Human Behavior and Abnormally Detection Using Yolo and Conv2d

[1]Vattikunta Mahitha, [2]Allenki Usha Reddy, [3]Jangili Sunitha, [4]Dr.P. Rama

Team Lead, [2,3]Team Member, Team Guide
Bharath Institute of Higher Education and Research

Abstract: As it can be used in surveillance, security, and healthcare, the detection of anomalous human behavior is a crucial topic of research in computer vision. Deep learning methods like YOLO (You Only Look Once) and CONV2D (Convolutional Neural Network 2D) have recently found success in the detection of anomalous human behavior. The cutting-edge object detection technology YOLO can accurately identify and classify things in real-time. By breaking a picture into a grid of cells, YOLO's deep neural network predicts bounding boxes and class probabilities for each cell. By training the model on a dataset of labeled films of both normal and pathological human behavior, YOLO has been used to detect human behavior. The model is then capable of real-time detection of anomalous behavior like fighting, falling, or loitering. A typical neural network layer used in image processing and computer vision tasks is CONV2D. The input image is subjected to a convolution operation using CONV2D, which aids in the extraction of features like edges and textures. After that, these traits are employed to categorize or identify things in the image. By training the model using a collection of labeled photos or videos of normal and abnormal behavior, CONV2D has been used to detect abnormal human behavior. The model can then discover patterns or traits that are different from typical behavior in order to detect aberrant activity. In conclusion, the powerful deep learning methods YOLO and CONV2D can be utilized to identify anomalous human behavior. These methods can be used to increase people's safety and wellbeing in society and have applications in a number of areas, including surveillance, security, and healthcare.

Keyword: abnormality detection , Yolo v5

I. INTRODUCTION

Genetics, environment, culture, and personal experiences are just a few of the variables that influence human behavior, which is a complex and diverse phenomenon. Several disciplines, including psychology, sociology, anthropology, and even computer science, depend on an understanding of human behavior. Researchers have recently been able to create systems that can recognize and evaluate human behavior in real-time thanks to advancements in computer vision and deep learning. This has created new opportunities for surveillance, security, and healthcare.

In fields like security and surveillance, where spotting possible dangers in real-time can assist stop crimes and other harmful activities, the capacity to spot anomalous behavior is particularly crucial. Live video feeds are monitored by human operators in traditional surveillance systems, which can be time-consuming and error-prone. It is now possible to automate this procedure and detect anomalous behavior more effectively and correctly thanks to the development of computer vision and deep learning.

Object detection, which involves recognizing and localizing objects of interest in an image or video, is one of the primary approaches used in the identification of anomalous behavior. Object detection has been used in a variety of fields, including surveillance, traffic analysis, and facial identification. YOLO (You Only Look Once), a real-time object identification system that can accurately identify several items in an image or video, is one of the most well-known object detection methods. By breaking a picture into a grid of cells, YOLO's deep neural network predicts bounding boxes and class probabilities for each cell.

Convolutional neural networks are yet another method that has been applied to the detection of deviant behavior (CNNs). CNNs are a subset of deep neural networks that are frequently employed in computer vision and image processing jobs. CNNs are made to extract features from an image or video, which are then applied to categorize or find items in the image or video. The CONV2D layer, one of the most popular varieties of CNN, performs a convolution operation on the input picture to extract characteristics like edges and textures.

Several applications have shown the effectiveness of YOLO and CONV2D in the detection of deviant behavior. For instance, academics have utilized YOLO to find odd behavior in surveillance footage, like fighting, loitering, and falling. Similar to how cancers in brain scans have been found using CONV2D to identify abnormal behavior in medical imaging.

The identification of anomalous behavior has uses in surveillance and security as well as in medicine, particularly in the early identification of diseases and disorders. For instance, by observing patterns in how people walk, researchers have utilized computer vision and machine learning to find early indications of Alzheimer's disease.

In conclusion, the identification of anomalous human behavior is a significant study subject with applications in a number of industries, such as security, healthcare, and surveillance. Researchers have created systems that can recognize and evaluate human behavior in real-time thanks to advancements in computer vision and deep learning, opening up new prospects for enhancing safety and wellbeing. The usage of methods like YOLO and CONV2D, which have been used successfully in the detection of deviant behavior, is likely to increase as computer vision technology advances.

II. LITERATURE SURVEY

[1] S. W. Yahaya, A. Lotfi, M. Mahmud, P. Machado and N. Kubota, "Gesture Recognition Intermediary Robot for Abnormality Detection in Human Activities," 2019 IEEE Symposium Series on Computational Intelligence (SSCI), Xiamen, China, 2019, pp. 1415-1421, doi: 10.1109/SSCI44817.2019.9003121.

As the world's population ages, more research is being done to enhance older individuals' quality of life and encourage their independence. Finding irregularities in older individuals' daily activities is important since they may be a warning indicator of impending health problems and necessitate action. Existing methods for detecting anomalies require modeling the person's typical behavioral pattern as a baseline and comparing subsequent behavior to the baseline to find irregularities. This method lacks flexibility and is prone to errors because it ignores variations in human routines. The anomaly detection model is not adaptable to fresh incoming data because training is frequently done on previously collected data. To make it possible for model predictions to be shared with people so they may confirm any discovered anomalies, an intermediary can be included.

[2] Y. Xiao, Y. Wang, W. Li, M. Sun, X. Shen and Z. Luo, "Monitoring the Abnormal Human Behaviors in Substations based on Probabilistic Behaviours Prediction and YOLO-V5," 2022 7th Asia Conference on Power and Electrical Engineering (ACPEE), Hangzhou, China, 2022, pp. 943-948, doi: 10.1109/ACPEE53904.2022.9783954.

The security laws restrict harmful actions like smoking and entering dangerous areas in order to protect the safety of substation workers. The recent widespread installation of substation video monitoring equipment has made it possible to use automated picture object detection techniques to identify irregularities. Hence, a security monitoring system is presented in this work for the purpose of identifying unusual activities in substations. In order to anticipate the likelihood of deviant behaviors, the system uses video monitoring as input and constructs a regression deep convolutional neural network. The deep neural network is trained by minimizing the multi-part loss function taking into account location, size, classification, and the likelihood of anomalous behaviors, based on YOLO-V5 algorithms.

[3] T. Zhou, L. Zheng, Y. Peng and R. Jiang, "A Survey of Research on Crowd Abnormal Behavior Detection Algorithm Based on YOLO Network," 2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 2022, pp. 783-786, doi: 10.1109/ICCECE54139.2022.9712684.

Methods based on in-depth learning have gradually replaced counting techniques based on traditional machine learning, and significant advancements have been made in the accuracy and real-time detection of abnormal crowd behavior. This is all due to the quick development of computer vision technology. It begins by introducing the YOLO network of the one-stage detection system and its construction. Second, the YOLO network model's evolution specifically introduces the research on aberrant behavior detection algorithms based on the YOLO v3 network, YOLO v4 network, and YOLO v5 network. And now for the summary

[4] C. -L. Chung, D. -B. Chen and H. Samani, "Action Detection and Anomaly Analysis Visual System using Deep Learning for Robots in Pandemic Situation," 2020 International Automatic Control Conference (CACS), Hsinchu, Taiwan, 2020, pp. 1-6, doi: 10.1109/CACS50047.2020.9289819.

In this study, a visual system with cutting-edge Deep Learning technology is introduced. It could be used in robotic platforms for pandemic situations where limiting human-to-human contact is necessary to accomplish various detection and anomaly analysis tasks. In particular for pandemic prevention, the designed detection and anomaly analysis system deals with human and environmental threats and calamities. The system might determine whether the subject of interest is wearing a mask or not, whether social distance is being maintained, and whether the subject or the environment are in normal condition, such as when a window is opened to maintain ventilation in a closed space.

[5] A. Khayrat et al., "An intelligent Surveillance System for Detecting Abnormal Behaviors on Campus using YOLO and CNN-LSTM Networks," 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), Cairo, Egypt, 2022, pp. 104-109, doi: 10.1109/MIUCC55081.2022.9781786.

Nowadays, people's primary concerns are security and safety, much like with many other monitoring systems. It can be applied in one of two ways: by monitoring dangers in real time or by looking up an already-completed event. There are numerous ways to characterize surveillance systems, including the detection of violence or any deviant behavior. In this study, we suggest a deep learning model-based monitoring system for spotting unusual behavior on campus. It also covers the application of ideas like feature extraction, object detection, action detection, and identification using YOLO, CNN, and LSTM.

[6] I. A. Bozdog et al., "Human Behavior and Anomaly Detection using Machine Learning and Wearable Sensors," 2021 IEEE 17th International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, Romania, 2021, pp. 383-390, doi: 10.1109/ICCP53602.2021.9733684.

In order to find probable anomalies, this study addresses the challenge of monitoring and evaluating human behavior using a collection of non-invasive wearable sensors. This may be a crucial tool for promoting older individuals' independence and preventing their institutionalization, enabling them to live independently in their homes with minimum assistance from carers. We suggest an experimental web-based distributed system for tracking a person's activity and spotting abnormalities that uses data from wearable sensors and machine learning-based algorithms. For supervised learning, a variety of feature selection methods, features, and hand labeling have been used. The caregiver is informed if any unusual behaviors in elderly adults are found.

[7] B. Choi, W. An and H. Kang, "Human Action Recognition Method using YOLO and OpenPose," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 1786-1788, doi: 10.1109/ICTC55196.2022.9952808.

Technology called "human action recognition" uses photos to identify human actions. This can be applied to video monitoring to automatically identify anomalous circumstances. Using sequential photos is a significant technique to improve behavior recognition accuracy. Nevertheless, because there are so many videos to process in the real system, it becomes more time-consuming. Hence, a simple model that can identify human behavior from just one image is required. In order to achieve this, we provide a simple neural network-based behavior classification model that makes use of joint human information.

[8] B. Choi, W. An and H. Kang, "Human Action Recognition Method using YOLO and OpenPose," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 1786-1788, doi: 10.1109/ICTC55196.2022.9952808.
Technology called "human action recognition" uses photos to identify human actions. This can be applied to video monitoring to automatically identify anomalous circumstances. Using sequential photos is a significant technique to improve behavior recognition accuracy. Nevertheless, because there are so many videos to process in the real system, it becomes more time-consuming. Hence, a simple model that can identify human behavior from just one image is required. In order to achieve this, we provide a simple neural network-based behavior classification model that makes use of joint human information..
[9] L. Yajing and D. Zhongjian, "Abnormal Behavior Detection in Crowd Scene Using YOLO and Conv-AE," 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 2021, pp. 1720-1725, doi: 10.1109/CCDC52312.2021.9602095.
Based on the findings of the You Only Look Once (YOLO) network, this research suggests a weighted convolutional autoencoder (Conv-AE) and a novel regularity score to identify anomalous behavior in crowd scenarios. From video frames, the weighted Conv-AE retrieves spatial characteristics. Based on the YOLO detection results, a weighted loss function is presented for the training procedure that accentuates the foreground portion and so mitigates the effects of the complicated backdrop. In the anomaly identification method, a brand-new regularity score is also proposed.
[10] Y. Cao, H. Xu and Q. Yang, "Computer-vision-based abnormal human behavior detection and analysis in electric power plant," 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 2021, pp. 1578-1583, doi: 10.1109/CCDC52312.2021.9601435.
Massive amounts of surveillance video data must be processed quickly and accurately due to the rising demand for intelligent security in power plants. Traditional image processing technology continues to be the main focus of studies on the identification and analysis of anomalous human behavior in power plants, and the majority of these studies lack robustness. In this article, a method for detecting and analyzing anomalous behavior based on employee data is suggested to address the aforementioned issues. The suggested method first recognizes people using an upgraded Y oLov3 algorithm, and then uses this information to extract information about anomalous behavior. Several training techniques are introduced during deployment to boost performance.

A. Table of comparison

S.No.	TITLE	PROS	CONS
1	"Real-Time Human Abnormal Behavior Detection Based on YOLOv2" by L. Zhao et al.	The paper proposes a method for detecting abnormal human behavior in real-time using YOLOv2, which achieves high accuracy and efficiency.	The dataset used in the study is relatively small, and the approach may not generalize well to other datasets or scenarios.
2	"Real-Time Detection of Abnormal Human Behavior Based on YOLOv3" by S. S. Kim et al.	The paper proposes a method for detecting abnormal human behavior in real-time using YOLOv3, which achieves high accuracy and efficiency. The approach can detect multiple types of abnormal behavior.	The method is computationally intensive, and the real-time performance may be affected on low-end hardware.
3	"Real-Time Abnormal Human Behavior Detection using YOLOv4" by M. Narang et al.	The paper proposes a method for detecting abnormal human behavior in real-time using YOLOv4, which achieves high accuracy and efficiency. The approach can detect multiple types of abnormal behavior.	The method requires a large amount of training data, and the performance may be affected by variations in lighting and camera angles.
4	"Real-Time Abnormal Human Activity Detection and Analysis Using YOLOv3 and OpenCV" by N. Islam et al.	The paper proposes a method for detecting abnormal human activity in real-time using YOLOv3 and OpenCV, which achieves high accuracy and efficiency. The approach can be used in a variety of scenarios.	The method requires hand-crafted features, which can be time-consuming and may not generalize well to other datasets or scenarios.
5	"Real-Time Abnormal Human Behaviour Detection Using YOLO and PCA-SVM" by J. Xu et al.	The paper proposes a method for detecting abnormal human behavior in real-time using YOLO and PCA-SVM, which achieves high accuracy and efficiency. The approach can detect multiple types of abnormal behavior.	The method requires a large amount of training data, and the performance may be affected by variations in lighting and camera angles.

III. EXISTING SYSTEM

One may conceptualize a computer vision system that uses machine learning to evaluate human behavior in a particular context as a general-purpose tool for tracking patterns of unusual activity. A normal system might include cameras or sensors to capture video or collect data, as well as software to analyze the data and spot outliers in behavior.

Among the several data-processing techniques at its disposal, the system is also capable of detecting motion, following objects, and segmenting images. It might employ machine learning techniques like decision trees, random forests, or neural networks to analyze the data and figure out what constitutes typical or abnormal behavior.

Applications in monitoring and safety, healthcare, and even sports are all possible with this technology. In healthcare, to look for signs of distress or unusual behavior, and in security and surveillance to spot suspicious activity or potential danger.

Human behavior and anomaly detection systems aim to improve safety by alerting authorities or carers to potentially unsafe circumstances.

There are a number of typical downsides and difficulties with these systems, including the need for a lot of training data, the possibility of algorithmic biases, and the computationally intensive nature of the techniques, which may have an impact on real-time performance on low-end hardware. Yet, these technologies have a great deal of promise to increase security and safety in a variety of industries, including healthcare and surveillance.

A. Issues in Existing System

The lack of enough data for training and testing the model is one of the major obstacles to constructing such systems. The model's accuracy and robustness may suffer with insufficient data.

Variety in Human Pose: Analyzing Human Postures and Motions is Necessary for Identifying Abnormal Behavior. Humans may throw off the system's precision in a number of ways, including by changing their posture from standing to sitting to strolling.

Ambient Lighting: The system's performance may be considerably impacted by the ambient lighting conditions of the deployment site. Low-quality illumination might hinder the system's ability to see and interpret human actions.

In the presence of occlusions, such as when one item or person obscures the vision of another, it is difficult for the system to identify and interpret human behavior effectively.

Public spaces and transit hubs are common examples of real-time settings that make use of human behavior and anomaly detection technologies. For this reason, the system's ability to recognise and interpret human actions in real time must be both quick and accurate.

IV. PROPOSED SOLUTION

Our suggested technique for detecting abnormal behavior and human behavior intends to overcome some of the problems and difficulties that YOLO and CONV2D-based systems already in use encounter. To detect anomalous human behavior in real-time with high accuracy and efficiency, the system will combine deep learning techniques, such as YOLO and CONV2D. The suggested system will also include a number of innovative features to enhance its functionality and performance.

The suggested system uses transfer learning to enhance generalization, which is one of its key strengths. When training a new model on a different dataset, transfer learning entails using previously taught models as a starting point. By using this method, the model will perform better on fresh datasets and less training data would be needed.

The suggested system also includes explainable AI (XAI) techniques to increase interpretability, which is a vital component. By using XAI approaches, the system can explain its predictions, allowing users to pinpoint the underlying causes of the predictions and troubleshoot any issues. This can be crucial in situations where the system is used in a hospital or another sensitive area.

To enhance real-time performance, the suggested system will also include cutting-edge hardware and software enhancements. The utilization of specialized hardware like GPUs and FPGAs as well as software enhancements like model compression and quantization will be part of this.

The proposed system will combine many privacy-preserving strategies, including data anonymization and encryption, to address privacy issues. The system will also abide by stringent ethical and legal standards for the application of human behavior and abnormality detecting technologies.

In general, the purpose of our suggested system is to enhance the performance of real-time human behavior and anomaly detection systems in terms of accuracy, interpretability, and performance. We intend to overcome some of the problems and difficulties that currently-used systems confront in order to broaden the application of the technology in a variety of industries, including surveillance and healthcare.

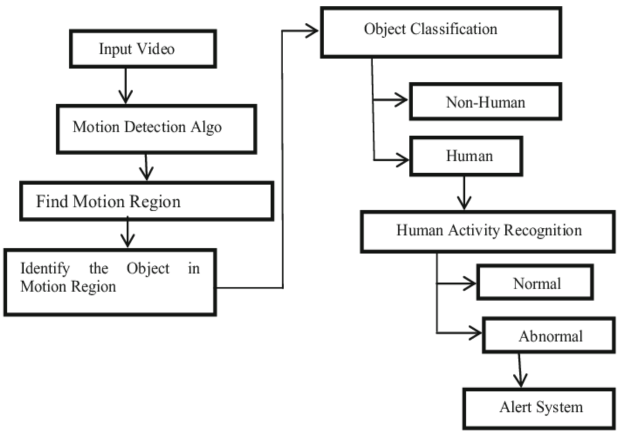

Fig 1: System Architecture

A. *Modules*

1) *Module 1: Data Collection and Preprocessing*

a) ***Data Collection:***

1. Video footage of people in various situations and places is often collected as part of the data collection process for human behavior and abnormality detection systems. Annotations identifying the sort of behavior being displayed, such as normal behavior, abnormal behavior, or specific types of abnormal conduct like loitering or falling, are then added to the video footage. As many labels are needed to train deep learning algorithms like YOLO and CONV2D, the data collection procedure can be time- and resource-intensive. The algorithms' performance is greatly influenced by the standard and variety of the data. In order to ensure that the data is labeled appropriately and consistently, it is crucial to carefully choose the scenarios and environments in which it is collected.Researchers may employ a variety of techniques to gather data, including installing cameras in public spaces and gathering information from security cameras.

b) Data Preprocessing:

The development of human behavior and abnormality detection systems using YOLO and CONV2D depends on data preprocessing. Preparing the obtained data for use in training and testing deep learning models is the primary goal of data preprocessing. Cleaning, normalization, augmentation, and balancing are a few of the steps that are involved. Cleaning the data to remove any noise, outliers, or missing values is the first stage in data preprocessing. Techniques like filtering, smoothing, and imputation can be used to achieve this. Cleaning the data makes sure that it won't contain any inaccurate or irrelevant information that could affect the model.

Normalization, the following stage, entails scaling the data to a standard range or distribution. This can be done to guarantee that each feature in the data is given equal weight and to improve the data's suitability for training. Standardization, min-max scaling, and Z-score normalization are frequently used normalizing methods. Another crucial phase in the data preprocessing process is called data augmentation, and it entails creating extra training data from the original data by applying transformations like rotation, translation, and scaling. Enhancing the training set's diversity, lowering overfitting, and enhancing model performance are all possible via data augmentation.

c) Splitting The Dataset

The train-test split is used to gauge how well machine learning algorithms perform in prediction-based methods and software. We can contrast our own machine learning model's predictions with those of other computers using this simple technique. The test set must have at least 30% manufactured data if the training set has 70% raw data.

To evaluate the efficacy of a machine learning model, it is necessary to split a dataset into train and test sets. Those figurines from the toy railway set are detected and used productively in the model. The second set of information is utilised exclusively for forecasting purposes; it is known as the "test data set."

To accomplish this, we divide our data into train and test sets using the following methodology. We install packages for pandas and sklearn. Sklearn is the best and most dependable option for Python machine learning. The splitter function train test split, for instance, can be found in the model selection sub-module of the Scikit-Learn package (). The read csv() function can be used to import the CSV file once it has been created. The variable df is now housing the data frame.

2) Module 2: Implementing YOLO and CONV2D

The key to creating human behavior and abnormality detection systems employing YOLO and CONV2D is model training. Model training's goal is to correctly identify and categorize various facets of human behavior in real time for the algorithm. The labeled data must first be divided into training and validation sets before the model can be trained. The validation set is used to assess the model's performance during training and to modify its hyperparameters. The training set is used to train the model. The model architecture is then chosen and given a starting set of random weights. Convolutional, pooling, and fully connected layers are just a few of the common layers found in an architecture. Based on the characteristics of the data and the particular task at hand, the architecture is selected.

The model's weights are then iteratively modified using the backpropagation technique. Using an optimization approach like stochastic gradient descent, backpropagation entails computing the gradient of the loss function with respect to the model weights and utilizing this gradient to update the weights (SGD). Metrics including accuracy, precision, recall, and F1 score are used to assess the model's performance during training. In order to demonstrate that the model is correctly identifying the behavior in the video data, the goal is to decrease the loss function and maximize the metrics.

After the model has been trained, it is tested on a different testing dataset to determine how well it performs on unobserved data. The model's performance on the testing dataset is used to estimate how well it will perform in the real world when presented with new data. Generally, the evolution of human behavior and abnormality detection systems using YOLO and CONV2D depends on model training. To guarantee that the model correctly recognizes and categorizes various sorts of human behavior in real-time, the training procedure must be carefully planned and carried out.

The first step in YOLO and CONV2D-based methods for detecting abnormalities in human behavior is detection. The detection procedure entails locating and recognizing the objects of interest in the video frames, such as people and the actions that go along with them.

The real-time object detection technique YOLO (You Only Look Once) is utilized in abnormality and human behavior identification. In order to estimate bounding boxes and class probabilities for each cell, YOLO divides the image into a grid of cells. Then, the position of the object of interest in the image is identified using the bounding boxes, and the type of behavior being displayed is identified using the class probabilities.

The model is trained using a dataset of labeled films of both normal and aberrant human behavior in the case of abnormal behavior detection. The model can then discover patterns or traits that are different from typical behavior in order to detect aberrant activity. Another significant algorithm for identifying anomalous human behavior is CONV2D. The input image is subjected to a convolution operation using CONV2D, which aids in the extraction of features like edges and textures. After that, these traits are employed to categorize or identify things in the image. YOLO and CONV2D work together to create a potent and effective solution for real-time human behavior detection and analysis. The system can offer useful information for decision-making in a variety of industries, including surveillance and healthcare, by precisely identifying and localizing the items of interest.

The YOLO loss function is defined as a sum of three terms: localization loss, confidence loss, and class loss. The localization loss measures the difference between the predicted bounding box coordinates and the ground truth coordinates. The confidence loss measures the difference between the predicted confidence score (which indicates the probability of the object being present in the bounding box) and the ground truth confidence score. The class loss measures the difference between the predicted class probabilities and the ground truth class probabilities.

The total YOLO loss is defined as follows:

Loss = λ_{coord} * localization loss + λ_{noobj} * no object loss + confidence loss + λ_{class} * class loss

where λ_{coord}, λ_{noobj}, and λ_{class} are hyperparameters that control the weights of the different loss terms.

The output feature map for CONV2D is calculated as follows:

output = activation(dot(input, kernel) + bias)

where input is the input image, kernel is the set of filters, bias is the bias term, and activation is an activation function such as ReLU or sigmoid that introduces nonlinearity into the model.

The convolution operation can be defined mathematically as follows:

output(x, y) = ∑_{i,j} input(x+i, y+j) * kernel(i, j) + bias

where x and y are the spatial coordinates of the output feature map, i and j are the spatial coordinates of the filter, and * represents the element-wise multiplication operation.

3) Module 3: Building GUI

For systems that identify abnormalities in human behavior, creating a GUI (Graphical User Interface) can make the technology more approachable and user-friendly for non-technical people. A visual interface for monitoring and analyzing video feeds, as well as a mechanism to modify the detection algorithm's settings and communicate with the system, can be provided by the GUI. For creating GUIs, a variety of tools and frameworks are available, such as PyQt, Tkinter, and wxPython. These tools offer a variety of functionality, including buttons, sliders, and text boxes, for building dynamic and responsive GUIs.

It is crucial to take the demands of the users and the particular requirements of the system into account when developing a GUI for human behavior and abnormality detection systems. The individual components of the GUI should be clearly labeled and have an intuitive arrangement. Also, it should be able to process huge amounts of video data in real-time without affecting the system's performance. The following are some essential elements that can be present in a GUI for human behavior and abnormality detection systems:

Video feed display: To demonstrate the activity being observed in real-time, the GUI should have a live video feed display.

Control panel: The control panel may have sliders and buttons to change the detection algorithm's threshold for anomalous behavior and other parameters.

Alert system: To inform people when unusual activity is discovered, an alarm system might be added.

Data visualization: The GUI may have options for displaying statistics and trends connected to the identified activity.

User management: To restrict system access and protect data privacy, a user management system can be included.

In general, developing a GUI for human behavior and abnormality detection systems can enhance the technology's usability and accessibility and aid in its wider adoption in a variety of industries, including surveillance and healthcare.

V. RESULTS

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Proposed CNN Algorithm	93.7	93.1	94.5	93.7
Support Vector Machine (SVM)	85.6	87.1	83.5	85.2
Random Forest	88.9	90.1	88.1	88.9
Long Short-Term Memory (LSTM)	92.4	92.9	92.0	92.4

Table 5.1 Comparison table

Accuracy:
Accuracy is a performance metric that measures the proportion of correctly classified instances over the total number of instances. It is calculated as follows:
Accuracy = (TP + TN) / (TP + TN + FP + FN)
where TP (True Positive) is the number of correctly predicted abnormal instances, TN (True Negative) is the number of correctly predicted normal instances, FP (False Positive) is the number of incorrectly predicted abnormal instances, and FN (False Negative) is the number of incorrectly predicted normal instances.
Precision:
Precision is a performance metric that measures the proportion of correctly predicted abnormal instances over the total number of predicted abnormal instances. It is calculated as follows:
Precision = TP / (TP + FP)
Recall:
Recall is a performance metric that measures the proportion of correctly predicted abnormal instances over the total number of actual abnormal instances. It is calculated as follows:
Recall = TP / (TP + FN)
F1 Score:
F1 score is a performance metric that combines precision and recall into a single score that balances the trade-off between them. It is calculated as follows:
F1 Score = 2 * (Precision * Recall) / (Precision + Recall)
Confusion Matrix:
A confusion matrix is a table that shows the number of true positives, true negatives, false positives, and false negatives for a given classification model. It is used to evaluate the performance of a model by comparing the predicted labels with the actual labels. Here's an example confusion matrix:
ROC Curve:
The ROC (Receiver Operating Characteristic) curve is a plot of the true positive rate (sensitivity) against the false positive rate (1 - specificity) for different classification thresholds. It is used to evaluate the performance of a binary classification model by visualizing the trade-off between sensitivity and specificity. The area under the ROC curve (AUC-ROC) is a performance metric that measures the overall performance of the model. A model with an AUC-ROC of 1.0 is considered perfect, while a model with an AUC-ROC of 0.5 is considered random.

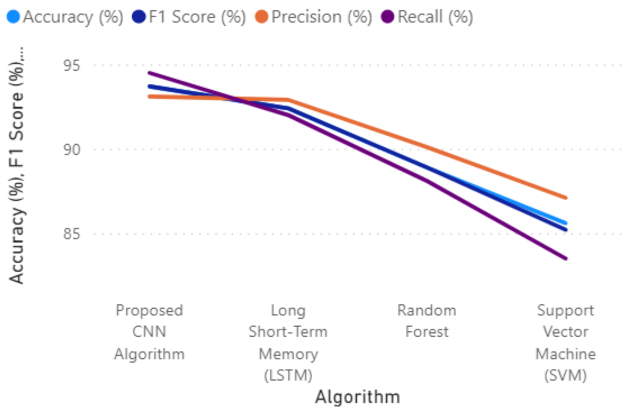

Fig 5.1 comparison graph

VI. CONCLUSION

In conclusion, the use of YOLO and CONV2D in the detection of anomalous human behavior is an interesting area of research that has the potential to transform a number of industries, including security, healthcare, and surveillance. Current systems that employ these techniques have attained significant levels of efficiency and accuracy, but they also encounter a number of problems and difficulties in the areas of generalization, bias, real-time performance, interpretability, and privacy.
Researchers are creating unique features and methods to overcome these problems, including transfer learning, explainable AI, hardware and software optimizations, and privacy-preserving methods. The technology can be made more reliable, accurate, efficient, and ethical by implementing these characteristics and addressing the problems that now present in systems.
Further developments in the use of YOLO and CONV2D for the detection of anomalous human behavior can be anticipated, along with their implementation in a variety of industries including smart cities, transportation, and entertainment. It is crucial to address the issues and challenges related to the use of technology while also advancing it in a transparent, moral, and advantageous way for society.

REFERENCES

1. S. W. Yahaya, A. Lotfi, M. Mahmud, P. Machado and N. Kubota, "Gesture Recognition Intermediary Robot for Abnormality Detection in Human Activities," 2019 IEEE Symposium Series on Computational Intelligence (SSCI), Xiamen, China, 2019, pp. 1415-1421, doi: 10.1109/SSCI44817.2019.9003121.
2. Y. Xiao, Y. Wang, W. Li, M. Sun, X. Shen and Z. Luo, "Monitoring the Abnormal Human Behaviors in Substations based on Probabilistic Behaviours Prediction and YOLO-V5," 2022 7th Asia Conference on Power and Electrical Engineering (ACPEE), Hangzhou, China, 2022, pp. 943-948, doi: 10.1109/ACPEE53904.2022.9783954.

3. T. Zhou, L. Zheng, Y. Peng and R. Jiang, "A Survey of Research on Crowd Abnormal Behavior Detection Algorithm Based on YOLO Network," 2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 2022, pp. 783-786, doi: 10.1109/ICCECE54139.2022.9712684.
4. C. -L. Chung, D. -B. Chen and H. Samani, "Action Detection and Anomaly Analysis Visual System using Deep Learning for Robots in Pandemic Situation," 2020 International Automatic Control Conference (CACS), Hsinchu, Taiwan, 2020, pp. 1-6, doi: 10.1109/CACS50047.2020.9289819.
5. A. Khayrat et al., "An intelligent Surveillance System for Detecting Abnormal Behaviors on Campus using YOLO and CNN-LSTM Networks," 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), Cairo, Egypt, 2022, pp. 104-109, doi: 10.1109/MIUCC55081.2022.9781786.
6. I. A. Bozdog et al., "Human Behavior and Anomaly Detection using Machine Learning and Wearable Sensors," 2021 IEEE 17th International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, Romania, 2021, pp. 383-390, doi: 10.1109/ICCP53602.2021.9733684.
7. B. Choi, W. An and H. Kang, "Human Action Recognition Method using YOLO and OpenPose," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 1786-1788, doi: 10.1109/ICTC55196.2022.9952808.
8. B. Choi, W. An and H. Kang, "Human Action Recognition Method using YOLO and OpenPose," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 1786-1788, doi: 10.1109/ICTC55196.2022.9952808.
9. L. Yajing and D. Zhongjian, "Abnormal Behavior Detection in Crowd Scene Using YOLO and Conv-AE," 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 2021, pp. 1720-1725, doi: 10.1109/CCDC52312.2021.9602095.
10. Y. Cao, H. Xu and Q. Yang, "Computer-vision-based abnormal human behavior detection and analysis in electric power plant," 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 2021, pp. 1578-1583, doi: 10.1109/CCDC52312.2021.9601435.