RESEARCH ARTICLE OPEN ACCESS

Literature Survey for Women Safety and Threat Monitoring System

MD Faiz[1], Rishi Sharma[2], S Madan[3], Tejashwini[4]
Department of CSE-AI, Ballari Institute of Technology and Management, Ballari, Karnataka, India
Email: [1]faiz31807@gmail.com, [2]rishisharma21950@gmail.com,[3]smadanchinna7@gmail.com,[4]tejashwini09072005@gmail.com

--*************************----------------------------------

Abstract:

Women's safety is a serious issue in many countries. Crimes like harassment and assault are rising year after year. Over the past five years, many researchers have tried to use modern computer methods to help with this problem. This paper reviews fifteen research studies that focus on two related areas. The first area is detecting violent actions from video footage automatically. The second area is predicting which locations are more likely to become crime hotspots using past crime records. We have gone through each study carefully. For every study, we note what the authors did, what data they used, what results they got, and what limitations their work had. After reading all fifteen papers, we see that most of them work well inside a lab or a controlled setting. But very few systems have been tested in real streets or with real users. Also, most methods need a stable internet connection and a good quality camera. This review shows that there is still room for a simple, affordable, and automatic safety system that can work without internet and without any manual action from the user. Students and researchers who want to start work in this field will find this review useful

Keywords — **Women safety, violence detection, literature review, surveillance..**

--*************************----------------------------------

I. INTRODUCTION

Women's Every day we read news about crimes against women. Many of these crimes happen in public places like bus stops, parks, and empty streets. Governments have installed CCTV cameras in many cities. But these cameras only record what happens. Nobody watches them all the time. By the time someone watches the recording, the victim is already hurt.

In the last few years, researchers have tried to make computers watch these videos automatically. The idea is simple. If a computer can learn to recognize a violent action like punching or kicking, it can raise an alarm immediately. Similarly, if a computer can look at where crimes happened in the past, it can tell us which areas are dangerous today. A woman can then avoid walking through those areas.

This paper is only a review. We are not building any system here. We have collected fifteen research papers published between 2024 and 2026. Some papers focus on video violence detection. Some focus on crime hotspot prediction. A few papers try to combine both. For each paper, we explain what the authors did, what dataset they used, what numbers they reported, and what problems they faced. At the end, we summarize what the research community has achieved so far and what is still missing.

II. LITERATURE REVIEW

The literature review gives an overview of different methods used for women's safety, violence detection, and crime hotspot prediction. Past approaches were primarily reliant on manual CCTV monitoring, traditional image processing, and basic motion detection algorithms, whereas recent research is more centered on deep learning models to enhance accuracy and response time. To achieve improved results, researchers have considered different methods like CNN, LSTM, hybrid CNN-LSTM models, transformer networks, and spatial-temporal analysis for crime forecasting. This section brings out the key developments as well as demonstrates the necessity of having a simple and effective system such as the one presented in this project.

A. Merit et al. (2025) - Temporal Fusion for Violence Detection

Merit and his team published their work in Acta Polytechnica in 2025. They took three different CNN models — MobileNet V3, VGG16, and InceptionV3 — and attached an LSTM layer on top of each. The LSTM helped the model understand changes across multiple video frames. They trained everything on a dataset called RLVS, which contains real surveillance clips of fights and normal activities. The MobileNet V3 combined with LSTM gave the best result: 91.03% accuracy and an F1 score of 90.90%. The good thing about this work is that they focused on making the model fast. The weak point is that they only tested on one dataset. We do not know how the same model would perform on videos from a different city or a different camera angle

B. Explainable Framework for Violence Detection (Scientific Reports, 2026)

This paper came out in Nature's Scientific Reports in 2026. The authors added two extra ideas. First, they used an attention mechanism inside their CNN so that the model pays more attention to important parts of the video frame. Second, they used an unsupervised method to pick only the most important frames instead of processing every single frame. This made the model faster. They tested on five different datasets including RLVS and Hockey Fight. Their model reached an average accuracy of 94.6% and an F1 score of 93.9%. These numbers are better than many older models. One drawback is that the attention mechanism makes the model more complex. It may not run well on a cheap phone or a low-power device.

C. ResNet for Surveillance Violence (Journal of Telecommunications and IT, 2025)

This study compared three versions of ResNet — ResNet50V2, ResNet101V2, and ResNet152V2. Each ResNet was combined with a BiGRU layer instead of a regular LSTM. The authors found that deeper ResNet models gave better accuracy but were also slower. ResNet152V2 achieved the highest accuracy of 92.3% on their test set. However, the model took nearly twice as long to process each video clip compared to ResNet50V2. The authors concluded that for real-time use, ResNet50V2 is a better choice if speed matters more than a small gain in accuracy. This trade-off is important for anyone building a real system.

D. CNN-LSTM for Crime Forecasting (arXiv preprint, 2025)

This paper moved away from violence detection and looked at crime forecasting instead. The authors used a CNN-LSTM model to predict how many crimes would happen in different parts of a city on a given day. They tried different ways of grouping the crime counts. For example, they grouped counts into 5 bins, 10 bins, and 20 bins. The best results came from 10 bins. Their model achieved a hit rate of 87.5% when predicting the top 5% of risky areas. A limitation of this work is that they used only one city's data. Crime patterns can be very different from one city to another.

E. Hybrid STResNet + LSTM for Crime Hotspots (Scientific Reports, 2025)

This paper added an interesting new input to their model. Besides the usual crime location and time data, they also added the distance from each crime spot to the nearest park. Why parks? Because many violent crimes in their dataset happened near parks. Their model combined STResNet (a special CNN for spatial-temporal data) with LSTM. The extra park-distance channel improved their hit rate to 88.2% at a fine resolution of 500 meters. Without that extra channel, the hit rate was only 83.5%. So this shows that adding relevant external data helps a lot. The downside is that you need good maps and park location data for every city where you want to use the model.

F. Hybrid Transformer for Violence Recognition (Int. J. of Advanced Research, 2025)

Most violence detection papers use CNNs and LSTMs. This paper tried something different. They built a hybrid transformer model. Transformers are normally used for text, but here they adapted them for video. Their model had three parts: a multi-scale CNN to extract spatial features, a BiLSTM to handle time sequences, and a cross-attention transformer to connect everything. The authors claimed that this three-part design performed better than plain CNN-LSTM on their dataset. However, they did not report exact accuracy numbers in a clear way. Also, the model is very large. Training it required a powerful GPU with 24 GB of memory. This makes it hard for small labs or individual researchers to reproduce.

G. Intelligent Surveillance System (IGI Global book chapter, 2025)

This was a book chapter rather than a journal paper. The authors described a complete surveillance system that does four things: anomaly detection, violence recognition, object tracking, and person re-identification. For violence recognition, they used a ResNet50 followed by a 3D CNN. For anomaly detection, they used a Variational Autoencoder with LSTM. The chapter is useful because it shows how different modules can work together. But because it is a book chapter, there are no experimental results with exact numbers. It is more of a design proposal than a tested system.

H. IoT-Based Smart Safety for Women (Scientific Reports, 2025)

This paper took a completely different approach. Instead of using cameras, they used wearable sensors like heart rate monitors, temperature sensors, GPS, and accelerometers. The idea is that when a woman is in danger, her heart rate goes up and her body moves suddenly. A machine learning model (they used SVM) can detect this pattern and send an alert. They reported an extremely high accuracy of 99.7% with only 3 seconds of delay. That sounds impressive. But the study was done in a controlled environment with volunteers acting out stressful situations. Real fear and real attacks may produce different sensor signals. Also, wearing multiple sensors every day is not practical for most women.

I. Review of Deep Learning for Crime Prediction (ScienceDirect, 2025)

This is a review paper, not an original research paper. The authors collected and summarized more than 50 papers on crime prediction against women. They pointed out several common problems. First, most datasets are from Western cities like Chicago or Los Angeles. There are very few datasets from Asian or African cities. Second, many models work well on the training data but fail when applied to a different city or a different year. Third, very few papers have actually deployed their models in the real world. This review is helpful because it tells us what the research gaps are. It also warns about ethical issues — for example, a crime prediction model might unfairly label a poor neighborhood as dangerous just because more police patrol there and report more crimes.

J. MobileNet V2 for Women's Safety (IEEE Conference, 2025)

This conference paper focused on building a light-weight model that can run on a mobile phone. The authors used MobileNet V2 (a small CNN designed for phones) and trained it to classify video clips as violent or non-violent. They achieved 89% accuracy on a small dataset of their own collection. The main contribution is the low computational cost. The model ran at 25 frames per second on a mid-range Android phone. That is fast enough for real-time use. However, their dataset was small (only 500 clips) and may not represent the variety of real-world violence. Also, they did not test the model in low light or crowded scenes.

K. Crime Hotspot Prediction with XGBoost (RIT Thesis, 2025)

This master's thesis from Rochester Institute of Technology compared three machine learning models — Random Forest, Gradient Boosting, and XGBoost — for predicting violent crime hotspots. The dataset had 769,680 violent crime incidents. The author used kernel density estimation to create heatmaps first, then applied the three models. Random Forest gave the best AUC score of 0.9365. XGBoost was very close at 0.9312. The thesis is thorough and well documented. But it is a thesis, not a peer-reviewed paper. Also, the method predicts hotspots at the level of large areas (census blocks) which may be too big for a woman trying to decide whether a particular street is safe.

L. CHART Framework (ScienceDirect, 2024)

CHART stands for something like Crime Hotspot Alert and Real-Time tracking. The authors used Kernel Density Estimation to generate heatmaps and Random Forest to predict crime likelihood. They tested on three different crime detection tasks. Their accuracy numbers were very high: 95.24%, 96.12%, and 94.68% across the three tasks. One interesting feature of CHART is that it updates its predictions in near real-time as new crime reports come in. But the paper does not clearly explain how fast this update happens. Is it every minute, every hour, or every day? Without this detail, it is hard to call it "real-time".

M. Crime Hotspot Classification Using K-Means (IEEE, 2025)

This IEEE paper compared two popular methods for finding crime hotspots: kernel density estimation (KDE) and K-means clustering. Surprisingly, K-means gave almost the same results as KDE but was much faster to compute. The authors also added a time component — they looked at which hours of the day crimes were most common in each hotspot. For example, one hotspot might be dangerous only between 10 PM and 2 AM. This temporal information is very useful for a woman planning her travel. The limitation is that the paper only tested on one year of data from a single city. More testing is needed.

N. Real-Life Violence Detection with Attention (arXiv preprint, 2025)

This preprint (not yet peer-reviewed) introduced a new dataset called RLVSD (Real Life Violence Surveillance Dataset) which contains 2,500 video clips of real fights, robberies, and assaults recorded from actual security cameras. The authors also trained a model with spatial attention and temporal attention. Their model achieved 93.1% accuracy on their new dataset. The main contribution is the dataset itself — most other papers use hockey fights or movie scenes, which are not the same as real surveillance footage. However, because the paper is only a preprint, the work has not been verified by independent reviewers.

O. Combined Violence Detection and Crime Prediction (No clear venue, 2025)

This is the only paper in our list that tried to do both violence detection and crime prediction in a single system. The authors built a web-based dashboard. One part of the dashboard takes live video from a webcam and runs a CNN-LSTM model to detect fights. The other part shows a color-coded map of the city based on historical crime data. A user can click on any location and see a "safe" or "unsafe" label. The system was tested briefly with 10 users who gave positive feedback. But the paper lacks proper evaluation metrics. No accuracy numbers are reported for either the violence detector or the crime predictor. So it is more of a demonstration than a rigorous study.

III. CONCLUSIONS

After reviewing these fifteen papers, we can see a clear pattern. Many researchers have built good models for detecting violence from videos. Many others have built good models for predicting crime hotspots. A few have even tried to put both together. But the models that work well in a lab setting often fail when taken outside. They need good lighting, high-quality cameras, and fast internet. Also, most systems still require the user to do something — like opening an app or wearing a special device.

Another gap we noticed is the lack of real-world testing. Very few papers have deployed their system on actual streets or given it to real women to use. Without such testing, we cannot be sure that these models will actually help in an emergency. There is also a shortage of datasets from developing countries where women's safety is a big problem.

In short, the research community has done a lot of good work. But there is still a long way to go before a woman can truly rely on an automatic, offline, and affordable safety system. We hope this review helps other researchers find direction for their future work.

REFERENCES

[1] K. Merit, M. Beladgham, and A. Taleb-Ahmed, "Temporal fusion strategy for violence detection: utilising convolutional and LSTM neural

networks for surveillance videos," *Acta Polytechnica*, vol. 65, pp. 306–319, 2025.
[2] "An explainable deep learning framework for video violence detection using unsupervised keyframe selection and attention-based CNN," *Scientific Reports*, vol. 16, Article 11098, 2026.
[3] "AI-based Violent Incident Detection in Surveillance Videos to Enhance Public Safety," *Journal of Telecommunications and Information Technology*, 2025.
[4] L. Mao, "Crime Forecasting: A Spatio-temporal Analysis with Deep Learning Models," *arXiv preprint arXiv:2502.07465*, 2025.
[5] "Hybrid ST-ResNet and LSTM approach for precise crime hotspot prediction," *Scientific Reports*, vol. 15, Article 40754, 2025.
[6] "Innovative Hybrid Transformer Model for Intelligent Violence Recognition in Surveillance Systems," *International Journal of Advanced Multidisciplinary Research*, 2025.
[7] M. Evany Anne, M. Brindha, and N. Sivakumaran, "Enhancing Intelligent Surveillance: Hybrid Deep Learning for Threat Detection, Violence Recognition, and Person Re-Identification," in *Computer Vision and Internet of Everything (IoE) for Societal Needs*, IGI Global, 2025, ch. 13.
[8] "Evaluation of IoT based smart safety systems for women and children using machine learning techniques," *Scientific Reports*, vol. 16, Article 87, 2025.
[9] "Crime prediction against women in surveillance videos using deep learning models: A review," *ScienceDirect*, 2025.
[10] "Smart Safety Solutions: Utilizing MobileNet v2 for Violence Recognition to Enhance Women's Safety," *IEEE Conference*, May 2025.
[11] M. Alhammadi, "Predicting Violent Crime Hotspots," M.S. thesis, Rochester Institute of Technology, 2025.
[12] "CHART: Intelligent Crime Hotspot Detection and Real-Time Tracking Using Machine Learning," *Computers, Materials and Continua*, vol. 81, no. 3, pp. 4171–4194, Dec. 2024.
[13] "Crime Hotspot Classification Using Machine Learning," *IEEE Conference*, 2025.
[14] "Real-Life Violence Detection with Spatial and Temporal Attention," *arXiv preprint arXiv:2506.13910*, 2025.
[15] N. Parveen et al., "WOMEN SAFETY ANALYTICS – PROTECTING WOMEN FROM SAFETY THREAT," *Lex Localis*, 2025