International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 2, Mar-Apr 2025
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Human Trespassing Detection System through Deep Learning

Mr.Ashraf ali shaikh[1], Jagtap Nandini[2], Gatkal Pranali[3], Kathale Apurva[4], Rode Amruta[5]
[*1]Asst.prof. Department of Computer Technology, Padmashri Dr. Vitthalrao Vikhe Patil Institute of Technology & Engineering(Polytechnic), Loni, Maharashtra, India
[*2,3,4,5]Department of Computer Technology, Padmashri Dr. Vitthalrao Vikhe Patil Institute of Technology & Engineering(Polytechnic), Loni, Maharashtra, India

--************************----------------------------------

Abstract:
Using CCTV footage allows for real-time danger monitoring. In order to deter criminal activity, video surveillance is used to keep an eye out for suspicious conduct. Nowadays, most store owners have CCTV cameras to catch these suspicious activities, but these systems can't tell when someone is trespassing. The lack of a faster and more effective method to detect theft would make the task difficult. At this time, the market offers systems that include trip wires, sensors, and CCTV. This proposal details methods for detecting and catching shoplifters in confined spaces. In this proposal, we want to transform passive monitoring into active surveillance by utilizing AI and deep learning approaches. We will assess photos, identify unwanted things, and alert them using neural networks.

Keywords: Sensor, CCTV, Artificial Intelligence, Deep Learning, Neural Network.

--************************----------------------------------

I. INTRODUCTION

The global incidence of theft and other forms of intrusion has been steadily rising over the past few years. Because these thefts can cause financial instability and, in extreme cases, bankruptcy, they are extremely harmful to both individuals and organizations. A number of businesses and homes use CCTV cameras to dissuade would-be intruders and keep an eye on things. As static ways of intrusion detection, these technologies have served their purpose well.

Static approaches are those that only capture the act effectively and do not offer any countermeasures. In order to avoid these kinds of attacks, capturing tasks is not very useful. An active mechanism is required to reliably identify when someone is trying to get into the area. Present methods have proven successful in delivering monitoring, however this is insufficient for the identification of any questionable behavior. The paradigm of image processing provides the best opportunity for doing activity detection. By utilizing image processing techniques, the system can be equipped to detect and identify the intruder. The intruder identification is one of the most difficult techniques that can be very problematic for the computer to perform automatically. These problems arrive with the inconsistencies and the large number of human postures that need to be identified with accuracy. The detection model must be trained accurately and should be able to identify the human intruder in any position or posture.

The image processing alone cannot be tasked with such a great responsibility. The combination of image processing along with machine learning techniques improves the detection accuracy. The increase in accuracy entails the effective technique that can be highly reliable for the detection of the intruder. the image processing approach effectively provides surveillance along with the detection of the intruder. Whereas the machine learning approaches determine the suspicious activities that are being performed by the detected intruder. The proposed methodology employs the use of Convolutional Neural Networks along with the TensorFlow framework to facilitate the implementation of an effective intruder detection approach.

[1] The necessary Human motions must be recognized in order for An novel and thorough integration of the Hungarian job allocation system with the H-net neural network has produced frame capturing model, which can be influenced by the work of Shubham gade et al. In order to improve the outcomes of the slot allocation, the author will now utilize the sequence list. Thanks to this vital information, the scheduling method used to find the closest electric vehicle charging station may be fine-tuned in real-time. In this step of the choice tree procedure, author acquire a sorted list. The decision-tree approach uses this list to find the closest electric vehicle charging station and see whether there are any available spaces. Frame grabbing to identify the human motion is scheduling at good speed is one of the short-term and long-term uses of these technologies.

[2] A novel approach to heart rate detection via 77 GHz FMCW radar was put out by ZEKUN CHEN et al. The author suggests using the APSEM method— Adaptive Parameter Selection for Expectation- Maximization—to deal with artifacts caused by human motion. By supplementing the conventional EM approach with an adaptive initial distribution parameter selection, this method makes motion artifacts filtering more accurate and resilient. The authors also provide an innovative sub-framework for heart rate detection, the KFRV approach, to deal with interference after the APSEM method. To successfully suppress random disturbances like impulse noise and unpredictable body motions, this sub-framework integrates Kalman filtering with the Rife spectral analysis approach and the VME algorithm. Hence, the accuracy of heart rate estimation can be enhanced even further. The author proposes a new technique for heart rate detection that uses MCA, APSEM, and KFRV as its foundation.The author conducted thorough experiments to confirm the efficacy of the proposed strategy. Utilizing heartbeat signals collected from 25 volunteers while subjected to human motion, the suggested method is assessed, and the outcomes demonstrate that the MAE of heart rate detection is below 2.5 bpm.

[3] DCNLSTM, developed by CHHAYA GUPTA et al., is a neural network that integrates multiple methods. In addition to authorll, the study examines various methods. A number of datasets have been utilized to compare different methods with the suggested model; these include UCF101, NTU- RGB-D, KITTI, and NTU-RGB-D 120.Additionally, a fused dataset was created using data fusion techniques. This dataset has 79282 photos from 148 classes and 184680 videos from 281 classes. Additionally, the suggested model is contrasted with others that have been trained on the same datasets: DTR-HAR, CNN with BiLSTM, Classic CNN, and YOLOv6. In terms of accuracy, the proposed model is superior to all previous models. The study is structured into five parts. We present a comprehensive overview of ITS in the second section. A concise introduction to deep learning and its uses is given in Section 3. Section 4 details how smart cities and ITS make use of deep learning to detect pedestrians, summarizes the work of several academics in the area, and lays out the obstacles to further study in each subfield. Next, in Section 5, we shall present the final verdict.

II. LITERATURE SURVEY

[4] A method for detecting unusual human trajectories inside buildings was suggested by DOI THI LAN et al. via a deep learning model. The author's suggested model relied on the Transformer encoder and SOM, which were trained together to learn normal trajectory representations and the latent space clusters of these representations. To understand the relationships between each trajectory's points and its sequence information, the Transformer encoder

equipped with a self-attention mechanism was utilized. The input trajectory was rebuilt from its latent representation by a decoder. The SOM layer was also employed to acquire latentspace cluster representations of normal trajectories. The test trajectory's anomaly score was calculated using the trained model for the purpose of anomaly trajectory detection. Both the quantization error on the SOM and the trajectory reconstruction errors from the latent space were included in the anomaly score of the trajectory. An anomaly was identified in the test trajectory if its anomaly score was greater than a predetermined threshold. An innovative metric, WS, was also suggested by the author as a means of calculating the anomalous threshold; it is a weighted combination of recall and precision. A suitable anomalous threshold was chosen in particular by comparing the validation set's maximum value for WS.

[5] Archangel, a novel UAV-based object identification dataset introduced by YI-TING SHEN et al. [5], aims to inspire further research into improving UAV-based object detection methods by combining dataset metadata with synthetic data. An exhaustive investigation is meticulously planned to demonstrate how to employ Archangel to completely optimize a cutting-edge object detector using a hybrid fine-tuning dataset that combines actual and synthetic data. By comparing the model's performance across various object postures and UAV placements, the author further demonstrates the significant benefit of using the dataset metadata during model evaluation. Although the authors acknowledge that Archangel has a long way to go, they are hopeful that it will serve the machine learning community well and pave the way for future developments in unmanned aerial vehicle (UAV) perception. Though it already has all these great qualities, Archangel still has a long way to go before it reaches its full potential.

[6] Pekka Pääkkönen et al. describe the study's overarching research objective, which sought to assess the practicality and efficacy of technologies for detecting objects and recognizing human poses in order to facilitate robotic applications' context awareness in an edge computing setting. Answering the research topic and making a new contribution might be achieved by evaluating the feasibility and efficiency of Yolov5-based object identification and Move net-based human pose recognition on the Jetson AGX Xavier platform. Possible supplementary contributions include algorithms for activating the concepts of objects and human stance.

[7] A method for detecting human sitting posture using FMCW radar was proposed by GUOXIANG LIU et al. The technique achieves an average detection accuracy of 98.07% by using a human sitting posture feature extraction algorithm to extract distance, angle, and Doppler information from human sitting posture data, and then by using support vector machines to classify five frequent states of human sitting posture. Concurrently, we suggest a single-target human sitting posture recognition approach to enhance target detection accuracy while reducing phase error due by distance FFT bin drift. The article's suggested method successfully identifies five different sitting positions and provides a sitting history, according to experimental data. Limitation\Future Scope— The author plans to utilize deep learning to recognize more complicated sitting positions and to further broaden the sorts of sitting postures.

[8] A major step forward in the area of Human Action Recognition (HAR), which combines computer vision with wireless computing, is MISHA KARIM et al. Current approaches and data sets have improved HAR significantly, but they aren't without its flaws. In order to tackle these issues, the author has put up the HADE framework, which builds upon the strengths of earlier methods. The HADE architecture incorporates a large dataset mainly from smartphone cameras and is known for its groundbreaking jump in HAR. This method records a broad range of human motions that have been processed by the innovative HADE I and HADE II models. To improve HAR systems' accuracy, precision, recall, and F1-score, these

models use GPU parallel processing and sophisticated machine learning methods.

[9] A framework for human-object interaction in real-life situations, such sports and other activities, is described by ISRAR AKHTER et al. [9]. So, using the MPII dataset, the author accomplished a remarkable accuracy of 88.46% in human body component detection. The UCF aerial dataset has it at 82.00%, while the wild data set for sports video has it at 88.30%. Classification accuracy ranges from 92.60% in the UCF_aerial dataset to 92.42% in the wild sports video dataset, with the MPII dataset having the highest accuracy at 92.71%. The author plans to incorporate additional composite activities from different contexts, such as smart homes, workplaces, and medical centers, in future work. The author plans to recognize complicated motion patterns in many circumstances by fusing more feature engineering techniques from diverse fields.

[10] In this review, the authors DONGHYEON NOH et al. sought to classify the most up-to-date innovations in HMR techniques into two broad groups: VS methods and WS methods. The research was organized chronologically within each area, with additional subcategorization into sensors, algorithms, datasets, target body parts, gesture kinds, and recognition performances. While this paper's method-focused review style is helpful for understanding current research trends, future reviews should center on how these methods might be applied in real-world industries. The majority of the HMR studies presented here either failed to account for potential commercial applications or relied too much on dataset-based simulations without also validating their results experimentally using physical systems.

[11] In their discussion of human mobility monitoring with Ultra-Wide Band (UWB) technology, thottempudi pardhu et al. cite a number of different approaches. Researchers have pushed the boundaries with approaches based on CNN and more advanced RMDL approaches based on SGWO. Yet, the road has not been easy. Researchers have attempted to overcome a number of obstacles, including the complexity of motion state overlaps, the need for higher spatial resolution, the difficulty of detecting against noisy backgrounds, and the requirement to learn from varied data formats. Based on the comparison of the various methods, the RMDL method based on SGWO is the most effective, outperforming the others in terms of accuracy, TNR, TPR, and Mean Squared Error (MSE). Their exceptional speed and precision in flaw detection after the aforementioned procedure has also set a new standard. Finding the best technologies for detecting human movements will definitely continue from here. current scope. The capacity of security and surveillance systems will be improved, and new opportunities for innovation in smart home tech, disaster response, and healthcare monitoring will be created, as this area continues to advance. While the TwSense model was novel, it was limited in its ability to detect human breathing patterns since it needed extensive validation of the through-the-wall scenario.

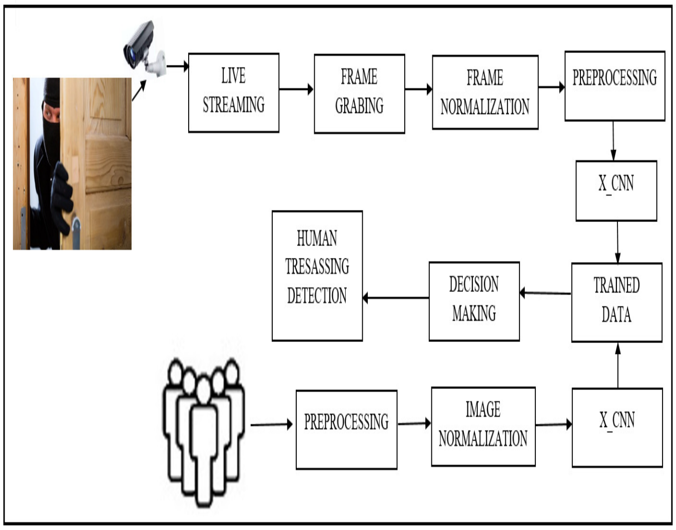
[12] There has been a lot of discussion in the literature about the effects of HAR on everyday living, real-time circumstances, and collaborative activities (MISHA KARIM et al., 2012). The present status of HAR systems is thoroughly examined in this article. Some important points are brought to light by the results from the different sources that are referenced in this article. Tabulated in Table 7 is also a synopsis of the HAR method and technique taxonomy. Several facets of HAR have been the subject of research. These include user behavior, lifestyles, and both static and dynamic living activities. There has been little focus on real-time HAR for healthcare security, though. Problems with developing HAR systems include data scarcity, technical constraints, and the complexity of real-time activities. It takes a lot of processing power to keep up with the demands of reacting to changing surroundings and public or populated locations in real time. Current research also fails to address real- time activities, and there is a lack of high-quality real-time data.

[13] Human occupancy detection has many uses, including security, intrusion detection, energy management systems for smart buildings, and e-health systems (REZA SHAHBAZIAN et al., 2013). Various sensor-and vision-based approaches to human presence identification in indoor settings have been investigated in the literature. These approaches use either simple detection methods or algorithms based on machine learning. While vision- based approaches often lead to privacy concerns, sensor-based systems typically require supplementary hardware, data gathering and recording tools, and analysis software for decision- making. Problems arise in the real world when trying to install and maintain sensors in the best possible way. Wi-Fi and Bluetooth, which rely on radio waves for sensing, have recently attracted a lot of interest from researchers. This study delves into the current standards, techniques, and system designs of radio frequency occupancy detection, as well as their guiding concepts. utilizing open-source code, existing datasets, and data-gathering methodologies for performance evaluations was the subject of the author's investigation into the practical challenges. According to the author's analysis of research conducted between 2018 and 2022, Wi-Fi sensing remains a hot subject in the field of applied wireless networking.

[14] In their description of novel perception metrics, Ken T. Mori et al. outline the development of interpretable pass/fail criteria as author thresholds. In order to establish interpretable analytical perception standards related to safety, human performance is taken into account as the author's conservative estimations. We present and implement a motion prediction network-based validation approach. Although the results back up the claims made in this paper, they also demonstrate that the existing metrics are invalid. Current detector evaluations reveal that, for most items in a modern dataset, the criteria are not satisfied. Matching failures predominate in the data, even though the localization criteria are looser han typical AP matching standards. Using the criteria defined in this work, the disparities between different modalities and detectors are minimal. Current measures may place too much emphasis on fine-grained localization, according to this. This demonstrates how important it is to evaluate and develop perception algorithms with safety in mind. Future research should focus on developing detectors that take use of this fact as an author, while other details of the review process are left for now.

[15] In their publication Uniss-FGD, PIETRO RUIU et al. present a new dataset that records eye movements in relation to facial pictures. Neuroscience, physiology, psychology, human-robot interaction, and computer science are just a few of the many scientific disciplines that could benefit greatly from the collected data. When it comes to activities like recognition, humans are incredibly efficient compared to automated computer systems. Therefore, recording human gaze can be incredibly helpful for training and optimizing neural network-based automatic recognition systems.

III. METHODOLOGY



The method that has been suggested to establish a Human Tress passing System depicted in the system overview in Figure 1 up top. The suggested method was based in part on the execution of the procedures detailed below.

Step 1: Data collection and Preprocessing: You can find a dataset that has been properly labeled as human or non-human at this URL: https://www.kaggle.com/datasets/constantinwerner/ human-detection-dataset/data. Both human and non- human sorts of photos are abundant in this dataset. There are two directories for this dataset: test and train. Each directory has a number and contains photos from the dataset. The photographs are resized according to their folder hierarchy using the python programming language once they have been classified. Here, the paths to the directories are obtained by recursively traversing them. Using the CV2 object of OpenCV, the images in the obtained path are scaled to 128 X 128 dimensions. In order to learn correctly, the objects in the scaled images are turned into grayscale. To get the right dataset for training the X-Convolution neural network, the photos are re-stored in the same directory after being resized and grayscaled.

Step 2: Training through X- convolution neural network: The first step in training a human dataset is to establish the test and train routes. Then, for 32 batches and 20 epochs, a total of 921 test and train images are utilized. To make test and train data objects, an image data generator object is made with a depth ratio of 1:255. With this ratio, the pixels will be analyzed to their furthest depth. The train and test data objects have the necessary parameters inserted into them. The size of the objects is 128 X 128 and the batch size is 23, and the mode of operation is grayscale. Class mode with a categorical value is loaded as the last argument. For images with 128 x 128 dimensions, a convolutional neural network model is constructed using a sequential neural networkas a representation for the grayscale color channel denoted as 1. This is followed by the first layer's model, which has 32 3 X 3 kernels. Then, a second convolutional layer with 64 3X3 kernels is added to the second layer. The "Relu" Activation Function powers the top and bottom layers, respectively. To gather the neurons in a 2D matrix, the following step is to add a single maxpooling 2D layer to the model, with a kernel size of 2 X 2. A dense layer with a dropout rate of 25% ends this max pooling 2D layer. The third layer consists of a max-pooling 2D layer with a 2X2 kernel size and 128 3 X 3 kernels activated using the "Relu" function. Finally, a Dropout layer with a dropout percentage of 25% is applied in the same way to the fourth layer. The next step is to flatten the network as a whole such that the neurons can be collected in a dense layer of 1024 by 50% dropout. In the end, we use the activation function "Softmax" to gather training data for the two categories of humans and non-humans. While training with the data, an optimizer named "Adam" uses category cross entropy to estimate the losses. Finally, the trained data is obtained in an.h5 file by invoking the model for 20 epochs using the fit generator function. In picture 14, we can see the convolution neural network architecture.

LAYER	ACTIVATION
conv 2D 32 (3X3)	Relu
conv 2D 64 (3X3)	Relu
MaxPooling 2D 2X2	
DropOut 0.25	
conv 2D 128 (3X3)	Relu
MaxPooling 2D 2X2	
conv 2D 128 (3X3)	Relu
MaxPooling 2D 2X2	
DropOut 0.25	
Flatten	
Dense 1024	Relu
DropOut 0.5	
Dense 2	Softmax
Optimizer	Adam
batch_size	32

Figure 2: X- CNN network Architecture

Step 3: *Human TressPassing detection and alert generation* : In this stage, the Python software uses the Droid Cam app, which is compatible with both laptops and mobile phones, to record video and, by extension, frames from the mobile phone's camera. Intruders can be identified in the live streaming frames using the trained model file.h5. The positions of the intruders in the upper left rectangular area are retrieved from this file. This role is responsible for keeping an eye out for human intruders, taking pictures of them, and then sending them to the owner of the premises using WhatsApp with the help of the Pywhatkit library. On top of that, a voice alert goes out to tell the burglars to get out of the building right away.

IV. RESULTS AND DISCUSSIONS

The Anaconda framework for Python, and the Spyder IDE were used to construct the suggested Deep Learning approach for detecting human trespassing. The development computer features eight gigabytes of main RAM and one terabyte of secondary memory. The viability of the deployed idea has been assessed by taking into account a variety of elements. Here, we describe the experimental study's findings in detail through Accuracy an:d loss of the model

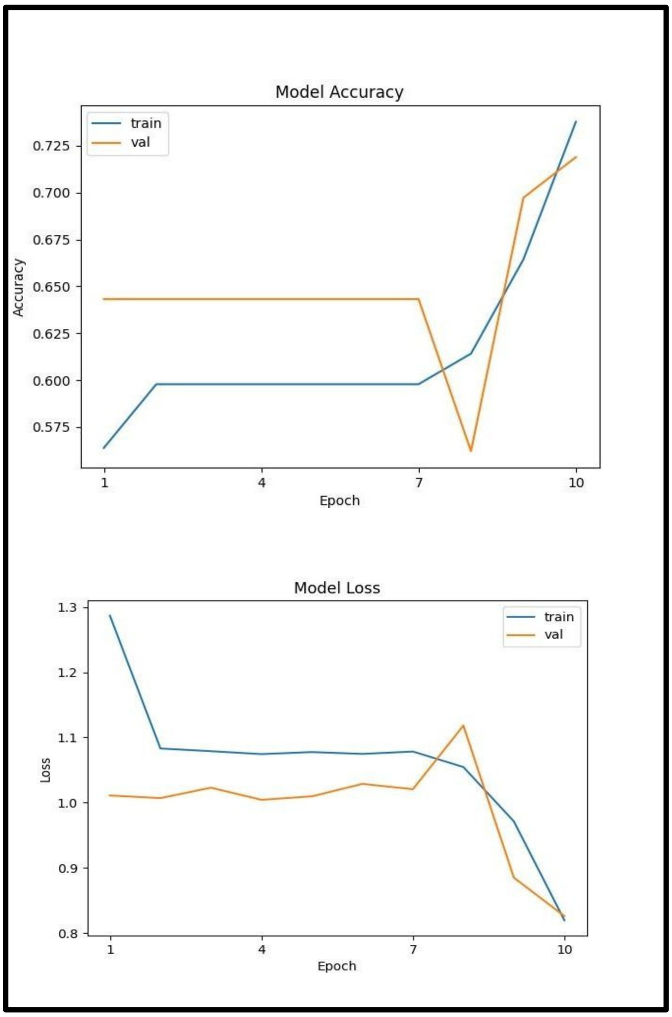


Figure 4 : accuracy and loss model

V. CONCLUSION AND FUTURE SCOPE

The growing demand for video surveillance has prompted new developments in the detection of human intrusion/trespassing and theft. Using extracted surveillance footage with recognized human intrusion, the proposed methodology attempts to efficiently identify trespassing and human intrusion and warn the owner. To train the model, the suggested technique has made use of two separate datasets, each of which contains more than 2.5K photos. Human intrusion detection can be seen as a binary classification problem involving the identification of non-human entities using the Virgate dataset. Various body positions can be recognized with the aid of the Human Pose dataset, which also facilitates the recording of whole actions into time-stamped images. Data pretreatment steps, including picture cleaning and segmentation and grayscale conversion, are a part of the suggested methodology. The suggested technique achieves an average accuracy score of over 90% when identifying human intrusion/trespassing using a Convolutional Neural Network (CNN) with a pre-processed dataset. Potential for Future Development: Implementation of a real-time home security system is within the system's purview ¬ Multi-location functionality through cloud-based live frame streaming is also within the system's purview

REFERENCES

[1] Shubham Gade, Amita Singh, Shubham Sarote, " Efficient H-net model-based slot assignment solution to accelerate the EV charging station searching process ",ISSN: 2349-6002,Volume:11 ,Issue: 6,PageNo: 2590-2597

[2] zekun chen ,yunxue liu, chenhong sui , min zhou and yuqing song, "A Novel Scheme for Suppression of Human Motion Effects in Non-Contact Heart Rate Detection", Received 21 July 2023, accepted 2 August 2023, date of publication 7 August 2023, date of current version 14 August 2023.

[3] chhaya gupta, nasib singh gill, preeti gulia, sangeeta yadav, giovanni pau, mohammad alibakhshikenari, and xiangjie kong, "A Real-Time 3-Dimensional Object Detection Based Human Action Recognition Model", Received 3 November 2023; accepted 12 November 2023. Date of publication 20 November 2023.

[4] doi thi lan, seokhoon yoon, P. Kothari and A. Alahi, "Anomalous Indoor Human Trajectory Detection Based on the Transformer Encoder and Self-Organizing Map", Received 18 October 2023,accepted 16 November 2023, date of publication 21 November 2023, date of current version 29 November 2023.

[5] yi-ting shen, yaesop lee, heesung kwon, damo m. conover, shuvra s. bhattacharyya, nikolas vale, joshua d. gray, g. jeremy leong, kenneth evensen, and frank skirlo, "Archangel: A Hybrid UAV-Based Human Detection Benchmark With Position and Pose Metadata", Received 2 June 2023, accepted 16 July 2023, date of publication 27 July 2023, date of current version 7 August 2023.

[6] PEKKA PÄÄKKÖNEN, DANIEL PAKKALA,F. Sherwani, M. M. Asad, M. R. Endsley, "Evaluation of Human Pose Recognition and Object Detection Technologies and Architecture for Situation-Aware Robotics Applications in Edge Computing Environment", Received 7 August 2023,accepted 19 August 2023, date of publication 28 August 2023, date of current version 1 September 2023.

[7] guoxiang liu, xingguang li , chunsheng xu, lei ma, and hongye li, "FMCW Radar-Based Human Sitting Posture Detection", Received 20 August 2023, accepted 2 September 2023, date of publication 5 September 2023, date of current version 26 September 2023.

[8] MISHA KARIM, SHAH KHALID, ALIYA ALERYANI, NASSER TAIRAN, ZAFAR ALI, AND FARMAN ALI, "HADE: Exploiting Human Action Recognition Through Fine-Tuned Deep Learning Methods", Received 18 February 2024, accepted 9 March 2024, date of publication 18 March 2024, date of current version 26 March 2024.

[9] ISRAR AKHTER, NAIF AL MUDAWI, BAYAN IBRAHIMM ALABDULLAH, MOHAMMED ALONAZI, AND JEONGMIN PARK, "Human-Based Interaction Analysis via Automated Key Point Detection and Neural Network Model", Received 31 July 2023, accepted 6 September 2023, date of publication 11 September 2023, date of current version 20 September 2023.

[10] DONGHYEON NOH, HOJIN YOON, AND DONGHUN LEE, "A Decade of Progress in Human Motion Recognition: A Comprehensive Survey From 2010 to 2020", Received 10 November 2023, accepted 1 January 2024, date of publication 8 January 2024, date of current version 12 January 2024.

[11] THOTTEMPUDI PARDHU, VIJAY KUMAR, PRAVEEN KUMAR, AND NAGESH DEEVI, "Advancements in UWB-Based Human Motion Detection Through Wall: A Comprehensive Analysis", Received 31 March 2024, accepted 2 May 2024, date of publication 6 May 2024, date of current version 5 July 2024.

[12] MISHA KARIM, SHAH KHALID, ALIYA ALERYANI, JAWAD KHAN, IRFAN ULLAH, AND ZAFAR ALI, "Human Action Recognition Systems: A Review of the Trends and State-of-the-Art", Received 12 February 2024, accepted 26 February 2024, date of publication 7 March 2024, date of current version 13 March 2024.

[13] REZA SHAHBAZIAN, AND IRINA TRUBITSYNA, "Human Sensing by Using Radio Frequency Signals: A Survey on Occupancy and Activity Detection", Received 28 February 2023,accepted 19 April 2023, date of publication 24 April 2023, date of current version 28 April 2023.

[14] Ken T. Mori and Steven Peters, "SHARD: Safety and Human Performance Analysis for Requirements in Detection", IEEE

TRANSACTIONS ON INTELLIGENT VEHICLES, VOL. 9, NO. 1, JANUARY 2024.

[15] PIETRO RUIU, MAURO FADDA, ANDREA LAGORIO, SETH NIXON, MATTEO ANEDDA, ENRICO GROSSO, AND MARINELLA IOLE CADONI, “Uniss-FGD: A Novel Dataset of Human Gazes Over Images of Faces”, Received 11 May 2024, accepted 25 May 2024, date of publication 28 May 2024, date of current version 4 June 2024.

*****