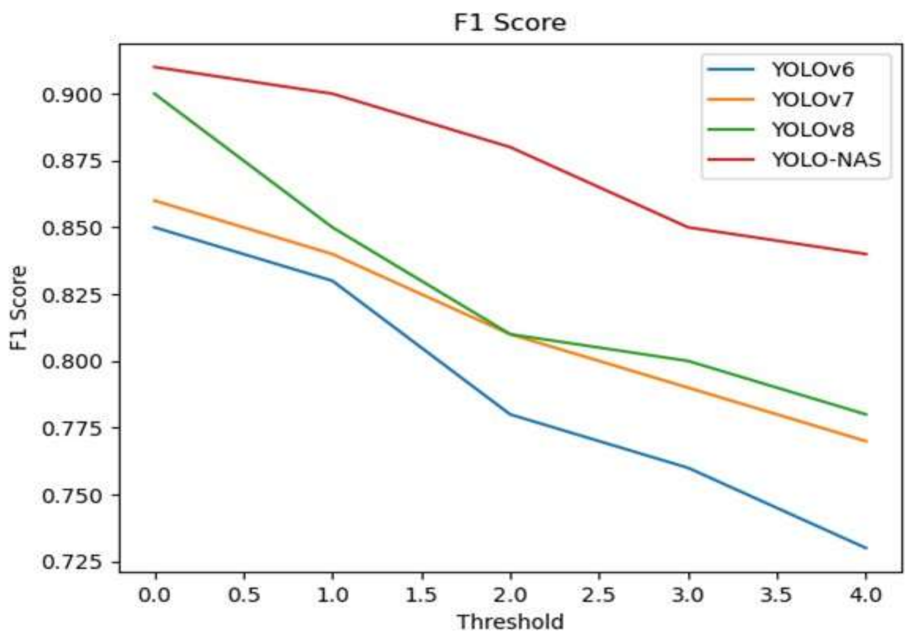

Figure 4.9: F1 score of the model

Moreover, the radar chart as shown below in figure 4.10, also illustrates the result to understand predicted data. This radar chart compares the actual labels with the predicted labels for each weapon category like guns, pistols, rifles, rockets, and knives. The more out from the center point means the F1 score is higher which shows the model is more accurate. These both cobwebs should be close to each other. In this case, the model has a high F1 score for pistols, rifles, and guns. The model has a lower F1 score for knives and rockets. The model needs to work more on this scenario because it has a lower F1 score for the rockets and knives, but it is still very accurate for the other classes as mentioned above which is quite well result.

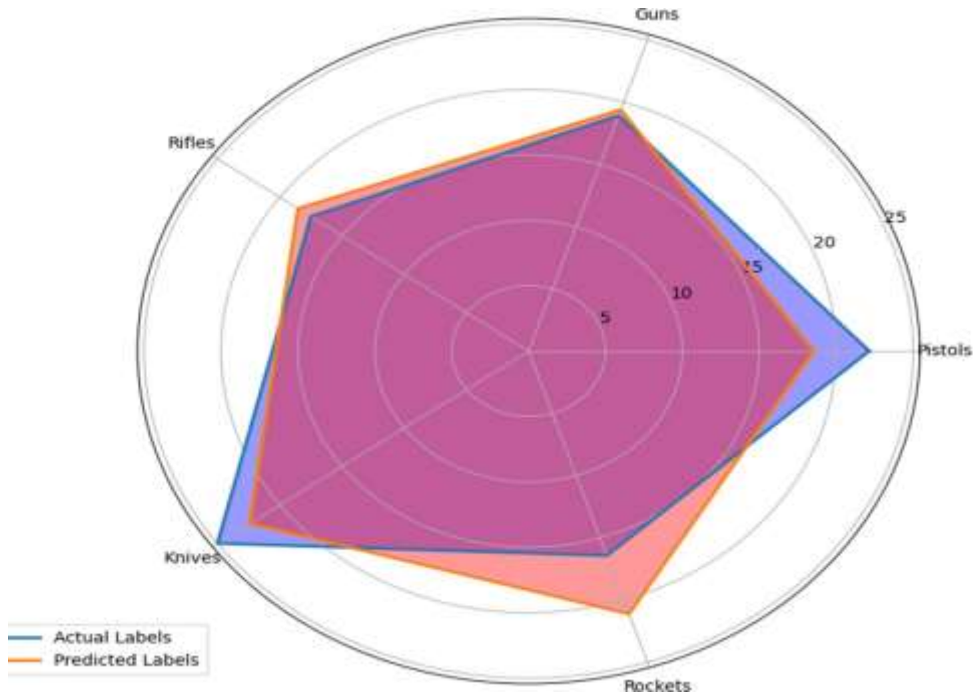

Figure 4.10: Radar plot of the model

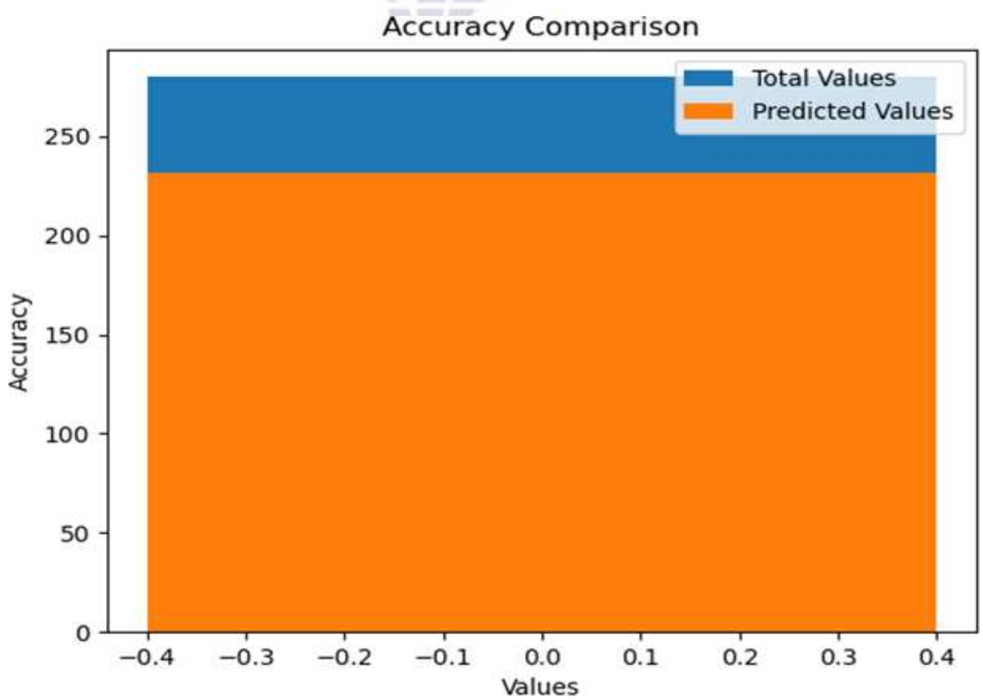

Figure 4.11: Accuracy score of the model

The graph 4.11 shows an accuracy chart comparing the total number of weapons in a dataset which is 280 as total to the number of weapons detected by a YOLO NAS model which is 232. This chart clearly visually represents the accuracy of the weapon detection model where the model detected a weapons almost 82.6% of the total. The difference between the total number of weapons and the significant portion of the weapons in the dataset but not all of them. In percentage, it detects number detected represents the "false negatives" or weapons that were missed by the model.

YOLO-NAS comparison result with and without skip connection

Using skip connections led to better performance across most metrics: higher precision, recall, and F1-score, fewer false positives and false negatives, and faster detection time. This makes the system more reliable and efficient for weapon surveillance. Details are shown in below table 4.2

Metric	With Skip Connections	Without Skip Connections
True Positives (TP)	210	208
True Negatives (TN)	22	21
False Positives (FP)	23	21
False Negatives (FN)	25	22
Precision	0.80	0.77
Recall	0.95	0.93
F1-Score	0.87	0.83
Detection Time (ms)	53	50

Table 4.2: Comparison Result With and Without Skip Connection

The table 4.2 compares how well your weapon surveillance and detection system performs when it uses skip connections versus when it uses YOLO-NAS. It draws attention to several crucial variables that show how effectively and accurately the system detects things as shown in figure 4.12.

True Positives (TP): These represent the correctly identified weapons. The system with skip connections detected 210 weapons correctly while without skip connections, it detected 208. As a result, it shows a minor improvement in detection accuracy when using skip connections.

True Negatives (TN): In these cases, the model accurately detected the absence of a weapon. Compared to 22 systems without skip connections, the system with skip connections obtained 25 true negatives, demonstrating improved performance in identifying non-weapon cases.

False Positives (FP): When the system mistakenly labels a non-weapon instance as a weapon, this is known as a false positive. Skip connections resulted in 21 false positives as opposed to 22 without, indicating a decrease in false alarms.

False Negatives (FN): These are the situations in which a weapon was not detected by the system. Improved sensitivity was seen in the 22 false negatives the skip connection method had compared to the 23 without.

Precision: Precision quantifies the proportion of actual positive cases that were expected. The precision is higher with skip connections (0.80%) than without (0.77%), indicating that the system produced fewer false positives.

Recall: The number of true positive cases that were accurately recognized is measured by the recall. Although both models exhibit good performance, the system with skip connections missed fewer guns, as evidenced by its slightly higher recall of 0.95 compared to 0.93 without skip connections.

F1-Score: The harmonic mean of recall and precision is the F1-score. A better balance between precision and recall is indicated by the F1-score of 0.87 with skip connections, which is higher than 0.83 without skip connections.

Detection Time: The detection time of without the skip connection system is 50 milliseconds, which is faster than the 53 millisecond detection time of the skip connection system. This demonstrates both the accuracy gains and the efficiency improvement.

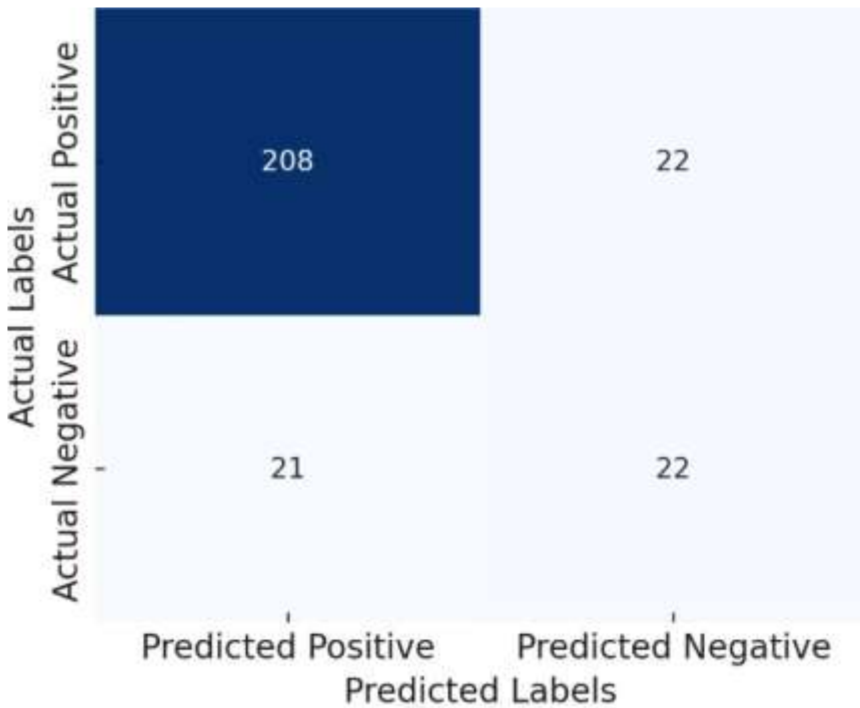

Figure 4.12: Confusion matrix without skip connections

Skip connections have been proved to enhance the performance of the weapon surveillance and detection system by improving m precision, recall, and F1-score as shown in figure 4.13. They help the model detect more accurately, reducing false positives and false negatives. However, in terms of detection time, the model without skip connections performs slightly faster as compared to with skip connection process while skip connections improve accuracy, they may introduce a small computational overhead that impacts processing speed.

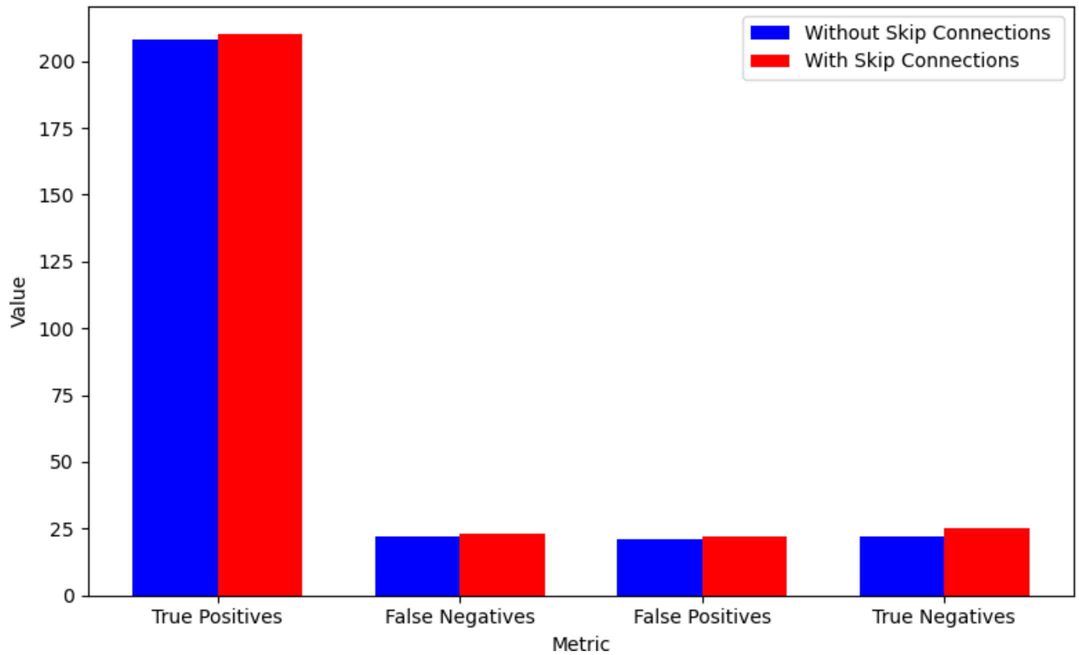

Figure 4.13: Comparison with and without skip connections

Summary

The proposed methodology used for weapon detection and surveillance system shows that more accuracy will be obtained by using the pre-trained model or by training on different hidden layer sizes with different combinations of hyper parameters with more resource consumption. Besides all of these deficiencies, the model learns many patterns converges with more data points, and predicts the quantile values precisely.

Chapter 5 Conclusion, Limitations, and Future Work

This chapter will cover the technical difficulties we encountered, the findings, and how we may enhance the models' predicting ability by making various changes to each one. An assessment of the outcomes and our expectations will be made at the conclusion.

Challenges

The employment of a weapon surveillance system subjects questions about the privacy involved. As it is, the dilemma of finding a balance between public safety and privacy is always difficult. It will be key for future implementations, which should necessarily prioritize transparency, safeguarding privacy, and adhering to the ethics mentioned above, to address these concerns. Flexibility to the advancing or persisting threats and changes in techniques is extremely important. The YOLO-NAS model must be improved constantly to make sure it is up to speed with the new generation of weapons armory and avoid being defeated by the emerging detection evasion plots of the enemies.

Limitations

YOLO-NAS is a prominent sophisticated object identification method known for its speed and accuracy in real-time object recognition applications. It is critical to recognize the study's limitations, which include the accuracy of data sources. Furthermore, the study findings and researchers will be advised on how to enhance weapon monitoring and detection systems utilizing this technology. In addition, future studies might look at the efficacy of other object detection models in weapon detection systems and multiple objects nearby. When numerous weapons are close together or overlap, YOLO-NAS may have trouble processing the situation as it might see them as one combined entity. Moreover, environmental conditions like weather, dim lighting, or cluttered backdrops might make it difficult for YOLO- NAS to detect weapons with efficiency.

There is more that can be done in the future. There is a need to investigate differences between the YOLO-NAS model and other models, and the accompanying challenges to optimize the YOLO-NAS-based Weapon Surveillance and Detection System, address specific problems, and lead the future research direction in the weapon detection domain. Nonetheless, it is important to address a few limitations:

1. **Effectiveness of the Data Sources:** Although YOLO-NAS performs with great accuracy, the network's results depend decisively on the level of data quality and accuracy of the training stages. By rectifying wrong or partial representations in the training data, training an accurate model is never easy and hence, future research may give prominence to the use of more curated and diverse datasets to make the AI systems adaptive.
2. **Environmental Conditions:** Remission of the area's chemical environment (adverse weather, dim lighting, clutter backgrounds) may be a source of system deficiency. Our studies for the future should be concerned with methods that should be applied to improve the model reliability under the complex situations that may arise otherwise with the decrease in the output accuracy.
3. **Depth errors caused by mispositioning of occluded and overlapping object:** The model might show some inefficiency in bottleneck situations when there are many weapons around or where video recordings overlap. Scrutinizing cutting-edge methods that address occlusions and correctly measure weapons in busy scenes is also necessary in this regard.
4. **Privacy and Ethics Considerations:** The integration of surveillance systems with weapons detection poses ethical issues as well as privacy consequences. Security and privacy are necessarily a continuum and that balance. It's crucial to implement transparent policies, and anonymization techniques, and stick to general guidelines to deal with privacy issues.

Future Work

The recent outcome opened new pathways for future work, consisting of modeling improvements, investigation of the advanced object detection methods, and assessment of the model's applicability for diverse purposes. The findings of the implicit task guide the study of new weapon surveillance systems. Thus, this is the section that brings the final point of the study, displaying the achievement of the model proposed in the project, which is the YOLO-NAS-based Weapon Surveillance and Detection System by showing its efficiency. These implications, however, stretch further than just technological innovation, cross pressing ethical concerns, public safety, and areas of technological advances in weapon detection will be concerns later on.

Recommendations for Improvement

In the future, more can be accomplished. To enhance the Weapon Surveillance and Detection System based on YOLO-NAS, it is necessary to examine these recommendations to bring improvement in the system:

5. **Enhanced Data Collection:** Correct and diverse training data can be achieved by making it an imperative recommendation. Activities should focus on assembling databases incorporating environmental situations, such as weapons and wave profiles. Such variability can also belong to the sample in the model and help it learn better generalization.
6. **Advanced Model Fine-tuning:** The gradual refinement of the YOLO- NAS's fine-tuning algorithm will also give the algorithm a high efficiency. Discovering the right use of elaborate fine-tuning approaches, parameter adjustments, and arch combination can be an answer to the fact that is desirable to get the best results in the detection of weapons.

Summary of Findings

The research process included the creation (the cutting-edge weapon surveillance and detection system which is based on YOLO-NAS) and the evaluation of this new and advanced weapon detection model. Key findings include: The YOLO-NAS model which signified high precision with high speed will be used for real- time weapon detection. The task was made of three steps: data pre-processing, initialization of the model with pre-trained weights, and fine-tuning to get the system working better. By contrast, comparative analysis between YOLO-NAS and other methods proves the YOLO-NAS-based systems to be superior in terms of accuracy, adaptability, and real-time operation. Weak points that diminish the functionality of the system under extreme weather and opacity have been identified. These limitations provide an informative basis for future growth.

Conclusion

The summing up paragraph demonstrates the results of the research works, out- lines the contributions, and application of the YOLO- NAS-based Weapon Surveil- lance and Detection System in the real-time domain, as well as concluding remarks on the research. This research article analyzes the Weapon Surveillance and identification System for the YOLO-NAS object identification method. To detect weapons, work has completed a fact-finding analysis. We start with the two versions' overview and examine the architecture and advancements of the earlier iterations. The dataset was split into training and testing sets, each of which included versions that were used to measure performance on a particular dataset and train on it. The performance is estimated using predefined parameters, including Precision, Recall, F1 Score, Quality, mAP, and so on. This research demonstrated that YOLO-NAS works better in terms of performance than other object detection methods. The researchers can take a deeper look at the situation and offer insights into how little changes could have a bigger effect. We will devise a plan to enhance the number of courses and add more photos to our collection to broaden the scope of weapon detection.

YOLO-NAS, a state-of-the-art object detection architecture, has shown promising results in the field of weapon detection and surveillance systems. The key advantages of YOLO-NAS include its ability to detect small objects, its high accuracy, and its efficient performance even on resource constrained devices. Experiments have demonstrated that YOLO-NAS can be effectively trained to detect a variety of weapons, including firearms, knives, and other dangerous objects, with a high degree of accuracy. This makes it a valuable tool for security and surveil- lance applications, where the rapid and reliable identification of potential threats is crucial. Furthermore, the YOLO-NAS architecture's efficient design allows it to be deployed on a wide range of hardware platforms, from edge devices to cloud- based systems, enabling real-time weapon detection and alerting capabilities. This flexibility is particularly important in surveillance applications, where the ability to monitor and respond to threats in a timely manner is essential. Overall, the integration of YOLO-NAS into weapon detection and surveillance systems has the potential to significantly enhance public safety and security, by providing a reliable and efficient means of identifying and addressing potential threats. As the technology continues to evolve, we can expect to see even more advanced and capable weapon detection solutions based on YOLO-NAS and similar deep learning architectures.

REFERENCES

J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," A rXiv preprint arXiv:1804.02767, 2018.
A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," ArXiv, vol. abs/2004.10934, 2020.
W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multi box detector," in Computer Vision – ECCV 2016, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham: Springer International Publishing, 2016, pp. 21–37.
R. M. Alaqil, J. A. Alsuhaibani, B. A. Alhumaidi, R. A. Alnasser, R. D. Alotaibi, and H. Benhidour, "Automatic gun detection from images using faster R-CNN," in 2020 First International Conference of Smart Systems and Emerging Technologies (SMARTTECH), 2020, pp. 149–154.
Z.-W. Sun, Z.-X. Hua, H.-C. Li, Z.-P. Qi, X. Li, Y. Li, and J.-C. Zhang,
"Fbd-sv-2024: Flying bird object detection dataset in surveillance video," 2024.

S. Axelsson, "Intrusion detection systems: A survey and taxonomy," 2000.

J. Beel and B. Gipp, "Academic search engine spam and google scholar's resilience against it," Journal of Electronic Publishing, vol. 13, 12 2010.

M. Andrejevic and N. Selwyn, "Facial recognition technology in schools: crit- ical questions and concerns," Learning, Media and Technology, vol. 45, no. 2,
pp. 115–128, 2020.

N. Z. M. of Health and N. Z. N. E. A. Committee, Ethical Guidelines for Observational Studies: Observational Research, Audits and Related Activities. New Zealand Government - Ministry of Health, 2006.

A.-E. Benrazek, "Internet of things: Analysis of suspicious behaviour in a surveillance camera network," Ph.D. dissertation, 10 2021.

A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intru- sion detection systems: techniques, datasets and challenges," Cybersecurity, vol. 2, 12 2019.

G. Publication, "Weapon detection for smart surveillance system using yolov6," vol. 2, pp. 199–206, 10 2022.

M. Pullakandam, K. Loya, P. Salota, R. M. R. Yanamala, and P. K. Javvaji, "Weapon object detection using quantized yolov8," in 2023 5th International Conference on Energy, Power and Environment: Towards Flexible Green Energy Technologies (ICEPE), 2023, pp. 1–5.

S. Narejo, B. K. Pandey, D. E. Vargas, C. R. Rodriguez, and M. R. Anjum, "Weapon detection using yolo v3 for smart surveillance system," Mathematical Problems in Engineering, vol. 2021, p. 9975700, 2021.

S. Wu, Z. Li, S. Li, Q. Liu, and W. Wu, "Static gesture recognition algorithm based on improved yolov5s," Electronics, vol. 12, p. 596, 01 2023.

J. Terven and D.-M. Cordova-Esparza, "A comprehensive review of yolo: From yolov1 to yolov8 and beyond," 04 2023.

R. Manuel, B. Ioana, and A. Gavrilas, "Application for threat surveil- lance using python and yolo-v7," in 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH), 2023, pp. 1–5.