**PROJECT PROPOSAL- TEAM DATA DIVERS**

**Violence Detection in videos using Computer Vision and Neural Networks**

We live in a world where we rely heavily on usage of surveillance cameras for our ensuring high levels of security. Even though these smart cameras are easily available and used everywhere, these monitoring is highly ineffective and unreliable due to a number of reasons. The most important being monitoring of huge quantities of video footage to be able to detect any incidents of violence. Although Violence detection has been receiving significant attention due to its numerous practical use cases, it lacks in-depth research due to more focus on action recognition based on detecting simple actions like clapping, walking, or jogging etc.

The problem statement for the project consists of two parts: feature extraction. and classification. The approach applied for feature extraction will be using deep convolutional neural networks-based representations (CNNs). Once feature extraction is completed, Long Short-term Memories (LSTM) will be used for modeling the temporal information, as they help to find out relationships between the consecutive frames through their memory ability. The proposed model will use a spatial feature extractor followed by Long Short-Term Memory (LSTM) as temporal feature extractor and sequence of fully connected layers for classification purpose with attention mechanism that would help enhance the performance of the LSTM. In summary, CNN + LSTM network will be used for this action recognition due its proven high performance in previous models. Below is a diagram of the proposed system:



Some of the potential datasets that we plan to use are: INRIA IXMAS, which show an individual kicking or punching or datasets like CAVIAR, BEHAVE or CareMedia contain some instances of people engaged in aggressive behaviors or RWF-2000 which is an Open Large Scale Video Database for Violence Detection that contains both-fight and non-fight sequences for detection.