PROJECT REPORT

SURVEILLANCE ACTIVITY DETECTION USING DEEP LEARNING

KALASALINGAM UNIVERSITY - INTERNSHIP PROGRAM

Submitted by

ADHAVAN M (sit20cs003@sairamtap.edu.in) (7550194049)

HARISH KUMAR S (sit20cs007@sairamtap.edu.in) (9176941079)

SARAN R (sit20cs064@sairamtap.edu.in) (9499040940)

YAKESH KARTHIKEYAN B(sit20cs012@sairamtap.edu.in) (8903046081)

FROM: SRI SAIRAM INSTITUTE OF TECHNOLOGY CHENNAI.

INTRODUCTION:

The analysis of human activities is one of the most interesting and important open issues for the automated video surveillance community. In order to understand the behaviors of humans, a higher level of understanding is required, which is generally referred to as activity recognition. While traditional approaches rely on 2D data like images or videos, the development of low-cost depth sensors created new opportunities to advance the field. In this paper, a system to recognize human activities using data of CCTV cameras is proposed. The proposed system focuses on recognizing human activities. And detect any Violence and accident events occurs or not Human activities take place over different time scales and consist of a sequence of sub-activities (referred to as actions). The proposed system recognizes learned activities via trained InceptionV3 Model. Experimental results on two human activity recognition benchmarks show that the proposed recognition system outperforms various detctions.

ABSTRACT:

In the last few Decades surveillance cameras also known as CCTV(Closed Circuit Television) is used to view what are the activities happening on the Environment. In this developing environment CCTV cameras are not updated yet., Our idea is to make a CCTV cameras into next version by applying deep learning technologies. In this way violence, accidents happened on the environment will be detected by the CCTV camera and an alert message will be sent. By this way the chance for risk will be reduced and CCTV cameras will be well effectively used.

REQUIREMENTS:

Hardware requirements:

- CCTV camera

Software Requirements:

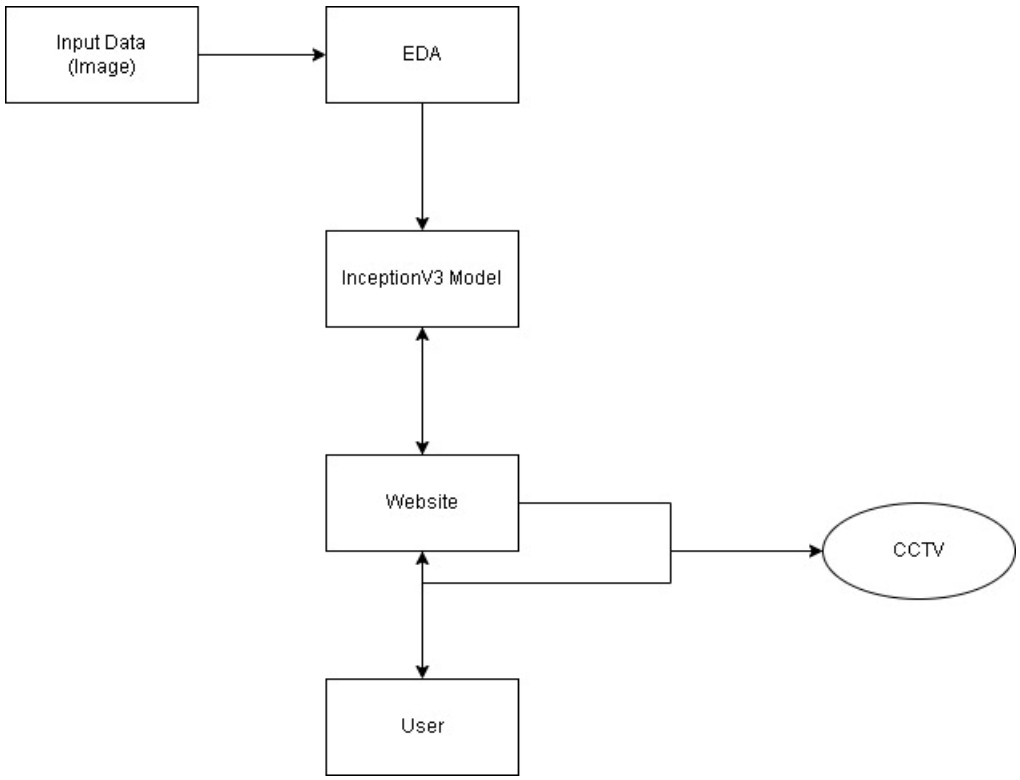
- Python
- OpenCV
- TensorFlow
- Django

SYSTEM DESIGN:

In this project we created a deep learning model which is constructed based on video data it consists of (Violence,Non-Violence,Accident) videos are trained by using InceptionV3Algorithm . The Algorithm works based on deep learning approach that stores image data in neural network construction and saves the data this process is known as training the Model. After training completes the model is tested and verify whether the model is detecting action are not. It is the First Phase of the project. Afterwords the Website is developed for the model the deep learning model runs in the background of the website continuously and detects whether any unknown actions on the Surroundings. If any violence activities are accident events occurs that event will be immediately detected by the model in the way of CCTV cameras. And an Alert message will be send to the user .Now we build a website that will accept a video as input and process it produce a labeled output video. Based on that output user can detect where the violence or

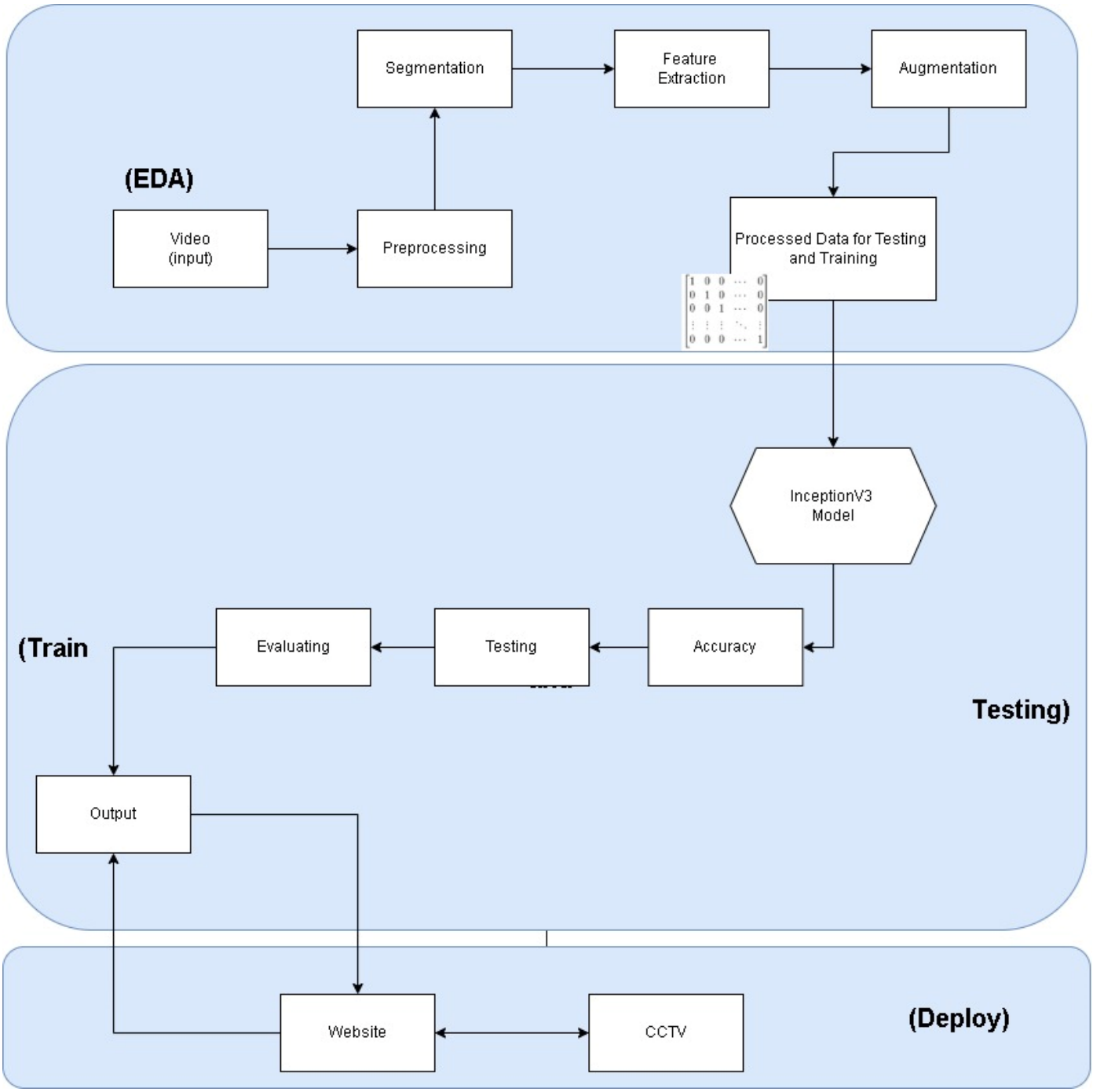
accident event occurs and increase the productivity gradually.

MODEL ARCHITECTURE:



The Model the constructed based on input images that consists of violence ,Nonviolence and fire images. And EDA process takes place on it, in EDA(pre-processing, segmentation, feature extraction) process takes place. After afterwords data is splitting for training and testing. And the processed for model training, Model training,labels are stored while model training. After model is trained then tested based on test data and validated. Now the accuracy of the model is calculated. The deployment phase starts here by deploy the trained model to a website that will run the model on the background and user can interact with it easily, by this way globally anyone can interact with the model and use it directly.

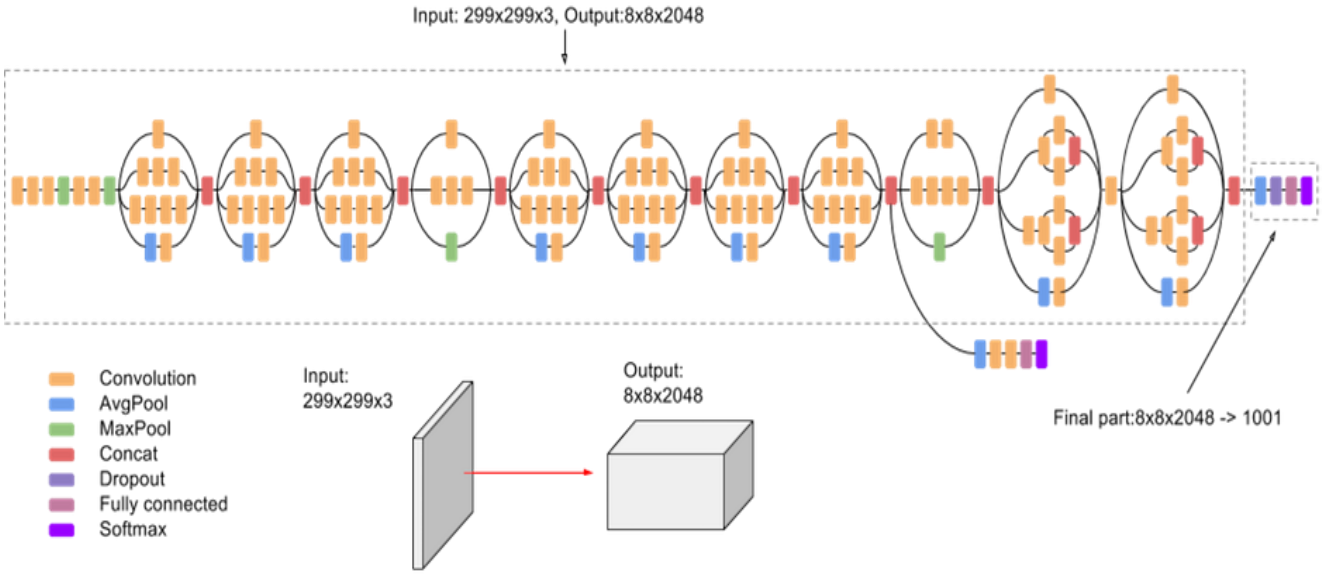
SYSTEM ARCHITECTURE:



ALGORITHM:

InceptionV3:

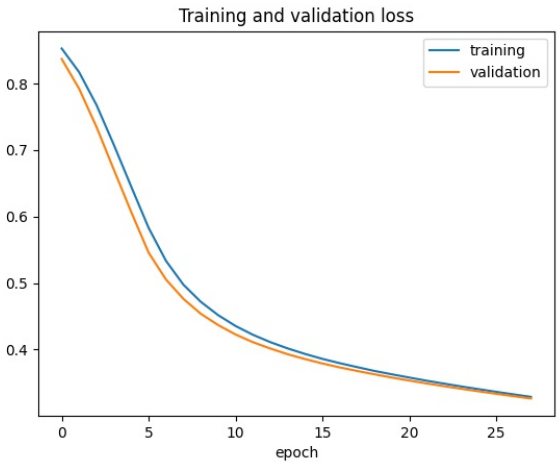
Inception v3 is an image recognition model that has been shown to attain greater than 78.1% accuracy on the ImageNet dataset. The model is the culmination of many ideas developed by multiple researchers over the years. It is based on the original paper: "Rethinking the Inception Architecture for Computer Vision" by Szegedy, et. al. The model itself is made up of symmetric and asymmetric building blocks, including convolutions, average pooling, max pooling, concatenations, dropouts, and fully connected layers. Batch normalization is used extensively throughout the model and applied to activation inputs. Loss is computed using Softmax. A high-level diagram of the model is shown in the following screenshot:

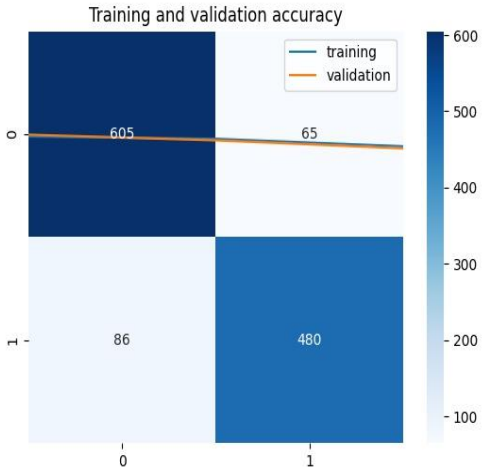


In comparison to VGGNet, Inception Networks (GoogLeNet/Inception v1) have proved to be more computationally efficient, both in terms of the number of parameters generated by the network and the economical cost incurred (memory and other resources). If any changes are to be made to an Inception Network, care needs to be taken to make sure that the computational advantages aren't lost. Thus, the adaptation of an Inception network for different use cases turns out to be a problem due to the uncertainty of the new network's efficiency. In an Inception v3 model, **several techniques for optimizing the network have been put suggested to loosen the constraints for easier model adaptation. The techniques include factorized convolutions, regularization, dimension reduction, and parallelized computations.**

INPUT IMAGES:

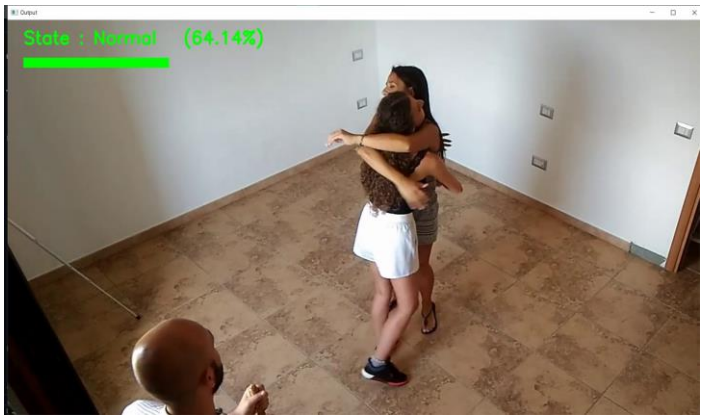
MODEL PERFORMANCE:



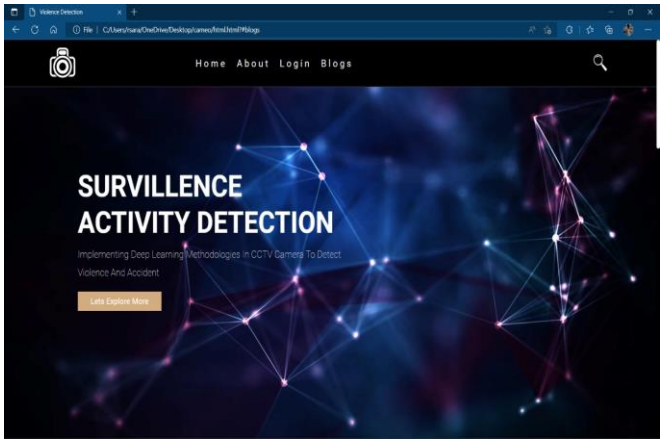


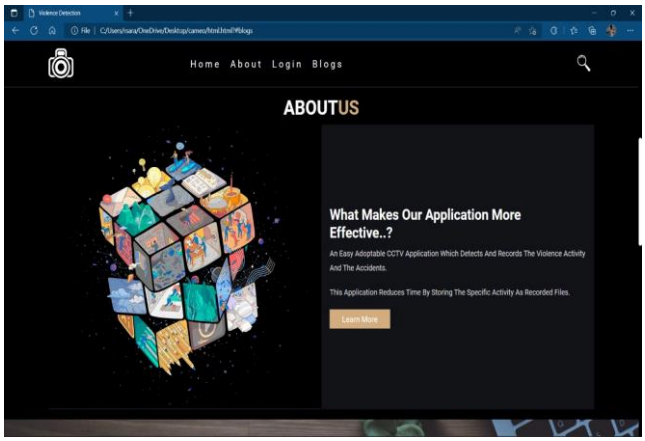
	precision	recall	f1-score	support
NonViolence	0.96	0.98	0.97	1196
Violence	0.97	0.95	0.96	898
accuracy			0.97	2094
macro avg	0.97	0.97	0.97	2094
weighted avg	0.97	0.97	0.97	2094

OUTPUT RESULTS:



WEBSITE:





ADVANTAGES OF THE SYSTEM:

- CCTV cameras were effectively used.
- Chance for damages to common people and properties will be

reduced.

- Productivity will be improved.

DISADVANTAGES OF THE SYSTEM:

- Good Quality Cameras are required.
- Need to maintain and update system properly.
- Video should be in good quality

CONCLUSION:

The Surveillance Activity detector is built with correct algorithm with efficient techniques, the prediction by reducing the size the image without losing the information needed for making predictions. In future optimization techniques can be applied so as to decide the number of layers and filters that can used in a model. As of now for the given dataset the InceptionV3Algorithm proves to be the better technique in predicting the presence of activities.

REFERENCE:

[1] Kavita V. Bhaltilak, Harleen Kaur, Cherry Khosla, "Human Motion Analysis with the Help of Video Surveillance: A Review," In the

International Journal of Computer Science Engineering and Technology

(IJCSET), Volume 4, Issue 9, pp. 245-249, September 2014.
[2] Chen Change Loy, "Activity Understanding and Unusual Event Detection in Surveillance Videos," PhD dissertation, Queen Mary University of London, 2010.
[3] Mao Ye, Qing Zhang, Liang Wang, Jiejie Zhu, Ruigang Yang, Juergen Gall, "A Survey on Human Motion Analysis from Depth Data," Lecture Notes in Computer Science, Springer Berlin Heidelberg, Volume 8200, pp 149-187, 2013.
[4] Lulu Chen, Hong Wei, James Ferryman, "A survey of human motion analysis using depth imagery," In Pattern Recognition Letters, Elsevier Science Inc., Volume 34, Issue 15, pp. 1995-2006, November 2013.
[5] Ronald Poppe, "A survey on vision-based human action recognition," In the International Journal of Image and Vision Computing, Volume 28, Number 6, pp.976-990, June 2010
[6] Maaike Johanna, "Recognizing activities with the Kinect," Master thesis, Radboud University Nijmegen, Nijmegen, Netherlands, July 2013.
[7] Ferda Ofli, Rizwan Chaudhry, Gregorij Kurillo, René Vidal, and Ruzena Bajcsy, "Sequence of the Most Informative Joints (SMIJ): A New Representation for Human Skeletal Action Recognition," In proceedings of the IEEE Computer Vision and Pattern Recognition Workshops (CVPRW), Providence, Rhode Island, USA, PP. 8-13, June 2012.