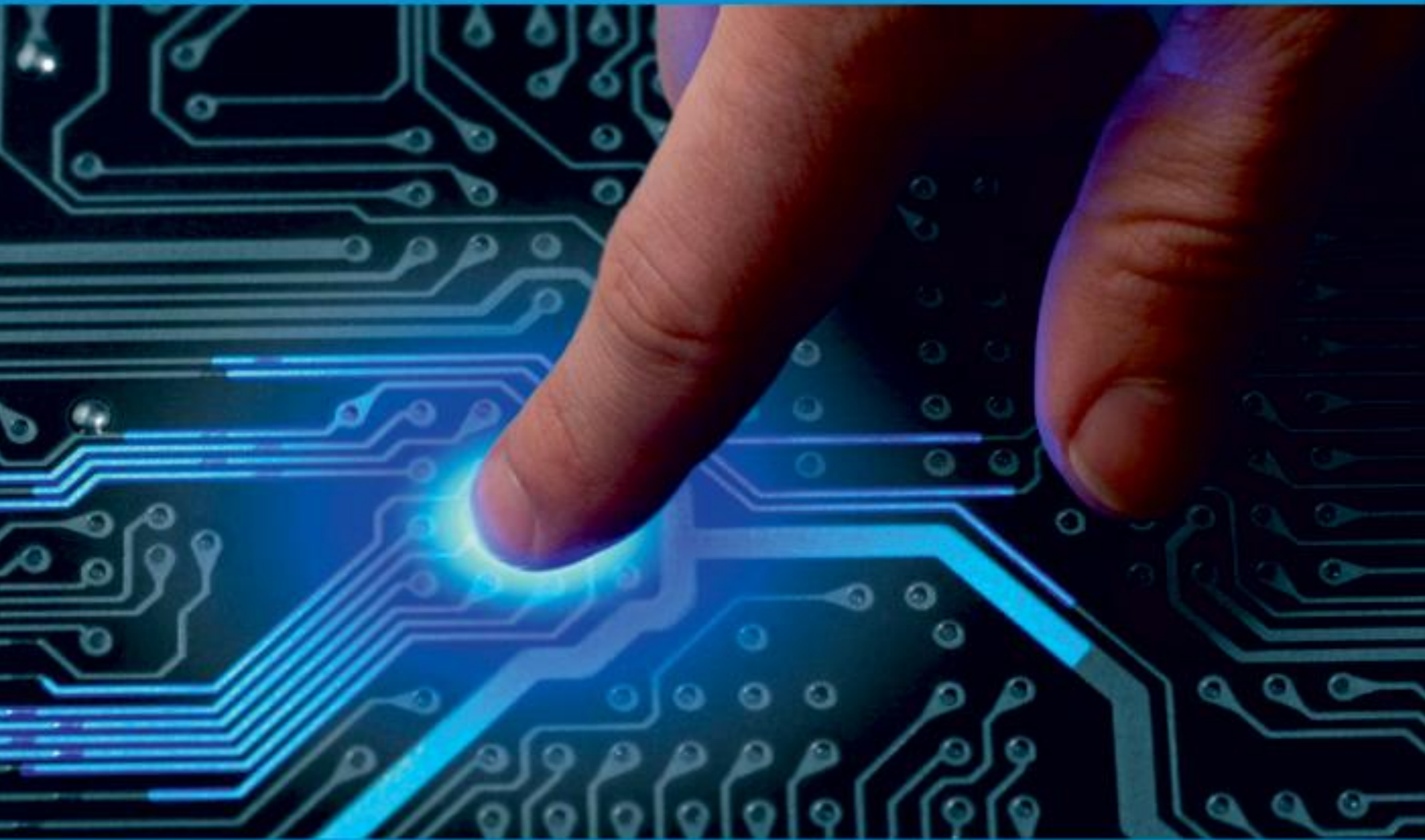
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 4, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

9940 572 462 6381 907 438 ijircce@gmail.com www.ijircce.com

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

DOI: 10.15680/IJIRCCE.2024.1204184 |

A Comprehensive Survey on Exploring Deep Learning Models for Violence Detection

Nadia Anjum, Dendi Varsha, Samudrala Shalini, Prabhavati

Assistant Professor, Department of Computer Engineering, Stanley College of Engineering, Hyderabad, India

U.G Student, Department of Computer Engineering, Stanley College of Engineering, Hyderbad, India

U.G Student, Department of Computer Engineering, Stanley College of Engineering, Hyderbad, India

U.G Student, Department of Computer Engineering, Stanley College of Engineering, Hyderbad, India

ABSTRACT: The ability to detect and flag violent content in real-time is essential for preventing crimes, identifying potential threats, and protecting vulnerable individuals. This task presents significant challenges, as violence can take many forms, and distinguishing between violent and non-violent content requires sophisticated analysis of visual and auditory cues. This paper focuses on violence detection in videos using multiple deep-learning models, including ResNet50v2, DenseNet 121, MobileNetV2, VGG16, I3D, and C3D. The objective is to perform a comparative analysis of the model's performance on various datasets, including RLVS, bus violence, hockey violence, movie violence, fight surveillance violence, and RWF2000. The study aims to understand how different deep learning models produce results for different datasets, with the ultimate goal of understanding the diversity of the data and their performance over light weight models

I. INTRODUCTION

Exploring deep learning models for violence detection involves the development and application of advanced artificial intelligence (AI) algorithms to identify, analyze, and predict violent behaviors or actions in various contexts. This field leverages the computational power of deep learning—a subset of machine learning—to process and interpret complex data, such as images, videos, text, and audio, in a way that can recognize patterns indicative of violence. The goal is to create automated systems that can assist in early warning, real-time surveillance, and post-event analysis to enhance safety and security measures.
Deeplearning models, such as Convolutional Neural Networks (CNNs), Recurrent Neural ***prabhavati***

Networks (RNNs), and their variations,are particularly suited for this task due to their ability to learn hierarchical representations of data, making them effective in handling the spatial and temporal complexities present in violent actions. For instance, in video surveillance, CNNs can be Life Violence Situations (RLVS), instances of violence on buses, violent episodes in hockey, scenes depicting violence in movies, surveillance footage of physical altercations, and the RWF2000 dataset.

used to detect physical movements that match violent behaviors, while RNNs can analyze sequential data to identify patterns over time.Applications of violence detection using deep learning models include, but are not limited to:

1. **Public Safety and Surveillance:** Implementing real-time monitoring systems in public spaces, schools, or commercial areas to detect and prevent violent incidents.
2. **Social Media and Online Content Moderation:** Automatically identifying and filtering out violent content or behavior on digital platforms to maintain community standards and user safety.
3. **Crime Prevention and Law Enforcement:** Assisting police and security agencies in identifying potential threats and responding more effectively to incidents.
4. **Domestic Violence and Bullying Detection:** Analyzing audio, video, or textual communications for signs of abusive behavior or aggression.

Developing effective violence detection systems involves addressing several challenges, including the ethical considerations of privacy and surveillance, the technical hurdles of achieving high accuracy and low false positive rates, and the need for large, annotated datasets for training and testing the models.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

DOI: 10.15680/IJIRCCE.2024.1204184 |

Research in this area is ongoing, with scientists and engineers continuously working to improve the performance, reliability, and applicability of deep learning models for violence detection, aiming to contribute positively to societal safety and well-being.

The overarching objective of this investigation is to unravel the intricate relationship between different deep learning models and their efficacy in violence detection across various datasets.

Through a rigorous examination of these models, the study seeks to elucidate their strengths aweaknesses, providing insights into their adaptability and robustness across diverse scenarios. By scrutinizing their performance on datasets of varying complexities, including real- life situations and staged movie violence, the research aims to shed light on the models' capacity to discern nuanced patterns and adapt to the diverse nature of violent content. Ultimately, this exploration is poised to contribute not only to the refinement of violence detection methodologies but also to the broader understanding of how different deep learning models respond to the multifaceted challenges posed by real-world scenarios.

Image Features

Color is the most popular visual feature of the image. Human eyes are very receptive to color images than grey level images. RGB color space uses Red, Green and Blue color intensities. There are various color feature extraction methods like color histogram, auto- correlogram, color moments etc. Color histogram is used to signify the color stuffing of the image. The image histogram displays the range of gray levels 0-255. To reduce this wide range of value, the color range is divided into bins. Color moments compute the color allocation in an image. If the color allocation in the two images is same then the two images are supposed to be a like.
Color correlogram is a technique that evaluates the distribution of color pair as a function of the space connecting pixel pair. The method describes spatial correlation of duo of colors alters with space. It calculates spatial correlation between like colors.

Shape Feature

Shape feature of image, determine the edge determination in an image. Shape feature categorized into two groups:

a. **Exterior border of the shape.**
b. **Region of the shape.**

To evaluate the above two categories Fourier descriptor and Moment invariant methods are used.

Texture Feature

Texture feature of image, determine regular replication of a component or pattern on a surface. Texture feature depicts the structural collection of plane and their association with the neighboring environment. The texture feature can be taken out using GLCM (Grey level co-occurrence matrix), Wavelets, Fourier transform, entropy, correlation methods. Gabor and wavelet transform extracts statistical distribution of image. The six texture properties are coarseness, contrast, roughness, regularity, directionality and line likeness.

CBIR Benchmark Dataset

Different types of dataset are available online to work with the imaging system. These dataset contain wide variety of images from Tribal, Transport to Buildings. Some of the datasets, we have illustrated in this section. Total eleven datasets are discussed in this paper [22].

Table 1: Dataset Description

Type of dataset	**Description**
Wang Dataset	Used in feature extraction

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

DOI: 10.15680/IJIRCCE.2024.1204184 |

Corel Dataset	Thousand of images from various categories. Each category has 100 images of JPEG format.
GHIM Dataset	Contains only 20 categories. Each category has 500 images of JPEG format.
CIFAR-10 Dataset	Contains images from 10 classes.
UW database	There are 18 categories, for example "spring flowers"
ZuBuD database	It contains images of buildings from different angles in different time of day and night.

II. LITERATURE SURVEY

• The paper titled "Violent interaction detection in video based on deep learning" by Zhou et al. (2017) explores the application of deep learning techniques for detecting violent interactions in videos. The study focuses on utilizing advanced neural network architectures to automatically identify and classify violent activities within video content. The authors likely discuss the methodology, results, and implications of employing deep learning for video-based violent interaction detection in this research.

The paper "Deep Learning for Multi-class Identification from Domestic Violence Online Posts" by Subramani et al. (2019) investigates the use of deep learning techniques for themulti-class identification of domestic violence content in online posts. The study likely delves into the development and application of neural network models to analyze textual data, aiming to enhance the automatic detection of various forms of domestic violence expressed in online communication. Insights from this research may contribute to the advancement of tools addressing the identification and monitoring of domestic violence-related content on digital platforms.

The paper by Wu et al. (2020) titled "Not Only Look, But Also Listen: Learning Multimodal Violence Detection Under Weak Supervision" investigates the integration of visual and auditory information for violence detection. The study emphasizes the utilization of weak supervision techniques to train models capable of identifying violent content across multiple modalities. This research contributes to advancing multimodal approaches in computer vision and audio analysis, aiming to enhance violence detection systems' accuracy and robustness.

The paper by Durães et al. (2023) titled "Violence Detection in Audio: Evaluating the Effectiveness of Deep Learning Models and Data Augmentation" assesses the efficacy of deep learning models for detecting violence in audio, with a specific focus on evaluating the impact of data augmentation techniques. The study likely investigates the performance of various deep learning architectures in identifying violent audio content and explores the effectiveness of augmenting training data to enhance model generalization. The findings may contribute insights to the field of audio-based violence detection and the optimization of deep learning methodologies.

The paper by Tang (2023) titled "AnimeNet: A Deep Learning Approach for Detecting Violence and Eroticism in Animated Content" introduces AnimeNet, a deep learning framework designed for the detection of violence and eroticism in animated content. The study likely presents the development and evaluation of this model, aiming to provide a specialized tool for content analysis in the context of animated media. The research contributes to the field by addressing specific challenges related to distinguishing and categorizing sensitive content in animated materials through deep learning techniques.

The paper by Subramani et al. (2018) titled "Domestic Violence Crisis Identification from Facebook Posts Based on Deep Learning" explores the application of deep learning techniques to identify domestic violence crises from posts on Facebook. The study likely involves the development and evaluation of a deep learning model tailored for recognizing patterns indicative of domestic violence within social media content. The research contributes to leveraging advanced technologies for early detection and intervention in domestic violence situations using online platforms.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

DOI: 10.15680/IJIRCCE.2024.1204184 |

The study by Menger et al. (2018) titled "Comparing Deep Learning and Classical Machine Learning Approaches for Predicting Inpatient Violence Incidents from Clinical Text" evaluates the performance of deep learning and traditional machine learning methods in predicting violence incidents using clinical text data. The research likely involves a comparative analysis of the two approaches, exploring their effectiveness in identifying patterns related to inpatient violence. The findings contribute insights into the suitability of deep learning versus classical machine learning for predictive tasks in the context of clinical text data and violence incidents.

The paper by Castorena et al. (2021) titled "Deep Neural Network for Gender-Based Violence Detection on Twitter Messages" introduces a deep learning approach for identifying gender-based violence in messages on Twitter. The study likely involves the development and evaluation of a deep neural network model tailored for detecting patterns indicative of gender-based violence within the social media platform. This research contributes to the advancement of computational methods for addressing social issues by leveraging deep learning techniques for the detection of gender-based violence in online communication.

The paper by Khan et al. (2024) titled "Detection of Violence Incitation Expressions in Urdu Tweets Using Convolutional Neural Network" focuses on leveraging a convolutional neural network (CNN) for identifying expressions inciting violence in Urdu tweets. The study likely involves the development and evaluation of the CNN model, emphasizing its effectiveness in recognizing patterns indicative of violence incitement in the Urdu language on social media. This research contributes to the application of deep learning techniques for detecting potentially harmful content in multilingual social media environments.

Eduardo Dias Gomes' dissertation from 2022 presents a deep learning-based algorithm designed for detecting violence in audio data. The study likely involves the development and assessment of the algorithm, showcasing its efficacy in identifying patterns indicative of violent audio content. This research contributes to the field of audio analysis by introducing a specialized deep learning approach for violence detection.

The paper by Vosta and Yow (2023) introduces "KianNet," a violence detection model employing an attention-based CNN-LSTM structure. The study likely focuses on combining convolutional neural networks (CNN) and long short-term memory (LSTM) networks with attention mechanisms for improved violence detection. This research contributes to the field of computer vision by proposing a novel model architecture designed to enhance the accuracy of identifying violent content in visual data.

The seminal paper by LeCun, Bengio, and Hinton (2015) titled "Deep Learning" provides a comprehensive overview of the deep learning field. It outlines the fundamental concepts, architectures, and challenges in training deep neural networksemphasizing the significance of deep learning in various applications. The authors highlight the role of unsupervised learning and deep neural network structures, contributing to the foundational understanding of deep learning in the scientific community.

The paper by Wang, Zhao, and Pourpanah (2020) titled "Recent Advances in Deep Learning" provides a concise overview of advancements in the field. It likely highlights key developments and trends in deep learning techniques, algorithms, and applications. The paper contributes to the understanding of the current state of deep learning, offering insights into the latest progress in the rapidly evolving field.

The book "Fundamentals of Deep Learning" by Buduma, Buduma, and Papa serves as a comprehensive guide to the foundational principles of deep learning. It likely covers essential concepts, techniques, and practical applications, providing readers with a fundamental understanding of the field. The book contributes to the educational resources available on deep learning, aiming to equip readers with the knowledge necessary to navigate and apply deep learning methodologies.

The survey by Dong, Wang, and Abbas (2021) titled "A Survey on Deep Learning and Its Applications" offers a comprehensive overview of deep learning and its diverse applications. It likely covers the evolution of deep learning, its underlying methodologies, and explores the broad spectrum of practical applications across various domains. This survey contributes to consolidating knowledge in the field, providing a valuable resource for understanding thecurrent landscape and potential future directions of deep learning applications.

The paper by Janiesch, Zschech, and Heinrich (2021) titled "Machine Learning and Deep Learning" in Electronic Markets discusses the integration and impact of machine learning and deep learning in contemporary business contexts. It likely addresses the evolving role of these technologies in electronic markets, exploring their applications, challenges, and potential benefits. The paper contributes to the understanding of how machine learning and deep

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

DOI: 10.15680/IJIRCCE.2024.1204184 |

learning are shaping electronic markets and influencing business processes.

The preprint by Patel (2021) titled "Real-Time Violence Detection Using CNN-LSTM" focuses on implementing a model combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks for real-time violence detection. The study likely explores the effectiveness of this architecture in identifying violent content, addressing the temporal and spatial aspects of video data. This research contributes to the development of efficient real-time violence detection systems by leveraging the capabilities of CNNs and LSTMs.

The study by Shim and Park (2015) focuses on a violence recognition system using Closed-Circuit Television (CCTV). It likely explores the development and evaluation of a system aimed at detecting violent incidents through CCTV footage. This research contributes to the field of video analytics by addressing violence recognition in real-world scenarios, emphasizing the application of technology for enhancing security measures.

Table 3: Techniques used in various research

S.No.	Author and Reference	Method	Dataset	Improvements	Findings
1	WEI SONG, DONGLIANG ZHANG, XIAOBING ZHAO, JING YU, RUI ZHENG, AND ANTAI WANG	3D CNN with multiple branches to extract features from different parts of the video, and then combining these features to make a final prediction.	• Hockey Fight • Movie Violence • Crowd Violence	This paper uses a 3D ConvNet it could be further improved by the use of adaptive deep networks.	This paper presents a 3D ConvNet model which compares 3 different datasets and uses separate strategies to suit the detection of violence. These comparisons signify that the particular approach was effective.
2	Muhammad Ramzan , Adnan Abid, Hikmat Ullah Khan, Shahid Mahmood Awan, Amina Ismail, Muzamil Ahmed, Mahwish Ilyas, Ahsan Mahmood	Involves using state-of- the-art violence detection techniques such as SVM,CNN and traditional machine learning algorithms and comparisons is done on various datasets to check the performance of the algorithms.	• Hockey Fight • UCF-101 • Movie Violence	This paper contributes in highlighting the techniques and methods of violence activity detection from surveillance videos.	In this paper, the methods of detection are divided into three categories that is based on classification techniques used: violence detection using traditional machine learning, using Support Vector Machine (SVM) and using Deep Learning.
3	Ming Cheng, Kunjing Cai, Ming Li	flow gated network which is divided into 4 parts the RGB channel, the Optical Flow channel, the Merging Block, and the Fully Connected Layer	• Hockey Fight • Movie Violence • Crowd Violence	This paper involves the datatset being tested only by 3 datasets which come under the RWF-2000 dataset because of the heterogeneity in the data format, annotation type etc. So either more datasets can be added or the size of it can be increased	This paper presents both a novel dataset and a method for violence detection in surveillance videos. A unique pooling mechanism is employed by optical flow, which could implement temporal feature pooling instead of human-designed strategies.

4	Simone Accattoli, Paolo Sernani, Nicola Falcionelli, Dagmawi Neway Mekuria & Aldo Franco Dragoni	This paper proposes a new method for violence detection in videos. The method combines the strengths of 3D Convolutional Neural Networks (3D-CNNs) and Support Vector Machines (SVMs) to improve the accuracy of violence detection.	• Hockey Fight • Crowd Fight • Movie Violence	The computational time taken by this algorithm is too high for a real time violence detection.a large portion of the computation time was due to the memory load of the C3D network, instead of the real output computation. So a better management would solve this issue.	The paper evaluates the performance of the proposed method on a publicly available dataset and compares it To state-of-the-art methods.
5	Fath U Min Ullah , Amin Ullah , Khan Muhammad , Ijaz Ul Haq and Sung Wook Baik	In this paper, a three-staged end- to-end framework is proposed for violence detection in a surveillance video stream. First stage persons are detected using CNN In the first stage, persons are detected using an ecient CNN model to remove unwanted frames, which results in reducing the overall processing time. Next, frames sequences with persons are fed into a 3D CNN model trained on three benchmark datasets, where the s patiotemporal features are extracted and forwarded to the Softmax classifier for final predictions.	• Hockey Fight • Movie Violence • Crowd Violence	The model resulted in providing a better accuracy than the state-of-the-art models. It could be further improved by implementing in resource constrained devices	This paper proposes a method for detecting violence in videos using a 3D convolutional neural network (CNN) with spati otemporal features. The method involves using a 3D CNN with multiple co nvolutional and pooling layers to extract spatiotemporal features from video frames.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

DOI: 10.15680/IJIRCCE.2024.1204184 |

6	MIN-SEOK KANG, RAE-HONG PARK AND HYUNG-MIN PARK	They proposed a novel violence detection pipeline that can be combined with the conventional 2-dimensional Convolutional Neural Networks (2D CNNs). frame- grouping is proposed to give the 2D CNNs the ability to learn spatio-temporal representations. Our proposed pipeline brings signicant performance improvements compared tothe 2D CNNs followed by the Long Short- Term Memory	• RFW-2000, • Hockey Fight • Movie Violence • Crowd Violence • Survillence Fight • RLVS	we can collect more data and explore a variety of data augmentation techniques to train a more robust model. Also, we can extend their work to address various action recognition tasks for a versatile use.	We demonstrated the efficiency of our proposed modules with efficient 2D CNN backbones through a variety of experiments and successfully implemented a real-time violence recognition system in a resource-constrained environment.
7	Fernando J. Rendó-Segador, Juan A. Álvarez-García, Fernando Enríquez and Oscar Deniz	They present a new deep learning architecture, using an adapted version of DenseNet for three dimensions, a multi-head self-attention layer and a bidirectional convolutional long short-term memory (LSTM) module, that allows encoding relevant spatio-temporal feature. An ablation study of the input frames, comparing dense optical flow and adjacent frames subtraction and the influence of the attention layer is carried out.	• HF • MF • Violent flow • RLVS	This model does not include any human features, and it works correctly given the input dataset, but to achieve a generalization of violence involving people it would be necessary to include pose estimation or at least face detection in our future work.	We have proposed Violence Net, a space-time encoder architecture for the detection of violent actions that improves the state of the art.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

DOI: 10.15680/IJIRCCE.2024.1204184 |

8	Enrique Bermejo Nievas, Oscar Deniz Suarez, Gloria Bueno Garc´ıa, and Rahul Sukthankar	The approach is in such a way that the video sequence is considered as a space-time volume using gradients, intensities, flows etc. It also involves spatio-temporal of 3D descriptors at multiple scales in highres videos.	• Hockey Fight • Movie Violence	This paper bag-of-words approach can accurately recognize fight sequences with approximately 90% accuracy. It signifies that it could be futher improved by increasing the number of datasets	This paper evaluates how state-of-the-art video descriptors can perform fight detection on two new datasets: a 1000-video collection of NHL hockey games and a smaller 200-clip collection of scenes from action movies

III. CONCLUSION

In this paper, we investigated the performance of six state-of-the-art deep learning models, DenseNet121, ResNet50v2, VGG16, MobileNetV2, C3D and I3D, for the task of violence detection in videos. We evaluated the models on seven datasets with varying levels of violence and complexity, including RLVS, RWF200, Fight Surveillance Data, Bus Violence Data, Crowd Violence Data, Movies Violence Data, and Hockey Fights dataset.

The study shows that deep learning models can effectively detect violence in videos, even in complex and diverse scenarios, and provides valuable insights into the state-of-the-art models for violence detection in videos. The performance of the models varies depending on the dataset and the type of violence depicted. Therefore, violence detection systems need to be trained on a diverse set of data to generalize well to different scenarios.

Transfer learning with the DenseNet121, ResNet50v2, VGG16, and MobileNetV2 models involves using the pre-trained models as feature extractors and training a classifier on top of the extracted features. The C3D and I3D models, which are specifically designed for video analysis tasks, involve fine-tuning the last few layers of the models on violence detection datasets. The study also investigates the performance of the models when trained on the RLVS and RWF2000 datasets separately and tested on the other datasets. The results show that the models did not perform as well in this scenario compared to training and testing on the same dataset.We can see that it is important to consider the domain and diversity of the data when developing violence detection models. Training on a single dataset may not be sufficient for achieving good performance on other datasets, particularly if there is significant variation in violence types and contexts. Therefore, it may be necessary to train on multiple datasets and use techniques such as domain adaptation to improve generalization to new data.

The experimental results suggest that transfer learning was particularly effective with the C3D and I3D models, which achieved the highest performance on most of the datasets. Future work could explore the use of ensemble methods to further improve the performance of violence detection models. Moreover, the integration of contextual information, such as audio, text, and social media data, could enhance the accuracy and robustness of violence detection systems in real-world applications.

REFERENCES

1. M. Ramzan et al., "A Review on State-of-the-Art Violence Detection Techniques," in IEEE Access, vol. 7,pp. 107560-107575, 2019, doi: 10.1109/ACCESS.2019.2932114.
2. W. Song, D. Zhang, X. Zhao, J. Yu, R. Zheng and A. Wang, "A Novel Violent Video Detection Scheme Based on Modified 3D Convolutional Neural Networks," in IEEE Access, vol. 7, pp. 39172-39179, 2019, doi: 10.1109/ACCESS.2019.2906275.
3. Ming Cheng, Kunjing Cai, Ming Li, 2019,'RWF-2000: An Open Large Scale Video Database for Violence Detection', arXiv:1911.05913
4. Serrano Gracia I, Deniz Suarez O, Bueno Garcia G, Kim TK. Fast fight detection. PLoS One. 2015 Apr 10;10(4):e0120448. doi: 10.1371/journal.pone.0120448. PMID: 25860667; PMCID: PMC4393294.
5. Bermejo, E., Déniz-Suárez, O., García, G.B., & Sukthankar,2011, 'Violence Detection in Video Using Computer Vision Techniques', International Conference on Computer Analysis of Images and Patterns

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

DOI: 10.15680/IJIRCCE.2024.1204184 |

6. Simone Accattoli, Paolo Sernani, Nicola Falcionelli, Dagmawi Neway Mekuria & Aldo Franco Dragoni,2020, 'Violence Detection in Videos by Combining 3D Convolutional Neural Networks and Support Vector Machines', Applied Artificial Intelligence,329-344, DOI: 10.1080/08839514.2020.1723876
7. Ullah FUM, Ullah A, Muhammad K, Haq IU, Baik SW. Violence Detection Using Spatiotemporal Features with 3D Convolutional Neural Network. Sensors (Basel). 2019 May 30;19(11):2472. doi: 10.3390/s19112472. PMID: 31151184; PMCID: PMC6603512
8. Jain and D. K. Vishwakarma, "Deep NeuralNet For Violence Detection Using Motion Features From Dynamic Images," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 826-831, doi: 10.1109/ICSSIT48917.2020.9214153
9. M. -S. Kang, R. -H. Park and H. -M. Park, "Efficient Spatio- Temporal Modeling Methods for Real-Time Violence Recognition," in IEEE Access, vol. 9, pp. 76270-76285, 2021, doi: 10.1109/ACCESS.2021.3083273.
10. Traoré and M. A. Akhloufi, "Violence Detection in Videos using Deep Recurrent and Convolutional Neural Networks," 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Toronto, ON, Canada, 2020, pp. 154-159, doi: 10.1109/SMC42975.2020.9282971
11. M. M. Soliman, M. H. Kamal, M. A. El-Massih Nashed, Y. M. Mostafa, B. S. Chawky and D. Khattab, "Violence Recognition from Videos using Deep Learning Techniques," 2019 Ninth International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 2019, pp. 80-85, doi: 10.1109/ICICIS46948.2019.9014714