

**UNIVERSITÀ DEGLI STUDI DI BARI “ALDO MORO”**



**DIPARTIMENTO DI INFORMATICA**

**Corso di Laurea in Informatica**

**RICONOSCIMENTO DI ANOMALIE DA  
TELECAMERA DI VIDEOSORVEGLIANZA**

**RELATORE**

**Prof. Donato Impedovo**

**Laureando**

**Marco Cappiello**

# INTRODUZIONE :

Le telecamere messe in commercio al giorno d'oggi non riescono ad identificare crimini autonomamente.

Attraverso tecniche di computer vision e machine learning è possibile identificare azioni sospette in maniera automatica.



# INTRODUZIONE :



Ma che cos'è un'azione anomala?

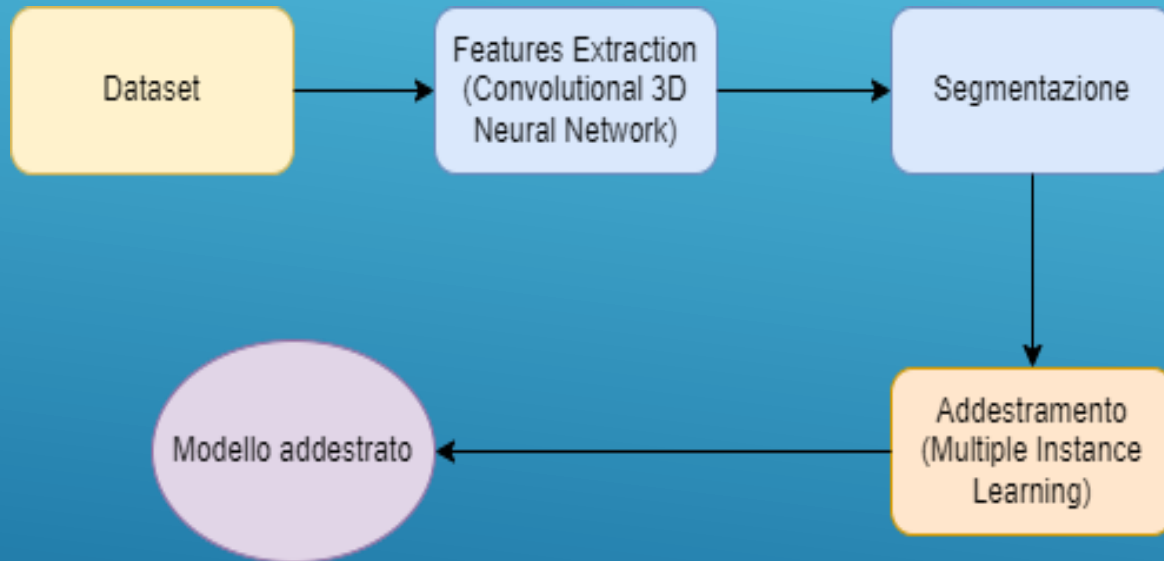
Intuitivamente:  
un'azione anomala è  
un'azione infrequente in  
un contesto pubblico.

# OBIETTIVO

Creare un sistema capace di identificare autonomamente situazioni sospette in un video di sorveglianza



# IDEA GENERALE



Creare un modello che sfrutta il multiple instance learning per assegnare punteggi di anomalia ai video ricevuti in input.

# DATASET 1

- Il primo dataset utilizzato è l'UCF-crime.
- Il dataset è composto da 14 classi di azioni
- 950 video normali e 950 video anomali per un totale di 1900

Categoria	Numero di video
Anomali	950
Normali	950
#Totale	<b>1900</b>

# DATASET 2

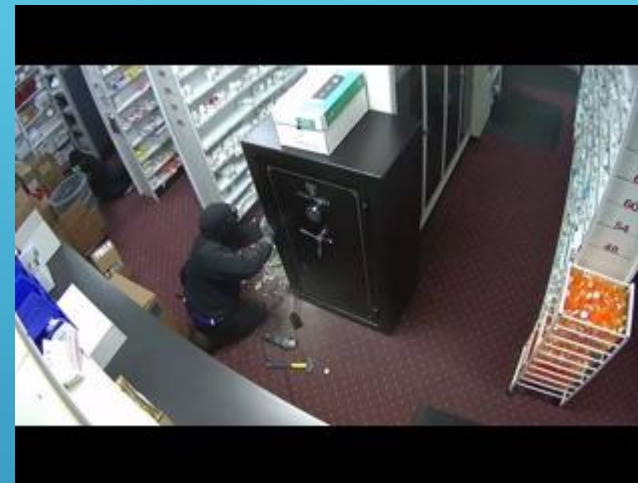
- Il secondo dataset utilizzato è il real life violence situations
- Il dataset è composto da 2 classi di azioni
- 1000 video normali e 1000 video anomali

Categoria	Numero di video per categoria
Anomali	1000
Normali	1000
#Totale	<b>2000</b>



# DATASET

- Tutti i video sono stati raccolti da contesti reali e sono di lunghezza variabile.
- Ogni video è stato fissato a 25 fps con risoluzione di 320x240 px.
- Di ogni video si conosce unicamente la classe di appartenenza e la lunghezza.

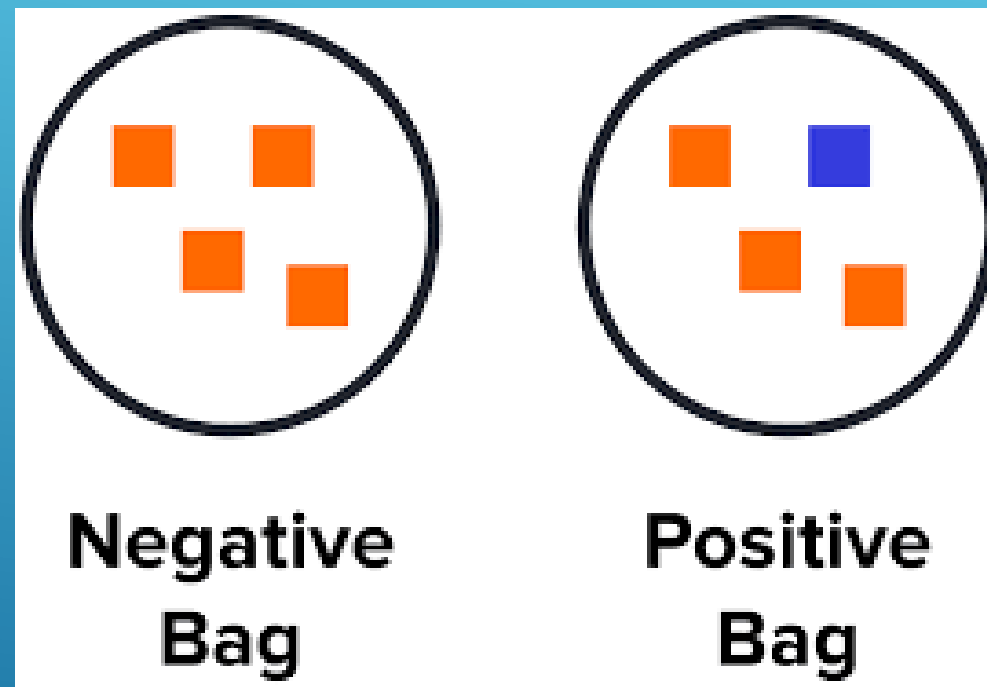




# MULTIPLE INSTANCE LEARNING

- Una bag è formata da istanze.
- Una bag è positiva se almeno una delle sue istanze è positiva
- Altrimenti la bag è negativa

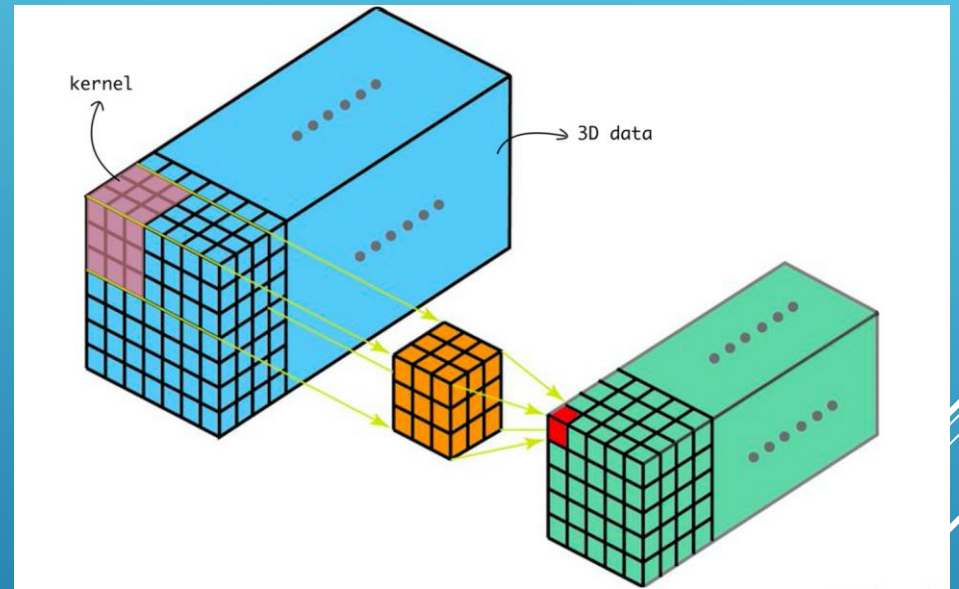
Nel nostro caso i video sono bag e i segmenti sono istanze.



# FEATURE EXTRACTION

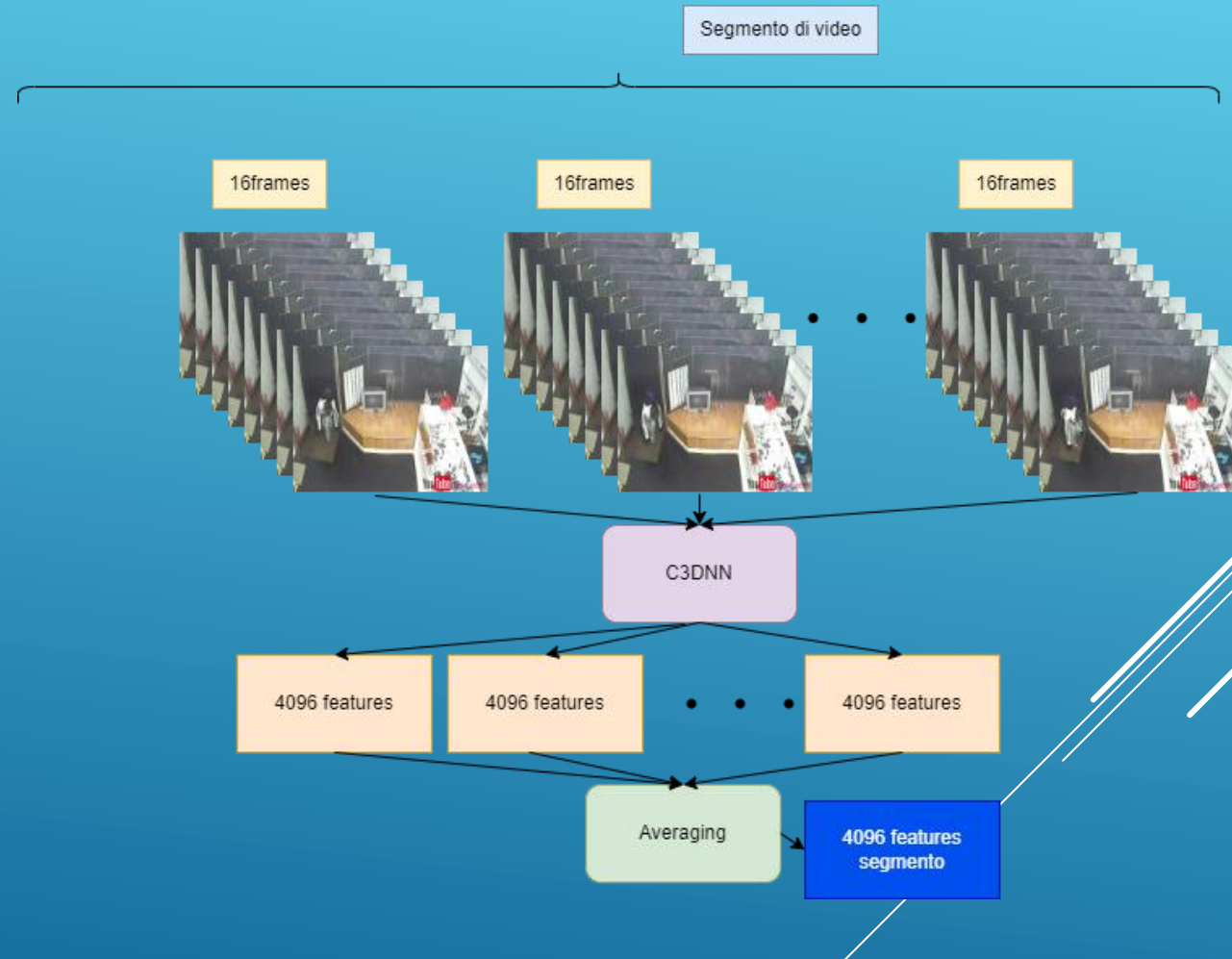
La features extraction è stata effettuata attraverso una convolutional neural network.

Il kernel utilizzato ha 3 dimensioni e estrae features da clip di 16 frames.

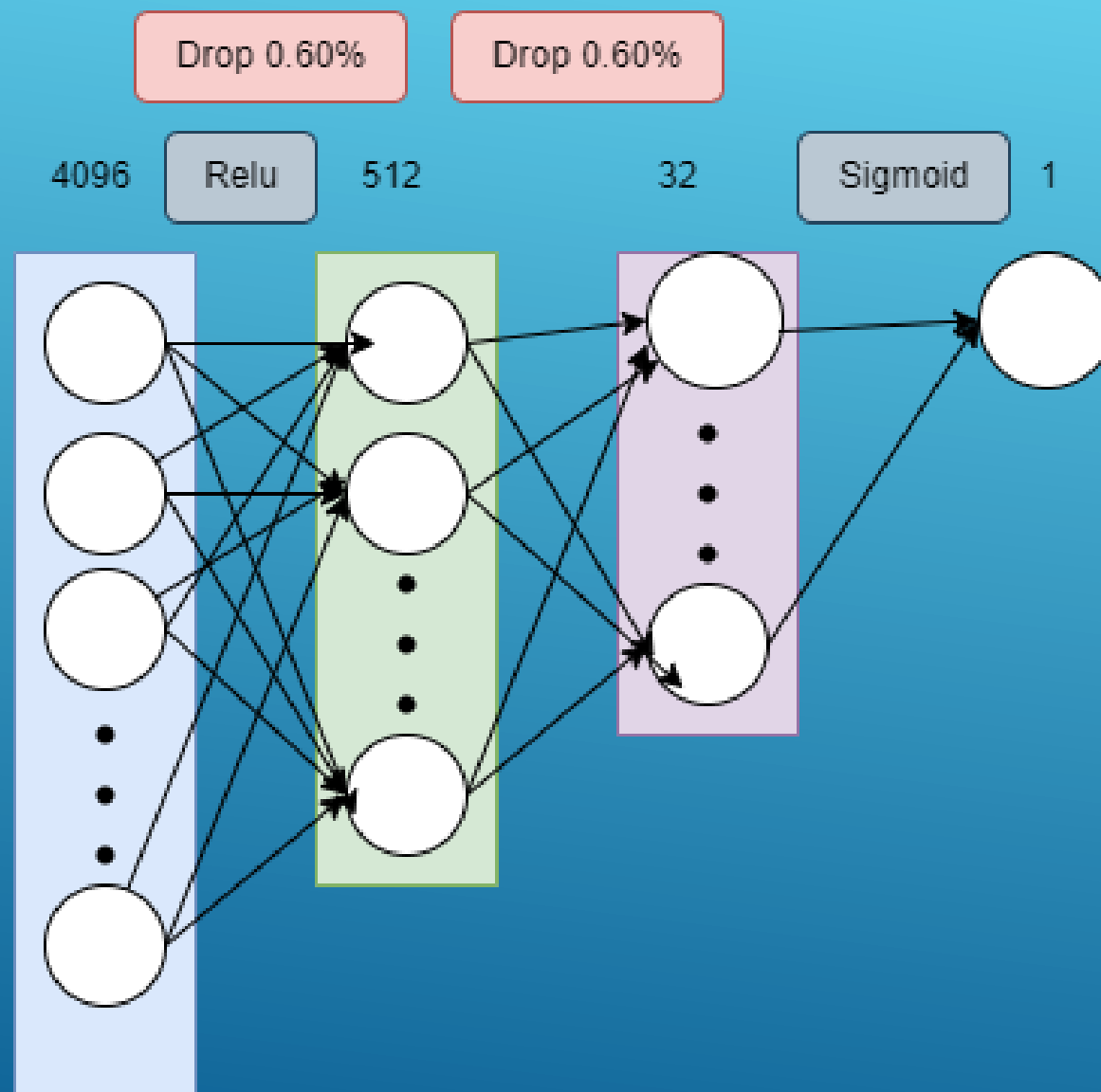


# SEGMENTAZIONE

- Ogni segmento è stato diviso in 32 segmenti
- Ogni segmento contiene il valore medio delle features delle clip di cui è composto



# MODELLO



# SPERIMENTAZIONE: PRIMO DATASET

Il dataset è stato diviso in 2 partizioni:

- 1620 video per il train set
- 280 video per il test set

Categoria	Numero di video nel train set	Numero di video nel test set
Anomali	810	140
Normali	810	140
#Totale	<b>1620</b>	<b>280</b>

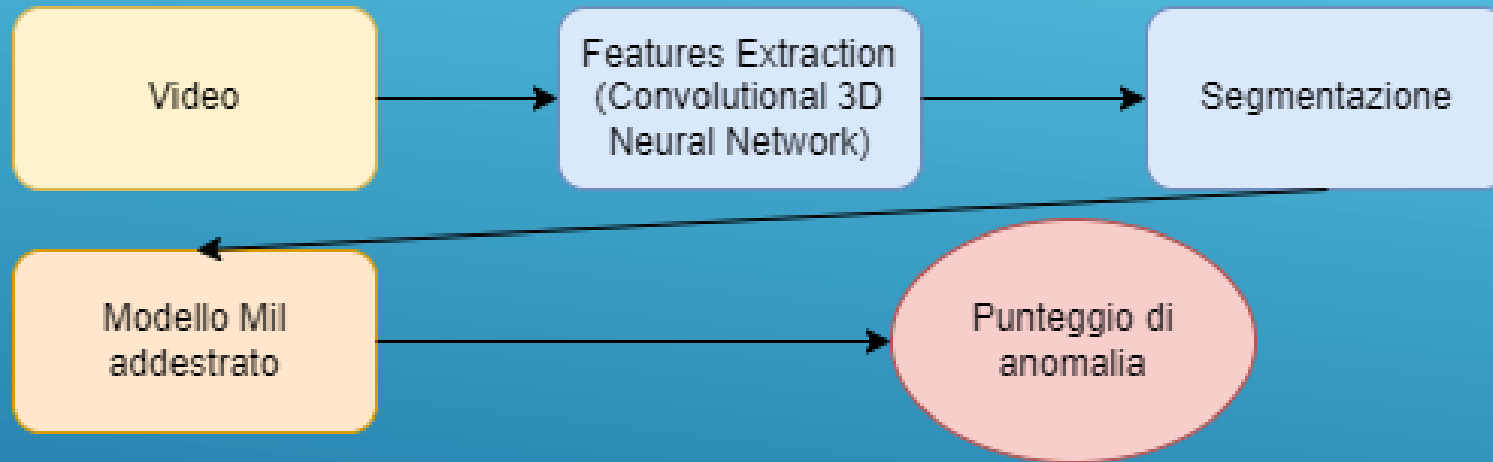
# SPERIMENTAZIONE: SECONDO DATASET

Il dataset è stato diviso in 2 partizioni:

- 1400 video per il train set
- 600 video per il test set

Categoria	Numero di video nel train set	Numero di video nel test set
Violenza	700	300
Normali	700	300
#Totale	<b>1400</b>	<b>600</b>

# SPERIMENTAZIONE



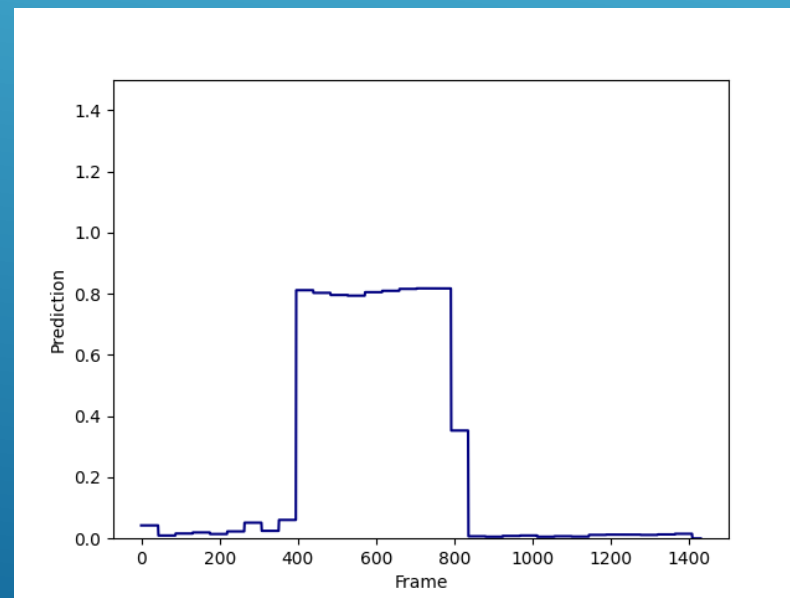
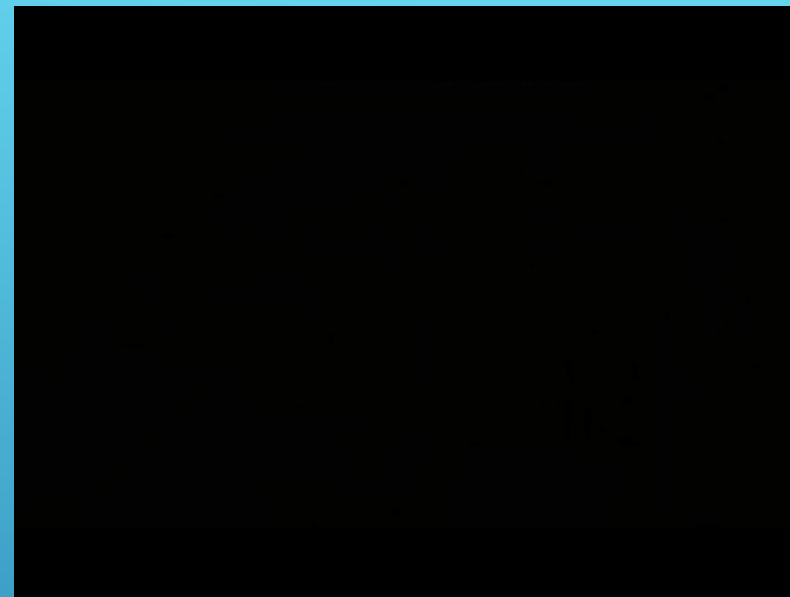
Ogni video presentato al sistema viene segmentato e viene effettuata la features extraction attraverso una C3DNN.



# SPERIMENTAZIONE

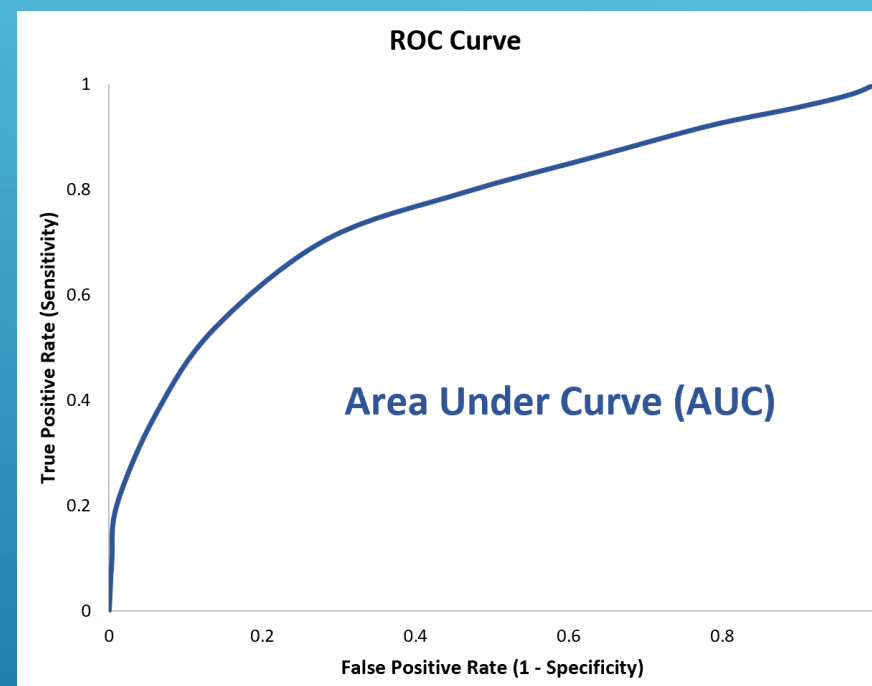
I punteggi di anomalia relativi ai singoli segmenti vengono riportati su un grafico che ha:

- punteggio di anomalia sull'asse delle ascisse
- frame sull'asse delle ordinate



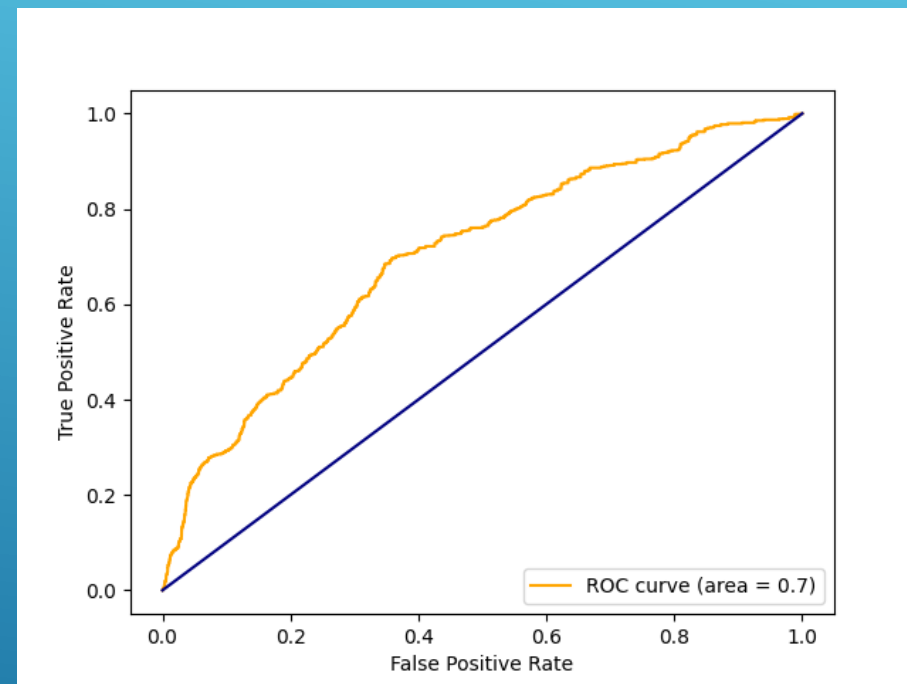
# METRICA DI VALUTAZIONE

- La metrica di valutazione utilizzata è la curva Receiver Operating Characteristic e la corrispondente AUC.



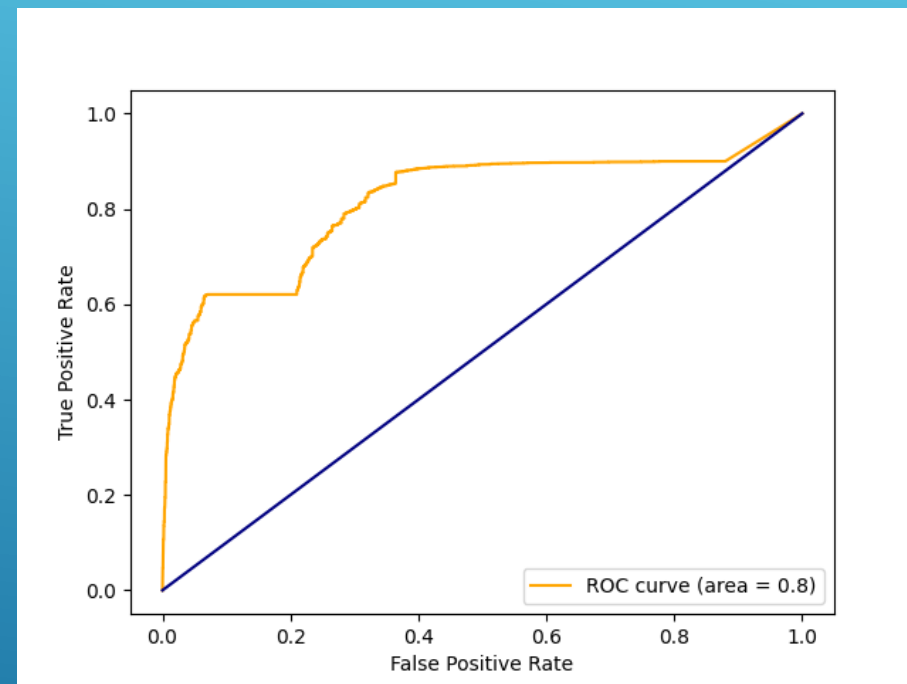
# SPERIMENTAZIONE 1

- Modello addestrato sull'UCF-Crime e testato sull'UCF-CRIME
- L'AUC raggiunto nella prima sperimentazione è del 70%.



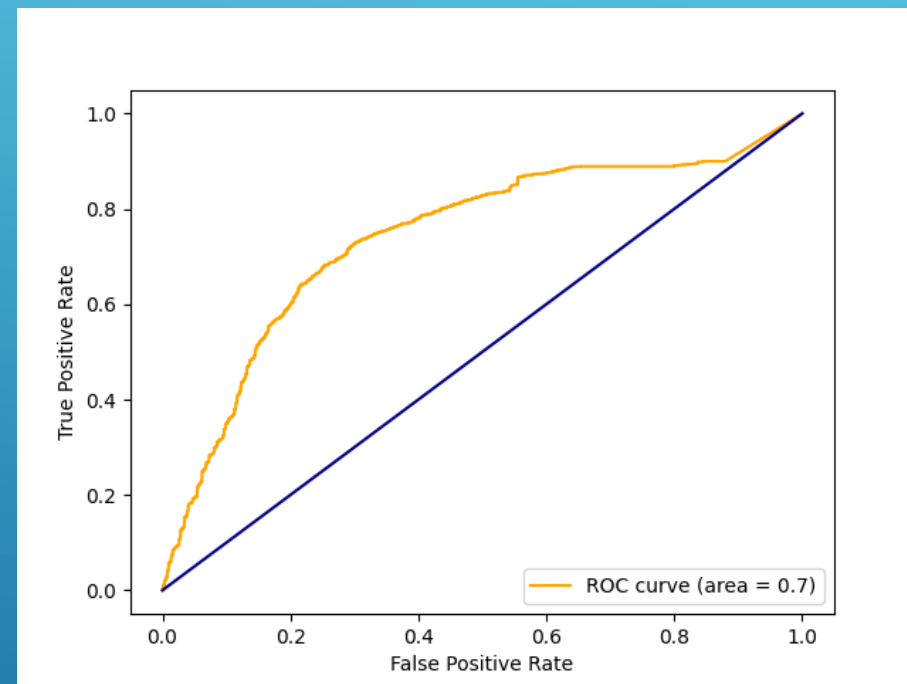
# SPERIMENTAZIONE 2

- Modello addestrato sul Real Life Violence Situations e testato sul Real Life Violence Situations.
- L'AUC raggiunto nella seconda sperimentazione è del 81%.



# SPERIMENTAZIONE 3

- Modello addestrato sull'UCF-Crime e testato sul Real Life Violence Situations
- L'AUC raggiunto nella terza sperimentazione è del 74%.



# CONCLUSIONI

Il sistema raggiunge un buon AUC in ogni sperimentazione.

La prima sperimentazione raggiunge un AUC del 70% avvicinandosi di molto all'AUC conseguito nello stato dell'arte del 75%.

Modello	AUC
Esperimento 1	70%
Esperimento 2	74%
Esperimento 3	84%
Stato dell'arte	75%

# SVILUPPI FUTURI

Possibili sviluppi futuri prevedono:

- ▶ L'utilizzo di un dataset con video di qualità migliore
- ▶ L'utilizzo di diverso tipo di features (es: I3D features)
- ▶ L'implementazione del sistema per riconoscimento di anomalie in tempo reale



**GRAZIE PER L'ATTENZIONE!**

Several thin, white, parallel diagonal lines extending from the bottom right towards the top right of the slide.