Android Phone Based Smart Video Surveillance System

Dalitso Sauli[1*], Rajkumar Kalimuthu[2]

1*UG Student, School of Computer Science and Information Technology, DMI St. John the Baptist University, Malawi.

2 Lecturer cum Head, School of Computer Science and Information Technology, DMI St. John the Baptist University, Malawi.

Abstract – The aim of this project is to automate video surveillance using a smart video surveillance device based on an Android phone. Since camera surveillance under manual control goes wrong the majority of the time, the project grasps its monumentality. However, the standard of free movement is subpar. It necessitates someone to supervise in front of the control equipment, which can be very inconvenient. This device can be easily and rapidly implemented using an Android phone. In contrast to conventional surveillance systems, this latest one replaces large and costly surveillance cameras with an Android phone. Furthermore, the conventional system's back-end equipment has lost its functionality. The standard system's back-end equipment has been shifted to the front-end equipment. The front-end equipment becomes more intelligent with an updated invasion detection algorithm. This configuration will greatly improve the system's mobility and flexibility. As a result, this device will work better in various surveillance situations when dealing with emergencies. Furthermore, this cutting-edge intelligent video surveillance device not only introduces a new surveillance tool, but also advances surveillance technique growth. Recent developments in object detection and face recognition have enabled the development of realistic video surveillance systems. Embedded object detection and face recognition capabilities that are reliable and quick enough for commercial applications.

Keywords: Android,Video surveillance,Video analytics, Deep learning, Face detection, Face recognition, Color temperature, Mobility, Machine learning, Object detection.

I. Introduction

As technology advances, people realize that monitoring the video of a business or any entity for security purposes is a waste of time. The masses want something to happen without them having to spend too much of their time. Users can view/share images, as well as upload and import footage files from anywhere in the world, using this system. We want the multitasking installation to be able to take place in any remote location [1]. Video can be tracked from a remote location using a cell phone with a web camera, which can be an inbuilt web camera or an external web camera using Android. Different methodologies, such as warning systems and PC-based video systems, were used in older technology to guarantee safety [1] Many people have problems when using old devices because it is difficult for a consumer to track the security position when they are out of the house. Using both of these devices requires the user to constantly track the video for security reasons, which has a number of drawbacks. It is possible to connect with someone at any time using the internet. People can watch videos of their location on their phones even though they are not there. The video is captured in the system using a web camera, and people can watch it on their phones, saving time and ensuring security.

Related work

Based on my study, users can verify a specific location protected by a camera range using video verification in the context of closed-circuit television. For example, such a useful mechanism enables security personnel to determine if a raised warning signal is a false positive or whether a severe situation exists that needs immediate attention. Typically, this means that a burglar alarm system has several cameras mounted, the cameras monitor the areas covered by the alarm system's sensors, and the security staff is alerted if there is an alarm. Rather than running to the sensor that was activated, they first check the cameras to see if they can see something the cause of the alarm the protection group can only proceed to the target area after this phase of verification, and only if there is a justification for the warning. In general, video verification is a procedure that allows a person to log into a workstation and view video footage. This person can then verify, or check, any fact or location by watching the video. Remote video verification basically means that the video verification process is accessible not only from a specified workstation, but also from a mechanism that allows the user to log into the system from almost any location, such as a mobile phone.

II. Survey

Every day, technology advances, and people expect multitasking installations to arrive at their doorstep. As a result, they won't be able to continuously track the device for security, which will require more time and effort. We are able to improve our situation by using a Video Surveillance system and a cell phone.

III. Methodology

Through the advancement of cell phones, everyone can connect with anyone else at anytime, anywhere on the planet, using internet technology. People may track the videos of their location or company using their mobile phone without having to spend time in the same room but outside of their working area.The video is collected and stored in the framework using a web camera, which the user can access at any time using their mobile phone. More protection will be given as a result of this. [1] A video surveillance system solves the security issue, despite the fact that cell phones have limited capacity. The solution to this issue is a content sharing framework. When we have less memory space on our cell phone to store data, we can move the data to a remote PC on the network and free up some memory on our phone by transferring some of the files to a remote PC. Instead of manually connecting our cell phone to a PC and transferring data, this allows us to easily access the data. We can migrate data from mobile to PC and from PC to mobile using the content sharing framework. We can access information from a remote PC via a cell phone from anywhere in the world. In addition, we are improving the security of the team viewer program by providing security for both easy users and administrators. Simple users may only access the drives that were assigned to them during their registration. Since access rights are assigned to users individually, and the administrator has complete access, one simple user cannot access the details of another simple user the project will be implemented using the following comprehensive models in order to achieve the outlined objectives.

1. Authentication

An authentication scheme is a module that implements a method for a user to authenticate themselves to Simple ID by comparing credentials provided by the user to credentials stored in a data store containing user information and determining if the credentials match those stored in the data store. If incorrect information is provided, the user's identity is not verified, and therefore no access is granted.

2. Video Capture

This video capture module is used on the Intelligent Video Surveillance System's server side. This module's main role is to capture video from the surveillance area. The picture sequences taken by the Android phone's camera will be shown. In addition, provide a method for retrieving Real-Time video data. The information will be included in the Intelligent Transportation System.

3. Video Storage

Video Storage modules are available on both the server and client sides of the Intelligent Video Surveillance System. The Video Storage module is used on the server to encrypt the video data and save the video clip to the SD card. The compression is done with Java Native Interface. When the Intelligent Video Surveillance module senses an invasion, only a 30-second video clip is saved to the SD card. It contains a 15-second video clip before the invasion is detected, as well as a 15-second video clip afterward. This system not only provides users with documentation of the invasion, but also guarantees that the video recording is not too massive to be stored on an Android phone's SD card. On the client's end. On the client side, the Video Storage module encodes real-time surveillance video using the Media Recorder class provided by Android APIs. Users have the option of turning this feature on or off.

4. Video surveillance With Intelligence

The Intelligent Video Surveillance module is an essential part of the overall framework. It examines each frame of video to see whether an invasion occurs based on a set of criteria. Context Subtraction has been used as the invasion detection algorithm of the Intelligent Video Surveillance System due to the limited resources of the Android phone and the fixed background of the realistic surveillance scenario. To model the Background frame in Intelligent Video Surveillance System, we use an updated algorithm. It only updates the static part of the background frame, unlike conventional background modeling.

5. Warning

In both the server and client sides of the Intelligent Video Surveillance System, there is an Alert module, as well as a Video Storage module. Real-Time Alert is the name of the Alert module on the server side. When the invasion occurs, the user will receive an SMS notification. Then a screenshot of the invasion will be saved to the SD card. Meanwhile, the Real-Time Communication module will send this snapshot to the client-side Android phone. The Warning module is known as the Alert Information Storage module on the client side. It saves the invasion photo and video clip that the server-side Android phone sends.

6. Appreciation

Faces delivered by the detection module will be recognized by this module. The Eigen faces algorithm and component analysis dimensionality reduction are used to accomplish this. Different groups of faces in the database will be needed for the training phases.

The recognition module will have an individual id, which will be used to extract data from the MySQL database.

8. Module For Object Detection

Detection of objects the detection models are present in the system, and the images in datasets are to be processed. The main advantage of practicing on a dataset with more object categories is that we are less likely to miss any relevant object classes for many video surveillance applications. In other words, even though the models are not trained on all available categories, the evaluation results obtained from models trained on datasets with more object categories would be more descriptive and representative in our case.

System Architecture

The process of defining the components, modules, interfaces, and data for a system in order to meet specific specifications is known as system architecture. The system's architecture is outlined below.

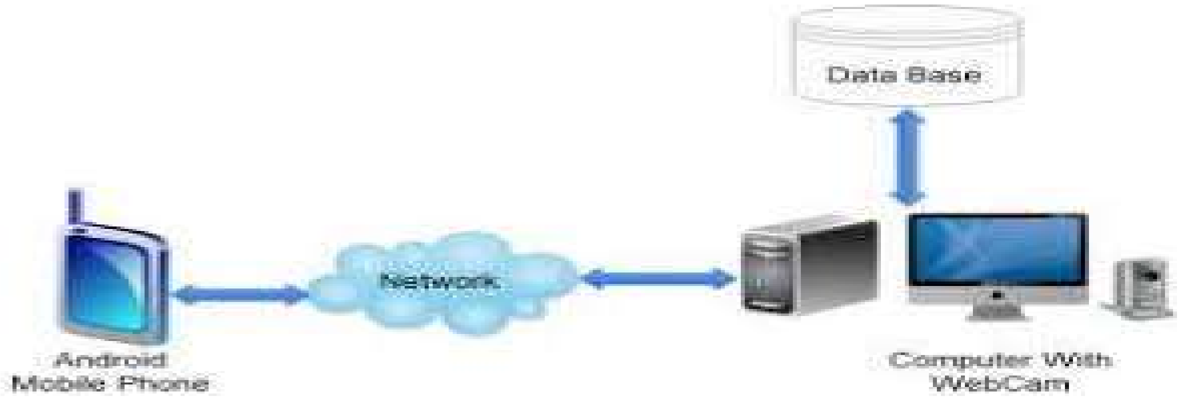

Figure 1: Overall System ARCHITECTURE

IV. Algorithm And Techniques Object Detection / Face Detection

Clearly, in today world we have couple of active object detection algorithms with promising accuracies. Spatial pyramid pooling networks [4], for example, are able to achieve up to 59.2% mAP on PASCAL VOC 2007 [3] for detection tasks, which is a bit better than the 58.5% mAP produced by region-based convolutional neural networks [5, 4]. More importantly, Spatial pyramid pooling networks was invented to achieve better efficiency over region-based convolutional neural networks by sharing computation. As a result, Spatial pyramid pooling networks can outperform region-based convolutional neural networks by sharing computation by 10 to 100× in terms of speed while still producing accuracy of similar level [4].

In order for us to acquire inclusive rating outcome on the performance of our face recognition models, datasets consisting of both static images and image sequences are needed. Therefore, we choose to train our face recognition models on two separate datasets, one with static images and the other with image sequences. Obviously, these two datasets will have face images of two different groups of people, and for our evaluation, it has to be the case that a subset of people is shared by both datasets. With that in mind, we discover that VGGFace2 [6] together with Tiktok Faces DB [7] makes a promising combination. VGGFace2 contains over 3.3 million face images of over 9000 people with an average of 362 images per subject [4]. Meanwhile, Tiktok Faces DB provides 3,425 videos of 1,595 different subjects with an average of 2.15 videos for each subject [7]. Although only a small subset of subjects is shared by VGGFace2 and Tiktok Faces DB, it is enough for carrying out our evaluation.

V. Conclusion

To sum it up Situation recognition is greatly supported by a smart video monitoring system. These systems convert video surveillance from a data-gathering method to a data-gathering and intelligence-gathering system. Smart surveillance systems will respond in real time thanks to real-time video analysis. Our system detects the intrusion and sends alerts to authorized individuals so that appropriate action can be taken.

VI. References:

1. D. Shiny Irene PG Scholar," Video Surveillance System and Content Sharing Between Mobile and PC Using Android",Dept of Computer Science and Engineering RMK Engineering College, Anna University of Technology, Kavaraipettai, Chennai, Pages(s), Year 2012.

2. Burczynski, T., Ku ´ s, W., Długosz, A., Poteralski, A., Szczepanik, M.: Sequential and ´ distributed evolutionary computations in structural optimization. In: Rutkowski, L., Siekmann, J.H., Tadeusiewicz, R., Zadeh, L.A. (eds.) ICAISC 2004. LNCS (LNAI), vol. 3070, pp. 1069–1074. Springer, Heidelberg (2004)

3. Gool LV, Williams CKI, John Winn AZ (2010) The pascal visual object classes (VOC) challenge. Int J Computer Vis 88:303–338

4. He K, Zhang X, Ren S, Sun J (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intel 37(9):1904–1916

5. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: The IEEE conference on computer vision and pattern recognition (CVPR)

6. Cao Q, Shen L, Xie W, Parkhi OM, Zisserman A (2018) VGGFAce2: A dataset for recognising faces across pose and age. In: International conference on automatic face and gesture recognition.

7. Wolf L, Hassner T, Maoz I (2011) Face recognition in unconstrained videos with matched background similarity. In: The IEEE conference on computer vision and pattern recognition (CVPR).