SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I4P108

Volume 11 Issue 4, 64-71, April 2024

Original Article

Intelligent Crime Scene Recognition: Advancing Public Safety through Deep Learning Architectures and Event Sequence Analysis

Anil Hingmire[1], Aman Sheikh[2], Saurabh Shukla[3], Rahul Shah[4], Sunayana Jadhav[5], Tatwadarshi P. Nagarhalli[6], Amrita Ruperee[7]

[1,2,3,4]Department of Computer Engineering, Vidyavardhini's College of Engineering & Technology, Maharashtra, India.
[5,7]Department of Electronics and Telecommunication Engineering, Vidyavardhini's College of Engineering & Technology, Maharashtra, India.
[6]Department of Artificial Intelligence and Data Science, Vidyavardhini's College of Engineering & Technology, Maharashtra, India.

[1]Corresponding Author : anil.hingmire@vcet.edu.in

Received: 06 February 2024 Revised: 06 March 2024 Accepted: 05 April 2024 Published: 30 April 2024

***Abstract** - In contemporary society, ensuring public safety is a paramount concern, and one of the significant challenges faced by law enforcement agencies is the swift detection and classification of criminal activities from surveillance footage. Current crime scene detection systems often lack real-time analysis and struggle with the prompt identification of criminal acts, hindering the timely response required for effective law enforcement. Consequently, there is a critical need for an advanced Crime Scene Detection System (CSDS) capable of classifying the type of crime occurring in real time, triggering immediate alarms, and aiding in the rapid identification of criminals captured within the surveillance footage. In this paper, we have highlighted various types' of techniques that can be used to detect crime scenes using Artificial Intelligence (AI). Crime detection using Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) involves the application of deep learning techniques to analyse and interpret complex patterns in crime-related data. The primary objective is to accurately identify and classify criminal activities, thereby assisting law enforcement agencies in taking proactive measures to ensure public safety. This multidimensional approach is essential for addressing the dynamic nature of criminal behavior, the diversity of criminal activities, and the need for real-time data processing. The analysis reveals that the LRCN model excels in accurately identifying crime events, achieving an impressive 94% accuracy. In contrast, the CNN-LRCN model lags behind with an 84% accuracy rate.*

***Keywords** - Artificial Intelligence, Crime Scene Detection System, Machine Learning, Deep Learning, Neural Network, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Chain snatching.*

1. Introduction

Artificial Intelligence (AI)-based crime detection is a revolutionary step forward for law enforcement, enabling organizations to respond to the changing nature of criminal activity with never-before-seen levels of efficiency and accuracy. The combination of cutting-edge algorithms, machine learning, and computer vision technologies is driving this paradigm change because it makes it possible to identify, analyze, and mitigate criminal activity in a proactive and adaptable manner.

The large volume and constantly changing nature of monitoring data presents one of the main obstacles to criminal detection. The sheer amount and complexity of information can be overwhelming for traditional approaches, which causes response and inquiry times to lag. Artificial Intelligence interventions seek to address these issues by utilizing real-time and predictive analytics. Using past crime data as fuel, predictive models can spot possible hotspots and patterns, giving law enforcement the ability to allocate resources and stop crimes before they start proactively. AI is essential to picture and video analysis in the field of visual data. The subtleties of conventional security film, which are typified by low lighting, warped viewpoints, and hidden features, present serious difficulties. Deep learning-capable computer vision systems are excellent at getting over these obstacles. Technologies like object recognition, anomaly detection, and facial recognition make it possible to quickly identify people and things of interest, which speeds up the process of solving cases.

Another aspect of AI-driven crime detection is pose estimation, which examines how people move and hold themselves in surveillance footage. Poses that are unusual or suspicious can be marked, indicating possible illegal activity. This feature is especially helpful in crowded public areas when other approaches might not be as effective.

Analyzing sequential data is essential for comprehending how events unfold over time. Recurrent Neural Networks (RNNs) of the Long Short-Term Memory (LSTM) kind perform exceptionally well when handling sequential input. LSTM models are useful in crime detection because they may examine movement, behavior, or interaction patterns across time and help identify suspicious activity.

The combination of LSTMs and Convolutional Neural Networks (ConvNets) improves crime detection much more. LSTMs are experts in temporal dependencies, whereas ConvNets are superior at extracting spatial features from images. When evaluating complicated spatiotemporal patterns in video data, such as tracking persons, identifying deviant behaviors over time, or identifying specific objects, this combination works well.

AI-based crime detection gains further depth with the use of Generative Adversarial Networks (GANs). In order to supplement small datasets used for training, GANs can produce synthetic data, which increases the generalization and robustness of models. This is especially helpful in situations where it is difficult to collect a large and diverse dataset.

In conclusion, by utilizing deep learning, computer vision, and predictive analytics, AI-powered crime detection transforms conventional law enforcement strategies. These developments provide authorities with the ability to better respond to new threats quickly, prevent them from happening in the first place, and build safer, more secure communities. In addition to overcoming conventional constraints, the combination of several AI methodologies puts law enforcement at the forefront of utilizing technology for public safety in the twenty-first century.

2. Related Work

A number of researchers have put all their innovations and ideas for developing such a system. V. Mandalapu et al. offer a comprehensive analysis of over 150 articles on using machine learning and deep learning for crime prediction, showcasing their efficacy in identifying crime patterns and trends [1].

Machine learning algorithms like decision trees and support vector machines, along with deep learning methods such as convolutional and recurrent neural networks, demonstrate their potential in accurately predicting crime patterns and analyzing surveillance video footage for improved real-time monitoring and response.

Jenga K et al. have done a systematic literature review on machine learning in crime prediction, categorizing research objectives and addressing key questions to assist law enforcement. It highlights diverse data sources, including crime records, traffic data, location data, visual data, and text data from social media. It emphasises Artificial Neural Networks as the best-performing algorithm, especially when combined with ensemble methods and boosting parameters for enhanced real-time predictions [2].

P, Ashokkumar et al. suggest the use of an optimized system that integrates advanced machine learning and big data analytics to enhance crime detection, enabling proactive identification and prevention of criminal activities. By facilitating real-time monitoring and predictive analytics, it aims to empower authorities with actionable insights for strategic decision-making, fostering safer communities [3].

Nazir A et al. proposes a novel approach for detecting shoplifting through video analysis, employing a two-stage algorithm that utilizes YOLOv5 with Deep Sort for object detection and time-series classification models like Inception Time and Xception Time for efficient and accurate suspicious behavior detection. It emphasizes the significance of precision and robustness, aiming to outperform existing methods in speed and accuracy [4].

Mahmud, Sakib et al. aims to predict crime rates using machine learning and data mining techniques, employing algorithms such as K-means clustering, linear regression, Naïve Bayes, and K-nearest neighbors for crime analysis and hotspot identification. The dataset, collected through fieldwork, includes key features such as crime types, areas, victim details, and perpetrator information, which are preprocessed and transformed for effective prediction and analysis [5].

Mohammad Reza Keyvanpour et al. highlight the significance of data mining in crime analysis, advocating for the use of artificial intelligence methods to extract knowledge from intricate crime datasets. It outlines entity extraction challenges in narrative reports, introduces a lookup table-based approach, and discusses the application of Self-Organizing Maps (SOM) and AGNES algorithms for crime data clustering. Additionally, the utilization of Multi-Layer Perceptron (MLP) neural networks for crime pattern recognition and the implementation of binary encoding for crime matching are addressed [6].

Shah N et al. talks about the integration of machine learning and computer vision in law enforcement enables proactive crime prediction, efficient data analysis, and automated surveillance, leading to faster response times and improved crime prevention. These technologies aid in identifying suspects, analyzing crime scenes, and predicting future criminal activities based on learned patterns, significantly enhancing law enforcement capabilities [7]. M. Petty et al. have proposed a methodology to detect missing, planted, and displaced objects in pre- and post-contaminated images of the same crime scene. It incorporates ASIFT, an extension of SIFT that is more robust to affine distortion. The methodology involves feature extraction, matching, and registration of the images, followed by the creation of a different image and image segmentation to isolate areas of difference.

The methodology also includes feature matching between cropped object images and evidence images to identify missing or planted objects. Additionally, it estimates the displacement of objects within the scene. The methodology is tested using both SIFT and ASIFT registration techniques [8].

Boukabous et al. highlight the use of YOLOv5 for real-time crime object detection, emphasizing its high speed and accuracy. It stresses the importance of ethical considerations in the deployment of AI for law enforcement, providing a comprehensive methodology for implementing deep learning-based object detection systems in crime prediction [9].

Sreenu G et al. suggest an approach of intelligent video surveillance and its relevance in big data applications, focusing on deep learning techniques for crowd analysis. It highlights the importance of automated systems and discusses various applications such as theft identification and violence detection. The authors delve into deep learning methods, emphasizing their advantages and challenges, and call for further research in real-time processing and accuracy improvement [10].

Jiang, YG et al. explore complex event recognition in multimedia data, focusing on videos. It discusses techniques for recognizing objects, scenes, and human activities, addressing challenges like event localization and scalability. The paper emphasizes the use of low-level features and knowledge-based techniques, along with classification strategies and graphical models. It also highlights the need for better benchmarks and multidisciplinary knowledge from fields like computer vision. The paper concludes by emphasizing the importance of further research in event localization and the integration of prior knowledge [11].

Tzelepis et al. address high-level event detection in videos and propose a framework for learning event detectors from textual descriptions or limited positive video examples. It tackles challenges in bridging the gap between low-level features and semantic-level actions and the scarcity of positive training samples.

The author introduces a learning framework and explores various design choices, demonstrating effectiveness through large-scale dataset experiments. The paper emphasizes the importance of combining textual information and limited video examples for improved event detection, offering valuable insights into overcoming training sample limitations [12].

Roberto Arroyo et al. proposed an intelligent framework for detecting multiple events in surveillance videos. The framework modularizes the surveillance problems into variables comprising regions of interest, classes, attributes, and notions. By leveraging the principle of compositionality, the framework is able to optimize the reasoning of complex events and detect multiple events simultaneously. Experimental results demonstrate the effectiveness and robustness of the proposed framework in detecting multiple events in surveillance videos [13].

Anima Pramanik et al. discusses a conceptual framework for a traffic surveillance system aimed at improving road safety. The framework involves analyzing traffic pre-events, determining suitable locations for CCTV cameras, developing algorithms for modeling traffic pre-events, training and validating these algorithms, and testing the system's operation. The developed algorithms can detect various traffic violations such as speed violations, one-way traffic, overtaking, illegal parking, and wrong drop-off locations of passengers.

The algorithms are trained and validated using video data and are tested on live video streams. The system generates alarms in the control room to take preventive measures in case of potential road accidents. The paper has also mentioned a comparative study conducted to demonstrate the effectiveness of the developed algorithms compared to state-of-the-art algorithms [14].

J. Chen et al. describe a Distributed Intelligent Video Surveillance (DIVS) system that utilizes Deep Learning (DL) algorithms and is deployed in an edge computing environment. The system consists of monitoring terminals, edge nodes, and a cloud server. The authors propose a multi-layer edge computing architecture and a distributed DL training model for the DIVS system. They address the challenges of parallel training, model synchronization, and workload balancing. Task-level and model-level parallel training methods are proposed to accelerate video analysis. A model parameter updating method and a dynamic data migration approach are also proposed to achieve model synchronization and workload balance among edge nodes. Experimental results demonstrate the efficiency and scalability of the DIVS system [15].

S. Sathyadevan et al. suggested that the idea of auto-face detection from surveillance cameras and CCTVs is highly relevant due to the increasing installation of such devices. To simplify face recognition, a system dynamically updates its database with newly detected faces from camera views. The labeling of faces can be done at a later time, if at all.

This system does not rely on a pre-existing image database and instead generates its collection of images, tracking future occurrences of those images. The study done by the authors has explored various face recognition algorithms, including Eigenface, fisherface, LBP histograms, and SURF, with SURF yielding the best results [16].

Figure 1 illustrates the distribution of various types of crimes reported in India from the years 2018 to 2022. The data, sourced from official records, highlights trends and fluctuations in crime rates over the specified period. This visualization serves to provide insight into the prevalence and nature of criminal activities across different regions and time frames, thereby contributing to a comprehensive understanding of law enforcement and public safety challenges in India.

Figure 1 depicts the state-wise distribution of reported crimes against women in India, sourced from data.gov.in. Our project utilizes this data to deepen our understanding of gender-based violence patterns. By pinpointing high-risk areas, our initiative aims to support targeted interventions and community-driven efforts to improve women's safety. Through data-driven insights, we empower stakeholders to make informed decisions in tackling this pressing societal issue.

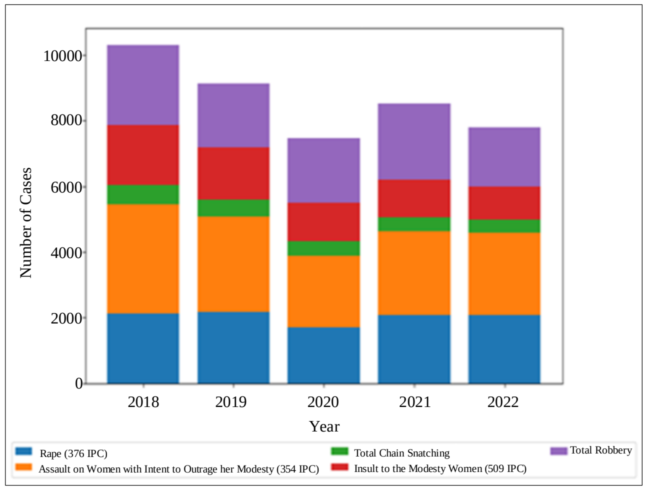

Fig. 1 Crime distribution by year

Table 1. Crime against women

State / Year	2017	2018	2019	2020	2021
Andhra Pradesh	17909	16438	17746	17089	17752
Arunachal Pradesh	337	368	317	281	366
Assam	23082	27687	30025	26352	29046
Bihar	14711	16920	18587	15359	17950
Chhattisgarh	7996	8587	7689	7385	7344
Goa	369	362	329	219	224
Gujarat	8133	8329	8799	8028	7348
Haryana	11370	14326	14683	13000	16658
Himachal Pradesh	1246	1633	1636	1614	1599
Jharkhand	5911	7083	8760	7630	8110
Karnataka	14078	13514	13828	12680	14468
Kerala	11057	10461	11462	10139	13539
Madhya Pradesh	29788	28942	27560	25640	30673
Maharashtra	31979	35497	37144	31954	39526
Manipur	236	271	266	247	302
Meghalaya	567	571	558	568	685
Mizoram	301	249	170	172	176
Nagaland	79	75	43	39	54
Odisha	20098	20274	23183	25489	31352
Punjab	4620	5302	5886	4838	5662
Rajasthan	25993	27866	41550	34535	40738
Sikkim	163	172	125	140	130
Tamil Nadu	5397	5822	5934	6630	8501
Telangana	17521	16027	18394	17791	20865
Tripura	972	907	1070	874	807
Uttar Pradesh	56011	59445	59853	49385	56083
Uttarakhand	1944	2817	2541	2846	3431
West Bengal	30992	30394	29859	36439	35884

The data presented in Table 1 illustrates the incidence of crimes against women in various states of India over the span of five years, from 2017 to 2021. Uttar Pradesh consistently records the highest numbers, with fluctuation. Maharashtra also consistently ranks high, showing an upward trend. Rajasthan saw a spike in 2019. Odisha and West Bengal display fluctuations, while Goa, Mizoram, and Nagaland have lower numbers. Kerala and Tamil Nadu show upward trends.

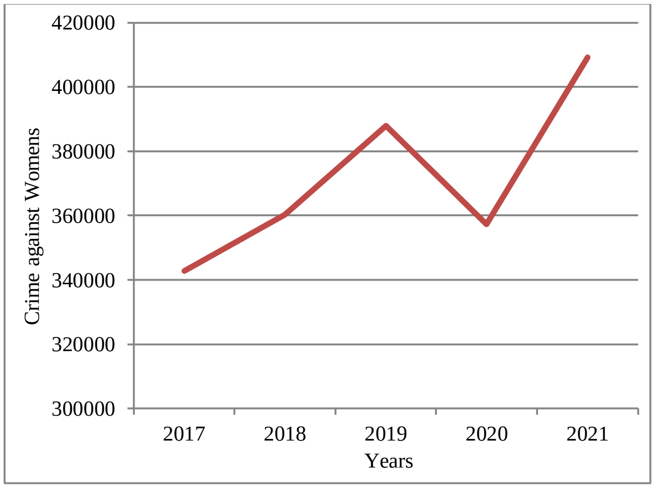

Fig. 2 Average crime against women

Figure 2 illustrates the average number of crimes against women in India from 2017 to 2021. The graph indicates a notable increase from 342,860 to 409,273 cases over the five-year period. The data emphasizes the need for ongoing efforts to address and prevent crimes against women, focusing on understanding underlying factors.

3. Problem Statement, Objectives & Motivation

3.1. Problem Statement

The current landscape of law enforcement faces a critical challenge in swiftly detecting and classifying criminal activities from surveillance footage, hampering timely response and effective crime prevention. Existing Crime Scene Detection Systems (CSDS) often lack real-time analysis capabilities, hindering the prompt identification of criminal acts and impeding law enforcement’s ability to ensure public safety.

Consequently, there is an urgent need for an advanced CSDS that utilizes Artificial Intelligence (AI) techniques, such as Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNNs), to accurately classify crime scenes in real-time and trigger immediate alarms for rapid intervention.

This paper aims to address this challenge by exploring the application of deep learning techniques to analyze complex patterns in crime-related data, facilitating proactive measures by law enforcement agencies to combat criminal behavior and enhance public safety.

3.2. Objectives

To review the current challenges faced by law enforcement agencies in detecting and classifying criminal activities from surveillance footage. To explore the potential of Artificial Intelligence (AI) techniques, specifically Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNNs), in enhancing crime scene detection and classification.

To investigate the effectiveness of deep learning algorithms in analyzing complex patterns within crime-related data for real-time identification of criminal acts. To design and develop an advanced Crime Scene Detection System (CSDS) capable of accurately identifying and classifying various types of criminal activities in surveillance footage.

To evaluate the performance of the proposed CSDS in terms of accuracy, speed, and scalability compared to existing systems. To assess the practical feasibility and potential challenges associated with implementing the proposed CSDS in real-world law enforcement scenarios. To provide recommendations for the integration and optimization of the CSDS within law enforcement agencies to enhance public safety and crime prevention efforts.

4. Methodologies

4.1. Long-Term Recurrent Convolutional Networks

Figure 3 represents a sophisticated deep-learning architecture designed for effectively handling video data. It draws its strength from the fusion of two prominent neural network types: Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. The key innovation in Long-term Recurrent Convolutional Networks (LRCN) is the integration of both spatial and temporal information processing.

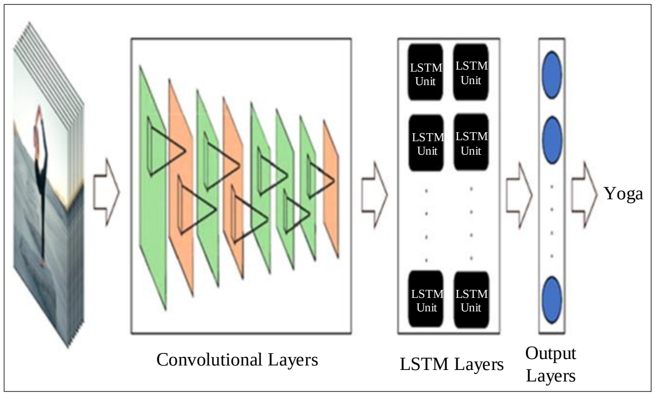

Fig. 3 LRCN architecture

LRCN, the process begins with the CNN component, which is primarily responsible for extracting spatial features from each frame of a video. CNNs excel at capturing patterns and structures within images, making them ideal for recognizing objects, textures, and spatial information within video frames. By applying the CNN to every frame, the model can create a set of spatial feature representations.

However, what makes LRCN truly powerful is its LSTM component. Unlike traditional CNNs, LSTM networks specialize in capturing temporal dependencies and sequential information. In the context of video data, LSTM is utilized to understand how these spatial features evolve over time.

It processes the sequence of spatial feature representations obtained from the CNN across different video frames. This temporal modeling allows LRCN to capture the motion, changes, and interactions between objects within the video, making it suitable for tasks such as action recognition, video captioning, and even anomaly detection.

In summary, LRCN bridges the gap between spatial and temporal understanding of video data. The CNN component extracts spatial information from individual frames, while the LSTM network processes these spatial features over time, enabling a holistic and context-aware analysis of video content. This architecture is particularly valuable in applications where both spatial and temporal aspects are crucial for accurate interpretation, and it has found use in various fields, including computer vision, surveillance, and video analytics.

The fusion of CNN and LSTM in the LRCN architecture exemplifies the synergy between image processing and sequential data analysis. This dynamic combination is instrumental in various video-related tasks, from recognizing activities in surveillance footage to generating descriptive captions for video content.

LRCN’s ability to bridge the gap between spatial and temporal understanding has paved the way for more accurate, context-aware, and efficient video data analysis, making it a valuable tool in advancing the field of computer vision and enhancing the capabilities of applications such as video content indexing, action prediction, and real-time event detection.

4.2. CNN-LSTM

Figure 4 illustrates a robust deep learning architecture known as CNN-LSTM, which combines Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to process sequential data, such as videos effectively. This fusion of CNNs and LSTMs enables the model to simultaneously capture spatial features and temporal dependencies within the video data, allowing for comprehensive understanding and analysis.

The CNN component of the CNN-LSTM architecture serves as the initial stage, where spatial features are extracted from each frame of the video. CNNs excel at identifying patterns, objects, and spatial relationships within images, making them well-suited for processing video frames. By applying CNNs to the individual frames, the model generates a set of spatial feature representations that capture important visual information.

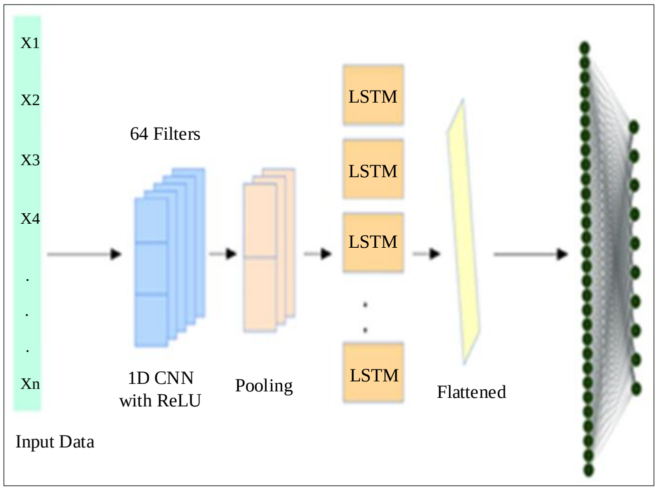

Fig. 4 CNN-LSTM architecture

However, what sets CNN-LSTM apart is its integration of LSTM networks, which specialize in capturing temporal dependencies and sequential information. In the context of video data, LSTM networks are crucial for understanding how spatial features evolve over time. The LSTM component processes the sequence of spatial feature representations obtained from the CNN across different video frames, allowing the model to capture the dynamics, motion, and interactions within the video content.

By integrating both CNNs and LSTMs, CNN-LSTM bridges the gap between spatial and temporal understanding in video data. The CNN component focuses on extracting spatial information from individual frames, while the LSTM network models the temporal evolution of these spatial features. This holistic approach enables the model to perform context-aware analysis of video content, making it suitable for tasks such as action recognition, video captioning, and anomaly detection.

In summary, CNN-LSTM leverages the strengths of both CNNs and LSTMs to process sequential data effectively. The CNN component extracts spatial features, while the LSTM component models temporal dependencies, enabling comprehensive analysis of video data. This architecture has applications in various fields, including computer vision, surveillance, and video analytics, where accurate interpretation of both spatial and temporal aspects is essential for task performance. The fusion of CNNs and LSTMs in CNN-LSTM exemplifies the synergy between image processing and sequential data analysis, facilitating advancements in video-related tasks and enhancing the capabilities of applications such as video content indexing, action prediction, and real-time event detection.

4.3. Data Collection and Preprocessing

The model was trained on a dataset comprising 150 videos, each averaging five seconds in duration, designed to classify scenes into four distinct classes. To ensure consistency and optimize computational efficiency, images extracted from these videos were resized to 244x244 pixels and normalized by dividing pixel values by 255, thereby standardizing inputs between 0 and 1.

In this study, YouTube videos were utilized as a primary source for data collection, contributing to the establishment of our dataset. The videos were systematically curated based on predefined criteria relevant to our research objectives. This approach allowed for the inclusion of real-world examples and perspectives, enriching the dataset with varied and authentic data sources. Leveraging YouTube as a data repository facilitated access to a vast array of multimedia content, providing valuable insights and enhancing the depth of analysis in our research.

5. Results and Discussion

Fig. 5 LRCN result

Figure 5 captures a chain-snatching incident from surveillance footage, showing a perpetrator snatching a chain from a pedestrian while riding a motorcycle. This still frame provides data for spatial and temporal feature extraction using the LRCN architecture, which combines CNNs and RNNs to analyze such scenes.

The LRCN model identifies patterns like the motorcycle's approach, the chain snatch, and the getaway, facilitating real-time alerts and post-event analysis for crime prevention and investigation. The LRCN model, in tandem with a CCTV camera, recognizes chain-snatching incidents by first passing the preprocessed image through a CNN component.

This extracts high-level spatial features crucial for tasks like object detection. Subsequently, the output is fed into an LSTM network, which learns temporal dependencies in sequential data. This enables the model to understand the chronological order of events in the video frames, such as reaching for the chain, grabbing it, and pulling it away. This process illustrates how the LRCN model effectively recognizes complex scenes like chain-snatching incidents in CCTV footage. Training the model with diverse and representative data is paramount to its accuracy and generalizability in real-world applications.

Fig. 6 CNN-LSTM result

Table 1. Performance of different models

Deep Learning Model	Training Dataset Size	Accuracy
CNN-LRCN	600 Videos	84.0%
LRCN	600 Videos	94.0%

Figure 6 captures a chain-snatching incident, emphasizing the swift motorcycle-based theft from a pedestrian. The victim's surprised expression and motion blur highlight the abruptness of the crime. In utilizing the CNN-LSTM architecture for crime detection, this image is pivotal for extracting both spatial and temporal features. The CNN-LSTM model, combining Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, thoroughly analyzes frames to identify crucial patterns. CNNs extract spatial features, recognizing elements like the approaching motorcycle and the act of chain snatching.

Simultaneously, LSTM networks capture temporal dependencies, discerning the sequential progression of events, including the perpetrator's approach, the snatch, and the subsequent getaway. Table 1 presents the performance analysis of two implemented models: CNN-LRCN and Long-term Recurrent Convolutional Networks (LRCN). The LRCN model demonstrates 94% accuracy in correctly detecting crime events, surpassing the CNN-LRCN model, which achieves an 84% accuracy. Both models were trained on a dataset consisting of 600 videos.

6. Conclusion

Both the Long-term Recurrent Convolutional Networks (LRCN) and CNN-LSTM architectures represent innovative hybrids, combining the power of Convolutional Neural Networks (CNNs) with Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks. Recent advancements in LSTM and CNN technologies have significantly bolstered crime detection efficacy, driving forward a dynamic research landscape aimed at fostering robust and effective systems for crime prevention and intervention. While the CNN-LSTM hybrid model showcases a respectable accuracy rate of 84%, it is the LRCN model that truly shines. By leveraging CNNs for spatial features and LSTMs for temporal sequence processing, the LRCN model is uniquely poised to tackle tasks demanding both spatial and temporal analysis, such as video understanding and action recognition. Demonstrating its superiority, the LRCN model achieves an outstanding accuracy of 94%, solidifying its position as a frontrunner in the realm of crime detection and analysis.

References

[1] Varun Mandalapu et al., "Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions," *IEEE Access*, vol. 11, pp. 60153-60170, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Karabo Jenga, Cagatay Catal, and Gorkem Kar, "Machine Learning in Crime Prediction," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 2887–2913, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Ashokkumar Palanivinayagam et al., "An Optimized Machine Learning and Big Data Approach to Crime Detection," *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1-10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Amril Nazir et al., "Suspicious Behavior Detection with Temporal Feature Extraction and Time-Series Classification for Shoplifting Crime Prevention," *Sensors*, vol. 23, no. 13, pp. 1-19, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Sakib Mahmud, Musfika Nuha, and Abdus Sattar, "Crime Rate Prediction Using Machine Learning and Data Mining," *Soft Computing Techniques and Applications*, vol. 1248, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Mohammad Reza Keyvanpour, Mostafa Javideh, and Mohammad Reza Ebrahimi, "Detecting and Investigating Crime by Means of Data Mining: A General Crime Matching Framework," *Procedia Computer Science*, vol. 3, pp. 872-880, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[7] Neil Shah, Nandish Bhagat, and Manan Shah, "Crime Forecasting: A Machine Learning and Computer Vision Approach to Crime Prediction and Prevention," *Visual Computing for Industry, Biomedicine, and Art*, vol. 4, pp. 1-14, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] Mark Petty, Shyh Wei Teng, and Manzur Murshed, "Improved Image Analysis Methodology for Detecting Changes in Evidence Positioning at Crime Scenes," *2019 Digital Image Computing: Techniques and Applications (DICTA)*, Perth, Australia, pp. 1-8, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Mohammed Boukabous, and Mostafa Azizi, "Image and Video-Based Crime Prediction Using Object Detection and Deep Learning," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 3, pp. 1630-1638, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] G. Sreenu, and M.A. Saleem Durai, "Intelligent Video Surveillance: A Review through Deep Learning Techniques for Crowd Analysis," *Journal of Big Data*, vol. 6, pp. 1-27, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[11] Yu-Gang Jiang et al., "High-Level Event Recognition in Unconstrained Videos," *International Journal of Multimedia Information Retrieval*, vol. 2, pp. 73-101, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[12] Christos Tzelepis et al., "Learning to Detect Video Events from Zero or Very Few Video Examples," *Image and Vision Computing*, vol. 53, pp. 35-44, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[13] Roberto Arroyo et al., "Expert Video-Surveillance System for Real-Time Detection of Suspicious Behaviors in Shopping Malls," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7991-8005, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[14] Anima Pramanik, Sobhan Sarkar, and J. Maiti, "A Real-Time Video Surveillance System for Traffic Pre-Events Detection," *Accident Analysis & Prevention*, vol. 154, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Jianguo Chen et al., "Distributed Deep Learning Model for Intelligent Video Surveillance Systems with Edge Computing," *IEEE Transactions on Industrial Informatics*, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[16] Shiju Sathyadevan et al., "Identifying Moving Bodies from CCTV Videos Using Machine Learning Techniques," *2014 First International Conference on Networks & Soft Computing (ICNSC2014)*, India, pp. 151-157, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[17] Year-wise Details of the Number of Cases Reported against Women: Harassment, Rape, Chain Snatching, Theft, and Robbery from 2018 to 2022, A Digital India Initiative, 2023. [Online]. Available: https://data.gov.in/resource/year-wise-details-number-cases-reported-against-women-harassment-rape-chain-snatching

[18] State/UT-wise Details of Cases Registered under Crime against Women in the Country from 2019 to 2021, A Digital India Initiative, 2023. [Online]. Available: https://data.gov.in/resource/stateut-wise-details-cases-registered-under-crime-against-women-country-2019-2021