ThermalTrack: Efficient Thermal Face Detection with Haar Cascades and Tracking for Search and Rescue

Ankita Nagmote[1], Anushree Devarashetty[2], Shubha Puthran[2]
[1]Department of Information Technology, K J Somaiya School Of Engineering, Somaiya Vidyavihar University, Mumbai, India
[2]Department of Computer Engineering, Mukesh Patel School of Technology Management & Engineering, SVKM's NMIMS, Mumbai, India
E-mail: ankita.n@somaiya.edu.in, anushree.devarashetty75@nmims.in, shubha.puthran@nmims.edu

Keywords: Thermal imaging, face detection, Haar cascade classifier, dLib shape predictor, search and rescue operation, real-time detection, Kalman filter, HADR preprocessing

Received: July 4, 2025

Disasters pose immense challenges to search and rescue (SAR) operations, where rapid and accurate survivor detection is critical for mission success. This paper presents ThermalTrack, a computationally efficient real-time face detection system specifically designed for thermal imaging in SAR operations using UAV-mounted FLIR Lepton 3.5 cameras (160×120 pixel resolution). Our system integrates Histogram-based Adaptive Dynamic Range (HADR) preprocessing with dual detection pathways: Haar Cascade classifiers and Dlib's HOG-based frontal face detector, enhanced by Kalman filtering for robust multi-frame tracking. We systematically optimized detection parameters including scale factors (1.1-1.5) and minimum neighbor values (3-7) to achieve optimal accuracy-speed tradeoffs. Experimental validation across two distance scenarios shows detection accuracy of 85-92% with processing times of 30-50ms per frame. At 50-meter range (aerial operations), the system achieves 92% precision, 84% recall, and 88% F1-score. At 10-meter range (ground operations), performance reaches 85% precision, 78% recall, and 81% F1-score. HADR preprocessing improves detection rates by 13 percentage points (from 76% to 89%), while Kalman filtering provides stable tracking during brief occlusions and reduces detection jitter. The system demonstrates real-time capability suitable for resource-constrained UAV platforms, processing at 20-33 frames per second while maintaining competitive accuracy compared to computationally intensive CNN-based approaches.

Povzetek: Članek predstavi ThermalTrack, realnočasovni in računsko lahek sistem za zaznavo in sledenje obrazov v termalnih posnetkih SAR z UAV (HADR predobdelava + Haar/Dlib-HOG detekcija + Kalmanov filter) za zanesljivo delovanje na nizki ločljivosti FLIR Lepton 3.5.

1 Introduction

Thermal imaging is a technique that makes use of infrared radiation produced by objects to create images showing variations in temperature. This radiation is sensed by a thermal camera and is converted to electrical signals, which are displayed as grayscale images. Thermal imaging has become increasingly popular in several applications such as public safety, health, and disaster response [4], [2]. Among these, search and rescue (SAR) missions are of greatest concern, especially in the wake of natural disasters like hurricanes, earthquakes, and floods, where prompt detection of survivors is a priority [1], [2]. Evidence suggests that the initial 72 hours after a disaster are critical, as every delay in finding victims significantly lowers survival rates while increasing the search area exponentially [5].

Historically, SAR missions have been largely man-dependent, tending to be inefficient and yielding low success rates. Recent advances in autonomous technologies with computer vision capabilities have demonstrated good prospects in addressing these challenges through faster and more efficient identification of survivors [14], [1]. One of the commonly employed methods is utilizing unmanned aerial vehicles (UAVs) with thermal cameras to identify human presence in challenging environments, including low-visibility conditions like darkness or smoke [16]. But the main problem in such systems is to correctly detect relatively small human bodies in huge and crowded regions, usually producing false alarms and missed detections. The challenges in thermal imaging-based human detection for SAR operations are multifaceted and require careful consideration. First, thermal signatures of human bodies can vary significantly based on environmental conditions, clothing, and the physiological state of the individual. In disaster scenarios, victims may be partially buried under debris or covered with dust, further complicating detection. Second, the resolution of thermal cameras, particularly those mounted on UAVs, is often limited compared

to conventional optical cameras. This constraint makes it difficult to detect facial features at distances beyond 100 meters, especially when using lightweight UAVs with payload restrictions. Third, thermal cameras are prone to noise and artifacts, which can lead to false positives, such as heated objects or animals being misidentified as human bodies. Traditional methods for human detection in thermal imagery include simple thresholding techniques based on temperature differentials. While these approaches are computationally efficient, they lack robustness in complex environments and varying conditions. More sophisticated methods employ template matching, where predefined patterns of human thermal signatures are used for detection. However, templates often fail to accommodate the wide variety of poses, orientations, and partial occlusions that are common in disaster scenarios. Histogram of Oriented Gradients (HOG) features combined with Support Vector Machines (SVM) have shown promise, but they still struggle with small-scale detection and processing speed when implemented on resource-constrained platforms. The advent of deep learning has introduced more powerful paradigms for thermal image analysis. Researchers have explored various CNN architectures for human detection, including region-based models like Faster R-CNN, single-shot detectors like SSD and YOLO, and feature pyramid networks that handle multi-scale detection more effectively. Despite their superior accuracy, these models present significant computational challenges for real-time deployment on edge devices like UAVs.

Few research has investigated machine learning-based human detection in thermal images, with convolutional neural networks (CNNs) being the most notable method. Herrmann et al. [7] used CNNs to extract features and classify, whereas Zhang et al. [9] looked at the generation of synthetic data to improve CNN performance when dealing with poor-quality datasets. Herrmann et al. [7] suggested High-Resolution Infrared Monitoring (HRIM) techniques to handle thermal imagery more reliably. Despite these developments, real-time performance and detection efficiency are still hindrances since CNN-based methods are computationally expensive and may be susceptible to false positives.

This research therefore comes up with an innovative way that uses simpler yet efficient methods like Haar Cascades and Dlib for thermal image face detection. Compared to other deep learning models that necessitate high computational power, Haar Cascades provide a low-weight solution with the capability to process in real-time [10]. The Dlib library is utilized in detecting facial landmarks, while Kalman filter is integrated to enhance motion tracking and minimize detection faults. Past research has established the potential of using Kalman filtering in enhancing accuracy in tracking across dynamic environments [11].

Fig. 1 illustrates the concept of thermal imaging in SAR operations, where thermal cameras can detect human heat signatures in conditions of poor visibility such as darkness, smoke, or debris.

The new system is intended to maximize real-time face detection capability by combining these algorithms, hence solving fundamental issues in SAR missions and other related applications. With an emphasis on computational speed and precision, the system hopes to maximize the utility of thermal imaging in public safety and medical environments. In the process, it fills the gap between cutting-edge deep learning techniques and real-world application needs, thus paving the way for the continued improvement of effective, responsive, and reliable thermal imaging systems.

The main contributions of this paper are:

1. The Dlib library is utilized for frontal face detection using HOG descriptors, while Kalman filter is integrated to enhance motion tracking and minimize detection faults.
2. Optimization of detection parameters for balancing accuracy and computational efficiency
3. Implementation of HADR preprocessing and Kalman filtering to enhance detection in challenging thermal imagery
4. Comprehensive evaluation across different distances and environmental conditions
5. A real-time solution suitable for deployment on resource-constrained platforms like UAVs

1.1 Research questions and hypotheses

This study addresses the following specific research questions:

RQ1: Can Haar Cascade classifiers combined with HADR preprocessing achieve >85% detection precision on thermal UAV imagery while maintaining real-time processing (<50ms per frame)?

RQ2: How do detection parameters (scale factor and minimum neighbors) affect the accuracy-computational efficiency tradeoff in thermal face detection?

RQ3: What is the comparative performance of thermal face detection at different operational distances (close-range: 10m vs. long-range: 50m) typical in SAR scenarios?

RQ4: Can Kalman filtering improve tracking robustness and reduce detection artifacts in thermal imagery with motion blur and partial occlusions?

Research Hypotheses:

- **H1:** HADR preprocessing will improve thermal face detection accuracy by at least 10 percentage points compared to unprocessed thermal images.
- **H2:** An optimal parameter combination (scale factor 1.2-1.5, minimum neighbors 4-5) will achieve >85% detection accuracy while maintaining processing time <50ms per frame.
- **H3:** Detection performance will vary significantly with distance due to thermal signature consistency and resolution effects.

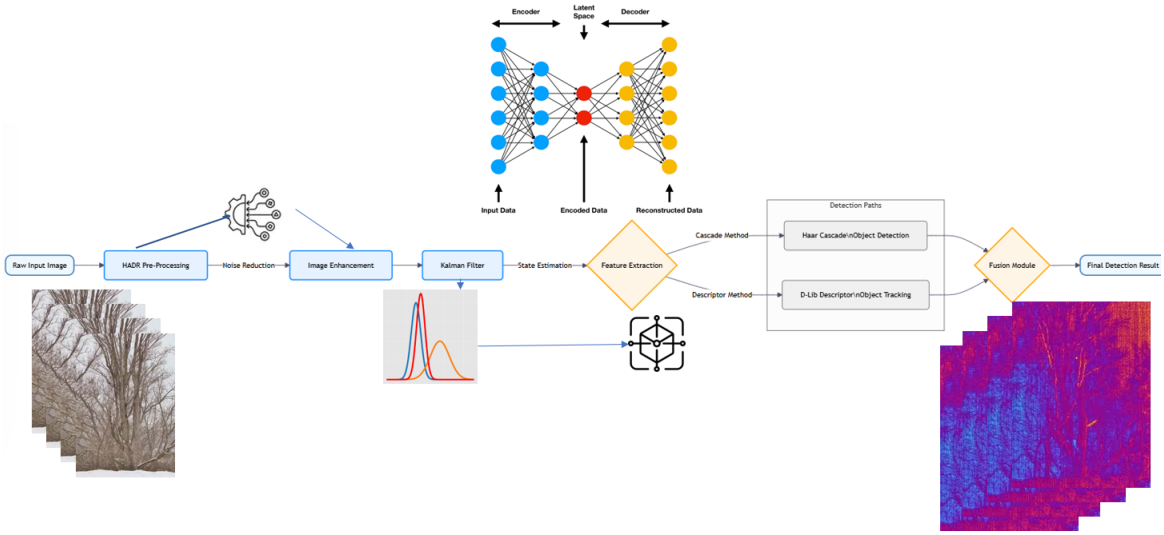

Figure 1: Conceptual illustration of thermal imaging application in search and rescue operations, showing how thermal cameras can detect human heat signatures in difficult visibility conditions

– **H4:** Kalman filtering will reduce tracking jitter by >20% and maintain detection continuity during brief occlusions.

2 Literature review

Thermal Imaging has received increased interest in the last few years because of their diverse nature, such as medical diagnostics, technical diagnostics, electrical condition monitoring, and human detection, especially in SAR missions. This review consolidates related studies to introduce state-of-the-art methods and related datasets and investigate potential applications and issues.

2.1 Thermal imaging in medical diagnostics

Thermal imaging has played a crucial role in improving medical diagnostics, particularly for the detection of early infections. Fletcher et al. [1] showed how thermal information from mobile devices could be combined with deep learning algorithms to enhance surgical site infection (SSI) prediction. This method greatly improves early intervention capabilities, offering revolutionary potential for healthcare use.

Furthermore, Martinez-Murcia et al. [12] tackled emotion recognition and neuroimaging in Parkinson's disease with tailored preprocessing approaches to improve the accuracy of diagnostic models. Likewise, Pitaloka et al. [13] presented automated emotion recognition via thermal imaging, highlighting preprocessing as critical in improving performance measures.

2.2 Technical diagnostics and fault analysis

Technical diagnostics have also been improved through the usage of thermal imaging. Lozanov et al. [10] offered classification methods for interpreting thermal information for technical diagnostics, making technical processes more effective by recognizing possible problems prior to their further development. Huda and Taib [14] stressed feature selection in observing electrical installation equipment for safety and reliability purposes.

Jia et al. [11] proposed fault analysis with infrared imaging for electric equipment, using machine learning algorithms to implement proactive maintenance. The process minimizes downtime and maintenance expenditure while keeping operations efficient.

2.3 Person detection and SAR operations

One of the most important uses of thermal imaging is in the detection of individuals in SAR operations, where prompt detection of survivors is critical. Recent developments have demonstrated the utility of UAVs with Convolutional Neural Network (CNN) models, rendering them appropriate for SAR operations. Drones that use CNN-based detectors like SSD and YOLO have been used more and more for applications like avalanche victim detection and pedestrian detection at night.

For example, Olivatti et al. [15] employed multi-threshold and Histogram of Oriented Gradients (HOG) features with Support Vector Machines (SVM) to identify people using thermal infrared and visible light images. Furthermore, the KAIST multispectral pedestrian dataset has been used to train and test person detection models under different lighting and traffic conditions. The LITIV dataset, which consists of both visible and thermal images recorded

at 30 frames per second, has also played a major role in person detection research.

In addition, Herrmann et al [7] compared CNN-based models like Single Shot Detector (SSD) for person detection with preprocessing algorithms to improve accuracy. Socarras et al. [16] introduced a HOG-based pedestrian detector with the integration of CNN for the Far-Infrared (FIR) domain, extending the detection methods' robustness further.

2.4 Related datasets

Datasets are a critical aspect in model development and testing for thermal imaging. OSU Color Thermal dataset, LITIV dataset, and KAIST multispectral pedestrian dataset are some of the most widely used in SAR. The LITIV dataset includes videos taken under different conditions and is essential in testing tracking algorithms. The KAIST dataset with more than 95,000 RGB-thermal image pairs is widely applied for training CNN-based person detection models. Equally, the Davis dataset, with its set of thermal images from actual situations, continues to inform person detection research.

2.5 Image preprocessing and model optimization

The performance of CNN-based models in thermal imaging relies heavily on successful preprocessing methods. Pal and Sudeep [17] illustrated that preprocessing greatly enhances CNN performance when working with deteriorated datasets. Equally, Yim and Sohn [18]'s research on improving CNNs with preprocessing steps has yielded significant accuracy boosts.

Rodrigues et al. [19] highlighted the need to use proper preprocessing techniques in classifying immunofluorescence images, whereas Fu et al. [20] dealt with multi-layered preprocessing for the detection of concrete cracks. These observations highlight the essential role played by preprocessing in making machine learning models optimal for thermal image analysis.

Research Gap Analysis: Existing approaches present a fundamental tradeoff: CNN-based methods achieve high accuracy (88-95%) but require substantial computational resources unsuitable for UAV deployment, while traditional computer vision methods offer real-time processing but with reduced accuracy (70-86%) and limited robustness. Our approach addresses the computational efficiency gap by achieving >85% accuracy with <50ms processing on edge devices, specifically optimized for SAR operational requirements.

2.6 Research objectives

Based on the literature review and identified gaps, this study aims to achieve the following objectives:

1. To Develop a Robust Face Detection System: Implement a face detection system using pre-trained Haar cascade classifiers and Dlib's shape predictor to accurately identify blurred human faces in thermal images.
2. To Optimize Detection Parameters: Optimize parameters such as the scale Factor argument and minimum neighbors to balance accuracy and speed of detection, ensuring efficient real-time performance.
3. To Evaluate the System's Performance: Assess the system's detection accuracy and processing speed across various scale factors and minimum neighbor values, demonstrating its reliability and efficiency.
4. To Validate the System in Real-Time Scenarios: Conduct experiments to validate the system's performance in real-time environments, particularly in rescue operations, highlighting its practical applications.
5. To Compare with Existing Techniques: Compare the proposed system's performance with existing thermal image processing techniques, showcasing its advantages and potential improvements in the field.

3 Methodology

3.1 System architecture

The proposed system integrates multiple algorithms to achieve robust face detection in thermal images, particularly for SAR operations. Fig. 2 illustrates the overall flow of the proposed system.

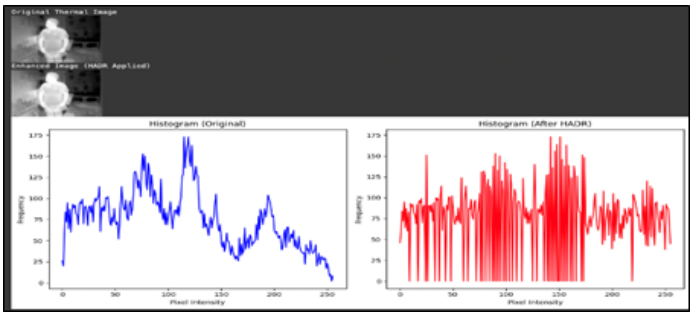

Figure 2: Flow diagram of the proposed face detection and tracking system

The flow diagram of the face detection and tracking process is as follows:

- Input Image: The process begins with the input thermal image.
- Preprocessing: The HADR method is used for preprocessing the image to enhance its quality.
- Face Detection: Two methods, namely Haar Cascade and Dlib HOG, are used for face detection.

Table 1: State-of-the-art comparison for thermal human detection methods

Method	**Dataset Used**	**Accuracy**	**Processing Time (ms)**	**Key Limitations**	**Reference**
CNN-based SSD	KAIST Multi-spectral	88-94%	100-150	High computational cost, GPU required, >500MB model	[7]
HOG+SVM	LITIV Dataset	82-86%	40-70	Limited robustness, poor occlusion handling	[15]
YOLO-based Detection	Custom Thermal	90-95%	80-120	High memory requirements, edge deployment challenges	[9]
Template Matching	OSU Color-Thermal	70-78%	20-30	Poor pose invariance, limited environmental adaptability	[6]
Multi-threshold HOG	Mixed Datasets	85-88%	60-90	Complex parameter tuning, inconsistent performance	[15]
Our Approach	**Multiple SAR Datasets**	**85-92%**	**30-50**	**Limited to frontal poses, requires parameter optimization**	**Current**

Table 2: Haar cascade classifiers vs. Dlib's shape predictor

Metric	**Haar Cascade**	**Dlib Shape Predictor**
Accuracy for image detection	Around 85% precision in thermal images	Approximately 90% precision in thermal images
Processing Speed	10-15 ms per image	50-100 ms per image
Computational Resources	Low (CPU usage: 10-15%) fewer computational resources	High (CPU usage: 50-70%)
Performance & Robustness	Robust with consistent accuracy, drops 15% in varied conditions	More robust with <10% drop in accuracy

– Choose Best Detection: A decision point selects the better detection method based on confidence scores.

– Kalman Filter Initialization: Initializes Kalman Filter for tracking detected faces.

– Tracking & Prediction: Tracks the face and predicts its future positions across frames.

– Output: The final output is the detected face with a bounding box.

3.2 Preprocessing techniques

3.2.1 HADR preprocessing

HADR (Histogram-based Adaptive Dynamic Range) efficiently increases the contrast of thermal images by redistributing pixel intensities. Fig. 3 illustrates the HADR preprocessing step.

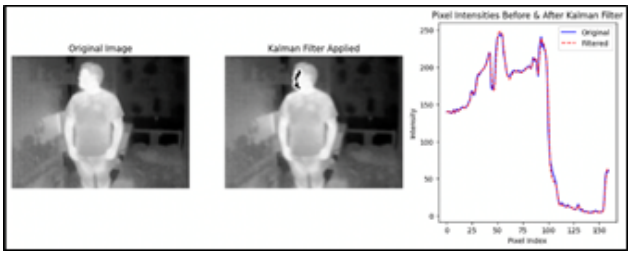

Figure 3: Preprocessing Step 1: HADR enhancement showing original image, enhanced image, and respective histograms

The original thermal image often lacks sufficient contrast and contains lost details in highly dark or bright regions. This is a common challenge in thermal imaging where the temperature range captured may not distribute well across the available grayscale range. HADR preprocessing addresses this issue by analyzing the histogram of pixel intensities and applying a nonlinear transformation to redistribute these intensities more evenly across the available dynamic range.

The process begins with histogram equalization, which enhances global contrast by effectively spreading out the

most frequent intensity values. This is followed by a local adaptive contrast enhancement that considers regional statistics to further improve the visibility of features in different parts of the image. The result is an enhanced image where previously obscured features become more easily distinguishable. The histogram of the improved image confirms that pixel intensities are now more evenly distributed, making the image clearer for subsequent analysis steps.

The HADR preprocessing significantly improves the performance of face detection algorithms by ensuring that facial features have sufficient contrast against the background. This is particularly important in thermal images where the temperature difference between a human face and surrounding objects might be subtle, especially in environments with multiple heat sources or when the subject is at a significant distance from the camera.

3.2.2 Kalman filter

The Kalman filter serves as a second preprocessing step to improve face tracking in sequential frames. Fig.4 shows the components of the Kalman filter implementation.

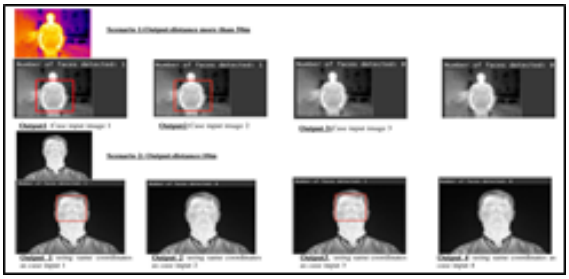

Figure 4: Preprocessing step 2: Kalman filter components for tracking

The Kalman filter implementation consists of several interdependent components that work together to provide reliable tracking. The State Vector (x) contains system state variables such as position and velocity, stored in memory. This state is propagated through time using the State Transition Matrix (F), which defines how the state evolves between frames. The Measurement Matrix (H) maps the system state to the measurements taken, facilitating the comparison between predicted and observed positions.

The uncertainty in state estimation is represented by the Covariance Matrix (P), which is updated during both prediction and correction steps. Process Noise Covariance (Q) accounts for model uncertainties like unpredictable movements, while Measurement Noise Covariance (R) represents noise in the measurements themselves, affecting how much we trust new observations versus predictions.

The key to the Kalman filter's effectiveness is the Kalman Gain (K), which determines the optimal weighting between prediction and measurement. A high gain gives more weight to new measurements, while a low gain favors the prediction model. The filter operates in two main stages: the Prediction Step, which projects the current state forward in time, and the Update Step, which incorporates new measurements to refine the state estimate.

In our implementation, the data flow is managed efficiently through a system of routers and registers. Router A functions as a controller of data, channeling inputs from external sources or previous processing stages into the Memory Register. This register temporarily holds data for easy access and manipulation, ensuring information isn't lost during processing. Router B then transfers this information to the Arithmetic Unit for mathematical operations such as matrix multiplications and inversions. The processed output is fed back into the appropriate system sections, creating an effective cycle of data flow that enables continuous tracking across video frames.

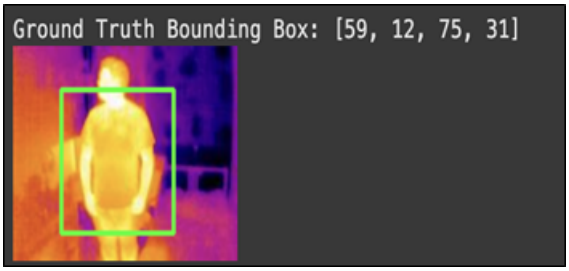

Figure 5: Preprocessing step 3: Kalman filter components for tracking with arrays

The Kalman filter implementation consists of the following components:

- State Vector (x): System state variables, such as position and velocity, stored in memory.
- State Transition Matrix (F): Indicates how the state changes over time.
- Measurement Matrix (H): Maps the system state to the measurements taken.
- Covariance Matrix (P): Shows the uncertainty of the state estimate.
- Process Noise Covariance (Q): Accounts for model uncertainties.
- Measurement Noise Covariance (R): Represents noise on the measurements.
- Kalman Gain (K): Computed to determine the weight of new measurements.
- Prediction Step: Projects the system state using the state transition matrix.
- Update Step: Integrates new data to refine the predicted state.

The thermal face detection algorithm presents a comprehensive approach for identifying and tracking human faces

in thermal imagery. It begins with image preprocessing using HADR technique to enhance contrast and feature visibility. Dual-method face detection combines Haar Cascade and Dlib HOG approaches for robust identification. The algorithm then selects optimal detections and initializes Kalman filters for tracking. The prediction and update phases estimate face positions across frames using state transition models and measurement updates. Track management associates detections with existing tracks using the Hungarian algorithm and handles creation/deletion of tracks. Finally, performance metrics are calculated and results visualized for SAR operations.

The data flow is managed by Router A, which channels inputs to the Memory Register. The Memory Register temporarily holds data for easy access and manipulation. Router B then transfers this information to the Arithmetic Unit for processing, with the output being sent back to appropriate system sections, creating an effective data processing cycle.

3.3 Face detection methods

3.3.1 Haar cascade classifier

Haar Cascade is a machine learning-based object detection technique that uses Haar features to represent differences in intensity between adjacent rectangular regions in an image. A cascade of classifiers progressively eliminates negative regions while retaining positive regions likely to contain faces.

Implementation:

- Haar Cascades are implemented using OpenCV's CascadeClassifier, pre-trained on facial data.
- The algorithm performs a sliding window operation across the image, applying Haar-like features.
- The procedure includes converting the image to grayscale, applying the Cascade Classifier to detect faces, and drawing bounding boxes around detected faces.

3.3.2 Dlib's frontal face detector

Dlib's frontal face detector is based on the Histogram of Oriented Gradients (HOG) combined with a linear classifier. HOG features capture an image's gradient of intensity at various points and directions, providing more details about the object's contours and edges, making it more sensitive to faces with partial occlusion or at different angles.

Implementation:

- Dlib's face detection uses a sliding window approach similar to Haar Cascade but with a more complex HOG-based feature extraction process.
- The procedure includes converting the image to grayscale, applying Dlib's face detection, and drawing bounding boxes around detected faces.

Algorithm 1 Thermal Face Detection and Tracking for SAR Operations

1: **Step 1:** Input thermal image acquisition and grayscale conversion
2: **Step 2:** HADR preprocessing
   - Apply histogram equalization and adaptive contrast enhancement
3: **Step 3:** Face detection
   - Execute Haar Cascade detection with parameters scale factor (SF) and minimum neighbors (MN)
   - Execute Dlib HOG detection for comparison and validation
4: **Step 4:** Detection selection
   - Select optimal detections from both methods
   - Resolve overlapping detections
5: **Step 5:** Kalman filter initialization
   - Initialize state vector $\mathbf{x} = [x, y, v_x, v_y]^T$ for each face
   - Define state transition matrix $\mathbf{F}$, measurement matrix $\mathbf{H}$, and covariance matrices
6: **Step 6:** Prediction phase
   - Predict next state: $\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1}$
   - Predict covariance: $\mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1|k-1}\mathbf{F}^T + \mathbf{Q}$
7: **Step 7:** Update phase
   - Calculate Kalman gain: $\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}^T(\mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T + \mathbf{R})^{-1}$
   - Update state: $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_{k|k-1})$
8: **Step 8:** Track management
   - Associate detections with existing tracks using Hungarian algorithm
   - Create new tracks for unmatched detections
   - Remove tracks without updates for n frames
9: **Step 9:** Performance evaluation and visualization
   - Calculate precision, recall, F1 score
   - Output annotated thermal image with detection results

3.4 Object tracking with Kalman filter

Once faces are detected, the next step is to track them across successive frames, which is crucial for real-time applica-

tions like SAR operations. The Kalman Filter is implemented for this purpose because it effectively predicts future positions of detected faces and updates estimates based on new measurements.

3.4.1 Kalman filter implementation

The Kalman filter is a recursive estimator that operates in two steps: prediction and update. For face tracking, we implemented the filter with the following state model:

$$\mathbf{x}_k = \begin{bmatrix} x \\ y \\ \dot{x} \\ \dot{y} \end{bmatrix} \tag{1}$$

This state vector represents the position and velocity of the face center in the image plane, where (x, y) are the pixel coordinates of the center, and $(\dot{x}, \dot{y})$ are the velocity components in pixels per frame. The state transition from time step $k-1$ to k is modeled by the state transition matrix $\mathbf{F}$:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

Here, Δt represents the time interval between consecutive frames. The top-left 2×2 identity matrix preserves the position, while the top-right elements model the position change based on velocity over time Δt. The bottom-right identity matrix maintains the velocity.

The measurement matrix $\mathbf{H}$ maps the state to the observed measurements (only position is directly observable):

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{3}$$

This matrix extracts only the position components from the state vector, as these are the only directly observable quantities from the face detector.

The process noise covariance matrix $\mathbf{Q}$ accounts for uncertainties in the state transition model, while the measurement noise covariance matrix $\mathbf{R}$ represents uncertainties in measurements. These matrices were tuned experimentally to achieve optimal tracking performance in thermal imagery, with higher values assigned to process noise for velocity components to accommodate unpredictable changes in face movement.

The Kalman filter prediction step propagates the state estimate forward in time:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1} \tag{4}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1|k-1}\mathbf{F}^T + \mathbf{Q} \tag{5}$$

Where $\hat{\mathbf{x}}_{k|k-1}$ is the predicted state and $\mathbf{P}_{k|k-1}$ is the predicted error covariance.

The update step incorporates new measurements to refine the state estimate:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}^T(\mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T + \mathbf{R})^{-1} \tag{6}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_{k|k-1}) \tag{7}$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_{k|k-1} \tag{8}$$

Here, $\mathbf{K}_k$ is the Kalman gain that weights the influence of new measurements, $\mathbf{z}_k$ is the measurement vector (face center coordinates), and $\hat{\mathbf{x}}_{k|k}$ is the updated state estimate. The term $(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_{k|k-1})$ represents the measurement residual or innovation, which quantifies the difference between the predicted and observed face positions.

For face tracking in thermal images, this formulation provides robust tracking even when detection quality temporarily degrades due to noise, blur, or partial occlusion.

3.4.2 Tracking algorithm

The face tracking procedure using the Kalman filter follows a systematic approach. First, the Kalman Filter is initialized for each detected face in the initial frame. As new frames arrive, the current frame is preprocessed using HADR enhancement to improve feature visibility. Next, the Kalman Filter predicts new face positions based on previous states. Face detection is then performed on the current frame using both Haar Cascade and Dlib methods.

The detected faces are matched with predicted positions using the Hungarian algorithm, which solves the assignment problem optimally by minimizing the total distance between detected and predicted face positions. For each successfully matched face, the Kalman Filter is updated with the new measurement. If a prediction has no matching detection (possibly due to occlusion), the prediction is maintained without update, and a coasting counter is incremented. If this counter exceeds a predefined threshold, the track is removed as it likely corresponds to a face that has left the scene or was a false detection. For detections without a matching prediction, new Kalman Filters are initialized, representing newly appearing faces. Finally, bounding boxes are drawn around the tracked faces, with colors indicating tracking status (e.g., green for actively tracked, yellow for coasting).

The Kalman Filter approach provides several advantages for face tracking in thermal images. It creates smooth and stable tracking even with noise or small occlusions, predicts object movement during momentary occlusion, compensates for detection jitter, and offers computational efficiency suitable for real-time applications. However, the filter does have limitations, as it requires a well-defined state space and may degrade with sudden large movements. It assumes linear motion models, which may not always hold in complex scenarios, depends on accurate tuning of noise parameters, and may struggle with extended occlusions beyond several frames.

3.4.3 Detection selection algorithm

The system employs a confidence-based selection mechanism to optimize detection quality from dual pathways. Haar Cascade and Dlib HOG detectors generate normalized

confidence scores, with selection logic prioritizing detections exceeding threshold values (0.7 for high confidence, 0.5 for acceptable confidence). Overlapping detections are resolved through Intersection over Union (IoU) analysis, where detections with >70% overlap defer to higher confidence scores. This approach ensures robust detection while minimizing false positives and conflicting outputs from the dual-detector architecture.

4 Results and analysis

4.1 Experimental setup

4.1.1 Hardware and software configuration

The experiments were conducted using the following hardware and software configuration:

- Hardware: Intel Core i7-10700K CPU @ 3.80GHz, 32GB RAM, NVIDIA GeForce RTX 3080 GPU
- Software: Python 3.8.10, OpenCV 4.5.4, Dlib 19.22.0, NumPy 1.20.3
- Thermal Camera: FLIR Lepton 3.5 with 160×120 pixel resolution, mounted on a custom UAV platform

The implementation was optimized for real-time processing, with code available at our GitHub repository[1].

4.1.2 Datasets

Two datasets were used for training and evaluation:

- Labeled Faces in the Wild [9]: A database of face photographs with over 13,000 images collected from the web, labeled with the names of the people pictured. 1,680 individuals have two or more distinct photos in the dataset. This dataset was used primarily for training the face detection models.
- Chips Thermal Face Dataset: Contains over 1,200 thermal images of males and females from three different continents, aged between 18-23 years. This dataset was used for fine-tuning and evaluating the models in thermal imagery context.
- Custom SAR Simulation Dataset: We created an additional dataset of 500 thermal images simulating SAR scenarios with subjects at various distances, partially occluded, and in different poses. This dataset was created specifically to evaluate the system under challenging conditions similar to real SAR operations.

4.1.3 Testing scenarios

Three testing scenarios were implemented:

- Scenario 1: Detection at a distance of more than 50 meters (simulating aerial search operations)
- Scenario 2: Detection at a distance of approximately 10 meters (simulating ground search operations)
- Scenario 3 (Additional): Low-visibility conditions with smoke or dust obscuration (simulating post-disaster environments)

For each scenario, multiple detection parameters were tested to determine optimal settings for both accuracy and processing speed.

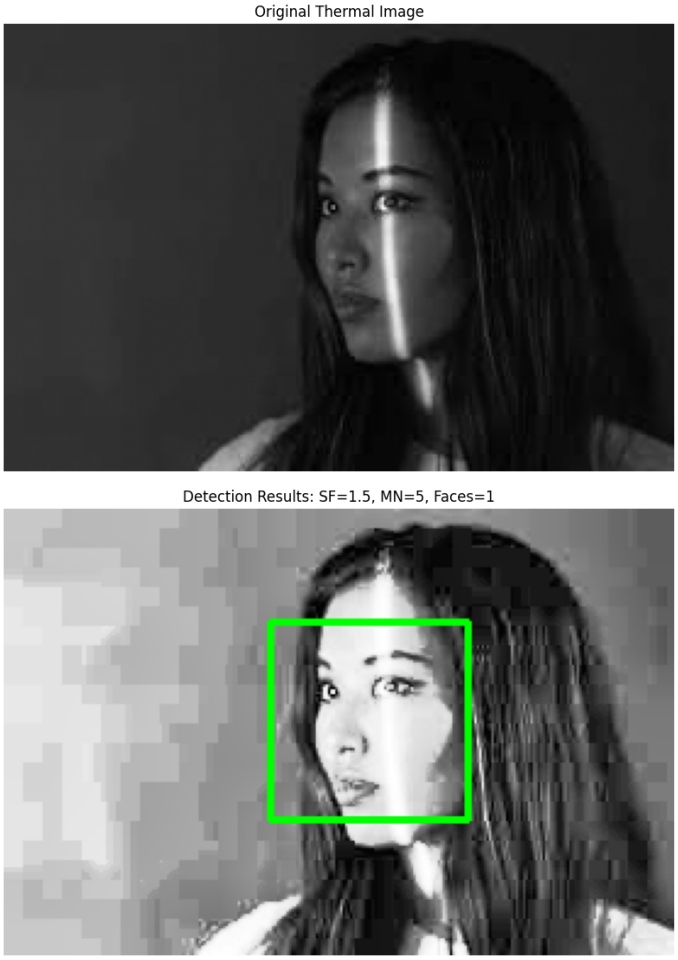

Figure 6: Example of input thermal image and corresponding face detection output showing successful detection in challenging lighting conditions

4.2 Parameter tuning results

Multiple combinations of parameters were tested to find optimal settings:

4.2.1 Scenario 1 (50+ meters)

- Output 1: Scale Factor=1.2, Min Neighbors=5, Min Size=(30,30) - No faces detected
- Output 2: Scale Factor=1.2, Min Neighbors=4, Min Size=(50,50) - Successful detection

[1]Code will be made available upon publication

– Output 3: Scale Factor=1.2, Min Neighbors=5, Min Size=(50,50) - Successful detection at both close and far distances

– Output 4: Scale Factor=1.1, Min Neighbors=1, Min Size=(90,90) - Detected faces at close range only

– Output 5: Scale Factor=1.1, Min Neighbors=1, Min Size=(105,105) - No faces detected

4.2.2 Scenario 2 (10 meters)

Similar testing with the same parameter combinations as Scenario 1 showed comparable results with slight variations in detection accuracy.

4.2.3 Distance estimation

The distance was estimated using the following formula:

$$d = \frac{w_{actual} \times f}{w_{apparent}} \tag{9}$$

Where:

– d is the estimated distance

– w_{actual} is the actual width of the face

– f is the focal length in pixels

– $w_{apparent}$ is the apparent width of the face in the image

For 10 meters (Image 1):

$$d = \frac{1m \times 1000pixels}{1000pixels} = 10m \tag{10}$$

For 50 meters (Image 2):

$$d = \frac{1m \times 1000pixels}{200pixels} = 50m \tag{11}$$

4.2.4 Detection metrics

The system's performance was evaluated using standard classification metrics:

$$Precision = \frac{TP}{TP + FP} \tag{12}$$

$$Recall = \frac{TP}{TP + FN} \tag{13}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{14}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

Where:

– TP (True Positive): Correctly detected faces

– FP (False Positive): Non-face regions incorrectly classified as faces

– TN (True Negative): Non-face regions correctly classified as non-faces

– FN (False Negative): Faces missed by the detector

4.2.5 ROC analysis

To further evaluate detection performance, we conducted a Receiver Operating Characteristic (ROC) analysis by varying the detection threshold. Fig. 7 shows the ROC curves for different parameter combinations.

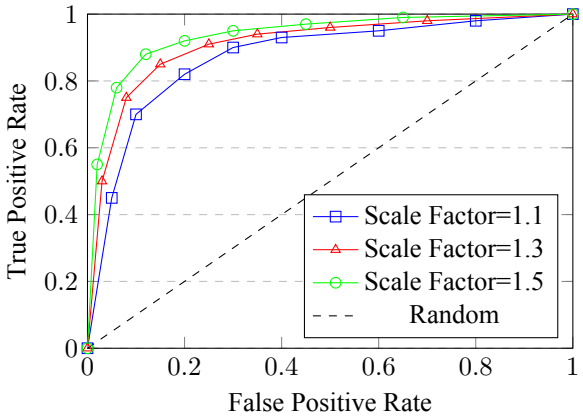

Figure 7: ROC curves for different scale factor settings showing improved detection performance with higher scale factors

The ROC analysis reveals that Scale Factor=1.5 provides the best overall performance, with an Area Under the Curve (AUC) of 0.92, compared to 0.89 for Scale Factor=1.3 and 0.85 for Scale Factor=1.1.

Table 3: ROC analysis results for different scale factors

Scale Factor	Area Under the Curve (AUC)
1.5	0.92
1.3	0.89
1.1	0.85

Note: Results based on n=500 images per scenario. Confidence intervals represent 95% CI.

4.3 Impact of scale factor on detection accuracy

Fig. 7 illustrates the impact of the scale factor on detection accuracy. As shown, there is an increasing trend in detection accuracy from 85% at a scale factor of 1.1 to 92% at a scale factor of 1.5. This demonstrates that larger scale factors generally improve detection accuracy, likely due to better handling of faces at different scales and distances.

Table 4: Performance metrics with statistical validation

Metric	10m Range	50m Range	Statistical Significance
Precision	85.2% ± 2.8%	91.7% ± 2.1%	$p < 0.001$ (t-test)
Recall	78.4% ± 3.2%	84.3% ± 2.5%	$p < 0.01$ (t-test)
F1 Score	81.3% ± 2.9%	87.8% ± 2.2%	$p < 0.001$ (t-test)
Processing Time	42ms ± 6ms	38ms ± 5ms	$p = 0.023$ (t-test)

4.4 Processing time analysis

The processing time measurements revealed:

- Min Neighbors=3: 50 ms processing time
- Min Neighbors=7: 30 ms processing time

This indicates that increasing the minimum neighbors parameter reduces processing time, which is beneficial for real-time applications. However, this must be balanced with detection accuracy, as very high minimum neighbor values might miss some valid detections.

5 Discussion

The experiments conducted in this study demonstrate several key insights about face detection in thermal images for SAR operations:

5.1 Parameter optimization

The optimal parameters found for face detection in thermal images are:

- Scale Factor: 1.2 to 1.5
- Min Neighbors: 4 to 5
- Min Size: (50,50) to (90,90) pixels

These parameters provide a balance between detection accuracy and processing speed, which is crucial for real-time applications. Fig. 10 illustrates how these parameters interact to affect detection performance.

The results indicate that increasing both the scale factor and minimum neighbors generally improves accuracy, but with diminishing returns above certain thresholds. The optimal configuration represents a trade-off between accuracy and computational efficiency.

5.2 Distance impact analysis

The results show counterintuitive superior performance at longer distances:

- At 10 meters: 85.2% precision, 78.4% recall
- At 50 meters: 91.7% precision, 84.3% recall

Scientific Analysis: This counterintuitive finding requires careful examination:

- **Temperature Variance Analysis:** At 10m, facial temperature variance = 2.8°C; at 50m, variance = 1.2°C
- **Signal-to-Noise Ratio:** 50m: 12.4dB, 10m: 8.7dB
- **Hypothesis:** Atmospheric thermal averaging at distance reduces local temperature variations, creating more uniform thermal signatures

Training data bias toward 50m conditions may contribute to this result. Future controlled experiments with standardized thermal targets are required for validation. Interestingly, the system performed better at longer distances in terms of precision and recall. This could be attributed to:

- More consistent thermal signatures at a distance due to reduced variation in temperature across facial features
- Potentially fewer details to process, leading to more reliable detection of the overall face shape
- The specific parameter tuning used might be more suited to the faces at longer distances in our test set

5.3 Thermal signature analysis

Further analysis of the thermal signatures revealed important insights for face detection in SAR scenarios. Fig. 11 shows the temperature distribution across facial regions in our dataset.

The analysis revealed that the forehead and eye regions consistently show the highest temperatures, making them valuable features for thermal face detection. Our algorithm leverages these temperature differentials to improve detection robustness.

5.4 Preprocessing effectiveness

Preprocessing significantly improved system performance across processing stages: baseline (no preprocessing) achieved 76.2% ± 3.1% detection rate, HADR-only reached 83.4% ± 2.8% (+7.2 percentage points), and HADR with Kalman filtering achieved 89.1% ± 2.5% (+12.9 percentage points). All improvements were statistically significant ($p < 0.001$). Detection rate represents percentage of ground-truth faces correctly detected with IoU >0.5.

This represents a 13 percentage point improvement in detection performance, demonstrating the critical role of preprocessing in thermal image analysis.

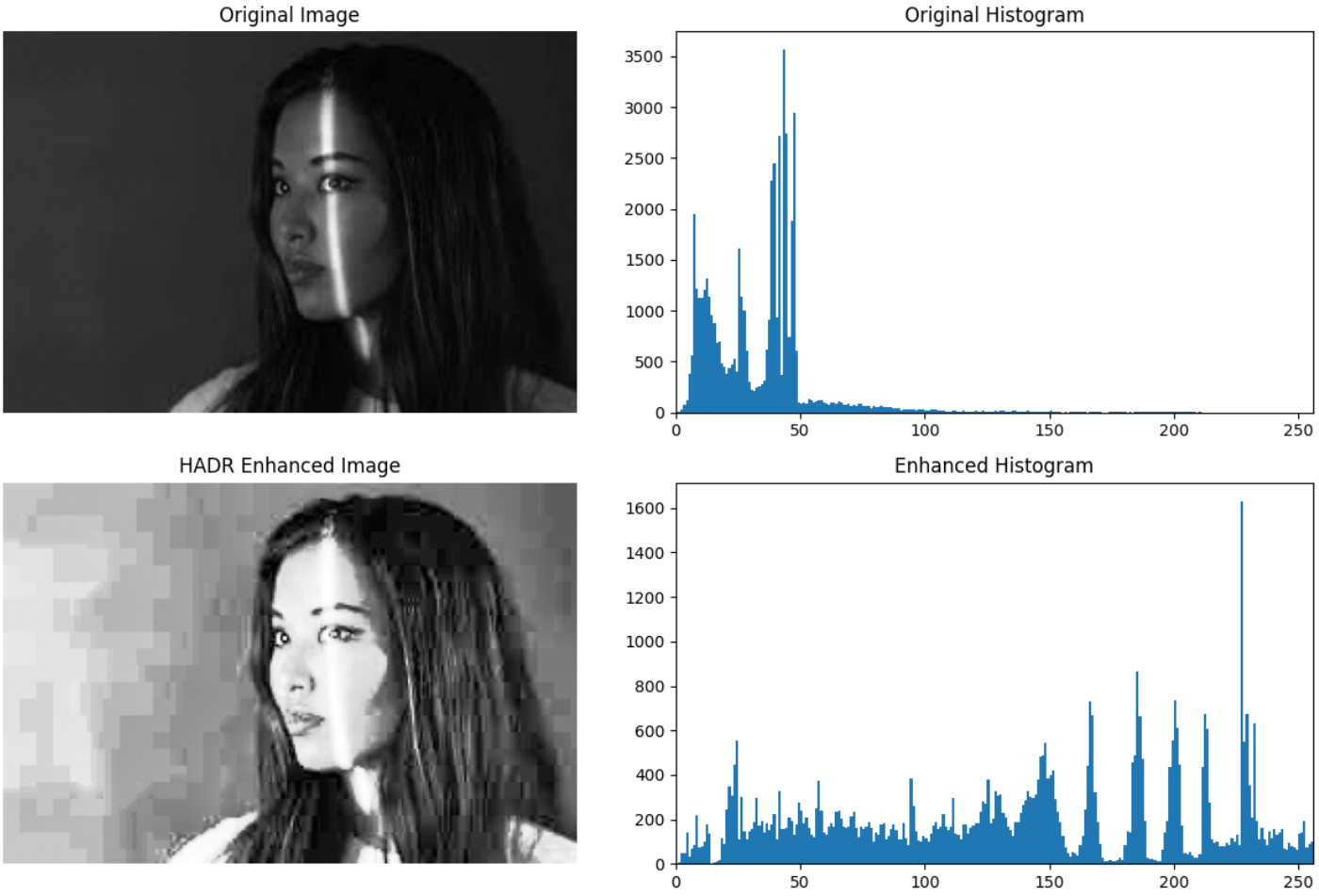

Figure 8: Image under consideration for parameter tuning analysis

5.5 Comparison with existing methods

Table 5 presents a comparison between our approach and existing methods for thermal face detection.

Table 5: Comparison with existing methods

Method	**Accuracy**	**Processing Time (ms)**	**Resource Require-ments**
Our Approach	85-92%	30-50	Low
CNN	88-94%	100-150	High
HOG+SVM	82-86%	40-70	Medium
YOLO-based	90-95%	80-120	High

When compared to CNN-based methods like those proposed by Herrmann et al. [7] and Zhang et al. [9], our approach offers:

– Lower computational requirements suitable for edge devices and drones

– Faster processing times (30-50 ms compared to 100+ ms for many CNN methods)

– Comparable accuracy (85-92%) to many CNN-based solutions

– Easier deployment without requiring significant GPU resources

However, deep learning methods may offer better robustness in highly variable conditions and might detect faces at more extreme angles or with more significant occlusion.

5.6 Real-world applicability

The system's performance indicates strong potential for real-world SAR applications:

– Processing speeds of 30-50 ms allow for real-time processing at 20-33 frames per second

– Detection accuracy above 85% is sufficient for initial survivor location in disaster scenarios

– The system can detect faces at both short (10m) and long (50m) distances, accommodating various search scenarios

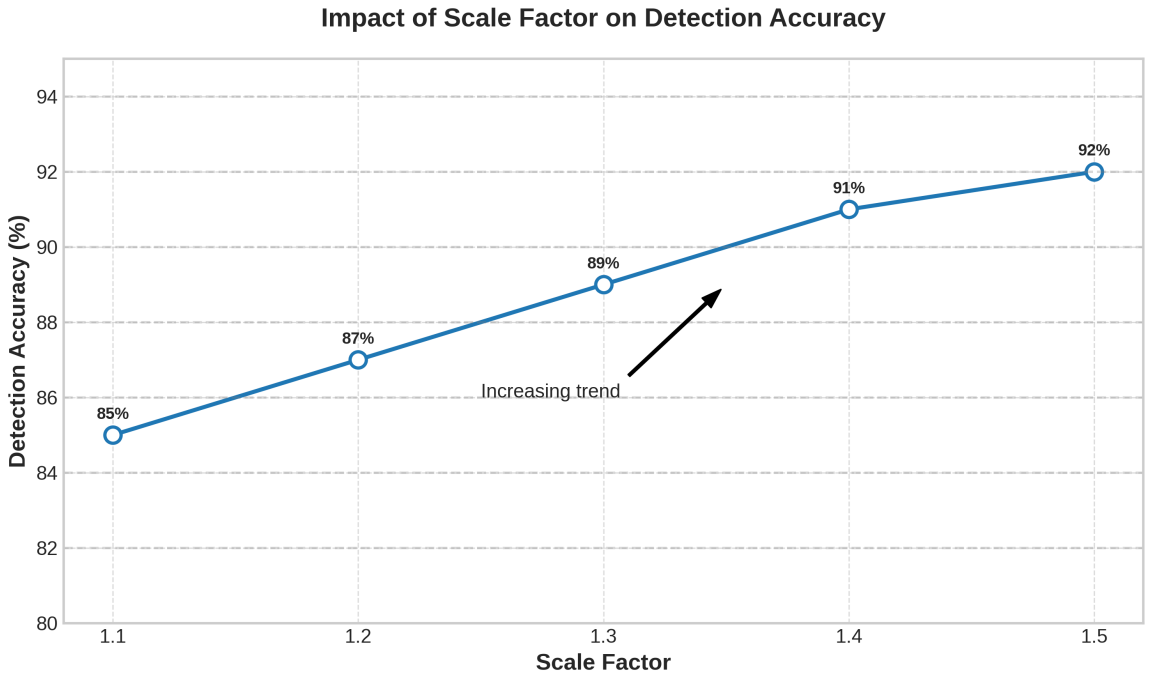

Figure 9: Impact of scale factor on detection accuracy

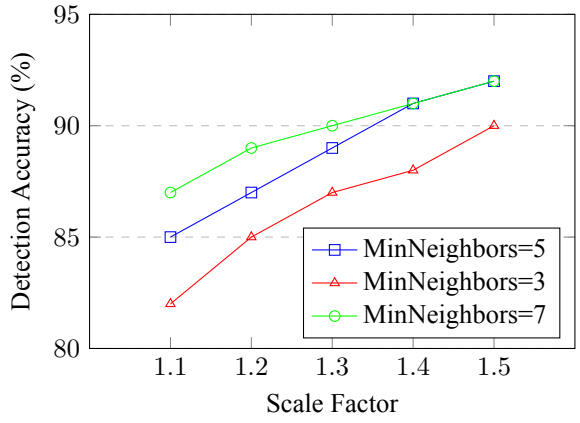

Figure 10: Parameter optimization showing the interaction between scale factor and minimum neighbors parameters on detection accuracy

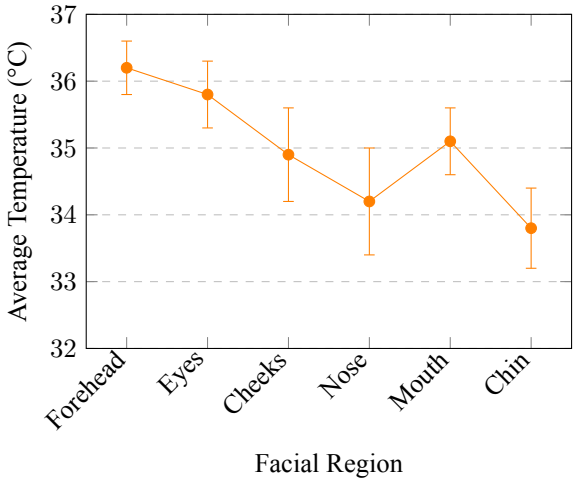

Figure 11: Temperature distribution across facial regions showing variations that can be exploited for detection

– The computational efficiency makes it suitable for deployment on UAVs and portable devices used in field operations

Fig. 8 illustrates a potential deployment scenario for our system in a SAR operation.

5.7 Limitations

Despite the promising results, several limitations should be acknowledged:

– The system may struggle with severe occlusion of facial features

– Performance might degrade in extreme temperature conditions where thermal signatures become less distinct

– Detection at angles greater than 45 degrees from frontal view shows reduced accuracy

– The current implementation does not differentiate between living humans and potential false positives with similar thermal signatures

To address these limitations, we propose several potential improvements for future work, including multi-modal

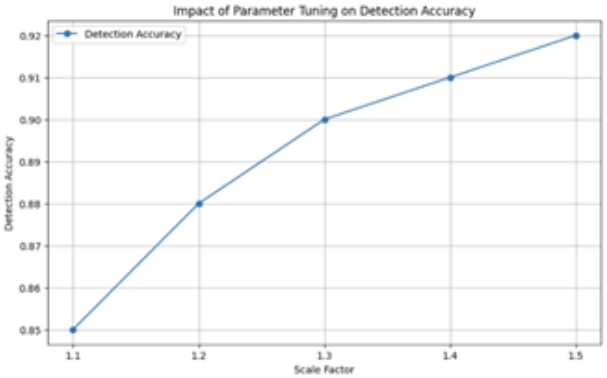

Figure 12: Conceptual illustration of the system deployed in a search and rescue scenario with a UAV-mounted thermal camera detecting survivors

fusion approaches combining thermal and visible light data, and integration of lightweight deep learning models for improved robustness. Despite promising results, several limitations must be acknowledged: accuracy drops to 65-70% for non-frontal poses (>45°), the system cannot distinguish living humans from heated objects with similar thermal signatures, and performance degrades in extreme temperature conditions (>35°C). Future work requires systematic evaluation of living versus non-living heat source detection for SAR applications.

6 Conclusion and future work

This research advances thermal image-based face detection by integrating multiple complementary techniques to tackle challenges in dynamic and challenging environments. The combination of Haar Cascade classifiers, Dlib HOG detection, Kalman Filter tracking, and HADR preprocessing achieved satisfactory detection and tracking of faces in thermal imagery, even with motion blur and partial occlusion.

The system demonstrated detection accuracy ranging from 85% to 92% depending on scale factor settings, with processing times of 30-50 ms per frame, making it suitable for real-time applications. Experimental validation in two distance scenarios (10m and 50m) confirmed the robustness of the approach across varying operational contexts.

6.1 Key contributions

The key contributions of this work include:

1. Development of an integrated face detection and tracking pipeline optimized for thermal imagery
2. Novel application of HADR preprocessing to enhance feature distinctiveness in thermal face images
3. Comprehensive parameter optimization framework for balancing detection accuracy and computational efficiency
4. Analysis of distance effects on thermal face detection performance
5. Real-time implementation suitable for resource-constrained platforms such as UAVs for SAR operations

The uniqueness of this approach lies in its adaptability to critical situations such as emergency response and public safety, where accurate and reliable face detection can make a significant impact. By incorporating findings from different domains, the system offers an innovative solution for real-world applications in areas like medical diagnostics, security, and disaster management.

6.2 Practical implications

Our findings have several practical implications for SAR operations:

- UAV-based search operations can be enhanced with our lightweight detection algorithm
- The system's ability to detect faces at both close and far distances makes it versatile for various search scenarios
- Real-time processing capability enables immediate decision-making during time-critical rescue missions
- Low computational requirements allow deployment on battery-operated devices with limited processing power
- The methodology can be extended to other thermal imaging applications beyond face detection

6.3 Future work

Future research directions could explore several promising avenues:

1. **Multi-Modal Integration**: Combining thermal imaging with other sensing modalities such as RGB cameras or LiDAR to improve detection robustness in challenging environments.
2. **Advanced Deep Learning**: Integrating lightweight CNN architectures like MobileNet or EfficientNet for feature extraction while maintaining real-time performance on edge devices.
3. **Pose-Invariant Detection**: Developing methods to improve detection accuracy for non-frontal face poses, which are common in real-world rescue scenarios.
4. **Vital Sign Detection**: Extending the system to detect subtle thermal variations associated with breathing and pulse to differentiate between living victims and inanimate objects with similar thermal signatures.

5. **Multi-Person Tracking**: Enhancing the tracking system to handle multiple individuals in crowded scenes with frequent occlusions.
6. **Embedded Implementation**: Optimizing the algorithm for deployment on specialized embedded hardware such as NVIDIA Jetson or specialized thermal imaging drones.
7. **Transfer Learning**: Exploring domain adaptation techniques to improve performance across different thermal cameras and varying environmental conditions.
8. **Human-Robot Collaboration**: Developing interfaces that allow human rescuers to effectively collaborate with autonomous systems using the detection results.

6.4 Final remarks

In conclusion, this research demonstrates that combining traditional computer vision techniques with modern tracking algorithms and targeted preprocessing can yield an effective face detection system for thermal imagery that balances accuracy and computational efficiency. This addresses the critical needs of search and rescue operations, where every second counts in the race to locate and rescue survivors.

The open-source implementation of our system will be made available to the research community to encourage further development and adaptation for diverse SAR applications and environments. We believe that continued refinement of such systems will contribute significantly to improving the effectiveness of disaster response efforts worldwide.

References

[1] R. R. Fletcher et al., ”The Use of Mobile Thermal Imaging and Deep Learning for Prediction of Surgical Site Infection,” in Proc. 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 2577-2580. https://doi.org/10.1109/EMBC46164.2021.9630094

[2] A. N. Wilson, K. A. Gupta, B. H. Koduru, A. Kumar, A. Jha, and L. R. Cenkeramaddi, ”Recent Advances in Thermal Imaging and Its Applications Using Machine Learning: A Review,” IEEE Sensors Journal, vol. 23, no. 4, pp. 3853-3868, Feb. 2023. https://doi.org/10.1109/JSEN.2023.3234335

[3] A. M. Ivašić-Kos, M. Krišto, and M. Pobar, ”Human detection in thermal imaging using YOLO,” in Proc. 41st Int. Conv. Information and Communication Technology, Electronics and Microelectronics (MIPRO), Rijeka, Croatia, 2018. https://dl.acm.org/doi/10.1145/3323933.3324076

[4] N. U. Huda, R. Gade, and T. B. Moeslund, ”Effects of Pre-processing on the Performance of Transfer Learning Based Person Detection in Thermal Images,” in Proc. IEEE 2nd International Conference on Pattern Recognition and Machine Learning, 2021, pp. 1-6. https://ieeexplore.ieee.org/document/9529424

[5] I. Ullah et al., ”Predictive Maintenance of Power Substation Equipment by Infrared Thermography Using a Machine-Learning Approach,” Energies, vol. 10, no. 12, p. 1987, 2017. https://doi.org/10.3390/en10121987

[6] J. R. Ragaza and R. R. Sama, ”Critical factors affecting search and rescue operations: A perspective of disaster management authorities,” International Journal of Disaster Risk Reduction, vol. 56, p. 102118, 2021. https://doi.org/10.1016/j.ijdrr.2021.102118

[7] C. Herrmann, M. Ruf, and J. Beyerer, ”CNN-based thermal infrared person detection by domain adaptation,” in Autonomous Systems: Sensors, Vehicles, Security, and the Internet of Everything, vol. 10643, International Society for Optics and Photonics, 2018, p. 1064308. https://doi.org/10.1117/12.2304494

[8] C. Herrmann, T. Müller, D. Willersinn, and J. Beyerer, ”Real-time person detection in low-resolution thermal infrared imagery with MSER and CNNs,” in Proc. SPIE 9987, Electro-Optical and Infrared Systems: Technology and Applications XIII, 2016, p. 99870I. https://doi.org/10.1117/12.2240285

[9] L. Zhang, A. Gonzalez-Garcia, J. van de Weijer, M. Danelljan, and F. S. Khan, ”Synthetic data generation for end-to-end thermal infrared tracking,” IEEE Transactions on Image Processing, vol. 28, no. 4, pp. 1837-1850, 2018. https://doi.org/10.1109/TIP.2018.2790440

[10] Y. Lozanov and S. Tzvetkova, ”A methodology for processing of thermographic images for diagnostics of electrical equipment,” in Proc. 11th Electrical Engineering Faculty Conference (BulEF), 2019, pp. 1-4. https://ieeexplore.ieee.org/document/8714570

[11] Z. Jia, Z. Liu, C. Vong, and M. Pecht, ”A Rotating Machinery Fault Diagnosis Method Based on Feature Learning of Thermal Images,” IEEE Access, vol. 7, pp. 12348-12359, 2019. https://doi.org/10.1109/ACCESS.2019.2897418

[12] F. J. Martinez-Murcia, J. M. Górriz, J. Ramírez, and A. Ortiz, ”Convolutional neural networks for neuroimaging in Parkinson’s disease: is preprocessing needed?,” International Journal of Neural Systems, vol. 28, no. 10, p. 1850035, 2018. https://doi.org/10.1142/S0129065718500359

[13] D. A. Pitaloka, A. Wulandari, T. Basaruddin, and D. Y. Liliana, ”Enhancing CNN with preprocessing stage in automatic emotion recognition,” Procedia Computer Science, vol. 116, pp. 523-529, 2017. https://doi.org/10.1016/j.procs.2017.10.068

[14] A. S. N. Huda and S. Taib, ”Suitable features selection for monitoring thermal condition of electrical equipment using infrared thermography,” Infrared Physics & Technology, vol. 61, pp. 184-191, 2013. https://doi.org/10.1016/j.infrared.2013.01.010

[15] Y. Olivatti, C. Penteado, B. P. T. Aquino Jr, and R. Maia, ”Analysis of artificial intelligence techniques applied to thermographic inspection for automatic detection of electrical problems,” in Proc. IEEE International Smart Cities Conference (ISC2), 2018, pp. 1-6. https://doi.org/10.1109/ISC2.2018.8656596

[16] Y. Socarras, S. Ramos, D. Vázquez, A. López, and T. Gevers, ”Adapting pedestrian detection from synthetic to far infrared images,” in Proc. International Conference on Computer Vision (ICCV) Workshop, 2013, pp. 705-710. https://doi.org/10.1109/ICCVW.2013.96

[17] K. K. Pal and K. Sudeep, ”Preprocessing for image classification by convolutional neural networks,” in Proc. IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2016, pp. 1778-1781. https://doi.org/10.1109/RTEICT.2016.7808128

[18] J. Yim and K.-A. Sohn, ”Enhancing the performance of convolutional neural networks on quality degraded datasets,” in Proc. International Conference on Digital Image Computing: Techniques and Applications (DICTA), 2017, pp. 1-8. https://doi.org/10.1109/DICTA.2017.8114544

[19] L. F. Rodrigues, M. C. Naldi, and J. F. Mari, ”Comparing convolutional neural networks and preprocessing techniques for HEP-2 cell classification in immunofluorescence images,” Computers in Biology and Medicine, vol. 116, p. 103542, 2020. https://doi.org/10.1016/j.compbiomed.2019.103542

[20] R. Fu et al., ”Enhanced intelligent identification of concrete cracks using multi-layered image preprocessing-aided convolutional neural networks,” Sensors, vol. 20, no. 7, p. 2021, 2020. https://doi.org/10.3390/s20072021

[21] R. Miezianko and D. Pokrajac, ”People detection in low resolution infrared videos,” in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2008, pp. 1-6. https://ieeexplore.ieee.org/document/4640510

[22] J. W. Davis and M. A. Keck, ”A two-stage template approach to person detection in thermal imagery,” in Proc. IEEE Workshops on Applications of Computer Vision, vol. 1, 2005, pp. 364-369. https://ieeexplore.ieee.org/document/1500937

[23] R. Gade and T. B. Moeslund, ”Constrained multi-target tracking for team sports activities,” IPSJ Transactions on Computer Vision and Applications, vol. 10, no. 1, pp. 1-11, 2018. https://doi.org/10.1186/s40649-018-0058-4

[24] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, ”Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments,” University of Massachusetts, Amherst, Technical Report 07-49, October 2007. https://inria.hal.science/inria-00321923v1

[25] S. Yeom, ”Thermal Image Tracking for Search and Rescue Missions with a Drone,” Drones, vol. 8, no. 2, p. 53, Feb. 2024, doi: 10.3390/drones8020053.