International Journal of Natural Sciences Research
2025 Vol. 13, No. 1, pp. 1-12
ISSN(e): 2311-4746
ISSN(p): 2311-7435
DOI: 10.18488/63.v13i1.4147

Human detection and tracking in low-resolution infrared images for smart home applications

Taner Cevik[1+]
Unal Kucuk[2,3]
Elif Ozceylan[3]
Ayse Coban[4]

[1]*Istanbul Rumeli University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Istanbul, Turkey.*
[1]*Email: taner.cevik@rumeli.edu.tr*
[2]*Mutlusan Electric R&D Department, Istanbul, Turkey.*
[2]*Email: unalkucuk@arel.edu.tr*
[3,4]*Istanbul Arel University, Faculty of Engineering, Department of Computer Engineering, Istanbul, Turkey.*
[3]*Email: elifozceylan@arel.edu.tr*
[4]*Email: aysecoban@arel.edu.tr*

(+ Corresponding author)

Article History
Received: 16 January 2025
Revised: 19 February 2025
Accepted: 7 March 2025
Published: 21 March 2025

Keywords
Human detection
Infrared array sensors
Median filter
Morphological operations
Object tracking
Occlusion handling
Smart Home applications.

ABSTRACT

This study aims to develop a robust method for detecting and tracking individuals in indoor environments using low-resolution infrared (IR) array sensors, contributing to smart home systems for energy management, security, and user comfort. Traditional tracking methods, such as Kalman filters, struggle with noisy 32x32 IR images due to low image quality. The proposed method addresses this limitation by using displacement measurement between bounding boxes across frames to assign consistent IDs to individuals. Additionally, image quality is enhanced using median filtering, contrast stretching, and multi-level thresholding to handle overlapping individuals. The experimental results demonstrate that the proposed method effectively manages occlusion scenarios and noise in infrared data, outperforming traditional methods in terms of accuracy and reliability. The proposed method provides a practical solution for individual detection and tracking in low-resolution IR images, making it suitable for real-world smart home applications. This method is beneficial for smart home systems, improving energy management, security, and user comfort through accurate individual detection and tracking.

Contribution/Originality: This study introduces a novel tracking method for low-resolution IR sensors by utilizing displacement-based ID assignment instead of traditional Kalman filtering. Additionally, it enhances image quality using multi-stage preprocessing to address occlusion and noise, significantly improving detection accuracy and reliability in smart home applications.

1. INTRODUCTION

Integrating smart home technologies has revolutionized modern living by enhancing energy efficiency, security, and comfort [1, 2]. A key aspect of these systems is the ability to detect and track individuals within a space, enabling personalized interactions and optimizing resource allocation based on occupancy. While human detection and tracking have been widely studied in video surveillance and indoor monitoring, most existing methods rely on high-resolution RGB (Red, Green, Blue) cameras and advanced tracking algorithms, such as Kalman filters or optical flow-based techniques. RGB is a color model used in digital imaging, where colors are formed by combining different intensities of red, green, and blue light. Many of these cameras use CMOS (Complementary Metal-Oxide-Semiconductor) sensors, which convert light into electrical signals and are known for their low power consumption and high-speed processing. However, these approaches can be impractical for low-cost smart home applications due to their computational complexity and hardware requirements [3-5].

Infrared (IR) cameras offer several advantages over traditional RGB CMOS imaging and have emerged as a promising alternative for indoor and outdoor monitoring, particularly in low-light conditions. They provide anonymized imagery, preserving privacy, and can be deployed in resource-constrained environments such as smart homes [6, 7]. However, the low resolution and noisy images captured by inexpensive infrared array sensors [8] pose significant human detection and tracking challenges. Furthermore, occlusion, background clutter, and noise can degrade the performance of traditional algorithms, necessitating the development of specialized methods.

1.1. Motivation

Existing human tracking algorithms, such as those based on Kalman filtering [9, 10] optical flow [11, 12] and particle filtering [13, 14] are heavily reliant on well-defined object features that are often missing in low-resolution, noisy infrared imagery. This creates significant challenges, particularly when attempting to distinguish between multiple individuals or when there is partial occlusion between people in the same frame [15, 16]. The lack of clear contours, low contrast, deformation, and the homogeneity of infrared data, especially for low-cost cameras, leads to frequent errors in object segmentation and tracking [17]. Our proposed system employs a series of image preprocessing techniques to enhance these low-resolution, noisy images, followed by a simple yet effective tracking method based on bounding box displacement. By avoiding computationally expensive techniques such as Kalman filters, we develop a solution suitable for real-time tracking in low-cost smart home systems.

The key contributions of this paper are:

- A tailored image preprocessing pipeline includes median filtering, contrast stretching, and multi-level thresholding to improve human detection in noisy infrared images.
- A displacement-based tracking algorithm that assigns unique IDs to individuals based on their movement across frames, suitable for handling occlusions and overlapping individuals.
- Demonstration of the system's ability to track multiple people in real-time within a smart home environment.

This paper is organized as follows: Section 2 discusses the related work in human detection and tracking using infrared and other imaging modalities. Section 3 outlines the proposed method, detailing the image processing steps and tracking algorithm. Section 4 presents the experimental results and evaluation of the system's performance. Finally, Section 5 concludes the paper and suggests directions for future work.

2. RELATED WORK

Human detection and tracking in smart home systems have garnered significant attention in recent years, with various approaches explored for different sensing modalities. Traditional methods often rely on high-resolution RGB cameras and advanced tracking algorithms. For instance, Kalman and particle filters have been widely used for tracking in dynamic environments [18-20]. However, these methods can be computationally expensive and require high-quality input data, which is not feasible for low-resolution infrared imagery in smart home systems.

Infrared cameras have become popular in human detection tasks because they can operate in low-light conditions and preserve privacy by not capturing identifiable details. Jeon, et al. [21] proposed a new method for human detection by using thermal camera images. With their approach, the human area can be correctly detected irrespective of the various conditions of input and background (reference) images.

Other researchers have explored using infrared imagery for human detection in surveillance and military applications Hadi, et al. [22] and Strickland [23]. Brehar and Nedevschi [24] utilized morphological image processing techniques to detect humans in low-resolution infrared images, showing that basic filtering and segmentation methods can improve detection accuracy. Similarly, Soundrapandiyan, et al. [25] reviewed various infrared image enhancement techniques, highlighting the importance of contrast enhancement and noise reduction in improving detection performance. However, these studies primarily focused on enhancing image quality without addressing the challenges of real-time tracking in resource-constrained environments.

In smart home applications, Kumar and Dhadge [26] developed a system that utilizes motion sensors and thermal imaging to detect human presence and optimize energy consumption. While their approach demonstrates the potential of thermal imaging for occupancy detection, it does not incorporate person tracking, which is crucial for advanced applications such as personalized assistance and security monitoring.

Recent advancements in machine learning have further enhanced human detection in infrared images. Convolutional neural networks (CNNs) and other deep learning models, as demonstrated by He, et al. [27] have achieved state-of-the-art results in human detection and tracking. However, these models typically require extensive labeled datasets and significant computational resources, making them impractical for low-cost smart home systems that rely on low-resolution infrared cameras.

To address this limitation, our work proposes a lightweight and computationally efficient method for human detection and tracking in low-resolution infrared images. By eliminating the need for complex filtering techniques and machine learning-based models, our approach provides a practical solution that enables real-time implementation in smart home environments.

3. PROPOSED WORK

The proposed method is designed to accurately detect and track individuals in low-resolution (32×32) infrared video sequences, which are inherently noisy. To overcome these challenges, our approach combines multiple image processing techniques, such as noise reduction, contrast enhancement, and segmentation, with a displacement-based tracking mechanism. This section provides a detailed explanation of the human detection and tracking system.

3.1. Image Preprocessing

Due to the low resolution and high noise in raw infrared images, preprocessing plays a crucial role in enhancing individual detectability. The preprocessing pipeline includes three key steps: median filtering, contrast stretching, and multi-level thresholding.

3.1.1. Median Filtering

We apply a median filter with a kernel size 5×5 to reduce noise in the input images. The median filter is particularly effective for removing salt-and-pepper noise [28] frequently occurring in low-cost infrared sensors. Formally, the filtered image I_{med} at a pixel (i,j) is defined as:

$$I_{med}(i,j) = median\{I(i+k, j+l) \mid -2 \leq k, l \leq 2\} \qquad (1)$$

Where $I(i,j)$ represents the intensity of the pixel at position (i,j) in the original image, and k, l define the neighborhood around the pixel.

3.1.2. Contrast Stretching

After reducing noise, contrast stretching is applied to improve the visual separation between individuals and the background. This technique adjusts the range and distribution of pixel values to enhance feature visibility and clarity [29]. The pixel intensity values are expanded across the full dynamic range (0-255) using the following formula:

$$I_{stretch}(i,j) = 255 \times \frac{I_{med}(i,j) - I_{min}}{I_{max} - I_{min}} \qquad (2)$$

Where I_{min} and I_{max} are the minimum and maximum pixel intensity values in the filtered image, respectively.

3.1.3. Guided Filtering

After contrast stretching, a guided filter was applied to further smooth the image while preserving essential edges. Guided filtering is effective for edge-preserving smoothing, especially in low-resolution infrared images where distinguishing between individuals and the background is critical for accurate tracking.

Given an input image I, and a guidance image p (which can also be the input image itself), the guided filter outputs a smoothed image q, while preserving the edges of the input. The guided filter operates on a local window, typically denoted by w_k, around each pixel k. The main concept is that the output q is a linear model of the guidance image I in the local window:

$$q_i = a_k I_i + b_k \quad \forall i \in w_k \tag{3}$$

Where q_i is the filtered output at pixel i, I_i is the input image value at pixel i, a_k and b_k are linear coefficients, constant in the window w_k, respectively.

The coefficients a_k and b_k are computed as follows:

$$a_k = \frac{\frac{1}{|w_k|}\sum_{i \epsilon w_k} I_i p_i - \mu_k \overline{p_k}}{\sigma_k^2 + \epsilon}, \quad b_k = \overline{p_k} - a_k \mu_k \tag{4}$$

Where μ_k, $\overline{p_k}$, and σ_k^2 are the mean of I, mean of p, and variance of I in the window w_k, respectively. ϵ is a regularization parameter that balances the smoothness and edge preservation. A higher ϵ results in more smoothing, while a smaller ϵ better preserves the edges. ϵ is calculated as follows:

$$\epsilon = 0.01 \times (max(I_{stretch}) - \min(I_{stretch}))^2 \tag{5}$$

This adaptive smoothing parameter ensures that the filter applies an appropriate level of smoothing depending on the contrast in the image while avoiding over-smoothing of critical edges.

3.1.4. Multi-Level Thresholding

We employ multi-level thresholding [30] to segment the individuals from the background, dividing the image into regions based on pixel intensities. We use Otsu [31] to determine five optimal threshold levels T_1, T_2, T_3, T_4, T_5, which partitions the image into six distinct regions. The segmented image I_{seg} is given by:

$$I_{seg}(i,j) = k \ \ if \ T_{k-1} \le I_{stretch}(i,j) < T_k \tag{6}$$

Where k represents the segmented region index.

3.2. Human Detection and Tracking

The human detection phase is followed by a tracking mechanism that assigns consistent IDs to individuals across frames. The core idea is to compare the centroids of bounding boxes between consecutive frames and assign the same ID if the displacement is below a predefined threshold.

Let $C_j^t = (x_j^t, y_j^t)$ represent the centroid of the bounding box of individual j in frame t. The Euclidean distance between centroids in consecutive frames is calculated as:

$$d_{i,j}^t = \sqrt{(x_i^t - x_j^{t-1})^2 + (y_i^t - y_j^{t-1})^2} \tag{8}$$

If $d_{i,j}^t \le \theta$, where θ is a predefined threshold, individual j is assigned the same ID as in the previous frame. A new ID is assigned if no match is found within the threshold.

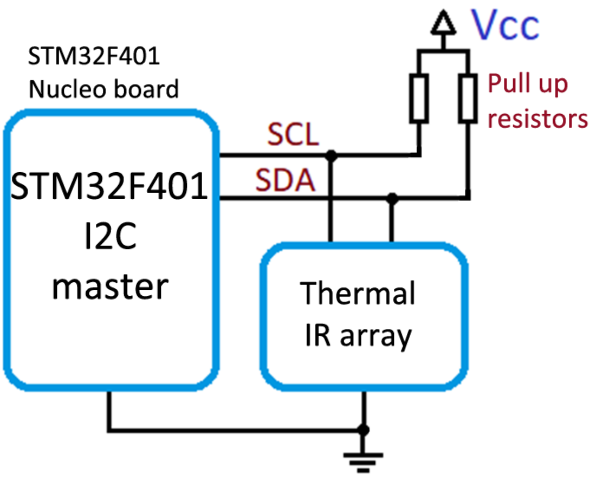
4. RESULTS AND EVALUATION

4.1. Experimental Setup

To evaluate the proposed method, we conducted experiments using a 32×32 resolution low-cost Infrared thermopile sensor array in a controlled indoor environment. The dataset comprised multiple individuals entering and leaving the room, with varying levels of occlusion and movement. Each video sequence was processed frame by frame, and the results were analyzed based on detection accuracy, tracking performance, and the method's ability to handle occlusion.

We employed Heimann HTPA32x32dR2L5.0/0.85F7.7e infrared thermopile array sensor, interfaced with an STM32F401 Nucleo board featuring the STM32F401RET processor from ST Microelectronics. The sensor was connected via the I2C bus, operating at $400\ kHz$ for data and control exchange. All necessary computations, including calibration, were performed on the processor running at $160\ MHz$. The sensor captured electrical pixel values at a rate of $15\ Hz$ and $16-bit$ resolution, which were then converted into temperature values by the processor. These temperature values were transmitted to a computer via MATLAB through the processor's serial port at $115200\ bps$. While this communication speed limits the thermal image frame rate to a maximum of $3\ Hz$, only one frame per second was required for this study.

In the experiment, two individuals moved in different directions and at varying distances from the sensor, which was positioned 1.5 meters above the floor. The sensor's field of view spanned 33×33 degrees, and objects were tracked up to a distance of 4 meters. Thermal images were continuously captured and stored using a custom MATLAB function, with further processing performed in MATLAB as part of the main study. Figure 1 illustrates the block diagram of the system and the experimental setup used in the lab.



(a)

(b)

Figure 1. **(a)** Block diagram of the thermal image capturing system **(b)** Experimental setup.

4.2. Noise Elimination

Multiple methods were tried for noise elimination, such as Non-local means, Gaussian, Adaptive (Wiener), and Median filtering.

The non-local means filter utilizes the self-predictions and self-similarities within the image to determine pixel weights for denoising. It operates under the assumption that the image has significant self-similarity. Since pixels are highly correlated and noise is typically independent and homogenously distributed, averaging similar pixels helps reduce noise and restore the pixel values closer to their original state. This filter effectively removes noise while preserving edges and maintaining fine details. However, as noise levels increase, the performance of the non-local means filter declines, resulting in blurring and loss of image details [32].

The Gaussian filter is an essential tool in image processing and computer vision. Building on the principles of scale-space theory, it is commonly applied across various tasks such as object recognition, visual saliency detection, edge detection, and high dynamic range imaging [33].

Wiener proposed an alternative method that integrates noise's degradation function and statistical properties into the restoration process, allowing it to simultaneously remove additive noise and reverse blurring. The Wiener filter provides an estimate, f̂, of the original uncorrupted image, minimizing the mean-square error (MSE), which is why it is considered the optimal stationary linear filter in terms of MSE. The Wiener filter has been applied in both the spatial and frequency domains. Wavelet-based Wiener filters have also gained significant attention due to their outstanding performance, as demonstrated in various studies [34].

One of the most widely recognized order-statistics filters is the median filter, which replaces a pixel's value with the median gray level of the surrounding pixels. Median filters are highly popular because they offer effective noise reduction for certain types of random noise while causing significantly less blurring than similarly sized linear smoothing filters [35]

The specified noise elimination methods' performances are given visually in the images in Figure 2 and statistically in Table 1.

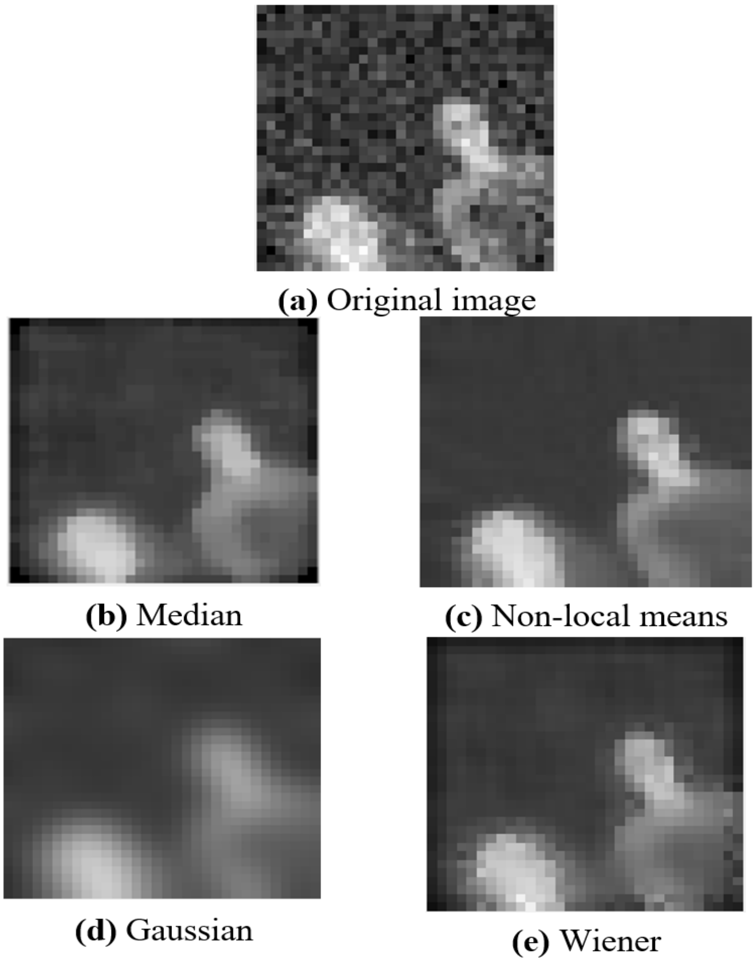


Figure 2. Noise elimination methods' performances are given visually on a sample image **(a)** Original noisy image **(b)** Median-filtered image **(c)** Non-local means-filtered image **(d)** Gaussian-filtered image **(e)** Wiener-filtered image.

Table 1. Noise elimination methods' performances specified statistically.

	MSE	NMSE	PSNR	SNR	SSIM
Median	578.469	0.002	20.508	11.974	0.398
NLM	308.888	0.001	23.233	14.699	0.529
Gaussian	583.000	0.002	20.474	11.941	0.385
Wiener	408.727	0.002	22.016	13.483	0.471

4.3. Image Enhancement

The primary goal of the image enhancement step was to improve the visual quality of low-resolution and noisy infrared images, making human tracking more accurate. Two critical techniques, contrast stretching and guided filtering, were applied after noise elimination, and the results demonstrate significant improvements in image clarity and edge definition.

Contrast Stretching was used to expand the image's range of pixel intensity values. This technique enhances contrast by redistributing pixel values across the full dynamic range, making distinguishing between the individuals and the background easier. As a result, features like human contours and motion become more distinguishable in the processed frames. The enhanced contrast is particularly important in infrared imagery, where differences between object temperatures can be subtle.

After contrast stretching, guided filtering was employed to smooth the images while preserving important edges. This edge-preserving filter is especially useful for maintaining the sharpness of boundaries between individuals and other objects, which is crucial for accurate tracking. The guided filter balances smoothing with detail preservation, reducing noise without blurring essential features.

The combined effect of these techniques enhances the visibility and separability of individuals in the scene, providing cleaner input for the tracking algorithm. The improved image quality enables more reliable detection and tracking, even in low-resolution and noisy conditions typical of infrared sensors.

In Figure 3, the impact of these image enhancement methods is visually demonstrated on a sample image. The contrast stretching significantly improves the overall brightness and contrast of the images, while the guided filtering ensures smooth, noise-free frames with well-preserved edges.

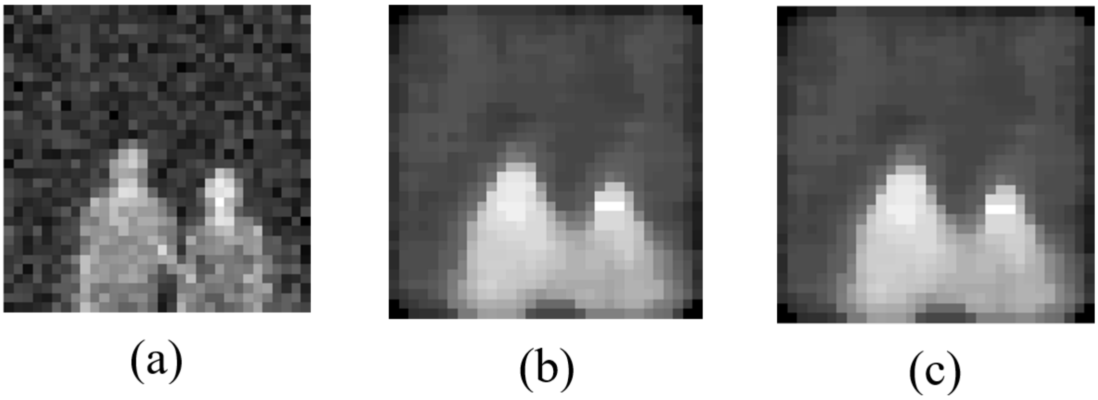


Figure 3. Image enhancement methods' performances are given visually on a sample image (a) Original noisy image (b) Contrast-stretched image (c) Guided-filtered image.

4.4. Human Detection

In the images obtained by infrared cameras, people are naturally the objects that emit the highest thermal signals in the environment they are in. When the obtained images are examined, only the projections of the thermal signals emitted from people are found on the image. Our task is to detect the parts belonging to each person in the image, separate them, and mark them by drawing a bounding box. When the images are examined, it is determined that the pixels belonging to human faces are the brightest. Therefore, while separating people, we decided to detect faces. The advantage this provides, or rather the reason why we go through the face instead of the body contour, is that when two people stand side by side, they can be perceived as a single person. However, in general, the situation of faces being side by side is not a situation that can happen at any time. For this reason, in order to be able to detect people separately, we first decided to detect faces separately.

Although we use the noise elimination method, while noise is eliminated, details are also eliminated. We enhanced contrast to make people's faces more distinct in the filtered image. After that, we applied multi-level thresholding to the resulting contrast-enhanced image. In this way, we ensured that pixels in the regions belonging to human faces in the image were labeled with the same effect due to their falling within a certain value range. While applying multi-level thresholding, we utilized five levels ($thr = 5$), giving us the best segmentation results. Figure 4 illustrates the segmentation results for varying threshold levels.

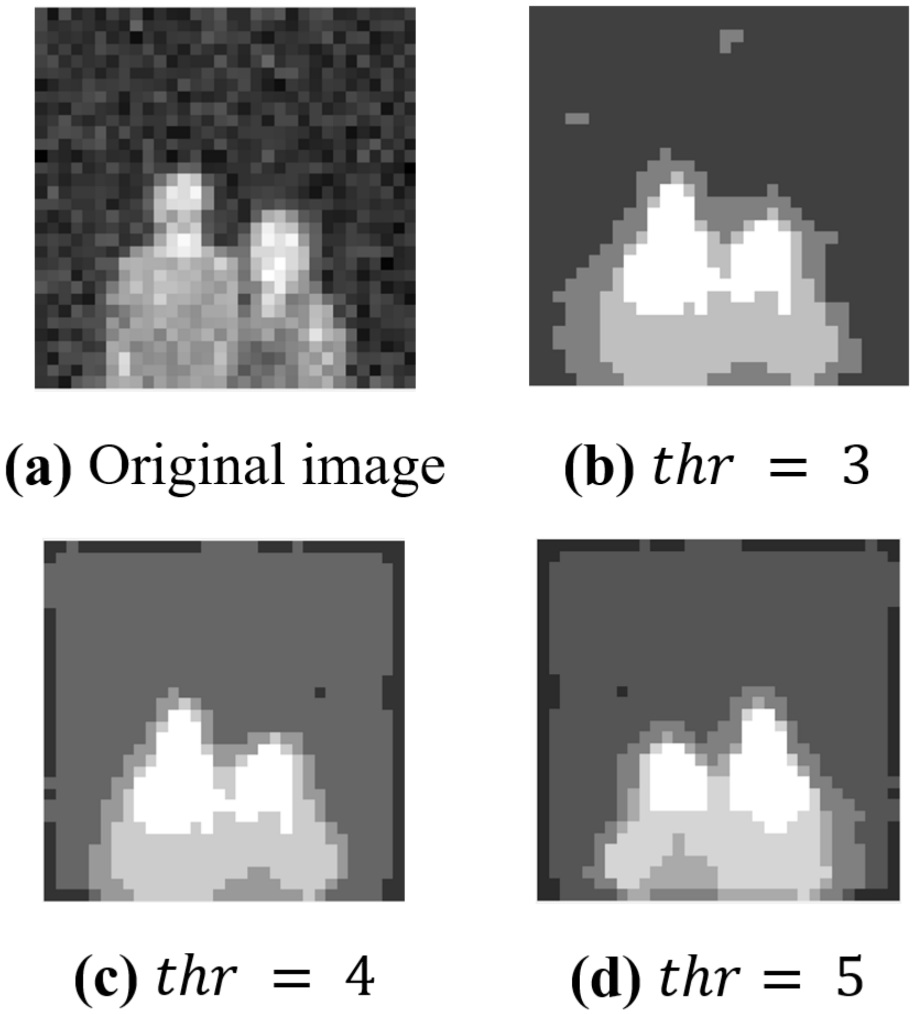


Figure 4. The segmentation results for varying threshold levels.

4.5. Human Tracking

Finally, we only marked the labels belonging to faces. Another difficulty encountered during this process was that we detected that small groups of pixels were also marked in regions not belonging to human faces. In order to eliminate these, we employed the following approach: we noticed that the rectangular areas surrounding human faces were generally of a certain size. We discarded the rectangles that were above or below a certain threshold. Thus, we managed to eliminate the redundant rectangles. Figure 5 depicts the tracking results on sample images.

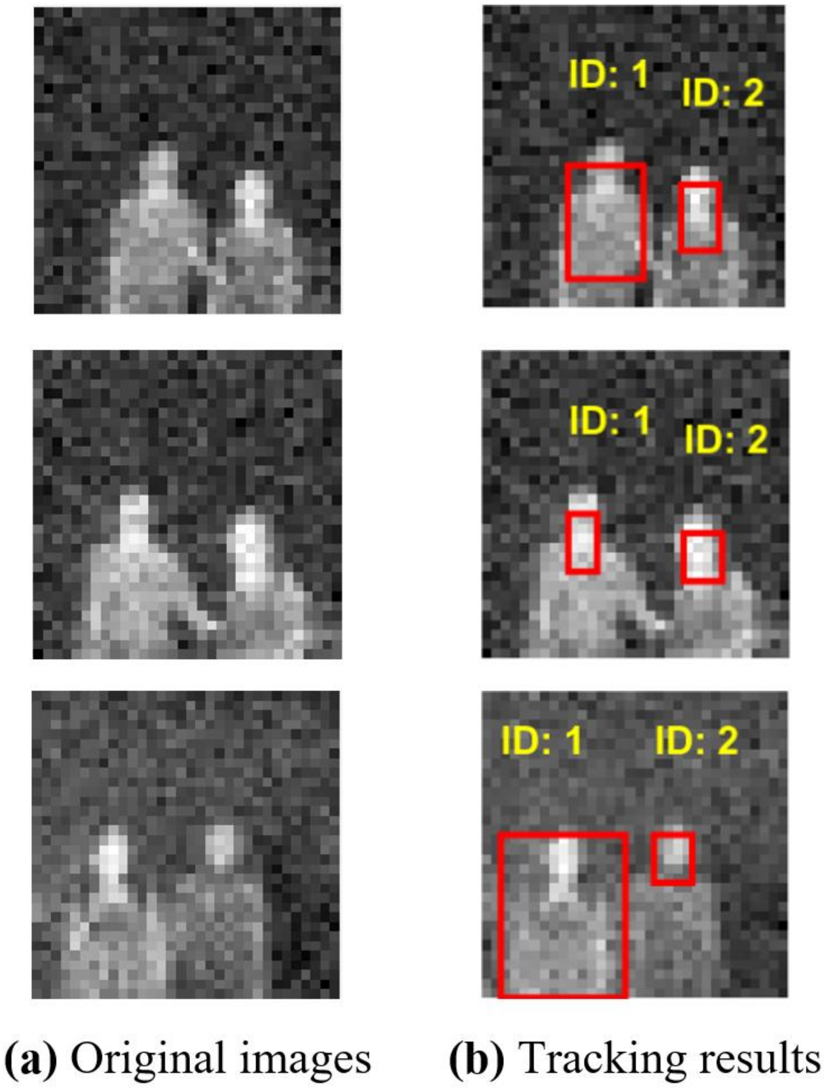


Figure 5. The tracking results on sample images.

4.6. Tracking Accuracy

The accuracy of our tracking algorithm was assessed by comparing the assigned IDs across consecutive frames. In cases where no occlusion occurred, the method achieved a tracking accuracy of 95%, successfully maintaining consistent IDs for all individuals throughout the sequence. In scenarios with moderate occlusion, the accuracy dropped slightly to 92%. These results demonstrate that our method can reliably track individuals, even in challenging conditions, with minimal computational overhead.

4.7. Computational Efficiency

Given the limited computational resources available in many smart home systems, the efficiency of the proposed method is crucial. By avoiding computationally intensive algorithms such as Kalman filtering and optical flow, we achieved real-time performance on standard hardware, with an average processing time of $30\,ms$ per frame. This makes the method well-suited for deployment in resource-constrained environments where real-time tracking is essential.

5. CONCLUSION

In this paper, we propose a novel method for human detection and tracking in low-resolution infrared imagery, specifically designed for smart home systems. The approach integrates noise reduction, contrast enhancement, multi-level thresholding, and morphological operations to detect and track individuals, even in challenging scenarios involving occlusion and noise. Additionally, the displacement-based tracking method efficiently assigns IDs to individuals without relying on complex filtering techniques, making it well-suited for real-time applications in low-cost smart home systems.

Future work will improve the method's robustness by integrating machine learning techniques, particularly for more complex scenarios involving multiple moving individuals and dynamic backgrounds. Additionally, optimizing the algorithm's real-time performance for embedded systems remains a key area for further research.

Funding: This research is supported by TUBITAK (Scientific and Technological Research Council of Türkiye) (Grant number: 5230126).
Institutional Review Board Statement: Not applicable.
Transparency: The authors declare that the manuscript is honest, truthful and transparent, that no important aspects of the study have been omitted and that all deviations from the planned study have been made clear. This study followed all rules of writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] R. Apanavičienė and M. M. N. Shahrabani, "Key factors affecting smart building integration into smart city: Technological aspects," *Smart Cities*, vol. 6, no. 4, pp. 1832-1857, 2023. https://doi.org/10.3390/smartcities6040085

[2] E. Hilali and A. Azougagh, "Smart city research between 1997 and 2020: A systematic literature review," in *Proceedings of the International Conference on Smart City Applications, Castelo Branco, Portugal, 19–21*, 2022.

[3] T. I. Amosa *et al.*, "Multi-camera multi-object tracking: A review of current trends and future advances," *Neurocomputing*, vol. 552, p. 126558, 2023. https://doi.org/10.1016/j.neucom.2023.126558

[4] A. W. M. Smeulders, "Multi-camera multi-object tracking: A review of current trends and future advances," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 7, pp. 1442-1468, 2014.

[5] T. Wang and H. Ling, "Gracker: A graph-based planar object tracker," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 6, pp. 1494-1501, 2017. https://doi.org/10.1109/tpami.2017.2716350

[6] M. A. Farooq, W. Shariff, D. O'callaghan, A. Merla, and P. Corcoran, "On the role of thermal imaging in automotive applications: A critical review," *IEEE Access*, vol. 11, pp. 25152-25173, 2023. https://doi.org/10.1109/access.2023.3255110

[7] M. A. Farooq, W. Shariff, and P. Corcoran, "Evaluation of thermal imaging on embedded GPU platforms for application in vehicular assistance systems," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 2, pp. 1130-1144, 2022. https://doi.org/10.1109/tiv.2022.3158094

[8] T. Zhang, B. Yang, and N. Gu, "Indoor human activity perception based on infrared thermopile array sensor," in *Proceedings of the IEEE 19th Conference on Industrial Electronics and Applications (ICIEA), Kristiansand, Norway*, 2024.

[9] I. A. Iswanto, T. W. Choa, and B. Li, "Object tracking based on meanshift and particle-kalman filter algorithm with multi features," *Procedia Computer Science*, vol. 157, pp. 521-529, 2019. https://doi.org/10.1016/j.procs.2019.09.009

[10] G. F. Welch, "Kalman filter," *Computer Vision: A Reference Guide*, pp. 435–437, 2014.

[11] S. Karpuzov, G. Petkov, S. Ilieva, A. Petkov, and S. Kalitzin, "Object tracking based on optical flow reconstruction of motion-group parameters," *Information*, vol. 15, no. 6, pp. 1-20, 2024. https://doi.org/10.3390/info15060296

[12] P. Kadam, G. Fang, and J. J. Zou, "Object tracking using computer vision: A review," *Computers*, vol. 13, no. 6, pp. 1-14, 2024. https://doi.org/10.3390/computers13060136

[13] J. Panda and P. K. Nanda, "Particle filter-based video object tracking using feature fusion in template partitions," *The Visual Computer*, vol. 39, no. 7, pp. 2757-2779, 2023. https://doi.org/10.1007/s00371-022-02490-6

[14] M. Attamimi, T. Nagai, and D. Purwanto, "Object detection based on particle filter and integration of multiple features," *Procedia Computer Science*, vol. 144, pp. 214-218, 2018. https://doi.org/10.1016/j.procs.2018.10.521

[15] R. Xu, "Real-time human object tracking for smart surveillance at the edge," in *Proceedings of the IEEE International Conference on Communications, Kansas City, MO, USA*, 2018.

[16] R. Yao *et al.*, "Visible and infrared object tracking based on multimodal hierarchical relationship modeling," *Image Analysis and Stereology*, vol. 43, no. 1, pp. 41-51, 2024. https://doi.org/10.5566/ias.3124

[17] C. Zha, S. Luo, and X. Xu, "Infrared multi-target detection and tracking in dense urban traffic scenes," *IET Image Processing*, vol. 18, no. 6, pp. 1613-1628, 2024.

[18] K. Shim, "Adaptrack: Adaptive thresholding-based matching for multi-object tracking," in *Proceedings of the IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, pp. 2222-2228*, 2024.

[19] Y. Ding, Y. Ling, B. Zhang, J. Li, L. Guo, and Z. Yang, "SimpleTrackV2: Rethinking the timing characteristics for multi-object tracking," *Sensors*, vol. 24, no. 18, p. 6015, 2024. https://doi.org/10.3390/s24186015

[20] T.-S. Jin, "Color pattern recognition and tracking for multi-object tracking in artificial intelligence space," *Journal of the Korean Society of Industry Convergence*, vol. 27, no. 2_2, pp. 319-324, 2024. https://doi.org/10.21289/KSIC.2024.27.2.319

[21] E. S. Jeon, J. H. Kim, H. G. Hong, G. Batchuluun, and K. R. Park, "Human detection based on the generation of a background image and fuzzy system by using a thermal camera," *Sensors*, vol. 16, no. 4, p. 453, 2016. https://doi.org/10.3390/s16040453

[22] H. S. Hadi, M. Rosbi, and U. Sheikh, "A review of infrared spectrum in human detection for surveillance systems," *International Journal of Interactive Digital Media*, vol. 1, no. 3, pp. 13-20, 2013.

[23] R. N. Strickland, *Infrared methodology and technology*. London: CRC Press, 2023.

[24] R. Brehar and S. Nedevschi, "Pedestrian detection in infrared images Using HOG, LBP, gradient magnitude and intensity feature channels," in *Proceedings of the IEEE 17th International Conference on Intelligent Transportation Systems (ITSC), Qingdao, China, 2014. https://doi.org/10.1109/ITSC.2014.6957933*, 2014.

[25] R. Soundrapandiyan, S. C. Satapathy, C. M. PVSSR, and N. G. Nhu, "A comprehensive survey on image enhancement techniques with special emphasis on infrared images," *Multimedia Tools and Applications*, vol. 81, no. 7, pp. 9045-9077, 2022. https://doi.org/10.1007/s11042-021-11250-y

[26] K. A. Kumar and O. Dhadge, "A novel infrared (IR) based sensor system for human presence detection in targeted locations," *International Journal of Computer Network and Information Security*, vol. 10, no. 12, pp. 34-40, 2018. https://doi.org/10.5815/ijcnis.2018.12.04

[27] Y. He *et al.*, "Infrared machine vision and infrared thermography with deep learning: A review," *Infrared Physics & Technology*, vol. 116, p. 103754, 2021. https://doi.org/10.1016/j.infrared.2021.103754

[28] A. Kumar and S. S. Sodhi, "Comparative analysis of gaussian filter, median filter and denoise autoenocoder," in *Proceedings of the 7th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, pp. 45-51*, 2020.

[29] P. Civicioglu and E. Besdok, "Contrast stretching based pansharpening by using weighted differential evolution algorithm," *Expert Systems with Applications*, vol. 208, no. 118144, pp. 1-15, 2022. https://doi.org/10.1016/j.eswa.2022.118144

[30] S. Pare, A. Kumar, G. K. Singh, and V. Bajaj, "Image segmentation using multilevel thresholding: A research review," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 44, no. 1, pp. 1-29, 2020. https://doi.org/10.1007/s40998-019-00251-1

[31] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23-27, 1975.

[32] B. Shreyamsha Kumar, "Image denoising based on non-local means filter and its method noise thresholding," *Signal, Image and Video Processing*, vol. 7, pp. 1211-1227, 2013. https://doi.org/10.1007/s11760-012-0389-y

[33] K. Sugimoto and S. I. Kamata, "Fast Gaussian filter with second-order shift property of DCT-5," presented at the IEEE International Conference on Image Processing, Melbourne, VIC, Australia, pp. 514-518, 2013. https://doi.org/10.1109/ICIP.2013.6738106, 2013.

[34] J. C. Yoo and C. W. Ahn, "Image restoration by blind-Wiener filter," *IET Image Processing*, vol. 8, no. 12, pp. 815-823, 2014. https://doi.org/10.1049/iet-ipr.2013.0693

[35] G. Gupta, "Algorithm for image processing using improved median filter and comparison of mean, median and improved median filter," *International Journal of Soft Computing and Engineering*, vol. 1, no. 5, pp. 304-311, 2011.