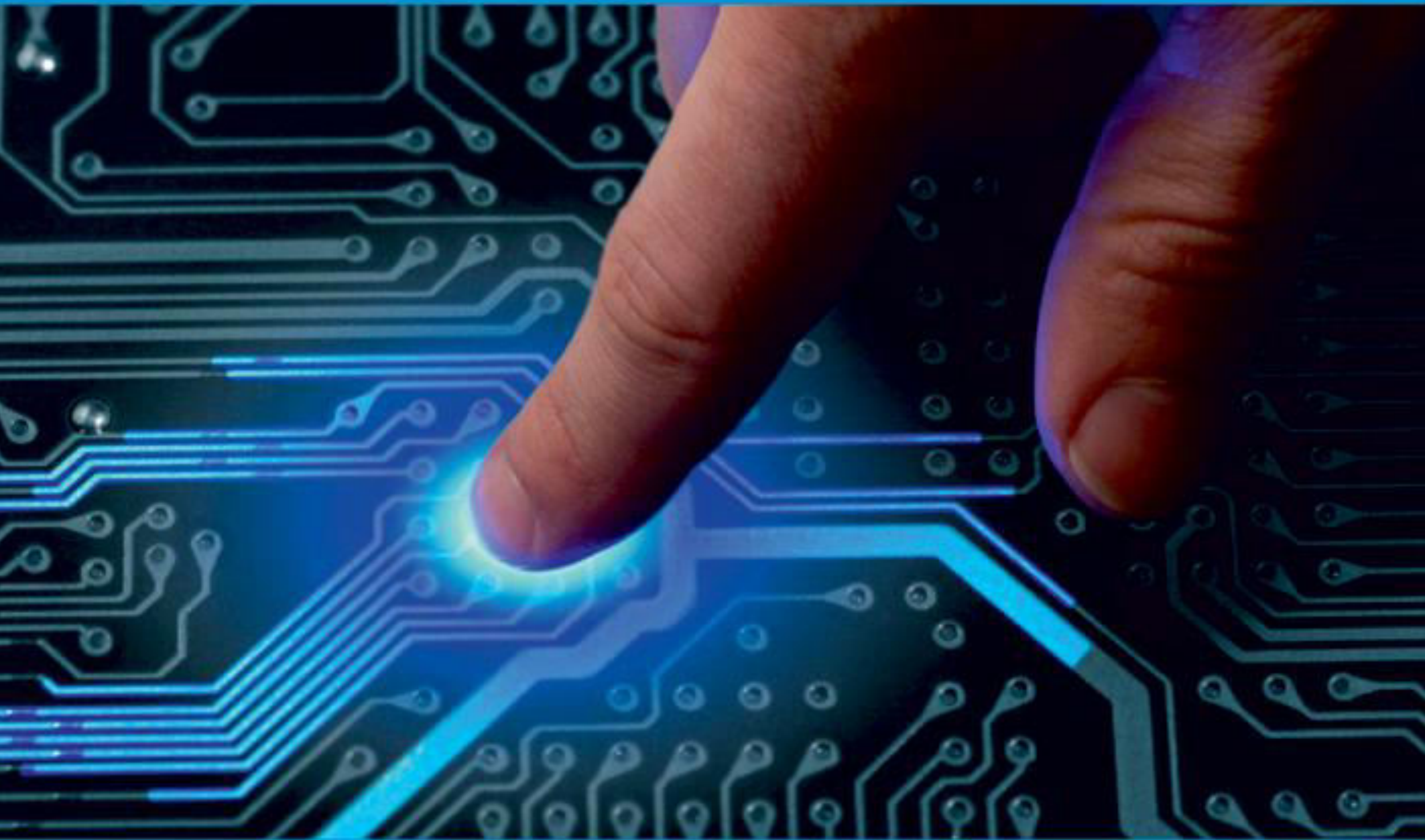
IJIRCCE

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

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.2024.1204130 |

Human Anomaly Detection Using Deep Learning

B Naga Manish, Jilla Sainandhan, T Sai Teja, NMS.Desai

Student, Department of Computer Science and Engineering, Anurag University, Hyderabad, India

Associate Professor, Department of Computer Science and Engineering, Anurag University, Hyderabad, India

ABSTRACT: In our current world the need for automation is increasing rapidly. We all are trying to replace the human activities with a machine, so that various events can be accomplished in no time and can achieve them with higher accuracy. As the various unusual human activities are increasing rapidly day-to-day, we will now try to automate the process to detect the various situations such as Chain Snatching, Theft, Accidents, Cardiac Arrest, Fighting, Violence etc.

We will use the various Image Processing techniques to pre- process the image and then apply the Machine Learning techniques to construct the model and use it for identifying the type of situation. We will not only use the model for classifying the situation but also to notify the anomaly situation.

KEYWORDS: Image Pre-processing, Machine Learning, DENSENET.

I. INTRODUCTION

The predominance of violent behaviors poses a severe threat to social stability and personal safety. A number of approaches have been tested to solve this problem and stop the violent behavior, one of which is the installation of surveillance equipment.

The use of surveillance systems is one of several strategies that have been tried to stop violent behavior. The ability of the monitoring systems to automatically identify aggressive behavior and send out alarm or warning signals will be very important. A series of steps can be used to implement the entire system. The system must first detect the presence of people in a video frame. The frames that are anticipated to feature aggressive behavior must then be removed. Now is the time to remove the unnecessary frames. Last but not least, the trained model recognizes aggressive behaviors, and these frames are individually preserved as photos.

The photographs are transmitted as a warning to the relevant authority together with other crucial information, like the time and place. The suggested solution uses deep learning to automatically identify violence in a video using different algorithms. However, utilizing only conventional network has the drawback of being less precise and taking a much time to compute. i.e., a pre-trained model called Mobile Net was created. It offers more accuracy and serves as the foundation for creating the full model.
Using the Telegram app, a warning is sent to the relevant authorities. we make use of different

The whole system can be implemented with a sequence of steps:

1. The installation of the surveillance systems;
2. The programming of the system to detect violent activities;
3. The setting of certain thresholds to determine whether an activity should trigger a warning or alert signal;
4. The regular maintenance and updating of the system to ensure it remains functional;

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.2024.1204130 |

5. The implementation of effective measures to help prevent violent activities from occurring in the first place. With all these steps taken, the surveillance systems can be effective in deterring and preventing violent activities, and thus help to ensure personal safety and social stability.

II. RELATED WORK

This paper related work has done uniquely, We will use the various Image Processing techniques to pre- process the image and then apply the Deep Learning techniques to construct the model and use it for identifying the type of situation. We will not only use the model for classifying the situation but also to notify the anomaly situation.

III. EXISTING METHOD

The CNN, or convolution neural network, has been used in our proposed method to identify unusual activity. Understanding the time information in the video is essential for correctly categorizing unexpected behaviors.
Recently, CNN has mostly been used to extract significant details from each video frame. The algorithm that best meets this requirement is or convolution neural network.
In the way to correctly classify the input, or convolution neural network must be able to recognize and extract different necessary information out the frame of videos, hence or convolution neural network must be able to do so.

IV. PROPOSED METHOD

We will develop a (LRCN) which is totally based system for monitoring CCTV surveillance footage and detecting normal activities such as jogging and walking, as well as abnormal activities such as fighting. If any suspicious activity is detected, this system can be used to generate an alert that will remind the user.
LRCN stands for Long-term Recurrent Convolutional Networks. LRCN is a neural network architecture that combines convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to process both spatial and temporal information in sequential data, such as videos or image sequences.
The LRCN architecture is designed to leverage the power of CNNs in extracting spatial features from individual frames of a video or image sequence. The output of the CNN is then fed into an RNN, typically an LSTM, to capture temporal dependencies and model the sequence of frames over time.

V. RESULTS

Our proposed model aims to identify abnormal behavior in video data, and we have achieved a commendable accuracy of 82% on our custom dataset.

Figure No 1: Output Screens -I

| 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.2024.1204130 |

Figure No 2: Output Screens -II

Figure No 3: Output Screens -III

VI. CONCLUSION AND FUTURE WORK

In our project, we've dedicated our efforts to crafting a sophisticated activity detection system specifically tailored for surveillance videos, with a primary focus on identifying behaviors like fighting, walking, and running. Our approach revolves around harnessing the power of the Long-term Recurrent Convolutional Network (LRCN) architecture, which enables us to capture both spatial and temporal features crucial for accurate classification.

Our dataset compilation involved meticulous selection and curation of 300 surveillance videos, each containing instances of the target activities. We adhered to ethical guidelines and considerations surrounding data collection, ensuring privacy and consent were paramount. Additionally, we undertook rigorous preprocessing steps to standardize

| 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.2024.1204130 |

the video format, including frame extraction, resizing, and potential normalization, thereby optimizing compatibility with the LRCN model.

Delving into the architecture design, we seamlessly integrated convolutional layers to extract spatial features from individual frames. These were complemented by recurrent layers that facilitated the capture of temporal dependencies across frames. Our model implementation drew inspiration from official documentation and research papers, alongside leveraging established deep learning frameworks like TensorFlow or PyTorch for efficient execution.

During the training and evaluation phase, we maintained transparency and reproducibility by adhering to best practices. This entailed dividing the dataset into training, validation, and testing subsets to comprehensively assess the model's generalization capabilities. To gauge performance, we employed a suite of evaluation metrics, including accuracy, precision, recall, and F1-score, selected for their relevance to our binary or multiclass classification task.

Our model yielded promising results, achieving an accuracy rate of approximately 84% on the test dataset. This underscores its efficacy in discerning activities of interest within surveillance footage. Looking ahead, we've identified avenues for refinement and enhancement based on insights gleaned from our endeavor. These include exploring advanced model architectures such as Temporal Convolutional Networks (TCNs), integrating data augmentation techniques to bolster robustness, and harnessing transfer learning from pre-trained models to further elevate performance.

REFERENCES

1. M. E. Issa, A. M. Helmi, M. A. A. Al-Qaness, A. Dahou, M. A. Elaziz, and R. Damaševičius, "Human activity recognition based on embedded sensor data fusion for the internet of healthcare things," *Healthcare*, vol. 10, no. 6, p. 1084, 2022. View at: Publisher Site | Google Scholar
2. G. Şengül, E. Ozcelik, S. Misra, R. Damaševičius, and R. Maskeliūnas, "Fusion of smartphone sensor data for classification of daily user activities," *Multimedia Tools and Applications*, vol. 80, no. 24, pp. 33527–33546, 2021. View at: Publisher Site | Google Scholar
3. M. T. Bhatti, M. G. Khan, M. Aslam, and M. J. Fiaz, "Weapon detection in real-time CCTV videos using deep learning," *IEEE Access*, vol. 9, pp. 34366–34382, 2021. View at: Publisher Site | Google Scholar
4. F. Afza, M. A. Khan, M. Sharif et al., "A framework of human action recognition using length control features fusion and weighted entropy-variances based feature selection," *Image and Vision Computing*, vol. 106, Article ID 104090, 2021. View at: Publisher Site | Google Scholar
5. P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea, "Machine recognition of human activities: a survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 11, pp. 1473–1488, 2008. View at: Publisher Site | Google Scholar
6. Y. ZhuZhu, N. M. Nayak, and A. K. Roy-Chowdhury, "Context-aware activity modeling using hierarchical conditional random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 7, pp. 1360–1372, 2015, Jul. View at: Publisher Site | Google Scholar
7. Yue-Hei Ng, M. Hausknecht, S. Vijayanarasimhan, O. Vinyals, R. Monga, and G. Toderici, "Beyond short snippets: deep networks for video classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4694–4702, Boston, MA, USA, June 2015. View at: Google Scholar
8. M. S. Ibrahim, S. Muralidharan, Z. Deng, A. Vahdat, and G. Mori, "A hierarchical deep temporal model for group activity recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1971–1980, Las Vegas, NV, USA, June 2016. View at: Publisher Site | Google Scholar