Object Detection in Low-Visibility Firefighting Scenarios using Thermal Cameras

Meghna Raje

Master of Science
School of Informatics
University of Edinburgh
2023

Abstract

This research project aims to develop an intelligent object detection algorithm optimized for firefighters operating in smoke-filled environments to enhance safety and operational efficiency during rescue missions. The algorithm's primary goal is to accurately identify critical objects in firefighting scenarios, such as doors, windows, and fire hydrants, leveraging thermal image data. The project also focuses on creating a diverse and annotated dataset of thermal images and RGB images, serving as the fundamental training and refinement source for the algorithm.

In this work, three distinct object detection methodologies are explored: the baseline YOLO model, transfer learning technique, and domain adaptation with transfer learning. Each object detection model undergoes a comprehensive evaluation using metrics such as precision, recall, mean Average Precision at an Intersection-over-Union (IoU) value of 0.5 (mAP50), and mean Average Precision across IoU thresholds ranging from 0.5 to 0.95 (mAP50-95).

A comparative visual analysis is carried out to highlight the respective strengths and limitations of each approach in identifying objects within thermal imagery. Through rigorous experimentation and evaluation, domain adaptation with transfer learning emerges as the most effective approach, exhibiting superior accuracy, F1 score, and probability scores for detected objects. The transfer learning approach is the most effective in enhancing detection capabilities as it leverages knowledge from the red-green-blue (RGB) domain to improve results. The domain adaptation technique significantly enhances accuracy and localization in thermal images, with precision rising from 0.728 (baseline) and 0.903 (transfer learning) to 0.935, recall improving from 0.673 (baseline) and 0.840 (transfer learning) to 0.859, mAP50 increasing from 0.743 (baseline) and 0.906 (transfer learning) to 0.919, and mAP50-95 growing from 0.393 (baseline) and 0.613 (transfer learning) to 0.673.

This method effectively addresses challenges related to data scarcity in the thermal domain, bolstering the algorithm's robustness and adaptability in low-visibility scenarios. This research lays a strong foundation for further advancements in intelligent systems for firefighting and emergency response teams.

Research Ethics Approval

This project was planned in accordance with the Informatics Research Ethics policy. It did not involve any aspects that required approval from the Informatics Research Ethics committee.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Meghna Raje*)

Acknowledgements

I express my gratitude to the University of Edinburgh, School of Informatics, for granting me the invaluable opportunity to undertake this dissertation project. Their unwavering support and abundant resources were pivotal in ensuring the successful completion of this work.

I am deeply appreciative of Prof. Chris Lu for his exceptional guidance and supervision. His profound insights and expert direction played a central role in shaping the trajectory of this research and enriching its outcomes significantly.

I extend my sincere thanks to Tianhang Zhang for his invaluable assistance in data collection and engaging discussions about the project's results. His contributions were of exceptional importance and had a significant impact on improving the overall quality of this work.

I am deeply grateful to my family and friends for their unwavering support and encouragement throughout this challenging journey. Their belief in my abilities kept me motivated to overcome obstacles and successfully complete this academic pursuit.

Table of Contents

Chapter 1

Introduction

1.1 Motivation

Firefighting is an inherently demanding and hazardous profession that necessitates specialized training, skills, and tools to effectively safeguard both people and property. The primary objective of firefighters is to extinguish fires and swiftly respond to emergencies across diverse settings, encompassing residential neighborhoods and commercial buildings [1]. Despite advancements in technology and safety protocols, firefighting remains an unpredictable and physically exhausting occupation [2]. Firefighters willingly subject themselves to significant personal risk, driven by their understanding of the inherent dangers involved, as they strive to save lives and extinguish fires [3]. The limited visibility in such environments makes it difficult to assess the fire's magnitude, locate potential victims, and make effective firefighting decisions and operations. Successfully overcoming these challenges demands adaptability, super-fast response, and a firm commitment to public safety.

Advancements in camera technology and object detection algorithms are poised to revolutionize firefighting in the near future [4]. While traditional visible light cameras have limitations in smoke-filled environments, they can still be valuable in certain firefighting scenarios, particularly during daytime or under relatively clear conditions [5]. However, smoke rapidly impairs vision, making visual sensors ineffective in the fire scenario. In such situations, thermal imaging cameras play a crucial role [6]. These cameras utilize infrared ray, which has a longer wavelength than visible lights and shows a stronger smoke penetration ability, to detect the heat signatures of objects, enabling firefighters to perceive their surroundings even through smoke and flames [7]. By providing enhanced visibility, thermal imaging devices assist in identifying individuals, animals, and other objects in smoke-filled environments, empowering firefighters to make informed decisions and take prompt and appropriate actions.

Firefighting demands rapid thinking, stamina, and attentiveness to navigate its hazardous environment [3]. Yet, firefighters encounter hindrances like low lighting and potential dangers when approaching burning structures, impeding their firefighting and life-saving efforts [8]. To overcome these challenges, the introduction of a smart device (e.g., helmet or potable detector) equipped with cutting-edge technologies offers a promising solution [9]. This innovative gear assists firefighters in efficiently navigating smoke-filled areas and promptly identifying vital elements such as doors, windows, and fire hydrants. By improving operational effectiveness and ensuring safety, the smart detector will greatly enhance the potential for saving lives during firefighting operations.

The main objective of this research is to create an intelligent object detection algorithm specifically designed for firefighters, capable of accurately identifying crucial objects like doors, windows, tables, and fire hydrants in a firefighting environment using thermal image data. To achieve this goal, the research involves collecting and annotating thermal image data for training purposes and focusing on developing and optimizing thermal detection algorithms that offer high accuracy and reliability. To overcome the challenge of limited data availability and enhance the algorithm's performance in real-world scenarios, the research incorporates domain adaptation and transfer learning techniques. These approaches enable the algorithm to generalize better to new, unseen data by leveraging knowledge from related domains or pre-trained models, ultimately improving the algorithm's adaptability and effectiveness in real firefighting situations.

The integration of the developed algorithm into the smart detection device empowers firefighters with real-time feedback, augmenting their spatial awareness and reducing the potential risks due to dangerous actions. This integration significantly improves the effectiveness and safety of firefighting tasks, making the proposed algorithm an invaluable tool for firefighters. It enables them to execute their duties with heightened confidence and efficiency, ultimately leading to more successful outcomes in firefighting missions.

1.2 Research Hypothesis and Objectives

The idea is based on the hypothesis that firefighters might benefit significantly from the usage of a real-time algorithm capable of recognising and precisely detecting essential objects. It is predicted that by enhancing their spatial awareness, navigation skills, and task execution, both firemen and victims will benefit.

This thesis presents a study into the critical issue of detecting objects, with an emphasis on enhancing firefighter operational effectiveness and safety during rescue missions. The main goal of this project is to create and implement an intelligent object detection algorithm that can work well in smoke-filled environments. This research aims to advance technologies dedicated to firefighter safety and emergency response by limiting the study to the development and evaluation of the object detection algorithm for field use by firefighters. This research may find applications in sectors that require precise object identification under demanding environmental conditions. Instead of addressing more general issues relating to fire safety, prevention, or emergency management, this thesis concentrates on the specific objective of improving object detection capabilities to increase the overall safety and effectiveness of firefighters during rescue missions.

1.3 Thesis Overview

The thesis aims to create an intelligent object detection algorithm tailored for firefighters, capable of accurately identifying crucial objects in a firefighting environment using thermal image data. It begins with a comprehensive introduction in Chapter 1, laying the groundwork for the study. Chapter 2 explores the background of thermal cameras for object detection and various techniques, including traditional methods, two-stage detectors, and one-stage detectors. It also covers domain adaptation and transfer learning, crucial for real-world algorithm enhancement. In Chapter 3, data collection is discussed, encompassing the setup for data acquisition and the methodology used to collect and annotate thermal image data. Bounding box calibration to thermal images is also addressed. Chapter 4 presents the methodology, starting with data and object class analysis. It further explores the development of baseline models, enhanced through transfer learning, and incorporates domain adaptation techniques for improved real-world performance. Chapter 5 provides a comprehensive analysis and evaluation of the developed object detection models. It assesses the baseline model's performance, the impact of transfer learning, and the effectiveness of domain adaptation, accompanied by a comparative visual analysis. Chapter 6 concludes the thesis by summarizing key findings and accomplishments while discussing potential future enhancements to the intelligent object detection algorithm for firefighters.

Chapter 2

Background

Object detection is a crucial task in computer vision, finding diverse applications [10]. Recent progress in deep learning algorithms and the availability of large datasets have significantly advanced object detection techniques [11]. One area gaining traction is thermal cameras for object detection, excelling in challenging conditions like low-light and smoky environments [12]. This chapter explores current research and developments in thermal camera usage for accurate and efficient object detection. The chapter covers various object detection techniques, including traditional methods, two-stage detectors like Faster Regions with Convolutional Neural Networks (R-CNN) and Mask R-CNN, and one-stage detectors like You Only Look Once (YOLO) and SSD. Transfer learning, a powerful approach, is also discussed for adapting pre-trained models to object detection tasks. Additionally, the concept of domain adaptation is examined to address the challenge of limited availability of data. By reviewing advancements in thermal object detection techniques, such as transfer learning, and domain adaptation, this chapter aims to identify state-of-the-art methods and discuss challenges and opportunities. The insights gained will aid future research in thermal object detection and domain adaptation, ensuring improved performance in real-world scenarios.

2.1 Thermal Cameras for Object Detection

Thermal cameras have gained considerable attention in recent years due to their exceptional performance in object detection, particularly in demanding environments where traditional visible cameras are limited [13]. Distinguishing them from conventional red-green-blue (RGB) cameras is their ability to capture and process the infrared radiation emitted by objects, thereby enabling the detection of heat signatures and facilitating enhanced visibility even in low-light scenarios, areas filled with smoke, and other visually challenging conditions [14]. This distinctive attribute empowers thermal cameras to prove remarkably effective across a diverse range of applications, encompassing surveillance systems, search and rescue missions [15], industrial inspections [16], and automotive safety measures [17].

Thermal cameras have emerged as the preferred option in firefighting due to their superior capabilities compared to RGB cameras [18]. In firefighting scenarios, visibility is often severely impaired by smoke and low-light conditions, making it challenging for firefighters to assess the situation accurately. Thermal cameras, equipped with infrared technology, overcome these obstacles by detecting heat signatures emitted by objects [19]. This enables firefighters to see through smoke and darkness, effectively identifying individuals, hotspots, and crucial elements like doors and windows. The real-time thermal imaging provided by these cameras enhances situational awareness, allowing firefighters to navigate hazardous environments and make informed decisions [20]. Moreover, thermal cameras enable the visualization of temperature variations, helping to detect hidden sources of fire and structural weaknesses. With their ability to improve visibility, enhance situational awareness, and aid decision-making, thermal cameras have proven to be an invaluable tool in firefighting operations, ultimately leading to more effective fire suppression and increased safety for firefighters and victims [21].

In the realm of object detection using thermal cameras, significant progress has been made in recent studies. One notable paper by Dai et al. [22] introduces an object detection algorithm specifically designed for thermal images, demonstrating impressive performance in detecting humans and vehicles. Similarly, other studies [23, 24, 25] have focused on the detection of pedestrians and vehicles using thermal cameras, showcasing promising results. However, it is worth noting that the majority of these research efforts [26, 27] have primarily concentrated on the identification of individuals and vehicles, with less emphasis on other critical objects such as doors and windows in firefighting scenarios. This presents an important gap in the existing literature, highlighting the need for further research and development to address the specific challenges faced by firefighters in identifying these crucial elements during firefighting operations.

Firefighter safety in smoke-filled environments is of paramount importance. Previous research by Tsai et al. [28] has proposed a deep learning-based approach that focuses on detecting humans in thermal images specifically in heavy smoke conditions. While this study addresses the crucial task of human detection, it overlooks the recognition of other vital objects that play a significant role in firefighting scenarios. The identification of objects such as doors, windows, and potential obstacles is essential for effective firefighting operations. Another study [29] addresses this gap by developing an object detection algorithm using thermal cameras, specifically designed to enhance firefighters' safety. However, this approach heavily relies on large datasets for training the object detection models, which can be challenging in real-world firefighting situations where annotated data is limited. Therefore, this research aims to explore alternative methods that can overcome the constraints of data availability and successfully detect various important objects in smoke-filled environments.

2.2 Object Detection Techniques

This section aims to explore the various approaches and advancements in object detection, a fundamental task in computer vision. Object detection involves the identification and localization of objects of interest within images or videos. Over the years, numerous techniques have been developed, each with its own unique strengths and limitations. To provide a comprehensive overview, object detection techniques are categorised into three broader groups.

2.2.1 Traditional Methods

Traditional object detection techniques, such as Viola-Jones [30], Histogram of Oriented Gradients (HOG) [31], and Deformable Parts Models (DPM) [32], have made significant contributions to computer vision [33]. The Viola-Jones algorithm efficiently detects objects using Haar-like features and cascaded classifiers, originally gaining popularity in face detection [34]. HOG-based detectors capture object appearance through local gradient distributions and employ classifiers like Support Vector Machine (SVM) [35]. DPM represents objects as deformable parts and models their spatial relationships using a pictorial structure framework [36].

These traditional methods have advantages such as computational efficiency and interpretability. They excel in scenarios with simpler object appearances or limited computational resources. However, they may struggle with complex object variations, requiring manually designed features that may not fully capture object complexities [37]. Deep learning-based approaches have surpassed traditional techniques by automatically learning discriminative features from data [38]. Nevertheless, traditional object detection methods laid the foundation for subsequent advancements and retain relevance in specific applications and research contexts.

2.2.2 Two-Stage Detectors

Two-stage detectors are a popular class of object detection algorithms that consist of two main components: a region proposal network (RPN) and a classifier. These detectors excel at accurately localizing and classifying objects in images, making them widely used in computer vision tasks [39]. Two-stage detectors are a popular class of object detection algorithms that consist of two main components: a region proposal network (RPN) and a classifier. These detectors excel at accurately localizing and classifying objects in images, making them widely used in computer vision tasks.

One prominent example of a two-stage detector is R-CNN [40]. R-CNN generates region proposals using a selective search algorithm, which identifies potential object regions in an image. These proposed regions are then individually classified using a convolutional neural network (CNN), enabling accurate object detection. However, R-CNN can be computationally expensive due to the independent processing of each proposed region [41].

To make it faster, Fast R-CNN [42] was introduced. It shares the work by using the same convolutional features for multiple region proposals at once. This reduces the time needed for computations and improves the speed of detection.

A further improvement came with Faster R-CNN [41]. It added the RPN, which generates region proposals and predicts objectness scores and bounding box offsets. By training the RPN and the detection network together, Faster R-CNN achieves better accuracy and efficiency.

These two-stage detectors are essential for various computer vision applications, such as self-driving cars [43] and surveillance systems [44]. They combine precise object localization with efficient processing, helping identification objects more effectively.

2.2.3 One-Stage Detectors

One-stage detectors, such as YOLO [45] and SSD (Single Shot MultiBox Detector) [46], are widely used object detection algorithms that directly predict object bounding boxes and class probabilities from fixed grids or anchor boxes. These detectors offer a different approach compared to two-stage detectors, simplifying the detection process and achieving real-time performance.

YOLO divides the input image into a grid and processes it as a whole, making predictions for bounding boxes and class probabilities for each grid cell in a single pass. This grid-based approach enables YOLO to efficiently detect objects, but it may struggle with accurately localizing small objects due to the coarser grid resolution.

SSD [46] emerged as a significant advancement in object detection by utilizing the VGG-16 [47] architecture and introducing additional layers for improved performance. It stood out as the first single-stage detector capable of matching the performance of traditional two-stage detectors in terms of accuracy metrics. However, despite its advantages over YOLO and faster R-CNN in terms of both speed and accuracy, SSD faced certain limitations. It encountered challenges in accurately detecting smaller objects within images. To overcome these limitations and further enhance object detection, subsequent versions of YOLO, including YOLOv3 [48], and YOLOv4 [49], were introduced. These iterations of YOLO offered even better speed, accuracy and addressed the specific issues that the earlier versions faced.

RetinaNet [50] revolutionized object detection with the introduction of its novel focal loss and feature pyramid network (FPN). The focal loss, a key component of RetinaNet, focuses the training process on challenging examples, leading to improved accuracy. By combining this focal loss with the ResNet [51] backbone, RetinaNet achieved increased accuracy and speed compared to traditional two-stage detectors. Additionally, RetinaNet is known for its ease of implementation, faster training convergence, and efficient utilization of the FPN, which effectively captures objects at various scales.

2.3 Domain Adaptation in Object Detection

In the domain of thermal object detection, the utilization of synthetic thermal images generated from visible images has become a prevalent technique to enhance detection performance. For instance, in the work presented by Liu et al. [52], an innovative unsupervised domain adaptation method for thermal object detection employs an image generation network to create synthetic thermal images. Similarly, Guo et al. [53] proposed a domain adaptation approach for pedestrian detection in thermal images using a CycleGAN-based (Cycle-Consistent Generative Adversarial Network) image translation technique to produce synthetic thermal images from visible light images. Both of these methods leverage synthetic images for either unsupervised or supervised training, leading to improved object detection performance compared to existing domain adaptation approaches.

In thermal object detection, Lee et al. [54] introduced a distinctive approach for domain adaptation. Emphasizing edge information preservation during image translation, it accurately transfers details from RGB to synthetic thermal images. This strategy bridges the domain gap, enhancing object detection effectiveness in thermal imagery. The project leverages this technique to explore better ways of improving thermal object detection and robustness in real-world scenarios.

Another notable study by Kieu et al. [55] introduced a compelling LSGAN-based (Least Squares Generative Adversarial Network) method that facilitates the transformation of visible spectrum photos into thermal spectrum representations. This technique, accomplished through mixing training, greatly enhances pedestrian detection in thermal images. The demonstrated success of these domain adaptation techniques in adapting object detection models to different domains, even with limited data, underscores their suitability and relevance for the proposed research project. By leveraging the power of synthetic thermal images, the project can explore novel and effective ways to improve thermal object detection performance and robustness in real-world scenarios.

2.4 Transfer Learning

Transfer learning is a machine learning technique that allows the transfer of knowledge gained from one domain to another, typically by using a pre-trained model on a source domain to improve performance in a target domain [56]. In the case of transferring learning from RGB to thermal imaging, the objective is to leverage the knowledge learned from RGB images (with three color channels) to enhance the analysis of thermal images (with a single channel representing heat signatures).

The Transfer learning from RGB to thermal imaging is valuable because RGB images capture visual information for object recognition, while thermal images reveal heat patterns and anomalies. By transferring learned features from RGB to thermal, models gain a general understanding of objects and improve tasks like object detection, anomaly detection, and activity recognition in thermal images. This approach leverages the strengths of both modalities, enhancing the analysis of thermal data and enabling insights that are not visible to the human eye alone.

Several recent papers have explored transfer learning from RGB to thermal imaging. For example, a study by Mantau et al. [57] proposed using the YOLOv5 framework and pre-trained models from the MS COCO dataset [58] to enhance object detection in surveillance systems employing RGB and thermal imaging from UAVs. A semi-supervised transfer learning method was proposed by Weipeng et al. [59] for infrared object detection, utilizing a pre-trained RGB model to improve efficiency and generalization on limited-scale IR datasets. These papers demonstrate the effectiveness of transfer learning from RGB to thermal imaging.

2.5 Summary

The exploration of thermal cameras for accurate and efficient object detection reveals their efficacy in challenging conditions such as low-light and smoky environments, making them highly valuable in firefighting, surveillance, and industrial inspection applications. This investigation highlights advancements in object detection techniques, categorizing them into traditional methods, two-stage detectors, and one-stage detectors. Notable algorithms like R-CNN, Faster R-CNN, YOLO, SSD, and RetinaNet are discussed, each offering unique strengths in detecting objects.

Additionally, the concept of domain adaptation is addressed, which deals with the challenge of deploying algorithms trained in one domain, such as RGB images, to effectively work in another domain, such as thermal images. Techniques like using synthetic thermal images and applying transfer learning from RGB to thermal imaging demonstrate promising potential in enhancing object detection performance in thermal imagery, even when faced with limited data availability.

2.6 Research Gaps

1. **Unavailability of Indoor Thermal Data**

   One of the major gaps in current literature is the absence of any open-source dataset that provides annotated indoor thermal data suitable for training an object detection algorithm. There are no existing studies that have tackled this specific issue, and therefore, there is no dataset available that adequately represents indoor thermal conditions. This severe data scarcity poses a significant challenge in developing a robust and accurate object detection model for indoor environments. Hence, the project aims to address this critical gap by exploring innovative and feasible data collection strategies to create a comprehensive and diverse dataset tailored to indoor thermal conditions.

2. **Transfer Learning from RGB to Thermal**

   This research area faces challenges in training an accurate object detection algorithm for indoor thermal environments, primarily due to the limited availability of annotated thermal data. While some existing studies have explored transfer learning techniques from RGB to thermal domains, they have predominantly focused on outdoor scenarios, leaving a significant gap in the literature for indoor thermal datasets. The key research gap lies in investigating effective transfer learning techniques that can leverage pre-trained models from RGB image domains to improve the performance of the object detection model on the collected indoor thermal dataset. By addressing this unexplored area, this project aims to develop a robust and efficient object detection algorithm tailored specifically for indoor thermal scenarios, where transfer learning from RGB to thermal has not been adequately studied yet.

3. **Domain Adaptation for Indoor Thermal Environments**

   The existing literature on thermal object detection has primarily focused on domain adaptation methods for outdoor scenarios, with limited exploration for indoor thermal environments. This research gap poses a challenge in achieving accurate and robust object detection in indoor thermal settings. Thus, this project tries to address this gap by adapting existing domain adaptation methods developed for outdoor scenarios to suit indoor thermal datasets. By exploring the effectiveness of these adapted domain adaptation techniques in indoor scenarios, the project aims to contribute to the advancement of indoor thermal object detection.

Chapter 3

Data Collection

Data collection is an essential aspect of training machine learning models, providing the necessary foundation for their development and evaluation. While datasets for outdoor environments exist, primarily consisting of objects like cars, they are not suitable for our needs. To address this gap and facilitate effective indoor thermal analysis, a dedicated dataset needs to be collected. The dataset should capture thermal properties and temperature variations of indoor objects, including doors, windows, and fire hydrants. This dataset will serve as the foundation for developing machine learning models capable of accurately detecting and analyzing these indoor objects.

This chapter presents a comprehensive overview of the methodology and procedures adopted for data collection. The details of the data acquisition setup, including the specialized hardware employed, are discussed. The techniques and processes involved in capturing thermal data along with RGB images for bounding box labeling are explained. Furthermore, the calibration procedure utilized to align the bounding boxes with the corresponding thermal images ensuring accurate object labeling is explored.

3.1 Data Acquisition Setup

The data acquisition setup utilized for this project involved the combination of an RGB camera and a FLIR Boson 640 camera as shown in Figure 3.1. These cameras were strategically positioned next to each other along the same axis to facilitate alignment and reduce the calibration effort. The RGB camera was placed above the FLIR Boson 640 camera, ensuring that both cameras captured the same scene to the largest extent.

The RGB camera used in the setup had a resolution of 1920x1080 pixels and an FPS (frame per second) of 30. It was capable of capturing high-quality color images,

Figure 3.1: Platform with RGB and Thermal sensor

providing visual details and context for the objects of interest. On the other hand, the FLIR Boson 640 camera offered a thermal imaging capability with a resolution of 640x512 pixels and an FPS of 9. This specialized camera was specifically chosen for its high sensitivity and ability to capture fine-grained thermal data.

Both the RGB and thermal cameras were connected to an NVIDIA processor, which played a crucial role in real-time data collecting, processing, and storage. The processor facilitated efficient communication and synchronization between the cameras, ensuring that the captured RGB and thermal data were properly aligned and timestamped.

During the data acquisition process, the synchronized RGB and thermal data were stored as '.bag' packages, a file format commonly used in the Robot Operating System (ROS) framework. '.bag' files are capable of storing time-stamped messages, making them ideal for storing and managing sensor data in ROS applications. From these '.bag' files, individual images were extracted and saved for further analysis and annotation.

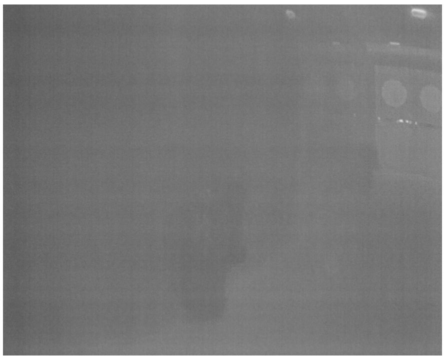
To ensure accurate pairing of RGB and thermal images, the images with nearby timestamps were carefully paired together. This allowed for synchronized visual and thermal data, ensuring a coherent dataset for subsequent analysis and modeling tasks. This setup facilitated the collection of comprehensive and properly aligned data for subsequent analysis, annotation, and machine learning model development.

3.2 Thermal Data Collection

During the data collection process, a variety of objects related to safety and security in indoor environments were included. This includes doors, windows, fire extinguishers, free pathways, and fire exits, as they are crucial for ensuring proper evacuation routes and identifying potential hazards. By incorporating these objects into the dataset, the project aims to develop comprehensive models that can accurately detect and analyze relevant objects in smoke-filled environments.

Figures 3.2 and 3.3 show some examples from the dataset. These examples showcase image pairs that combine thermal and RGB images, allowing to observe the captured data from both thermal and visual perspectives.

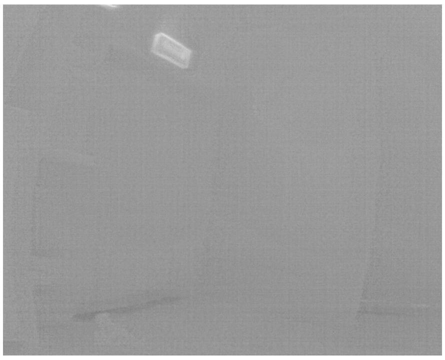
(a) RGB image of fire extinguisher



(b) Thermal image of fire extinguisher

Figure 3.2: Example of fire extinguisher image

(a) RGB image of exit sign



(b) Thermal image of exit sign

Figure 3.3: Example of exit sign image

Privacy protection was given utmost importance throughout the data collection process. As the dataset focused solely on objects and did not contain any personal information or images of humans, there was no risk to privacy. Data collection was conducted exclusively during daytime within the university premises. The selected locations for data collection included office corridors, lecture theaters, study rooms, and labs, providing a diverse set of indoor environments. These locations were specifically chosen to capture thermal patterns in areas commonly found within indoor settings.

To ensure a comprehensive dataset, a substantial number of images were collected for each object class. The data collection process yielded a dataset consisting of 920 thermal images. Each image represented a specific object class, such as doors, windows, fire extinguishers, free pathways, or fire exits. The dataset encompassed a wide range of thermal patterns and temperatures associated with these objects, enabling accurate analysis and modeling.

3.3 Bounding Box Labeling in RGB Images

The process of bounding box labeling in RGB images plays a pivotal role in generating annotated data. The RGB labeled data generated through the bounding box labeling process serves a specific purpose in this research project. Its primary role is to facilitate the acquisition of thermal labels for the corresponding objects.

RGB images are crucial for generating annotations as direct labeling in thermal images presents challenges. Thermal images primarily capture temperature differences, making it difficult to precisely identify and label objects. By utilizing RGB images as a reference, the labeling process can accurately assign annotations to corresponding thermal images. This approach ensures the training of an effective object detection algorithm for recognizing critical objects.

The first step in the labeling process involved acquiring RGB images, which served as the basis for generating annotations. These RGB images provided visual representations of the objects of interest in the indoor environments. Once the RGB images were obtained, they were processed using the LangSAM [60] algorithm, specifically designed for object detection and segmentation tasks.

The LangSAM algorithm utilizes text prompts to predict masks, bounding boxes, labels, and logits associated with the objects within the RGB images. By providing text prompts related to the objects of interest, such as “door”, “window”, “fire extinguisher”, “emergency sign”, and “clear pathway”, the algorithm can accurately identify and localize these objects within the images.

During the implementation phase, the LangSAM algorithm was applied to each RGB image, generating bounding boxes as outputs. These bounding boxes were then stored for further analysis and use. The next section focuses on the calibration process, which is undertaken to calibrate and annotate corresponding thermal images.

3.4 Calibration of Coordinate System of RGB-Thermal Images

Calibration is a crucial process that establishes a connection between the RGB and thermal domains, allowing for the transfer of labels from RGB images to corresponding thermal images. The calibration process involves the following steps:

1. **Preparation**

   The calibration process involved using a checkerboard pattern as seen in Figure 3.4 with a 4x3 grid, which had 48 points in total. This pattern was carefully prepared to work with both the RGB and thermal cameras. It helped establish a connection between the RGB and thermal images by providing a reference for matching points. This step was important for transferring labels from RGB to thermal images accurately.

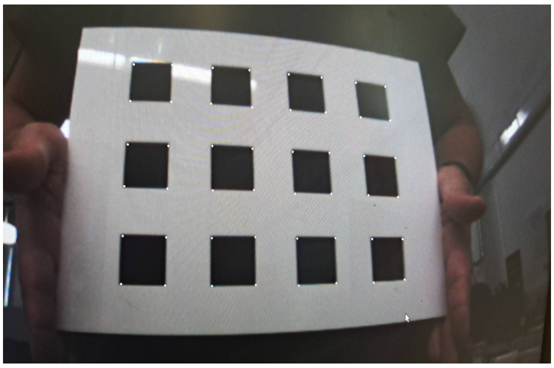


Figure 3.4: Checkboard Pattern

2. **Image Capture**

   Sets of calibration images were captured using both the RGB and thermal cameras. The checkerboard pattern was placed at various angles and positions within the camera frame to ensure visibility and accommodate different perspectives. Special attention was given to capturing clear and well-exposed images to ensure accurate calibration results.

3. **Point Selection**

   The subsequent step required the manual selection of corresponding points in both the RGB and thermal images. This was accomplished by using a mouse click program, which allowed for accurate marking of the corners on the checkerboard pattern. The program provided the coordinates of the mouse clicks, ensuring precise selection in both types of images. The use of specialized software and careful mouse clicking ensured the accurate identification of corresponding points for an effective calibration procedure.

4. **Point Correspondence**

   Once the points were selected in both the RGB and thermal images, point correspondence was established by pairing the corresponding points. Each pair represented the same physical location in the real world, but captured by different cameras. This correspondence formed the basis for subsequent calibration computations.

5. **Calibration Computation**

   The calibration step involved a calibration algorithm that determined the transformation between the RGB and thermal domains. This algorithm used manually selected corresponding points from the RGB and thermal images and applied techniques like least-squares estimation to calculate a transformation matrix. This matrix captured the geometric relationship between the two camera systems, enabling accurate calibration. It facilitated the transfer of labels and information between the RGB and thermal domains, ensuring precise correspondence and reliable label transfer. The calibration algorithm employed sophisticated mathematical calculations to achieve an optimal alignment between the RGB and thermal cameras.

6. **Calibration Transformation**

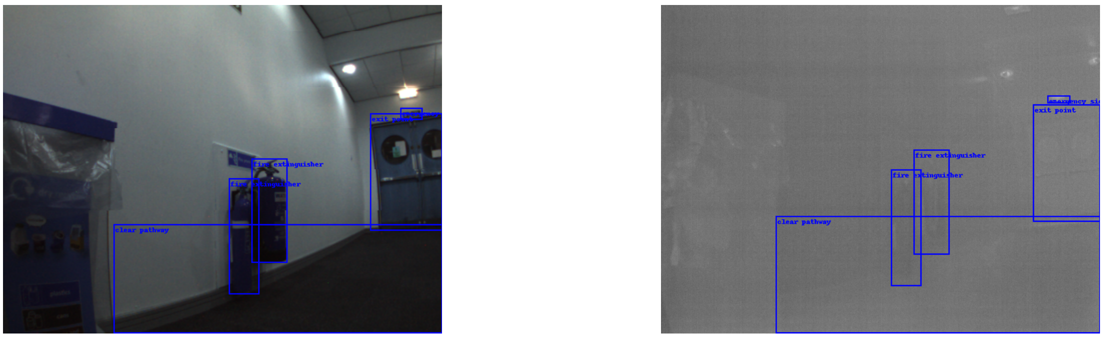
   The computed transformation matrix was then applied to map points from the RGB image domain to the thermal image domain. This step facilitated the transfer of labels from RGB images to corresponding thermal images based on the established point correspondence. By leveraging the calibration transformation, the labeled data in the RGB domain was effectively utilized and applied to the thermal domain.

Chapter 4

Methodology

4.1 Data Analysis

The dataset at hand comprises 920 RGB images and 920 thermal images, meticulously collected from diverse sources within a real-world environment. This comprehensive dataset encompasses a total of 3079 annotated objects. The dataset encompasses four distinct classes, namely “exit point”, “fire extinguisher”, “emergency sign”, and “clear pathway”. The “exit point” class encompasses annotated doors and windows, while the other classes refer to specific elements essential for safety and accessibility within the observed context. The objects within the dataset have been annotated using the widely adopted YOLO format, ensuring compatibility and ease of use for further analysis and research purposes. To provide a visual reference, an example image is presented in Figure 4.1. This example image showcases a scenario where a clear pathway, an exit point, and a fire extinguisher are present.



(a) RGB imager

(b) Thermal imager

Figure 4.1: Example of RGB and thermal image

4.2 Object Class Analysis

The dataset includes four distinct object classes: “exit point”, “fire extinguisher”, “emergency sign”, and “clear pathway”. The distribution of images per object class is visualized in Figure 4.2. As seen in the graph, the “exit point” class has the highest number of images, with 743 instances. Following that, the “emergency sign” class has 539 images, the “fire extinguisher” class has 499 images, and the “clear pathway” class has 429 images.

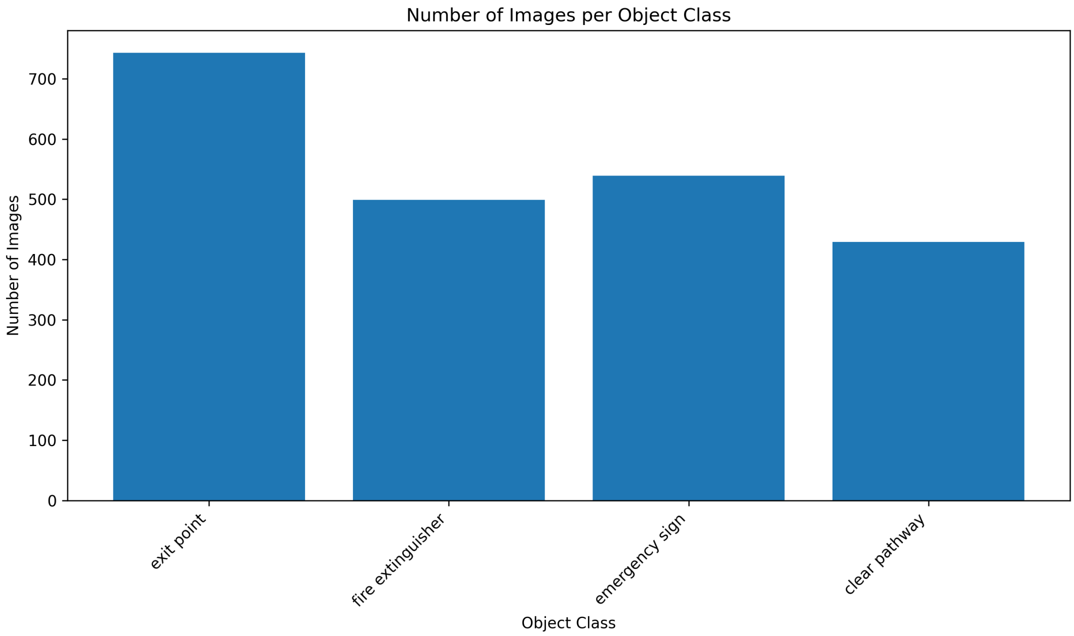


Figure 4.2: Distribution of Objects per Object Class

To further analyze the dataset, an examination of the number of objects per object class was conducted, as illustrated in Figure 4.3. This graph provides a visual representation of the distribution of annotated objects across the different object classes. The “exit point” class exhibits the highest count, with 1101 annotated objects, indicating its prominence within the dataset. Following this, the “fire extinguisher” class is represented by 739 objects, highlighting its significant presence. The “emergency sign” class comprises 540 objects, while the “clear pathway” class demonstrates an equal count of 429 objects.

The analysis of the object class distribution provides valuable insights into the relative importance and occurrence of safety-related objects within the observed environment. By examining the number of objects per class, we gain a deeper understanding of the dataset’s composition. This information is crucial for developing object detection algorithms as it guides the training process and ensures a balanced representation of object categories. Algorithms trained on such datasets can accurately detect and classify safety-related objects, enhancing safety planning, emergency response preparedness, and accessibility assessments. The object class analysis serves as a foundation for developing effective object detection algorithms that can recognize and understand safety-related objects in real-world environments.

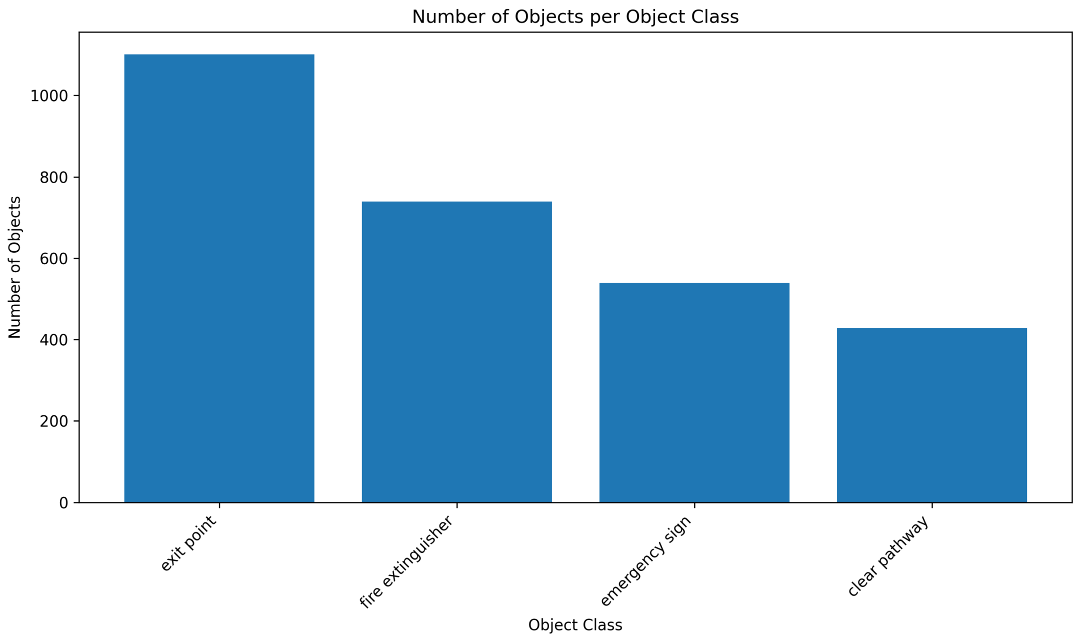


Figure 4.3: Distribution of Images per Object Class

4.3 Baseline Model

The baseline model selected for this research project is the YOLO algorithm [45], renowned for its robust object detection capabilities. The choice of YOLO as the baseline stems from its exceptional performance and suitability for the task at hand.

The YOLO algorithm offers several advantages that make it an ideal candidate for object detection. One of its notable strengths is its real-time performance, enabling efficient processing of live video or image streams. This real-time capability is of utmost importance in scenarios where quick and accurate object detection is required to make timely decisions and take appropriate actions.

The training process for the baseline YOLO model involved several crucial steps. Firstly, a dataset of annotated images specifically tailored to the object detection task was collected. These annotated images serve as the foundation for training and evaluating the YOLO algorithm.

The collected dataset was divided into three distinct subsets: the training set, the test set, and the development (or validation) set. This division allows for effective training, evaluation, and fine-tuning of the baseline YOLO model. The training set constitutes the largest portion and was used to train the YOLO model, enabling it to learn to recognize and localize objects of interest. The development set provides a separate set of images to fine-tune the model and assess its generalization capabilities. Finally, the test set remains unseen during training and serves as an unbiased benchmark to evaluate the final performance of the model.

During the training phase, the YOLO architecture was adapted and optimized to suit the object detection task. The network architecture, consisting of 157 layers, includes convolutional layers followed by fully connected layers that extract discriminative features from the input images. The network was trained to predict both the bounding box coordinates and class probabilities for the objects of interest.

Optimization techniques such as backpropagation, using stochastic gradient descent (SGD) algorithm [61], were employed to iteratively update the model's parameters. This process minimizes the discrepancy between the predicted bounding boxes and class labels and the ground truth annotations.

Once the baseline YOLO model was trained and optimized, its performance was rigorously evaluated using appropriate metrics. Commonly used metrics include precision, recall, and mean average precision (mAP). The evaluation process involved comparing the model's predictions against the ground truth annotations to measure its accuracy, reliability, and overall effectiveness in detecting objects.

The YOLO algorithm was selected as the baseline model due to its real-time performance and suitability for object detection tasks. The training process involved adapting and optimizing the YOLO architecture, dividing the dataset into training, test, and development sets, iteratively updating the model's parameters, and evaluating its performance using relevant metrics. This comprehensive approach ensured the baseline model's ability to accurately detect objects and lays the foundation for further enhancements and improvements in object detection algorithms.

4.4 Transfer Learning Model

Transfer learning is a powerful technique utilized in computer vision, specifically for object detection tasks [62]. When confronted with limited annotated data in the target domain, transfer learning enables the transfer of knowledge gained from a source domain with abundant labeled data.

One of the challenges in thermal object detection arises from the scarcity of annotated thermal data. However, by employing transfer learning, the unique characteristics of RGB images can be leveraged to enhance the performance of the thermal object detector. While RGB images encompass three color channels, thermal images predominantly have a single channel. By utilizing transfer learning, the distinctive features learned from RGB images can be harnessed to improve the performance of thermal object detection.

By utilizing a pretrained RGB object detector, the knowledge and feature extraction capabilities acquired during training on the RGB dataset are transferred to the thermal object detection task. The pretrained RGB weights provide a valuable starting point for the thermal detector, allowing it to benefit from the learned knowledge. Although the number of channels differs between the RGB and thermal images, the transferred weights serve as a foundation for effective feature extraction and object detection in the thermal domain.

This transfer of knowledge addresses the limitations posed by the scarcity of labeled thermal data. The pretrained RGB detector captures valuable insights from the RGB domain, which can be applied to enhance the performance of the thermal object detector.

Through the integration of transfer learning, the model leverages the strengths of the RGB and thermal domains. By initializing the thermal object detector with the pretrained weights of the RGB detector, the model benefits from the knowledge acquired in the RGB domain and gains the ability to detect objects effectively in the thermal domain. This approach enhances the thermal object detection system and expands its applicability in various domains, including surveillance, autonomous driving, and search and rescue missions.

4.5 Domain Adaptation with Transfer Learning Model

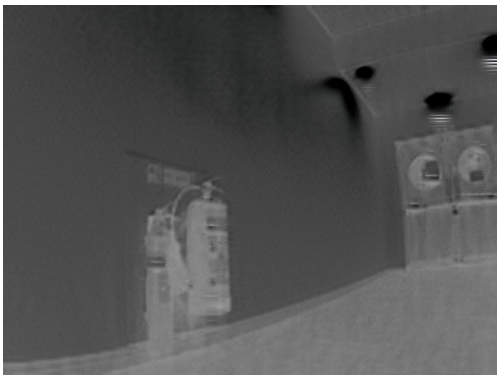
Incorporating domain adaptation techniques in thermal object detection research is crucial to address the challenges posed by the scarcity of thermal data, which can limit the system's performance. Leveraging the wealth of annotated data available in the source domain, such as RGB images, allows us to generate synthetic images in the thermal domain. Augmenting the training dataset with synthetic images enhances the thermal object detection system's performance, as it increases data size and improves the model's generalization capabilities.

To achieve domain adaptation, a specialized approach based on the method proposed by Lee et al. [54] was employed. This model was carefully designed with a strong emphasis on preserving edge information during the image translation process. By ensuring the accurate transfer of critical details from the original RGB images to the synthetic thermal images, the aim was to bridge the gap between the two domains, consequently boosting the effectiveness of the object detection system in thermal imagery.

Through the application of this translation model, a mixed dataset was carefuly constructed, intelligently combining approximately 10% synthetic thermal images with 90% real-world thermal images. This strategic dataset composition was crucial in effectively addressing the limited availability of labeled thermal data while preserving the authenticity and realism of the majority real-world samples.

Domain adaptation plays a pivotal role in this research, empowering the thermal object detection system to directly benefit from the specialized image translation method. By effectively bridging the gap between the RGB and thermal domains, this method significantly improved the system's accuracy and robustness. The integration of the proposed translation model and the mixed dataset offers a promising solution to tackle the challenges of thermal object detection, thus paving the way for further improvements in this critical area of computer vision research. Figure 4.4 shows some examples of synthetic images generated using the image translation method employed in this study. The images show corresponding pairs of original images and their corresponding synthetic thermal images. Notwithstanding the imperfections, these synthetic thermal images can be suitably employed for training purposes.

(a) Original RGB Image (Consisting of Fire Extinguisher)

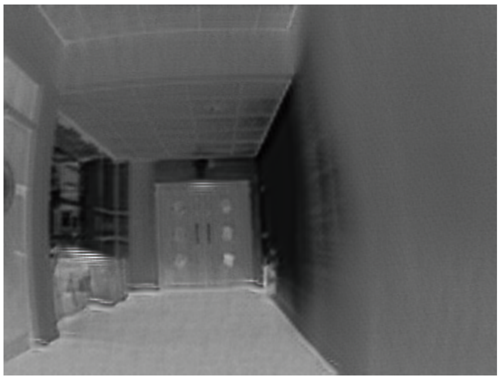


(b) Synthetic Thermal Image (Consisting of Fire Extinguisher)

(c) Original RGB Image (Contains Exit Point and Emergency Sign)

(d) Synthetic Thermal Image (Contains Exit Point and Emergency Sign)

(e) Original RGB Image (Contains Exit Points and Fire Extinguisher)



(f) Synthetic Thermal Image (Contains Exit Points and Fire Extinguisher)

Figure 4.4: Comparison of Original RGB Images and Synthetic Thermal Images

4.6 Performance Evaluation Metric

This section presents a method for comprehensive evaluation of the object detection models developed in this study using specific evaluation metrics. The metrics used for assessment include ‘precision’, ‘recall’, mean Average Precision at an Intersection-over-Union (IoU) value of 0.5 (mAP50), and mean Average Precision across IoU thresholds from 0.5 to 0.95 (mAP50-95) to thoroughly analyze the detection capabilities of each model.

‘Precision’ quantifies the models’ selectivity in identifying objects by measuring the ratio of correctly predicted instances to all instances predicted as a particular class. It can be calculated as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{4.1}$$

‘Recall’ evaluates the models’ sensitivity in capturing objects by calculating the proportion of correctly predicted instances to all instances of the class present in the ground truth. It can be calculated as:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{4.2}$$

The mAP50 metric gauges the models’ overall object detection performance, considering detections with an IoU of 0.5 or higher as correct.The mAP50-95 provides a comprehensive assessment across a broader range of IoU thresholds, and is calculated as:

$$\text{mAP50-95} = \frac{1}{|C|} \sum_{i=1}^{|C|} \text{AP}_i \tag{4.3}$$

where AP_i is the Average Precision for class i, and $|C|$ is the total number of classes.

By employing these evaluation metrics,the performance of each object detection model is analyzed on individual classes and across all classes, assisting in the identification of strengths and weaknesses. This analysis enables optimization and refinement of the algorithms, ultimately enhancing their object detection capabilities in real-world applications.

The results of this evaluation are further discussed in the next chapter, where a detailed comparison of the models’ performances is presented.

Chapter 5

Results Analysis and Evaluation

5.1 Baseline Model

This section shows the evaluation results of the trained baseline model for object detection. Its performance is demonstrated on 184 images, comprising of a total of 604 instances of objects. Precision and recall are crucial indicators of the model's performance, which shows an overall precision of 72.8% and a recall of 67.3%, as seen in Table 5.1. The mAP50 score, assessing bounding box localization accuracy, is 74.3% for this model. This signifies the model's capability to accurately predict object locations with an average IoU of 0.39.

Table 5.1: YOLOv5s Baseline Model Evaluation Results

Class	Images	Instances	Precision	Recall	mAP50	mAP50-95
All	184	604	0.728	0.673	0.743	0.393
Exit Point	184	241	0.701	0.614	0.690	0.380
Fire Extinguisher	184	150	0.823	0.753	0.795	0.416
Emergency Sign	184	120	0.824	0.908	0.925	0.502
Clear Pathway	184	93	0.564	0.417	0.562	0.274

The confusion matrix shown in Figure 5.1 provides a more detailed analysis of the model's performance for each class. As seen in the confusion matrix, the "emergency sign" class demonstrates a remarkable ability to be correctly identified by the model. However, the model faced challenges in accurately identifying instances of other classes. For example, the "exit point" class has relatively low accuracy, as it was misclassified in some cases. Similarly, while the model performed well in detecting the "fire extinguisher" class, it still had some false positives. The "clear pathway" class presented significant difficulties, with relatively low accuracy in detection.

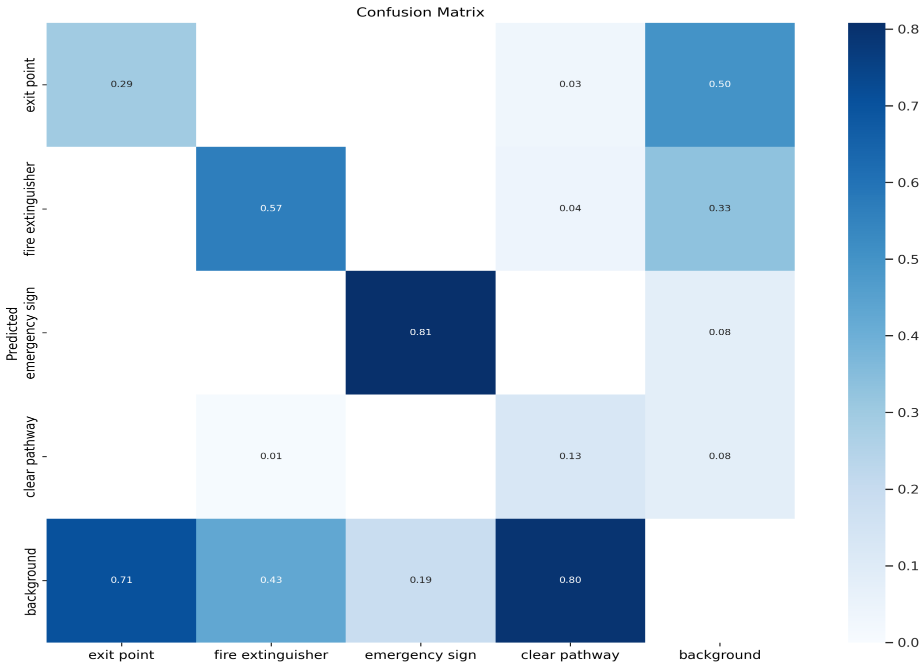


Figure 5.1: Confusion Matrix for Baseline Model

The confusion matrix offers valuable insights into the strengths and weaknesses of the model's predictions for different classes, helping identify areas for improvement. By addressing the challenges faced by the model, such as enhancing the detection of "clear pathway" instances, its overall performance can be further improved, achieving better object detection capabilities for all classes.

The F1 score is another essential metric that evaluates the model's accuracy and completeness of predictions. As seen in Figure 5.2, the trained model achieved an F1 score of 70% for all classes, at a confidence level of 0.129. The highest F1 score is obtained for the "emergency sign" class, showcasing impressive precision of 82.4% and recall of 90.8%. This indicates that the model's predictions for the "emergency sign" class are both accurate and comprehensive. However, the F1 score reveals that there is room for improvement for other classes, especially for the "clear pathway" class, where the F1 score is lower (56.4%), reflecting challenges in correctly identifying instances of this class.

The trained model shows promising results in detecting the "emergency sign" class. However, there is significant room for improvement for other object classes. The analysis from the confusion matrix and the F1 score highlights specific areas where

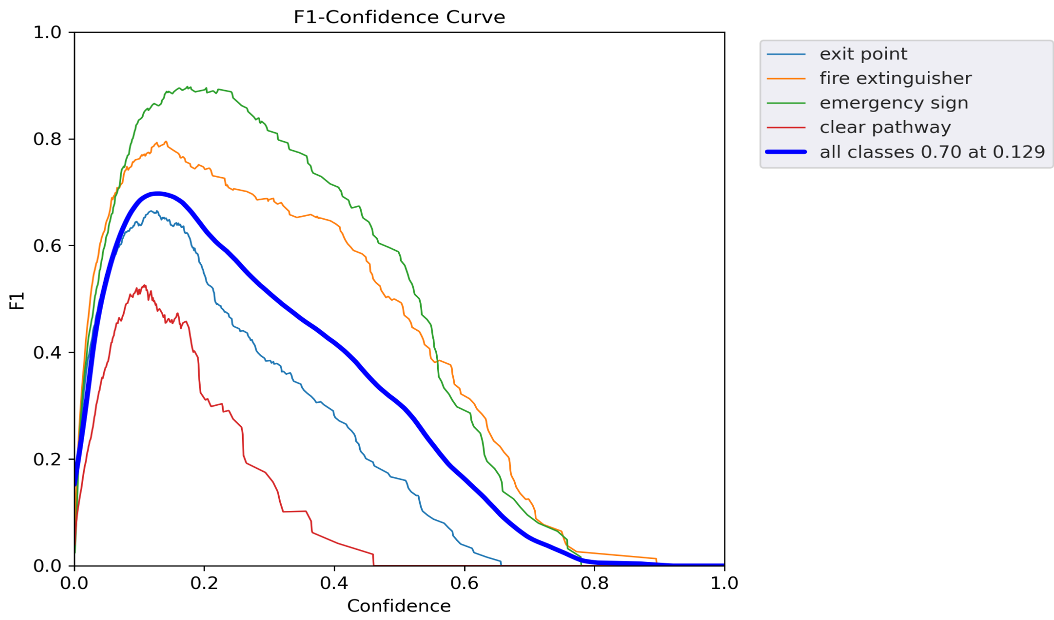


Figure 5.2: F1 Score for Baseline Model

the model's accuracy and completeness can be enhanced. By applying strategies like transfer learning and domain adaptation, the model's performance can be enhanced, surpassing the baseline and improving object detection across all classes.

5.2 Transfer Learning Model

This section investigates the integration of transfer learning to enhance the performance of the baseline thermal object detection model. The baseline model exhibited promising results, but its effectiveness was constrained by the scarcity of annotated thermal data. To overcome this challenge, transfer learning was employed by leveraging a pretrained RGB object detector, which had undergone extensive training on a diverse dataset containing rich visual information. By leveraging the pretrained RGB model's weights, the thermal object detection system was able to capitalize on the general features and patterns learned from RGB images, as illustrated by the precision and recall values in Table 5.2.

The transfer of knowledge from the RGB domain to the thermal domain significantly improves the model's ability to detect objects in thermal images, despite the differences in image characteristics. The transferred weights were fine-tuned to adapt them specifically to the thermal domain, refining the model's representations and enhancing its localization accuracy, as demonstrated in Table 5.2.

Table 5.2: YOLOv5s Baseline Model with Transfer Learning Evaluation Results

Class	Images	Instances	Precision	Recall	mAP50	mAP50-95
All	184	604	0.903	0.840	0.906	0.613
Exit Point	184	241	0.890	0.807	0.904	0.633
Fire Extinguisher	184	150	0.935	0.780	0.836	0.599
Emergency Sign	184	120	0.905	0.958	0.964	0.667
Clear Pathway	184	93	0.883	0.814	0.919	0.554

As a result, the enhanced model achieves an overall precision of 90.3% and a recall of 84.0%. The mAP50 score is enhanced to 90.6%, indicating the model's capability to accurately predict object locations with an average IoU of 0.613.

The F1 score shows significant improvement, reaching a value of 0.87. This shows a substantial advancement over the baseline model's performance, as seen in Figure 5.3. Classes that previously had lower accuracy rates, such as the "clear pathway" class, witness remarkable progress, showcasing the efficacy of transfer learning in enhancing object detection across all classes.

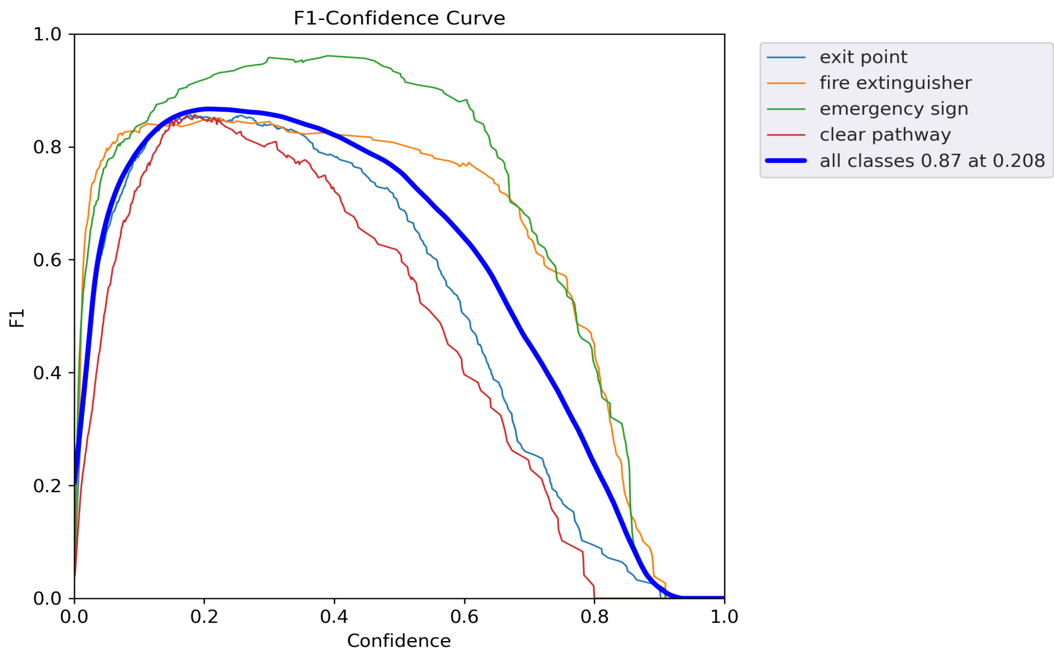


Figure 5.3: F1 Score for Transfer Learning

The model's performance can be analyzed in more detail using the confusion matrix shown in Figure 5.4. For instance, the "emergency sign" class demonstrates exceptional performance, with a high true positive count of 120 instances and very few misclassifications. On the other hand, the “exit point” class has a lower true positive count of 193 and a notable number of false negatives, indicating challenges in accurately identifying instances of this class. Similarly, while the model performs well in detecting the ”fire extinguisher” class, it still has some false positives.

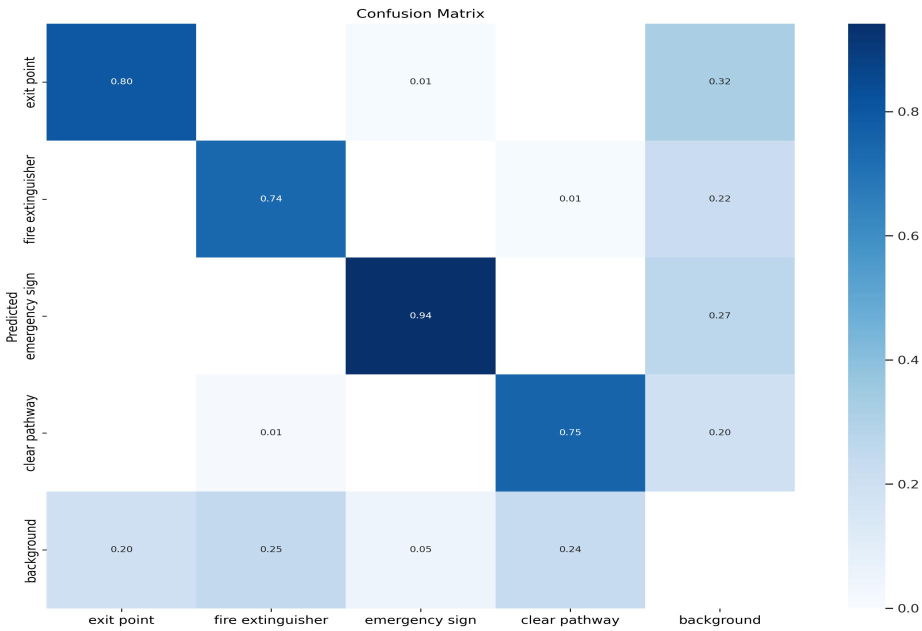


Figure 5.4: Confusion Matrix for Transfer Learning

Thus, the integration of transfer learning is shown to boost the capabilities of the thermal object detection system. Leveraging the knowledge acquired from the RGB domain for the thermal domain leads to substantial improvements in accuracy and detection capabilities.

5.3 Domain Adaptation with Transfer Learning Model

The results for the domain adaptation with transfer learning model are shown in Table 5.3. For the “exit point” class, the model exhibits enhanced performance with a precision of 92.6% and a recall of 88.2%. Similarly, the “fire extinguisher” class shows an increased precision of 97.2% and a recall of 78.7%, indicating more accurate and reliable detections.

The F1 score analysis as seen in Figure 5.5 reveals substantial improvements across all classes. The F1 score graph highlights that the domain adaptation technique achieves higher accuracy and completeness of predictions for all object classes, surpassing the

Table 5.3: YOLOv5s with Domain adpataion and Transfer Learning Evaluation Results

Class	Images	Instances	Precision	Recall	mAP50	mAP50-95
All	184	604	0.935	0.859	0.919	0.673
Exit Point	184	241	0.926	0.882	0.935	0.681
Fire Extinguisher	184	150	0.972	0.787	0.833	0.634
Emergency Sign	184	120	0.942	0.939	0.971	0.719
Clear Pathway	184	93	0.900	0.828	0.935	0.659

performance of both the baseline model and transfer learning. This demonstrates the model's ability to strike a balance between precision and recall for each class.

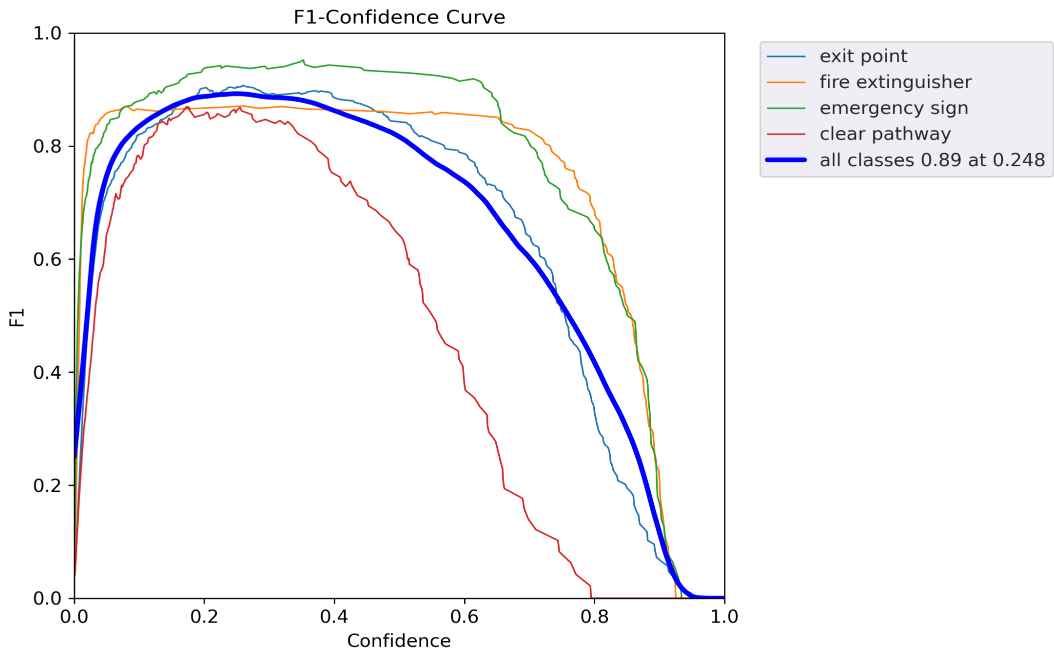


Figure 5.5: F1 Score for Domain Adaptation with Transfer Learning

The results from the confusion matrix in Figure 5.6 and the F1 score show the effectiveness of using domain adaptation in addressing challenges posed by data scarcity and domain differences. By generating extra training data that aligns with the target domain's characteristics, the model becomes more proficient at handling variations and uncertainties in the data. This synthetic data acts as a bridge, aiding the model in improving its overall performance and enhancing accuracy and robustness of thermal object detection. The ability to leverage knowledge from a different domain leads to more accurate and reliable predictions, further validating the benefits of domain adaptation in cross-domain computer vision tasks.

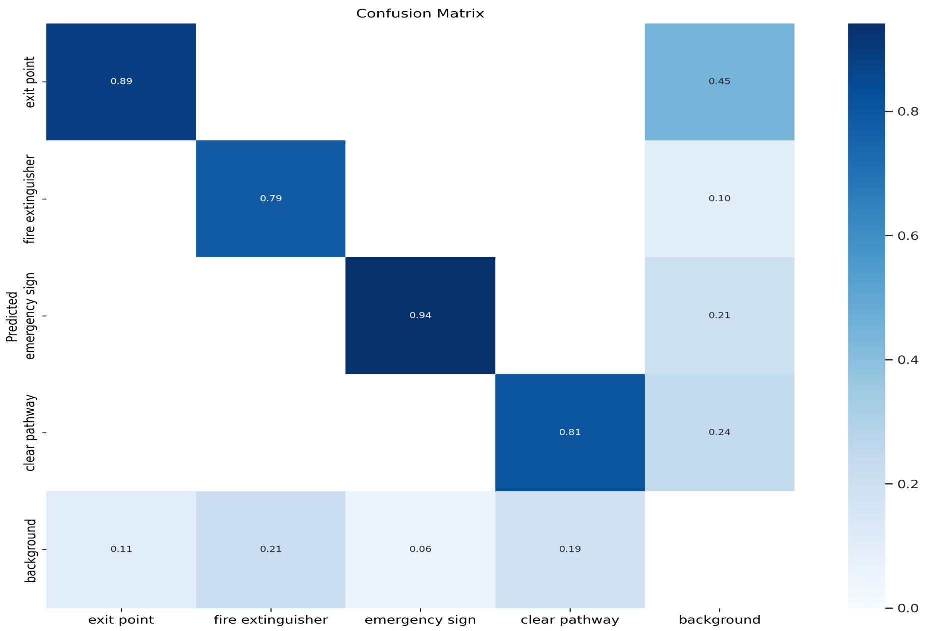


Figure 5.6: Confusion Matrix for Domain Adaptation with Transfer Learning

5.4 Comparative Visual Analysis

This section presents a comprehensive comparative analysis of three object detection models applied to a thermal imagery dataset. The primary objective is to assess the performance of each approach in accurately identifying and localizing objects of interest in thermal images. To facilitate a comprehensive evaluation, the analysis showcases ground truth annotations obtained through calibration from visible images to thermal images. The corresponding RGB images are displayed alongside the thermal results, providing visual context and aiding in the understanding of the real-world scenarios captured in the thermal data.

The following figures display examples of results obtained from each method, i.e., baseline model, transfer learning approach, and domain adaptation technique. This gives insights into the respective strengths and weaknesses of these models to identify the best approach for thermal object detection.

As seen in the first example of a fire extinguisher in Figure 5.7, both the baseline and transfer learning models show limitations in detecting the object, whereas the domain adaptation combined with transfer learning demonstrates superior performance in accurately identifying it. However, it is important to note that the ground truth for

Example 1 only comprises the fire extinguisher, and there were no additional points of interest present.

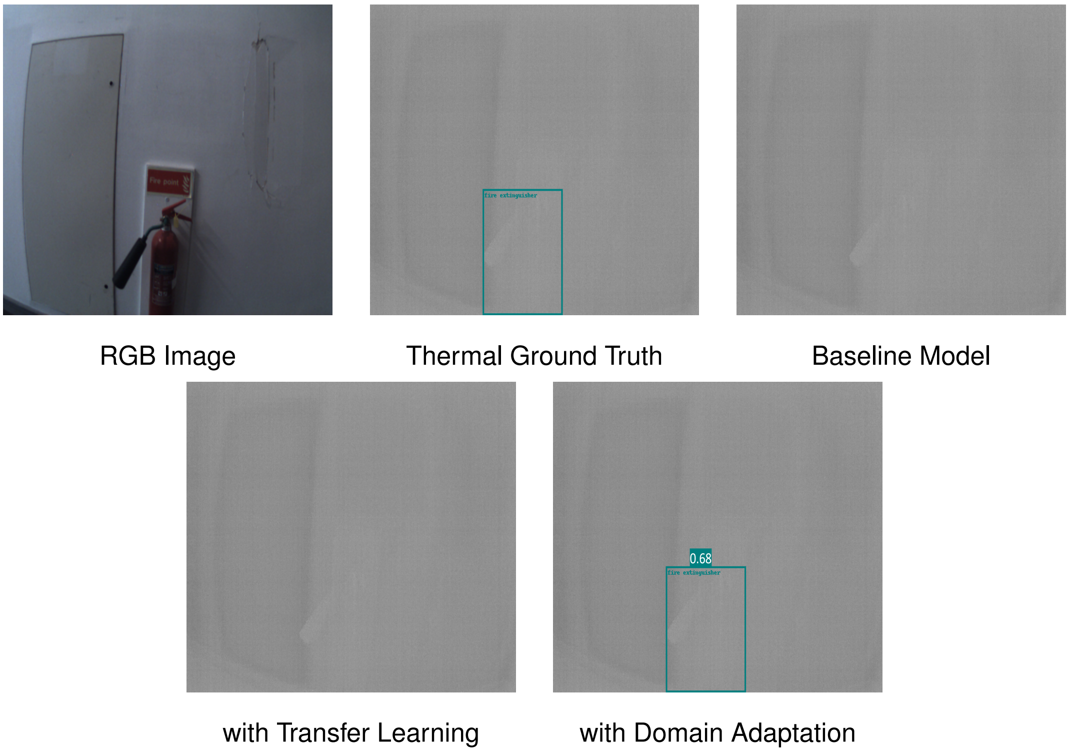


Figure 5.7: Comparison of different models - Example 1

In the more complex example in Figure 5.8, both the baseline and transfer learning models exhibit limited detection capabilities, where the baseline model does not detect any objects while the transfer learning model only detects the fire extinguisher while overlooking the exit points and clear pathway present in the ground truth. Conversely, the domain adaptation combined with transfer learning shows better accuracy, successfully detecting all objects in the scene, including the exit points and clear pathway.

The ground truth in the example in Figure 5.9 consists of a fire extinguisher, exit points, and a clear pathway. In this case, both the baseline and transfer learning models only identify the fire extinguisher, falling short of detecting the exit points and clear pathway. In contrast, the domain adaptation combined with transfer learning successfully detects all the ground truth objects in the scene.

In Figure 5.10, the ground truth consists of three exit points, two fire extinguishers, and one emergency sign. In this case, all the models detect one exit point, one fire extinguisher and the emergency sign. However, the transfer learning model shows better probability values for all three objects than the baseline model, and the domain adaptation with transfer learning model shows the highest probabilities.

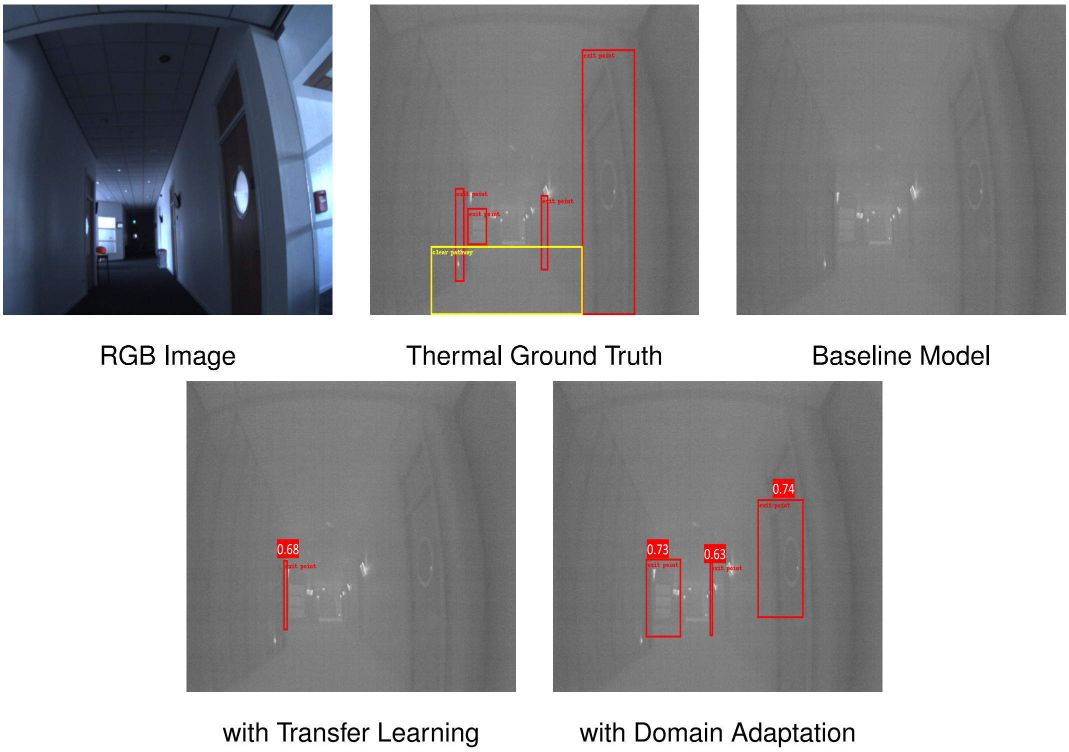


Figure 5.8: Comparison of different models - Example 2

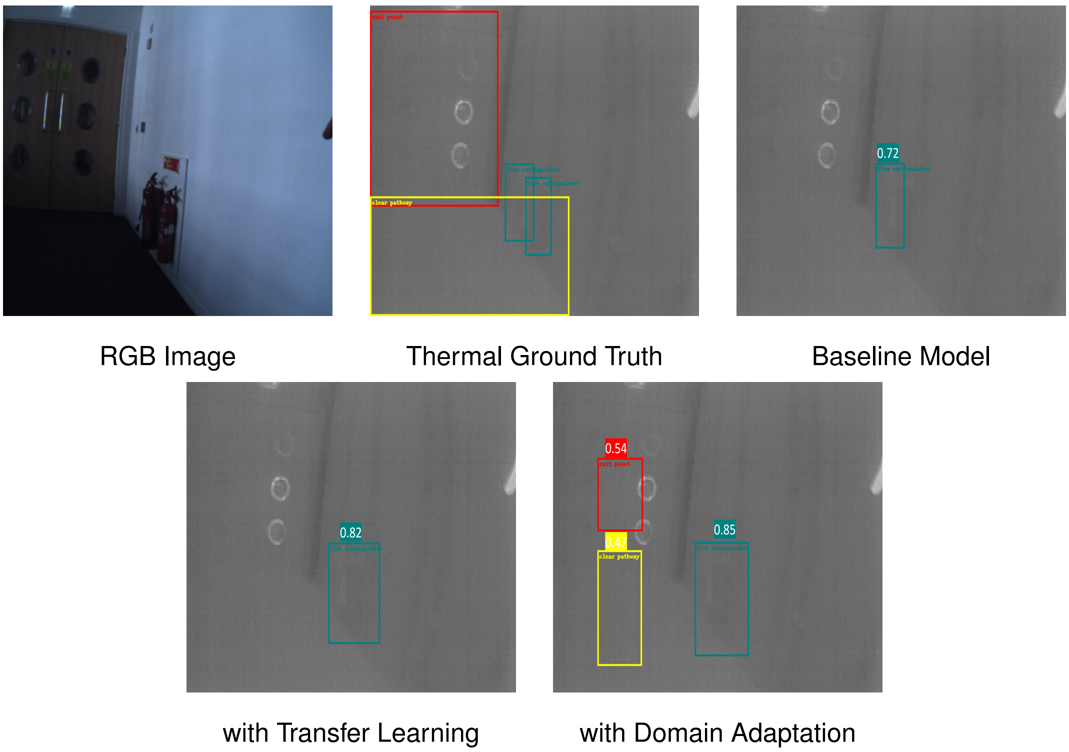


Figure 5.9: Comparison of different models - Example 3

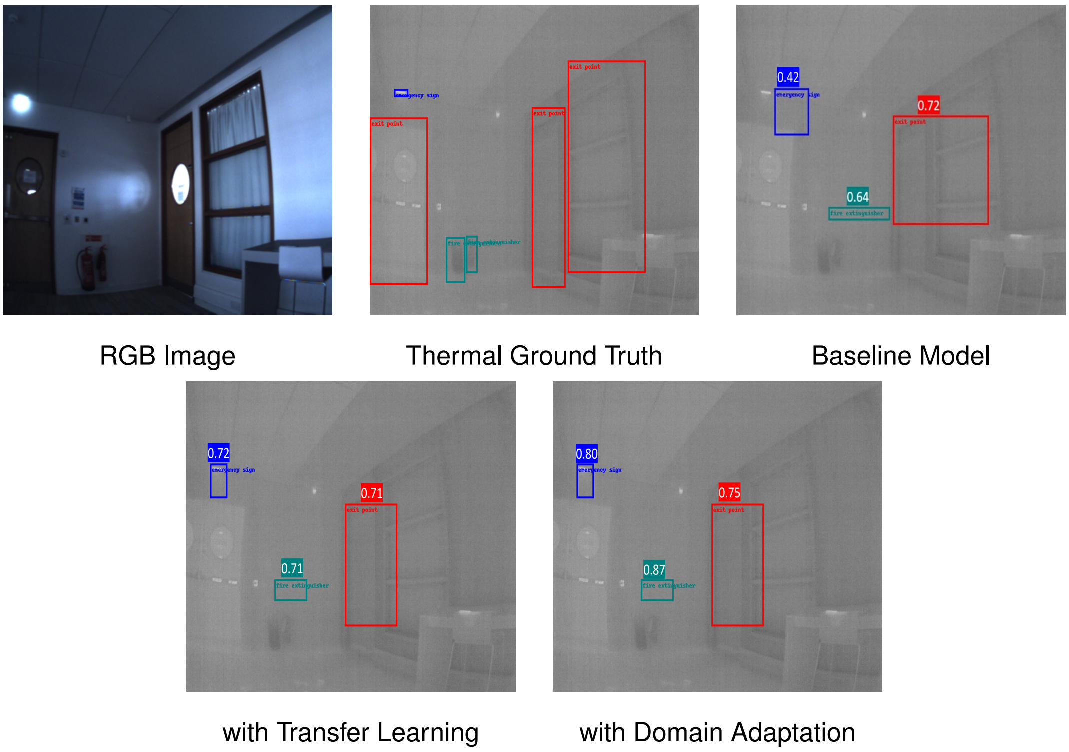


Figure 5.10: Comparison of different models - Example 4

Lastly, Figure 5.11 presents one instance of each class in the ground truth. The baseline model solely detects the emergency light, while the transfer learning model detects both the emergency light and the clear pathway with better probability. The domain adaptation combined with transfer learning detects all the objects in the scene with even higher accuracy.

In the comparative visual analysis, it is evident that the object detection results vary significantly among different models across several examples. It is evident that the transfer learning model and the domain adaptation method consistently outperform the baseline model, exhibiting significantly higher probability scores. This improvement in probability scores indicates a greater level of confidence and accuracy in the object detection results obtained through these advanced approaches. This enhancement is particularly valuable in practical applications where reliable and precise object detection is vital for making critical decisions and ensuring safety. The elevated probability scores of transfer learning and domain adaptation underscore their effectiveness and potential in addressing real-world object detection challenges.

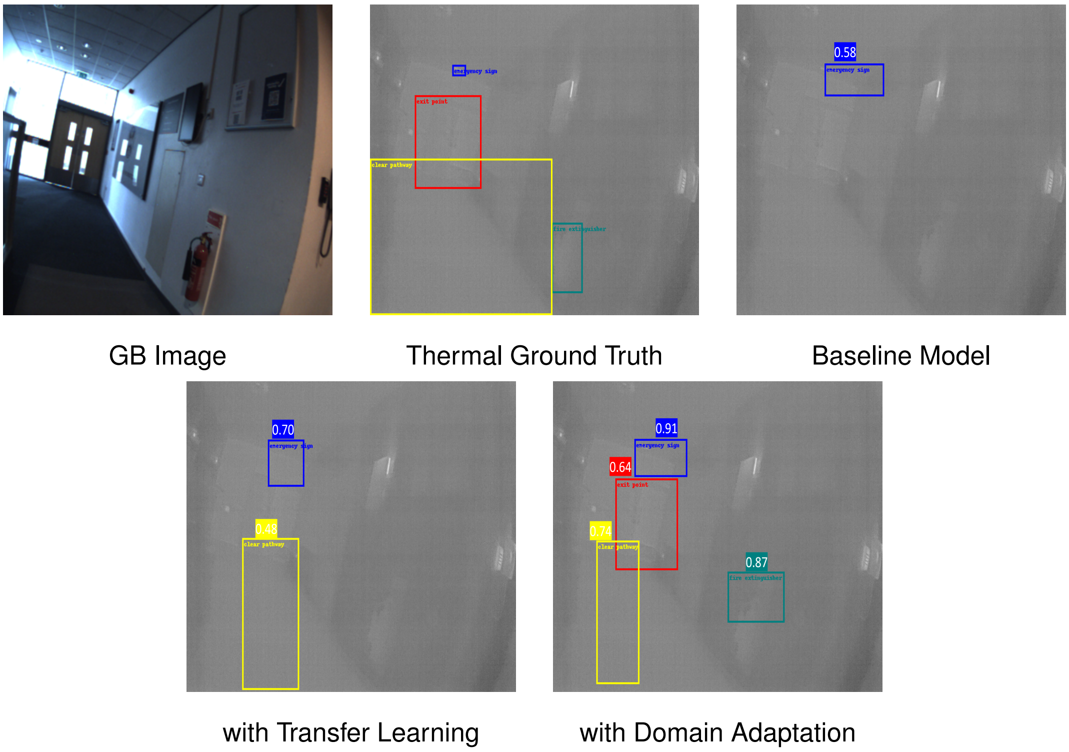


Figure 5.11: Comparison of different models - Example 5

5.5 Run-time Analysis

This section focuses on evaluating the run-time performance of the thermal object detection methods used. The aim here is to measure how fast and efficient each model is for processing thermal images for object detection. The evaluation considers the time required for the three key stages, i.e., pre-processing (image preparation), inference (model execution), and non-maximum suppression (result refinement).

Table 5.4 shows the processing times for the key detection stages in all the models. The pre-processing stage involves readying input thermal images via resizing, enhancing and formatting. The inference time is the time required for computation, i.e., for deciphering images and detecting objects. The non-maximum suppression (NMS) time is the time required for streamlining the results by retaining the most accurate bounding box. The total time helps compare the efficiency of each method.

The three thermal object detection models – baseline, transfer learning and domain adaptation – show comparable pre-processing and inference times, with slight variations in NMS times. The variations in total run-time among the three models arise from differences in their model complexities. The domain adaptation model, being more complex than the others, takes more time to complete the computations. While the

Table 5.4: Run-time Analysis of Thermal Object Detection Models

Model	**Pre-process (ms)**	**Inference (ms)**	**NMS (ms)**	**Total (ms)**
Baseline	0.2	4.0	0.4	4.6
Transfer Learning	0.2	4.0	0.5	4.7
Domain Adaptation	0.2	4.0	0.6	4.8

differences may appear small, they are significant in critical scenarios like real-time surveillance or autonomous driving, where even small time variations matter. The choice of model should therefore align with the needs of the application, prioritizing either speed or precision based on the context.

5.6 Discussion and Interpretation

The integration of domain adaptation with transfer learning is shown to be the best approach to enhance thermal object detection compared to the baseline model and transfer learning alone. Through domain adaptation, the model effectively leverages knowledge from the RGB domain and fine-tunes it for thermal object detection, addressing data scarcity challenges in the thermal domain. As a result, the domain adaptation technique exhibits improvements in accuracy and robustness, showcasing enhanced object detection capabilities across all classes. The higher values of evaluation metrics, such as precision, recall, and mAP at various IoU thresholds, along with consistently superior probability scores, establish domain adaptation as the most effective method among the evaluated approaches.

Chapter 6

Conclusions

6.1 Summary of Research

This research project centered around the development of an intelligent object detection algorithm tailored specifically for firefighters to enhance their safety and operational effectiveness during rescue missions in smoke-filled environments. The primary objective was to create an algorithm capable of accurately identifying crucial objects, such as doors, windows, tables, and fire hydrants, using thermal image data.

The motivation behind this research stemmed from the recognition of firefighting as a demanding and hazardous profession, where limited visibility posed significant challenges for firefighters in assessing the fire's magnitude, locating potential victims, and determining the best course of action. Traditional visible light cameras had limitations in smoke-filled environments, making them ineffective. Therefore, the integration of thermal imaging cameras, utilizing infrared technology, presented a promising solution to perceive surroundings even through smoke and flames, empowering firefighters with enhanced visibility.

To achieve the research objectives, the study focused on data collection and annotation for both thermal and RGB images. The study utilized RGB camera and a FLIR Boson 640 thermal camera for data collection. RGB images played a crucial role in calibrating and annotating the thermal data, and they were further utilized for domain adaptation. The dataset collected served as a foundation for the development and optimization of thermal detection algorithms with high accuracy and reliability.

The research employed a comprehensive approach, implementing three distinct methodologies: the baseline YOLO model, transfer learning, and domain adaptation with transfer learning.

The YOLO model was selected for its real-time object detection capabilities, a crucial feature for firefighting scenarios where timely decisions are essential. To improve the model's accuracy in low-visibility environments, transfer learning techniques were explored. A pre-trained model on RGB data served as the starting point, with its weights being used to initialize the training process for thermal data. This approach greatly enhanced the model's adaptability to thermal images, leading to more accurate object detection.

The domain adaptation with transfer learning proved to be the most effective methodology, significantly improving the model's performance in identifying crucial objects. This approach involved training the model with thermal data while leveraging knowledge from RGB data. Moreover, the incorporation of 10% synthetic data and the subsequent increase in data size during training contributed to the model's enhanced adaptability and accuracy. The domain adaptation with transfer learning produced the best results, exhibiting superior accuracy, F1 score, and probability scores for detected objects. The combined power of domain adaptation and transfer learning addressed data scarcity challenges in the thermal domain, leading to improvements in accuracy and robustness across all classes. The domain adaptation technique increased precision from 0.728 to 0.935, recall from 0.673 to 0.859, mAP50 from 0.743 to 0.919, and mAP50-95 from 0.393 to 0.673. This approach demonstrated improved performance in object detection, making it a promising and effective technique for enhancing object detection using thermal cameras.

Therefore, this thesis successfully implemented various methodologies for object detection using thermal cameras, with an emphasis on transfer learning and domain adaptation. The research demonstrated their significant potential in enhancing firefighter capabilities in low-visibility environments, ultimately leading to safer and more effective firefighting efforts. These findings contribute valuable insights to the broader field of object detection in challenging scenarios, with promising real-world applications such as search and rescue missions, industrial inspections, and automotive safety measures.

The research's impact lies in its potential to greatly improve firefighting operations through advanced object detection capabilities. The developed algorithm offers real-time feedback, assisting firefighters in challenging environments and enhancing rescue missions while protecting property. By utilizing cutting-edge technology, this research has the potential to make firefighting procedures safer, more efficient, and more effective for both firefighters and the public.

6.2 Future Scope

The scope for future research in this area is as follows.

1. **Multi-Sensor Fusion:** Future research can focus on multi-sensor fusion, integrating data from both thermal and RGB sensors, to enhance the thermal object detection model's robustness and adaptability. Existing work has shown the benefits of combining data from these modalities to enhance tracking accuracy and robustness [63]. Thermal cameras are excellent at detecting heat signatures in low-light conditions, while RGB sensors provide rich color and texture information. By combining these modalities, the model can overcome individual sensor limitations and make more informed and accurate predictions. This fusion will enable the system to perform effectively in challenging environments, such as adverse weather conditions, where thermal data alone might be insufficient. Multi-sensor fusion holds significant potential for advancing thermal object detection technology and its real-world applications.

2. **Real-Time Performance Optimization:** Real-time object detection is crucial in firefighting scenarios to provide immediate and actionable information to firefighters. Future research can focus on optimizing the thermal object detection model for faster inference and improved real-time performance. Techniques like model quantization, network pruning, and hardware acceleration can be explored to reduce the model's computational complexity and memory footprint without compromising detection accuracy.

Bibliography

[1] Wikipedia contributors. Firefighter — Wikipedia, the free encyclopedia, 2023. [Online; accessed 9-April-2023].

[2] Cody E Morris and Harish Chander. The impact of firefighter physical fitness on job performance: A review of the factors that influence fire suppression safety and success. *Safety*, 4(4):60, 2018.

[3] Ahad Heydari, Abbas Ostadtaghizadeh, Ali Ardalan, Abbas Ebadi, Iraj Mohammadfam, and Davoud Khorasani-Zavareh. Exploring the criteria and factors affecting firefighters' resilience: A qualitative study. *Chinese Journal of Traumatology*, 25(02):107–114, 2022.

[4] Hao Wu, Deyang Wu, and Jinsong Zhao. An intelligent fire detection approach through cameras based on computer vision methods. *Process Safety and Environmental Protection*, 127:245–256, 2019.

[5] DM Bulanon, TF Burks, and V Alchanatis. Image fusion of visible and thermal images for fruit detection. *Biosystems engineering*, 103(1):12–22, 2009.

[6] Magdalena Mazur-Milecka, Natalia Głowacka, Mariusz Kaczmarek, Adam Bujnowski, Miłosz Kaszyński, and Jacek Rumiński. Smart city and fire detection using thermal imaging. In *2021 14th International Conference on Human System Interaction (HSI)*, pages 1–7, 2021.

[7] Nermin K Negied, Elsayed E Hemayed, and Magda B Fayek. Pedestrians' detection in thermal bands–critical survey. *Journal of Electrical Systems and Information Technology*, 2(2):141–148, 2015.

[8] Mohammadreza Shokouhi, Khadijeh Nasiriani, Hamidreza Khankeh, Hosein Fallahzadeh, and Davoud Khorasani-Zavareh. Exploring barriers and challenges

in protecting residential fire-related injuries: a qualitative study. *Journal of Injury and Violence Research*, 11(1):81–92, 2019.

[9] E&T editorial staff. Smart helmets could help firefighters locate blaze victims, September 2022.

[10] Wikipedia contributors. Object detection — Wikipedia, the free encyclopedia, 2023. [Online; accessed 23-June-2023].

[11] Xiongwei Wu, Doyen Sahoo, and Steven C. H. Hoi. Recent advances in deep learning for object detection. *CoRR*, abs/1908.03673, 2019.

[12] Rohan Ippalapally, Sri Harsha Mudumba, Meghana Adkay, and Nandi Vardhan H. R. Object detection using thermal imaging. In *2020 IEEE 17th India Council International Conference (INDICON)*, pages 1–6, 2020.

[13] Mate Krišto, Marina Ivasic-Kos, and Miran Pobar. Thermal object detection in difficult weather conditions using yolo. *IEEE Access*, 8:125459–125476, 2020.

[14] Marina Ivašić-Kos, Mate Krišto, and Miran Pobar. Human detection in thermal imaging using YOLO. In *ICCTA '19: Proceedings of the 2019 5th International Conference on Computer and Technology Applications*, page 20–24, New York, NY, USA, 2019. Association for Computing Machinery.

[15] Piotr Rudol and Patrick Doherty. Human body detection and geolocalization for uav search and rescue missions using color and thermal imagery. In *2008 IEEE Aerospace Conference*, pages 1–8, 2008.

[16] Wai Kit Wong, Poi Ngee Tan, Chu Kiong Loo, and Way Soong Lim. An effective surveillance system using thermal camera. In *2009 International Conference on Signal Acquisition and Processing*, pages 13–17, 2009.

[17] Tim LeBeau. Thermal imaging for safer autonomous vehicles. In *Infrared Technology and Applications XLV*, volume 11002, pages 349–353. SPIE, 2019.

[18] Larson Electronics. Smoke penetrating technology: What light works best?, Nov 2017.

[19] Rikke Gade and Thomas B Moeslund. Thermal cameras and applications: a survey. *Machine vision and applications*, 25:245–262, 2014.

[20] Jong-Hwan Kim, Seongsik Jo, and Brian Y Lattimer. Feature selection for intelligent firefighting robot classification of fire, smoke, and thermal reflections using thermal infrared images. *Journal of Sensors*, 2016, 2016.

[21] A. N. Wilson, Khushi Anil Gupta, Balu Harshavardan Koduru, Abhinav Kumar, Ajit Jha, and Linga Reddy Cenkeramaddi. Recent advances in thermal imaging and its applications using machine learning: A review. *IEEE Sensors Journal*, 23(4):3395–3407, 2023.

[22] Xuerui Dai, Xue Yuan, and Xueye Wei. Tirnet: Object detection in thermal infrared images for autonomous driving. *Applied Intelligence*, 51:1244–1261, 2021.

[23] My Kieu, Andrew D Bagdanov, and Marco Bertini. Bottom-up and layerwise domain adaptation for pedestrian detection in thermal images. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 17(1):1–19, 2021.

[24] Chaitanya Devaguptapu, Ninad Akolekar, Manuj M Sharma, and Vineeth N Balasubramanian. Borrow from anywhere: Pseudo multi-modal object detection in thermal imagery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 0–0, 2019.

[25] My Kieu, Lorenzo Berlincioni, Leonardo Galteri, Marco Bertini, Andrew D Bagdanov, and Alberto Del Bimbo. Robust pedestrian detection in thermal imagery using synthesized images. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 8804–8811. IEEE, 2021.

[26] Pei-Fen Tsai, Chia-Hung Liao, and Shyan-Ming Yuan. Using deep learning with thermal imaging for human detection in heavy smoke scenarios. *Sensors*, 22(14):5351, 2022.

[27] Michael Ulrich, Thomas Hess, Sherif Abdulatif, and Bin Yang. Person recognition based on micro-doppler and thermal infrared camera fusion for firefighting. In *2018 21st International Conference on Information Fusion (FUSION)*, pages 919–926, 2018.

[28] Pei-Fen Tsai, Chia-Hung Liao, and Shyan-Ming Yuan. Using deep learning with thermal imaging for human detection in heavy smoke scenarios. *Sensors*, 22(14):5351, 2022.

[29] Manish Bhattarai and Manel Martinez-Ramon. A deep learning framework for detection of targets in thermal images to improve firefighting. *IEEE Access*, 8:88308–88321, 2020.

[30] Paul Viola and Michael J Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57:137–154, 2004.

[31] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893. IEEE, 2005.

[32] Pedro F Felzenszwalb, Ross B Girshick, and David McAllester. Cascade object detection with deformable part models. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2241–2248. IEEE, 2010.

[33] Chinthakindi Balaram Murthy, Mohammad Farukh Hashmi, Neeraj Dhanraj Bokde, and Zong Woo Geem. Investigations of object detection in images/videos using various deep learning techniques and embedded platforms—a comprehensive review. *Applied sciences*, 10(9):3280, 2020.

[34] Yi-Qing Wang. An analysis of the Viola-Jones face detection algorithm. *Image Processing On Line*, 4:128–148, 2014.

[35] Feng Han, Ying Shan, Ryan Cekander, Harpreet S Sawhney, and Rakesh Kumar. A two-stage approach to people and vehicle detection with hog-based SVM. In *Performance Metrics for Intelligent Systems 2006 Workshop*, pages 133–140, 2006.

[36] Pedro F Felzenszwalb, Ross B Girshick, David McAllester, and Deva Ramanan. Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1627–1645, 2009.

[37] Syed Sahil Abbas Zaidi, Mohammad Samar Ansari, Asra Aslam, Nadia Kanwal, Mamoona Asghar, and Brian Lee. A survey of modern deep learning based object detection models. *Digital Signal Processing*, page 103514, 2022.

[38] Licheng Jiao, Fan Zhang, Fang Liu, Shuyuan Yang, Lingling Li, Zhixi Feng, and Rong Qu. A survey of deep learning-based object detection. *IEEE access*, 7:128837–128868, 2019.

[39] MF Ansari and KA Lodi. A survey of recent trends in two-stage object detection methods. In *Renewable Power for Sustainable Growth: Proceedings of International Conference on Renewal Power (ICRP 2020)*, pages 669–677. Springer, 2021.

[40] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.

[41] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28, 2015.

[42] Ross Girshick. Fast R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1440–1448, 2015.

[43] Ramesh Simhambhatla, Kevin Okiah, Shravan Kuchkula, and Robert Slater. Self-driving cars: Evaluation of deep learning techniques for object detection in different driving conditions. *SMU Data Science Review*, 2(1):23, 2019.

[44] Sudan Jha, Changho Seo, Eunmok Yang, and Gyanendra Prasad Joshi. Real time object detection and trackingsystem for video surveillance system. *Multimedia Tools and Applications*, 80:3981–3996, 2021.

[45] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

[46] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, pages 21–37. Springer, 2016.

[47] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[48] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

[49] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020.

[50] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.

[51] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[52] Peng Liu, Fuyu Li, Shanshan Yuan, and Wanyi Li. Unsupervised image-generation enhanced adaptation for object detection in thermal images. *Mobile Information Systems*, 2021:1–6, 2021.

[53] Tiantong Guo, Cong Phuoc Huynh, and Mashhour Solh. Domain-adaptive pedestrian detection in thermal images. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 1660–1664. IEEE, 2019.

[54] Dong-Guw Lee, Myung-Hwan Jeon, Younggun Cho, and Ayoung Kim. Edge-guided multi-domain RGB-to-TIR image translation for training vision tasks with challenging labels. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 8291–8298. IEEE, 2023.

[55] My Kieu, Lorenzo Berlincioni, Leonardo Galteri, Marco Bertini, Andrew D Bagdanov, and Alberto Del Bimbo. Robust pedestrian detection in thermal imagery using synthesized images. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 8804–8811. IEEE, 2021.

[56] Wikipedia contributors. Transfer learning — Wikipedia, the free encyclopedia, 2023. [Online; accessed 7-July-2023].

[57] Aprinaldi Jasa Mantau, Irawan Widi Widayat, Jenq-Shiou Leu, and Mario Köppen. A human-detection method based on YOLOv5 and transfer learning using thermal image data from UAV perspective for surveillance system. *Drones*, 6(10):290, 2022.

[58] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects

in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, pages 740–755. Springer, 2014.

[59] Li Weipeng, Yang Xiaogang, Li Chuanxiang, Lu Ruitao, and Huang Pan. An improved semi-supervised transfer learning method for infrared object detection neural network. *Infrared and Laser Engineering*, 50(3):20200511–1, 2021.

[60] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. Segment anything. *arXiv:2304.02643*, 2023.

[61] Léon Bottou. Stochastic gradient descent tricks. In *Neural Networks: Tricks of the Trade: Second Edition*, pages 421–436. Springer, 2012.

[62] Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1):43–76, 2020.

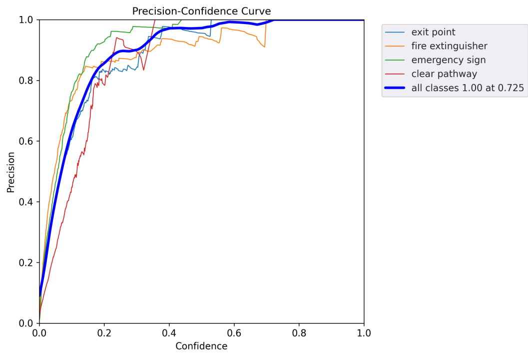
[63] Supriya Mangale, Meghna Raje, Nikhita More, Shivangi Sinha, and Madhuri Khambete. Object tracking by adaptive fusion of thermal infrared and visible spectrum imaging using particle filter and structural similarity index metric. *Journal of Applied Remote Sensing*, 15(4):044517–044517, 2021.

Appendix A

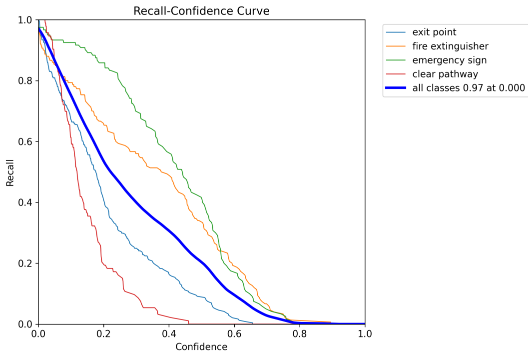
Additional Insights from Experiments

The graphs below illustrate the performance curves for the developed object detection algorithm under different methodologies: baseline, transfer learning, and domain adaptation with transfer learning. The P curve (Precision-Confidence Curve) showcases the algorithm's precision at varying confidence thresholds, indicating how confident the model is in its predictions. The R curve (Recall-Confidence Curve) depicts the algorithm's recall, indicating its ability to identify true positive instances while varying confidence thresholds. The PR curve (Precision-Recall Curve) visualizes the trade-off between precision and recall, revealing the algorithm's overall performance in identifying objects accurately. These curves offer valuable insights into the algorithms' performance across different scenarios and provide critical information for optimizing its precision and recall for real-world applications.

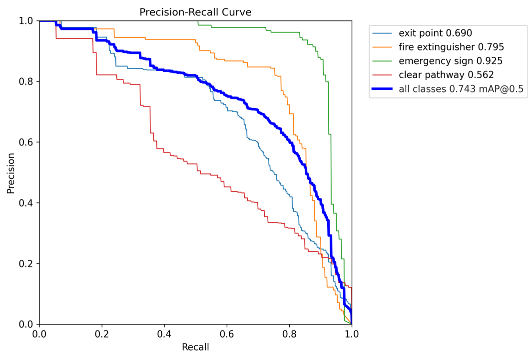
A.1 Baseline Performance Curves



(a) P Curve (Precision-Confidence Curve)



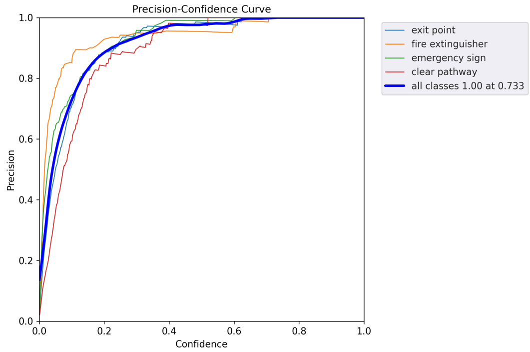
(b) R Curve (Recall-Confidence Curve)



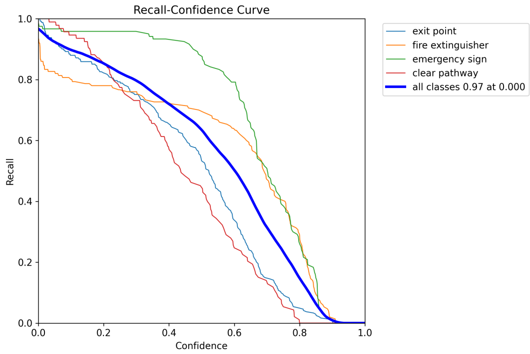
(c) PR Curve (Precision-Recall Curve)

Figure A.1: Baseline Performance Curves

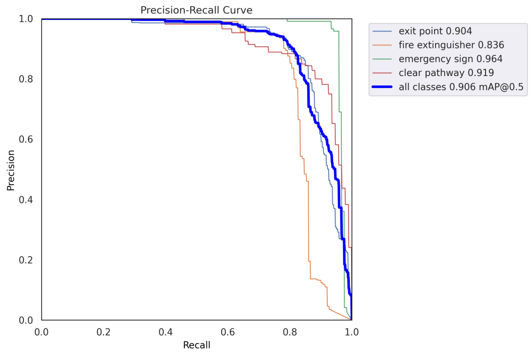
A.2 Transfer Learning Performance Curves



(a) P Curve (Precision-Confidence Curve)



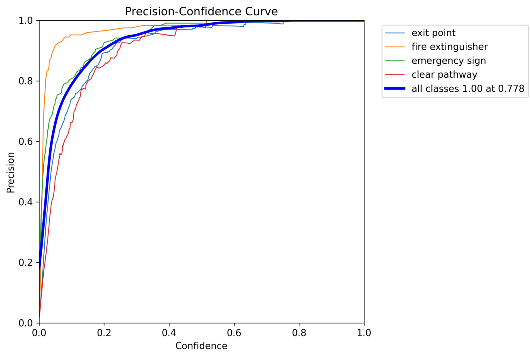
(b) R Curve (Recall-Confidence Curve)



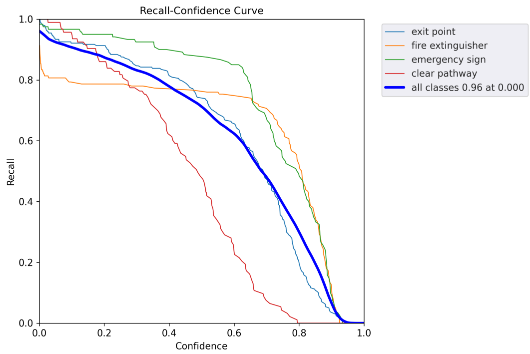
(c) PR Curve (Precision-Recall Curve)

Figure A.2: Transfer Learning Performance Curves

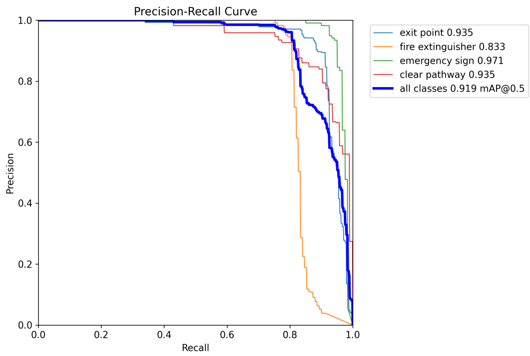
A.3 Domain Adaptation with Transfer Learning Performance Curves



(a) P Curve (Precision-Confidence Curve)



(b) R Curve (Recall-Confidence Curve)



(c) PR Curve (Precision-Recall Curve)

Figure A.3: Domain Adaptation with Transfer Learning Performance Curves

Appendix B

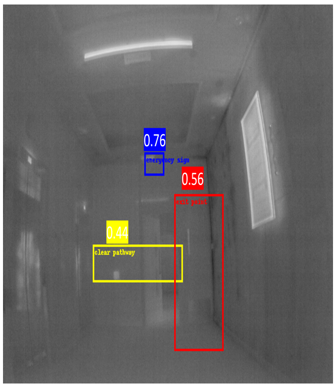
Additional Evaluation Results

This appendix contains additional experimental result outputs.

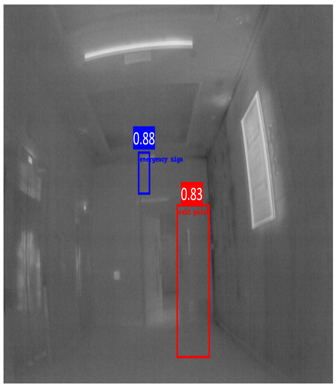
RGB Image

Thermal Ground Truth

Baseline Model



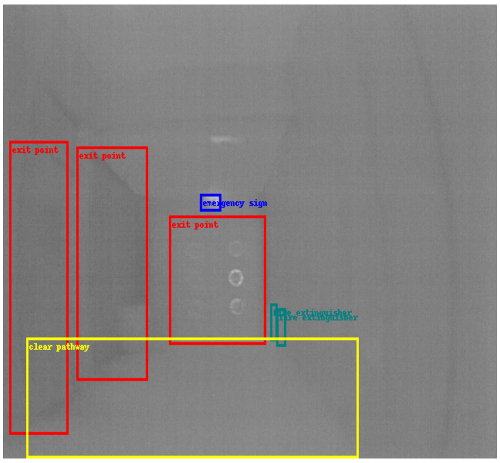
with Transfer Learning



with Domain Adaptation

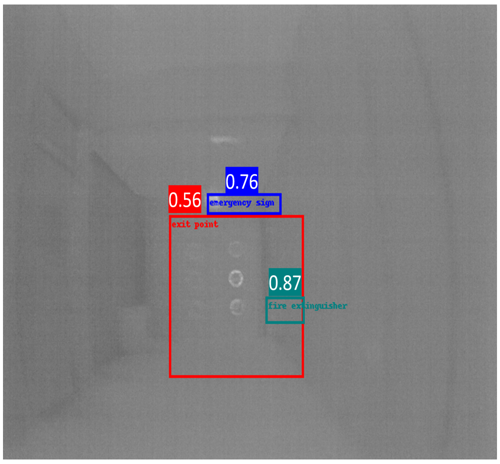
Figure B.1: Comparison of different methods - Example 1

RGB Image

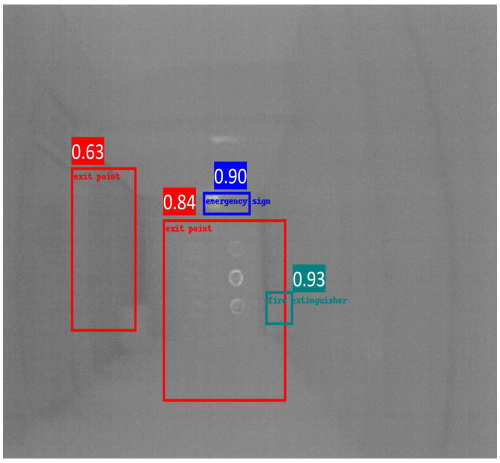


Thermal Ground Truth

Baseline Model



with Transfer Learning

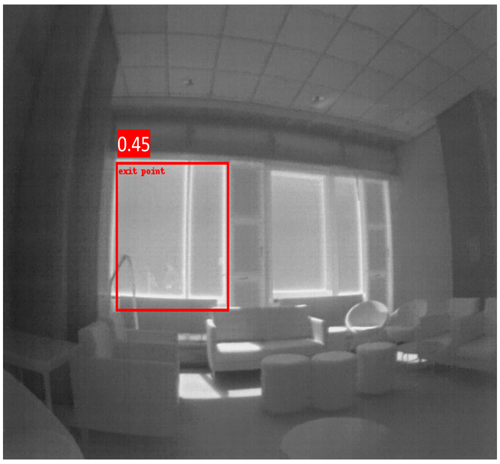


with Domain Adaptation

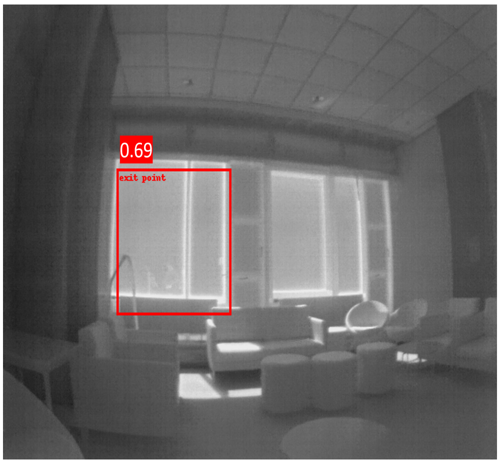
Figure B.2: Comparison of different methods - Example 2

RGB Image

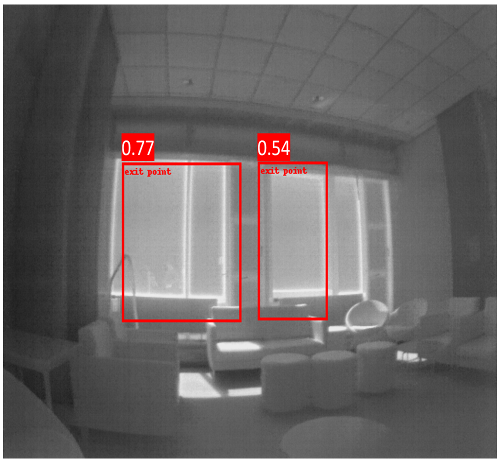
Thermal Ground Truth



Baseline Model



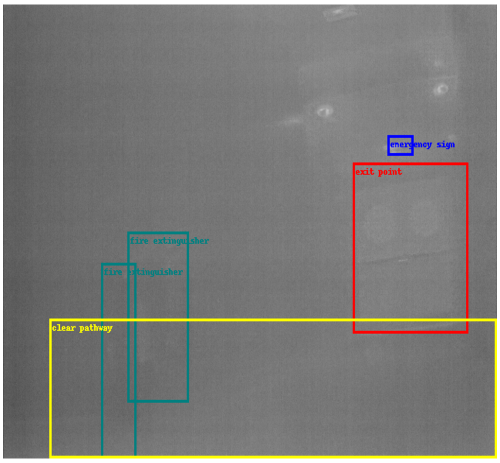
with Transfer Learning



with Domain Adaptation

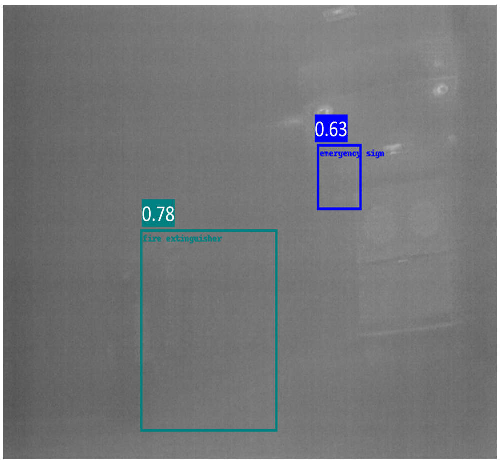
Figure B.3: Comparison of different methods - Example 3

RGB Image

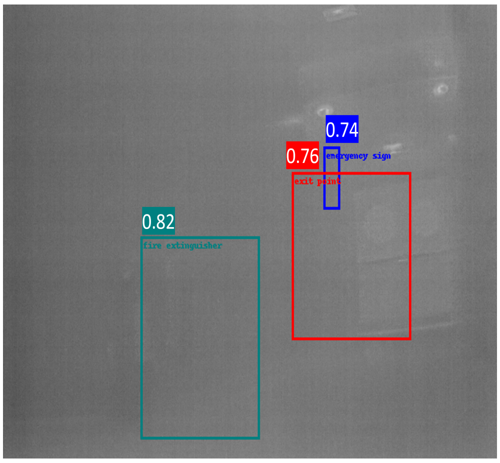


Thermal Ground Truth

Baseline Model



with Transfer Learning



with Domain Adaptation

Figure B.4: Comparison of different methods - Example 4