IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Violence Detection In Real Life Videos Using Pre-Trained Models

Gangu Basavaraj,Shilpa B.Kodli

Post Graduate Student, Assistant Professor
Department of Computer Science and Engineering (MCA),
Visvesvaraya Technological University, Kalaburagi, India

Abstract: Real-world video violence identification is a critical task with broad implications for guaranteeing security, content control, and public safety. In order to reliably identify violent scenes in real-time videos, this study suggests an automated approach for violence identification utilizing pre-trained models. Modern pre-trained convolutional neural networks (CNNs) for visual processing are combined in the system. The system provides robustness and higher accuracy in detecting violent content by combining the data from the two modalities. The progress is efficient enough to enable real-time processing, allowing quick analysis of video streams and quick detection of violent situations. The progress alerts the appropriate authorities or content moderators when it detects violent behaviour. To ensure prompt action, a timing and location of the identified violent incident are included in an email message sent to predetermined recipients

Index Terms **- Violence Identification, Deep Learning , convolutional neural networks, OpenCV, Firebase, JSON.**

I. INTRODUCTION

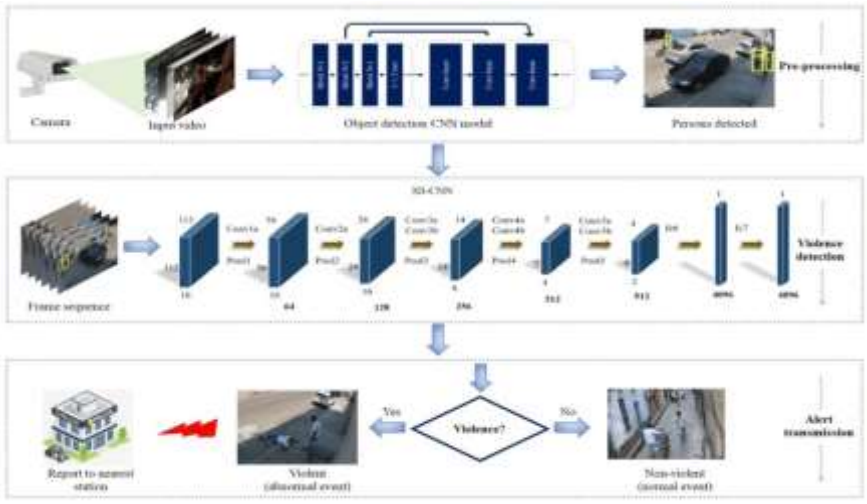
In today's digital era, the prevalence of video content across various online platforms has grown significantly, bringing attention to the need for efficient and automated violence detection systems. To secure user security, safeguard online communities, and help law enforcement authorities quickly respond to bad content, violence identification in real-world recordings is an essential duty. Researchers have looked to pre-trained models as a possible approach to tackle this problem, relying on their capacity to generalize from in-depth training on substantial datasets.

Public video observation systems are widely utilized worldwide and can give accurate and detailed information in various security applications [1]. Although the need to review hours of video material compromises the capacity to make decisions fast, which is essential in video surveillance for the prevention of crime and violence [2]. In sequence to relieve authorities of the burden of having to view hours of movies in sequence to identify events lasting only a few seconds, a number of research about the automatic detection of violent sequences in videos have been presented. Recent research has demonstrated the accuracy of deep learning-based algorithms in the identification of violence, in contrast to earlier studies that used hand-crafted features and flow descriptors typical of classical action recognition methods [3]. Due to their outstanding performance in image identification tasks, Convolutional Neural Networks (CNNs) are more typically utilized in computer vision [4]. In several research areas, CNNs are developing quickly, and that's anticipated that next solutions will accelerate the implementation of CNNs. These learning algorithms still have lots of room for progress given the accessibility of massive data and the rapid increase in computing power. The spatial CNNs, They are used for images recognition tasks, have recently been extended to the temporal domain for HAR in movies using a ways of effective methods [5].

Currently, an increasing amount of Internet traffic comes from video content. Short-form video services and live broadcasts have emerged as the newest Internet development trends. However, there are a lot of violent movies online that seriously jeopardize the growth of the network ecology. Additionally, it is extremely difficult to monitor spontaneous acts of violence in time. The result is , violent video detection is crucial. The term "violent video detection" often refers to the process of identifying violent acts and violent part in videos by first extracting their visual and acoustic data, then categorizing them using a linear classifier. Video

Violence Detection In Real Life Videos Using Pre-Trained Models Department of CSE (MCA), VTU's Center for PG Studies, Kalaburagi 2 action recognition can be noticed as a particular type of violent video detection in various aspects. [6].

Deep learning architectures for pre-trained models are learned on enormous datasets for prevalent CV duties like image or object recognition. To be able to capture crucial patterns and representations helpful for diverse visual identification tasks, they learn to extract high-level features from visual data. Researchers can efficiently detect violent events in real-world videos by modifying these pre-trained models using violence detection datasets. Pre-trained models and cutting-edge architectures, including 3D Convolutional Neural Networks (3DCNN), have attracted attention in recent years for violence detection applications. In order to detect motion patterns and temporal connections in films, 3DCNNs take considering both spatial and temporal dimensions. The violence detection algorithm can now distinguish between violent and non-violent activities with more accuracy and resilience thanks to this improvement [7].



In sequence to ensure public safety and content filtering, it is becoming increasingly difficult due to the prevalence of violent content in real-world videos on different digital platforms. We want to solve this problem by creating a deep learning system that spontaneously detects violent acts. putting in spot a violence detection system that is accurate and efficient, able to scan real-world films and distinguish between violent and non-violent content. Including a mail notification system that, when it identifies violence in a video, delivers quick alerts to a list of recipients. The aim of this work is to examine and develop a real-world film-based violence detection system. The study's objectives are as follows Create and improve already-trained models: The project will concentrate on choosing appropriate pre-trained models for violence detection, such as two-stream CNNs or 3D CNNs. Using annotated datasets of violent and nonviolent videos, these models will be refined and tailored to the unique properties of real-life settings. Improve accuracy and robustness: The study will look into ways to make violence detection models more accurate and resilient. To manage differences in video quality, lighting conditions, and scene complexity, this may involve investigating new model architectures, data augmentation approaches, and regularization techniques. Real-time processing: By enhancing the violence detection models for low latency, the study will address the problem of real-time processing. To provide prompt detection of violent actions in streaming or saved movies, effective model architectures, parallel processing strategies, and hardware acceleration techniques will be investigated. Evaluation of the model's performance will be done using extensive evaluation measures, which will be used to gauge how well the violence detection system is working. Metrics including precision, recall, F1 score, and accuracy in the study will be used to gauge how well the system performs in correctly classifying violent occurrences and minimizing false positives or false negatives. Identify ethical concerns: The study will consider the moral implications of bias and privacy in violence detection. We will investigate privacy-preserving methods to make sure the system respects peoples' private rights. Additionally, the models will undergo a thorough evaluation in order to spot and correct any potential biases and discrimination in the system's automated decision-making. Scalable deployment: To handle high amounts of movies effectively, the study will look into scalable infrastructure and distributed computing techniques.

II. Related Work

According to [8] and others investigated due to rising crime rates, a machine learning framework for automatically identifying and categorizing violence in video streams. It employs Inception-v3 and Gated Recurrent Units (GRUs) to identify violent acts in binary and multi-class scenarios and 3D Convolutional Neural Networks for initial detection. Transfer learning is used to enhance performance while preserving computational resources. The models are trained using a variety of datasets, including surveillance footage, human recordings, movies, and YouTube videos. As the [9] suggested a importance of automating the detection of violent behavior in surveillance systems. It draws attention to the drawbacks of manual labor-intensive and education-based techniques, particularly for videos with jerky camera motion. The suggested method outperforms current state-of-the-art algorithms in identifying aggressive human actions in violent sequences, attaining extraordinary accuracies of 99.28% and 99.97% on the Hockey and Movies datasets, respectively. According to [10] addresses the difficulty of recognizing violence, especially in surveillance systems where human oversight is prone to error. It presents attention-based LSTM and BiLSTM-based solutions and makes use of a distinctive database derived from security cameras and publicly available films (like those on YouTube and Facebook). The BiLSTM model outperforms the LSTM model in identifying battle situations, according to extensive testing on datasets like Hockey Fight and crowd scenes. As stated by the [11] and others suggested on reducing human monitoring in surveillance through real-time violence detection using deep learning. It combines DeepSort for effectiveness, YOLO for object identification, and Long Short-Term Memory for categorization. The system was tested and evaluated as a functioning violence detection solution after being trained on a dataset for violence detection and integrated into software frameworks. According to [12] addresses reportedly discusses the expanding need for automated technologies to identify violent and suspicious activities in public settings. It discusses cutting-edge methods for detecting violence based on convolutional neural networks (CNNs). The paper discusses many CNN variations, their uses, and the critical contribution of datasets to improving the precision of violence detection techniques. Based on the author view of [13] proposes a real-time violence detection system using deep learning on UAVs to address violence prevention. It ensures high-speed calculations, model accuracy, and compact size, achieving a 93.69% mAP with low computational cost and the ability to operate at 21-22 FPS for efficient monitoring and object tracking. According to [14] and others focuses on violence detection using pre-trained deep neural networks, aiming to simplify the process. Features extracted from models like ResNet-50 and VGG16 are fed into a fully connected network to identify violent actions. The method is evaluated on four public datasets, demonstrating its efficiency compared to more time-consuming recurrent network approaches. As the [15] introduces a Using two-stream networks, 3D CNNs, and Convolutional LSTM networks, a violence detection model specifically designed for video surveillance has been developed. It uses YOLO for human target detection and local feature extraction, CA attention to improve spatial and temporal feature extraction, and ConvLSTM to efficiently capture video timing relationships in place of several 3D convolutional layers. The model's excellent accuracy in detecting violence is demonstrated through evaluation on the RWF-2000 dataset, making it useful for surveillance and public safety reasons. According to [16] and others investigated a 3D convolutional neural network-based deep learning model for automated violence detection in surveillance videos. It uses effective bottleneck units and DenseNet architecture for motion pattern learning and feature improvement instead of manually created features and RNNs. The model is suited for applications like internet video filtering and public security protection because to its superior recognition accuracy and computational efficiency, which are demonstrated by several experiments. As the [17] and others also Due to the problem of data volume, emphasis is placed on automating violence detection in surveillance videos. It makes use of action recognition models that have already undergone training to gather spatial and temporal information using 3D convolutional networks. Evaluations across a variety of datasets show a 2% boost in accuracy over current approaches, and the method also shows to be robust in addressing common compression aberrations, notably in applications involving remote server processing.. According to [18] response to the The use of behavioral analysis in intelligent monitoring is becoming more common in order to create safer cities. With little attention paid to complicated behaviors like fighting or violence, existing approaches mostly concentrate on detecting simple activities like walking or jogging. By presenting the most recent techniques in this field, this paper tackles the pressing need for accurate violence detection in video surveillance, which is critical for maintaining social order and protecting people and property.

III. PROPOSED METHODOLOGY

When seen from a system viewpoint, violence detection in real-world films using trained models entails a number of parts and processes cooperating to get the intended result. Let's look at this solution from the standpoint of the system.

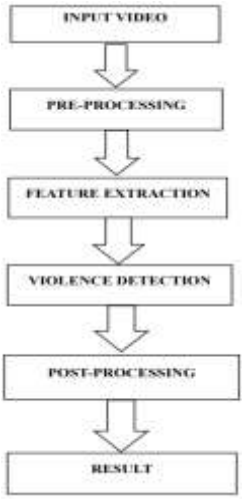


Figure 1 : Proposed System Architecture

Data Ingestion system obtains video data from a number of sources, such as web services, video security cameras, and recorded material. It must be capable of capturing videos in a range of formats and resolutions. Pre-processing incoming video data may need to be handled in order to manage compression, standardize formats, or normalize the video information. In this step, the data are prepared for additional analysis. The system makes use of trained models to retrieve relevant information from the video frames. These traits capture violent behavior as visual patterns, motion information, or aural cues. Violence Detection obtained features are included into algorithms that, through data analysis, pinpoint violent scenes. Machine learning techniques must be employed to compare the attributes to taught thresholds and patterns in order to accomplish this. Post-processing system employs post-processing techniques to enhance the outcomes of its algorithm for detecting violence. This can require getting rid of false positives, gathering detections over time, or utilizing contextual information.

IV. DATA STORAGE

Google created Firebase, a broad platform that offers a variety of cloud-based tools and services to aid developers in more effectively creating and maintaining online and mobile apps. For the creation of apps, Firebase provides a variety of capabilities, including as a real-time database, authentication, cloud services, hosting, storage, and more. You must establish a Firebase project on the Firebase Console in order to begin using Firebase. Once your project is configured, you can add your web or mobile app by adding it to the project. To establish a connection between your app and the Firebase services, Firebase provides configuration files that you insert in the code of your project. You can manage user authentication in your app with Firebase Authentication. It supports a variety of providers of authentication, including Google Sign-In, email and password, and others. User data security for your app is made simple with Firebase, which takes care of user registration, login, password resets, and identity verification. Realtime Database and Firestore are two NoSQL databases offered by Firebase for storing and syncing data in real-time. You may store and retrieve data, including user profiles, chat messages, and more, using these JSON-based databases. Changes to the data are quickly reflected across all linked devices because to Firebase's real-time capabilities.

Img 1:Firestore Database

Img 2: Storage

For the files in your app, including pictures, videos, and other media assets, Firebase Storage offers cloud storage. It makes it simple to securely upload, download, and serve files to users. Overall, Firebase offers a suite of cloud-based services that can be utilized alone or collectively to build, scale, and maintain your web or mobile applications, which simplifies many elements of app development.

V. Results and Discussion

The system is designed to detect the Violence in the real world. It uses the Convolutional Neural Networks (CNN) to detect the violence in the real world in both the real time as well as in the recorded video. The system Detects the violence and displays the True message if the violence is detected else False.

Img 3: Violence Detection is True

Img 4 :Violence Detection is False

Conclusion

Implementing pre-trained modules with mail notification to detect violence in real-world videos is a powerful way to use machine learning for content control and public safety. The method detects violent incidents in real-time movies with efficiency by using pre-trained convolutional neural networks (CNNs). The system's resilience and dependability are increased by the combination of visual and aural elements, which makes it possible for it to precisely identify violent content. The system's functionality is enhanced by the new capacity to send mail notifications when violence is detected. A safer online environment is maintained by using this real-time alert system to make sure that relevant authorities or content moderators are swiftly alerted of any potential violent situations. In terms of establishing a more responsible and secure online environment, violence detection with mail notification utilizing pre-trained modules has considerable promise. We can promote a safer online environment and help to create a more supportive and respectful online community by constantly developing such technology

References

[1] Z. Xu, C. Hu, and L. Mei, Multimedia, "Video structured description technology based intelligence analysis of surveillance videos for public security applications." Tools Appl., vol. 75, no. 19, pp. 12155–12172, 2016

[2] A. Castillo, S. Tabik, F. Pérez, R. Olmos, and F. Herrera, Neurocomputing, Brightness-guided preprocessing for automatic cold steel weapon detection in surveillance films with deep learning, vol. 330, pp. 151–161, Feb. 2019.

[3] T. Hassner, Y. Itcher and O. Kliper-Gross, "Violent flows: Real-time detection of violent crowd behavior", Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, pp. 1-6, Jun. 2012.

[4] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar and L. Fei-Fei, "Large-scale video classification with convolutional neural networks", Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1725-1732, Jun. 2014.

[5] U. Buchler, B. Brattoli and B. Ommer, "Improving spatiotemporal self-supervision by deep reinforcement learning", Proc. Eur. Conf. Comput. Vis. (ECCV), pp. 770-786, 2018.

[6] Wei Song, Dongliang Zhang," A Novel Violent Video Detection Scheme Based on Modified 3D Convolutional Neural Networks" 19 March 2019, Electronic ISSN: 2169-3536

[7] Sultani, W., Chen, C., & Shah, M. (2018). Real-World Anomaly Detection in Surveillance Videos. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), INSPEC Accession Number: 18325997.

[8] Maria Gadelkarim, Mazen Khodier, Violence Detection and Recognition from Diverse Video Sources, 2022 International Joint Conference on Neural Networks (IJCNN)DOI: 10.1109/IJCNN55064.202218-23 July 2022

[9] Aqib Mumtaz, Allah Bux Sargano, Violence Detection in Surveillance Videos with Deep Network Using Transfer Learning,018 2nd European Conference on Electrical Engineering and Computer Science (EECS)DOI: 10.1109/EECS45162.2018 20-22 Dec. 2018

[10] Himanshu Gupta, Syed Taqi Ali, Violence Detection using Deep Learning Techniques, 2022 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)DOI: 10.1109/ICETCI55171.202225-27 Aug. 2022

[11] Gul e. Fatima Kiani, Taheena Kayani, Real-time Violence Detection using Deep Learning Techniques, 2022 3rd International Conference on Innovations in Computer Science & Software Engineering (ICONICS)DOI: 10.1109/ICONICS56716.202214-15 Dec. 2022

[12] Aayush Jain, Dinesh Kumar Vishwakarma, State-of-the-arts Violence Detection using ConvNets, 2020 International Conference on Communication and Signal Processing (ICCSP)DOI: 10.1109/ICCSP48568.202028-30 July 2020

[13] Huy Hoang Nguyen, Quoc Trung Le, A novel violence detection for drone surveillance system, 2023 International Conference on Communication, Circuits, and Systems (IC3S)DOI: 10.1109/IC3S57698.2023
26-28 May 2023

[14] Narges Honarjoo, Ali Abdari, Violence detection using pre-trained models, 2021 5th International Conference on Pattern Recognition and Image Analysis (IPRIA)DOI: 10.1109/IPRIA53572.2021 28-29 April 2021.

[15] Chenglong Pan, Shumin Fei, Violence detection based on attention mechanism, 2022 41st Chinese Control Conference (CCC)DOI: 10.23919/CCC55666.202225-27 July 2022

[16] Ji Li, Xinghao Jiang, Efficient Violence Detection Using 3D Convolutional Neural Networks, 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)DOI: 10.1109/AVSS44038.201918-21 Sept. 2019

[17] Minglan Su, Chaoying Zhang, Deep Learning in Video Violence Detection, 2021 International Conference on Computer Technology and Media Convergence Design (CTMCD)DOI: 10.1109/CTMCD53128.202123-25 April 2021