IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 5, Issue 2, January 2025

Deep Learning Approach For Suspicious Activity Detection from Surveillance Video in Examination Hall

Shradha Auti[1], Priyanka Darade[2], Ankita Munje[3], Neha Patil[4], Prof. Anoop Khushwaha[5]

Final Year Students, Department of Computer Engineering[1,2,3,4]

Assistant Professor, Department of Computer Engineering[5]

Alard College of Engineering and Management, Hinjawadi, Pune, India

Savitribai Phule Pune University, Pune, Maharashtra, India

Abstract: *This paper presents a deep learning-based approach for detecting suspicious activities in examination halls using surveillance video. The traditional methods of monitoring students during examinations have limitations, particularly when it comes to real-time detection and alerting. To address this issue, a system is proposed that employs computer vision and deep learning techniques to identify suspicious behaviors such as cheating, unauthorized movements, or distractions. The system processes video feeds from surveillance cameras, analyzes the actions of students, and triggers real-time alerts when abnormal behavior is detected. Experimental results show that the proposed system can detect suspicious activities with high accuracy, providing an automated and reliable mechanism for monitoring exam environments. This paper highlights the potential of deep learning to enhance security and integrity in educational settings.*

Keywords: Suspicious activity detection, Surveillance video, Deep learning, Examination hall, Real-time alert

I. INTRODUCTION

Human behavior detection in real-world environments has numerous applications, such as intelligent video surveillanceand analyzing shopping patterns. Video surveillance, particularly through CCTV cameras, plays a vital role in maintaining security in public spaces. It enables the detection of suspicious activities without requiring human intervention, making it an essential tool for safety. This paper focuses on identifying suspicious activities through surveillance video and notifying the relevant authorities when such activities are detected. CCTV cameras have become integral to daily life, providing security and surveillance in various settings. As part of the Indian government's Digital India initiative, e-surveillance has gained significant attention in enhancing public safety and security.

1.1 Aim

The primary objective of this project is to detect suspicious activities using video surveillance and generate alerts or notifications for the concerned users. The system processes video inputs from a dataset and passes them through a Convolutional Neural Network (CNN) model to classify the activities as suspicious or not.

1.2 Scope of Project

The recognition of suspicious human activities in video surveillance plays a crucial role in preventing incidents such as theft in sensitive areas, including banks, hospitals, shopping malls, parking lots, bus and railway stations, airports, refineries, nuclear power plants, educational institutions, and border areas.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 5, Issue 2, January 2025

II. DETAILED DESCRIPTION OF OUR PROJECT

This project focuses on detecting suspicious activities in real-time surveillance video using deep learning. Video data is collected, preprocessed, and passed through a Convolutional Neural Network (CNN) for activity classification. The model is trained on a labeled dataset to distinguish between normal and suspicious behaviors. When suspicious activity is detected, the system triggers an immediate alert to the authorities. The real-time alert system ensures timely intervention in critical situations. Future improvements will focus on enhancing detection accuracy under various environmental conditions.

III. LITERATURE SURVEY

1. Paper Name:Real-Time suspicious Detection and Localization in Crowded Scenes
Author: Mohammad Sabokrou , Mahmood Fathy
Abstract **:** In this paper, we propose a method for real-time suspecious detection and localization in crowded scenes. Each video is defined as a set of non-overlapping cubic patches, and is described using two local and global descriptors. These descriptors capture the video properties from different aspects. By incorporating simple and cost-effective Gaussian classifiers, we can distinguish normal activities and anomalies in videos. The local and global features are based on structure similarity between adjacent patches and the features learned in an unsupervised way, using a sparse autoencoder. Experimental results show that our algorithm is comparable to a state-of-the-art procedure on UCSD ped2 and UMN benchmarks, but even more time-efficient. The experiments confirm that our system can reliably detect and localize anomalies as soon as they happen in a video.

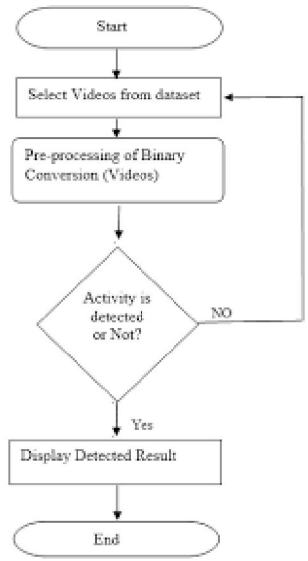
2.Paper Name: Learning Temporal Regularity in Video Sequences
Author: Mahmudul Hasan Jonghyun Choi
Abstract **:**Perceiving meaningful activities in a long video sequence is a challenging problem due to ambiguous definition of 'meaningfulness' as well as clutters in the scene. We approach this problem by learning a generative model for regular motion patterns (termed as regularity) using multiple sources with very limited supervision. Specifically, we propose two methods that are built upon the autoencoders for their ability to work with little to no supervision. We first leverage the conventional handcrafted spatio-temporal local features and learn a fully connected autoencoder on them. Second, we build a fully convolutional feed-forward autoencoder to learn both the local features and the classifiers as an end-to-end learning framework. Our model can capture the regularities from multiple datasets. We evaluate our methods in both qualitative and quantitative ways - showing the learned regularity of videos in various aspects and demonstrating competitive performance on suspecious detection datasets as an application.

3.Paper Name: Suspicious Detection in Video Using Predictive Convolutional Long Short-Term Memory Networks
Author::Jefferson Ryan Medel
*Abstract :*Automating the detection of anomalous events within long video sequences is challenging due to the ambiguity of how such events are defined. We approach the problem by learning generative models that can identify anomalies in videos using limited supervision. We propose end-toend trainable composite Convolutional Long Short-Term Memory (Conv-LSTM) networks that are able to predict the evolution of a video sequence from a small number of input frames. Regularity scores are derived from the reconstruction errors of a set of predictions with abnormal video sequences yielding lower regularity scores as they diverge further from the actual sequence over time. The models utilize a composite structure and examine the effects of 'conditioning' in learning more meaningful representations. The best model is chosen based on the reconstruction and prediction accuracies. The Conv-LSTM models are evaluated both qualitatively and quantitatively, demonstrating competitive results on suspecious detection datasets. Conv-LSTM units are shown to be an effective tool for modeling and predicting video sequences.

4.Paper Name: Abnormal Event Detection in Videos using Spatiotemporal Autoencoder
Author: : Yong Shean Chong
*Abstract :*We present an efficient method for detecting anomalies in videos. Recent applications of convolutional neural networks have shown promises of convolutional layers for object detection and recognition, especially in images.

However, convolutional neural networks are supervised and require labels as learning signals. We propose a spatiotemporal architecture for suspecious detection in videos including crowded scenes. Our architecture includes two main components, one for spatial feature representation, and one for learning the temporal evolution of the spatial features. Experimental results on Avenue, Subway and UCSD benchmarks confirm that the detection accuracy of our method is comparable to state-of-the-art methods at a considerable speed of up to 140 fps

3.1 Flowchart



Working of the Flowchart

1. Video Input: Surveillance video is captured and input into the system.
2. Preprocessing: The video is converted into frames for analysis.
3. Object Detection (YOLOv3): YOLOv3 detects objects like people in each frame.
4. Suspicious Activity Detection: The system analyzes detected behaviors for suspicious patterns.
5. Activity Classification (CNN): A CNN model classifies the activity as normal or suspicious.
6. Alert Generation: If suspicious behavior is detected, an alert is triggered.
7. Display of Result: The system displays whether the activity is normal or suspicious.
8. Action by Authorities: Authorities take necessary actions based on the alert.

IV. IMPLEMENTATION ALGORITHM

Machine learning (ML) is a branch of artificial intelligence (AI) that allows software systems to learn from historical data, improving their ability to make predictions or decisions without explicit programming for each scenario. Machine learning plays a central role in today's industries, such as in Facebook, Google, and Uber, where its application significantly boosts accuracy and efficiency. As machine learning evolves, it has become a vital tool for gaining competitive advantages in numerous fields, particularly in surveillance and security systems.

4.1 Convolution Neural Network

A Convolutional Neural Network (CNN) is a specialized deep learning algorithm used extensively for image classification and pattern recognition tasks. CNNs work by processing visual data, such as images and video frames, through a series of layers designed to automatically detect patterns like edges, shapes, and objects. What makes CNNs particularly powerful is their ability to learn these features directly from data, eliminating the need for manual feature engineering, which is required in traditional image processing techniques

CNNs are particularly effective for applications like human activity recognition, where understanding the context of an image or video frame is essential. By using CNNs, we can train the system to recognize suspicious behavior patterns in surveillance footage, which can trigger real-time alerts for intervention.

The process of detecting suspicious activities and triggering alarms involves the following steps:

Step 1: Convolution

The first operation in a CNN is convolution, where filters (also called kernels) are applied to the input image or video frame. These filters are small matrices that slide over the image to detect local features such as edges, corners, and textures. The result of this operation is a set of feature maps that represent different aspects of the input image, such as the presence of certain objects or shapes. The convolution operation reduces the input data size while preserving important information that is relevant for identifying patterns.

In the context of suspicious activity detection, the convolutional layers focus on detecting human figures, abnormal movements, and other key features in the surveillance video.

Step 2: Pooling

After convolution, pooling is performed to reduce the spatial dimensions of the feature maps. Pooling helps to summarize the features detected by the convolutional layers and reduces the computational load of the system. There are different types of pooling, such as max pooling and average pooling, but max pooling is commonly used in CNNs. It involves taking the maximum value within a specified region of the feature map, helping the model focus on the most prominent features in the image.

By reducing the size of the feature maps, pooling makes the system more efficient and robust to minor changes in the position or orientation of the features in the image. This is important for detecting suspicious activities, which may occur in various positions and angles in the surveillance footage.

Step 3: Flattening

Once the feature maps have been convolved and pooled, the next step is flattening. Flattening involves converting the two-dimensional feature map into a one-dimensional vector, which is then passed into the fully connected layers of the network. This vector contains the extracted features that will be used to classify the activity in the image as either normal or suspicious.

Step 4: Fully Connected Layer And Activity Classification

The flattened vector is fed into one or more fully connected layers, where each neuron is connected to every other neuron in the previous layer. These layers process the information and help classify the activity based on the learned features. In this step, the network learns to classify the pattern of activity as suspicious or not, based on training data.

The final output is a classification result that indicates whether the activity in the surveillance footage is suspicious. If the activity is classified as suspicious, the system proceeds to the next step: triggering an alarm.

Step 5: Real Time Alarm Trigging

In our implementation, an additional feature is added where, upon detecting suspicious activity, the system triggers an alarm in real-time. The system continuously monitors the live feed of surveillance video and, once it classifies an activity as suspicious, it sends an alert. The alarm can be in the form of a message, sound, or visual notification, depending on the system's configuration.

This real-time alerting mechanism is crucial for timely intervention by security personnel. It ensures that suspicious behavior is flagged immediately, allowing authorities to respond faster and take appropriate action. The system uses event-based triggers, ensuring that only relevant and critical suspicious activities generate alarms.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 5, Issue 2, January 2025

Step 6: Display And Action By Authorities

Once the alarm is triggered, security personnel or the monitoring system receives the alert, which includes important information such as the time, location, and nature of the suspicious activity. Authorities can then take necessary actions, including reviewing the footage, investigating the event further, or activating security protocols

This algorithm enhances the effectiveness of traditional surveillance systems by not only detecting suspicious activities but also providing a real-time mechanism to alert security personnel, improving response times and ensuring proactive security measures.

4.2 System Architecture

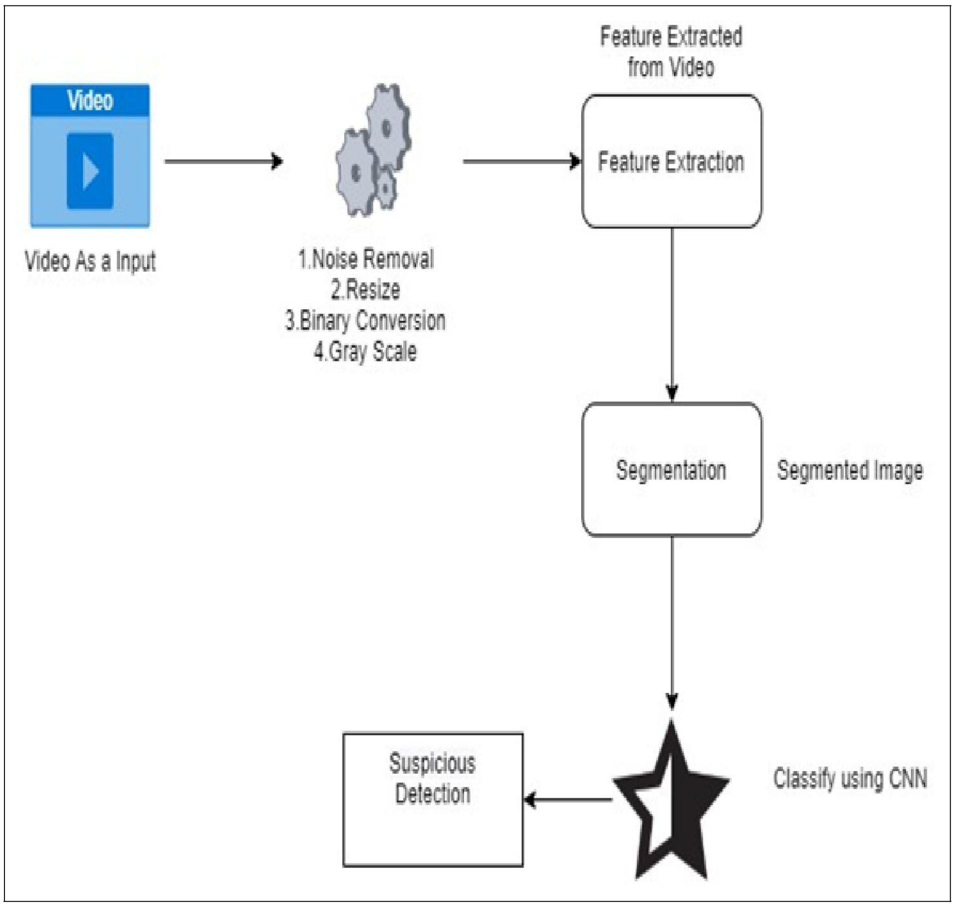


Figure 4.1: System Architecture

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 5, Issue 2, January 2025

4.3 Result

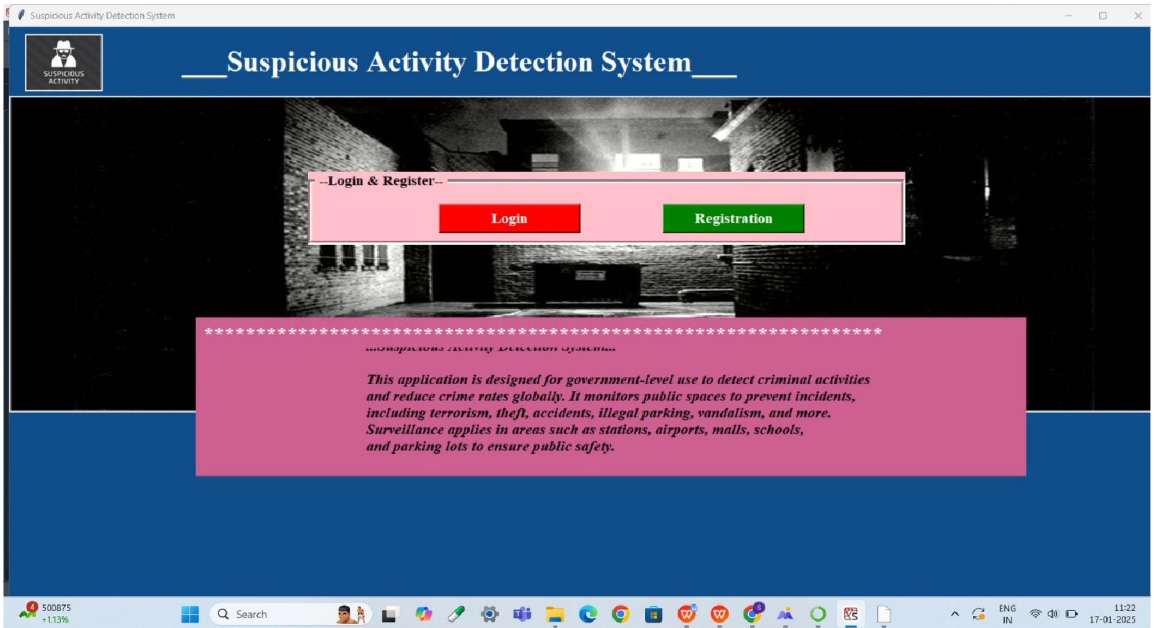


Fig.1 Home Page

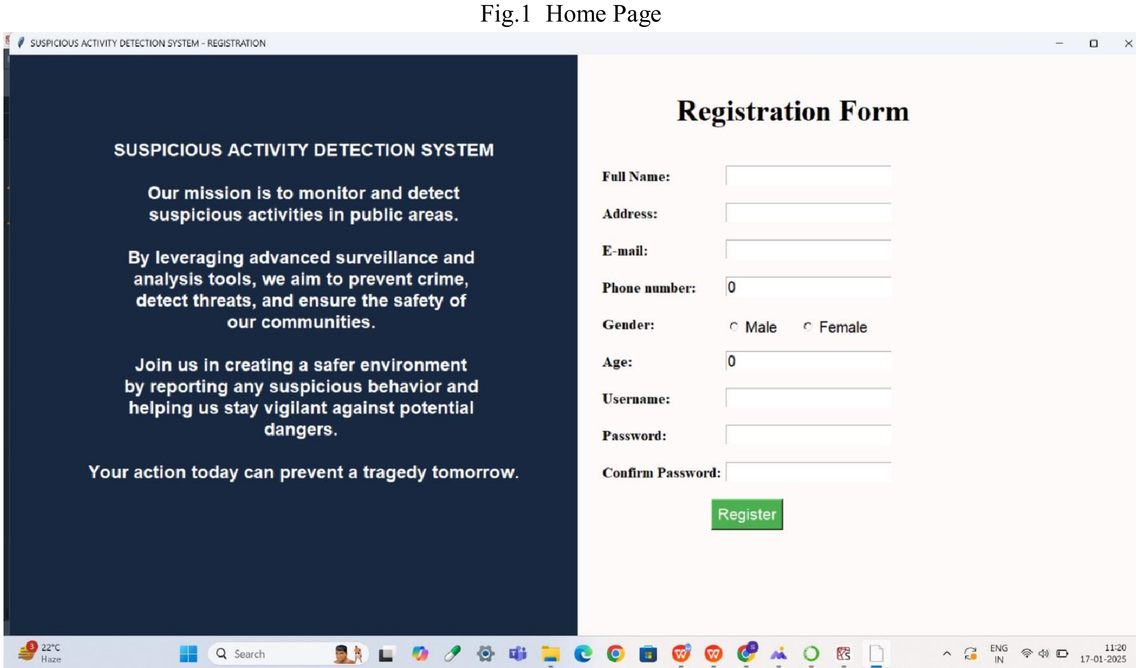


Fig.2 Registration Form

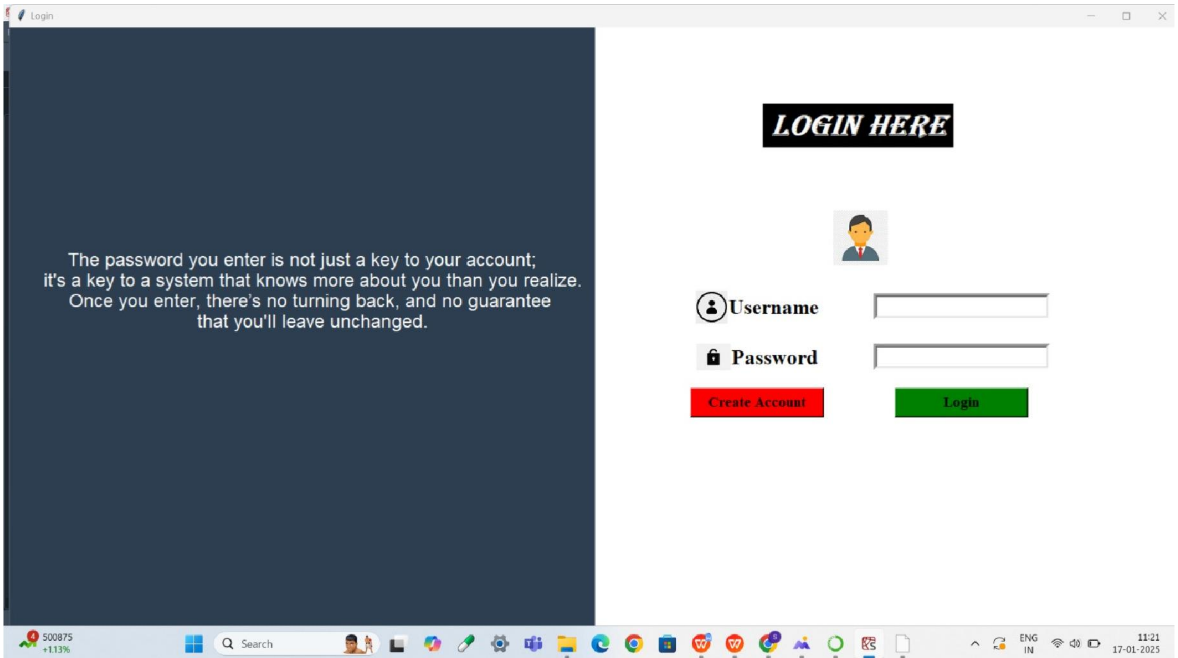


Fig.3 Login Page

V. CONCLUSION

A system to process real-time CCTV footage to detect any suspecious activity will help to create better security and less human intervention. Great strides have been made in the field of human suspecious Activity, which enables us to better serve the myriad applications that are possible with it. Moreover, research in related fields such as Activity Tracking can greatly enhance its productive utilization in several fields.

VI. ACKNOWLEDGMENT

We would like to express our profound gratitude to Prof. Anoop Khushwaha, our project guide, for his unwavering support, insightful feedback, and expert guidance throughout the development of this research work. His encouragement and valuable suggestions have played a pivotal role in shaping our understanding and enhancing the quality of our work.We extend our sincere thanks to ACEM College, Hinjawadi, affiliated with Savitribai Phule Pune University (SPPU), for providing us with an excellent academic platform and access to the resources required for the successful completion of this project. The institution's emphasis on research excellence and innovation has been a constant source of inspiration.Furthermore, we acknowledge the collaborative spirit of our peers and the continuous motivation from our friends and family, whose support has been instrumental in keeping us focused and committed throughout this research journey. Their encouragement has driven us to put forth our best efforts.

Finally, we express our appreciation to the researchers and authors whose work in the field of video surveillance and deep learning has served as a foundation for our study, inspiring us to contribute to the ongoing advancements in the domain.

REFERENCES

[1] Eralda Nishani, Betim Cico : "Computer Vision Approaches based on Deep Learning and Neural Networks" Deep Neural Networks for Video Analysis of Human Pose Estimation- 2017 6th MEDITERRANEAN CONFERENCE ON EMBEDDED COMPUTING (MECO), 11-15 JUNE 2017, BAR, MONTENEGRO

[2] Naimat Ullah Khan , Wanggen Wan : "A Review of Human Pose Estimation from Single Image"- 978-1-5386-5195-7/18/ 2018 IEEE

[3] Qiuhui Chen, Chongyang Zhang, Weiwei Liu, and Dan Wang, ”Surveillance Human Pose Dataset And Performance Evaluation For Coarse-Grained Pose Estimation”, Athens 2018.
[4] Baole Ai, Yu Zhou, Yao Yu : “Human Pose Estimation using Deep Structure Guided Learning”- 978-1-5090-4822-9/17 2017 IEEE DOI 10.1109/WACV.2017 .141
[5] Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh The Robotics Institute, Carnegie Mellon University ”Real time Multiperson 2D Pose Estimation using part affinity fields” - 1063-6919/17 2017 IEEE DOI 10.1109/CVPR.2017.143
[6] Hanguen Kim, Sangwon Lee, Dongsung Lee, Soonmin Choi, JinsunJu and Huyun Myung “Real-Time Human Pose Estimation and Gesture Recognition from depth Images Using Superpixels and SVM classifier.”- Sensors 2015, 15, 12410-12427; doi:10.3390/s150612410
[7] Tripathi, Rajesh and Jalal, Anand and Agarwal, Subhash(2017). ”Suspicious Human Activity Recognition: a Review”. Artificial Intelligence Review. 50.10.
[8] E. Eksioglu. Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3DMRI. Journal of Mathematical Imaging and Vision, 56(3):430–440, 2016.
[9] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, and D. Liang. Accelerating magnetic resonance imaging via deep learning. In Proceedings of the IEEE International Symposium on Biomedical Imaging, pages 514–517, 2016.
[10] L. Xu, J. Ren, C. Liu, and J. Jia. Deep convolutional neural network for image deconvolution. In Advances in Neural Information Processing Systems, pages 1790–1798, 2014.
[11] Y. Yang, J. Sun, H. Li, and Z. Xu. Deep ADMM-Net for compressive sensing MRI. In Advances in Neural Information Processing Systems, pages 10–18, 2016.
[12] Z. Zhan, J.-F. Cai, D. Guo, Y. Liu, Z. Chen, and X. Qu. Fast multiclass dictionaries learning with geometrical directions in MRI reconstruction. IEEE Transactions on Biomedical Engineering, 63(9):1850– 1861, 2016.

BIOGRAPHY

Miss. ShradhaAuti
B.E Computer Engineering Alard College Of Engineering And Management , Hinjawadi, Pune 41 Maharashtra, India

Miss. Priyanka Darade
B.E Computer Engineering Alard College Of Engineering And Management , Hinjawadi, Pune 41 Maharashtra, India

Miss.Ankita Munje
B.E Computer Engineering Alard College Of Engineering And Management , Hinjawadi, Pune 41 Maharashtra, India

Miss. Neha Patil
B.E Computer Engineering Alard College Of Engineering And Management , Hinjawadi, Pune 41 Maharashtra, India

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 5, Issue 2, January 2025

Assistant Professor

Prof. Anoop Khushwaha is a faculty member in the Department of Computer Engineering at ACEM College, Hinjawadi, affiliated with Savitribai Phule Pune University (SPPU). His expertise includes machine learning, deep learning, and computer vision, with a focus on intelligent video surveillance systems. He has guided several academic projects and continues to contribute to research and student mentorship