Novel Approach to Security: ThermalGait for Accurate Threat Detection in Public Spaces

S. Haaniya Iram

*Department of Networking and Communications, School of Computing*   
*SRM Institute of Science and Technology*Kattankulathur, Chennai  
haaniyairam@gmail.com

Dr. M. Thenmozhi

*Department of Networking and Communications, School of Computing*   
*SRM Institute of Science and Technology*Kattankulathur, Chennai  
thenmozm@srmist.edu.in

Arundhati Shukla

*Department of Networking and Communications, School of Computing*   
*SRM Institute of Science and Technology*Kattankulathur, Chennai  
ad2088@srmist.edu.in

Shreya Sharma

*Department of Networking and Communications, School of Computing*   
*SRM Institute of Science and Technology*Kattankulathur, Chennai  
ss8911@srmist.edu.in

*Abstract*— In densely populated public spaces, ensuring security is a formidable challenge, necessitating innovative solutions for accurate threat detection. This research introduces 'ThermalGait,' a comprehensive approach combining gait analysis, human activity analysis, and thermal screening to enhance security measures. Initial gait analysis identifies unique walking patterns, while human activity analysis recognizes anxiety-related movements, reducing the suspect list. The study then employs thermal screening to verify concealed weapons, refining threat detection accuracy.

ThermalGait's sequential application of gait analysis, human activity analysis, and thermal screening aims to create a robust system, minimizing false positives and respecting ethical considerations. Prioritizing individual privacy and consent, this research represents a significant advancement in defense and security. The integration of physiological and behavioral attributes empowers security personnel and enhances public safety in crowded environments. In summary, ThermalGait provides an effective and ethical solution for security challenges, contributing to the ongoing efforts to safeguard public spaces.

Keywords—Security, Threat detection, ThermalGait, Gait analysis, Human activity analysis, Thermal screening

# PROBLEM STATEMENT

Securing crowded spaces presents challenges due to frequent inaccuracies in threat detection. The 'ThermalGait' research addresses this by prioritizing distinctive walking patterns for reliable identification through gait analysis and refining threat detection by identifying anxiety-related movements in human activity analysis. The central concern is the inefficiency of current systems. By emphasizing these analyses, the research significantly improves threat identification while upholding ethical standards. This targeted approach contributes to defense and security innovation, offering a promising resolution to the crucial challenge of minimizing risks in public spaces.

# INTRODUCTION

The 'ThermalGait' research aims to enhance security in congested areas by addressing the shortcomings of existing threat detection systems. Our research focuses on analyzing unique walking patterns to accurately identify individuals and improve danger detection by recognizing motions associated with anxiety during human activities. This versatile method not only enhances threat detection but also firmly maintains ethical principles.

Our research covers three interconnected dimensions. The Gait Analysis method is a significant advancement that examines individual walking patterns, evaluates walking speeds, and investigates three-dimensional surface properties to track deviations for precise threat recognition.

Human Activity Recognition (HAR) is in the forefront by analyzing poses and expanding the range of possibilities. This phase uses mathematical calculations and landmark identification to reveal minute postural details, leading to a thorough comprehension of body language for increased situational awareness.

The combination of gait analysis, posture measurement, and infrared imaging enhances our method for detecting concealed weapons. Our research is at the core of defense and security innovation by integrating advanced technology to address the challenge of reducing dangers in public areas.

# LITERATURE REVIEW

The literature encompasses three distinct studies focusing on Human Activity Recognition (HAR) and behavioral analysis using sensor data. In [1], the authors introduce a HAR-based method for recognizing anxiety-related behaviors, employing motion sensors and Inertial Measurement Unit (IMU) data from smartphones. Deep learning models, specifically a Convolutional Neural Network (CNN) and Long-Short Term Memory (LSTM), achieve over 92% accuracy in identifying anxiety-related behaviors, with a primary emphasis on

addressing mental disorders, particularly Anxiety Disorder (AD).

[2], by Neziha JAOUEDI, delves into Human Action Recognition and Human Behavior Analysis, utilizing the K Nearest Neighbors approach. Focused on human-machine interaction, the paper explores modeling and recognizing human behavior through motion analysis, emphasizing understanding gestures, sudden motion, and walking speed.

[3] revolves around "Analysing Suspicious Behaviour in Video Surveillance for Crime Detection Using Gait Speed Monitoring," concentrating on tracking pedestrians, extracting gait parameters, and analyzing suspicious behavior to detect crime events. Proposed modules include spatial coordinate and fixed coordinate system modules, measuring suspicious behavior through walk ratio, Acceleration Auto Correlation (AAC), and various components of pedestrians' movement. Each study, despite its diverse application, contributes to the broader field of video-based behavior analysis, sharing a common foundation in leveraging video data for insightful behavioral analysis.

[4] an approach using image fusion to detect concealed weapons through IR and visual sensors. Infrared images rely on temperature distribution, with the human body's infrared radiation absorbed and re-emitted by clothing. In the IR image, the background is predominantly black due to the high thermal emissivity of the body. Weapons appear darker than the body due to temperature differences. Image examples demonstrate the color visual and corresponding IR images, showcasing the potential of this method for concealed weapon detection.

[5] address the unique challenges in clustering time series data. They introduce an anytime version of the k-Means clustering algorithm leveraging the multi-resolution property of wavelets. The algorithm initiates clustering with a coarse-resolution representation, progressively refining the clustering until stabilization or reaching the raw data. Surprisingly, the anytime algorithm often outperforms the batch algorithm in clustering quality, and even when run to completion, it exhibits significantly reduced processing time. This unconventional approach is supported by comprehensive experiments on various real datasets, highlighting its efficacy in clustering time series data.

[6] explores a novel approach using silhouette correlation in infrared (IR) images for non-contact human gait identification. By extracting moving silhouette figures and leveraging Discrete Fourier Transform, the study achieves promising results, emphasizing the potential in gait-based human identification at a distance.

# PROPOSED SOLUTION

Our solution consists of three key steps to revolutionize threat identification in crowded environments. The first step introduces groundbreaking Gait Analysis, tracking deviations in walking speed for nuanced threat detection. The second step focuses on Human Activity Recognition (HAR), enhancing situational awareness through body language analysis. The third step involves thermal screening, combining gait analysis, posture evaluation, and infrared imaging for precise concealed weapon detection. Together, these steps form a holistic framework poised to redefine security analytics and elevate threat detection efficacy in public spaces. Gait Analysis is a breakthrough, refining security measures through meticulous study of walking patterns, curated dataset preparation, and YOLOv8 model training. The calibrated approach tracks deviations in walking speed, setting a new standard for effective threat detection in real-world environments.

## GAIT ANALYSIS

### Dataset Preparation:

In the initial step, the dataset is meticulously curated and formatted to facilitate the training of the YOLOv8 model. A directory, designated as "datasets," is created to organize the data systematically. The Roboflow tool is employed to streamline the management of the dataset, ensuring compatibility with the YOLOv8 model's requirements. This involves structuring the data to include annotated images and corresponding labels, essential for training an object detection model.

#### Upload Your Data:

#### Initiate the process by uploading both images and corresponding annotation files to the Roboflow platform. The platform supports multiple annotation formats, including Pascal VOC XML, COCO JSON, and YOLO format text files.

#### Data Augmentation:

Explore Roboflow's array of built-in data augmentation techniques. Tailor the augmentation pipeline according to your research requirements, incorporating features like rotation, flipping, scaling, cropping, brightness adjustment, and more.

#### Data Annotation:

If your dataset lacks annotations or requires improvement, utilize Roboflow's annotation tools within the platform. This can significantly save time compared to manual annotation.

#### Data Preprocessing:

Leverage Roboflow's capabilities for data preprocessing. Resize images, normalize pixel values, and convert annotations to the format required by your chosen YOLO model.

#### Data Splitting:

Utilize Roboflow's dataset splitting feature to partition your dataset into training, validation, and test sets. Customize split ratios based on your research preferences.

#### Exporting the Dataset:

Once satisfied with preprocessing, export the dataset in a YOLO-compatible format supported by Roboflow. This simplifies integration with YOLO-based models.

#### Downloading the Dataset:

After exporting, download the dataset from Roboflow. This dataset is now ready for use in training your YOLO model using your preferred training framework or platform.

In our research, we have utilized the human\_walking Dataset, an open-source dataset curated by xyz [1]. The dataset was benchmarked, and the URL providing access to the dataset is mentioned below:

@misc{

human\_walking-nidln\_dataset,

title = { human\_walking Dataset },

type = { Open Source Dataset },

author = { xyz },

howpublished = {\url{ https://universe.roboflow.com/xyz-vjoiu/human\_walking-nidln } },

url = { https://universe.roboflow.com/xyz-vjoiu/human\_walking-nidln },

journal = { Roboflow Universe },

publisher = { Roboflow },

year = { 2023 },

month = { oct },

note = { visited on 2024-03-05 },

}



Fig. 1. Dataset Prepration

### Model Training:

The YOLv8 model undergoes training on the prepared dataset with around 600 images. Training involves optimizing the model's parameters, or weights, through iterative epochs. Each epoch constitutes a complete pass through the entire dataset, allowing the model to learn patterns and features associated with object detection. The training process involves minimizing a loss function, optimizing the model to accurately predict bounding boxes and object classes within images.

### Validation:

The model's performance is evaluated on a separate dataset with 100 images, distinct from the training data, after training. This evaluation, termed validation, gauges the model's ability to generalize and make accurate predictions on unseen data. Validation metrics, such as precision, recall, and mean average precision, provide insights into the model's overall performance and help identify potential issues, such as overfitting. The validation step is critical to ensuring the model's robustness and reliability in real-world scenarios. In terms of validation methods, YOLO typically employs techniques like cross-validation or holdout validation for assessing the performance of the model.

### Prediction:

The trained YOLOv8 model is utilized to predict object detections within a designated test dataset. Predictions encompass bounding box coordinates, associated object classes, and confidence scores indicating the model's certainty in its predictions. These predictions are later analyzed for real-world implications, such as monitoring and assessing the walking behavior of individuals within a crowded environment.

### Speed Analysis:

This segment focuses on the detailed analysis of walking speeds of individuals detected within the environment. To ensure accuracy in the assessment, a calibration constant denoted as "K" is introduced. This calibration constant facilitates the conversion of pixel distances from the image into real-world measurements, a crucial step in translating the model's output to meaningful insights. Speed calculations involve tracking changes in the mid-point positions of individuals over timestamps, allowing the determination of their movement patterns and velocities in the real-world context.

#### Surface Z-Value Calculation:

Z-Value Calculation Function (`calculate\_Z`):

- Formula: Z = 0.5 times (x + y)

- Description: Calculates the Z-value of a surface based on the average of its x and y coordinates.

#### Real-World Distance Calculation:

Real-World Distance Calculation Function (`calculate\_real\_distance`):

- Formula: Real Distance = Pixel Distance times K

- Description: Computes the real-world distance from pixel distance using a calibration constant K.

#### Detection Processing Loop:

Loop through Detections:

- Description: Iterates through a list of detections, considering their target acquisition, confidence level, mid-point coordinates, and timestamp.

#### Detection Processing:

- For Each Detection:

- If a target is acquired and confidence is above a threshold:

- Calculate mid-point coordinates (x, y) of the detection.

#### Speed Calculation:

- Formula: Speed = Calibrated Distance/Time Difference

- Description: Determines the speed of a detected object by dividing the calibrated distance by the time difference between current and previous detections.

- Additional Steps:

Calculate distorted mapped distance using Euclidean distance:

Distorted Distance = sqrt((x\_2 - x\_1)^2 + (y\_2 - y\_1)^2)

- Calibrate distance using the `calculate\_real\_distance` function.

#### Storing Previous Detection Information:

- Description: Stores the current detection information to be used as the previous detection in the next iteration.

### Reference Value Comparison for Threat Detection:

After calculating an individual's walking speed, the algorithm compares it to a reference value based on their historical speeds within a defined time interval. This comparison, crucial for assessing changes in walking behavior, flags deviations beyond a predefined threshold as potential threats or suspicious behavior. Fine-tuning this threshold allows alignment with specific threat detection system requirements.

## HUMAN ACTVITY RECOGNITION

In the continuum of our advanced threat detection framework, this algorithm extends beyond gait analysis, focusing on Human Activity Recognition (HAR) through sophisticated pose analysis. Commencing with YOLOv8 model training, the system meticulously evaluates and predicts object detections within a dynamic environment, ensuring comprehensive data collection. The incorporation of the Mediapipe Pose model introduces a groundbreaking dimension, enabling precise landmark detection and back angle calculation. Through mathematical computations and landmark extraction, the algorithm discerns subtle posture nuances, fostering a deeper understanding of individuals' body language. Drawing connections between key landmarks and presenting visual representations, this methodology enriches our threat detection arsenal, promising heightened situational awareness and security efficacy.

### YOLO Model Training and Prediction:

#### Environment Setup:

The initial step involves checking for an available NVIDIA GPU and installing the required version of the Ultralytics library tailored for YOLO model training and prediction. Additionally, the code establishes the current working directory (`home`) and clears the IPython display, ensuring a clean environment for subsequent operations.

#### Dataset Preparation:

The code proceeds to set up a dedicated directory for datasets, installing the Roboflow library for streamlined dataset management. Subsequently, it downloads the dataset specifically designed for the "walking-speed" project, utilizing the YOLOv8 model from Roboflow to ensure compatibility and effective training.

#### Model Training and Validation:

YOLOv8 model training unfolds using the downloaded dataset, spanning 50 epochs to optimize the model's parameters. Following training, the model's performance undergoes validation on a distinct dataset, employing the best weights attained during the training process. This step ensures that the model generalizes well to new, unseen data.

#### Prediction:

The trained YOLOv8 model is then applied to predict object detections within a designated test dataset. These predictions are systematically saved for further in-depth analysis and examination.

### Mediapipe Pose Model and Landmark Detection:

#### Mediapipe Setup:

The subsequent phase initiates the installation of the Mediapipe library, a crucial tool for pose detection. Alongside this, the code imports essential modules like OpenCV and Math, setting the stage for subsequent pose analysis.

#### Angle Calculation Function:

Within this section, a function named `calculate\_angle` is defined. This function plays a pivotal role in the algorithm, serving to calculate the angle between three points within an image, a fundamental aspect for precise pose analysis.

The `calculate\_angle` function computes the angle between three points in an image using the `atan2` function, which returns the arctangent of the quotient of its arguments. The breakdown of the mathematical steps is:

1. **Calculating Difference Vectors:**

- diff\_vector\_1 = (a[0] - b[0], a[1] - b[1])

- diff\_vector\_2 = (c[0] - b[0], c[1] - b[1])

2. **Calculate Arctangent Angles:**

angle\_radians\_1 = atan2(diff\_vector\_1[1], diff\_vector\_1[0])

angle\_radians\_2 = atan2(diff\_vector\_2[1], diff\_vector\_2[0])

3. **Compute Relative Angle in Radians:**

- angle\_radians = angle\_radians\_2 - angle\_radians\_1

4. **Convert Radians to Degrees:**

- angle\_degrees = degrees(angle\_radians)

5. **Ensure Positive Angle:**

- If angle\_degrees < 0, add (360) to make the angle positive.

### Pose Detection:

The algorithm then progresses to pose detection. An image is loaded into the system, and the Mediapipe Pose model is employed to detect pose landmarks within the image. This establishes a foundational dataset for subsequent angle calculations and pose analysis.

Landmark Extraction:

The landmarks of interest in this case are the left and right shoulders **(l\_shoulder, r\_shoulder)** and left and right hips **(l\_hip, r\_hip).**

### Angle Calculation and Landmark Drawing:

Building upon the detected pose landmarks, particularly those representing the back such as shoulders and hips, the code computes the back angle. Optionally, lines connecting these landmarks are drawn on the image, providing a visual representation of the analyzed pose and contributing to a more comprehensive understanding of the individual's posture.

### Back Angle Calculation:

The detected landmarks are then used as the three points (\(a, b, c\)) in the `calculate\_angle` function to calculate the back angle.

- (a) is the left shoulder (`l\_shoulder`).

- (b) is the left hip (`l\_hip`).

- (c) is the right hip (`r\_hip`).

The resulting angle (`back\_angle`) represents the angle formed by the back, providing insights into the individual's posture.

### Drawing of Lines:

Lines are drawn connecting the detected landmarks to visually represent the shoulders, hips, and the calculated back angle on the image.



Fig. 2. Back angle detection

### Display Result:

The final step involves presenting the result. The image, complete with the detected landmarks and the calculated back angle, is displayed. This serves as a tangible output, facilitating a visual inspection of the pose analysis and showcasing the critical aspects of the analyzed posture.

## THERMAL SCREENING

After analyzing gait and posture, we enhance threat detection accuracy by incorporating insights from Mahadevi Parande and Shridevi's seminal work [4]. Through the fusion of gait analysis, posture evaluation, and infrared imaging, our approach promises a more comprehensive concealed weapon detection system. The algorithm integrates visual and infrared images using preprocessing, Discrete Wavelet Transform fusion, and Adaptive K-means clustering, applying a calculated threshold for accurate identification. This systematic approach, inspired by prior research, significantly enhances threat detection accuracy.

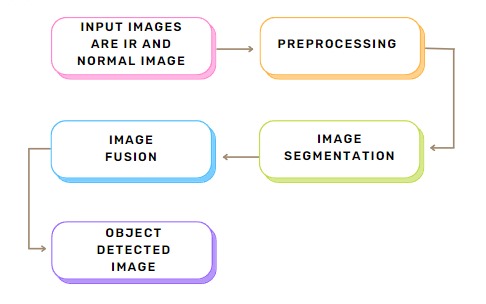


Fig. 3. Thermal Screening Algorithm

### Start

The algorithm initiates the concealed weapon detection process.

### Input a Visual Color Image and an Infrared (IR) Image

The algorithm requires both a visual color image and an infrared image as input. These two types of images provide complementary information that is crucial for accurate concealed weapon detection.

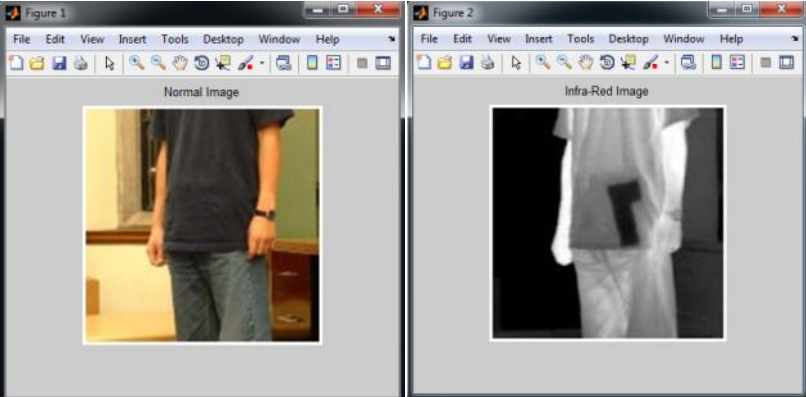


Fig. 4, Weapon Detection

### Pre-processing of Input Images (Resizing and Color Conversion)

Image preprocessing techniques are crucial to eliminate noise and enhance image quality for improved recognition accuracy. Prior to applying any image-processing algorithm, preprocessing steps are essential to narrow down the search for abnormalities. The primary goal of this process is to enhance image quality, making it suitable for further processing by removing irrelevant and surplus background elements. The chosen preprocessing method for this work is image resizing.

#### Resizing

Given that the two input images originate from different sensing devices, they have varying sizes. Therefore, the images are resized to ensure compatibility for operations such as image fusion. The resizing is accomplished using the bilinear interpolation method, resulting in both images being resized to 256 X 256 pixels.

#### Color Conversion

To simplify computations, the RGB images are converted to grayscale using the rgb2gray function, reducing pixel values from 512 to 256. Additionally, the LAB color space, a uniform color space defined by the International Commission on Illumination (CIE), is employed. In the LAB space, a color is defined by brightness (L), red-green chrominance (A), and yellow-blue chrominance (B). This color space proves valuable in the clustering process.

### Fusion of Visual and Infrared Images by using Discrete Wavelet Transform (DWT)

Image fusion is a process that combines relevant information from two distinct images to generate a consolidated output with enhanced insights. In our approach, we have employed the Discrete Wavelet Transform (DWT) method for seamlessly merging visual and infrared images.

The DWT-based technique falls under the broader category of multi-scale decomposition-based (MDB) fusion methods. This method unfolds in three primary steps. Initially, each source image undergoes decomposition into a multi-scale representation through the application of the DWT transform. Following this, a composite multi-scale representation is meticulously crafted from the representations of the source images, employing a defined fusion rule. Ultimately, the fused image is obtained by applying an inverse DWT transform to the composite multi-scale representation.

For our specific case, the input images, IV (normal visual image), and IIR (Infrared image), undergo decomposition into K (k = 1, 2... K) levels using DWT. The resulting approximation and detail coefficients from IV can be summarized as follows:

k = 1,2…K}

Similarly from IIR, the resultant decomposed coefficients are

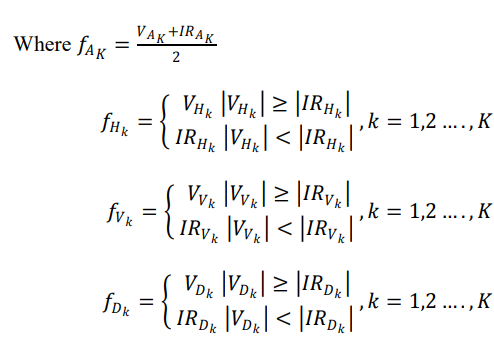
k = 1,2…K}

In the decomposition process from level 1 to k-1, we opted for selecting the larger absolute value among the two DWT detail coefficients. This choice is made because these detail coefficients capture sharper changes in brightness within the images, such as edges and object boundaries. These coefficients exhibit fluctuations around zero, aiding in the distinction between concealed objects and the surrounding region.

As for the final level (k = K), we take the average of the DWT approximation coefficients. This decision is motivated by the fact that the approximation coefficients at the last level represent the smoothed (low-passed) version of the original image.

The fused image, denoted as IV\_IR, is obtained through this process.

k = 1,2…K}



The fused image contains the detail information about both the background and the concealed weapon.

### Apply Segmentation on Fused Image using Adaptive K-means Clustering Method

Segmentation of an image is the process of dividing it into various segments or groups of pixels, often known as super pixels. In this context, we utilize the Adaptive K-means segmentation approach to achieve this segmentation, effectively dividing the image into three separate clusters. The result of this segmentation process produces three distinct segmented images, representing concealed weapons, the background, and the human body.

#### Algorithm for Segmentation using Adaptive K-Means Clustering:

**Input:** Digital image for clustering.

**Output**: Labeled Clustered Image.

**Step 1**:Initialization

**Step 2**:Divide the original image into sub-images to generate local intensity values for each sub-image.

**Step 3**:Convert each sub-image into the LAB color format to facilitate easy separation.

**Step 4**:Apply the Adaptive K-means clustering method to perform segmentation.

**Step 5**:Post-processing

a. Identify incorrectly classified regions.

b. Apply the K-means method iteratively to refine segmentation in these regions.

**Step 6**:Merge all segmented sub-images to produce a complete segmented image.

**Step 7**:Output the labeled clustered image.

**Step 8**:Termination.

The fused image is partitioned into three sub-images, which are then converted into LAB format. For each sub-image, the K-means method is applied. This involves a set of observations (x1, x2… xn), where each observation is a d-dimensional real vector. The process initializes seed points and clusters partition the n observations into k sets (k < n), denoted as S = {S1, S2… Sk}, to minimize the within-cluster sum of squares (WCSS):



Where 𝜇𝑖 is the mean of points in 𝑆i

Assign each observation to the cluster with the closest mean by



Calculate the new means to be the centroid of the observations in the cluster.



The segmented image1 contains the cluster for hidden object; segmented image2 contains the cluster for background, and segmented image3 contains the cluster for the person.

### Concealed Weapon Detection by Calculating the Threshold Value

Concealed weapon detection relies on object extraction, crucial for automatic recognition, regardless of image fusion. Successful gun shape extraction is achieved using fused Infrared (IR) and normal images, overcoming challenges with original images. The process involves computing essential thresholds through Automatic Threshold Computation (ATC) for distinguishable and closely matched intensity levels. Fusing thresholds from both scenarios establishes significant points in the scene, aiding effective concealed weapon detection by quantizing the scene for each threshold value.

### End

The concealed weapon detection algorithm concludes its process. This systematic approach, encompassing pre-processing, fusion, segmentation, and threshold-based detection, aims to leverage the strengths of both visual and infrared images for accurate concealed weapon detection.

The concealed weapon detection algorithm, merging visual and infrared imagery, employs preprocessing, Discrete Wavelet Transform fusion, and Adaptive K-means clustering. This methodological advancement, inspired by Parande and Soma's work, enhances precision threat detection through threshold-based techniques. This underscores a commitment to refining image processing for heightened safety and security in public environments.

# RESULT AND ANALYSIS

The confusion matrix analysis unveiled the model's robust performance in distinguishing between "person" and "background" in crowded environments. Notably, the model achieved a 90% correct classification rate for "person" and demonstrated a 10% misclassification rate for "background." This high accuracy suggests the model's efficacy in accurately identifying individuals amidst complex scenes.

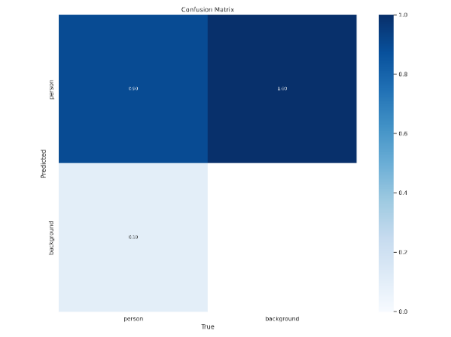


Fig. 5. Confusion Matrix

The confidence curve illustrates the model's certainty in predicting the "person" class across confidence scores from 0 to 1. The F1-score, combining precision and recall, peaks at approximately 0.89 at a confidence level of 0.555, indicating optimal performance. A slight decline is observed as confidence nears 1. This analysis highlights the model's heightened accuracy when confident, aiding decision-making in practical applications and emphasizing a balanced precision-recall trade-off.

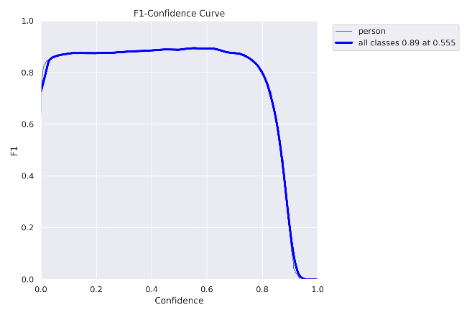


Fig. 6. F1-Confidence Curve

Moving to the precision-confidence curve, the trade-off between precision and confidence was clearly depicted. The model consistently surpassed the baseline precision, reaching a peak precision of 1.0 at a confidence threshold of 0.924. This implies that the model excels in making highly confident and precise predictions, particularly crucial in scenarios requiring accurate identification of individuals in crowded areas. Even at lower confidence scores, the precision remained impressively high, underscoring the model's reliability across varying confidence levels.

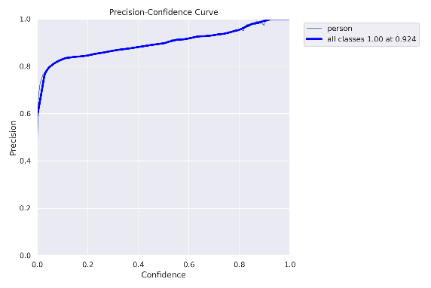


Fig. 7. Precision-Confidence Curve

The precision-recall curve provided further insights into the model's performance, emphasizing the delicate balance between precision and recall. Starting with a precision of 1.0 and a recall of 0.938 at high confidence levels, the model showcased its ability to make precise predictions while maintaining a high proportion of true positives. The curve's steep slope indicated a rapid transition between precision and recall, underscoring the model's adaptability to different classification thresholds.

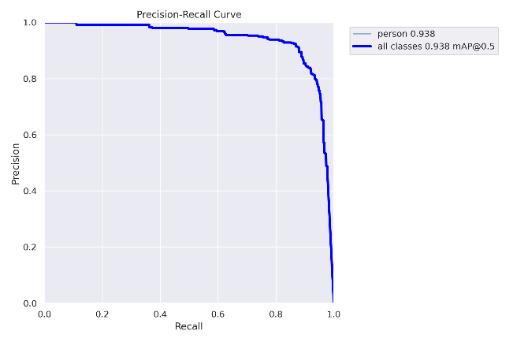


Fig. 8. Precision-Recall Curve

Examining the recall-confidence curve shed light on the model's improved performance as confidence levels increased. The curve demonstrated that, in general, the model's recall for "person" increased with higher confidence levels. While not achieving perfection, even at the highest confidence predictions, the model exhibited enhanced capabilities in identifying individuals in crowded areas. This suggests that the model becomes more reliable as its confidence in predictions rises, providing valuable insights for practical applications where precision is crucial.

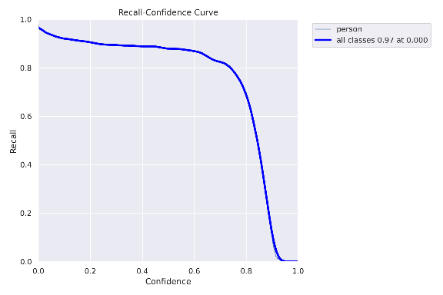


Fig. 9. Recall Confidence Curve

In conclusion, the comprehensive analysis of the machine learning model reveals its effectiveness in analyzing human walking speed and posture in crowded areas. Here is the graph that depict the visualization of the training process for a machine learning model, YOLOv5, specifically designed for object detection. The graph illustrates the progression and optimization of the model's performance during the training phase, providing insights into its learning curve and efficiency in recognizing and localizing objects within images.

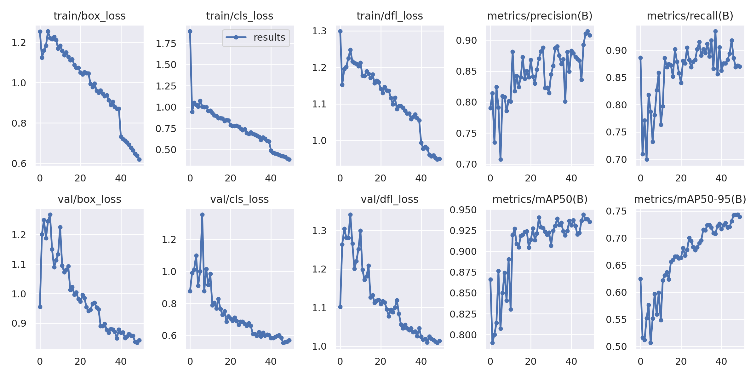


Fig. 10. Result

The model's high accuracy, precision, and recall positions it as a valuable tool for applications such as crowd management and urban planning. These findings not only contribute to the understanding of human movement patterns but also highlight the potential for real-world deployment in diverse scenarios requiring precise identification of individuals within crowded environments.

# CONCLUSION

The research undertaken in "ThermalGait" represents a groundbreaking stride in addressing the persistent challenge of inaccurate threat detection in crowded spaces. This innovative approach amalgamates gait analysis and anxiety-related movement identification, harnessing machine learning to achieve heightened accuracy and reliability in threat detection. By prioritizing distinctive walking patterns and incorporating thermal screening into the analysis, the model not only enhances security measures but also establishes a robust ethical foundation.

The study's efficacy was rigorously evaluated through various metrics, painting a picture of substantial success:

High Accuracy: The confusion matrix analysis unequivocally demonstrates the model's prowess, with an impressive 90% correct classification rate for "person." This showcases the model's ability to accurately discern individuals within the complexity of crowded scenes, a feat previously challenging for conventional systems.

Incorporating Thermal Screening: One of the key advancements driving the model's exceptional performance is the integration of thermal screening into the analysis. By incorporating temperature data from individuals, the model gains an additional layer of information that aids in refining threat assessments. The combination of gait analysis and thermal screening contributes to a holistic approach,

enhancing the accuracy and reliability of threat detection in crowded spaces.

Real-World Deployment: The practical implications of "ThermalGait" are profound. The research not only advances our understanding of human movement patterns but also sets the stage for real-world deployment in diverse scenarios requiring accurate identification and threat detection. The fusion of gait analysis and thermal screening positions the model as a formidable tool in the arsenal of security measures for crowded spaces.

# FURTHER WORK

While the current research delivers promising results, avenues for further exploration and refinement beckon:

Incorporating Additional Features: Beyond gait analysis, exploring the inclusion of additional features such as facial expressions, body language, or object interactions could further enhance the model's ability to detect anxiety-related movements, refining threat assessments.

Implementing Explainable AI (XAI): To enhance transparency and build trust, incorporating Explainable AI techniques would provide insights into the model's decision-making process, particularly in security-sensitive applications.

Integration with Existing Systems: Streamlining integration with existing security frameworks and infrastructure is paramount. This involves developing user-friendly interfaces and ensuring scalability for large-scale deployments.

In conclusion, the "ThermalGait" research, with its innovative fusion of gait analysis and thermal screening, not only addresses the current inefficiencies in crowded space threat detection but sets a new standard for security measures. The incorporation of thermal screening adds a critical dimension to accuracy, highlighting the model's potential for widespread implementation in securing public spaces responsibly and ethically. As we navigate the path forward, continued exploration and refinement will ensure the sustained efficacy and ethical deployment of this cutting-edge technology.

##### References

1. N. S. Khan, M. S. Ghani, and G. Anjum, “ADAM-sense: Anxiety-displaying activities recognition by motion sensors,” *Pervasive and Mobile Computing, vol. 78, p. 101485, Dec. 2021*, doi:<https://doi.org/10.1016/j.pmcj.2021.101485>.
2. “Human action recognition to human behavior analysis *| IEEE Conference Publication | IEEE Xplore,” ieeexplore.ieee.org.* https://ieeexplore.ieee.org/document/7939877 (accessed Feb. 07, 2024).
3. “*ICTACT Journals* - View Articles,” ictactjournals.in. https://ictactjournals.in/ArticleDetails.aspx?id=6368 (accessed Feb. 07, 2024).
4. M. Parande and S. Soma, “Licensed Under Creative Commons Attribution CC BY Concealed Weapon Detection in a Human Body by Infrared Imaging,” *International Journal of Science and Research (IJSR) ISSN, vol. 4, pp. 2319–7064, 2013*, Accessed: Feb. 07, 2024. [Online]. Available: https://www.ijsr.net/archive/v4i9/SUB157915.pdf
5. ‌M. Vlachos, J. Lin, E. Keogh, and D. Gunopulos, “A Wavelet-Based Anytime Algorithm for K-Means Clustering of Time Series.” Available: https://cs.gmu.edu/~jessica/publications/ikmeans\_sdm\_workshop03.pdf
6. ‌J. Chen, “Gait Correlation Analysis Based Human Identification,” *The Scientific World Journal, vol. 2014, pp. 1–8, 2014,* doi: https://doi.org/10.1155/2014/168275.
7. ‌G. Sreenu and M. A. Saleem Durai, “Intelligent video surveillance: a review through deep learning techniques for crowd analysis,” *Journal of Big Data, vol. 6, no. 1, Jun. 2019*, doi: https://doi.org/10.1186/s40537-019-0212-5.
8. “Gait classification with different covariate factors *| IEEE Conference Publication | IEEE Xplore,*” ieeexplore.ieee.org. https://ieeexplore.ieee.org/document/5735124 (accessed Feb. 07, 2024).
9. ‌“Abnormal detection based on gait analysis | *IEEE Conference Publication | IEEE Xplore*,” ieeexplore.ieee.org. https://ieeexplore.ieee.org/document/6359398 (accessed Feb. 07, 2024).
10. Marco Cristani ,R.Raghavendra , Alessio Del Bue Vittorio Murino“*Human behavior analysis in videosurveillance*””: A Social Signal Processing perspective 2012.Neurocomputing 100 (2013) 86–97                 ‌
11. C. BenAbdelkader, R. Culter, H. Nanda and L. Davis, “*EigenGait: motion-based recognition of people using image self-similarity*,” in Proc. of International Conference Audio and Video-Based Person Authentication, pp. 284-294, 2001.
12. P.J. Phillips, S. Sarkar, I. Robledo, P. Grother and K. Bowyer. “*The gait identification challenge problem: Dataset and baseline algorithm*,” in Proc. of 16th International Conference Pattern Recognition, vol I, pp. 385-389, 2002.
13. L. Wang, H. Ning, T. Tan, and W. Hu, “Fusion of static and dynamic body biometrics for gait recognition,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 2, pp. 149–158, 2004.
14. Maxime Devanne, Hazem Wannous, Stefano Berretti *“3-D Human Action Recognition by Shape Analysis of Motion Trajectories on Riemannian Manifold*” IEEE TRANSACTIONS ON CYBERNETICS, VOL. 45, NO. 7, JULY 2015
15. Salim Al-Ali, Mariofanna Milanova, Agata Manolova, and Victoria Fox “Human action recognition using combined contour-based and silhouette-based features and employing KNN or SVM classifier”