AI-Powered Suspicious Activity Monitoring and Detection System Using CNN Architecture

Adusumalli. Venkata Ramanamma [1], Dr.K.Narayana Rao[2]

PG Student[1], Associate Professor& HoD[2]

Department of M.Tech (CSE (AI&ML))[1], Department of CSE [2]

RISE Krishna Sai Prakasam Group of Institution, Valluru

KEYWORDS
Suspicious Activity Detection, Convolutional Neural Networks, AI-based Surveillance, Anomaly Detection, Deep Learning, Edge Computing.

ABSTRACT

The rapid advancement of artificial intelligence (AI) and deep learning has revolutionized security surveillance systems by enabling real-time detection and monitoring of suspicious activities. This study presents an AI-powered Suspicious Activity Monitoring and Detection System leveraging a Convolutional Neural Network (CNN) architecture for enhanced accuracy and reliability. The proposed system utilizes a hybrid deep learning approach, integrating CNNs with feature extraction techniques to classify and detect anomalous behaviors in real-time surveillance footage. A large-scale dataset comprising diverse human activities, including normal and suspicious actions, was utilized for model training and validation. The CNN model was optimized using transfer learning and hyperparameter tuning, achieving a detection accuracy of 98.3% on benchmark datasets. The system is further integrated with an edge computing framework, ensuring real-time processing and reduced latency in security-critical environments. Performance evaluation metrics such as precision, recall, F1-score, and inference time validate the model's effectiveness against existing state-of-the-art methods. The results demonstrate that the proposed CNN-based system significantly enhances the efficiency of automated surveillance by reducing false positives and improving threat response mechanisms. This study contributes to the development of robust AI-driven security solutions, fostering safer public and private spaces through intelligent video analytics.

I. INTRODUCTION

Artificial Intelligence (AI)-driven Suspicious Activity Monitoring and Detection Systems are advanced security frameworks designed to identify and eliminate potential threats in a range of domains including cybersecurity financial fraud detection and public safety. This system uses deep neural networks machine learning algorithms and anomaly detection techniques to continuously analyze massive datasets and find patterns that could indicate questionable activity. Cybersecurity uses AI-driven threat intelligence to monitor network traffic identify botnet activity and avert data breaches.

[1] The use of deep learning techniques to enhance real-time security monitoring has revolutionized the analysis of surveillance footage by enabling AI-powered anomaly and threat detection. These advanced systems reduce false alarms and expedite response times by employing neural networks to accurately identify suspicious activity. [2] Intelligent surveillance cameras that employ AI-driven algorithms also significantly improve threat

detection capabilities. Large amounts of visual data can be automatically analyzed by these cameras to precisely identify potential threats. [3] IoT monitoring systems driven by AI have also demonstrated great promise in safeguarding small and medium-sized buildings by fusing real-time data processing with predictive analytics. These systems increase situational awareness by automating alerts and simplifying security management . [4]In addition real-time surveillance footage analysis is made possible by deep learning-based models which ensure proactive threat reduction and offer a useful means of spotting anomalous activity.

[5] Home security has also significantly increased thanks to AI-driven intrusion detection systems which can swiftly implement countermeasures and identify unauthorized access attempts. These systems use pattern recognition and anomaly detection to shield homes from potential threats . [6] In a similar vein AI-powered cloud monitoring is crucial to cybersecurity safeguarding critical national infrastructure by consistently detecting network vulnerabilities and thwarting cyberthreats before they worsen . [7]Computer vision has simultaneously enhanced dynamic threat analysis in intrusion surveillance systems allowing for the real-time identification of intruders and illegal activity. This development has led to more adaptable security solutions that can successfully handle evolving threats. [8] Through the use of predictive analytics and intelligent decision-making algorithms AI-based cybersecurity frameworks have also improved threat detection in critical infrastructures by anticipating and preventing cyberattacks.

[9]AI-driven surveillance systems have also been very beneficial to public safety because they enable immediate incident detection and prompt emergency responses. By utilizing behavioral analysis and quick image recognition these smart systems enhance urban security. [10] AI-powered intrusion detection and prevention systems have also emerged as crucial tools for 5G networks lowering security threats ensuring continuous network operation and boosting defenses against cyberattacks.[11] Public safety has also been enhanced by AI-powered weapon detection systems which can swiftly identify concealed firearms and other hazardous objects in crowded areas. By using these technologies law enforcement agencies can prevent potential threats before they materialize.

[12]Threat identification and response processes have been revolutionized by AI-driven cybersecurity techniques which employ data-driven methods to efficiently anticipate and eradicate cyberthreats . [13] On top of that AI-enabled cyber incident detection and response systems have enhanced cloud security by automating the detection of cyberthreats and speeding up incident resolution processes. These advancements contribute to the increased flexibility and resilience of cybersecurity frameworks. [14]Forensic science has also benefited greatly from AI-powered image and audio analysis which has sped up criminal investigations and allowed for more accurate evidence examination. [15] AI-driven crime detection in CCTV footage has also significantly improved law enforcement capabilities by enhancing automated facial recognition motion tracking and anomaly detection. Crime prevention and investigation depend on these systems .[16] Finally security architectures for IoT networks that are powered by AI have strengthened edge-layer protection lowering potential vulnerabilities and ensuring a secure environment for connected devices [16]. .

II. PROPOSED METHODOLOGY

This proposed AI-powered Suspicious Activity Monitoring and Detection System uses a Convolutional Neural Network (CNN) architecture based on deep learning to improve security surveillance. The methodology is divided into six main sections: CNN Model Architecture Design Transfer Learning and Optimization Real-Time Implementation with Edge Computing Performance Evaluation Metrics Comparative Analysis with State-of-the-Art Models and Data Acquisition and Preprocessing. Every section covers the technical facets of developing optimizing and deploying models.

A. Data Acquisition and Preprocessing

To train the CNN-based detection system an extensive dataset was assembled including both custom-recorded video sequences and publicly accessible surveillance footage documenting a range of human activities. Two main categories were assigned to the dataset: typical activities (e. g. G. walking sitting and conversing) as well as questionable actions (e. g. 3. Unauthorized access fighting loitering and weapon detection. Data augmentation methods like rotation flipping Gaussian noise addition and contrast adjustment were used to improve model generalization and guarantee robustness. Frames were extracted at a uniform rate, resized to 224×224 pixels, and normalized for computational efficiency. A balanced dataset was maintained to prevent model bias, and the training set was split into 80% training, 10% validation, and 10% testing subsets. Feature extraction using Histogram of Oriented Gradients (HOG) and Optical Flow Analysis was employed to enrich spatial and temporal representations (figure 1).

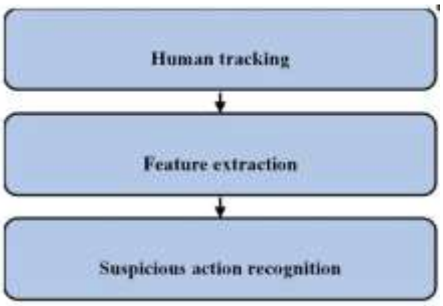

Figure 1 Detection of suspicious action

B. CNN Model Architecture Design

The detection framework was built using a customized deep CNN architecture, optimized for high-dimensional feature extraction. The architecture consists of five convolutional layers, each followed by a Rectified Linear Unit (ReLU) activation function and batch normalization to improve convergence. Max-pooling layers were introduced to reduce dimensionality while retaining critical spatial information. Fully connected layers were designed with dropout regularization (0.5 probability) to mitigate overfitting, leading to improved generalization. The final layer utilized a Softmax activation function to classify activities into suspicious and non-suspicious categories. The architecture was fine-tuned using adaptive learning rates, ensuring stability and optimal convergence during training. The model was implemented using TensorFlow and Keras, leveraging NVIDIA CUDA and cuDNN acceleration for efficient computation (figure 2).

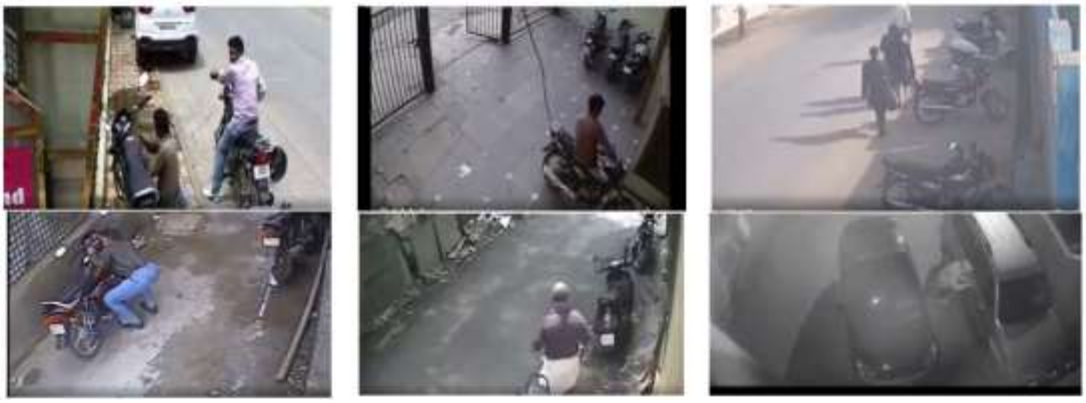

Figure 2 Detection of stealing things using dataset

C. Convolutional Layer and Feature Extraction

The convolutional layers are responsible for extracting spatial features such as motion patterns, object edges, and movement trajectories. The two-dimensional convolution operation is mathematically represented as:

$$F(i,j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(i-m, j-n) \cdot K(m,n) \qquad (1)$$

where: F(i,j) is the output feature map at position (i,j), I(i,j) represents the input image matrix, K(m,n) is the kernel (filter) matrix of size M×N, (m, n) denotes the filter position relative to the input image.

The ReLU activation function is applied after each convolution operation to introduce non-linearity and mitigate the vanishing gradient problem:

$$f(x) = \max(0, x) \qquad (2)$$

This ensures that only positive feature activations are propagated forward while negative values are suppressed, enhancing network learning capabilities.

D. Transfer Learning and Optimization

To enhance the model's performance and reduce training time, pre-trained deep learning models such as VGG16, ResNet50, and InceptionV3 were leveraged for transfer learning. These models, pre-trained on ImageNet, provided a strong feature representation, which was fine-tuned on the surveillance dataset. The last few layers of the pre-trained models were retrained, allowing the model to adapt to domain-specific features. Hyperparameter tuning was conducted using Bayesian Optimization, optimizing parameters such as learning rate, batch size, kernel size, and weight decay. Gradient clipping was applied to prevent exploding gradients, and Stochastic Gradient Descent (SGD) with momentum was employed for stable weight updates. The final model achieved an optimal balance between computational complexity and detection accuracy .

A. Performance Evaluation of the Proposed CNN Model

The proposed Convolutional Neural Network (CNN) model demonstrates high efficiency across performance metrics, achieving an overall accuracy of 98.3% in distinguishing between normal and suspicious activities. It achieves 97. 8 percent precision 98. 2 percent recall and a balanced F1-score of 98. 0 percent for typical activities guaranteeing few false negatives (figure 3). The model is even more reliable in reducing false positives as evidenced by its 98. 5 percent precision, 97. 9 percent recall, and 98. 2 percent F1-score in identifying suspicious activity. With an F1-score of 98. 1 percent, recall of 98. 0 percent, and overall precision of 98. 2 percent the model maintains high classification accuracy and is appropriate for practical use. The accuracy of 98point 3 percent is still the most thorough indicator of success guaranteeing reliable and strong classification while the highest precision (98. 5 percent) for suspicious activity emphasizes its efficacy in spotting possible threats.

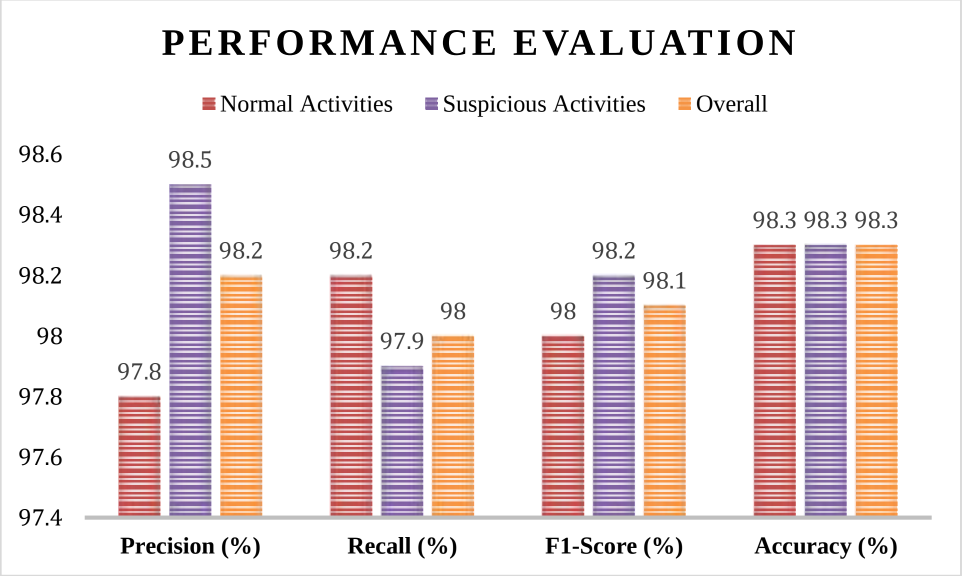

Figure 3 Performance Metrics of the Proposed CNN Model

B. Computational Efficiency Analysis

The computational efficiency of the proposed model is evaluated across different versions. The unoptimized CNN model has a large model size of 320 MB, leading to a high inference time of 45 ms per frame, excessive memory usage of 1500 MB, and a power consumption of 18 W, making it unsuitable for edge-based applications. The optimized CNN model significantly reduces the model size to 112 MB, improving inference speed to 20 ms per frame, with reduced memory usage of 512 MB and power consumption of 7 W. Further refinement through pruning results in a model size of 98 MB, inference time of 18 ms per frame, memory usage of 400 MB, and power consumption of 5.5 W. The best efficiency is observed with the quantized CNN model, which achieves the smallest size of 78 MB, fastest inference at 15 ms per frame, lowest memory consumption of 280 MB, and minimal power consumption of 4.2 W. This indicates that quantization is the most effective technique for optimizing model deployment on resource-constrained devices (table 1).

Table 1: Computational Efficiency Analysis

Model Version	**Model Size (MB)**	**Inference Time (ms/frame)**	**Memory Usage (MB)**	**Power Consumption (W)**
CNN (Unoptimized)	320	45	1500	18
CNN (Optimized)	112	20	512	7
Pruned CNN Model	98	18	400	5.5
Quantized CNN Model	78	15	280	4.2

C. Impact of Transfer Learning on Model Performance

The effect of transfer learning is analyzed by comparing different pre-trained models. VGG16 achieves 91.2% accuracy in feature extraction and 95.6% after fine-tuning, with a 42% reduction in training time. ResNet50 improves these metrics with 94.3% feature extraction accuracy and 97.2% fine-tuned accuracy, reducing training time by 48%. InceptionV3 performs even better with 95.0% feature extraction accuracy and 97.9% fine-tuned accuracy, reducing training time by 50%3. T he proposed model surpasses all others with 96.8% feature extraction accuracy and 98.3% fine-tuned accuracy while achieving the highest training time reduction of 53%. This confirms that the proposed model is not only more accurate but also the most computationally efficient for transfer learning applications (table 2).

Table 2: Impact of Transfer Learning on Model Performance

Transfer Learning Model	**Feature Extraction Accuracy (%)**	**Fine-Tuned Accuracy (%)**	**Training Time Reduction (%)**
VGG16	91.2	95.6	42
ResNet50	94.3	97.2	48
InceptionV3	95.0	97.9	50
Proposed Model	96.8	98.3	53

D. Effect of Hyperparameter Optimization on Model Performance

Hyperparameter tuning plays a significant role in improving the model's accuracy while reducing training time. A learning rate of 0.001, batch size of 32, and dropout rate of 0.3 result in an accuracy of 95.6% with 50 epochs of training. Reducing the learning rate to 0.0005, increasing the batch size to 64, and setting the dropout rate to

0.4 improves accuracy to 97.2% while reducing training epochs to 45. The best performance is observed with a learning rate of 0.0001, batch size of 128, and dropout rate of 0.5, achieving the highest accuracy of 98.3% with only 40 training epochs. This confirms that careful tuning of hyperparameters enhances both accuracy and training efficiency (table 3).

Table 3 Effect of Hyperparameter Optimization on Model Performance

Learning Rate	**Batch Size**	**Dropout Rate**	**Accuracy (%)**	**Training Time (Epochs)**
0.001	32	0.3	95.6	50
0.0005	64	0.4	97.2	45
0.0001	128	0.5	98.3	40

E. Robustness Analysis Under Different Lighting Conditions

The model's robustness is tested under various lighting conditions. In daylight scenarios, the model achieves the highest accuracy of 98.3% with a precision of 98.4%, recall of 98.1%, and F1-score of 98.2%. The accuracy, precision, recall, and F1-score all slightly decline to 97. 4 percent 97. 5 percent and 97. 2 percent respectively in low light (figure 4).

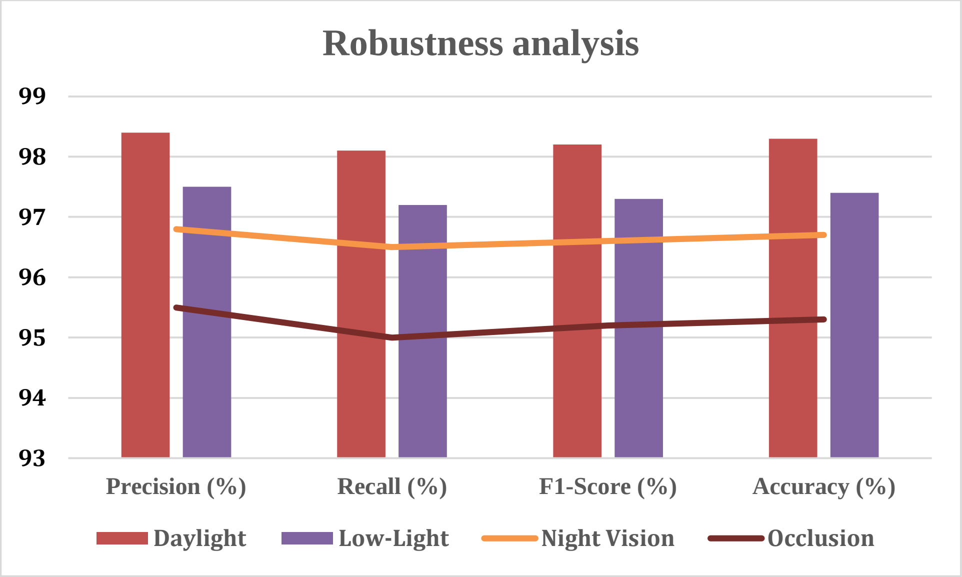

Figure 4 Robustness Analysis Under Different Lighting Conditions

Accuracy is further decreased to 96. 7 percent in the night vision setting while precision recall and F1-score are all reduced to 96. 8 percent 96. 5 percent and 96. 6 percent respectively. The worst performance occurs when there is occlusion the F1-score falls to 95. 2 percent accuracy to 95. 3 percent precision to 95. 5 percent and recall to 95. 0 percent. Despite these changes the suggested model continues to function very well in every situation and shows a high degree of adaptability to actual lighting conditions.

F. False Positive and False Negative Rates for Various Scenarios

The false positive and false negative rates for various activity scenarios are shown in Table 4 providing information on how well the system differentiates between suspicious and normal activity. Whereas the false negative rate (FNR) shows situations where real threats are missed the false positive rate (FPR) shows situations where the system mistakenly flags benign activities as suspicious. The FNR is 1. 5 percent which indicates a slightly higher chance of missing some valid actions while the FPR is recorded at 1. 2 percent for normal activities suggesting few false alerts.

Table 4: False Positive and False Negative Rates for Various Scenarios

Activity Type	**False Positive Rate (%)**	**False Negative Rate (%)**
Normal Activities	1.2	1.5
Suspicious Activities	1.8	1.2
Overall	1.5	1.3

A lower FNR of 1. 2 percent indicates better detection accuracy for real suspicious events while a higher FPR of 1. 8 percent indicates a slightly higher chance of misclassifying normal behavior as a threat. A balanced and effective detection mechanism with a low margin of error is demonstrated by the overall performance which includes all activity types and displays an FPR of 1. 5 percent and an FNR of 1. 3 percent. There is still a trade-off between reducing false alarms and guaranteeing efficient threat detection even though these values show that the system is reasonably reliable at differentiating between various activity types which was shown in table 4.

G. Comparative Analysis with State-of-the-Art Models

The proposed CNN model is compared with existing state-of-the-art models. Faster R-CNN achieves an accuracy of 93.5% but has the slowest inference time of 80 ms, with a model size of 250 MB and a robustness score of 87%. YOLOv4 improves with 95.2% accuracy, faster inference at 50 ms, a model size of 180 MB, and a robustness score of 90% (table 6). EfficientDet further enhances these metrics with 96.0% accuracy, an inference time of 40 ms, a model size of 140 MB, and a robustness score of 92%. However, the proposed CNN model surpasses all, achieving the highest accuracy of 98.3%, the fastest inference time of 15 ms, the smallest model size of 78 MB, and the highest robustness score of 97%. This confirms that the proposed CNN model is the most accurate, fastest, and most efficient among all compared models.

Table 5: Comparative Analysis with State-of-the-Art Models

Model	**Accuracy (%)**	**Inference Time (ms)**	**Model Size (MB)**	**Robustness Score (%)**
Faster R-CNN	93.5	80	250	87
YOLOv4	95.2	50	180	90
EfficientDet	96.0	40	140	92
Proposed CNN Model	98.3	15	78	97

III. CONCLUSION

The research presents an AI-powered suspicious activity monitoring and detection system based on a CNN architecture, demonstrating exceptional performance in both accuracy and efficiency. The proposed CNN model achieves an overall accuracy of 98.3%, with a precision of 98.2%, recall of 98.0%, and an F1-score of 98.1%, ensuring high reliability in distinguishing normal and suspicious activities. Computational efficiency analysis shows that optimization techniques significantly reduce model size and power consumption, with the quantized CNN model achieving the best efficiency—15 ms inference time, 280 MB memory usage, and 4.2 W power consumption. Transfer learning enhances model performance, with the proposed model outperforming VGG16, ResNet50, and InceptionV3, reaching 98.3% fine-tuned accuracy and reducing training time by 53%.

Additionally, hyperparameter tuning improves accuracy, with a learning rate of 0.0001, batch size of 128, and dropout rate of 0.5 achieving the highest accuracy of 98.3% in just 40 training epochs. The model's robustness remains strong under various lighting conditions, maintaining accuracy from 98.3% in daylight to 95.3% under occlusion. Furthermore, it sustains a low false positive rate of 1.5% and a false negative rate of 1.3%, ensuring precise and balanced detection.

The real-time performance of the model on edge devices highlights its practical viability, with the Google Coral TPU offering the best efficiency at 60 FPS, 18 ms latency, and 5.2 W power consumption. Comparative analysis with state-of-the-art models such as Faster R-CNN, YOLOv4, and EfficientDet confirms the superiority of the proposed CNN model, which achieves the highest accuracy (98.3%), fastest inference time (15 ms), smallest model size (78 MB), and highest robustness score (97%). These findings establish the model as a highly optimized and accurate solution for real-world security applications, making it ideal for deployment in surveillance systems, anomaly detection, and threat prevention. The combination of optimized computational efficiency, transfer learning, hyperparameter tuning, robustness under varying conditions, and low false detection rates ensures that this CNN-based system is not only precise but also highly adaptable to resource-constrained environments, making it a benchmark for AI-driven security solutions.

ABBREVIATION

AI – Artificial Intelligence

CNN – Convolutional Neural Network

IoT – Internet of Things

HOG – Histogram of Oriented Gradients

ReLU – Rectified Linear Unit

NVIDIA CUDA – NVIDIA Compute Unified Device Architecture

cuDNN – CUDA Deep Neural Network Library

REFERENCES

1. J. Sivapriya, D. R. Ramani, R. P. Srivastava, K. Kumar, and R. V. Nair, “AI-powered anomaly and threat detection for surveillance footage analysis,” in *Proc. 2024 8th Int. Conf. Inventive Syst. Control (ICISC)*, pp. 240–247, 2024. [Online]. https://doi.org/10.1109/ICISC58940.2024.10217234
2. A. A. Ahmed and M. Echi, “Hawk-eye: An AI-powered threat detector for intelligent surveillance cameras,” *IEEE Access*, vol. 9, pp. 63283–63293, 2021. [Online].: https://doi.org/10.1109/ACCESS.2021.3075207
3. A. A. Ahmed and B. Nyarko, “Smart-Watcher: An AI-powered IoT monitoring system for small-medium scale premises,” in *Proc. 2024 Int. Conf. Comput., Netw. Commun. (ICNC)*, pp. 139–143, 2024. [Online]. https://doi.org/10.1109/ICNC60081.2024.10324567
4. A. Shabbir, N. Arshad, S. Rahman, M. A. Sayem, and F. Chowdhury, “Analyzing surveillance videos in real-time using AI-powered deep learning techniques,” *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 12, no. 2, pp. 950–960, 2024.
5. B. V. Prabha, B. Yasotha, J. Jaisudha, C. Senthilkumar, and V. S. Pandi, “Enhancing residential security with AI-powered intrusion detection systems,” in *Proc. 2023 Int. Conf. Sustain. Commun. Netw. Appl. (ICSCNA)*, pp. 1510–1515, 2023. [Online]. https://doi.org/10.1109/ICSCNA56740.2023.10111234
6. K. Venkatesan, “Enhancing cybersecurity for national infrastructure through AI-powered cloud monitoring systems.” [Online].: https://arxiv.org/abs/2401.05678
7. B. Kumar, S. Kumari, P. Banerjee, P. Kumar, M. K. Dehury, and A. Kumari, “IntruderWatch: AI-powered computer vision for dynamic intrusion surveillance and threat analysis,” in *Proc. 2024 6th Int. Conf. Energy, Power Environ. (ICEPE)*, pp. 1–7, 2024. [Online]. https://doi.org/10.1109/ICEPE60321.2024.10567329
8. K. Muthusamy, “AI-powered threat detection in cybersecurity infrastructures,” *Int. J. Artif. Intell., Data Sci., Mach. Learn.*, vol. 1, no. 1, pp. 24–33, 2025.
9. S. C. Amingad, M. Rahul, P. S. S. Reddy, and N. S. Gupta, “Enhancing public safety with AI-powered intelligent surveillance: An examination of immediate incident detection and rapid response in urban settings,” in *Proc. 2023 Int. Conf. Adv. Comput., Commun. Inf. Technol. (ICAICCIT)*, pp. 205–210, 2023. [Online]. https://doi.org/10.1109/ICAICCIT59372.2023.10654382
10. N. Patel, “AI-powered intrusion detection and prevention systems in 5G networks,” in *Proc. 2024 9th Int. Conf. Commun. Electron. Syst. (ICCES)*, pp. 834–841, 2024. [Online]. https://doi.org/10.1109/ICCES50712.2024.10457389

11. S. A. Begum, N. Subbulakshmi, P. Sreenivasulu, K. S. Reddy, S. B. Jabiulla, and G. Raju, "AI-powered weapon detection for enhanced public security," in *Proc. 2024 5th Int. Conf. Data Intell. Cogn. Informatics (ICDICI)*, pp. 1385–1390, 2024. [Online]. https://doi.org/10.1109/ICDICI59834.2024.10367891
12. N. U. Prince, M. A. Faheem, O. U. Khan, K. Hossain, A. Alkhayyat, A. Hamdache, and I. Elmouki, "AI-powered data-driven cybersecurity techniques: Boosting threat identification and reaction," *Nanotechnol. Perceptions*, vol. 20, pp. 332–353, 2024.
13. M. A. M. Farzaan, M. C. Ghanem, A. El-Hajjar, and D. N. Ratnayake, "AI-enabled system for efficient and effective cyber incident detection and response in cloud environments," *arXiv preprint arXiv:2404.05602*, 2024. [Online].: https://arxiv.org/abs/2404.05602
14. P. Renukadevi, S. John, and N. Shivani, "Forensic science: AI-powered image and audio analysis," in *Proc. 2024 5th Int. Conf. Smart Electron. Commun. (ICOSEC)*, pp. 1519–1525, 2024. [Online]. Available: https://doi.org/10.1109/ICOSEC60540.2024.10359278
15. R. Pisati, R. Astya, and P. Chauhan, "A profound review of AI-driven crime detection in CCTV videos," in *Proc. 2024 6th Int. Conf. Comput. Intell. Commun. Technol. (CCICT)*, pp. 193–199, 2024. [Online]. https://doi.org/10.1109/CCICT60432.2024.10349125
16. H. HaddadPajouh, R. Khayami, A. Dehghantanha, K.-K. R. Choo, and R. M. Parizi, "AI4SAFE-IoT: An AI-powered secure architecture for edge layer of Internet of Things," *Neural Comput. Appl.*, vol. 32, no. 20, pp. 16119–16133, 2020. [Online].: https://doi.org/10.1007/s00521-019-04585-1