ATM BEHAVIOUR ACTIVITY DETECTION

Komal Desai, Shreelekha Jagtap, Khushi Vishwakarma, Anjali Gupta
Siddhant college of Engineering
Sudumbare,
Department of computer Engineering

---***---

Abstract - ATM behavior Activity is predicting the body part or joint locations of a person from an image or a video. This project will entail detecting ATM behavior Activity from real-time CCTV footage using neural networks. Human ATM behavior Activity is one of the key problems in computer vision that has been studied for more than 15 years. It is important because of the sheer number of applications which can benefit from Activity detection. For example, human pose estimation is used in applications including video surveillance, animal tracking and behavior understanding, sign language detection, advanced human-computer interaction, and marker less motion capturing. Low-cost depth sensors have limitations like limited to indoor use, and their low resolution and noisy depth information make it difficult to estimate human poses from depth images. Hence, we plan to use neural networks to overcome these problems. ATM behavior human activity recognition from surveillance video is an active research area of image processing and computer vision. Through the visual surveillance, human activities can be monitored in sensitive and public areas such as bus stations, railway stations, airports, banks, shopping malls, school and colleges, parking lots, roads, etc. to prevent terrorism, theft, accidents and illegal parking, vandalism, fighting, chain snatching, crime and other ATM behavior activities. It is very difficult to watch public places continuously, therefore an intelligent video surveillance is required that can monitor the human activities in real-time and categorize them as usual and unusual activities; and can generate an alert.

.

1.INTRODUCTION

We plan to build an application for detection of ATM behavior activity of people in public places in real time. Our application can be used in surveillance at places like malls, airports, railway stations, etc. where there is a risk of robbery or a shooting attack. We will be using deep learning and neural networks to train our system. This model will then be deployed as a mobile and desktop app which will take real time CCTV footage as input and send an alert on the administrator's device if some ATM behavior pose is found. Human ATM behavior activity is related to identifying human body parts and possibly tracking their movements. Real life applications of it vary from gaming to AR/VR, to healthcare and gesture recognition. Compared to image data domain, there is relatively little work on applying CNNs to video classification. This is because, a video is more complex than images since it has another dimension - temporal. Unsupervised learning exploits temporal dependencies between frames and has proven successful for video analysis. Some ATM behavior activity approaches use CPU instead of GPU so that ATM behavior activity can run-on low-cost hardware like embedded systems and mobile phones. Low-cost depth sensors are another new technology in computer vision. They are present in gaming consoles like the Kinect for Xbox 360. They are motion sensors which allow the user to interact with the console without a game controller, through just hand gestures. These are RGB-D sensors that obtain depth information by structured light technology. The structured light sensors infer the depth values by projecting an infrared light pattern onto a scene and analyzing the distortion of the projected light pattern. However, these sensors are limited to indoor use, and their low resolution and noisy depth information make it difficult to estimate human poses from depth images.

2. LITERATURE SURVEY

2.1 STUDY OF RESEARCH PAPER

1. Paper Name: Real-Time ATM behavior Detection and Localization in Crowded Scenes

Author: Mohammad Sabokrou, Mahmood Fathy

Abstract: In this paper, we propose a method for real-time ATM behavior 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 autoencoder. 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.

2. Paper Name: Learning Temporal Regularity in Video Sequences

Author: Mahmudul Hasan Jongh Yun Choi

Abstract: 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 autoencoders for their ability to work with little to no supervision. We first leverage the conventional handcrafted spatio-temporal local features and learn a fully connected autoencoder on them. Second, we build a fully convolutional feed-forward autoencoder 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. We evaluate our methods in both qualitative and quantitative ways - showing the learned regularity of videos in various aspects and demonstrating competitive performance on ATM behavior detection datasets as an application.

3. Paper Name: ATM behavior Detection in Video Using Predictive Convolutional Long Short-Term Memory Networks

Author: Jefferson Ryan Medal

Description: Automating the detection of anomalous events within long video sequences is challenging due to the ambiguity of how such events are defined. We approach the problem by learning generative models that can identify anomalies in videos using limited supervision. We propose end-to-end trainable composite Convolutional Long Short-Term Memory (Conv-LSTM) networks that are able to predict the evolution of a video sequence from a small number of input frames. Regularity scores are derived from the reconstruction errors of a set of predictions with abnormal video sequences yielding lower regularity scores as they diverge further from the actual sequence over time. The models utilize a composite structure and examine the effects of 'conditioning' in learning more meaningful representations. The best model is chosen based on the reconstruction and prediction accuracies. The Conv-LSTM models are evaluated both qualitatively and quantitatively, demonstrating competitive results on ATM behavior detection datasets. Conv-LSTM units are shown to be an effective tool for modelling and predicting video sequences.

4. Paper Name: Abnormal Event Detection in Videos using Spatiotemporal Autoencoder

Author: Yong Shean Chong

Description: We present an efficient method for detecting anomalies in videos. Recent applications of convolutional neural networks have shown promises of convolutional layers for object detection and recognition,

especially in images. However, convolutional neural networks are supervised and require labels as learning signals. We propose a spatiotemporal architecture for ATM behavior detection in videos including crowded scenes. Our architecture includes two main components, one for spatial feature representation, and one for learning the temporal evolution of the spatial features. Experimental results on Avenue, Subway and UCSD benchmarks confirm that the detection accuracy of our method is comparable to state-of the-art methods at a considerable speed of up to 140 fps.

5. paper Name: A Revisit of Sparse Coding Based ATM behavior Detection in Stacked RNN Framework

Author: Weixin Luo

Abstract: Motivated by the capability of sparse coding-based ATM behavior detection, we propose a Temporally-coherent Sparse Coding (TSC) where we enforce similar neighboring frames be encoded with similar reconstruction coefficients. Then we map the TSC with a special type of stacked Recurrent Neural Network (sRNN). By taking advantage of sRNN in learning all parameters simultaneously, the nontrivial hyper-parameter selection to TSC can be avoided, meanwhile with a shallow sRNN, the reconstruction coefficients can be inferred within a forward pass, which reduces the computational cost for learning sparse coefficients. The contributions of this paper are two-fold: i) We propose a TSC, which can be mapped to a sRNN which facilitates the parameter optimization and accelerates the ATM behavior prediction. ii) We build a very large dataset which is even larger than the summation of all existing dataset for ATM behavior detection in terms of both the volume of data and the diversity of scenes. Extensive experiments on both a toy dataset and real datasets demonstrate that our TSC based and sRNN based method consistently outperform existing methods, which validates the effectiveness of our method

3. ARCHITECTURE DIAGRAM

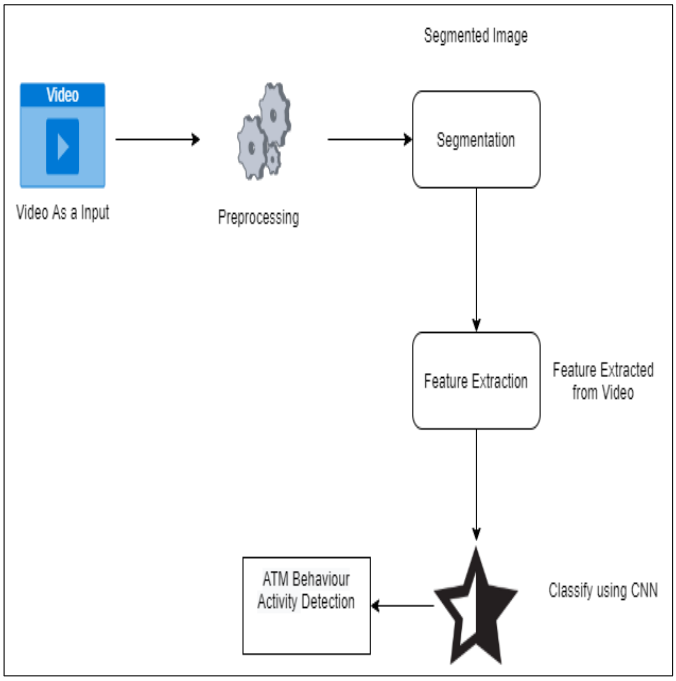


Figure 5.1: System

Fig -1: Figure

4. Objective

- Detect the ATM behavior activity
- Anyone can use this application easy.

5. EXTERNAL INTERFACE REQUIREMENT

User Interface

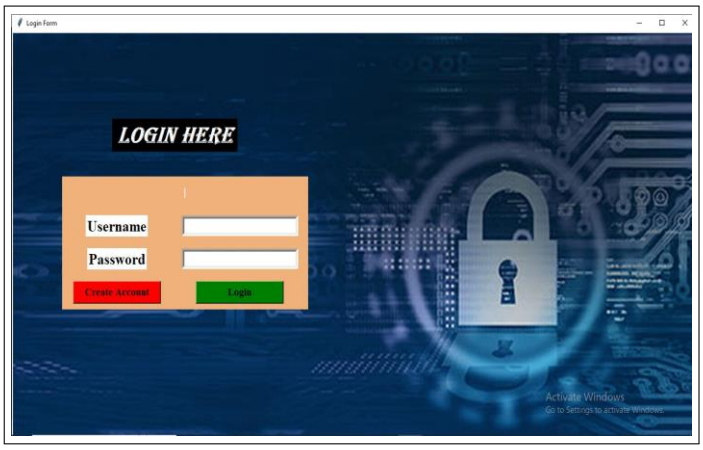
- **Application Based On ATM behavior Activity Detection.**

Hardware Interfaces

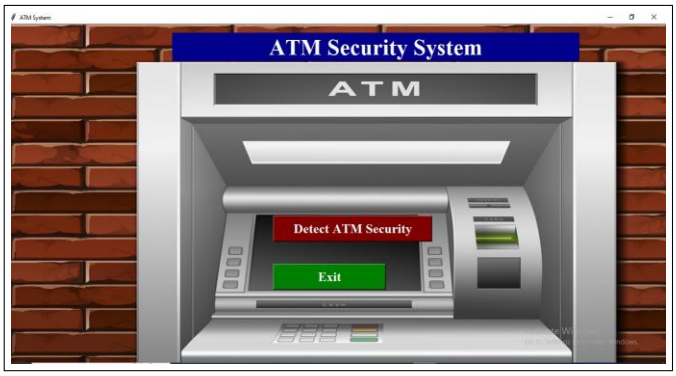
- **Hardware:** intel core
- **Speed:** 2.80 GHz
- **RAM:** 8GB
- **Hard Disk:** 40 GB
- **Key Board:** Standard Windows Keyboard

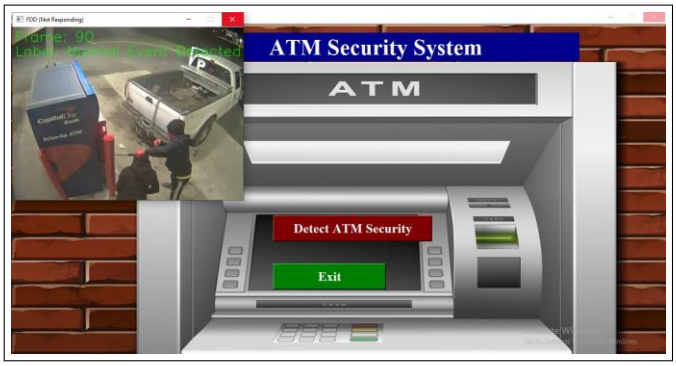
Software Interfaces

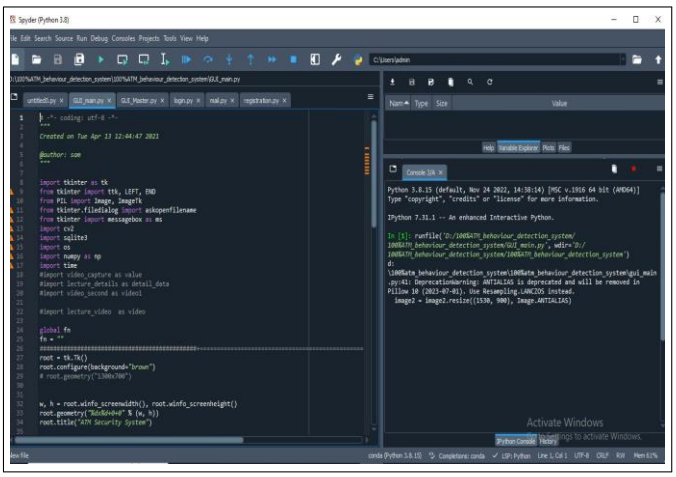
- **Operating System:** Windows 10
- **IDE:** Anaconda
- **Programming Language:** Spyder



6. RESULT







7.CONCLUSIONS

A system to process real-time CCTV footage to detect any ATM behavior activity will help to create better security and less human intervention. Great strides have been made in the field of human ATM behavior Activity, which enables us to better serve the myriad applications that are possible with it. Moreover, research in related fields such as Activity Tracking can greatly enhance its productive utilization in several fields.

8.. REFERENCES

1 Eralda Nishani, Betim Cico: "Computer Vision Approaches based on Deep Learning and Neural Networks" Deep Neural Networks for Video Analysis of Human Pose Estimation- 2017 6th MEDITERRANEAN CONFERENCE ON EMBEDDED COMPUTING (MECO), 11-15 JUNE 2017, BAR, MONTENEGRO

2 Naimat Ullah Khan, Wanggen Wan: "A Review of Human Pose Estimation from Single Image"- 978-1-5386-5195-7/18/ 2018 IEEE

3 Qiuhui Chen, Chongyang Zhang, Weiwei Liu, and Dan Wang," Surveillance Human Pose Dataset And Performance Evaluation For Coarse-Grained Pose Estimation", Athens 2018.