International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 03 | Mar 2020 www.irjet.net p-ISSN: 2395-0072

Using Convolutional Neural Network in Surveillance Videos for Recognizing Human Actions Based On Machine Learning in Examination

Sheetal Lole[1] Mrs. Indhra Muthuswamy[2], Ketki Kulkarni[3], Rutika Birajdar[4]

[1,3,4](Student, Dept. of Computer Engineering, G H Raisoni College of Engineering and Management, Maharashtra, India)
[2](Professor, Dept. of Computer Engineering, G H Raisoni College of Engineering and Management, Maharashtra, India)

Abstract - *Video analytics is the method of processing a video, gathering data and analysing the data for getting domain specific information. In the current trend, besides analyzing any video for information retrieval, analyzing live surveillance videos for detecting activities that take place in its coverage area has become more important. Such systems will be implemented real time. Automated face recognition from surveillance videos becomes easier while using a training model such as Artificial Neural Network. Hand detection is assisted by skin color estimation. This research work aims to detect suspicious activities such as object exchange, entry of a new person, peeping into other's answer sheet and person exchange from the video captured by a surveillance camera during examinations. Nowadays, people pay more attention to fairness of examination, so it is meaningful to detect abnormal behavior to ensure the order of examination. Most current methods propose models for particular cheating behavior. In this system, we extract the optical flow of video data and propose a 3D convolution neural networks model to deal with the problem. This requires the process of face recognition, hand recognition and detecting the contact between the face and hands of the same person and that among different persons. Automation of 'suspicious activity detection' will help decrease error rate due to manual monitoring.*

Keywords—*Video Surveillance, Anomaly detection, Artificial neural network based sparsity learning, suspicious activities.*

1. INTRODUCTION

Human face and human behavioural pattern play an important role in person identification. Visual information is a key source for such identifications. Surveillance videos provide such visual information which can be viewed as live videos, or it can be played back for future references. The recent trend of 'automation' has its impact even in the field of video analytics. Video analytics can be used for a wide variety of applications like motion detection, human activity prediction, person identification, abnormal activity recognition, vehicle counting, people counting at crowded places, etc. In this domain, the two factors which are used for person identification are technically termed as face recognition and gait recognition respectively. Among these two techniques, face recognition is more versatile for automated person identification through surveillance videos. Face recognition can be used to predict the orientation of a person's head, which in turn will help to predict a person's behavior. Motion recognition with face recognition is very useful in many applications such as verification of a person, identification of a person and detecting presence or absence of a person at a specific place and time. In addition, human interactions such as subtle contact among two individuals, head motion detection, hand gesture recognition and estimation are used to devise a system that can identify and recognize suspicious behavior among pupil in an examination hall successfully.

This system provides a methodology for suspicious human activity detection through face recognition. Video processing is used in two main domains such as security and research. Such a technology uses intelligent algorithms to monitor live videos. Computational complexities and time complexities are some of the key factors while designing a real-time system. The system which uses an algorithm with a relatively lower time complexity, using less hardware resources and which produces good results will be more useful for time-critical applications like bank robbery detection, patient monitoring system, detecting and reporting suspicious activities at the railway station, exam holes etc.

2. LITERATURE SURVEY

I. Bin Zhou,Li Fei-Fei,Eric P. Xing "Online Detection of Unusual Events in Videos via Dynamic Sparse Coding ". Author propose an improved Real-time unusual event detection in video stream has been a difficult challenge due to the lack of sufficient training information, volatility of the definitions for both normality and abnormality, time constraints, and statistical limitation of the fitness of any parametric models. They propose a fully unsupervised dynamic sparse coding approach for detecting unusual events in videos based on online sparse reconstructibility of query signals from an

atomically learned event dictionary, which forms sparse coding bases. Based on an intuition that usual events in a video are more likely to be constructible from an event dictionary, whereas unusual events are not, our algorithm employs a principled convex optimization formulation that allows both a sparse reconstruction code, and an online dictionary to be jointly inferred and updated. Our algorithm is completely unsupervised, making no prior assumptions of what unusual events may look like and the settings of the cameras. The fact that the bases dictionary is updated in an online fashion as the algorithm observes more data, avoids any issues with concept drift. Experimental results on hours of real world surveillance video and several YouTube videos show that the proposed algorithm could reliably locate the unusual events in the video sequence, outperforming the current state-of the art methods.

II. Mohammad Sabokrou, Mahmood Fathy, Mojtaba Hoseini, Reinhard Klette, "Real-Time Anomaly Detection and Localization in Crowded Scenes". In this paper, we propose a method for real-time anomaly detection and localization in crowded scenes. Each video is defined as a set of non-overlapping cubic patches, and is described using two local and global descriptors. These descriptors capture the video properties from different aspects. By incorporating simple and cost-effective Gaussian classifiers, we can distinguish normal activities and anomalies in videos. The local and global features are based on structure similarity between adjacent patches and the features learned in an unsupervised way, using a sparse auto encoder. Experimental results show that our algorithm is comparable to a state-of-the-art procedure on UCSD ped2 and UMN benchmarks, but even more time-efficient. The experiments confirm that our system can reliably detect and localize anomalies as soon as they happen in a video.

III. Cewu Lu. Jianping Shi, Jiaya Jia, "Abnormal Event Detection at 150 FPS in MATLAB". Speedy abnormal event detection meets the growing demand to process an enormous number of surveillance videos. Based on inherent redundancy of video structures, we propose an efficient sparse combination learning framework. It achieves decent performance in the detection phase without compromising result quality. The short running time is guaranteed because the new method effectively turns the original complicated problem to one in which only a few costless small-scale least square optimization steps are involved. Our method reaches high detection rates on benchmark datasets at a speed of 140150 frames per second on average when computing on an ordinary desktop PC using MATLAB.

IV. Mahmudul Hasan, Jonghyun Choiy, Jan Neumanny, Amit K. Roy-Chowdhury, Larry S. Davisz, "Learning Temporal Regularity in Video Sequences". Perceiving meaningful activities in a long video sequence is a challenging problem due to ambiguous definition of 'meaningfulness' as well as clutters in the scene. We approach this problem by learning a generative model for regular motion patterns (termed as regularity) using multiple sources with very limited supervision. Specifically, we propose two methods that are built upon the auto encoders for their ability to work with little to no supervision. They first leverage the conventional handcrafted spatio temporal local features and learn a fully connected auto encoder on them. Second, we build a fully convolutional feed-forward auto encoder to learn both the local features and the classifiers as an end-to-end learning framework. Our model can capture the regularities from multiple datasets. they evaluate our methods in both qualitative and quantitative ways - showing the learned regularity of videos in various aspects and demonstrating competitive performance on anomaly detection datasets as an application.

3. DESIGN & IMPLEMENTATION OF PROPOSED SYSTEM

In this system, we extract the optical flow of video data and propose a Convolution Neural Networks model to deal with the problem. The proposed system extracts the spatial and temporal features from video data and these features can be directly feed into the classifier for model learning or inference. The experiments on our own made dataset show that the proposed model achieves superior performance in comparison to current methods.

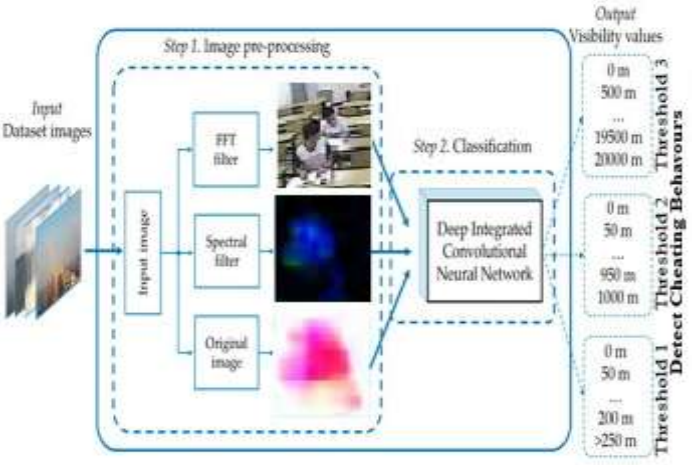

Fig.1: Proposed System Architecture

A. Abnormal Behavior

The abnormal behavior can be identified as irregular behavior from normal one. During examination, we pay more attention to the abnormal behaviors which mainly include leaning, reaching out hand, turning around, entering and leaving the classroom halfway. If there exist frequent abnormal behaviors in a period, it is very likely that some problems appear in the examination. For example, if students can't hear English listening clearly due to equipment problem, they would lean and turn around to confirm the problem. Meanwhile, the electronic monitor would detect the abnormal behaviors and notify the supervisor to deal with the problem in time. So it is essential and significant to detect abnormal behavior in examination surveillance video.

B. Our 3D CNN Model

We learn that 3D CNN has the ability to extract spatial and temporal information of video clips. In this paper, we proposed a new C3D model for detecting abnormal behavior in examination surveillance video. Table I shows the architecture of our 3D CNN model. As our CNN model is a binary classification, the precision and recall rate are the common evaluation indexes. We compare our methods with other methods. Our method1 adopts the first method to obtain "flow image", and our method2 adopts the second method to obtain "flow image". The experiments demonstrate that our methods have a better performance in comparison to other methods. Our method2 is better in accuracy and precision, and our method1 is better in recall rate. To summarize, our model has the ability to deal with many kinds of abnormal behaviors and performs better in comparison to current methods. There are four kinds of behaviors detected. If prediction of the sub-region sample is positive, it will be marked in corresponding testing clip by red box, and system would store the testing clip then.

4. RESULT & DISCATION

The result for proposed system is to identify Human action using image processing (CNN), detect & identification of Abnormal Behavior in Examination section and study extract the optical flow of video data and propose a convolution neural networks model to deal with the problem.

methods	Accuracy	Precision	Recall rate
Motion blob[4]	-	-	82%
Template matching[5]	-	86%	-
Skin+SVM[7]	84%	-	-
Our method1	87.6%	80.4%	**84.3%**
Our method2	**89.8%**	**86.5%**	83.2%

Fig.2: Performance of Different Methods

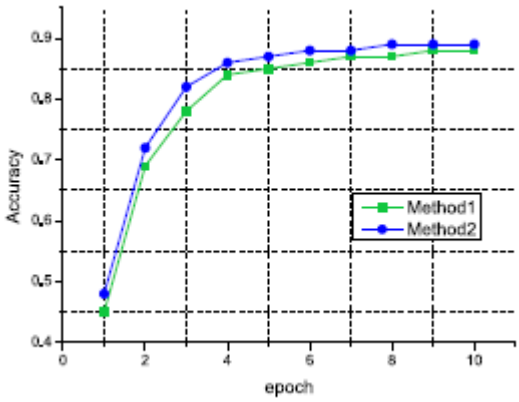

Fig.3: The performance of two "flow image" on 3D CNN

5. CONCLUSION

We propose a unified deep learning based framework for abnormal event detection from exam hall. The proposed system consists of three blocks which are designed to achieve three keys of abnormal detection in neural networks. In short, the motion fusion block is designed to keep the temporal and spatial connection between the motion and appearance cues. The feature transfer block is used to extract discriminative features by exploiting the transferability of the neural network from different tasks/ domains. The coding block is a novel LSTM to achieve fast sparse coding, which could enjoy fast inference and end-to-end learning. Extensive experiments show the promising performance of our method in image reconstruction and abnormal events detection in surveillance.

Acknowledgment

I would prefer to give thanks the researchers likewise publishers for creating their resources available. I'm conjointly grateful to guide, reviewer for their valuable suggestions and also thank the college authorities for providing the required infrastructure and support.

References

[1] C. Lu, J. Shi, and J. Jia, "Abnormal event detection at 150 fps in matlab," in Proceedings of the IEEE international conference on computer vision, 2013, pp. 2720–2727.

[2] M. Sabokrou, M. Fathy, M. Hoseini, and R. Klette, "Real-time anomaly detection and localization in crowded scenes," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2015

[3] M. Hasan, J. Choi, J. Neumann, A. K. Roy-Chowdhury, and L. S. Davis, "Learning temporal regularity in video sequences," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 733–742.

[4] B. Zhao, L. Fei-Fei, and E. P. Xing, "Online detection of unusual events in videos via dynamic sparse coding," in Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011, pp. 3313– 3320

[5] W. Luo,W. Liu, and S. Gao, "A revisit of sparse coding based anomaly detection in stacked rnn framework," in The IEEE International Conference on Computer Vision (ICCV), Oct 2017.

[6] S. Wu, B. E. Moore, and M. Shah, "Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes," in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010, pp. 2054–2060.

[7] V. Mahadevan,W. Li, V. Bhalodia, and N. Vasconcelos, "Anomaly detection in crowded scenes," in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010, pp. 1975–1981.

[8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol.1. IEEE, 2005, pp. 886–893.

[9] L. Kratz and K. Nishino, "Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models," in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009, pp. 1446–1453.

[10] N. Dalal, B. Triggs, and C. Schmid, "Human detection using oriented histograms of flow and appearance," in European conference on computer vision Springer, 2006, pp. 428–441. G