Smart surveillance advance system using machine learning and CNN to learn suspicious activity

Nayyab Kanwal

University of Wolverhampton - University of Wolverhampton
*Corresponding Author: Nayyab Kanwal. Email: N.kanwal@wlv.ac.uk

Abstract: Nowadays we can find security cameras everywhere, from roads to malls, we are under surveillance. A lot of people install security cameras in their houses and the sole reason for installing security cameras is safety. Everyone is worried about the safety of someone, some people are worried about themselves, some about their family, and some are concerned about the safety of their friends. In short, everyone cares about being safe. But how about increasing the security up a notch? How about people installing software with their security cameras that will automatically detect suspicious activities and report authorities? The primary goal of this Research is to develop a system that is capable to detect any anomalous and suspicious activities from given video frames captured by surveillance cameras and reporting them to authorized personnel for immediate actions. It is very common in surveillance that anomalies go unnoticed by the guards on the spot. Thousands of cameras are installed on streets and roads that record everything, but that recording is only used later because no one was watching a live video stream at that exact moment. But what if the software is always watching these thousands of live video streams? This means that every anomalous activity or crime that occurs and is recorded by these cameras will be reported instantly. This software will use an advanced Machine learning model to recognize suspicious activities and report them to proper authorities.

Keywords: Anomaly Detection, Surveillance System, Machine Learning, Real-time Monitoring

1 Introduction

In 1888, a French inventor shot the world's first video. Nowadays shooting a video is nothing. Dozens of devices can shoot a video. Some devices are just made for making videos while others have this feature with several others. In the present day, everyone wants everything to be saved so that they can see it later. If a person witnesses a fight between two people, they will use their mobile to record a video of the fight to share it later with friends or anyone else. Some people might record a fight so that they give it to the police later, a lot of big corporations install cameras on their premises so that their premises are under surveillance, and if anything like a robbery happens, they have video proof of the robbery or fight. Almost everywhere in public, we are under surveillance. There are cameras either hidden or in plain sight. From malls to roads, there are cameras installed.

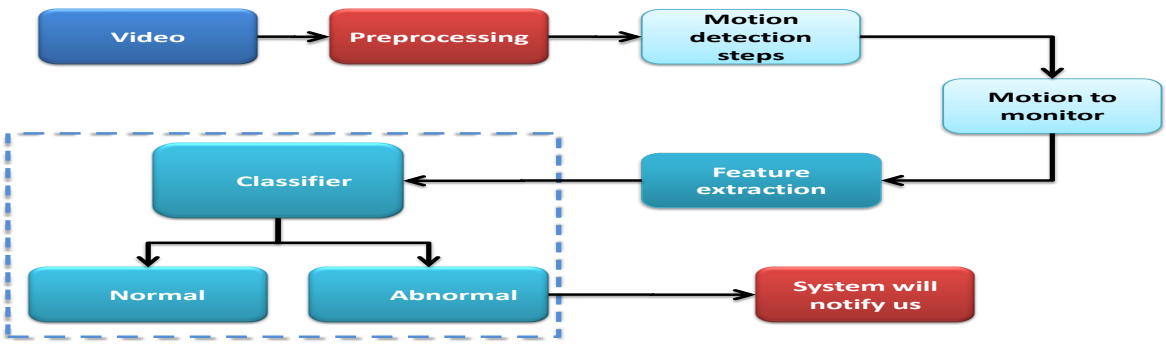
The reason for installing these cameras is safety. A thief might not rob a house or shop if there are cameras installed there. A lot of people install cameras in their houses for the same reason. As technology advances, people invent newer and better ways to keep the public safe. Security cameras were introduced for the same reason. One might feel safe leaving their house at night if they have security cameras and alarms installed. The owner of the jewelry shop will install security cameras in their jewelry store to keep their store safe from robbers. Malls install security cameras to monitor everyone and if anyone is behaving suspiciously,

they can report that person or monitor them and keep the mall safe. Cameras are installed with traffic lights so that people will be less likely to break traffic rules. In short, monitoring has made the world a safer place. People tend to break laws and rules when they think that no one can see them. Cameras and monitoring have discouraged this type, everyone knows that a camera might catch them in act of breaking the law or rule, so they won't try to break it.

Security cameras always record everything. Their recorded videos are stored somewhere. We can still see surveillance videos that were recorded 50 years ago. These videos are saved forever. But these videos are just what their names suggest videos, which can be seen later. If a camera is recording something and no one is watching the live stream of the camera, then the only use of the recorded video will be to analyze it later. Live streams still depend on people like security personnel to be useful at the time of suspicious activity. That security personnel will see the live footage and then report it to their senior or the department. If no one is watching the live stream from security cameras, then the only benefit of the footage is to watch it later.

With technology advancing, we want technology that makes life easier for us. We want systems that will act as humans so that we can focus on something more productive. We should have systems that will automate things that can be tedious for us. Watching live feed is something tedious but most of the places still have no automated systems for it. This is what we want to develop, an automated system that will make detect anything suspicious from the live stream of security surveillance cameras. The system can distinguish normal activities from suspicious activities and report them to the correct authorities.

The topic of this study will automate tedious work as monitoring live streams at malls. Malls and different stores hire security personnel whose job is to just watch a live stream from security cameras. This job can become tedious which can lead to negligence because 95% of the time, there is nothing to look at, which can lead to the person neglecting the live stream. Now, this negligence might cost as the person might neglect the stream at a moment when there is some suspicious activity occurring. On the other hand, a system can work 24/7. This can lead to automated systems which will reduce costs for organizations that they could utilize somewhere productive. This product will give them 24/7 surveillance and whenever any anomaly would occur, they would be notified instantly so that appropriate action could be taken immediately.



Objectives

The goal of this Research is to develop software that will detect activities and then separate activities which are normal and activities that are anomalous in any way. These anomalous activities then will be sent to their respective department.

The software will constantly analyze video streams from security cameras and will look for anomalous activities.

- The software will recognize activities that are anomalous and which activities are not.
- If there is an anomalous activity detected, then it will be reported to proper authorities.
- The software will have a location of each security camera and will send the location of the camera on which the video of anomaly activity was captured.
- Users can flag and change the type of a report in case the model predicts a wrong type
- The software will also cut a segment of a live stream in which suspicious activity is occurring and then send the video to the authorities as well.
- Both location and video segments will be sent together along with the type of anomaly (i.e. abuse or violence or road accident, etc.)
- The software will also save these clips in a repository so that they can be viewed later.

Problem Statement

The general idea of our Research is that this Research will help the organizations like Punjab Safe City Authority (PSCA) which is a Pakistani autonomous government body whose purpose is to make Punjab and soon the whole of Pakistan safer from criminal activity. Organizations like these have access to all the cameras installed in a specific region which is the whole Punjab in the case of PSCA. All these cameras are constantly monitoring the streets of Punjab and in all fairness, it is quite difficult, costly, and impractical to hire personnel to watch the live stream from each camera. This means these cameras are pretty much useless for watching criminal activity in real-time. Although these cameras can be quite helpful to watch the videos of criminal activities later for investigation or any other reason, they are not quite helpful for detecting crime in real-time.

Following problems are being faced currently all around the world:

Lack of personnel or budget to hire personnel to watch live streams.

Crime is increasing despite securities cameras installed as criminals think that no one is watching live footage in real-time and later on, the authorities won't recognize us

It is really important to solve these problems as the increase in crime means people won't believe in their government and the authorities created by the government to protect these people. Our software proposes the most cost-efficient and time-efficient solution. Our software will automatically analyze the live streams from the security cameras and will automatically detect any anomaly and then also report it to proper authorities within seconds. [9]

Assumptions and constraints

The authorities who will be connected to the main server of our software will have relatively better and consistent internet service as their systems will be connected to the software through the internet. It is assumed that the admin(s) who is maintaining the software will make more storage for anomalous reports if it is running low on storage. [5]

A major constrain of our Research is low-resolution cameras installed which will be connected to our system. These low-resolution cameras can reduce the accuracy of the system as it becomes difficult for the software to detect anomalies with lower-resolution footage. Another constraint which authorities can face is internet disconnection, as the authorities will be connected to our system via a web application, internet disconnection can be dangerous as the system of getting the report of the anomaly relies on the civilians again.

Research Scope

- The ultimate goal is to make an automated traffic surveillance system
- The system would detect anomalous activities happening on the road
- It will detect 8 types of anomalies
- It will then send the report which would consist of video clip, date and time, and location to the relevant users in relevant departments
- All the data and core activities would be done at the backend python implemented server
- The users can flag a report to be of any other type if the model predicts incorrectly
- The admin can add or delete more users and cameras
- These features would enable departments to take suitable action as soon as the anomaly occurs
- The task is to make the backend server capable of linking everything between the database and the web app on live time
- The cost would be minimal as the cameras are already installed on the roads and are just needed to be linked with the server
- All of this would eventually help the Safe city Research to move on to better and automated surveillance

Literature Review

Nowadays, from washing a car to predicting the survivability of another planet, everything is automated and is done by computers and complicated algorithms. These algorithms have different ways to predict, and each algorithm differs from other algorithms. One algorithm might have more parameters and another algorithm might use more pooling layers. But at the end of the day, all these models are used to predict an outcome from the given input. [7]

Our Research aims to detect a type of event to which authorities like police or ambulance or fire brigade reacts. For example, if there is a road accident, our software should report it instantly to the police and hospital so that they can act accordingly. The same is with arson, if there is arson somewhere, our software should report fire brigade and hospital. This explanation raises a question, and this question is, how the software will know about a dangerous event like a road accident. The answer to this question is security cameras, which have been installed by the government, but they do not have the budget to hire dozens and dozens of personnel to continuously monitor these cameras in case of some dangerous activity. This software will watch all the live streams 24/7 and will detect any anomalous event which it has been trained to detect. [10]

Another question arises with the previous explanation, and that question is, how the software will know that there is some anomalous activity in the live stream. This will be done by training a machine learning model to detect these anomalies. Our model will be trained with the online available dataset UCF_Crime [1], which provides hundreds of videos of activities like road accidents, shootings, vandalism, burglary, explosion, etc. After training our model, it will be able to detect anomalies correctly. For the model itself, we have used the C3D algorithm which is one of the most efficient algorithms when training and testing are related to videos. [8]

Our model will be trained on the UCF_Crime dataset and a separate dataset that will only contain normal videos. Normal videos mean exactly what the name suggests, it is a dataset containing videos in which there are no anomalous events, just daily life work, life people sitting around or walking on a street or driving a car, etc. These videos will make sure that our model does not always detect some anomaly because if a model is only trained on anomalous videos, it will try to detect an anomaly even though there would be none in the video. The model will predict anomalies correctly and will not constantly detect anomalies even if there is no anomaly.[5]

Our model will have a backend database on which all the previous anomalies will be stored as a report. This report will contain the video, type of anomaly, location of anomaly the department it was sent to. This report will be stored after sending the anomaly with all the information to the respective department. [6] The front-end interface of our software will be a web app that will have options like watching a live stream, seeing previously-stored reports, etc. All these requests will be verified through the server before giving access to the user. The admin interface will have options like add or delete videos from the database, add more users in the login authentication system of the software, which is connected to the database, add more cameras, etc. When an anomaly is detected by the model, a notification will be sent to the users who should be notified of the situation.

Detecting Pickpockets

This study shows the working and methodology of software that can detect pick pocketing through security cameras. The camera will automatically detect if a pickpocket is pick pocketing someone and then report it to security so that they can stop the pickpocket. This study validated their model by identifying 20 different scenarios of pick pocketing which was successfully detected by the model. This study showed that pickpockets can be detected easily with low failure rates [2]

Crime Detection

This study focused on detecting crime from a video stream. The characteristics used for detecting crime in this model were the movement and size of the object moving and the distance between two objects. The distance was used to determine that if the distance between two objects is closing and one object is moving suspiciously, then it can raise a red flag that a crime might happen. The model designed in this study can differentiate different human behaviors and analyze them to check any criminal activity. [3]

Dangerous Motion Detection

This study focused on dangerous motion and movement in big crowds that gather in big events like concerts or marriages. The system explained in this study detected dangerous movements in a crowd by recognizing sudden movement patterns and the flow of the crowd to distinguish individuals (s) who are moving suspiciously. This study also shows how their system detects sudden movement of crowd and change of flow of the crowd. [4]

Scanner.ia

This is a fully developed software from a Czech company that detects sudden motion changes using the trajectory change of an object. This software works on live video streams from cameras which can detect crime activities by detecting the movement of objects and dangerous objects that an object might possess. This software features functionality like if two people are chasing each other, this is done by analyzing the movement patterns of objects and how fast they are moving through frames. Sudden movement change like a person falling or running by tracking their trajectory movement and how it changes. [5]Software Development Life Cycle Model

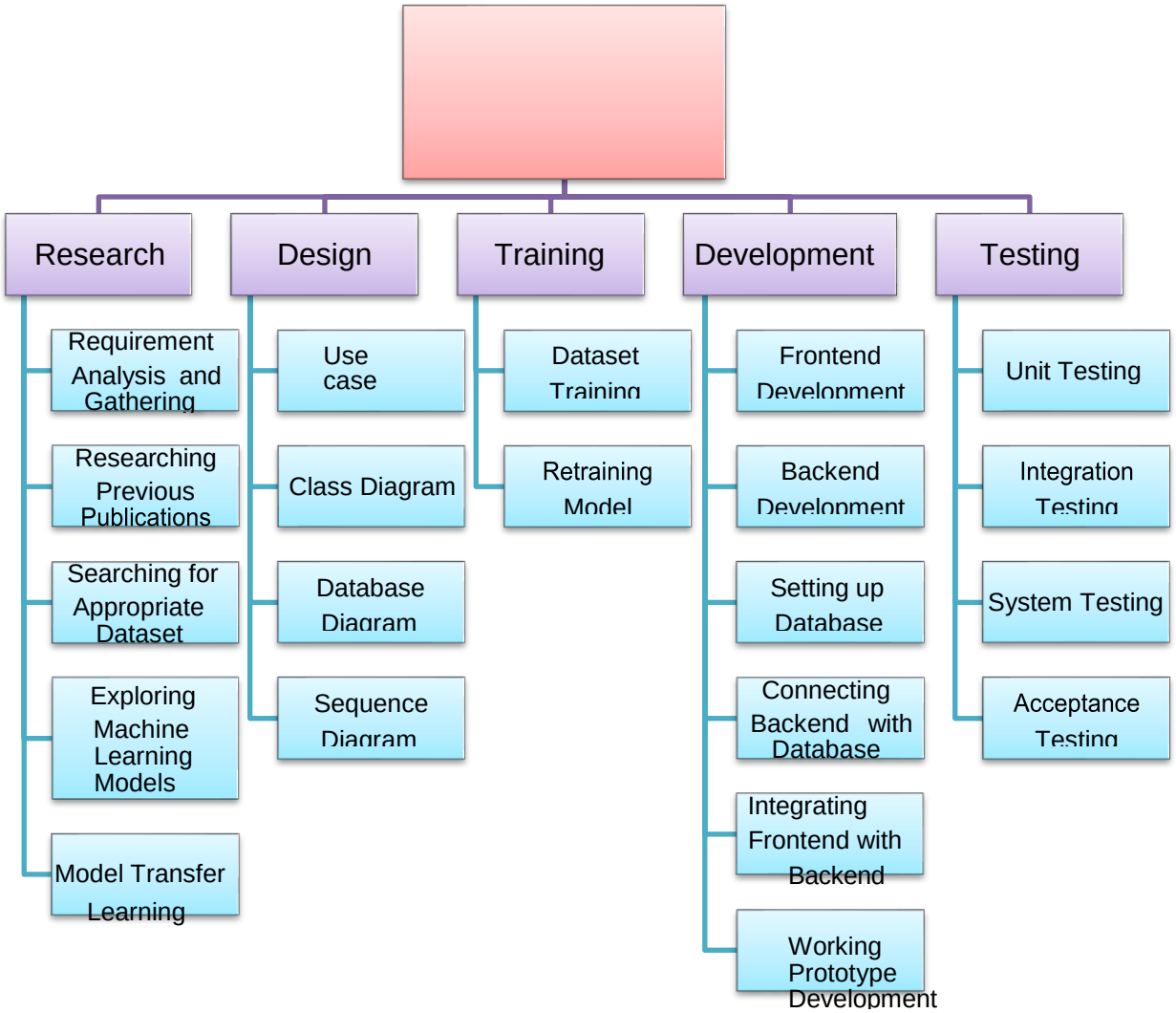
The SDLC model which has been used during the previous development and will be used in further development is an iterative model. We selected an iterative model because we wanted to test everything no matter how small it is. Implementation will be done in phases, and in each phase, a new feature will be implemented and tested. If later in the development, the software crashes due to some new features, we can always roll back to previously working implementation. Or if we want to resolve the problem of crashing, we will know that the crash was caused by the new implementation as previous modules were working completely fine. This frequent testing of the implementation also meant that if we made a mistake later, we would know where to roll back as we would know how many modules ago, our code was working completely

Fine. We can easily roll back to the correct working iteration and then try to implement the same thing which we failed last time, but this time with a new and better approach. [7]

Methodology

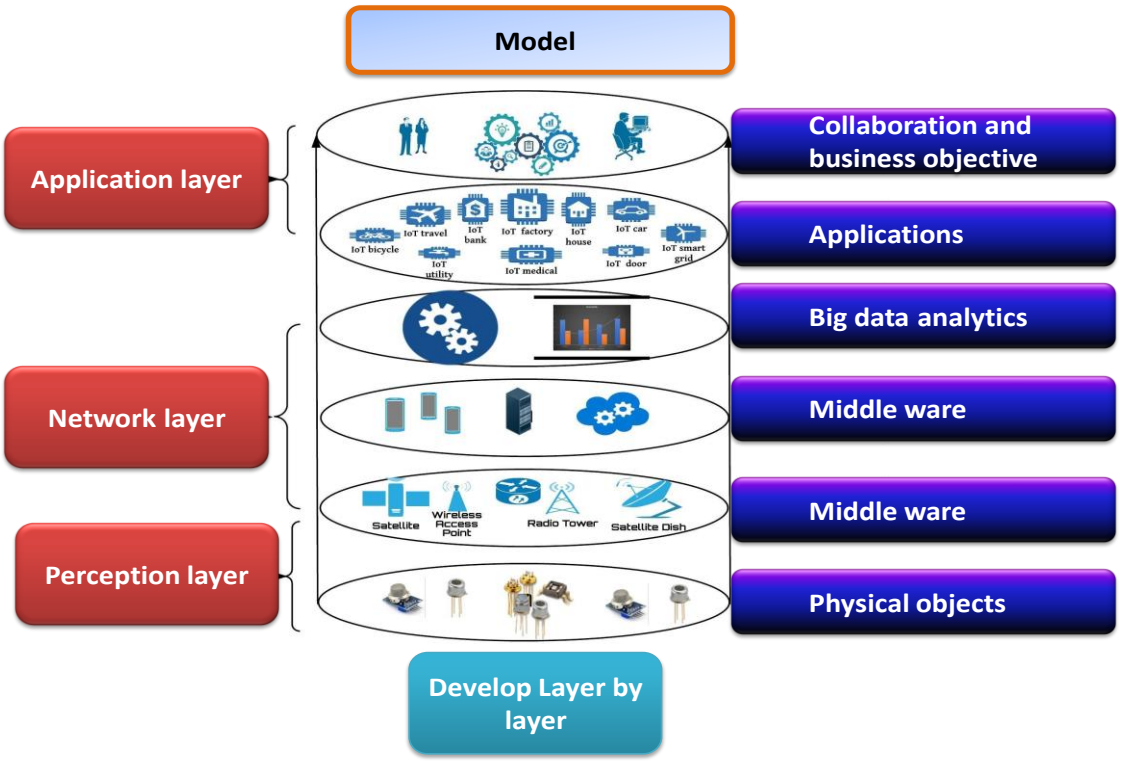
Work Breakdown Structure (WBS)
The Work Breakdown Structure for the system is displayed in Figure 4.

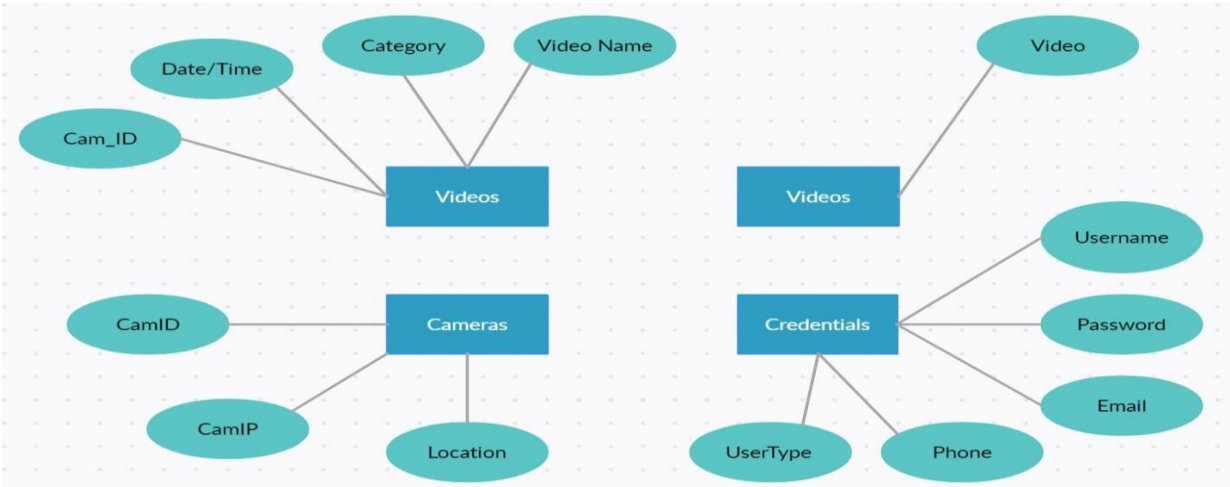
Figure 4 Work Breakdown Structure WBS



A work breakdown structure diagram of Smart surveillance system which is shown in figure# 4, is used to show the hierarchy of work done required to complete a Research. Work done is shown in tasks. For this Research, the first task will be divided into five parts: The research part of this Research, the Design part of this Research, the Training part of this Research, the actual development part of this Research, and the

Testing part of this Research. The first task is further divided into sub-tasks showing the functionalities required to complete the whole Research.





The database Diagram is shown in figure 9. The database being used is Google firebase. It contains separate tables for cameras, users, and records, and in its storage are the record videos.

TC-17) Admin clicks delete button against a user is displayed in Table 20.

Unit Testing

Unit Testing is done by testing each unit separately. Login is tested separately, Live stream is tested separately, retrain model is tested separately and similarly all the other unit features are tested separately in unit testing.

Integration Testing

This is done once each unit was tested, they were integrated and once that was done, it was checked that whether they performed with each other as expected and every feature would be tested with each other, and minor bugs or errors arise during the integration testing which was eliminated.

Acceptance Testing

Our Research will be tested on small scale with some departments and with very few cameras and once the Research is accepted on that scale and it performs according to the expectations and satisfies the users, our acceptance testing will be completed.

Conclusion

There were some problems that we faced during this Research. Those were as follows. First of all, while training our model C3D, we realized that we would need a GPU to train our model otherwise if we would have trained on our CPU it would have taken a lot more time. So our systems did not have GPU so then we used the Google Colab service which allots free available resources to the users and we then trained our model there so we learned that we should have known earlier that we would need a GPU to train the model because a lot of our time was wasted on trying to train on our CPU which would stop as soon as any power failure occurred.

Another problem that we faced was that when we decided to train on Google Colab, we were to upload our dataset to google drive and we started to upload small parts of our dataset on the drive and started training on that but as soon as the data increased, we realized that google drive allows uploading only 15GB of data for free and then more storage has to be purchased which also halted our training process so we learned that we should always check the amount of storage available and whether it would be sufficient for our use before actually using it.

Research Summary

This Research is intended to be a complete package for traffic surveillance. It would detect any anomaly or crime committed on the road and would send the report that would consist of the video clip of the event and the camera location to the respective department which would deal with that problem. For that different departments would have to log in to our web app and then get the desktop notifications of any such events.

The system design on ReactsJS and will connect the user to the server where all the main activities shall be performed and data shall be stored. The server shall consist of python files doing these operations and our model will detect anomalies from all the live streams is a 3d Convolution Network or C3D and is trained on the UCF crime dataset.

And after all this design is implemented, we would have to connect the traffic cameras to our system and register users from different departments and that would be all. Those registered users would then get notifications if any anomaly relevant to their department occurs. The admins can also add new videos or new datasets as a sample for the model to be retrained and become more efficient.

Future Work

There is a future for this Research in house and shops CCTV systems. The point where the system looks for a department to send the report would be modified and different people can take that place. For example, if it is for house surveillance then the contact information of everyone would be saved and then notification would be sent to anyone if something strange or criminal activity is happening around their house or if it is for a business or shop and the owners are mostly away, they can use this service. Data set would be different as traffic crimes are different than home crimes but, this Research has a future in this particular area.

References

1. Kaggle. (2018). UCF-Crime. Available: https://www.kaggle.com/mission- ai/crimeucfdataset [Accessed 29 December 2021]
2. H. Bouma, J. Baan, G. J. Burghouts, P. T. Eendebak, J. R. van Huis, J. Dijk, and J. H. van Rest, "Automatic detection of suspicious behavior of pickpockets with track-based features in a shopping mall," in Optics and Photonics for Counterterrorism, Crime Fighting, and Defence X; and Optical Materials and Biomaterials in Security and Defence Systems Technology XI, 2014, vol. 9253, p. 92530F: International Society for Optics and Photonics. [Accessed 14 November 2021].
3. B. Krausz and C. Bauckhage, "Automatic detection of dangerous motion behavior in human crowds," in 2011 8th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2011, pp. 224-229: IEEE. [Accessed 14 November 2021].
4. K. Goya, X. Zhang, K. Kitayama, and I. Nagayama, "A method for automatic detection of crimes for public security by using motion analysis," in 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009, pp. 736-741: IEEE. [Accessed 2 December f2].
5. Scannera. (2017). Scannera.ai. Available: https://scannera.ai. [Accessed 2 December 2021].
6. S. Gradari, A. Pallé, K. R. McGreevy, Á. Fontán-Lozano, and J. L. Trejo, “Can Exercise Make You Smarter, Happier, and Have More Neurons? A Hormetic Perspective,” Frontiers in Neuroscience,

vol. 10. 2016, [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2016.00093.

7. S. N. S. Al-Humairi and A. A. A. Kamal, “Opportunities and challenges for the building monitoring systems in the age-pandemic of COVID-19: Review and prospects,” Innov. Infrastruct. Solut., vol. 6, no. 2, p. 79, 2021, doi: 10.1007/s41062-020-00454-0.
8. N. Hanley, C. Boyce, M. Czajkowski, S. Tucker, C. Noussair, and M. Townsend, “Sad or Happy? The Effects of Emotions on Stated Preferences for Environmental Goods,” Environ. Resour. Econ., vol. 68, no. 4, pp. 821–846, 2017, doi: 10.1007/s10640-016-0048-9. [1] L. W. Yang and C. Y. Su, “Low-Cost CNN Design for Intelligent Surveillance System,” in 2018 International Conference on System Science and Engineering (ICSSE), 2018, pp. 1–4, doi: 10.1109/ICSSE.2018.8520133.
9. R. M. H., V. Upadhya, V. V Holla, S. S. Shetty, and V. Tantry, “CNN based Smart Surveillance System: A Smart IoT Application Post Covid-19 Era,” in 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020, pp. 72–77, doi: 10.1109/I-SMAC49090.2020.9243576.
10. Z. S. Sabri and Z. Li, “Low-cost intelligent surveillance system based on fast CNN,” PeerJ Comput. Sci., vol. 7, p. e402, Feb. 2021, doi: 10.7717/peerj-cs.402.