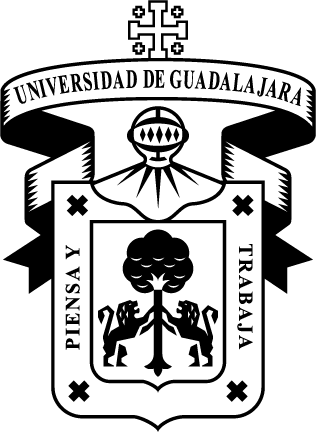
Universidad de Guadalajara

Centro Universitario de los Valles



**Automatic recognition of violent physical actions in video surveillance systems.**

Tabla of content

[Abstract 3](#_Toc19209750)

[Purpose 3](#_Toc19209751)

[Scope and Limitations 3](#_Toc19209752)

[Scope 3](#_Toc19209753)

[Limitations 4](#_Toc19209754)

[Audience 4](#_Toc19209755)

[Modules 4](#_Toc19209756)

[Creation of data set 4](#_Toc19209757)

[Requirements 5](#_Toc19209758)

[Pose extraction 5](#_Toc19209759)

[Requirements 5](#_Toc19209760)

[Classifier of violent action 5](#_Toc19209761)

[Requirements 5](#_Toc19209762)

[Detection of violent action in real time 5](#_Toc19209763)

[Requirements 5](#_Toc19209764)

[Technologies to use 6](#_Toc19209765)

[Work team 6](#_Toc19209766)

[Budget 6](#_Toc19209767)

# Abstract

Video surveillance systems are considered technological tools that help safeguard people's safety. Currently these systems have a great acceptance in daily life, for example, the monitoring of public areas such as airports, department stores, educational institutions, etc. With the application of current advances in digital image processing and wireless communication technology, systems can be created for the visualization and monitoring of areas of interest, capable of capturing and processing images, intelligently sending only the necessary information for the interpretation of different activities.

# Purpose

The interest in developing video surveillance systems to safeguard the security and protection of people grew significantly after the 11th September attack (9/11), so these systems emerge as a natural solution to detect and prevent attacks or situations that put at risk the integrity of people. Video surveillance systems have evolved over time. In 1960, the first generation of video surveillance systems emerged, which, directly focused on visual detections made by human operators, said operators had to analyze, interpret and classify the results of supervision. These types of systems did not guarantee rigorous surveillance because the attention in most individuals declines after 20 minutes of evaluating the screens. Current video surveillance systems called fourth generation (4GSS), seek automation in the detection of relevant events during the monitoring of regions of interest by implementing software systems based on artificial vision algorithms. These systems are developed as support tools, to help the operator to efficiently monitor different areas of interest.

With the next project is to design and implement a tool video surveillance with the ability to detect violent activities such as fights, the system will have the ability to highlight the 2 humans who are participating in the fight with the objective of providing aid for the visual detection of the participants in this violent action

# Scope and Limitations

The CONACYT (Consejo Nacional de Ciencia y Tecnologia) requires a computer vision system that is capable of automatically detecting violent physical actions, such as kicks and punches. Helping security guards detect these violent actions more quickly, the main features of the system are the following:

## Scope

• The system will allow the detection of people using a camera.

• The system will perform the characterization of violent actions by extracting poses in image sequences.

• The system will allow the training of new models to detect violent actions.

• The system will allow handling different types of scenarios.

## Limitations

• The system will not handle scenarios with poor lighting.

• The camera resolution must not be low in order to have visible images

• The system will be developed for a static camera

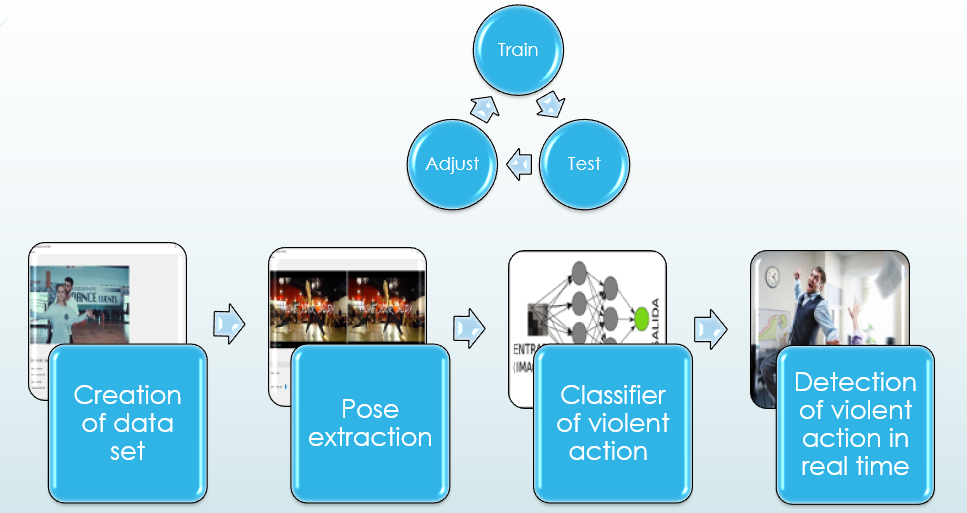
• The problem of detecting violent actions in crowds will not be addressed.

# Audience

The software is focused on being a support tool for video surveillance systems, so it can be implemented in homes, classrooms, shopping centers, hospitals, etc.

# Modules

The system has 4 modules in which:



## Creation of data set

This module is responsible for editing videos, this to obtain small fragments of video (around 15 frames) in order to create a fairly complete data set and then be used as a resource for training the model of fight detection.

### Requirements

|  |  |
| --- | --- |
| **Requirement Identification:** | RF001 |
| **Name of the requirement:** | Search video |
| **Features:** | Search in any operating system directory |
| **Description of the requirement:** | This module requires the option to search for video files in the operating system folders. Only avi , mp4, mov type files can be opened .  Handle exceptions when trying to open a different file. |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | high |

## 

|  |  |
| --- | --- |
| **Requirement Identification:** | RF002 |
| **Name of the requirement:** | Play video |
| **Features:** | Play the video in a window. |
| **Description of the requirement:** | This module requires that the playback of the video on a system screen be displayed with a (Play) button. This button will be able to resume video playback. |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | high |

|  |  |
| --- | --- |
| **Requirement Identification:** | RF003 |
| **Name of the requirement:** | Pause and stop video |
| **Features:** | Buttons to stop and play a video |
| **Description of the requirement:** | In this module you have to develop two buttons one to pause the video, this button has to have the ability to save the current position of the frame that is being displayed. The stop button will restart the video to the first frame of the video. |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | high |

|  |  |
| --- | --- |
| **Requirement Identification:** | RF004 |
| **Name of the requirement:** | Time bar |
| **Features:** | Create a playback time bar. |
| **Description of the requirement:** | Create a slider that shows the total number of frames to be played and as the indicator reproduces, it must reach the right side of the slider .  The slider has the facility to position itself in any frame that the user indicates. Showing the selected frame information |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | medium |

|  |  |
| --- | --- |
| **Requirement Identification:** | RF005 |
| **Name of the requirement:** | Crop video |
| **Features:** | Trim a number of frames , to save it on the computer as video (. Avi ) |
| **Description of the requirement:** | The user must select the frame where the video should start cutting and where that cut should end . You must create a textlabel where the user added the amount of frames to crop.  You must have a save button and select the address where you want to save this new video |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | medium |

## Pose extraction

As a second step, after obtaining the data set, this will serve to obtain the characteristics of each of the people involved in the fights, with characteristics we refer to the movements of the limbs that occur when they do a violent physical action . This information will be stored in a CSV document, which will obtain labels for violent action or non-violent action.

### Requirements

|  |  |
| --- | --- |
| **Requirement Identification:** | RF006 |
| **Name of the requirement:** | YOLO network implementation |
| **Features:** | Implement the YOLO feature descriptor on a screen |
| **Description of the requirement:** | It is required to use the YOLO pose characteristics descriptor, displaying that information on the screen. |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | high |

|  |  |
| --- | --- |
| **Requirement Identification:** | RF007 |
| **Name of the requirement:** | Compare normal video vs feature descriptor |
| **Features:** | Show normal video and video with feature descriptor |
| **Description of the requirement:** | The user must visualize both the normal video and the video already processed marking the detection of people in the frames . |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | high |

|  |  |
| --- | --- |
| **Requirement Identification:** | RF008 |
| **Name of the requirement:** | Save information in a csv |
| **Features:** | Save all movements of people with the variables obtained from the characteristics descriptor |
| **Description of the requirement:** | Save all movements of people with the variables obtained from the characteristics descriptor in a CSV file. Label when a violent action is saved or when it is a non-violent action. |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | medium |

## Classifier of violent action

As a third module of the system there is the classifier of violent physical action, this module is responsible for working with the information collected from the CSV, with this information we will adjust the classifier to obtain a reliable one (80% accuracy in detections of violent physical action).

### Requirements

|  |  |
| --- | --- |
| **Requirement Identification:** | RF009 |
| **Name of the requirement:** | Train Neural Network |
| **Features:** | Train neural network to identify violent physical actions |
| **Description of the requirement:** | Take all the CSV files obtained from the pose extraction module and train a neural network to identify violent physical actions.  Hand punches and kicks will be identified |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | high |

|  |  |
| --- | --- |
| **Requirement Identification:** | RF010 |
| **Name of the requirement:** | Adjust neural network to identify violent physical actions |
| **Features:** | Add more networks or remove networks until you get an acceptable percentage of fight recognition |
| **Description of the requirement:** | Save all movements of people with the variables obtained from the characteristics descriptor in a CSV file. Label when a violent action is saved or when it is a non-violent action. |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | high |

|  |  |
| --- | --- |
| **Requirement Identification:** | RF011 |
| **Name of the requirement:** | Test neural network |
| **Features:** | Test neural network to identify violent physical actions |
| **Description of the requirement:** | Prove that the neural network performs well by identifying violent and non-violent actions |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | high |

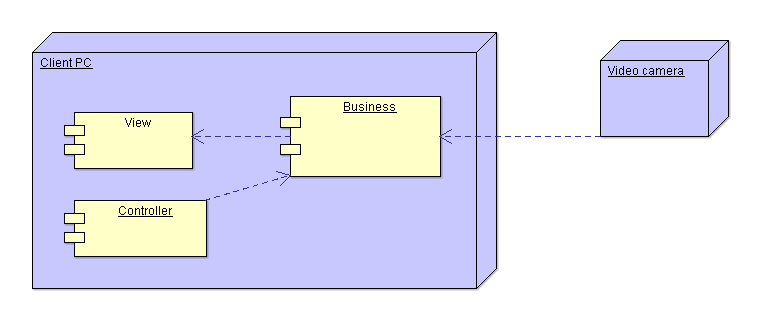
## Detection of violent action in real time

Module in charge of detecting violent actions in real time, with the images captured by the video surveillance camera will show an alarm when a violent physical activity is occurring in the region of interest. In this module there will be a historical panel where it will show all the events that have occurred, this to have greater control over all the violent physical actions captured by the system.

### Requirements

|  |  |
| --- | --- |
| **Requirement Identification:** | RF012 |
| **Name of the requirement:** | Online detector |
| **Features:** | With the information obtained from a camera, physical violent actions are detected |
| **Description of the requirement:** | With the information obtained from a video camera, the software will identify if it is a violent action, since the information will be continuously sent to the neural network to detect these actions |
| **Non-functional requirement:** |  |
| **Priority of the requirement:** | high |

## Deployment diagram

The deployment diagram shows the physical components of the application. In the illustration we can see that the operation is based on a desktop application infrastructure.

# Technologies to use

For this system the following technologies will be used:

* Python
* Open CV
* scikit learn
* PyQT

# Work team

• 3 Programmers  
• 2 Testers  
• 1 Technical Analyst (vision specialist)  
• 1 Project leader (vision specialist)

# Budget

6 month project budget:

* $ 744.23 monthly programmer
* $ 694.61 monthly tester
* $ 1240.38 monthly analyst
* $ 1736.53 monthly project leader

In the development of the application you have a budget of $ 39,592.77 plus 30% profit of the software $ 11877.83.

You have a total of the project investment of: $51470.60