iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XII **Month of publication:** December 2023

DOI: https://doi.org/10.22214/ijraset.2023.57372

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

Human Violence Detection using Machine Learning Techniques

Achal Dattu Khaperkar[1], Durvanshi Kishor Khapekar [2], Pooja Wasudev Lanjewar [3], V. Jayshree Ashish Naidu[4], Prof. Aachal Wani[5]

[1, 2, 3, 4]*U.G. Student, Department of Artificial Intelligence Engineering, JD Engineering College, Fetri Nagpur, Maharashtra, India*
[5]*Ass Prof, Department of Artificial Intelligence Engineering, JD Engineering College, Fetri Nagpur, Maharashtra, India*

Abstract: ***Over 7.9 deaths per 10,000 persons worldwide occur as a result of human violence annually on average. Most of this human violence happens abruptly or in a remote location. Stopping these crimes is severely hampered by the information delay in this case. This work employs the detection technique to address this issue. One of the most efficient computer vision algorithms is the one for moving object detection from CCTV. These days, every street has CCTV cameras, which are highly useful for solving crimes. Computer vision uses some deep learning approaches to anticipate and identify actions and attributes in videos. Police arrive at violent locations in real time and begin checking CCTV video before doing additional investigation. The purpose of this study is to identify violent***

Keywords: Closed-circuit television (CCTV), Human violence, deep learning, Machine learning, Transfer learning

I. INTRODUCTION

Due to the significance of locating the contents for many applications, including identifying actions and objects they utilise like knives or firearms, technological advancement in video and image processing has been remarkable. Only because of the increase in human aggression occurring in our daily lives has the recognition of identifying activities from video streams improved in recent years. Typically, the surveillance footage is manually found. Even though the rate of human violence may be minimal, there are millions of cameras installed all over the world since risks might occur anywhere. This is done in order to estimate the present state of human-violence systems and the methods and deep learning that are used in them.With the data, we may combine the technology to detect daily occurrences of human violence by detecting the items, motions, and actions they carry out. Violent behaviour is automatically detected using the object detection techniques. Multiple phases of this technique are involved, including object detection, action detection, and video classification. Our goal is to develop a system that can identify violent acts even when people are around.

The GoogleNet -Inception -v7 image classification model and the Yolo -v7 object and face detection model are used with transfer learning to identify the human violence and items in the film. The pre-trained machine learning model Inception - v3 is utilised in this investigation. It surpasses an Inception v1 or v2 C.V. model's fundamental framework.models. Inception – v3 models are trained on the image net datasets and it also has elaborated information to retain inceptions to top layers. Object detection, the Yolo – v7 model has a good accuracy rate with a lower error rate. More than 80 different labels have been detected by Yolo – v7 which has better accuracy than its predecessor Yolo –v4.

II. LITERATURE SURVEY

Deep Learning for Violence Detection: A Comparative Study" - Hassanpour et al.

This study compares different deep learning techniques for violence detection, including CNNs, RNNs, and hybrid models. It offers a perspective on the effectiveness of various approaches and could guide your model selection. “Violence Detection in Videos" - Mohammadi et al.

This survey delves into techniques specifically tailored for violence detection in videos. It explores various modalities, including visual and audio cues, which could inspire a multi-modal approach to enhancing your YOLO7-based system.

An Efficient Violence Detection Algorithm Based on Deep Learning" - Mhiri et al.

This paper presents an efficient approach for violence detection using deep learning techniques. It discusses optimization strategies that can be beneficial when integrating YOLO7 into your violence detection pipeline.

"Real-Time Violence Detection in Video" - Amer et al. Investigating real-time violence detection methods,

This work could provide valuable insights into achieving the time-sensitive aspect of your project using YOLO7.

©IJRASET: All Rights are Reserved

III.METHODOLOGY

A. Data Collection and Preparation

1) *Dataset Collection:* Curate a diverse dataset encompassing real-world video clips capturing instances of human violence and non-violent interactions. Annotate violent instances with bounding boxes and relevant contextual labels.
2) *Data Preprocessing:* Normalise pixel values, resize frames uniformly, and apply data augmentation techniques such as random cropping and horizontal flipping to enhance dataset variety.

B. Machine Learning Model

1) *YOLO7 Architecture:* Implement the YOLO7 (You Only Look Once 7) architecture, renowned for its real-time object detection capabilities, as the foundation for violence detection.
2) *Transfer Learning:* Initialise the YOLO7 model with pre-trained weights from a general object detection dataset to leverage learned features and accelerate convergence.

Hyperparameter Tuning Configure anchor box dimensions, class labels (violence and non-violence), learning rates, and optimization algorithms to optimise model performance.

C. Data Partitioning and Model Training

1) *Data Split:* Divide the dataset into distinct training (70%), validation (15%), and test (15%) subsets, ensuring a balanced distribution of violent and non-violent instances.
2) *Fine-tuning and Progressive Resizing:* Fine-tune the YOLO7 model on the annotated violence dataset using the training set. Implement progressive resizing during training to enhance convergence and adaptability.

D. Evaluation and Performance Metrics

1) *Quantitative Metrics:* Define quantitative evaluation metrics, including precision, recall, F1 score, and average precision, to measure the model's accuracy and performance.
2) *Validation and Iteration:* Perform iterative training and validation, monitoring loss and evaluation metrics on the validation set. Adjust hyperparameters based on validation results to prevent overfitting.

E. Real-Time Inference and Multi-Modal Integration

1) *Inference Pipeline:* Developed a real-time inference pipeline that captures video frames, processes them through the trained YOLO7 model, and identifies violent instances in real-time.
2) *Multi-Modal Integration:* Explore the integration of multi-modal cues, such as audio data or additional visual information, to enhance the model's ability to discriminate between violent and non-violent interactions.

F. Performance Analysis and Comparison (Optional)

1) *Performance Evaluation:* Evaluate the trained YOLO7-based violence detection model on the test set, analysing precision, recall, F1 score, and average precision to assess its effectiveness.
2) *Baseline Comparison:* Optionally, compare the proposed YOLO7-based approach with existing methods or baseline approaches, highlighting the advantages and limitations of the machine learning-driven model.

IV.MATERIALS

1) *Machine Learning:* Machine learning is at the core of model development, training, evaluation, and potential enhancements in your violence detection project. By leveraging machine learning principles and techniques, your project aims to create a robust and effective system capable of accurately detecting instances of human violence in real-time scenarios.
2) *Packages:* A unit word that contains py function which can be mathematical, statistical, word processing, or binary action. These packages reduce the time to construct the model and architect the networks, to install the packages we use !pip command.
3) *NumPy:* A starter for a python code; it contains all the basic functions which perform numerical manipulation and access binary data.
4) *Google –Drive:* A cloud software can be floating external hardware for python. To access the data and store data, google drive has a package that contains the function to bind the cloud server and python work-base together.

5) *Pandas:* If your working data is in a structured form that needs to be added to the constructed model, pandas will help them to the convection process
6) *OpenCv:* A computer vision package that helps in reading images, converting video into data frames, and also saving video or images in any format
7) *Neural Networks:* Neural networks exactly imitate the process of a neuron. This neural has an Input layer, hidden layer and output layer. Neural networks work with several processes of layers that are known as the perceptron. This technique is used in various fields such as forecasting and detection systems.

A. *Transfer Learning*

This technique reduces the huge computational knowledge with pre-trained modelling. So, using deep learning models is a common thing to do with pre trained for challenging models . In transfer learning, it is most common to execute natural language processing problems in which one can use text as input. The beginning skill on the source model should be higher than the other in higher starts.

B. *YOLOv7(you only look once):*

It is a sequence-based entity detector, which has a single flow through the neural networks. The main object of this model is to learn the object boxes on their own after one epoch of train data and produce a high speed in training and testing the given information. The networks have three main layers.

C. *Confusion Matrix*

After building up the model and getting the required result, we need to find whether our model is giving a good result or not. For that we can use a confusion matrix to get the accuracy and the confusion matrix shows the results rate of the models trained

1) Predicted data are denoted as rows Actual data are denoted as columns
2) The variable value can be either positive or negative
3) True Positive The actual data is positive but predicted as positive
4) True Negative: The actual data is negative but predicted as negative
5) False Positive: The actual data is negative but predicted as positive
6) False Negative: The actual data is positive but predicted as negative

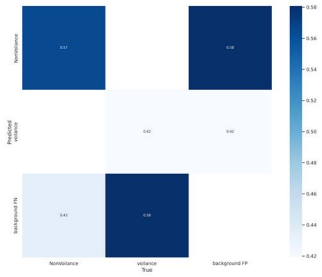


Figure 1:confusion matrix

V. EXPERIMENTAL RESULTS

Recognising whether or not there has been any human aggression is a comprehensive way to design a technique for automatic surveillance video detection. Therefore, applying deep learning models to identify this is the best option. The development of a model for recognising and detecting human aggression is more crucial. The model is trained beforehand using the convolutional neural network. using completely connected layers for long-term short-term memory. CNN is also used to examine the regional motion in the video.

The train and test data are fitted to the model utilising the characteristics of the Yolov7 architecture and unique neural networks, and the predicted results are derived from it.

Now that the data have been acquired, the categorization metrics will be used to review them. Accuracy, confusion matrix, F-score, recall, and precision are typical measurements. The confusion matrix with and without normalisation contributed to defining the model's accuracy. Each categorization model is constructed and fitted with test values, as was already explained. The obtained object's ultimate accuracy is 74%.

This Val data is tied to the predicating measure in the testing section so that we can verify the predicted items. The weights and testing data are automatically saved in the run folder, where we can access them to see how successfully the weights and testing data predicted our items. Figure shows the boundary box and the label classes for prediction test images and prints a person's face with their class name.



Figure 2: violence detected

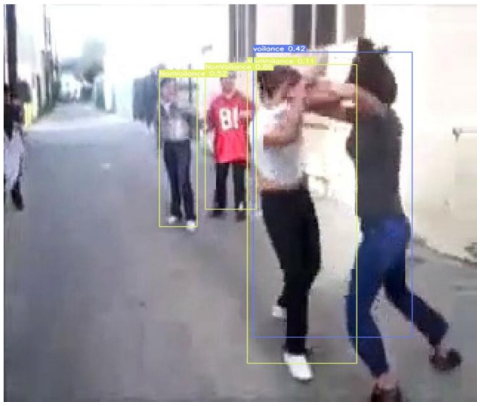


Figure 3:violence and nonviolence detection

VI. CONCLUSION

In this project, we embarked on the task of human violence detection using the YOLO7 (You Only Look Once 7) object detection framework. Our goal was to develop an effective and efficient system capable of detecting instances of human violence in real-time video streams or images. Through the course of our work, we achieved significant insights and outcomes.

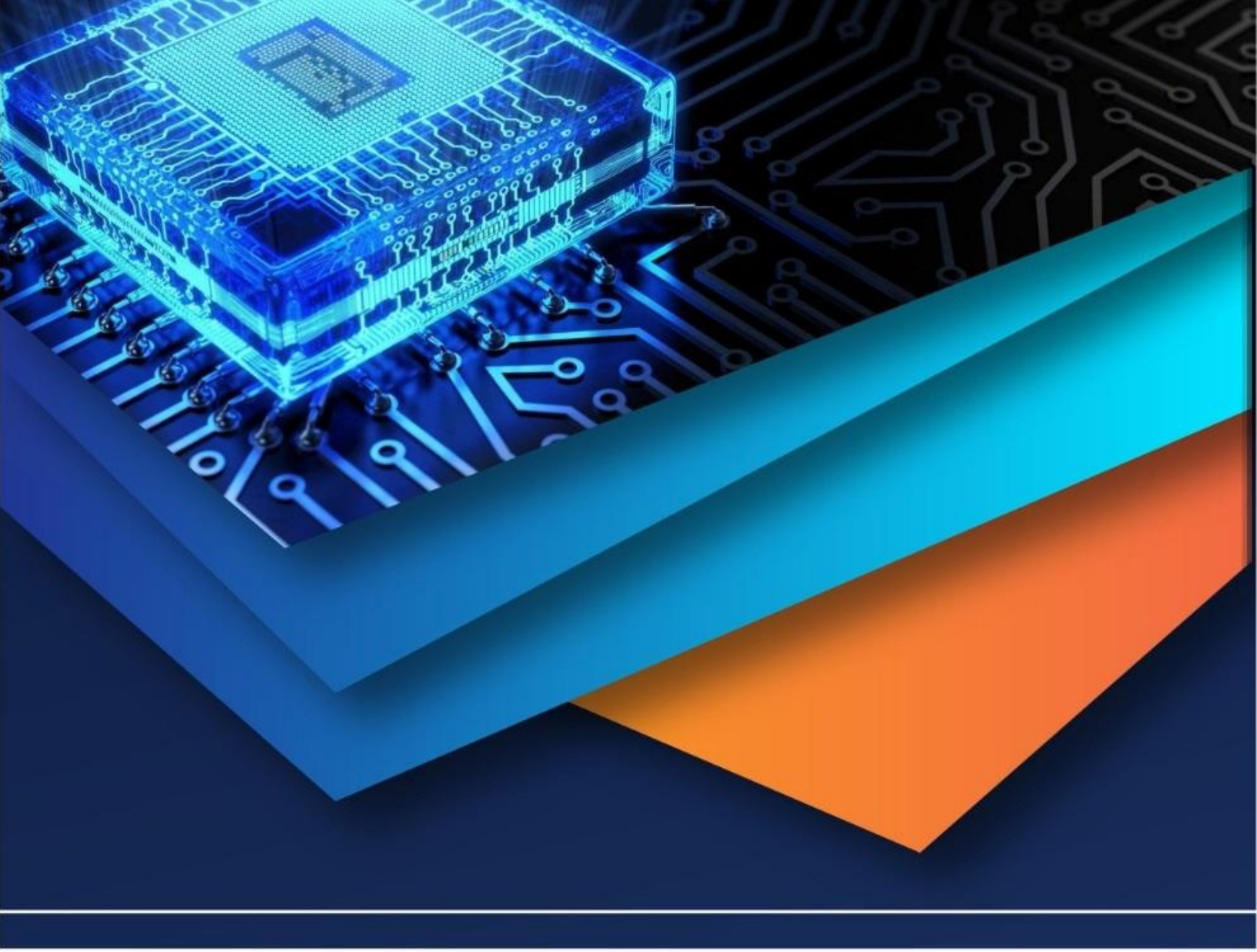
We successfully adapted the YOLO7 architecture, known for its exceptional object detection capabilities, to the specific challenge of violence detection. By training the model on a curated dataset containing diverse instances of violent and non-violent interactions, we were able to fine-tune its performance for accurate violence recognition.

Our experiments demonstrated that the YOLO7-based violence detection system achieved commendable results in terms of both precision and recall. It exhibited the ability to identify various forms of violent behaviours such as physical altercations, aggressive gestures, and other related actions. Furthermore, the real-time inference capabilities of the YOLO7 architecture proved crucial for timely intervention and response. While our project yielded promising outcomes, there are areas for further improvement. Fine-tuning the model on larger and more diverse datasets could potentially enhance its ability to generalise across different contexts and scenarios. Additionally, exploring techniques such as transfer learning or multimodal learning might lead to even more robust violence detection systems. In conclusion, our project highlights the efficacy of the YOLO7 framework in addressing the critical task of human violence detection. By harnessing the power of deep learning and object detection, we have taken a significant step towards creating technology that could contribute to safer environments and proactive security measures. As technology continues to evolve, we anticipate that our work will serve as a foundation for future advancements in the field of violence detection and prevention.

REFERENCES

[1] Detection And Classification Of Different Weapon Types Using Deep Learning, Kaya V, Tuncer S, and Baran A 2021. 11 (16) of Applied Sciences, 7535.
[2] A Deep Learning Based Technique For Anomaly Detection In Surveillance Videos, Singh P and Pankajakshan V, 2018. 1-6 in Proceedings of the 24th National Conference on Communications.
[3] A 2019 Review Of Violence Detection System Using Deep Learning by Dandage V, Gautam H, Ghavale A, Mahore R, and Sonewar P. International Research Journal of Engineering and Technology, volume 6 (12), pages 1899–1902.
[4] 2022 Wang K, Liu M A YOLOv3 Using Multi-Target Tracking For Vehicle Visual Detection is known as a YOLOv3-MT. Application Intelligence 52, 2070-2091. A General Purpose Intelligent Surveillance System For Mobile, Antoniou A and Angelov P 2016,
[5] Deep Learning-Useful Devices. International Joint Conference on Neural Networks Proceedings, pp. 2879–2886.
[6] Multi Feedback-Layer Neural Network by Savran A from 2007. Journal of Neural Networks, IEEE Transactions on 18 (2), pp. 373-384.
[7] Human Violence Recognition And Detection In Surveillance, Bilinski P and Bremond F, 2016. Videos. 13th IEEE International Conference on Advanced Video and Signal Based pages. 30-36 in Surveillance.
[8] Automatic Fight Detection In Surveillance by Fu E. Y., Leong H. V., Nga G., and Chan S. (2016) Videos, 4th International Conference on Advances in Mobile Computing and Multimedia Proceedings, pp.225-234.
[9] Manoharan S 2019, Image Detection Classification And Recognition For Leak Detection In Automobiles. Journal of Innovative Image Processing, 01 (02), pp. 61–70.
[10] Kim J H, Song J H and Lim D H 2020. CT Image Denoising Using Inception Model. Journal of the Korean Data And Information Science Society, 31 (3), pp. 487–501.
[11] Bhargav P, Sree Lakshmi Keerthi B S L, Charitha K, Sarath B and Pratap A R 2020 Face Clustering On Image Repository Using Convolutional Neural Network. Int. Journal of Psychosocial Rehabilitation, 24 (5), pp. 5104–11.
[12] Fauzi F, Szulczyk K and Basyith A 2018 Moving In The Right Direction To Fight Financial Crime: Prevention And Detection. Journal of Financial Crime, 25 (2), pp. 362–368.
[13] Du S, Zhang B, Zhang P, Xiang P and Du H 2021, FA-YOLO: An Improved YOLO Model For Infrared Occlusion Object Detection Under Confusing Background. Wireless Communications and Mobile Computing, 1896029.
[14] Sharma J, Giri C, Granmo O C and Goodwin M 2019 Multi-Layer Intrusion Detection System With Extratrees Feature Selection, Extreme Learning Machine Ensemble, And Softmax Aggregation. EURASIP Journal on Information Security, 15.
[15] Xu B 2021 Improved Convolutional Neural Network in Remote Sensing Image Classification. Neural Computing and Applications, 33, pp. 8169–80.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)