ISSN NO: 9726-001X

Volume 13 Issue 02 2025

Machine Learning Approaches for Cloud Monitoring and Dynamic Managements

[1]TEKU NAGADEVI, [2]PASUPULETI BHANU KIRAN, [3]PENMETSA JAGADEESH PHANI VARMA, [4]VELPURI SYAM SAI DURGA, [5]Dr. B.V.S. VARMA

[1234]Student Department of CSE, DNR College of Engineering & Technology, Balusumudi, Bhimavaram, India.

[5] Professor, Department of CSE, DNR College of Engineering & Technology, Balusumudi, Bhimavaram, India.

Abstract—

It is certain that the COVID-19 pandemic has served as a warning to all of us about the potential global impact of an infectious disease epidemic. Public spaces need to have suitable crowd detection and monitoring technologies installed if we want to stop epidemics and improve healthcare. We can significantly reduce the number of new infections by efficiently executing social distancing strategies. An concept for a social distancing system that detects and monitors crowds in real time was born out of this. If schools want to keep tabs on their pupils better, they may use the system proposed in this paper—a completely autonomous system for real-time crowd detection and monitoring. With the help of a Support Vector Machine (SVM) detector and an OpenCV-based Histogram of Oriented Gradients (HOG), this system can identify and count the number of persons present at every given instance. If the number of persons in the cluster exceeds the permissible level, an alarm will be raised to remind everyone to follow the regulations. Machine learning, object detection, crowd detection, support vector machine, histogram of oriented gradients, and other related terms

INTRODUCTION

Many locations, such as stations, malls, places of worship, airports, public events, and so on, see large crowds of people. In these types of venues, video monitoring is crucial since it gives the necessary data for crowd control. Public safety, traffic monitoring, studying crowd behavior, preventing riots, creating public spaces, keeping people at a safe distance during pandemics, and many more uses are possible with the ability to detect and monitor crowds. Because of this, scientists have been working on models to help with things like counting, estimating densities, monitoring movements, and detecting behaviors. Object recognition and human classification in video streams is the first step in developing this person detection system. Machine learning is a branch of AI that teaches computers to act like humans by analyzing data and using algorithms. One subfield of computer vision, object detection seeks for and labels various visual elements. Object detection is the process of identifying and classifying things based on their characteristics, such as whether they are people, animals, trees, or vehicles. What is The MCA program at the CMR Institute of Technology in Bengaluru, India is headed by Vakula Rani J. object location data to computer vision algorithms (vakula.r@cmrit.ac.in). Performing this operation manually was the norm in the past, but it was laborious, expensive, and error-prone. Automated detection and tracking algorithms developed for computers have reduced the need for human oversight, allowing for better, more cost-effective real-time performance. You may find a lot of human detection algorithms in books and online.

The first form of object detection methodology is based on classical machine learning, which does not use neural networks. The second type, which uses neural networks and deep learning, is shown in Figure 1. To train models to differentiate between people and other objects and finish the pedestrian identification job, non-neural techniques first utilize artificially built feature extractors in traditional human detection approaches to extract the main human traits. Some examples of non-neural based techniques are: Scale-invariant feature transform (SIFT) [1], ViolaJones object identification framework employing Haar features [2], Histogram of oriented gradients (HOG) features [3], etc. Object detection is carried out using neural networks or deep learning techniques by identifying patterns in images via the use of several levels, such as input, hidden, and output layers. Algorithms based on region proposals, such as R-CNN, SPP-NET, and Fast R-CNN, and regression-based algorithms, such as

ISSN NO: 9726-001X

Volume 13 Issue 02 2025

JOURNAL OF CURRENT SCIENCE

YOLO and SSD, are the two main types of deep learning-based target identification methods. When compared to traditional object detection methods, deep learning approaches are not as efficient or need more code to solve a problem. The standard algorithms are quite broad and apply in the same manner to all images. But deep neural network features are dataset specific, so they won't operate well with photos that aren't in the training set if they're not well-designed. Training massive datasets can do these tasks, but it would be impractical and take too much time for a closed application. You may test your solution's efficacy outside of a training setting thanks to the total transparency of traditional object identification methodologies.

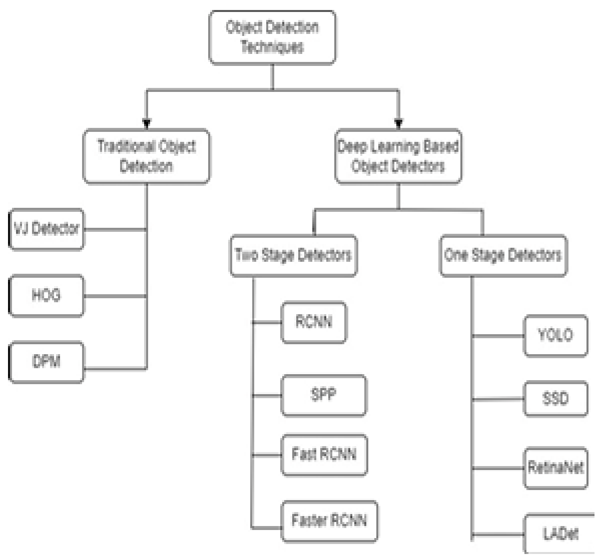

Fig. 1. Classification of Object Detection

Methods But OpenCV has a feature that can detect whether someone is a person. To identify people in photos and videos, it comes with a pre-trained HOG (Histogram of Oriented Gradients) and a Linear Support Vector Machine (SVM) model. The basic human anatomy—two arms, two legs, a brain, etc.—is used to build AI models. The trained model may then be used to the task of people detection in picture and video streams. This article takes a look at how well HOG works when combined with an SVM classifier to identify the current population size. There are five main parts to this paper: The experimental findings are shown in Section IV, the approach utilized to create a real-time CDMS is detailed in Section III, and the previous work on crowd detection is discussed in Section II. Finally, Section V covers the topic of future scope and closing comments. part two. Companion Tasks In order to accomplish pedestrian detection, several researchers have used various feature extraction and classification algorithms. The research on pedestrian detection in [4] is used as a basis for this article. Based on studies, the HOG feature is the most discriminative standalone feature utilized in the literature [3]. A large part of its benefit comes from its ability to reliably capture local edge/gradient information and its inherent resistance to changes in illumination. The study found that HOG pedestrian detection produced the best results with the fewest false positives [5]. Nearly all current detectors use HOG or a version of it in conjunction with SVM to get satisfactory performance results. One common classification approach is Support Vector Machines (SVM) [6]. This includes pedestrian classification. Estimating crowd size and counting individuals is a common goal of the many studies that have published their results in the literature. Using HOG and SVM, the method suggested by [7] creates a crowd detection system that sounds an alert in the event of aggressive crowd behavior. Using Light weight Convolutional Neural Networks (LW-CNN), B. Vivekanandam's architecture [8] enables the application of crowd computing in any public environment, leading to more precise counting. Since the COVID-19 pandemic began, researchers have used a variety of methods to study social distance. Using Deep Learning for social distance categorization, [9] and [10] developed crowd control systems to halt the spread of sickness. By monitoring people's distance management strategies, their system can determine when social distancing rules are being violated and trigger an alarm accordingly. One constant structure that Yang et al. [11] suggested for tracking social distance is an AI monocular camera. In order to manage access to the area of interest and avoid congestion, the proposed technique makes use of a crucial social density. The literature study concludes that HOG is the best supervised classification standalone descriptor and that SVM provides the best performance with the least amount of work to implement.

Methods

In Figure 2, we can see the suggested structure of a real-time CDMS that would help with social distance on campus. Several IP cameras strategically placed throughout campus provide the video and still photographs. The server will save these video recordings for future use in analysis. The video scene's pedestrians are identified using machine learning algorithms. Background Subtraction, Feature

ISSN NO: 9726-001X

Volume 13 Issue 02 2025

Extraction using HOG and SVM to classify the detected pedestrian as human or not, and finally, counting the number of detected humans to send to the social distancing monitoring system that triggers an alarm in case social distancing norms are violated are the main tasks that the system is expected to accomplish.

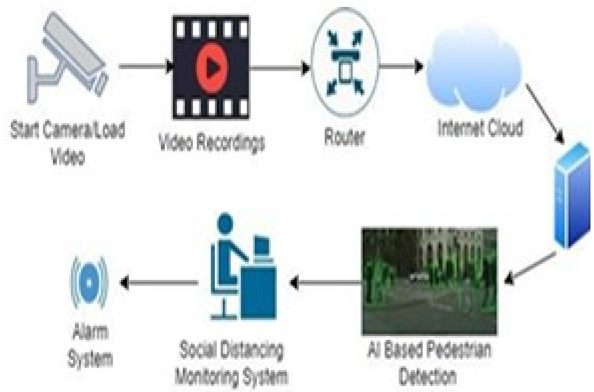

Fig. 2. Proposed Framework

Section A: Removing the Background The background removal approach is usually necessary to detect motion in real-time applications since HOG requires a lot of processing speed [12] [13] [14]. When motion is detected, frames are processed selectively instead of each frame being analyzed individually. By comparing the pixel locations in two photos, the background subtraction technique can ascertain the true intensity disparity in relation to displacement. It is assumed that the pixels in the foreground are moving while the pixels in the backdrop remain static if the background is eliminated from the picture. The HOG with SVM Classifier is set to identify humans when the background subtraction approach finds movement. B. HOG, or Histogram of Oriented Gradients Analysis A popular computer vision approach, Histogram of Oriented Gradients (HOG) [3][15][16] has a high success rate when it comes to object detection and feature extraction. In order to identify and characterize objects, HOG makes use of edge directions or intensity gradients as feature descriptors. In order to create a gradient histogram for the pixels in each cell, it first partitions the picture into cells. Equations (1) and (2) demonstrate how to use the Sobel operation to assess these gradients:

$$Sx\,(y,x) = Y\,(y,x+1) - Y\,(y,x-1) \tag{1}$$

$$Sy\,(y,x) = Y\,(y+1,x) - Y\,(y-1,x) \tag{2}$$

in which the pixel intensity with the coordinates value (x, y) is represented by Y(y, x). Sy (y, x) represents the vertical gradient while Sx (y, x) represents the horizontal gradient. You may get the gradient's magnitude (S) and direction () using Equations 3 and 4.

$$S = \sqrt{S_x^2 + S_y^2} \tag{3}$$

$$\theta = \arctan\left(\frac{S_y}{S_x}\right) \tag{4}$$

Next, the cell is sorted based on its gradient direction. The gradients of the two axes, horizontal and vertical, may be determined using filtering. [1, 0, 1] is used for horizontal filters while [1, 0, 1] is used for vertical filters. A description is provided by the cumulative histogram of all the cells. To improve accuracy, normalization is applied to all areas in the detection window by providing a measure of the local histogram across larger windows of defined geographical regions. One other name for this last vector is a feature vector, and it may be used for object detection. Furthermore, these feature vectors are used to create the support vector machine classifier, which helps to determine whether the resultant item is human or not. When it comes to classifying different types of data, SVM is a great machine learning technique. The goal of the support vector machine classifier [6][17][18] method is to increase the maximum possible difference between two classes. When training on sparse datasets, support vector machines (SVMs) excel in high-dimensional domains. Classification and regression issues are tackled with the help of SVM. Support vector machines (SVMs) excel in both linear and non-linear contexts, making them ideal for a wide range of practical problems. In cases when hyperplane and straight line segmentation of data points is straightforward, linear support vector machines (Fig. 3) are used. When you can't use a straight line to separate the data, you should employ Non-Linear SVM (Fig. 4). Kernel routines are available to handle this. They make a space linear after it was nonlinear.

ISSN NO: 9726-001X

Volume 13 Issue 02 2025

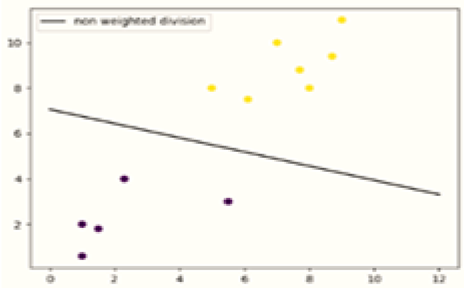

Fig. 3. : Linear SVM

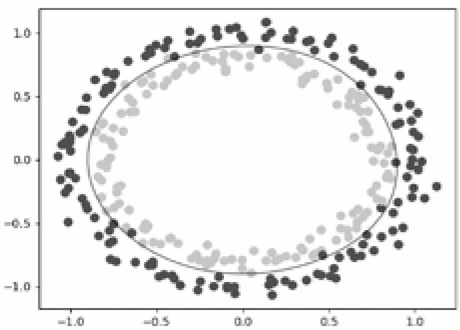

Figure 4: Support Vector Machine with Non-Linear Regressions Section D. CDMS, or Crowd Detection and Monitoring System In order to identify people clusters, a Real-Time CDMS monitors public areas. Figure 5 depicts the whole system process. The first step is to collect video sequences from the many security cameras that have been strategically placed. After that, we take still photos from films and use a technique called "Background subtraction" to crop out the unwanted parts of the backdrop. This way, we can focus on processing the foreground items. Then, to make sure social distance standards are met, the HOG and SVM algorithms determine the number of people in a cluster and use that information to identify them.

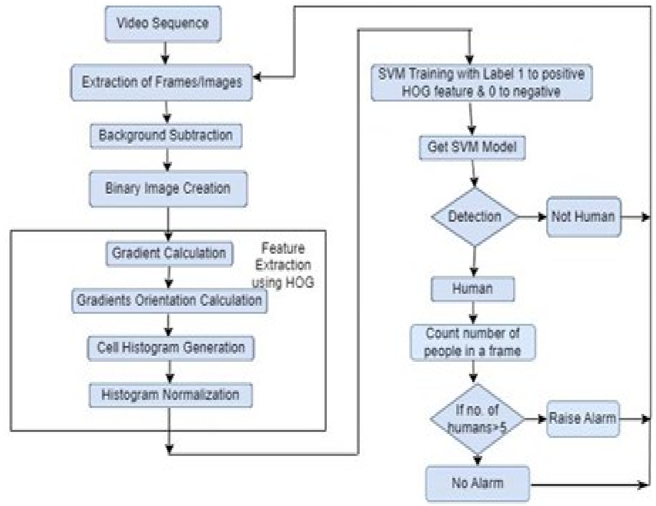

Fig. 5: Work Flow of Human Crowd Detection System The alarm system is activated when the human count is greater than the threshold value in the cluster. Here we have taken threshold value as 5 as mentioned in COVID-19 guidelines [19] [20]. Algorithm to Detect Humans

Algorithm:

Step 1: cv2.HOGDescriptor() method is used and the descriptor object is set as HOGCV to extract the relevant information from the image.

Step 2: The SVM detector is initialized using the HOG Descriptor function

setSVMDetector(cv2.HOGDescriptor getDefaultPeopleDetector()).

Step 3: The function cv2.imread() is used to read images from a specified path.

Step 4: The function detectMultiScale(), is used to identify items in a picture and get their x, y, height, and width.

Step 5: We draw a bounding box around any recognized humans using the cv2.rectangle method.

Step 6: The count argument keeps track of the number of humans found.

ISSN NO: 9726-001X

Volume 13 Issue 02 2025

JOURNAL OF CURRENT SCIENCE

EXPERMIMENTS AND RESULTS

Training and testing the model in this experiment is done using the MIT pedestrian dataset [21]. Pedestrians in metropolitan environments are the subjects of 509 training photographs and 200 test shots, together with their left-to-right reflections. The only options are front or back views, and there aren't many positions to choose from. Improved accuracy is achieved by fine-tuning hyperparameters. At the CMRIT college campus, the system was constructed and experiments were carried out. Under typical weather conditions, the algorithm accurately predicted the human population. Figures 6 and 7 provide the screenshots of the output. Both the low-light and normal-light person detection scenarios are shown in Figures 6 and 7, respectively. Ensuring a healthy environment to avoid diseases, the system looked for the number of pupils in a cluster and generated an alert if the count was discovered to be greater than 5.

Fig. 6: Detection in low light

Detection in natural illumination (Fig. 7) Table 1 displays the results obtained in various situations. The method successfully predicted the number of students in well-lit environments. The projected number of pupils, however, was inaccurate when lit up. In wet circumstances, the system's student count predictions were spot on, but in foggy conditions, they were off. Consequently, in cases of poor illumination and unfavourable weather, discrepancies were noted between the anticipated and actual results.

TABLE 1: OUTPUT PREDICTION UNDER DIFFERENT SCENARIOS

S.No	Output under different scenarios	Condition	Expected Output	Predicted Output
1	Image with 3 persons	Normal	3 persons	3 persons
2	Image with 3 persons	High Brightness	3 persons	3 persons
3	Image with 4 persons	Low Brightness	4 persons	3 persons
4	Image with 4 persons	Raining	4 persons	4 persons
5	Image with 4 persons	Fog	4 persons	3 persons

A typical technique to indicate how well the models operate is by showing the average test results as measures below (Table 2). Accuracy is defined as the proportion of correctly predicted outcomes from the test data. To what extent does a machine learning model fail to detect false positives when it displays false positives? This is known as the model's sensitivity. The specificity of a model is defined as the proportion of false negatives it properly identifies.

ISSN NO: 9726-001X

Volume 13 Issue 02 2025

The ratio of correct forecasts to total positive predictions is called precision. The F1 score is calculated by summing the accuracy and recall measures.

TABLE 2: PERFORMANCE METRICS

Measure	Values
Sensitivity	0.666
Specificity	0.918
Precision	0.40
Accuracy	0.90
F1 Score	0.498

Table 2 shows that the HOG + SVM combination achieves good accuracy and yields promising outcomes.

CONCLUSION AND FUTURE WORK

We developed a crowd detection system for a school using Python and OpenCV. Locating and tallying individuals is the function of this paradigm. For this job, we used the built-in SVM classifier in OpenCV and the HOG Descriptor Algorithm to identify individuals in pictures. Despite extensive testing in adverse conditions (such as poor light, fog, rain, etc.), we were able to achieve a 90% success rate. There are still major challenges that need fixing before a completely automated, real-time operational surveillance system can be implemented, even if advances in computer vision and related areas have been noticeable recently. Some of these issues are more specific, such where to put the cameras, how much bandwidth is needed for the network, and how to ensure comprehensive coverage; others are more general, like how long the cameras will last in typical weather and lighting conditions, how much it will cost to install, privacy concerns, and so on. This research presents an accurate and desirable outcome for crowd detection. Using this method in conjunction with CCTV to screen people during pandemics is possible. Train stations, bus stops, marketplaces, roadways, entrances to retail centers, schools, and colleges are all examples of densely populated areas where this method may be used for effective mass screening. In order to broaden the scope of the endeavor and boost its efficacy, the system presents several intriguing possibilities. From real-time CCTV video, it may be used for a plethora of different reasons, such coordinating public events, counting people, detecting falls, etc. Improving the model's overall performance may be the goal of future research into more advanced Deep Learning algorithms that would allow for more precise results even when weather conditions are less than ideal.

REFERENCES

[1]. Tony Linderberg. “Scale Invariant Feature Transform”, Scholarpedia, (2012)

[2]. Viola P. and Jones M. “Rapid Object Detection using a Boosted Cascade of Simple Features “, CVPR (2001)

[3]. Dalal N., Triggs B., and Schmid C. “Human Detection Using Oriented Histograms of Flow and Appearance”, Proceedings of European Conference on Computer Vision (2006)

[4]. Yadav R. P, Senthamilarasu V, Kutty K, Vaidya V, Ugale S. P “A Review on Day-Time Pedestrian Detection” World Congress, SAE (2015)

[5]. Solichin, Achmad & Harjoko, Agus & Putra, Agfianto. “A Survey of Pedestrian Detection in Video”, IJACSA (2014)

[6]. Junil Tao, Klette R, “Vision Based Pedestrian Detection Improvement and Verification of feature Extraction Methods and SVM-Based Classification”, ITSC (2011)

[7]. Singh, D. K., Paroothi, S., Rusia, M. K., & Ansari, M. A. “Human Crowd Detection for City Wide Surveillance”, Procedia Computer Science, (2020)

[8]. Vivekanandam, B. "Speedy Image Crowd Counting by Light Weight Convolutional Neural Network." Journal of Innovative Image Processing (2021)

[9]. Sangeeta Yadav, Preeti Gulia, Nasib Singh Gill, Jyotir Moy Chatterjee, "A Real-Time Crowd Monitoring and Management System for Social Distance Classification and Healthcare Using Deep Learning", Journal of Healthcare Engineering, (2022)

[10]. Ansari MA, Singh DK. “Monitoring social distancing through human detection for preventing/reducing COVID spread”. Int J Inf Technol (2021)

[11]. Yang, E. Yurtsever, V. Renganathan, K.A. Redmill, U. Ozg¨uner, “A vision-based social distancing and critical density detection system for covid-19”, arXiv e-prints pp. (2020)

[12]. O. Javed, K. Shafique, and M. Shah.” A Hierarchical Approach to Robust Background Subtraction Using Color and Gradient Information.” Proc. IEEE Workshop on Motion and Video Computing, IEEE CS Press, (2002)

[13]. Pushkar Protik Goswami, Dushyant Kumar Singh.” A hybrid approach for real-time object detection and tracking to cover background turbulence problem.” Indian Journal of Science and Technology (2016)

[14]. Ruolin Zhang, Jian Ding, “Object Tracking and Detecting Based on Adaptive Background Subtraction”, International Workshop on Information and Electronics Engineering (2012)

[15]. O. Kliper-Gross, T. Hassner, and Liu. Beyond.” Pixels: Exploring New Representations and Applications for Motion Analysis.” PhD thesis, Massachusetts Institute of Technology (2009)

[16]. Surasak, Thattapon & Takahiro, Ito & Cheng, Cheng-hsuan & Wang, Chi-en & Sheng, Pao-you.” Histogram of Oriented Gradients for Human Detection in Video” ICBIR, (2018) [17] Tzotsos, D. Argialas “Support Vector Machine Classification for Object-Based Image Analysis”, ISBN: 978-3-540-77057-2, (2008)

[17]. Ilao, Joel & Cordel, Macario II. “Crowd Estimation using Region- Specific HOG with SVM”, JCSSE, (2018)

[18]. https://economictimes.indiatimes.com/news/newsblogs/latest-daily news-and-live-updates-december-24/liveblog/88464087.cms(Dec, 2021)

[19]. https://www.outlookindia.com/website/story/india-news-coronavirus gathering-of-more-than-5-people-banned-will-lock-down-delhi-if needed-says-kejriwal/349213, (Mar, 2020) [21] Papageorgiou and Poggio, “A trainable system for object detection”, IJCV (2000)