Violence Detection using CNN + LSTM neural network

Today, the amount of public violence has increased dramatically. This has resulted in the ubiquitous use of surveillance cameras. But almost all systems today require the human-inspection of these videos.

which is virtually inefficient. It is therefore necessary to have such a practical system that can automatically monitor and identify the surveillance videos. The development of various deep learning techniques, especially violence and fighting. And that’s what we want to apply.

In this project, we have explored a one strategy from different strategies to find out the saliency of the features from pretrained models in detecting violence in videos. A dataset has been Collected raw surveillance videos from YouTube, sliced them into clips within 5s at 20 fps, and labelled each clip as Violent or Non-Violent Behaviour. we got 1600 clips and 160,000 frames as a new data set for real-world violent behaviour detection under surveillance camera.

unfortunately, i couldn't use more than 1,600 videos. because to train this model on a larger amount of data than this needs super machine, and when training the model I noticed that the more data the amount >>> the more accurate the model. That is why it would be great to use a larger amount of data with a super computer

The method consists of extracting a set of frames belonging to the video, sending them to a pretrained network called VGG16, obtaining the output of one of its final layers and from these outputs train another network architecture with a type of special neurons called LSTM. These neurons have memory and are able to analyse the temporal information of the video, if at any time they detect violence, it will be classified as a violent video.

In the end, this strategy provide an accuracy of **74.58** % on the **training** set and **76.25** % on the **validation** set. And i have used **regularization** technique to avoiding the risk of **Overfitting**