ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

www.ijasem.org

Vol 18, Issue 1, 2024

REAL TIME VEDIO BASED VOILENCE DETECTION SYSTEM IN PUBLIC AREA USING NEURAL NETWORKS

Ms. P.Udaya Sri, Assistant Professor, Department Of ECE SICET, Hyderabad

CH.Sindhuja, B.Harshitha, D.Sushanth, B.Premanand

UG Student, Department Of ECE, SICET, Hyderabad

ABSTRACT

Crime harms our environment and creates discontent in society. From street fights to shootings and mass murders, the number of terrorist acts continues to increase worl dwide. In order to avoid negative problems, these violations must be displayed and r eported in a certain way. CCTV helps protect public health and safety by recording ill egal behavior so law enforcement can take action. Security cameras are installed in businesses, train stations, shopping malls, banks, ATMs, etc. It can be used easily in some crowded or public places such as. [1] We compare the accuracy of different pr e-
trained deep learning (DL) models and describe the training of models using XAI. Ac cording to the inconsistencies in the literature, it has been observed that the referenc e focuses on the investigation of violence, without any clear evidence or explanation of the investigation carried out. However, our article focuses on the findings and inter pretations of the crowd.

INTRODUCTION

This project creates new unique insights by obtaining real CCTV evidence of armed and unarmed violent crime as well as anti-
violence attacks on YouTube. Optimize inference latency and simplify video delivery tasks to support smart city sustainability. For video recognition and video classificatio n purposes, the results show that the image distribution is equivalent to the existing SOTA 3D network, capable of identifying and distinguishing armed and unarmed viol

ISSN2454-9940

www.ijasem.org

Vol 18, Issue 1, 2024

ence. The models used are Deep Convolutional Neural Networks (DCNN or CNN), Visual Geometry Group (VGG) and Alex Net. Thus, VGG16 achieved an average accuracy of 97%, ResNet50 achieved an average accuracy of 98%, and DenseNet121 achieved an average accuracy of 98%. The biggest problem with this model is overfitting [2]. The average accuracy achieved by each model is 90%, indicating that each model is superior [5]. A general perception model based on human morphological and dynamic characteristics that can detect human crowds has been developed. Ability of the model to perform better by highlighting human body content and dynamic features obtained from adjacent frames. From the feature map, spatial features can be extracted by 3D CNN framework and temporal data can be extracted using long-term memory (LSTM) model called HD-Net. References include Hockey, Real Life Violence Situations (RLVS), and Violence Broadcast. This document uses different metrics to measure the overall capabilities of HD Net. This general ability is validated by comparison with other classical crime detection algorithms. Literature

REVIEW

This study compares Bayesian video classification neural networks with traditional video classification neural networks in terms of performance and capability. For each classification, the Bayesian video network outperformed the comparable non-Bayesian video network in terms of accuracy. Bayesian Neural Networks (BNN) are used to measure uncertainty, Relational Neural Networks (RNN) are used to classify image sequences (movies), and CNNs are used to classify images. The non-Bayesian equivalent root algorithm is used for comparison. The use of hybrid methods leads to a significant improvement in the classification of exposures. The accuracy of the Bayesian front-end network is 61.0%. The accuracy of the core network increased by 52% and 8% respectively.

ISSN2454-9940

www.ijasem.org

Vol 18, Issue 1, 2024

It has CNN front-end and RNN back-end respectively. The accuracy of the RNN backend dropped by 100%, while the accuracy of the Bayesian frontend increased by 192% [4]. The network architecture used, the application process and the data used are the subject of this review. To make the research independent, the progress in deep learning for video display was carefully evaluated and summarized. The report also outlines metrics and performance metrics for video distribution. This study offers new directions for further research into which of the three commonly used methods works best for video classification and will be determined by this study. The research also aims to examine the ins and outs of creating a reliable video distribution system and how these systems impact/continue to improve video rankings. This new study meets research needs by providing guidance to researchers in the region and providing a comparative analysis of three methods, showing their strengths and weaknesses. The findings were brought together through extensive analysis using hard documentation and large videos (29,000 videos). This will provide information for researchers in the field to choose the best game plan for their video ranking studies. The results show that deep learning models have the best performance in terms of direct connection to machine learning and video classification.

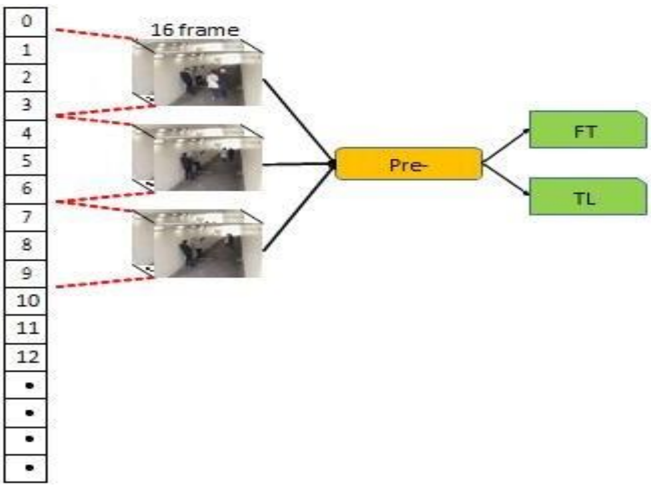
METHODOLOGY

This post justifies the long-term study debate 3 on whether this method is better than video distribution [6]. In this study, researchers used large numbers of people to investigate crime. The pipeline is derived from the video dataset and feature extraction. The author chose Close-Up Competition because of the large number of shots taken with crowd participation. 3D CNN Res Net architecture adopts the following features. An equal number of 246 training videos containing rigid and non-rigid content were used, with an average length of 14.76 minutes. A prediction from 0 to 1 is used to determine whether a video falls into the violent or non-violent category; where 0 means no violence and 1 means violence. They create a link to their submission and include a "search" option that allows users to browse video files or select video streams. Click the "Scan" button to review the video. Results ar

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN2454-9940
www.ijasem.org
Vol 18, Issue 1, 2024

e shown by displaying frames with bounding boxes and percentages from 0% to 100 %; where 0 represents no violence and 100 represents bad action [7]. The authors studied the effectiveness of CNNs in classifying multiple videos. The data used is 1 million YouTube videos from 487 sports teams. They found that CNN architectures can outperform traditional methods by learning robust features from poorly labeled data. The results are surprising. Ways to speed up and improve crime detection in automated video surveillance include using advanced techniques to increase the speed and accuracy of crime detection. Key elements of this approach may include:
Deep learning models: Use state-of-the-
art deep learning models such as convolutional neural networks (CNN) or convolutional neural networks (RNN) for extraction and recognition. patterns in video frames. This can improve the performance of the model by limiting the number of saved Pdata . Tracking: Combined with object tracking algorithms to track and identify the movements of people in images. This can improve the ability to identify potential conflicts in real time. This increases accuracy and reduces false positives. This makes brute force detection increasingly faster.

Continuous Learning: Implement mechanisms for continuous model updates based on new data, enabling the system to adapt to evolving patterns of violence and



improve over time.

ISSN2454-9940
www.ijasem.org
Vol 18, Issue 1, 2024

Privacy: Integrate privacy techniques such as blurring or anonymizing irrelevant parts of the video to address privacy concerns while maintaining effective tracking capabilities.

By combining these methodologies, an accelerated and improved automated video surveillance violence detection system can effectively increase public safety and security.

RESULT AND ANALYSIS

Accelerating and improving the detection of violence in automated video surveillance using a convolutional neural network (CNN) model can yield several results and analyses. Here are some potential results:

Higher accuracy: Implementing a CNN model to detect violence in video surveillance is likely to result in increased accuracy compared to traditional methods. CNNs are particularly effective at extracting hierarchical features from visual data, making them suitable for image and video analysis tasks.

Reducing false alarms: CNNs can help reduce false alarms, which are cases where the system incorrectly identifies non-violent actions as violent. This is key to ensuring that the surveillance system provides security personnel with reliable and actionable information.

Real-time detection: Depending on the effectiveness of the CNN model and the optimization techniques used, real-time detection of violence can be achieved. This is important for quick response and intervention in security-sensitive scenarios.

Scalability: CNN models can be scaled to handle a large number of surveillance cameras simultaneously. This is essential for deploying the system in large-scale surveillance networks, such as networks in smart cities or large public spaces.

Improved robustness: CNNs are known for their ability to generalize well across different types of data. This means that the violence detection model could work effectively in different lighting conditions, camera angles and other environmental factors.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

www.ijasem.org

Vol 18, Issue 1, 2024

Deep Feature Extraction: A CNN model would automatically learn and extract deep features from video images that indicate violent activities. This can lead to a more nuanced and sophisticated understanding of patterns of violence and thus improve overall detection capability.

Model explainability: Analysis of a CNN model can provide insight into the features and patterns it identifies as indicative of violence. This can contribute to a better understanding of the model's decision-making process and help address any biases or ethical issues.

Optimization Techniques: Research or implementation may include exploring optimization techniques such as transfer learning, fine-tuning, or model compression to increase the effectiveness of the violence detection system.

It's important to note that the success of violence detection using a CNN model relies on the quality and diversity of the training data, the architecture of the CNN, and the optimization strategies employed. Regular updates and maintenance of the model are

also crucial to adapt to evolving scenarios and maintain high detection accuracy.

Analysis of CNN algorithm

Convolutional neural networks (CNNs) have revolutionized the field of computer vision with their remarkable ability to automatically extract and learn hierarchical features from raw image data. At the heart of CNNs are convolutional layers, where learnable filters are applied to input images to extract low-level features such as edges and textures. Through multiple layers of convolutions and pooling operations, CNNs can gradually learn more abstract and complex features, allowing them to capture complex patterns and structures in images.

CONCLUSION

The project dataset is trained with a CNN model. All model predictions are explained using LIME, the Explainable AI framework. What can be learned from the CNN explanation is that more than just training a CNN model is needed to deploy it for a real-world scenario. Models need to be trained using multiple CCTV footage that clearly shows that violence occurs when there is a group of people and that violence does not occur even when there is a group of people. Therefore, it is important to carefully consider and mitigate these drawbacks when developing and deploying deep learning models for mob violence detection and using XAI techniques to explain the results.

REFERENCE

[1] Barredo Arrieta, S. Gil-Lopez, I. La˜na, M. N. Bilbao, and J. Del Ser, "On post-hoc explainability of deep echo networks for time series prediction, image and video classification, Neural Computing and Applications, St. 34, No. 13, pp. 10257–10277, August 2021, doi: 10.1007/s00521-021-06359-y. — Katna documentation," https://katna.readthedocs.io/en/latest/ [17] K. Joshi, V. Tripathi,

C. Bose and C. Bhardwaj, "Robust Sports Image Classification Using InceptionV3 and Neural Networks," Procedia Computer Science, vol. 167, pp. 2374–2381, 2020, doi: 10.1016/j.procs.2020.03.290.

[2] Towards SmartCity Security: Violence and Weaponised Vi olence Detection with DCNN Aremu, T., Zhiyuan, L., & Alameeri, R. (2022). "Any object is a potential weapon! Detecting gun violence using a salient image." arXiv. https://doi.org/10.48550/arXiv.2207.12850

[3] Z. Chexia, Z. Tan, D. Wu, J. Ning, and B. Zhang, "A General Model for Crowd Violence Detection Focusing on Human Contour and Dynamic Features", 2022 22nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid), 2022, pp. 327-335, doi: 10. 1109/ CCGrid54584. 2022. 00042.

[4] Swize, Emmie & Champagne, Lance & Cox, Bruce & Bihl, Trevor. (2022). Bayesian Augmentation of Deep Learning to Prove Video Classification. 10.24251/HICSS.2022.264.

[5] Rehman, Atiq; Belhaouari, Samir Brahim (2021): Deep Learning for Video Classification: A Review.

TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.15172920.v1

[6] Saddam Bekhet, Abdullah M. Alghamdi (2021): A Comparative Study of Video Classification Techniques Using Direct Feature Comparison, Machine Learning and Deep Learning; Volume 56, Issue 4; Journal of Southwest Jiaotong University; ISSN 02582724; 10.35741/issn.0258-2724.56.4.63

[7] Gkountakos, Konstantinos & Ioannidis, Konstantinos & Tsikrika, Theodora & Vrochidis, Stefanos & Kompatsiaris, Ioannis. (2021). Detection of mob violence from video footage. 1-4. 10. 1109/ CBMI50038. 2021. 9461921