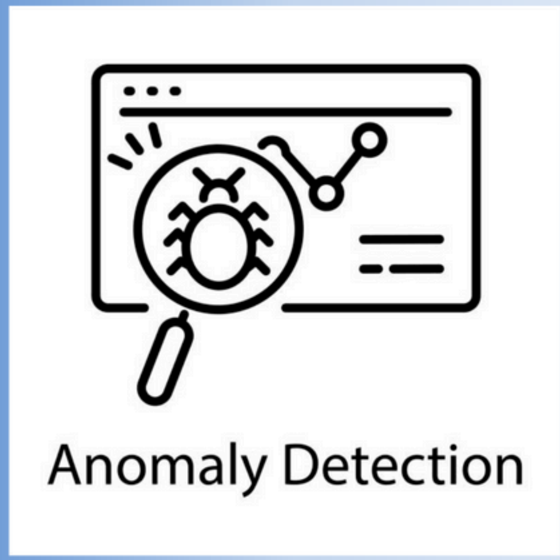
Real Time Violence Detection And Alert System



Guide -M.s Tanya Sharma

Presented by

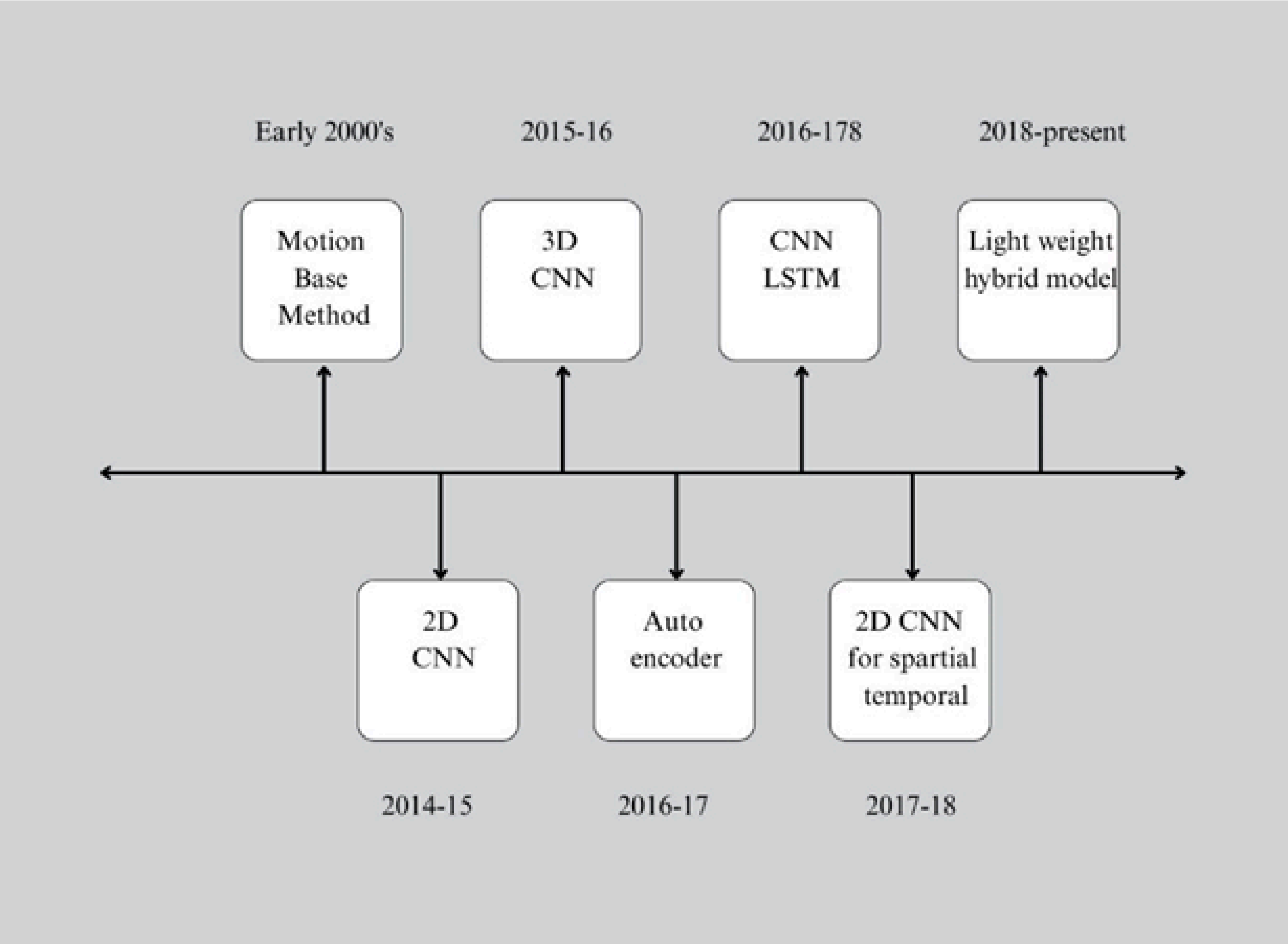
Neeraj Upadhyay- 2100300130088
Mohit Karakoti -2100300130085

Problem Statement

- Increasing public violence demands automated surveillance systems.
- Manual monitoring is error-prone and not scalable.
- Detecting violence in real-time from video feeds is a critical safety need.
- The challenge lies in recognizing complex, fast, and subtle violent actions accurate

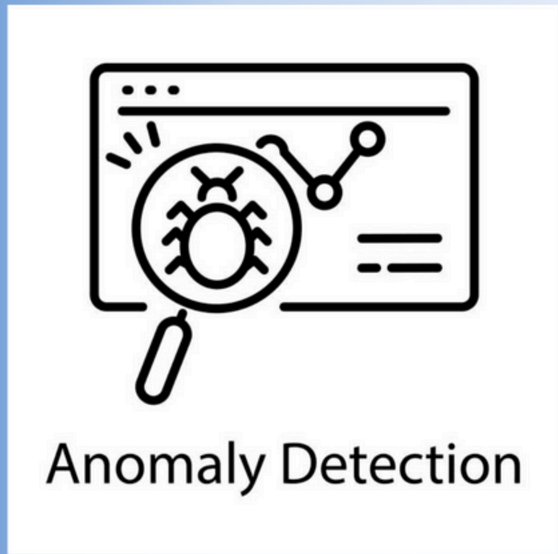
Related Work

- 2D CNNs: Good for spatial features but lack temporal context.
- Autoencoders: Used for anomaly detection, but not task-specific.
- 3D CNNs: Capture temporal data, but are computationally expensive.
- CNN + LSTM: Combines spatial and sequential learning.
- Gaps: Many models struggle with real-world variability and lightweight real-time deployment.

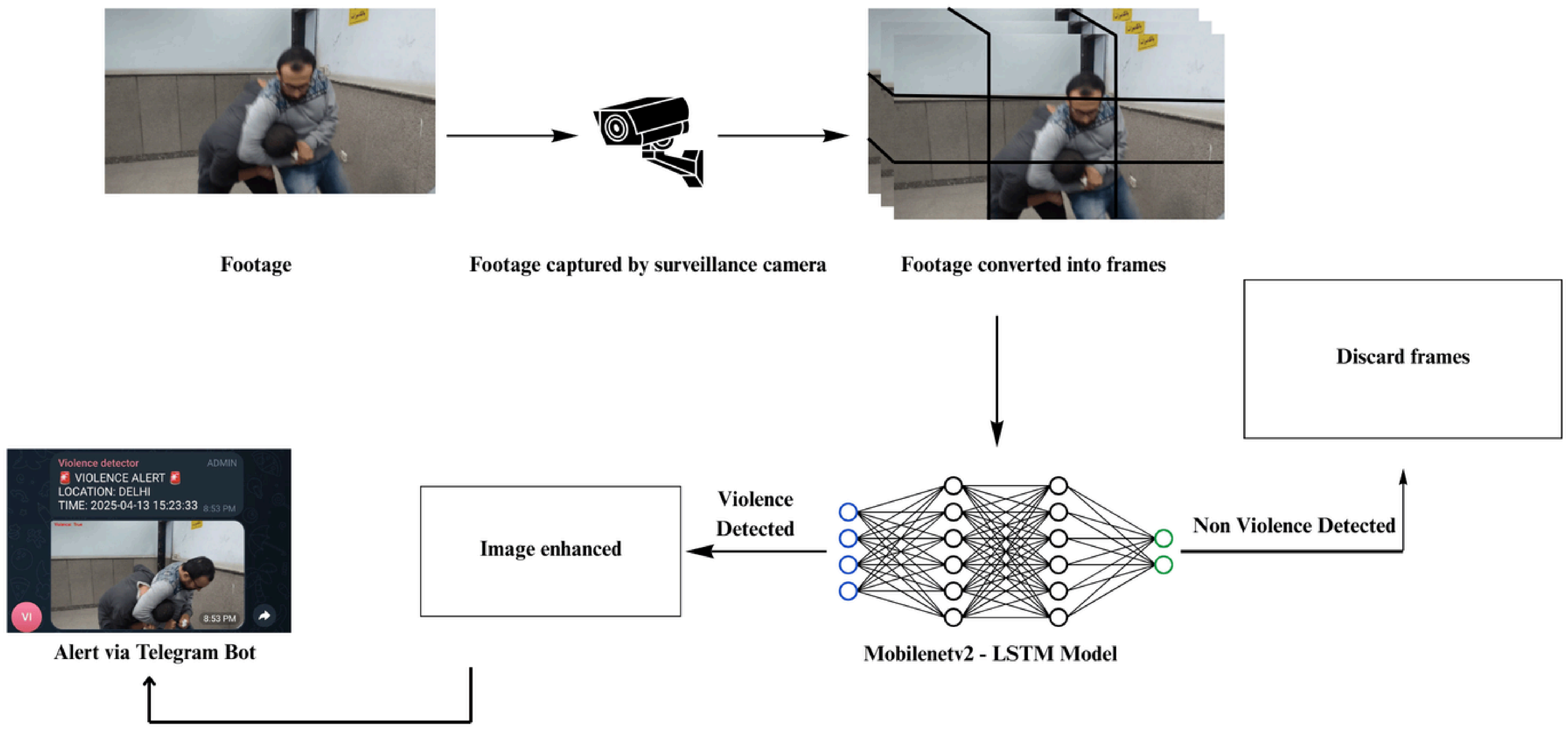


Proposed system

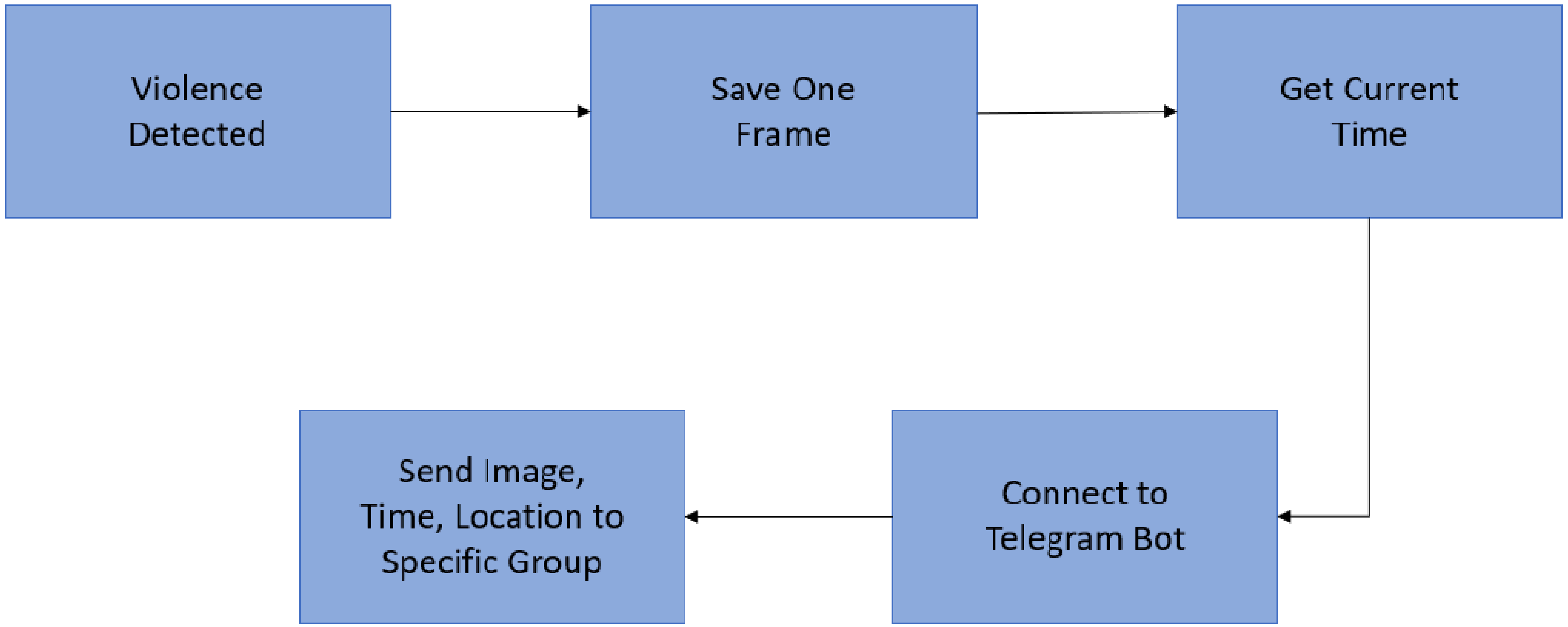
- The recommended technique is used for recognizing violent and non-violent incidents in real-time security camera footage is thoroughly examined in this section.
- In our system, we use MobileNetV2 and LSTM, along with hyperparameters that accurately identify the presence of violence in videos.
- When violence is detected, messages are sent to a telegram group via a telegram bot for real real-time



Architectural Diagram



Alert Module



Methodology

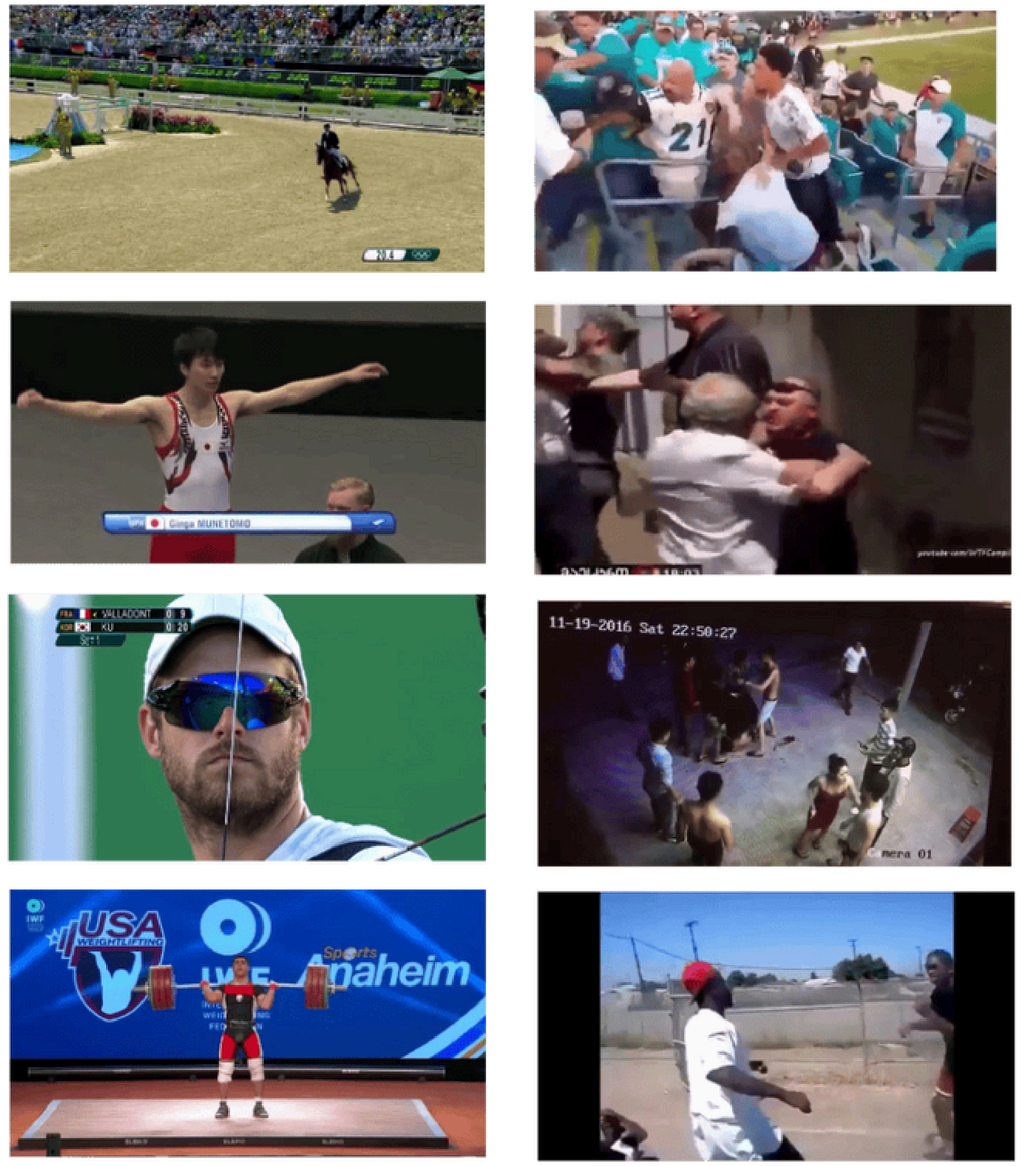
A. Dataset Collection
B. Data Preprocessing
C. Transfer Learning
D. Hyperparameter initialization
E. Performance Evaluation

Methodology

Data collection

- To test our methodology, we work with these two datasets.

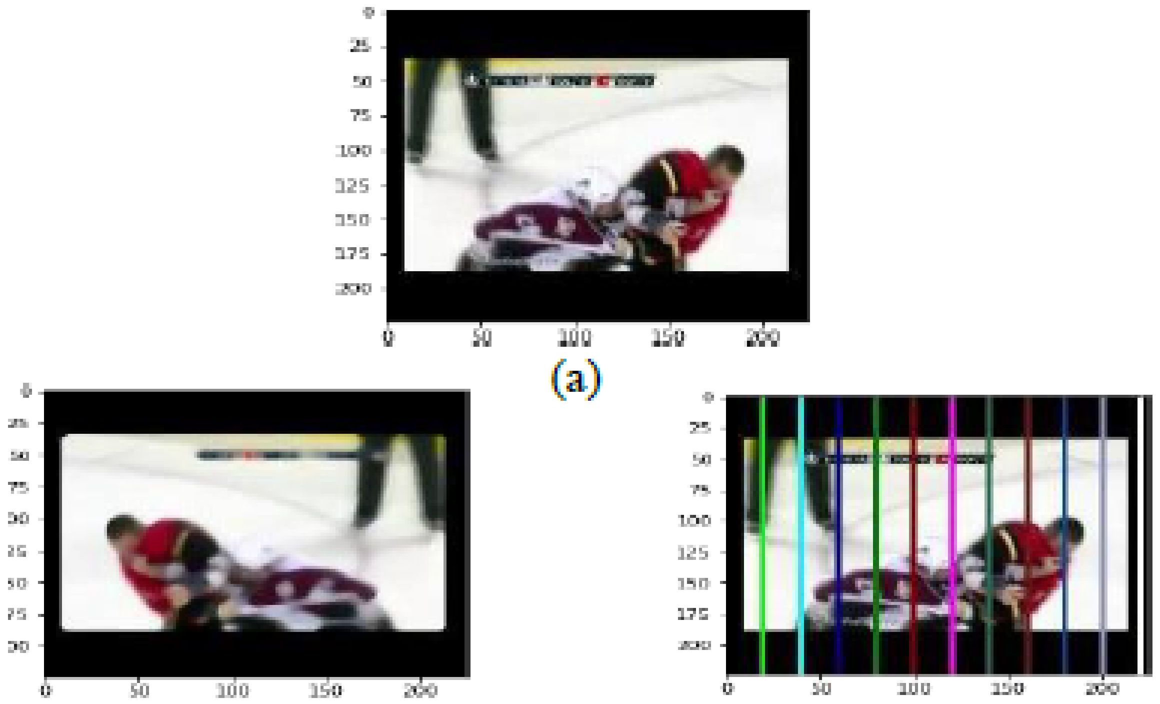
- Hockey Fight Dataset , which is standard in SOTA of violence detection, and Real Life Violence Situations Dataset, which is a relatively new dataset but has shown some promising results.



Methodology

Preprocessing

- Video clips are converted into frames and now data augmentation is applied.
- Techniques such as cropping, random rotate, horizontal flip, motionblurr , gaussianblur, etc, are applied via Albumentations Pipeline
- Also the both datasets are splitted into training set and testing set in 80% -20% ratio where 80% clips are used to train model and rest 20% clips are used to test model
- Few frames are skipped to avoid duplication.


(a)

Methodology

Tranfer Learning

- The suggested model is a binary image classification model
- MobileNetV2 act as a backbone supported by LSTM

.

- RelU introduces non-linearity, allowing the model to learn complex patterns also it makes training faster and avoids issues like vanishing gradients .
- Sigmoid in the final layer for binary classification output (Violent = 1, Non-Violent = 0)

Methodology

Hyperparameter initialization

- The initial learning rate (start_lr) was set to 0.00001, whereas the maximum learning rate (max_lr) used was 0.00005).
- To counteract overfitting, l2 regularization (0.005) was applied.
- ReduceLROnPlateau with an adaptive decreasing learning rate scheme if validation loss is not decreasing.
- ModelCheckpoint was used to track the weights of the best model in terms of validation accuracy.
- Batch size set to 16.
- epoch set to 50.
- Early stopping is also added to model .

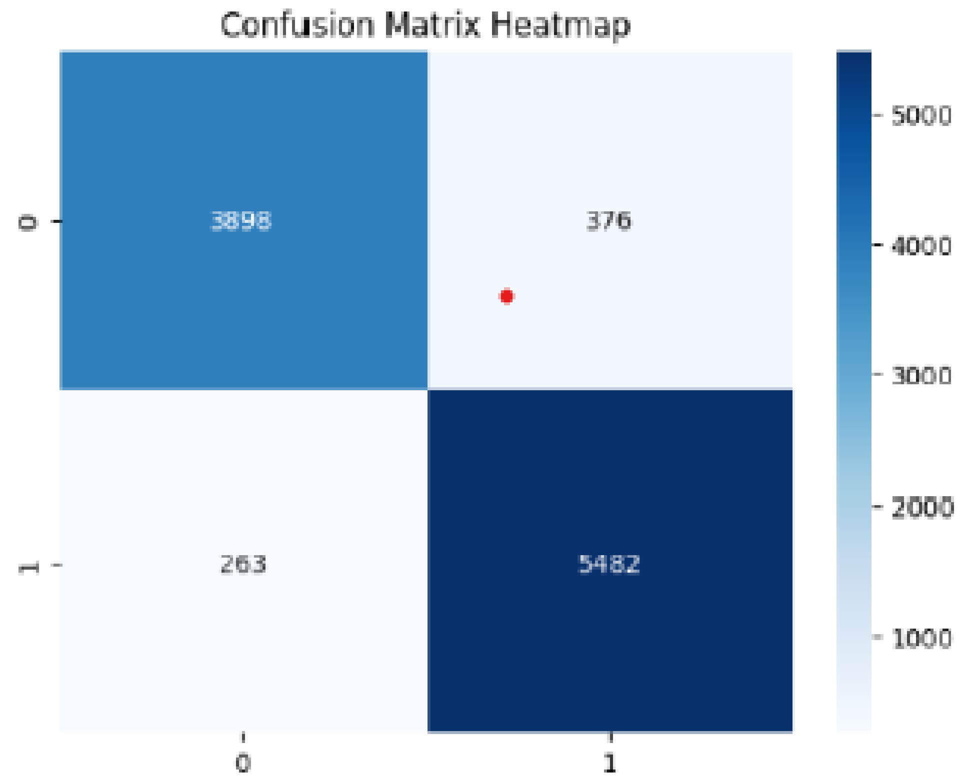
Methodology

Performance Evaluation

- For performance evaluation, the system uses
- confusion matrix
- accuracy
- precision
- recall
- Loss

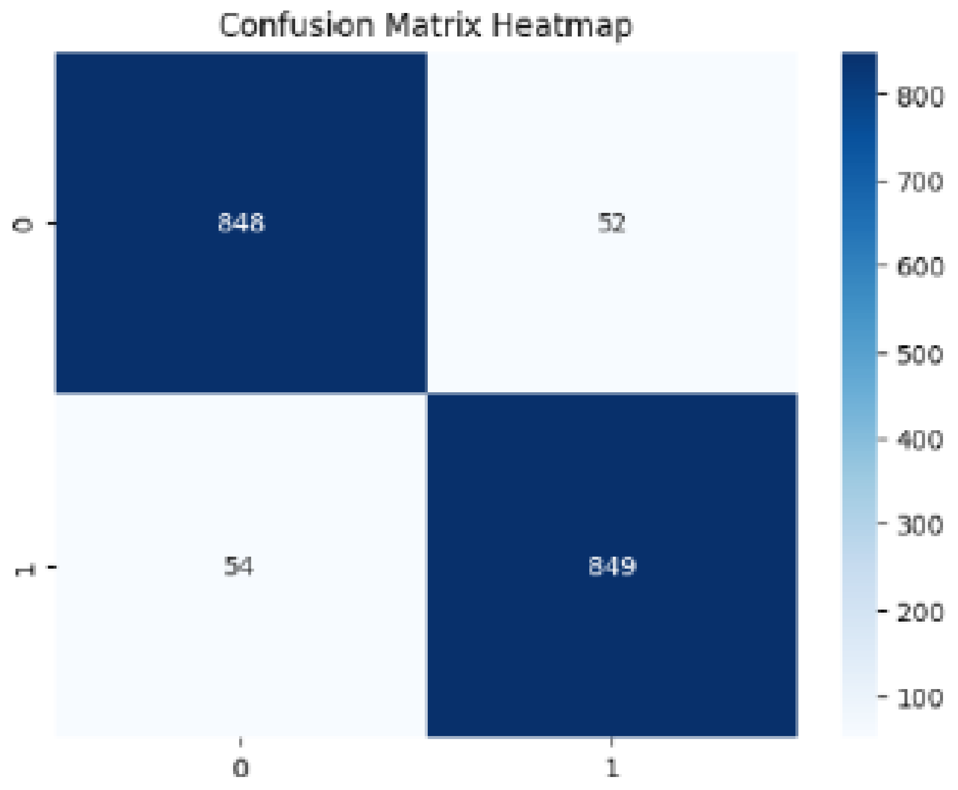
Result

> Correct Predictions: 9380
> Wrong Predictions: 639



	precision	recall	f1-score	support
NonViolence	0.94	0.91	0.92	4274
Violence	0.94	0.95	0.94	5745
accuracy			0.94	10019
macro avg	0.94	0.93	0.93	10019
weighted avg	0.94	0.94	0.94	10019

> Correct Predictions: 1697
> Wrong Predictions: 106

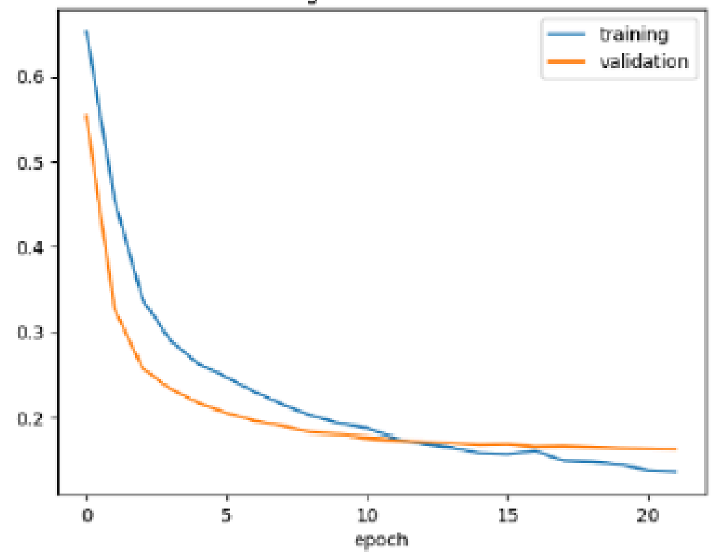


	precision	recall	f1-score	support
NonViolence	0.94	0.94	0.94	900
Violence	0.94	0.94	0.94	903
accuracy			0.94	1803
macro avg	0.94	0.94	0.94	1803
weighted avg	0.94	0.94	0.94	1803

Result

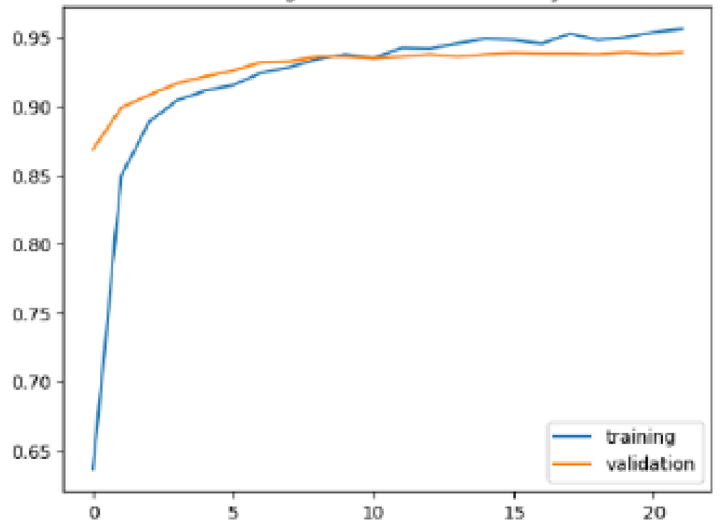
```
Accuracy on train: 0.9526641368865967   Loss on train: 0.14642678201198578
Accuracy on test: 0.941209077835083     Loss on test: 0.16338203847408295
```

Training and validation loss



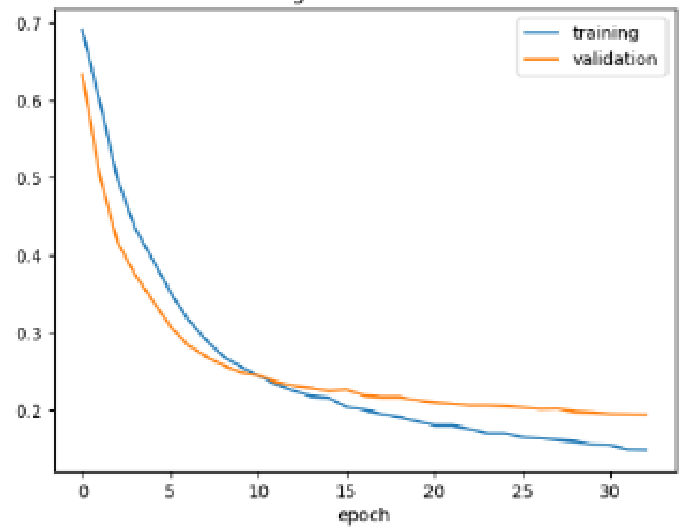
<Figure size 640x480 with 0 Axes>

Training and validation accuracy



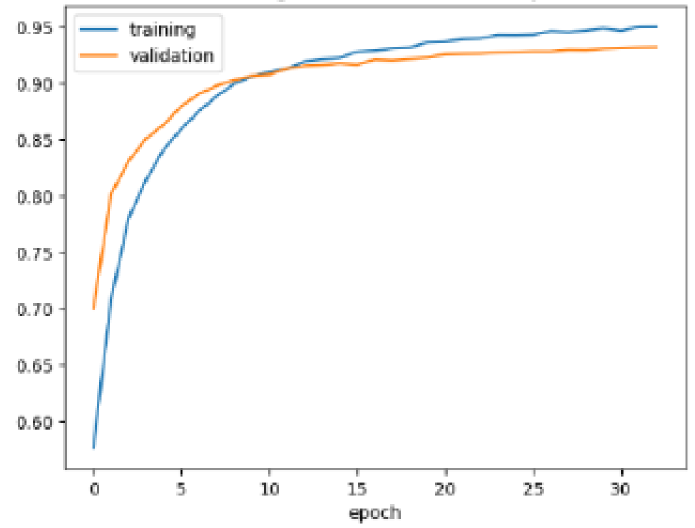

```
Best Epochs:  33
Accuracy on train: 0.9511913657188416   Loss on train: 0.14584754487405853
Accuracy on test: 0.936221182346344     Loss on test: 0.18665364384651184
```

Training and validation loss

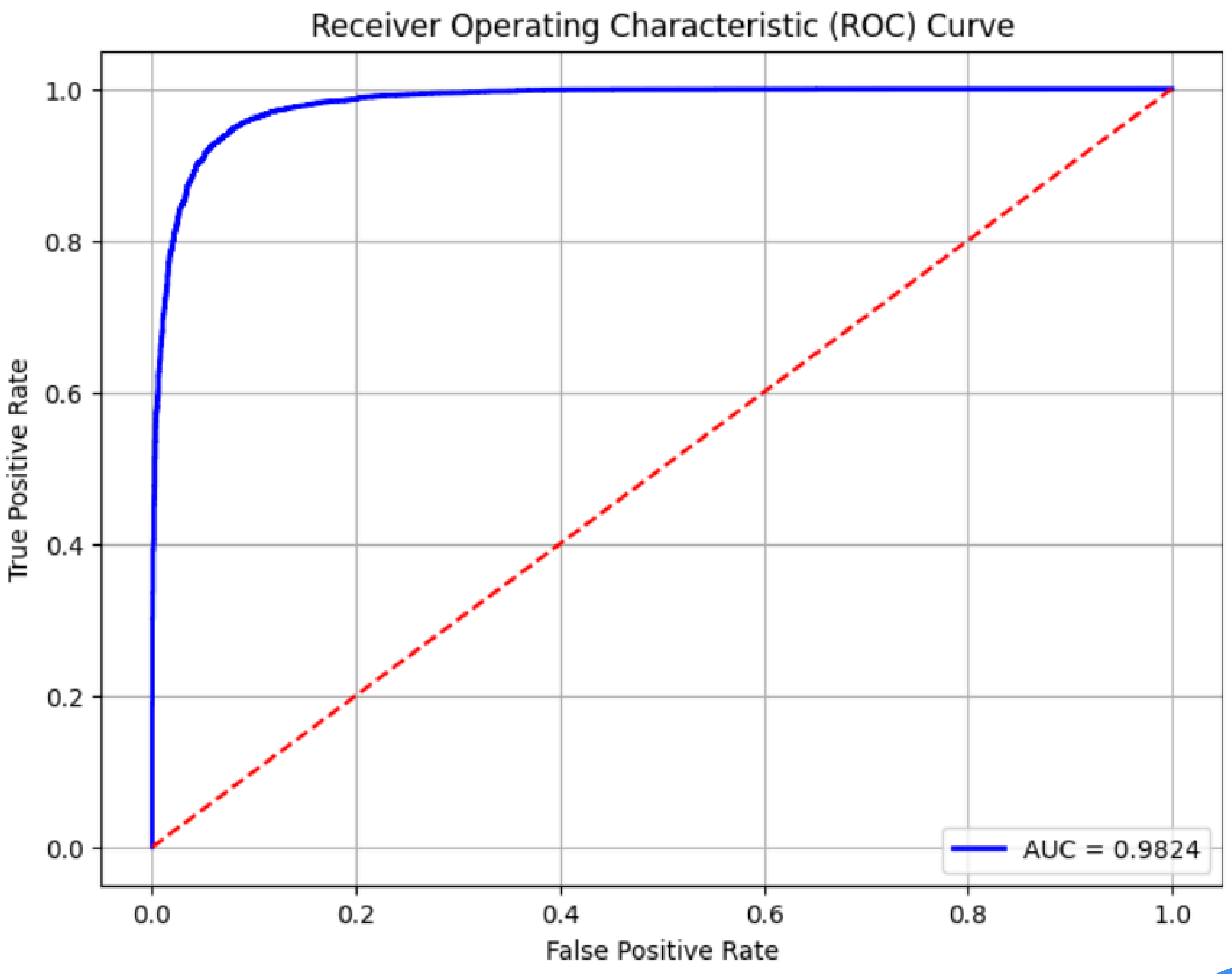


<Figure size 640x480 with 0 Axes>

Training and validation accuracy

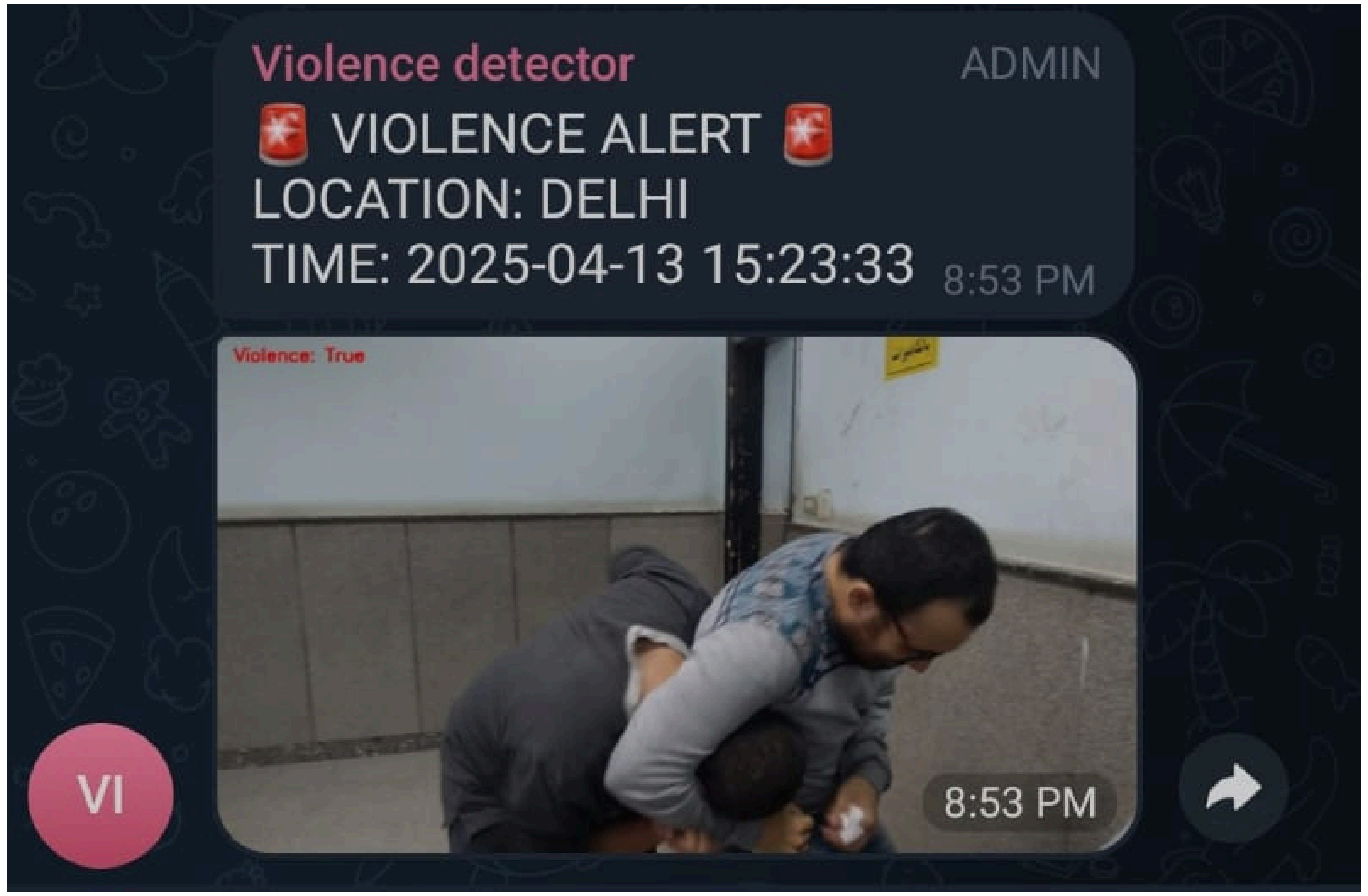


Result



Receiver Operating Characteristic (ROC) Curve
1.0
0.8
0.6
0.4
0.2
0.0
True Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate
AUC = 0.9824

Result



Violence detector
ADMIN
VIOLENCE ALERT
LOCATION: DELHI
TIME: 2025-04-13 15:23:33
8:53 PM
Violence: True
8:53 PM
VI

Refrences

[1] https://keras.io/

[2] https://www.tensorflow.org/

[3]https://opencv.org/

[4] https://matplotlib.org/

[5] Khan SU, Haq IU, Rho S, Baik SW, Lee MY. Cover the Violence: A Novel Deep-Learning-Based Approach Towards Violence-Detection in Movies. Applied Sciences. 2019;9(22):4963. https://doi.org/10.3390/app9224963J

Refrences

6] R. Parlika and A. Pratama, "The Online Test Application Uses Telegram Bots Version 1.0," Journal of Physics: Conference Series, vol. 1569, p. 022042, 2020. doi: 10.1088/1742-6596/1569/2/022042.

[7] M. -S. Kang, R. -H. Park and H. -M. Park, "Efficient Spatio-Temporal Modeling Methods for Real-Time ViolenceRecognition," in IEEE Access, vol. 9, pp. 76270-76285,2021, doi: 10.1109/ACCESS.2021.3083273.

[8] Ullah FUM, Ullah A, Muhammad K, Haq IU, Baik SW.Violence Detection Using Spatiotemporal Features with 3DConvolutional Neural Network. Sensors (Basel). 2019 May30;19(11):2472. doi: 10.3390/s19112472.

[9] A. -M. R. Abdali and R. F. Al-Tuma, "Robust Real-TimeViolence Detection in Video Using CNN And LSTM," 20192nd Scientific Conference of Computer Sciences (SCCS),2019, pp. 104-108, doi: 10.1109/SCCS.2019.8852616.

THANK YOU