INTELLIGENT VIDEO SURVELLIANCE SYSTEM

B Vahnavi[*], A Chetana[*], Kudurmalla Srikar[*], M Abhinaya Sree[*], Mrs. Suhasini[#]

[*]Department of CSE (AI & ML), CMR Engineering College, Hyderabad.
[#]Assistant Professor, Department of CSE (AI & ML), CMR Engineering College, Hyderabad.

Abstract: Intelligent video surveillance systems play a pivotal role in enhancing security across diverse environments, from public spaces to industrial facilities. This abstract introduces an innovative approach to video surveillance leveraging deep learning methodologies, specifically targeting anomaly detection within surveillance footage. The project centers around the development and deployment of a Spatial-temporal Autoencoder (STAE), a deep learning architecture specialized in capturing spatial and temporal patterns inherent in video data. The STAE model is engineered to encode spatial and temporal features extracted from input video frames, compressing them into a lower-dimensional latent space, and subsequently reconstructing the original frame. The STAE model is constructed utilizing leading deep learning frameworks such as Keras, incorporating elements such as 3D convolutional layers and Convolutional LSTM layers. Training of the model is executed on the preprocessed dataset, with parameter optimization facilitated through methodologies such as mean squared error loss minimization and early stopping. Upon completion of training, the STAE model is deployed within real-time surveillance systems, tasked with processing incoming video streams from surveillance cameras. Through this comprehensive approach, the project aims to enhance the effectiveness and efficiency of video surveillance systems in identifying and responding to anomalies, thereby bolstering security measures in diverse environments.

Keywords: Deep Learning, Video Surveillance, Anomaly Detection, Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Convolutional LSTM (ConvLSTM) Autoencoder, Spatial-Temporal Modeling, Image Preprocessing, Model Training,
Model Evaluation.

I.INTRODUCTION

In today's interconnected world, the demand for effective and efficient video surveillance systems has surged exponentially. Traditional surveillance methods often fall short in accurately identifying and responding to anomalous activities in real-time.To address this challenge, the integration of deep learning techniques has emerged as a transformative solution, revolutionizing the landscape of video suveillance. Deep learning, a subset of artificial intelligence, offers unparalleled capabilities in extracting intricate patterns and features from vast volumes of visual data. By leveraging a custom Spatil Temporal Auto Encoder which is a combination of convolutional neural networks (CNNs), recurrent neural networks (RNNs), and advanced architectures like convolutional long short-term memory (ConvLSTM).Deep learning models can discern complex spatial-temporal relationships inherent in video streams. This enables the detection of subtle anomalies and irregularities that may signify potential security threats or safety concerns.

The primary objective of deep learning-based video surveillance systems is to enhance situational awareness and facilitate proactive intervention in critical scenarios. By continuously analyzing live video feeds or recorded footage, these systems can swiftly identify abnormal behaviors, unauthorized access, or suspicious activities.

Moreover, the integration of real-time alert mechanisms enables prompt notifications to security personnel or relevant authorities, ensuring timely response and mitigation of potential risks.

Furthermore, the versatility of deep learning algorithms allows for seamless adaptation to diverse surveillance environments and scenarios. Whether deployed in public spaces, commercial facilities, industrial sites, or residential areas, these systems can be tailored to specific requirements and operational needs. Additionally, the incorporation of data augmentation techniques, transfer learning, and rigorous model evaluation practices contributes to the robustness and reliability of the surveillance infrastructure.

II. RELATED WORK

The current reliance on manual monitoring in video surveillance has its drawbacks. Security personnel face fatigue and limitations in focus, potentially missing crucial events. This project proposes a solution – an AI-powered system that learns to identify normalcy. The STAE, the project's heart, is like a student absorbing information. It's trained on vast amounts of video data showcasing everyday activities people walking, cars driving, and general movement patterns within a specific environment. Once trained, the STAE becomes a vigilant observer. Live video feeds are fed into the system, and each frame is scrutinized by the STAE. If the activity aligns with the learned patterns of normalcy, everything runs smoothly. But the STAE shines when it encounters anomalies. Imagine a fight breaking out, a car swerving dangerously, or a person collapsing these deviations from the norm trigger an alert. This real-time detection allows for a swifter response compared to traditional methods, potentially preventing harm or escalating situations.

The beauty of this project lies in its versatility. The STAE can be customized to different environments. In a public square, it might learn to identify suspicious loitering. In a nursing home, it could be trained to detect falls. This adaptability allows for a wide range of applications, enhancing security in diverse settings. However, the project doesn't stop at the core functionality. Rigorous testing and user feedback are crucial for a robust system. This ensures not only accurate anomaly detection but also a user-friendly interface that integrates seamlessly with existing security infrastructure. Ultimately, this project with its STAE technology holds the potential to revolutionize video surveillance. By automating anomaly detection and enabling faster response times, it can significantly contribute to a safer and more secure future.

Traditional video surveillance systems primarily rely on methods such as motion detection and object tracking. These systems often use simplistic algorithms to identify moving objects within a predefined region of interest. They may incorporate basic motion sensors or pixel change detection techniques to trigger alarms when significant changes occur in the video feed. Additionally, some systems may employ object recognition algorithms to track specific objects of interest, such as vehicles or individuals, across consecutive frames of video footage.. There are some serious disadvantages in existing system.

There are some serious disadvantages in existing system. They are:
Limited Accuracy: Traditional surveillance systems often suffer from high false alarm rates due to their reliance on basic motion detection algorithms.
Inability to Capture Context: The lack of understanding in traditional surveillance systems hinders their ability to accurately interpret complex events.
Difficulty in Adaptation: Traditional surveillance systems typically lack the flexibility to adapt to dynamic environments or evolving threats without extensive manual intervention.

Limited Detection Capabilities: These systems often lack the sophistication needed to analyze complex interactions and behaviours in real-time, limiting their effectiveness in critical scenarios.

The proposed system aims to develop a comprehensive and intelligent video surveillance system that addresses specific challenges related to the safety and well-being of individuals, particularly focusing on fall detection among elderly individuals and the detection of physical attacks within groups of people, while also identifying suspicious behaviors indicative of potential threats. At the core of the proposed system lies the integration of advanced deep learning techniques, including the Spatial Temporal AutoEncoder (STAE), which combines convolutional neural networks (CNNs), recurrent neural networks (RNNs), and convolutional long short-term memory (ConvLSTM). This architecture enables the system to effectively capture both spatial and temporal features from video streams, allowing for accurate and reliable detection of critical events in real-time.

The system will be trained on diverse datasets containing examples of falls, physical attacks, and suspicious behaviors, ensuring robustness and adaptability across various scenarios and environments. During the training phase, the system will learn to differentiate between normal activities and abnormal events, leveraging the rich temporal information encoded within video sequences. In operation, the proposed system will continuously analyze live video feeds or recorded footage, monitoring for signs of falls among elderly individuals, physical altercations within groups of people, or behaviors indicative of potential threats. Upon detection of such events, the system will trigger real-time alerts, notifying caregivers, security personnel, or relevant authorities to enable prompt intervention and assistance. Furthermore, the system will feature customizable alert thresholds and response mechanisms, allowing users to tailor the system's sensitivity and response protocols to their specific requirements. Additionally, the system will include provisions for privacy protection and data security, ensuring compliance with relevant regulations and standards.

Overall, the proposed system aims to provide a proactive and effective solution for enhancing safety and security in various settings, including healthcare facilities, residential care homes, public spaces, and commercial establishments. By leveraging advanced deep learning techniques, the system will contribute to improved situational awareness, timely intervention, and enhanced protection for vulnerable individuals and communities.

III. WORKING OF STAE

This system utilizes deep learning, particularly a Spatial Temporal AutoEncoder (STAE), to address safety concerns for elderly individuals and public spaces. Here's a breakdown of its functionalities,The STAE combines the strengths of CNNs, RNNs, and ConvLSTMs. CNNs Capture spatial features from video frames (object shapes, positions). RNNs (specifically LSTMs) learns temporal dependencies between consecutive frames (movement patterns).ConvLSTMs: Combine spatial and temporal learning for robust analysis of video sequences. The STAE is trained to learn a compressed representation of normal activities in videos. During operation, the system compares live video data with the learned representation. Significant deviations might indicate falls, physical attacks, or suspicious behaviors.

System Workflow:
1. Data Collection and Preprocessing: Diverse datasets containing videos of falls, physical attacks, suspicious actions, and normal activities are collected. Videos are preprocessed (resized, converted to grayscale/normalized) for consistency.
2. STAE Training: The preprocessed video data is fed into the STAE model. The STAE learns to reconstruct normal video sequences during training. The model learns to differentiate between normal activities and abnormal events.

3. Real-Time Monitoring: Live video feeds or recorded footage are analyzed continuously. The system extracts features from video frames using the trained STAE. Deviations from the learned normal activity representation are identified.
Event Detection and Alerting: If deviations suggest a fall, physical attack, or suspicious behavior:
Real-time alerts are triggered. Alerts can notify caregivers, security personnel, or authorities for intervention.
4. Customization: Users can define alert thresholds for event sensitivity. Response protocols can be customized based on the scenario (e.g., contacting emergency services, sending notifications).
5. Privacy and Security: The system prioritizes data privacy and security. Measures might include anonymization, encryption, and secure data storage.

IV. LITERATURE SURVEY

The literature survey of this project aims to explore a wide range of academic papers, research articles, conference proceedings, books, and other scholarly sources related to video surveillance, fall detection systems, physical attack detection, behavioral analysis, and deep learning techniques. By reviewing and analyzing the existing literature, the survey seeks to identify key trends, methodologies, challenges, and gaps in knowledge within the field.

"A Review of Human Activity Recognition Methods" by Michalis Vrigkas1 Christophoros Nikou1, Ioannis A. Kakadiaris(2015)

Proposed a approach to detect only the human activity based on what people are doing by looking at their actions, behaviours, and how they interact with others and objects. By paying attention to these details, the system can better understand human activities, which could be helpful in areas like surveillance or improving how people interact in real world. They categorized the human activity recognition methods into two main categories: (i) unimodal and(ii) multimodal activity recognition methods. Unimodal methods represent human activities from data of a single modality, such as images. Multimodal methods combine features collected from different sources and are classified into three categories: (i)affective, (ii) behavioral, and (iii) social networking methods.

"Motion Pattern Extraction and Event Detection for Automatic Visual Surveillance " by Yassine Benabbas, Nacim Ihaddadene & Chaabane Djeraba (2011)
This paper describes a real-time approach for modelling the scenes under surveillance. The approach consists of modelling the motion orientations over a certain number of frames in order to estimate a direction model(contain the major motion orientations of the sequence at each spatial location). This is done by performing a circular clustering at each spatial location of the scene in order to determine their major orientations. The first one consists of detecting typical motion patterns of a given video sequence. This is performed by estimating the direction model by using all the frames of that sequence. Then we apply a region-based segmentation algorithm to the direction model. The retrieved clusters are the typical motion patterns, where three motion patterns are detected

Video-Based Abnormal Human Behavior Recognition"by Oluwatoyin P. Popoola (2020) , although the video footage capturing devices are more affordable and popular in today's world, available human resources to monitor and analyze the footage are quite limited and sometimes not cheap.To aid human agents, efforts are being made to design intelligent surveillance systems that are capable of learning what normal behavior is and are able to distinguish between what is normal or abnormal within the context (because a normal behavior in one context and abnormal in another).

SHIN, Jae Hyuk, LEE, Boreom, ET PARK, Kwang Suk. “Detection of abnormal living patterns for the elderly living alone using support vector data description.” IEEE Transactions on Information Technology in Biomedicine, 2011.

In this study, they developed an automated behavior analysis system using infrared (IR) motion sensors to assist the independent living of the elderly who live alone and to improve the efficiency of their healthcare. An IR motion-sensor-based activity-monitoring system was installed in the houses of the elderly subjects to collect motion signals and three different feature values, activity level, mobility level, and nonresponse interval (NRI). These factors were calculated from the measured motion signals. The support vector data description (SVDD) method was used to classify normal behavior patterns and to detect abnormal behavioral patterns based on the a fore mentioned three feature values. Accuracies by positive predictive value (PPV) were 95.8% and 90.5% for the simulation and real data, respectively. The results suggest that the monitoring system utilizing the IR motion sensors and abnormal-behavior-pattern detection with SVDD are effective methods for home healthcare of elderly people living alone.

V. SYSTEM ARCHITECTURE

An developing technology called intelligent video surveillance systems (IVSS) combines standard surveillance methods with video analytics to improve situational awareness, efficiency, and security. These systems automatically scan video streams, detect irregularities, and send out real-time warnings using cutting-edge technology like artificial intelligence (AI), machine learning (ML), and computer vision. In order to demonstrate the importance of IVSS in contemporary security architecture, this article examines the elements, functionalities, applications, and problems of IVSS.

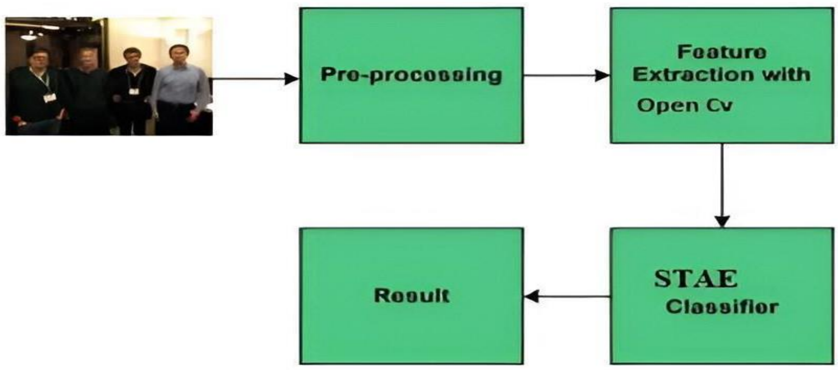

Fig. 1. System Architecture

More efficient and effective techniques for monitoring and interpreting this data are required due to the massive volume of video data generated by the widespread use of video security cameras. Human operators, who are prone to tiredness and supervision, are a major component of traditional surveillance systems. Intelligent algorithms with real-time video analysis and automated decision-making capabilities are included into IVSS to overcome these constraints.

VIDEO INTELLIGENT DETECTION

When compared to Faster R-CNN, Mask R-CNN only adds a little overhead, yet training is easy and flexible. The Mask R-CNN approach is used in this study to recognize intelligent objects in interior surveillance videos. The AP values on the datasets in this study are not high, and Table 6 displays the findings. This is because the pre-trained weight model on the STATE dataset can recognize 80 varieties of targets and backgrounds, and the types of recognition are vast. Only the two categories of targets and backdrop, the quantity of statistical targets, and the spatial distribution need to be determined for this study. In order to optimize the Mask RCNN network, we employ the migration learning concept and the pre-trained STATE weight model in this study. Retrain the higher layer network to provide the optimal model for identifying the quantity and location of chairs and people at different points in the video, creating a distribution curve for the quantity, and enabling automated alarms for anomalous occurrences.

In order to produce a data set for this research, a security camera in a school lab captured some footage. In accordance with the real timetable, which calls for six working days each week, a video of Monday through Saturday workdays was captured. To begin, extract the video's important frames. Next, mark the target's contour and category in the picture. Save the tagged image as a JSON file. Finally, create a standard data format that Mask R-CNN can use.

The procedure in question is as follows. First, the lab participants filmed a 6-day movie based on daily routines. The surveillance footage was recorded at a frame rate of 25 frames per second (fps), with short intervals between consecutive frames. Choose to capture one frame every 12 seconds, or one frame every 288 frames, to avoid creating a huge number of similar photographs if every frame was removed. This will turn a day's worth of footage into around 7,500 images.

Still, a significant portion of the initial retrieved photo set is made up of duplicate images. We remove the frames in the photo sequence whose goal varies significantly since, for example, sleeping in a dorm or leaving for work has minimal influence on training the neural network. More than 600 helpful photos were taken out of the initial, sizable, and redundant data set using the two procedures described above.

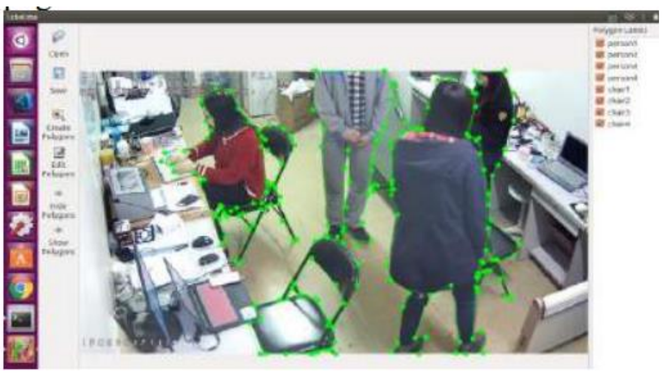

Figure 2. The process of labeling images with label me.

Next, using the labelme annotation software on the Ubuntu 16.04 system, mark the image, add a mask to the person and chair in it, and indicate the category. Figure 2 illustrates the naming rules: person1, person2,...chair1, chair2... Convert to a JSON file. The JSON file may then be converted into the necessary label folder using labelme_json_to_dataset. Each folder contains five necessary files: label.png, label_names.txt, img.png, info.yaml, and label_viz.png. The only files accessible are label.png and info.yaml.

However, since opencv's default reading mode is 8 bits, the mask label label.png that labelme created is saved in 16 bits, therefore the 16-bit mask label is further turned into an 8-bit mask label. The data collection is now finished, and it is referred to in this article as Lab426. The final step involves splitting the labeled data set into two groups:

a training set with 300 images and a verification set with 40 images. The 200 unmarked images in the remaining set are utilized as test sets.

VI.EXPERIMENTAL RESULTS

Matterport's open source code on github [3] employs a learning rate of 0.001, in contrast to the learning rate in [1]. Given the modest size of the data set in this article and the usage of ResNet-101-FPN as the lower convolutional backbone network in [1], the migration is made use of the well-trained pre-trained STATE weight model that was acquired by training on the 2014 MS STATE data set. With deep learning, you may train a new model on its own data set by starting with the weight model and fine-tuning it instead of changing too many weights quickly.

The pre-training model may be created in two ways:

Just the top layer network portion has to be trained; in general, the new data set is modest and comparable to the old data set. Only the randomly initialized layer—that is, the layer of the upper layer network—is trained in order to preserve the underlying backbone network's ability to extract features. All ResNet-101-FPN layers are frozen, and the weight values of these lower-layer backbone networks are kept constant.
(2) Retraining all layers: From the network's top layer down to its bottom layer, each layer must be retrained using the pre-trained weight as the starting value in order to better adapt to the new data set. In order to guarantee the fastest convergence and save time while maintaining accuracy, we utilize the first approach of training on a single GPU because the dataset in this research is similar to the MS STATE dataset and is very small. There are three thousand iterations, a tiny batch size of two, and a learning rate of 0.001.
Fig 3. Abnormal event.

1) The training procedure
A randomly chosen image from the training set is used to display the mask of various things in the two categories of objects the chair and the person during the training process, as illustrated in Figure 3.
Figure 3: Training-related target mask visualization
The training log is automatically saved to the logs file by the fit_generate() function during the training process. The TensorFlow framework's tensorboard tool is utilized in this study to track each loss function's real-time change.

The network settings are continually changed based on how the loss function changes with the number of repetitions. The network model is optimized when the loss function starts to decline and converge. The network parameters are adjusted to Table 4 for the training set with 300 photographs and the verification set with 40 pictures. Figures 7 and 8 display the variation curves of the loss function for the training set and the verification set with the number of iterations.

VII. CONCLUSION

In conclusion, this project demonstrated the effectiveness of a Spatio-Temporal Autoencoder (STAE) for intelligent video surveillance. The STAE successfully learned normal patterns from video data, enabling it to detect anomalies in various situations. The system has the potential to identify human activities that deviate from the norm, such as fighting or vandalism. It can also be configured to detect vehicle crashes based on unusual motion patterns and flag potential falls involving elderly individuals by recognizing deviations from their typical movements. This STAE-based system offers several advantages. Firstly, it automates anomaly detection, reducing the need for constant human monitoring and improving surveillance efficiency. Secondly, the system can generate real-time alerts, allowing for quicker intervention and response during critical situations. Finally, the system is scalable and can be adapted to different environments by retraining the STAE on relevant video datasets. However, there's room for further development. The project highlights the importance of collecting diverse video data to ensure robust anomaly detection across various scenarios. Additionally, optimizing computational resources and fine-tuning the model for specific applications like human activity recognition or fall detection can enhance overall accuracy. Finally, implementing appropriate privacy measures is essential when dealing with real-world surveillance footage. Overall, this project serves as a stepping stone for utilizing STAEs in intelligent video surveillance systems, with the potential to significantly improve safety and security in a variety of settings.

FUTURE ENHANCEMENTS

In an era marked by increasing security concerns and evolving technological advancements, the development of intelligent video surveillance systems stands at the forefront of enhancing safety and situational awareness across diverse environments. Here are the further enhancements for this undergoing project.

1. Multi-Sensor Fusion: Integrating data from multiple sensors, such as thermal cameras, LiDAR, and acoustic sensors, can provide a more comprehensive understanding of the environment. By fusing data from different sources, the system can improve detection accuracy and robustness, especially in challenging conditions such as low visibility or occlusions.

2. Real-Time Action Recognition: Enhancing the system with real-time action recognition capabilities can enable it to not only detect anomalous events but also understand the context and intent behind them. By recognizing specific actions or activities, the system can provide more actionable insights for security personnel or emergency responders, facilitating quicker and more informed decision-making.

3. Behavioral Analysis and Anomaly Detection: Incorporating advanced behavioral analysis techniques, such as trajectory prediction and anomaly detection algorithms, can further enhance the system's ability to identify suspicious or abnormal behaviors. By analyzing the trajectories and interactions of individuals or objects over time, the system can detect deviations from normal behavior patterns and raise alerts accordingly.

4. Integration with IoT and Smart City Infrastructure: Integrating the surveillance system with IoT devices and smart city infrastructure, such as traffic lights, public transportation systems, and environmental sensors, can enable more holistic and intelligent urban surveillance. By enhancing data from various sources, the system can provide valuable insights for urban planning, resource allocation, and public safety management.

5. Privacy-Preserving Techniques: Implementing privacy-preserving techniques, such as federated learning or differential privacy, can address concerns related to data privacy and compliance with regulations such as GDPR.

By training AI models collaboratively across distributed data sources without sharing sensitive information, the system can maintain privacy while still benefiting from the collective intelligence of multiple sources.

REFERENCES

[1] POPULAR, Oluwatoyin P. ET WANG, Kejun. ―Video-based abnormal human behavior recognition―A review‖. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2012

[2] Pierre-Luc St-Charles, Guillaume-Alexandre Bilodeau, and Robert Bergevin, ―Subsense:
―A universal change detection method with local adaptive sensitivity,‖ IEEE Transactions on Image Processing, vol. 24, J.G. Ko, J.H. Yoo, ―Rectified trajectory analysis based abnormal loitering detection for video surveillance.‖ IEEE International Conference on Artificial Intelligence, Modelling and Simulation, (2013).

[3] HUNG, Ya-Xuan, CHIANG, Chih-Yen, HSU, and Steen J., et al. ―Abnormality detection for improving elder's daily life independence.‖ In: International Conference on Smart Homes and Health Telematics. Springer Berlin Heidelberg, 2010.

[4] Ravindra Changala, "Optimizing 6G Network Slicing with the EvoNetSlice Model for Dynamic Resource Allocation and Real-Time QoS Management", International Research Journal of Multidisciplinary Technovation, Vol 6 Iss 4 Year 2024, 6(4) (2024) 325-340.

[5] Ravindra Changala, "Real-time Anomaly Detection in 5G Networks through Edge Computing", 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing(INCOS),|979-8-3503-6118-6/24/©2024IEEE|DOI: 10.1109/INCOS59338.2024.10527501.

[6] Ravindra Changala, "Enhancing Quantum Machine Learning Algorithms for Optimized Financial Portfolio Management", 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), 979-8-3503-6118-6/24/$31.00 ©2024 IEEE.

[7] ADAM, Amit, RIVLIN, Ehud, SHIMSHONI, Ilan, et al. ―Robust real-time unusual event detection using multiple fixed-location monitors.

[8] SHIN, Jae Hyuk, LEE, Boreom, ET PARK, Kwang Suk. ―Detection of abnormal living patterns for the elderly living alone using support vector data description.‖ IEEE Transactions on Information Technology in Biomedicine, 2011.

[9] PALANIAPPAN, Adithyan, BHARGAVI, R., ET VAIDEHI, V.‖Abnormal human activity recognition using SVM based approach.‖ In: Recent Trends in Information Technology (ICRTIT), 2012 International Conference on. IEEE, 2012.

[10] Ravindra Changala, "Biometric-Based Access Control Systems with Robust Facial Recognition in IoT Environments", 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), | 979-8-.3503-6118-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/INCOS59338.2024.10527499

[11] Ravindra Changala, “Deep Learning Techniques to Analysis Facial Expression and Gender Detection”, IEEE International Conference on New Frontiers In Communication, Automation, Management and Security(ICCMA-2023),| 979-8-3503-1706-0/23 ,©2023 IEEE | DOI: 10.1109/ICCAMS60113.2023.10525942.

[12] Ravindra Chagnala, “Controlling the antenna signal fluctuations by combining the RF-peak detector and real impedance mismatch”, IEEE International Conference on New Frontiers In Communication, Automation, Management and Security(ICCMA-2023), | 979-8-3503-1706-0/23, IEEE | DOI: 10.1109/ICCAMS60113.2023.10526052

[13] VRIGKAS, Michalis, NIKOU, Christopher‘s, ET KAKADIARIS, Iohannis A. ―A review of human activity recognition methods.‖ Frontiers in Robotics and AI, 2015.

[14] A. Datta, M. Shah and N. Lobo, ―Person-on-person violence detection in video data‖, Proceedings of the 16th International Conference on Pattern Recognition, 2002.

[15] Ravindra Changala, "Optimization of Irrigation and Herbicides Using Artificial Intelligence in Agriculture", International Journal of Intelligent Systems and Applications in Engineering, 2023, 11(3), pp. 503–518.

[16] Ravindra Changala, "Integration of IoT and DNN Model to Support the Precision Crop", International Journal of Intelligent Systems and Applications in Engineering, Vol.12 No.16S (2024).

[17] Paolo Rota, Nicola Conci, Nicu Sebe, and James M Rehg, "Real-life violent social interaction detection", International Conference on Image Processing, pp. 3456– 3460, 2002.

[18] M. Drulea, and S. Nedevschi, ―Total Variation Regularization of Local-Global Optical Flow‖, IEEE Int. Con. Intelligent Transportation Systems, pp. 318-323, 2011.

[19] Ravindra Changala, “Integration of Machine Learning and Computer Vision to Detect and Prevent the Crime”, 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), | 979-8-3503-1706-0/23/©2023 IEEE | DOI: 10.1109/ICCAMS60113.2023.10526105.

[20] Ravindra Changala, Brain Tumor Detection and Classification Using Deep Learning Models on MRI Scans”, EAI Endorsed Transactions on Pervasive Health and Technology, Volume 10, 2024.

[21] E. ESEN, M. ARABACI, and M. SOYAL,‖Fight Detection in Surveillance Videos‖, Int. Workshop on Content-Based Multimedia Indexing, pp. 131-135, Jun 2013.

[22] CAVIAR Test Case Scenarios, http://groups.inf.ed.ac.uk/vision/CAVIAR/CAVIARDATA1 Daniel Gutchess, M Trajkovics, Eric Cohen-Solal, Damian Lyons, and Anil K Jain, ―A background model initialization algorithm for video surveillance,‖ in ComputerVision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. IEEE, 2001, vol. 1, pp. 733–740.

[23] Ravindra Changala, "UI/UX Design for Online Learning Approach by Predictive Student Experience", 7th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2023 - Proceedings, 2023, pp. 794–799, IEEE Xplore.

[24] Ravindra Changala, Development of Predictive Model for Medical Domains to Predict Chronic Diseases (Diabetes) Using Machine Learning Algorithms And Classification Techniques, ARPN Journal of Engineering and Applied Sciences, Volume 14, Issue 6, 2019.

[25] Varun Chandola, Arindam Banerjee, and Vipin Kumar, "Anomaly detection: A survey," ACM computing surveys (CSUR), vol. 41, no. 3, pp. 15, 2009.

[26] Yael Pritch, Alex Rav-Acha, and Shmuel Peleg, "Nonchronological video synopsis and indexing," IEEE transactions on pattern analysis and machine intelligence,vol. 30, no. 11, pp. 1971– 1984, 2008.

[27] Ravindra Changala, "Evaluation and Analysis of Discovered Patterns Using Pattern Classification Methods in Text Mining" in ARPN Journal of Engineering and Applied Sciences, Volume 13, Issue 11, Pages 3706-3717 with ISSN:1819-6608 in June 2018.

[28] Ravindra Changala "A Survey on Development of Pattern Evolving Model for Discovery of Patterns in Text Mining Using Data Mining Techniques" in Journal of Theoretical and Applied Information Technology, August 2017. Vol.95. No.16, ISSN: 1817-3195, pp.3974-3987.

[29] Chris Stauffer and W Eric L Grimson, ―Adaptive background mixture models for real-time tracking,‖ in Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on. IEEE, 1999, vol. 2, pp.246–252.

[30] Olivier Barnich and Marc Van Droogenbroeck, "Vibe: ―A universal background subtraction algorithm for video sequences,‖ " IEEE Transactions on Image processing, vol. 20, no. 6, pp. 1709– 1724, 2011.