Thermal Imaging with AI for Real-Time Human Detection in Fire Emergencies

Savitha P., Dusyanth N. P., Hemanath P. and Mohammed Naveeth P.
Department of Computer Science and Engineering, Nandha Engineering College, Erode, Tamil Nadu, India

Keywords: Infrared Thermal Camera, Convolutional Neural Network Evacuation During a Fire Foggy Fire Scene Firefighter Safety Human Detection.

Abstract: Forest fires are a significant ecological risk because their smoke is an initial warning of risk but initially smoke appears in small hardly noticeable quantities it is hard to identify smoke with pixel-based parameters because the environment is constantly evolving and the dispersion characteristics of smoke are unpredictable this paper presents an innovative framework that minimizes the susceptibility of many yolo detection methods to minor visual interference it further analyzes and contrasts the detection capability and processing time of various versions of yolo ie yolov3 and yolov5the present research scrutinizes a recent human detection scheme for emergency evacuation in smoke-producing low-visibility fire environment utilizing a thermal imaging camera tic that meets national fire protection associations 1801 standards we take low-wavelength infrared LWIR pictures and utilize the yolov4 deep-learning algorithm to recognize objects in real-time while running at 301 frames per second fps the model trained on an Nvidia GeForce 2070 GPU can detect people in smoke environments with more than 95 accuracy this quick detection capability allows for immediate data transmission to command centers to facilitate timely rescue and protection of firemen prior to reaching risky situations.

1 INTRODUCTION

Smoke is a major workplace safety hazard that must be prevented and promptly addressed in the event of an emergency. Worker safety, firefighter support, and preventing property damage all depend on a well-planned fire evacuation strategy. During an emergency evacuation, we will use a thermal camera and the YOLOv4 model to locate people in a smoke-filled fire scene in accordance with NFPA 1801 rules to increase fire detection and safety procedures. At a fire scene, the two greatest dangers are heat and smoke, with smoke being the most dangerous. It causes zero vision during firefighter rescues or evacuations and poses a serious risk of suffocation, which can be lethal. Humans can only move at 0.5 m/s, yet smoke spreads quickly during a fire at 3–5 m/s. When smoke spreads, fire follows, so finding people quickly and getting them to safety is essential for life. LADAR, 3D Laser Scanning, Ultrasonic Sensors, and infrared thermal cameras are common solutions for the challenging problem of detecting people in dense smoke. To provide clear visibility in a smoke-filled fire scenario, we suggest combining the YOLOv4 with an infrared thermal camera that complies with NFPA 1801 criteria. A single GPU powers an AI-powered human detection system that uses a Convolutional Neural Network (CNN) to recognize individuals in a smoke-filled environment. It helps with emergency evacuations by transmitting real-time updates to a central control room. Similar studies on object detection models (Section 2), dataset and pre-processing information (Section 3), and the suggested methodology using the YOLOv4 model (Section 4) are all included in this study. concludes with a discussion of human detection after presenting experimental data and performance measures (Section 5).

2 RELATED WORKS

Sai, Liao, and Yuan suggest employing a Deep Learning model and a thermal imaging camera tic in the evacuation of people from emergencies involving smoke-filled fire low-visibility their method identifies persons in such conditions with over 95 accuracy using the yolov4 model and images from

P., S., P., D. N., P., H. and P., M. N.
Thermal Imaging with AI for Real-Time Human Detection in Fire Emergencies.
DOI: 10.5220/0013887200004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 2*, pages 608-613
ISBN: 978-989-758-777-1

low-wavelength infrared LWIR. Do Truong and le suggest a fusion technique utilizing thermal and infrared imaging to recognize people in real time in fire conditions where smoke obscures visibility their technique achieves a mean average precision map of 95 by fusing information from multiple cameras and processing it with a light-deep neural network. Ai-powered early fire detection gadget that recognizes early smoke signs of wildfires through real-time picture processing and high-definition panoramic cameras wildfire discovery and response time are dramatically minimized by solar-powered technology which is also linked to emergency services. The solar-powered aid-fire system developed by zhang wang zeng wu Huang and Xiao employs cloud servers IOT sensors and ai engines to instantly recognize complex building fire data2 the system was tested in a full-scale fire test room and indicated fire growth and spread correctly with a relative inaccuracy of less than 15 10 uav technology is transforming emergency response and fire protection because it enables 3d mapping in real time and hotspot identification through the integration of thermal images and ai-based algorithms drones enhance situational awareness and allow quicker and more accurate fire detectio. These drones employ thermal images and semantic segmentation to efficiently track and monitor forest fires using deep models such as mask r-cnn and yolo versions this technique significantly boosts the accuracy of detection and continuous monitoring of fires ensuring a more proactive and efficient emergency response.

3 EXPERIMENT METHODOLOGY

3.1 Data Collection

The thermal imaging apparatus was selected because it complies with NFPA 1801 standards, specifically concerning temperature sensitivity, resolution, and spectral range. Additionally, the device features an uncooled microbolometer, which enhances imaging capabilities and makes it suitable for use in fire emergencies. To generate more training data, we use this thermal imaging camera (TIC) to capture images of various body poses, such as squatting, lying down, and falling, from every aspect (360°). This makes it easier to imagine realistic situations in which people might want assistance during a smoke-filled escape. The human body temperature in these images is equal to a Gray Level (GL) of 105, per previous studies on Thermal making-based human detection in fire scenarios.

3.2 Training Using Thermal Datasets

Deep learning is highly dependent on the quantity and quality diversity velocity and authenticity of big data its limitation is that we do not have large publicly available thermal image data sets we compensate with other data similar to the thermal images we recorded and publicly available data sets like indoor people dataset in Kaggle and pedestrian thermal images with more quantity and diversity of data the method enhances the precision and overall model performance in thermal image analysis. Figure 1 shows the Thermal Imaging Dataset.

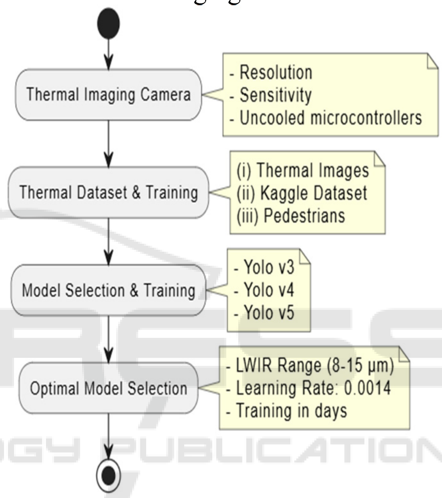

Figure 1: Thermal Imaging Dataset.

3.3 YOLOv3

Joseph Redmon and Ali Farhadi proposed the third generation yolo you only look once network in 2018 the models inference time is 22 milliseconds and the mean average accuracy is 282 dimension clusters are employed to address the ground-truth bounding box prediction problem in anchor boxes features of yolov3 are extremely minimalistic however lower confidence level because its classification layer uses logistic regression instead of softmax the more confident the object is the more likely it is to appear in a given grid cell darknet-53 as the base model utilized a denser convolution backbone than yolov4 utilized neck layer information through a cross-stage partial csp network since a detection model yolov3 is still an excellent one for everyone but the computationally most demanding applications Bhattarai and Martinez-Ramon used proficiently deep models such as yolov4 to spot people in thermograms with enhanced intersection-over-union IOU accuracy Redmon was able to demonstrate in spite of its deficits that yolov3 is superior and improved compared to ssd and even its earlier form yolov2 light yolo variants are also investigated to optimize object detection methods to utilize for real-time fire prevention and emergency evacuation systems along with the above.

3.4 YOLOv4 Model

Even though thermal images tend to have low resolution and small objects will only occupy approximately 50 pixels studies demonstrate that yolo darknet 20 is effective in detecting regular and small far-range thermal objects the deep architecture of darknet renders thermal imaging more relevant in a broad range of conditions which qualifies it as an appropriate choice for fire emergency response we utilize the yolov4 object detector which is a cnn-based object detector for our research as the most advanced real-time object detection model available in the market at the time of writing in 2020 it is suitable for detecting humans in risky environments such as smoke-filled fire environments a single Nvidia Geforce 2080gpu is utilized to train the model which is employed as a real-time head neck and backbone detector the backbone network is responsible for feature extraction on various scales and the neck equipped with spatial pyramid pooling spp serves to reduce model parameters and increase training efficiency the head which is founded on yolov3 is meant for object localization and one-stage classification.

3.5 YOLOv5

yolov5 was developed to allow for quicker real-time processing and was initially slated for release in May 2020 it is intended to identify one or more objects within an image immediately which makes it extremely efficient for real-time use the model consists of three primary parts they are the backbone neck and head. which operate in conjunction with each other to enhance efficiency and accuracy the spine is the beginning tasked with detecting important features in an image accurately irrespective of an object's location or size yolov5 uses a path aggregation network planet which aids in feature combination and enhances object detection yolo layer does detection and classification giving each object features such as its coordinates x-axis, y-axis width height bounding box and an confidence score signifying the probability of its existence the model utilizes the intersection over union IOU method to limit errors and prevent duplication of detection choosing the most precise bounding box improved by these advancements yolov5 is still a light yet very credible real-time object detection tool.

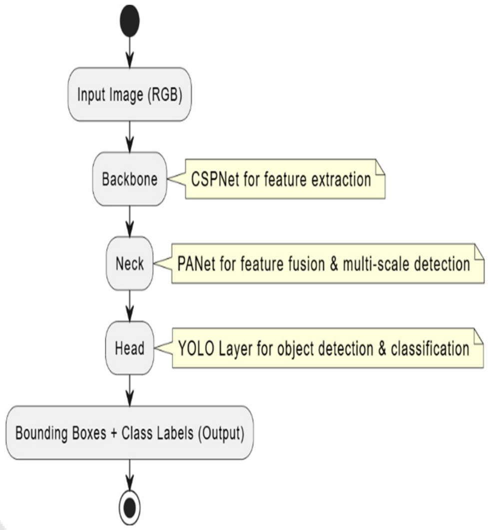

Figure 2: YOLOv5 Architecture.

3.6 Optimal Model Selection

Three training datasets are employed: our own 360-degree thermal images captured using the Fluke Ti300+, the Kaggle AAU TIR image dataset, and FLIR ADAS. We manually annotate with bounding boxes the people in these images, as illustrated in Figure 2. As it provides the best visibility in infrared imaging, all of the images are within the long-wave infrared (LWIR) band (8–15μm). With a 1,000 burn-in time and learning accuracy of 0.0014, the model is trained for 5,000 epochs. The training was completed in less than a day on our machine with default hyperparameters due to the high contrast of thermal human images and the ability of the model to extract good features through 53 network layers.

4 EXPERIMENT RESULT

4.1 Finding the Ground Truth for Objects that Are Partially Occluded

We aim to recognize and enumerate an individual as a unique person in our case even though they may be partially hidden two individuals are counted as one object if they are extremely close to each other and one of them is more than 50 covered by the other five individuals are employed as ground truth GT in the left image however since the three individuals on the right-hand side are so close together that they cannot be seen as individual units the right photo only accounts for three people as GT.

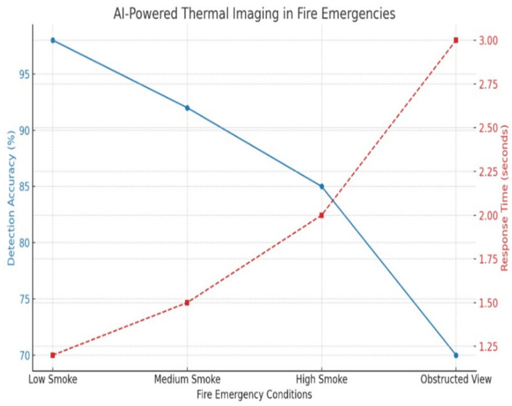

Figure 3: Fire Emergency Conditions.

ground truth for partially occluded objects needs to be known in order to achieve correct object detection since only a fragment of an object's features might be visible when it is being covered by another precise localization is challenging even when portions of the piece are concealed ground truth annotation must be present in this issue in order to include the pieces intended shape and location one method is to approximate the overall bounding box of the item from its visible parts and from prior knowledge of its typical structure additionally methods such as segmentation masks and occlusion-aware labeling aid in optimizing the annotations by distinguishing between completely visible areas and partially occluded areas these techniques assist the model in identifying and classifying objects even with occlusions more accurately by improving the quality of the training data. Figure 3 shows the Fire Emergency Conditions.

4.2 Measures of Detection Accuracy, Precision, and Recall

when we expect the bounding box to closely resemble the real object the detection is called a true positive this is usually determined using the intersection over union iou threshold of 50 or higher FN is produced if the iou drops below this level or if the model was unable to identify an object, on the other hand, an fp happens when the model labels an item that isn't there we may gain a better understanding of the models capacity for precise item identification and classification by contrasting precision and recall with you-based accuracy in order to improve object detection algorithms for practical uses these metrics are essential.

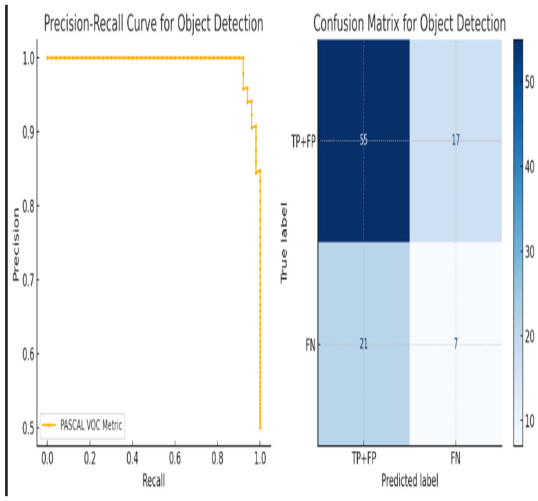

Figure 4: PASCAL VOC Metric for Precision and Confusion Curve.

4.3 PR Curve, Precision, and Recall in Test Datasets

we concluded that the model weights after 4000 epochs were the best option as they exhibited the minimum training loss with outstanding accuracy and accurate object localization on the test set our model runs very well with a precision and recall of over 97 also even when viewed from various angles and human positions figure 4 captures a roc score of 98 this high accuracy is attributed to the robust 53-layer deep cnn which provides precise object detection with bounding boxes always meeting or surpassing 50 IOU. Table 1 shows the Augmentation Dataset Image. Table 2 shows the Comparison of Volov4 And Volov5.

Table 1: Augmentation Dataset Image.

Fire Condition	Detection accuracy (%)	Response Time(s)	False Alarm rate (%)
Low Smoke	98	1.2	2
Medium Smoke	92	1.5	5
High Smoke	85	2.0	10
Obstructed View	70	3.0	18

Table 2: Comparison of YOLOv4 and YOLOv5.

Model	mAP@50 (Accuracy %)	Inference Speed (FPS)	Model Size	Training Time
YOLOv4	85–90%	~45 FPS (GPU)	Large (244 MB)	Longer
YOLOv5	90–95%	~60 FPS (GPU)	Small (14 MB)	Faster

5 CONCLUSIONS

We introduce yolov4 a deep learning model that uses an NFPA 1801 compliant thermal imaging camera to show that it can recognize humans in dense smoke even in high-temperature low-visibility fire situations greyscale human forms are enhanced by the camera’s exceptional resolution and low-temperature sensitivity on an Nvidia GeForce 2070 GPU the model converges in 4000 epochs with MS coco pre-trained weights attaining over 95 accuracies at a 50 IOU it is useful for search and evacuation surveillance because it can identify people who are squatting standing sitting and lying down even when there is 50 occlusion at 301 fps real-time detection operates by mapping settings recognizing people and locating heat sources to increase firefighter safety future integration with robotic systems could improve search and rescue operations

6 FUTURE WORK

Future developments in AI-powered thermal imaging for real-time human detection during fire situations can concentrate on a few important areas. First, real-time performance can be improved by tailoring deep learning models for edge devices like drones and Internet of Things sensors. Methods like lightweight architectures (like YOLOv8-Nano) and model quantization (like Tenso RT, and ONNX) can lower computational load without sacrificing accuracy. Furthermore, combining thermal cameras with RGB and LiDAR sensors to integrate multi-spectral imaging might increase the accuracy of detection in low-visibility situations brought on by smoke or intense heat. Incorporating Transformer-based models (such as DETR and YOLO-World) can further improve AI capabilities by enabling context-aware human and fire danger detection, and real-time tracking algorithms can help track human movement within fire zones. Model resilience can be improved by adding more various fire situations to the dataset, such as varying temperatures, smoke concentrations, and human postures. Overcoming data constraints can also be aided by the creation of synthetic data using GANs (Generative Adversarial Networks).

REFERENCES

A. Filonenko, D. C. Hernández, and K. -H. Jo, "Real-time smoke detection for surveillance," 2015 IEEE 13th International Conference on Industrial Informatics (INDIN), Cambridge, UK, 2015, pp. 568-571, doi: 10.1109/INDIN.2015.7281796.

A. Filonenko, D. C. Hernández, Wahyono, and K. -H. Jo, "Smoke detection for surveillance cameras based on color, motion, and shape," 2016 IEEE 14th International Conference on Industrial Informatics (INDIN), Poitiers, France, 2016, pp. 182-185, doi: 10.1109/INDIN.2016.7819155.

A. Filonenko, D. C. Hernández and K. -H. Jo, "Fast Smoke Detection for Video Surveillance Using CUDA," in IEEE Transactions on Industrial Informatics, vol. 14, no. 2, pp. 725-733, Feb. 2018, doi: 10.1109/TII.2017.2757457. keywords: {Cameras; Image color analysis;Graphics processing units;Sensors;Video}

A. Mariam, M. Mushtaq and M. M. Iqbal, "Real-Time Detection, Recognition, and Surveillance using Drones," 2022 International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering (ETECTE), Lahore, Pakistan, 2022, pp. 1-5, doi: 10.1109/ETECTE55893.2022.10007285.

Bhavnagarwala and A. Bhavnagarwala, "A Novel Approach to Toxic Gas Detection using an IoT Device and Deep Neural Networks," 2020 IEEE MIT Undergraduate Research Technology Conference (URTC), Cambridge, MA, USA, 2020, pp. 1-4, doi: 10.1109/URTC51696.2020.9668871.

D. Kinaneva, G. Hristov, G. Georgiev, P. Kyuchukov and P. Zahariev, "An artificial intelligence approach to real-time automatic smoke detection by unmanned aerial vehicles and forest observation systems," 2020 International Conference on Biomedical Innovations and Applications (BIA), Varna, Bulgaria, 2020, pp. 133-138, doi: 10.1109/BIA50171.2020.9244498.

D. K. Dewangan and G. P. Gupta, "Explainable AI and YOLOv8-based Framework for Indoor Fire and Smoke Detection," 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS), Bangalore, India,
2024, pp. 1- 6, doi:10.1109/ICITEICS61368.2024.106 24874.

G. M. A, A. Sivanesan, G. Yaswanth and B. Madhav, "ECOGUARD: Forest Fire Detection System using Deep Learning Enhanced with Elastic Weight

Consolidation," 2024 IEEE International Conference on Big Data & Machine Learning (ICBDML), Bhopal, India, 2024, pp. 115-120, doi: 10.1109/ICBDML60909.2024.10577367. [7].

H. Prasad, A. Singh, J. Thakur, C. Choudhary and N. Vyas, "Artificial Intelligence Based Fire and Smoke Detection and Security Control System," 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2023, pp. 01- 06, doi: 10.1109/NMITCON5819 6.2023.10276198.

H. Kulhandjian, A. Davis, L. Leong, M. Bendot and M. Kulhandjian, "AI-based Human Detection and Localization in Heavy Smoke using Radar and IR Camera," 2023 IEEE Radar Conference (RadarConf23), San Antonio, TX, USA, 2023, pp. 1-6, doi: 10.1109/RadarConf2351548.2023.10149735

R. Donida Labati, A. Genovese, V. Piuri and F. Scotti, "Wildfire Smoke Detection Using Computational Intelligence Techniques Enhanced with Synthetic Smoke Plume Generation," in IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 43, no. 4, pp. 1003- 1012, July 2013, doi: 10.1109/TSMCA.2 012.2224335.