VOL. 23, NO. 2, 2024, 44-52
elektrika.utm.my
ISSN 0128-4428

ELEKTRIKA
Journal of Electrical Engineering

Reviewing Approaches and Techniques for Detecting Suspicious Human Behavior: A Comprehensive Survey

Wong Khai Chiuan[1] and **Mohd Ridzuan Bin Ahmad**[1*]

[1]Department of Control and Mechatronics Engineering, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia.

[*]Corresponding author: mdridzuan@utm.my

Abstract: The paper aims to review related works that focus on detecting suspicious human behavior using machine-learning techniques. Suspicious human behavior refers to behaviors that may indicate involvement in or preparation for a crime. Detecting such behaviors before a crime is committed allows law enforcement to take early action and prevent criminal activities. One of the challenges in developing an effective detection system for suspicious human behavior is the absence of a well-defined definition for such behaviors. Different definitions can lead to various methods for designing the detection system. The paper explores different definitions and their implications on the design of detection systems and mentions two types of methods that can be used for detecting suspicious human behavior: image-based methods and saliency mapping. Image-based methods utilize image or video recognition techniques to analyze objects held by individuals or recognize specific activities. Saliency mapping, on the other hand, focuses on emphasizing the movement of individuals using techniques like optical flow calculation to generate saliency maps. Additionally, the paper highlights the increasing popularity of embedded machine learning, particularly on portable platforms. The use of embedded machine learning allows for the deployment of machine learning models on mobile or lightweight devices. This can be relevant for developing efficient and portable systems for detecting suspicious human behavior. Overall, the paper aims to provide an overview of existing works in the field of suspicious human behavior detection using machine learning, exploring different definitions and methods employed in the literature.

Keywords: Embedded machine learning, Lightweight, Suspicious Human Behavior Detection

Article History: received 16 January 2024; accepted 15 July 2024; published 29 August 2024

1. INTRODUCTION

Suspicious human behavior is behavior that can indicate an individual may be involved in a crime or before committing a crime [8].To maintain public safety and security in today's interconnected society, detecting and preventing suspicious human behavior plays an important role. As different technologies and approaches have been produced, there is a need to explore and understand the various approaches and techniques employed in this area. This survey aims to review and analyze the methodologies used for detecting suspicious human behavior ad providing a comprehensive overview of the current landscape and highlighting key advancements and challenges.

In recent years, the advent of advanced surveillance systems, machine learning algorithms, and behavioral analytics has significantly enhanced the ability to detect anomalies and identify potential suspicious activities. Smart surveillance systems have been used to detect potential crime activity [10]. Early detection of suspicious behavior also increases the success rate of controlling a crime activity [11]. To detect suspicious behavior, the definition of suspicious behavior is important or else the development of a well-performed detection system will become challenging [12]. Suspicious human behavior can be defined as a behavior that is different from normal behavior or can be defined as an action that a person tries to hide themselves to get noticed by other individuals.

This comprehensive survey will contain five main parts, which are the definition of suspicious human behavior, dataset, detection methods, algorithm of the model, and embedded machine learning approaches for suspicious human behavior detection. The definition of suspicious human behavior in a project will affect the dataset used for training and the detection method. The algorithm of the previous work was also reviewed to list the recent technology in this area.

2. DEFINITIONS OF SUSPICIOUS HUMAN BEHAVIORS

In the domain of suspicious behavior analysis, researchers have put diverse definitions, leading to the development of various methodologies for designing detection systems. This has resulted in the classification of the definition of suspicious human behavior in existing literature into the following categories and Table 1 shows the comparison between the definitions.

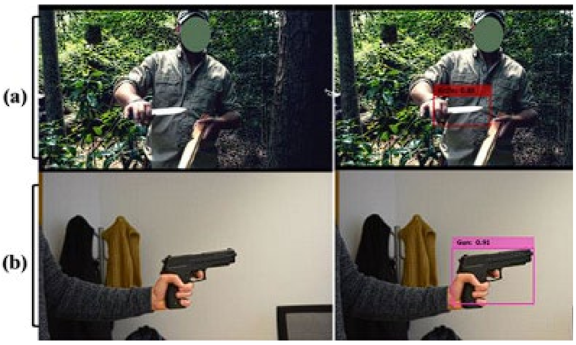

Figure 1. Unusual Object Detection [2]

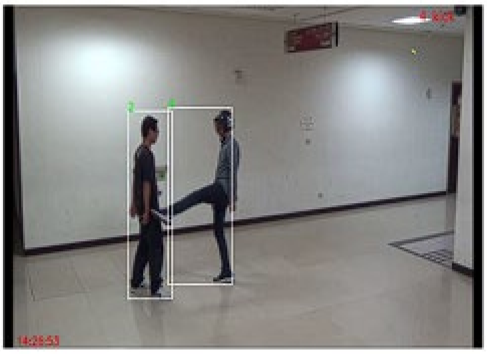

Figure 2. Unusual Activities (kicking) [5]

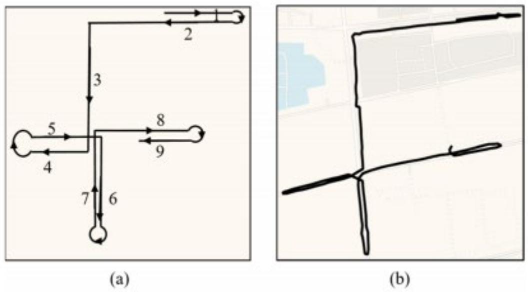

Figure 3. Unusual Movement Trajectory [8]

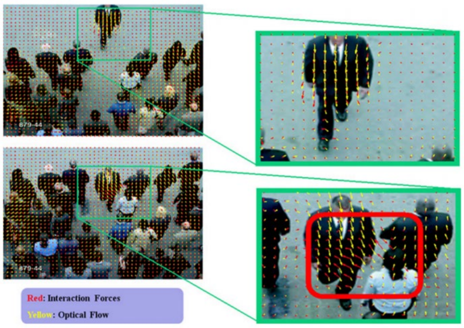

Figure 4. Individuals walk in different directions with the trend [9]

2.1 Unusual Object

To identify potential criminal activities, it is important to detect unusual objects that can be used as tools or weapons, such as crowbars or pliers, which pose a threat to the safety of others. Image-based or video-based detection techniques have been employed for this purpose [13]. By training a model using images of weapons, it becomes capable of examining each frame containing an individual holding a weapon and detecting any suspicious behavior [2]. Previous studies [14] utilized weapon image datasets to train their models for detecting suspicious human behavior. Another example is the development of a system [15] that can detect belt-shaped objects, which could be an indication of an individual tempted to remove an ATM from its location. Additionally, face accessories like masks are also considered suspicious behavior, as individuals may attempt to conceal their identity from surveillance cameras [14]. However, it should be noted that face mask detection may not be applicable in current situations due to the widespread use of face masks during the COVID-19 pandemic, which has become a commonplace occurrence. An example of unusual object detection is shown in Figure 1.

2.2 Unusual Activity

An unusual activity refers to an activity that deviates from the normal routine activities. Everyday actions like jogging, hugging, and waving hands are considered part of the normal routine. On the other hand, activities such as running in a crowd, falling, and snatching are regarded as unusual activities [5]. Figure 2 shows the unusual activity detection which is kicking other individuals.

In previous studies, various approaches have been employed to detect unusual activities. For instance, [9] focuses on detecting unusual activities by analyzing the motion of individuals that deviate from the general direction of others, which is considered abnormal behavior. In [10], normal events are categorized as part of the dominant set, and any event falling outside this set is classified as unusual activity. Similarly, [16] defines non-standard behavior as abnormal behavior, characterized by a pattern that differs from the majority of vents [17]. Additionally, a suspicious event can be defined as a rare, irregular or unexpected occurrence [18, 19].

2.3 Unusual Movement Pattern

The term “movement pattern” refers to the way individuals move, whether it is during a journey or while staying at a fixed location. When individuals plan to engage in criminal activities like burglary or theft, they often gather information about the entry points or potential opportunities at their target’s residence. In [8], trajectory data of individuals are analyzed to identify unusual movement patterns such as loitering, following, or visiting at unusual times. Detecting these movement patterns can help identify criminal activities even without the use of tools or weapons. Figure 3 shows an example of an unusual movement trajectory. However, this method has limitations in obtaining precise trajectory data for individuals. Additionally, recording trajectory data raises privacy concerns as it may expose an individual’s daily routines or addresses, leading to privacy breaches.

Moreover, deviating from the majority of crowd movement is also considered abnormal behavior [20]. Figure 4 shows the example of a person not following the trend of the movement of the crowd. This can be considered suspicious behavior because the person maybe

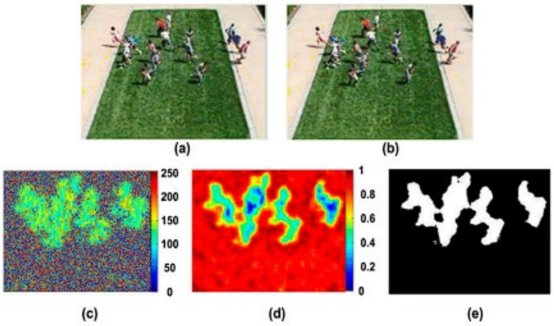

Figure 5. Energy Map [1]

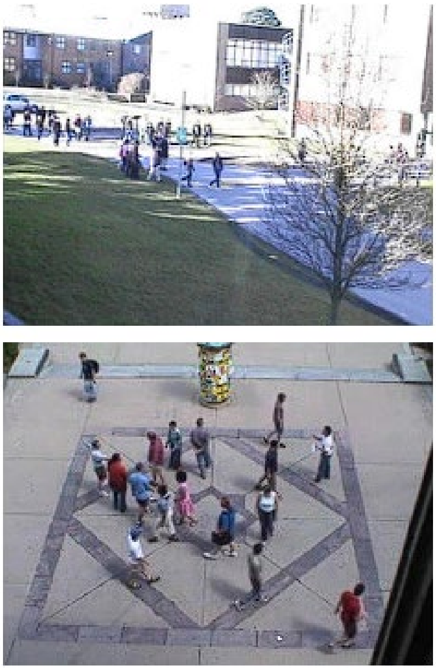

Figure 6. UMN dataset (people running) [4]

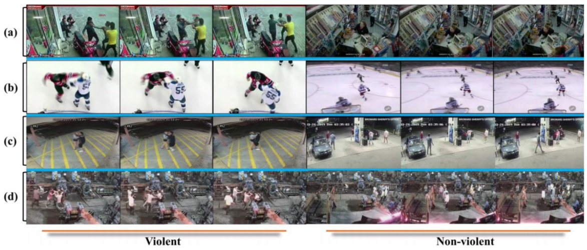

Figure 7. Hockey Fight Dataset [7]

is trying to do something inside the crowd which has the potential to have a criminal activity such as stealing or even murdering. By walking in different directions, there will be an interaction force that maybe causes injuries in the crowd.

Table 1. Comparison between Definitions

Unusual Objects	Unusual Activity	Unusual Movement Patterns
Refer to the items held or leaving from an individual.	Refer to the actions did by the individuals that make people uncomfortable.	Refer to the route of an individual that differ from usual.

2.4 Energy Level Map

The calculation of velocity applies to human movement, allowing for the identification of unusual activities such as running or falling. When such activities occur, there is a significant change in velocity, which can be quantified by applying the law of kinetic energy [1]. By obtaining the energy level, a kinetic energy model can be developed to monitor the energy levels of individuals. Parameters such as consistency can be established and when the consistency surpasses a predefined threshold, an alert can be triggered. The optical flow method can be employed to compute the velocity of individual pixels within video frames and based on these pixel velocities, a kinetic energy model can be constructed. Figure 5 shows an example of an energy map of a crowd running from a location.

3. DATASET

3.1 People Running

Several datasets contain people running or escaping. Figure 6 shows the UMN dataset [4], it contains 11 videos that showcase different escape scenarios within both indoor and outdoor settings. Each video begins by depicting normal behavior and transitions into abnormal examples, demonstrating various instances of deviating from typical actions. Behave dataset [21] contains behavior such as chasing, fighting, and running. Its primary focus is on aberrant behavior linked to criminal activity. It contains approximately 90000 frames that feature humans identified by bounding boxes. These individuals are involved in interacting groups, which are classified into six distinct behaviors. CUHK Avenue dataset [22] contains 16 training video clips and 21 testing video clips that are related to behaviors such as running, throwing objects, and loitering. There is a total of 30652 frames that capture the movement and behavior of pedestrians, cars, and cyclists.

3.2 Fights

As fighting can be detected as suspicious behavior, a dataset that contains fighting scenarios is used in various research. CAVIAR dataset [23] is a dataset that includes 28 videos of two different situations. Each video sequence is meticulously labeled with ground truth information, including frame-by-frame bounding boxes and a semantic description of the activities observed in each frame. The videos contain fighting scenarios of a leaving package in a public place. The hockey fight dataset [7] contains footage about fights in a hockey game and action movies specifically curated to depict violent behaviors observed in ice hockey matches. Figure 7 shows some examples of the hockey fight dataset. Each video clip in the dataset contains 50 frames and has been manually labeled as either

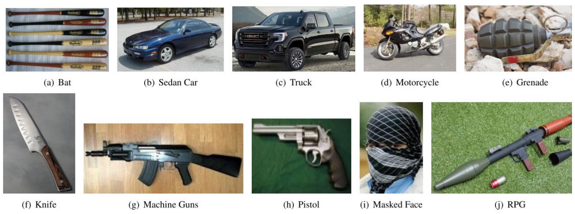
(a) Bat (b) Sedan Car (c) Truck (d) Motorcycle (e) Grenade (f) Knife (g) Machine Guns (h) Pistol (i) Masked Face (j) RPG

Figure 8. Cold and Hot weapons, vehicle dataset

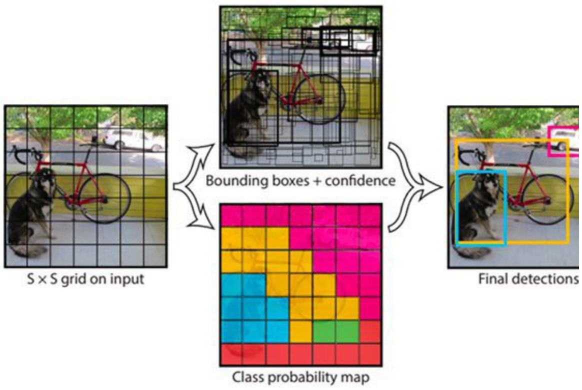

Figure 9. Yolo object detection

“fight” or “non-fight” to classify two types of behaviors. UCF-101 dataset [24] is a dataset that contains fight, robbery, and explosion scenarios. There is a total of 27 hours of footage, which are 13320 videos and covering 101 different action categories. The users uploaded videos with realistic camera movement and cluttered backgrounds, aiming to create a database that closely resembles real-world scenarios.

3.3 Weapons and Unusual Objects Dataset

[14] uses a weapon dataset consisting of 10,000 images with various types of dangerous objects or threatening objects. These objects include all kinds of hot weapons and cold weapons including machine guns, grenades, Rocket-Propelled Grenades (RPG), pistols, motorcycles, cars, knives, and bats. Figure 8 shows examples of dangerous objects used by the author. The author mentions that object detection is dependent on contextual information, so the images of the dataset are in their natural environments. Cars and trucks are also classified as dangerous objects because they car and truck maybe contain these dangerous objects. By considering masked robbers running away from the crime scene with a truck or car, their dataset also includes people with masked faces.

3.4 Dataset Limitation on Real-World Applications

Real world applications require a high quality and suitability of the dataset to play a crucial role in the outcomes and reliability of the results. Bias can occur due to the various sampling methods and collection of data. Dataset that collected from a specific region which not well perform and leads to skewed or in accurate results. Besides, the completeness of the data can impact the reliability of results in real-world applications. Noises and inaccurate data will be included in the dataset and leading to an incorrect result.

4. DETECTION METHODS

With the continuous improvement of machine learning, significant advancements have been made in the field of object detection [25]. Once suspicious human behavior is defined, the appropriate detection method can be selected. Various methods can be utilized for detecting suspicious human behavior and they can be classified as follows:

4.1 Image-based or Video-based Feature Extraction

4.1.1 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is a widely used architecture for extracting features from images [26]. CNN can contain large numbers of layers that learn to detect features of an image. Figure 10 shows the CNN structure. In the layer structure of CNN, there are three common layers, which are convolution, rectified linear unit (ReLU) and pooling. Convolution layer pass the images through a convolution layer to focus on the certain features from the image. ReLU is used to let the model learn complex patterns by introducing non-linearity. Pooling layers is used to reduce the spatial dimensions of the feature maps. In existing works, CNN has been employed for detecting suspicious human behavior, along with different architectures such as ToloV3 and Spatial-temporal CNN [5, 18]. Although CNN is capable of extracting image features, it alone may not be sufficient for detecting suspicious human behavior. Therefore, identification models like Long Short-Term Memory (LSTM), Convolutional LSTM (ConvLSTM), Gated Recurrent Unit (GRU), and improved SqueezeNet are utilized [5, 18, 27, 28]. A hybrid model combining CNN and LSTM has been employed for abnormal behavior detection [5]. Additionally, in some cases, mask-RCNN is used to detect objects in surveillance systems [29].

4.1.2 Long Short-Term Memory (LSTM)

LSTM is a model that designed to use for temporal sequences and long-range dependencies. LSTM is more useful in task that ongoing for a period of time, time series prediction and behavior analysis. LSTM contains memory cells and three types of gates which unique the structure of this structure. Memory cells maintain the information for

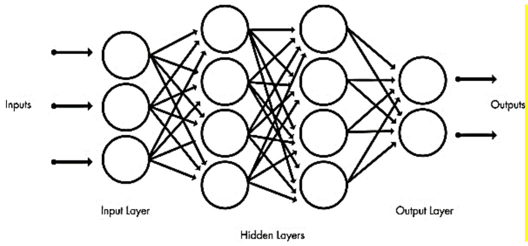

Figure 10. CNN

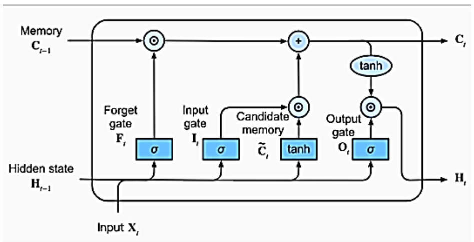

Figure 11. Structure of LSTM

long periods, each LSTM cell has the mechanisms to control the flow of information through gates. Forget gate decides the memory cell state to discard, input gate decides which value from the input is used to update the cell state and output gate controls the output from the memory cell. Figure 11 shows the structure of LSTM.

4.2 Saliency and Energy Level Map

The saliency map is a method that highlights the movement of individuals and it can be generated by calculating optical flow to obtain motion vectors of pedestrians. The optical flow method has been utilized for object detection [9]. Saliency map detection is particularly suitable for capturing behaviors that involve higher velocities such as running, falling, or jumping. It can also be applied to detect crowd behaviors. In [20], a saliency map is employed for detecting suspicious human behavior. Additionally, [1] introduces another type of saliency map known as the kinetic energy map. By calculating optical flow and obtaining velocity vectors, kinetic energy can be derived and used to generate an energy model.

4.3 Radar Detection

Radar sensor can be used in the usage of behavior detection due to its intrusive nature and privacy advantages over cameras [30]. Radar system emits radio wave and measure the time taken to bounce back after hitting an object. When the object is moving, doppler effect happens due to the frequency of the reflected wave changed. Radar sensors can measure the velocity of the object motion. Using CNN model to learn the doppler pattern collect from the radar sensor. Doppler information represent the rate of change over time, therefore doppler radar can suppress the stationary clutter and emphasize a human movement [31]. Authors of [32] introduce a Wi-Fi radar system to detect human behaviors. By using Wi-Fi radar, the targets are no needed to wear any sensors on the body to detect the behavior.

4.4 Challenges and Solutions of Detection Methods in Real-World Applications

In real-world applications, the detection methods face challenges and require solutions to solve the challenges. A problem of adaptability to the environment is a challenge of applying the detection in real-world. Various of weather and lighting will cause the struggling of the detection to adapt. To solve this, parameters based on real-time data to increase the accuracy of the detection. Data collection from different end device and analyze in a main server also can solve the various situation of environment.

When the real-time processing is needed for the real-time data for the adaptability of the detection, challenges occur to enable the timely decision-making. The solutions of the real-time processing are using high-performance computing hardware, and optimize the algorithm for efficiency such as pruning or parallel processing. These methods can enable the capabilities of real-time detection.

5. EMBEDDED MACHINE LEARNING

Embedded machine learning has gained significant attention among researchers, particularly in the context of unmanned aerial vehicles (UAVs). There is a growing need for novel solutions to enable embedded machine learning on these mobile platforms [33]. One of the main challenges in embedded Machin learning is achieving high output accuracy while minimizing power consumption [34]. Despite this challenge, the embedded machine-learning approach offers a mobile platform for deploying machine-learning models across diverse applications, including behavior detection and speech recognition.

5.1 You Only Look Once (YOLO)

Embedded machine learning offers the advantage of applying to tiny devices such as microcontrollers. However, microcontrollers often have limited memory size which necessitates the use of smaller and simpler model architectures.

The YOLO model is well-suited for object detection on embedded devices due to its fast speed and compact design [34]. YOLO is a real-time end-to-end object detection method developed by Joseph Redmon et al. Unlike CNN, YOLO not uses multiple stages but only uses a single CNN to predict bounding boxes and classes directly from full images in one evaluation. Figure 9 shows how the Yolo algorithm detects objects by a grid. The grid cells are resposible to detect the object's center that located in the cell. The grid cell will predicts bounding boxes and confidence scores in the image. Confidence scores stand as the accuracy of the bounding box and contains an object.

There are several versions of YOLO available, including YOLO-UAVlite [35] and YOLOv3-CSP [36]. YOLO is a fully CNN model that processes images using CNN as its backbone. It is a one-stage detector, meaning it analyzes the image in a single pass, making it computationally efficient and faster. However, YOLO has limitations in detecting small objects and may be less accurate compared to other methods. YOLO-UAVlite for example, is a model specifically designed to improve the detection of small objects from unmanned aerial vehicles, indicating that YOLO can still be enhanced by more effective small object detection.

5.2 SqueezeNet

SqueezeNet is a deep neural network that was designed in 2016 to achieve similar accuracy to AlexNet while having fewer parameters and a compressed size of less than 0.5MB. It is commonly used for recognition applications [37]. Compared to AlexNet, SqueezeNet is a smaller network, consisting of only 1/50 of AlexNet's parameters. Despite its smaller size, SueezeNet maintains its performance and is three times faster than AlexNet. It is a full CNN and is known for its lightweight architecture, making it suitable for various research applications. One of the techniques used in SqueezeNet to reduce computation is replacing 3x3 filters with 1x1 filters and utilizing 1x1 filters as bottleneck layers. This approach allows SqueezeNet to achieve comparable accuracy to larger and more complex models. SqueezeNet has been applied to various computer vision tasks, including image classification, object detection, and segmentation. For example, [38] combines SqueezeNet with MobileNetV2 for product quality inspection, while other researchers have modified the last layer of SqueezeNet to enable multiple object detection [39]. Figure 12 shows the model of a SqueezeNet algorithm.

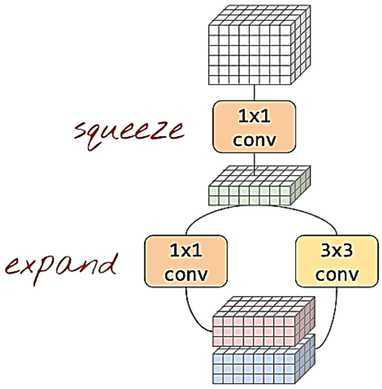

Figure 12. General model of a SqueezeNet

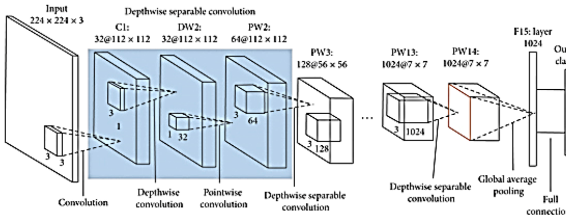

Figure 13. MobileNet Algorithm

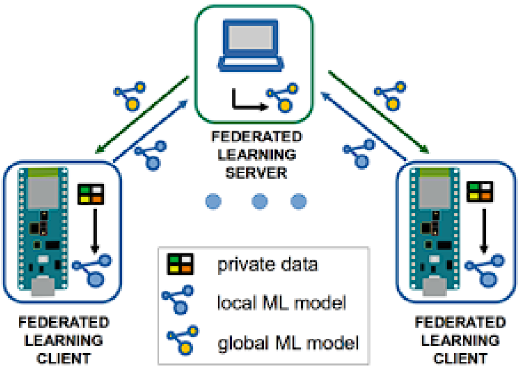
FEDERATED LEARNING SERVER
FEDERATED LEARNING CLIENT
FEDERATED LEARNING CLIENT
private data
local ML model
global ML model

Figure 14. Federated learning concept on microcontrollers [3]

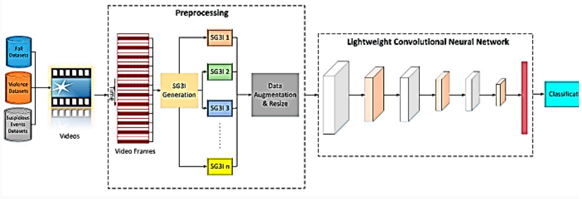

Figure 15. LightAnomalyNet Framework [6]

5.3 MobileNet

MobileNet is a lightweight model developed by Google researchers in 2017, specifically designed to enhance training speed and reduce convolutional computation [39]. It is specifically optimized for mobile and embedded devices that have limited computational resources. The MobileNet architecture achieves its lightweight nature by employing depth-wise separable convolutions and pointwise convolutions, resulting in higher accuracy for image classification and object detection tasks while reducing the number of parameters and computational complexity [39].

One of the significant advantages of MobileNet is its adaptability to memory and latency constraints, making it well-suited for resource-limited devices [40]. It offers small model sizes, low latency, and low power consumption, meeting the requirements of various tiny devices. MobileNet has several versions, including MobileNet V1, MobileNetV2, and MobileNetV3 [41, 42]. For example, in [43], a MobileNet-SSD architecture is utilized for tracking and detecting individuals. Overall, MobileNet's lightweight design and optimization make it an ideal choice for efficient and accurate image classification and object detection on mobile and embedded devices with limited resources. Figure 13 is the model of a MobileNet.

5.4 Federated Learning

Federated learning is a concept introduced by Google, where a machine learning model is constructed using a global dataset. This model is then distributed to individual end devices for local application. The key idea behind federated learning is to enable multiple parties to collaboratively train a shared model while keeping their local data decentralized. Rather than sharing raw data with a central server, participants remotely contribute their data to train a single deep-learning model.

The process involves updating the model locally on each end device and sending back the model's configuration to the cloud. These updates are then aggregated and averaged to create an improved central model. Importantly, federated learning maintains data privacy by not sharing the end device's dataset with the central server. Instead, only the model parameters are shared, allowing for the creation of a new model based on these parameters [44]. Federated learning has been utilized to train machine learning models on microcontrollers [3]. This approach ensures that the privacy of the end device's dataset is preserved, making it possible for companies to leverage valuable data for model training without uploading the data to a central cloud server. Federated learning offers a novel and secure method for collaborative model training while addressing privacy concerns. Figure 14 shows how federated learning works by transferring local model and global model.

5.5 LightAnomalyNet

LightAnomalyNet is a framework proposed by researchers [6] to provide an efficient solution with lower computational loads. It addresses the challenge of extracting features from video frames without relying on computationally intensive methods such as optical flow. The framework is designed to have a smaller memory size, making it suitable for embedded machine-learning applications.

LightAnomalyNet employs a lightweight convolutional neural network (CNN) architecture to differentiate between normal and abnormal events. Unlike conventional machine learning models, it doesn't require a large dataset for training and can achieve good performance even with small or medium-sized datasets. The model utilizes a low-cost approach for modeling motion features by stacking grayscale 3-channel images. Grayscale images with three channels are used because they exhibit lower occlusion in uncrowded scenes, allowing for effective capture of motion details. To reduce computational costs, LightAnomalyNet utilizes a 2D CNN structure instead of a 3D CNN. This design choice helps optimize the computational efficiency of the model. The researchers compared the performance of LightAnomalyNet with existing methods and achieved remarkable results, demonstrating the effectiveness of the framework. Figure 15 shows the framework of LightAnomalyNet.

Table 2. Summary of Existing Works in Suspicious Human Detection

Ref.	Title	Year	Definitions of suspicious behavior	Dataset	Detection Method	Limitations
[9]	Video crowd detection and abnormal behavior model detection based on machine learning method	2018	Unusual Movement Pattern	UMN Dataset	Vector Mapping	The method used is computationally heavy.
[1]	Energy Level-Based Abnormal Crowd Behavior Detection	2018	Unusual Movement Pattern	UMN Dataset	Energy Mapping	The method used is computationally heavy.
[14]	Hawk-Eye: An AI-Powered Threat Detector for Intelligent Surveillance Cameras	2021	Unusual Object	Weapon Dataset	Image Recognition	The experiment is tested in a controlled environment.
[20]	Temporal Saliency-Based Suspicious Behavior Pattern Detection	2020	Unusual Activities	UMN Dataset	Saliency Mapping	The method used is computationally heavy.
[8]	Early Detection of Suspicious Behaviors for Safe Residence from Movement Trajectory Data	2022	Unusual Movement Pattern	TucityLife Trajectory Dataset	Trajectory Data Analyze	Having risk on personal privacy.
[15]	ArchCam: Real-time expert system for suspicious behavior detection in ATM site	2018	Unusual Object	Mock ATM activities Dataset	Saliency Map	The method used is computationally heavy.
[6]	LightAnomalyNet: A Lightweight Framework for Efficient Abnormal Behavior Detection	2021	Unusual Activities	UR Fall, Avenue, Mini-Drone Video, and Hockey Fights Datasets	Image Recognition	The unusual activity only focuses on human movement and pattern.
[45]	Abnormal Crowd Behavior Detection Using Motion Information Images and Convolutional Neural Networks	2020	Unusual Activities	UMN Dataset	Motion Information Image Recognition	The paper focuses on crowd activity.
[46]	Dynamic Human Behaviour Pattern Detection and Classification	2019	Unusual Pattern	Human behavior pattern detection dataset	Image Recognition	The paper only detects human patterns.
[13]	Real-time abnormal behavior Recognition and Monitoring system based on Panoramic Video	2020	Unusual Pattern	-	Human Pose Recognition	The unusual activity only focuses on human movement and pattern.
[27]	Abnormal behavior detection algorithm based on multi-branch convolutional fusion neural network	2023	Unusual Activities	UCF-Crime dataset	Video Recognition	The method used is computationally heavy.
[5]	A hybrid CNN and LSTM-based deep learning model for abnormal behavior detection	2022	Unusual Activities	Fall Detection Dataset	Video Recognition	Maybe cannot detect crowd behavior or lots of people moving simultaneously

ACKNOWLEDGMENT

This work has been supported by Universiti Teknologi Malaysia through Hi-Tech (F4) Q.J130000.4623.00Q15 grant in the project F4 3.3: Edge Artificial Intelligence for Suspicious Human Activity Recognition.

REFERENCES

[1] Zhang, X., Zhang, Q., Hu, S., Guo, CS., Yu, H., *Energy Level-Based Abnormal Crowd Behavior Detection.* Sensors, 2018. **18**(2).

[2] Fath U Min Ullah, K.M., Ijaz Ul Haq, Noman Khan, Ali Asghar Heidari, *AI-Assisted Edge Vision for Violence Detection in IoT-Based Industrial Surveillance Networks.* IEEE Transactions on Industrial Informatics, 2022. **18**(8): p. 5359-5370.

[3] Llisterri Giménez, N., et al., *On-Device Training of Machine Learning Models on Microcontrollers with Federated Learning.* Electronics, 2022. **11**(4): p. 573.

[4] Mehran, R., A. Oyama, and M. Shah. *Abnormal crowd behavior detection using social force model.* in *2009 IEEE conference on computer vision and pattern recognition*. 2009. IEEE.

[5] Chang, C.W., C.Y. Chang, and Y.Y. Lin, *A hybrid CNN and LSTM-based deep learning model for abnormal behavior detection.* Multimedia Tools and Applications, 2022. **81**(9): p. 11825-11843.

[6] Mehmood, A., *LightAnomalyNet: A Lightweight Framework for Efficient Abnormal Behavior Detection.* Sensors, 2021. **21**(24).

[7] Bermejo Nievas, E., et al. *Violence Detection in Video Using Computer Vision Techniques*. in *Computer Analysis of Images and Patterns*. 2011. Berlin, Heidelberg: Springer Berlin Heidelberg.

[8] Cheng, J., Zhang, X., Chen, X., Ren, M., Huang, J., Luo, P., *Early Detection of Suspicious Behaviors for Safe Residence from Movement Trajectory Data.* ISPRS International Journal of Geo-Information, 2022. **11**(9).

[9] Xie, S.C., X.H. Zhang, and J. Cai, *Video crowd detection and abnormal behavior model detection based on machine learning method.* Neural Computing & Applications, 2019. **31**: p. 175-184.

[10] Alvar, M., et al., *Abnormal behavior detection using dominant sets.* Machine Vision and Applications, 2014. **25**(5): p. 1351-1368.

[11] Martinez, D., H. Loaiza, and E. Caicedo, *Algorithm For Early Threat Detection By Suspicious Behavior Representation.* Ieee Latin America Transactions, 2020. **18**(5): p. 825-832.

[12] Jebur, S.A., et al., *Review on Deep Learning Approaches for Anomaly Event Detection in Video Surveillance.* Electronics, 2023. **12**(1).

[13] Li, J.G., et al. *Real-time Abnormal Behavior Recognition and Monitoring System Based on Panoramic Video*. in *39th Chinese Control Conference (CCC)*. 2020. Shenyang, PEOPLES R CHINA.

[14] Ahmed, A.A. and M. Echi, *Hawk-Eye: An AI-Powered Threat Detector for Intelligent Surveillance Cameras.* Ieee Access, 2021. **9**: p. 63283-63293.

[15] Lee, W.K., et al., *ArchCam: Real time expert system for suspicious behaviour detection in ATM site.* Expert Systems with Applications, 2018. **109**: p. 12-24.

[16] Gorodnichev, M.G., et al. *Research and Development of a System for Determining Abnormal Human Behavior by Video Image Based on Deepstream Technology*. in *2022 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*. 2022.

[17] Marsiano, A.F.D., I. Soesanti, and I. Ardiyanto. *Deep learning-based Anomaly Detection on Surveillance Videos: Recent Advances*. in *2019 International Conference of Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*. 2019.

[18] Pawar, K. and V. Attar, *Deep learning approaches for video-based anomalous activity detection.* World Wide Web-Internet and Web Information Systems, 2019. **22**(2): p. 571-601.

[19] Bouma, H., et al. *Flexible human-definable automatic behavior analysis for suspicious activity detection in surveillance cameras to protect critical infrastructures*. in *Conference on Counterterrorism, Crime Fighting, Forensics, and Surveillance Technologies II*. 2018. Berlin, GERMANY.

[20] Cheoi, K.J., *Temporal Saliency-Based Suspicious Behavior Pattern Detection.* Applied Sciences-Basel, 2020. **10**(3).

[21] Blunsden, S. and R. Fisher, *The BEHAVE video dataset: ground truthed video for multi-person behavior classification.* Annals of the BMVA, 2010. **4**(1-12): p. 4.

[22] Lu, C., J. Shi, and J. Jia. *Abnormal event detection at 150 fps in matlab*. in *Proceedings of the IEEE international conference on computer vision*. 2013.

[23] Fisher, R.B. *The PETS04 surveillance ground-truth data sets*. in *Proc. 6th IEEE international workshop on performance evaluation of tracking and surveillance*. 2004.

[24] Soomro, K., A.R. Zamir, and M. Shah, *UCF101: A dataset of 101 human actions classes from videos in the wild.* arXiv preprint arXiv:1212.0402, 2012.

[25] Aqil, M., et al., *Rapid Detection of Hybrid Maize Parental Lines Using Stacking Ensemble Machine Learning.* Applied Computational Intelligence and Soft Computing, 2022. **2022**.

[26] Krizhevsky, A., I. Sutskever, and G.E. Hinton, *ImageNet Classification with Deep Convolutional Neural Networks.* Communications of the Acm, 2017. **60**(6): p. 84-90.

[27] Xu, Z. and Y.Y. Lu, *Abnormal behavior detection algorithm based on multi-branch convolutional fusion neural network.* Multimedia Tools and Applications.

[28] Kim, K., et al., *Lightweight and Energy-Efficient Deep Learning Accelerator for Real-Time Object Detection on Edge Devices.* Sensors, 2023. **23**(3).

[29] Mansour, R.F., et al., *Design of Integrated Artificial Intelligence Techniques for Video Surveillance on IoT Enabled Wireless Multimedia Sensor Networks.*

International Journal of Interactive Multimedia and Artificial Intelligence, 2022. **7**(5): p. 14-22.

[30] Jin, F., et al. *Multiple patients behavior detection in real-time using mmWave radar and deep CNNs*. in *2019 IEEE Radar Conference (RadarConf)*. 2019. IEEE.

[31] Kim, Y., S. Ha, and J. Kwon, *Human detection using Doppler radar based on physical characteristics of targets.* IEEE Geoscience and Remote Sensing Letters, 2014. **12**(2): p. 289-293.

[32] Zou, Y., et al., *Wi-Fi radar: Recognizing human behavior with commodity Wi-Fi.* IEEE Communications Magazine, 2017. **55**(10): p. 105-111.

[33] Martinez-Alpiste, I., et al., *Smartphone-based object recognition with embedded machine learning intelligence for unmanned aerial vehicles.* Journal of Field Robotics, 2020. **37**(3): p. 404-420.

[34] Mazzia, V., et al., *Real-Time Apple Detection System Using Embedded Systems With Hardware Accelerators: An Edge AI Application.* Ieee Access, 2020. **8**: p. 9102-9114.

[35] Liu, C., et al., *A Lightweight Object Detector Based on Spatial-Coordinate Self-Attention for UAV Aerial Images.* Remote Sensing, 2023. **15**(1).

[36] Zhang, G.R., et al., *Lightweight YOLOv3 Algorithm for Small Object Detection.* Laser & Optoelectronics Progress, 2022. **59**(16).

[37] Wentao, Z., G. Lan, and Z. Zhisong. *Garbage Classification and Recognition Based on SqueezeNet*. in *2020 3rd World Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM)*. 2020.

[38] Albanese, A., et al., *Tiny Machine Learning for High Accuracy Product Quality Inspection.* Ieee Sensors Journal, 2023. **23**(2): p. 1575-1583.

[39] Liu, X., et al., *Comparison between three convolutional neural networks for local climate zone classification using Google Earth Images: A case study of the Fujian Delta in China.* Ecological Indicators, 2023. **148**.

[40] Gorospe, J., et al., *A Generalization Performance Study Using Deep Learning Networks in Embedded Systems.* Sensors, 2021. **21**(4).

[41] Wang, L., et al., *Classification of Breast Lesions on DCE-MRI Data Using a Fine-Tuned MobileNet.* Diagnostics, 2023. **13**(6).

[42] Yin, X., et al., *Recognition of grape leaf diseases using MobileNetV3 and deep transfer learning.* International Journal of Agricultural and Biological Engineering, 2022. **15**(3): p. 184-194.

[43] Cob-Parro, A.C., et al., *Smart Video Surveillance System Based on Edge Computing.* Sensors, 2021. **21**(9).

[44] Yang, Q., et al., *Federated Machine Learning: Concept and Applications.* Acm Transactions on Intelligent Systems and Technology, 2019. **10**(2).

[45] Direkoglu, C., *Abnormal crowd behavior detection using motion information images and convolutional neural networks.* IEEE Access, 2020. **8**: p. 80408-80416.

[46] Wang, S.Q., et al. *Dynamic Human Behavior Pattern Detection and Classification*. in *5th IEEE International Conference on Big Data Computing Service and Applications (IEEE BigDataService) / Workshop on Big Data in Water Resources, Environment, and Hydraulic Engineering / Workshop on Medical, Healthcare, Using Big Data Technologies*. 2019. San Francisco, CA.