Human Recognition System Using Machine Learning

Swetha M, Sinthu R, Saranya C, Ebina Mathumitha A
Mrs. Sajila L, M.E
Department of Information Technology
Loyola Institute Of Technology and Science Kanyakumari District, Tamilnadu, India.

ABSTRACT

Technologies for human authentication are essential to intelligent systems, automated processes, and security. An Arduino PIR (Passive Infrared) sensor combined with machine learning is used in this project's human recognition system to improve accuracy in differentiating motion sources from human presence. Changes in infrared radiation, which indicate movement, are picked up by the PIR sensor. Machine learning algorithms are then used to interpret the data and determine whether the motion is human. An Arduino microcontroller is used to create the system; it receives signals from PIR sensors and sends information to a computer-based machine learning algorithm. To increase detection accuracy, the model examines motion patterns and ambient variables. To train the system using real-world data, a number of algorithms are investigated, such as Support Vector Machines (SVM), Logistic Regression, and Neural Networks. When the system detects a human being, it can initiate pre-programmed functions like turning on lights, giving out alarms, or recording the incident for safety reasons.

Keywords: Human Recognition, Machine Learning, Biometric Identification, Pattern Recognition, Computer Vision, Face Recognition

1. INTRODUCTION

Supermarkets produce enormous volumes of data in the current retail environment as a result of everyday transactions, consumer contacts, and systems for inventory control. Comprehensive analysis of this information can yield significant findings into market developments, consumer behavior, and purchase patterns. In order for supermarkets to execute decisions based on data that improve their client experience, improve inventory, and optimize economic viability, behavioral analysis of customers and predictive modelling are essential. Businesses may now use sophisticated predictive systems to foresee client needs and go beyond standard analytics thanks to the development of data mining and machine learning techniques. Supermarkets may employ such frameworks to learn which products are often purchased together, when consumer interest in particular products increases, including the shopping habits of specific clients. This makes it possible to achieve better demand forecasts, customized promotions, targeted marketing, and inventory optimization. The main goal of this project is to create an intelligent human identification system that minimizes false detections while properly detecting and classifying human presence using an Arduino PIR sensor and machine learning techniques.

2. LITERATURE SURVEY

Systems for person recognition are frequently utilized in domains like human-computer interaction, security, surveillance, and healthcare. These systems are now more precise, effective, and able to manage challenging jobs because to the development of machine learning (ML), particularly deep learning techniques. Significant improvements to the field's research are highlighted in this survey of the literature.

Machine learning has significantly improved facial recognition systems. suggested a technique for detecting facial identification in encrypted domains that makes use of Local Fisher Discriminant Analysis (LFDA). By employing Local Binary Pattern (LBP) features for quick the extraction process, their method seeks to improve interaction between humans and computers by effectively recognizing emotions. meanwhile there is a difference within the signal and the background, the system's performance suffers [1]. The paper presented Emonets, a multimodal deep learning system for video emotion recognition, in a different study [2].

For real-time recognition in cloud gaming, this method blends deep learning models with emotion-aware screen impacts. Although the system has a great degree of adaptability and lower latency, it has trouble addressing sensitivity delays and reliable cloud-based functionality. People are identified by their walking patterns using recognition of gait. A structure for deep learning that combines CNNs and RNNs for reliable gait recognition was proposed by the author [3].

Considering varied circumstances, such as shifting clothes and transporting circumstances, the model they developed performed better than conventional methods. Another powerful biometric technique is voice recognition. To generate voiceprints for speaker identification, the author employed deep neural networks that had been trained on big datasets. Their research exposes issues with noisy data and spoofing, but it also demonstrates remarkable accuracy in real-world settings [4].

Contemporary technologies combine several modalities to increase confidentiality and reliability. CNN-based characteristics from face and iris images were combined in the research to create a multimodal biometric recognition system that outperformed single-modality techniques [5].

Table 1:Literature review

Title of the Paper	Authors	Year	Methodology	Advantages	Limitations
Facial Expression Recognition in the Encrypted Domain Based on Local Fisher Discriminant Analysis	Caifeng Shan, Shaogang Gong, Peter W. McOwan	2023	Local Fisher Discriminant Analysis (LFDA) with Local Binary Pattern (LBP) for encrypted facial expression recognition	Fast feature extraction using LBP; suitable for real-time systems	Poor performance when signal and background variances differ
EmoNets: Multimodal Deep Learning Approaches for Emotion Recognition in Video	M. Shamim Hossain, Ghulam Muhammad	2022	Deep learning framework combining multimodal emotion-aware screen effects in cloud gaming	Low latency, high flexibility, and good resilience in video processing	Sensitivity delay and challenges in providing a good cloud gaming user experience
A Comprehensive Study on Cross-View Gait Based Human Identification with Deep CNNs	Zhirong Wu et al.	2017	Deep Convolutional Neural Networks (CNNs) for gait recognition across different views	Robust to clothing and view changes; suitable for surveillance	Sensitive to occlusions and environmental noise
VoxCeleb: A Large-Scale Speaker Identification Dataset	Arsha Nagrani, Joon Son Chung, Andrew Zisserman	2017	Deep learning for speaker identification using large-scale voice dataset	High accuracy in speaker verification; scalable to large datasets	Prone to spoofing and affected by noisy environments

A Multimodal Biometric System Using Deep Learning for Image Fusion	Y. Zhang, D. Song, Y. Yao	2019	Multimodal system using CNN for face and iris fusion	High accuracy; better than single-modality systems	Computationally intensive; requires high-quality input data

3. PROBLEM STATEMENT

Conventional PIR (Passive Infrared) sensor-based motion detection systems have become commonplace in monitoring, automation of homes, and security applications. The incapacity to distinguish between individuals and other things, as well as false detections brought on by non-human movements (such as pets, moving curtains, or temperature changes), are some of these systems' major drawbacks. This leads to security flaws, erroneous alerts, and wasteful energy use.

4. PROPOSED SYSTEM

The suggested method integrates the intelligence of machine learning with the effectiveness of hardware-based motion detection to provide a clever human presence detection approach. To identify any movement in the immediate surroundings, it makes use of a Passive Infrared (PIR) sensor that is coupled to an Arduino board. When the Arduino detects motion, it connects to a Python-based PC application, which then turns on a webcam to take a picture of the area. A machine learning model that has already been trained is then used to evaluate this image and assess whether the movement that has been noticed is indicative of humans being present. The method reduces inaccurate results brought on by non-human motion, such pets or moving objects, by combining sensor data with ocular confirmation. Because of its complex structure, which guarantees precise real-time decision-making, the technology can be used for intelligent surveillance, automatic lighting management, and security monitoring. The advantages include

- Enhanced Precision via Dual Confirmation reduces false positives by combining motion detection with machine learning-based image analysis to guarantee that only real human presence is verified.
- Minimal Power Use PIR sensors are particularly well-suited for energy-saving applications since they use little power and enable continuous monitoring.
- Economical Deployment makes use of reasonably priced parts that include an Arduino, a PIR sensor, and a common camera, making it suitable for small-scale implementations scholars, and students.

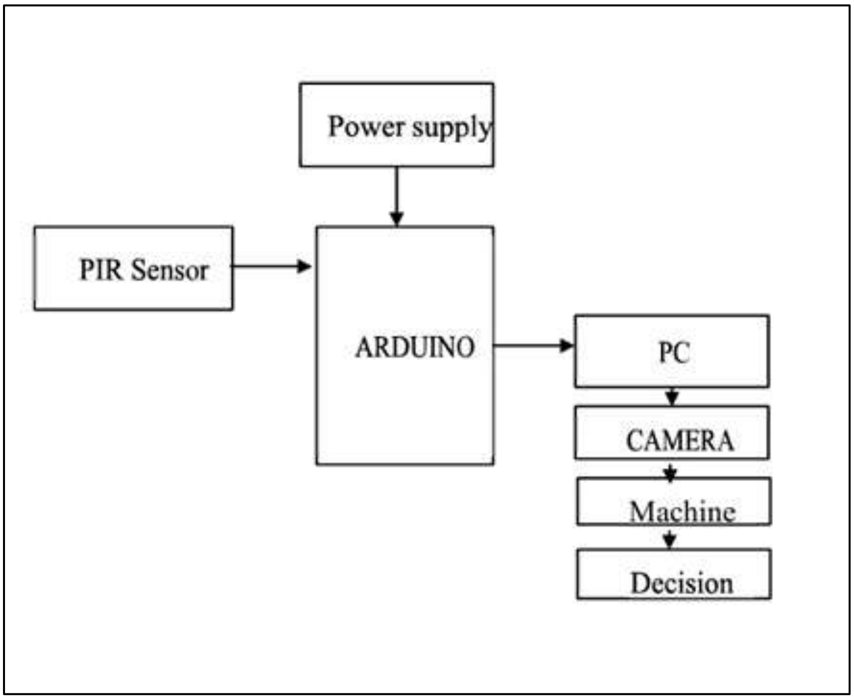

Fig 1: Proposed system

Modules:

- Human Presence Detection.
- Integration of Machine Learning.
- Data Gathering and Processing.
- Arduino System Implementation.
- Real-time Decision Making.
- Reduction of False Positives and Improvement of Accuracy.

Human Presence Detection

Passive Infrared (PIR) sensors, like the HC-SR501, are frequently employed for motion detection because of their capacity to detect variations in the infrared radiation that warm things, like people, release. Since these kinds of sensors are the first line of detection, they are essential to intelligent and environmentally friendly devices. A microcontroller like the Arduino UNO or Nano can receive a HIGH signal from a PIR sensor when it senses movement. Subsequent actions, such turning on a camera or an alert system, can then be initiated using this trigger. By eliminating the requirement for ongoing video surveillance or processing of data and only starting these processes when human movement is detected, the PIR sensor contributes to computing resource conservation. PIR sensors are therefore perfect for uses where quick detection and effective resource management are crucial, such automation projects, intelligent houses, and safety measures.

Machine Learning Integration

The linked PC turns on its camera to take a picture of the identified scene when it receives a HIGH signal from the Arduino that was set off by the PIR sensor. To ascertain if a human being is identifiable, this picture is subsequently examined using a pre-trained machine learning model like YOLOv5, MobileNet, or Haar Cascade. By excluding non-human motion triggers like animals, shadows, or moving things, such models greatly improve the precision of identification. For rapid human identification in confined spaces, Haar Cascade is the perfect technique because it is lightweight and speedy, especially for face detection jobs. Portable and effective, MobileNet performs well on platforms with low processing power and is appropriate for custom-trained classification jobs. In complicated settings, YOLOv5, which is renowned for its high accuracy and real-time performance, is excellent at identifying multiple objects, including people. By combining such models, the technique maximizes the utilization of resources and enhances the detection framework's overall dependability while guaranteeing accurate human recognition and reducing false positives.

Data Collection & Processing

The first step in developing a machine learning-based person recognition system is to manually or partially automatically classify webcam photos as "Human" or "Not Human." In order to ensure that the model can discriminate between pertinent and irrelevant inputs, this stage is essential for supervised learning. Following labeling, the dataset is preprocessed to improve model generalization and boost dataset diversity. This includes leveling pixel values, scaling images to a consistent dimension, and performing data augmentation techniques including flipping, rotation, and contrast modifications. After the data is ready, either a deep learning model that has been trained is adjusted for the particular purpose of person detection, or a binary classifier is trained. Gathering information through the webcam, extensive data cleaning and labelling, enrichment to increase variability, and model training and validation to assess performance are some of the crucial steps in the process. The model that has been trained is exported and used for immediate deduction if the results are satisfactory, allowing precise identification of humans in applications in the real world.

System Implementation On Arduino

An essential component of the human recognition system, the Arduino is mainly in charge of movement sensing and host PC connectivity. It is configured to constantly monitor the PIR sensor's output and, when movement is detected, transmit a signal to the PC via serial communication. As a result, the PC can start other processes, including taking a picture and performing machine learning-based segmentation. Users can also be given real-time feedback about observed movements by configuring the Arduino to turn on external devices like an LED or buzzer as instant alarm methods. By only turning on the image processing pipeline when motion is detected, the Arduino functions as a low-power and effective trigger mechanism, reducing needless computing strain on the PC and maximizing energy economy while enhancing the general efficacy of the human identification system.

Real-Time Decision Making

Python serves as the main controller of the human recognition system, coordinating sensor input from the Arduino and machine learning model output to allow for intelligent decision-making in real time. Python constantly tracks the serial data from the Arduino and, when it detects a "MOTION" signal, turns on the camera to take a picture. A trained machine learning model is then used to draw conclusions from this image. The system responds appropriately based on the model's confidence score, which may include sending an alert email for remote monitoring, recording the occurrence with a timestamp for future use, saving the image to disk for records, or setting off a buzzer for an instant alarm. The system employs multi-threading to guarantee responsive and fluid efficiency, enabling simultaneous model inference and video capturing without interfering with other system operations. Furthermore, the system's reliability and precision are increased by using confidence thresholding to weed out speculative predictions and guarantee that only trustworthy detections lead to actions.

Minimizing False Positives & Enhancing Accuracy

Several complementary tactics can be used to reduce false detections and increase the dependability of a human-recognition system. To start, a second-layer machine-learning algorithm that confirms PIR triggers before any action is taken helps reduce false positives brought on by non-human motion, such pets, swaying branches, or shifting shadows. Reliability is further increased by sensor fusion, which enables the entire system to validate competing occurrences and eliminate spurious signals by combining the PIR with thermal cameras, ultrasonic modules, or ordinary RGB cameras. The ML model must be continuously improved; retraining on difficult-to-classify samples and modifying confidence criteria progressively increase robustness. Scheduled inhibition at times known to provoke nuisances, such as pets being active at morning or tree shadows in the afternoon, can significantly reduce false alerts. Environmental filtering also helps. Hardware-wise, the PIR sensor's sensitivity and delay time can be adjusted to customize detection for the particular installation scenario. The best balance between responsiveness and precision is achieved by implementing temporal smoothing techniques, such as a voting mechanism across multiple frames or a rolling average of consecutive predictions, which guarantee that only reliable, high-confidence observations trigger downstream behaviors.

5. CONCLUSION

This study effectively illustrates a hybrid method of detecting human presence by combining clever, software-based machine learning validation with hardware-based motion sensing. A PIR sensor offers an inexpensive and environmentally friendly way to detect motion initially, and a PC-based video system with a trained machine learning model adds a second layer of validation to precisely establish the of humans. the use of flexible elements like Arduino-based control, real-time data gathering, machine learning categorization and decision-making logic, the system minimizes false positives while guaranteeing high accuracy. Additionally, the structure allows for flexibility and may be expanded to include safety devices, conservation of energy automated processes, and smart surveillance. All things considered, this clever, real-time recognition system provides a workable and affordable answer to contemporary presence detection requirements by bridging the gap between basic sensor-based techniques and sophisticated computer vision technologies.

REFERENCES

1. Rahulamathavan, Y., Phan, R.C.W., Chambers, J.A. and Parish, D.J., 2012. Facial expression recognition in the encrypted domain based on local fisher discriminant analysis. *IEEE transactions on affective computing*, *4*(1), pp.83-92.
2. Kahou, S.E., Bouthillier, X., Lamblin, P., Gulcehre, C., Michalski, V., Konda, K., Jean, S., Froumenty, P., Dauphin, Y., Boulanger-Lewandowski, N. and Chandias Ferrari, R., 2016. Emonets: Multimodal deep learning approaches for emotion recognition in video. *Journal on Multimodal User Interfaces*, *10*, pp.99-111.
3. Wu, Z., Huang, Y., Wang, L., Wang, X. and Tan, T., 2016. A comprehensive study on cross-view gait based human identification with deep cnns. *IEEE transactions on pattern analysis and machine intelligence*, *39*(2), pp.209-226.
4. Nagrani, A., Chung, J.S. and Zisserman, A., 2017. Voxceleb: a large-scale speaker identification dataset. *arXiv preprint arXiv:1706.08612*.
5. Zhang, Y., Song, D., & Yao, Y. (2019). A multimodal biometric system using deep learning for image fusion. Journal of Ambient Intelligence and Humanized Computing, 10(11), 4439–4450.
6. Rahulamathavan, Y., Phan, R.C.W., Chambers, J.A. and Parish, D.J., 2012. Facial expression recognition in the encrypted domain based on local fisher discriminant analysis. *IEEE transactions on affective computing*, *4*(1), pp.83-92.
7. Kahou, S.E., Bouthillier, X., Lamblin, P., Gulcehre, C., Michalski, V., Konda, K., Jean, S., Froumenty, P., Dauphin, Y., Boulanger-Lewandowski, N. and Chandias Ferrari, R., 2016. Emonets: Multimodal deep learning approaches for emotion recognition in video. *Journal on Multimodal User Interfaces*, *10*, pp.99-111.
8. Mehmood, R.M. and Lee, H.J., 2016. A novel feature extraction method based on late positive potential for emotion recognition in human brain signal patterns. *Computers & Electrical Engineering*, *53*, pp.444-457.
9. Ali, M., Mosa, A.H., Al Machot, F. and Kyamakya, K., 2016, July. EEG-based emotion recognition approach for e-healthcare applications. In *2016 eighth international conference on ubiquitous and future networks (ICUFN)* (pp. 946-950). IEEE.
10. Hochreiter, S., 1998. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *6*(02), pp.107-116.
11. Chen, M., Zhang, Y., Li, Y., Hassan, M.M. and Alamri, A., 2015. AIWAC: Affective interaction through wearable computing and cloud technology. *IEEE Wireless Communications*, *22*(1), pp.20-27.
12. Shan, C., Gong, S. and McOwan, P.W., 2009. Facial expression recognition based on local binary patterns: A comprehensive study. *Image and vision Computing*, *27*(6), pp.803-816.
13. Siddiqui, M.F.H., Dhakal, P., Yang, X. and Javaid, A.Y., 2022. A survey on databases for multimodal emotion recognition and an introduction to the VIRI (visible and InfraRed image) database. *Multimodal Technologies and Interaction*, *6*(6), p.47.
14. Rahman, M.A. and Muniyandi, R.C., 2020. An enhancement in cancer classification accuracy using a two-step feature selection method based on artificial neural networks with 15 neurons. *Symmetry*, *12*(2), p.271.
15. Redmon, J. and Farhadi, A., 2018. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.