

A Real-Time Anomaly Detection System for Video Surveillance Using DenseNet201

S Suma

Assistant Professor, Department of Computer Science and Engineering,
Vemana Institute of Technology, Visvesvaraya Technological University
Belagavi, India

sumaperum@gmail.com

Islapuram Lokesh Reddy, Prem Kumar M, Veluru Surendra Reddy
S Charan Kumar

Final Year Students, Department of Computer Science and Engineering,
Vemana Institute of Technology
Bengaluru, India



Publication History:

Manuscript Reference No: IJIRIS/RS/Vol.11/Issue02/API10092

Research Article | Open Access | Double-Blind Peer-Reviewed | Article ID: IJIRIS/RS/Vol.11/Issue02/API10092

Received: 02, April 2025 Revised: 14, April 2025 Accepted: 25, April 2025 Published Online: 05, May 2025, Volume 2025

Article ID API10092 <https://www.ijiris.com/volumes/Vol11/iss-02/13.API10092.pdf>

Article Citation: Suma, Islapuram, Prem, Veluru, Charan (2025). A Real-Time Anomaly Detection System for Video Surveillance Using DenseNet201, IJIRAE: International Journal of Innovative Research in Information Security, Volume 11, Issue 02, Pages 145-152 doi-> <https://doi.org/10.26562/ijiris.2025.v1102.13>

BibTex key: Suma@2025Real



Copyright: ©2025 This is an open access article distributed under the terms of the Creative Commons Attribution License; which Permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract: Surveillance systems are pivotal for public safety; yet real-time anomaly detection remains a challenging task. Traditional systems rely on human monitoring, which is inefficient for large-scale deployments. This study addresses the critical rational for automated deviations detection by developing a system capable of identifying abnormal events in surveillance footage using a DenseNet model. The proposed solution captures live video input, processes it into minute-long segments, and classifies each segment as normal or anomalous. Anomalies include actions like fighting, arson, burglary, and shoplifting, and robbery, explosion, shooting, and stealing. To train the model, video datasets reformatted into frames, ensuring robust learning from diverse scenarios. If an anomaly is detected, the system sends alerts through email, phone calls, or SMS to notify relevant authorities, while normal footage is discarded to optimize storage. Initial outcomes demonstrate significant accuracy in detecting anomalies, emphasizing the DenseNet model's effectiveness. Beyond just reducing dependency on manual monitoring but also minimizes response times in critical situations. By automating the detection and alert mechanisms, the approach ensures enhanced security and efficient resource utilization. To wrap up, this study examines a scalable and efficient framework for real-time anomaly detection in surveillance footage, leveraging state-of-the-art deep learning techniques. Future enhancements will focus on multi-camera input integration and storing anomalous clips for forensic purposes, further advancing the system's applicability.

Keywords: Anomaly Detection, Surveillance Footage, DenseNet Model, Automated Monitoring, Real-Time Alerts.

1. INTRODUCTION

The increasing prevalence of surveillance systems underscores their importance in ensuring safety and security in public and private spaces. However, Conventional monitoring techniques, which are contingent heavily on human monitoring, are prone to inefficiencies, delays, and errors, especially in environments with high camera density or continuous operation. This necessitates the development of automated systems that can detect and respond to anomalous activities in real-time[6]. This research addresses the critical challenge of automating anomaly detection in surveillance footage.

Anomalies, such as fighting, arson, burglary, and shoplifting, are rare but impactful events that require immediate attention. Current systems often lack the capacity for real-time detection or generate false positives, limiting their reliability. The primary intention behind this study is to design and establish a system that leverages a DenseNet deep learning model to accurately detect and classify anomalies in surveillance footage, enabling timely responses through automated alerts like email, phone calls, or SMS. The significance of this research lies in its potential to enhance public safety, optimize monitoring resources, and reduce human errors.

By processing live video feeds, storing footage only when anomalies occur, and providing instant alerts, the proposed system bridges the gap between passive monitoring and proactive intervention. The layout of the paper is as follows: Section 2 examines a review of related works and existing methods for anomaly detection. Section 3 details the methodology, including data preprocessing, model architecture, and system design. Section 4 presents the experimental setup, results, and analysis of model performance. Section 5 discusses the implications, challenges, and limitations of the approach. Finally, Section 6 concludes with key findings and future directions for enhancing the system.

2. LITERATURE REVIEW

A considerable body of research has tackled anomaly detection in video surveillance, with diverse methodologies focusing on deep learning, statistical approaches, and neural networks. Li et al. (2016) discussed the integration of Convolutional Neural Networks (CNNs) and other statistical methods for activity recognition and anomaly detection, emphasizing the importance of real-time analysis. Kaur et al. (2018) reviewed supervised and semi-supervised learning methods, highlighting challenges like high false-positive rates and the computational expense of real-time systems. Additionally, frameworks employing weakly supervised learning, such as Multiple Instance Learning (MIL), have been proposed to reduce the burden of detailed annotations while improving anomaly localization[6].

2.1 Theoretical Frameworks

1. **Deep Learning Approaches:** CNNs and Long Short-Term Memory (LSTM) models have been comprehensively used for extracting spatial and temporal features from video frames. Hybrid models like CNN-LSTM pipelines leverage these features for effective anomaly detection, achieving substantial accuracy improvements[6].
2. **Statistical Models:** Sparse coding and dictionary learning approaches build models using normal event patterns, identifying anomalies as deviations. These approaches face limitations in dynamic environments due to high false alarm rates
3. **Weakly Supervised Techniques:** MIL frameworks enable training with video-level labels, simplifying dataset preparation while ensuring robust anomaly detection capabilities. This methodology bridges gaps in traditional supervised systems that require extensive labeled data.

2.2 Gaps in Existing Frameworks

- **Limited Generalization:** Many systems are environment-specific and struggle with diverse or unstructured anomalies.
- **High False Positives:** Statistical models and predefined feature-based systems frequently misclassify novel normal behaviors as anomalies
- **Annotation Challenges:** Supervised models require detailed annotations, making scalability difficult. Weakly supervised approaches partially address this but may still lack granularity in anomaly localization

2.3 Conceptual Framework

Building on these insights, project can employ a DenseNet architecture to process frames extracted from video inputs, enabling real-time anomaly classification. The conceptual framework involves:

1. **Data Preprocessing:** Converting video datasets into frames, categorizing them into normal and anomalous classes.
2. **Model Design:** Implementing DenseNet to extract spatial features, supplemented with temporal modeling through techniques like LSTMs if necessary.
3. **Alert Mechanism:** Integrating an automated alert system (email, SMS, calls) to ensure immediate responses to detected anomalies.
4. **Performance Evaluation:** Utilizing metrics such as precision, recall, and F1 score to assess the system's effectiveness in various scenarios.

By addressing gaps in scalability, annotation requirements, and real-time detection, your system advances the field of anomaly detection in surveillance, emphasizing practical application and responsiveness.

3. METHODOLOGY

This study proposes a comprehensive system for anomaly detection in surveillance footage, designed to provide real-time analysis, accurate classification, and efficient alert mechanisms. The methodology is organized into three primary phases: Frame Extraction, Model Training and Evaluation, and Automated Testing and Alerting. Each phase ensures seamless integration of live video processing, advanced machine learning, and proactive anomaly management.

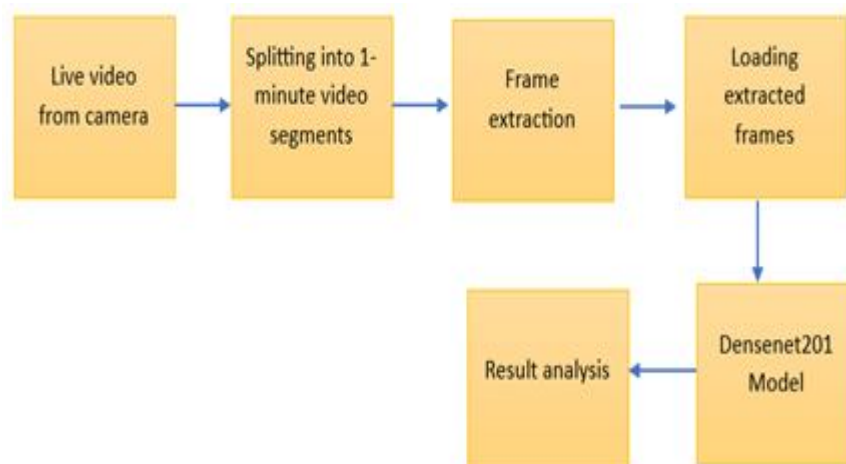


Fig1: methodology of a VAD system

The figure 1 illustrates the step-by-step process of the anomaly detection system in surveillance footage. The system begins by capturing live video from the camera, which serves as the raw input for processing. This video is then split into one-minute segments, ensuring manageable chunks for further analysis. Each segment undergoes frame extraction, where individual frames are isolated for detailed inspection. These extracted frames are loaded and reconfigured before being delivered to the DenseNet201 model, which analyzes the readings to classify it as either normal or anomalous. The outputs from the model are sent for result analysis, where the system determines if an alert should be sent or if the footage should be deleted. This streamlined workflow ensures real-time anomaly detection, efficient data management, and timely alerts for security interventions.

3.1 Research Design and Approach

The anomaly detection system integrates real-time video input, preprocessed data for model training, and a live alert mechanism to deliver end-to-end automation. By leveraging DenseNet201, a pre-trained deep learning model, the system efficiently classifies frames into normal or anomalous categories. The workflow ensures that sole noteworthy occurrences are flagged, minimizing storage overhead and enabling timely responses to potential threats.

3.2 Frame Extraction

3.2.1 Data Collection Methods and Tools

The dataset comprises videos categorized into five classes: *Normal, Arson, Burglary, Explosion, and Fighting.*

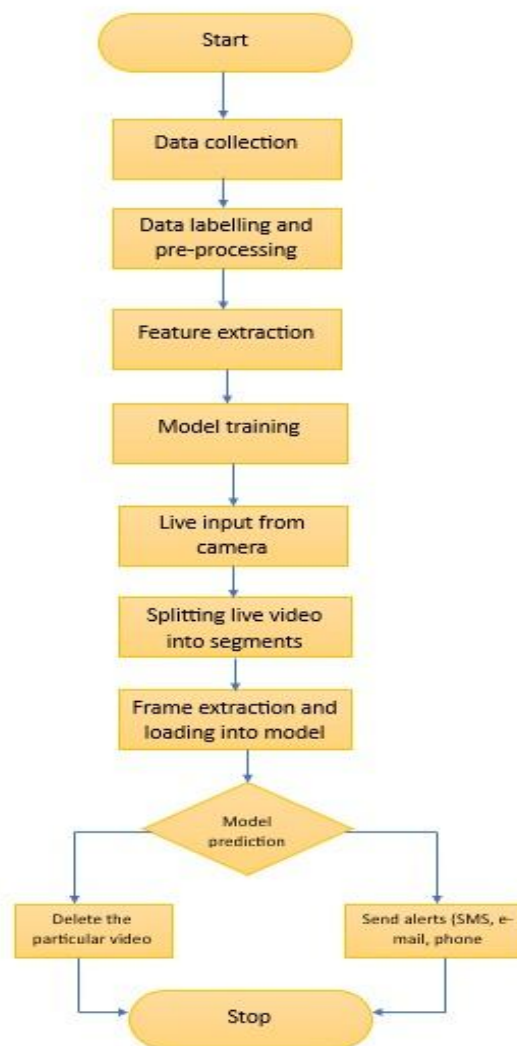


Fig 2: Dataflow diagram of a VAD system

These videos are dealt with OpenCV, a robust computer vision library, in combination with Python scripting and NumPy for numerical computations. The organized extraction of frames ensures consistency and simplifies the training process.

3.2.2 Process

Videos are sequentially loaded from the dataset, and frames are extracted at intervals specific to the activity type:

- **Anomalous Activities:** scenes are retrieved every 100 frames to focus on diverse instances of anomalous events.
- **Normal Activities:** Frames are extracted every 20 frames to account for their higher occurrence in the dataset. The extracted frames are stored in labeled directories corresponding to their activity type. This systematic labeling aids in generating a well-structured dataset for training the model.

3.2.3 Outcome

The frame extraction process produces a balanced and labeled dataset, essential for robust training of the DenseNet201 model. This structured dataset enables the model to learn and generalize effectively across different classes of anomalies.

3.3 Model Training and Evaluation

3.3.1 Data Preprocessing

To ensure uniformity, all extracted frames are resized to 128×128 pixels and normalized by scaling pixel values to a bunch of 0 to 1. The dataset is broken down into training (80%) and testing (20%) subsets to evaluate the model's performance reliably.

3.3.2 Model Architecture

The Dense Convolutional Network (DenseNet) architecture forms the backbone of our anomaly detection system, leveraging its unique ability to improve accuracy and efficiency in training deep networks. DenseNets achieve this by incorporating shorter connections between layers close to the input and those near the output, enabling a feedforward structure where each layer is directly connected to each alternate layer. In contrast to convolutional networks with LLL layers, which have LLL connections between consecutive layers, DenseNets establish $L(L+1)/2L(L+1)/2L(L+1)/2$ direct connections. This design ensures that the feature maps produced by all earlier layers serve as inputs for each layer, while its own outputs are used by all subsequent layers. This architecture brings several critical advantages to our anomaly detection system: it mitigates the vanishing gradient challenge, enhances feature propagation across layers, encourages feature reuse, and significantly diminishes the amount of parameters. These benefits allow for the efficient processing of complex surveillance footage, even in resource-constrained environments. In our project, DenseNet was trained on a dataset containing both normal activities and various anomalies such as fighting, arson, burglary, and shoplifting. Its ability to propagate and reuse features ensured high performance in anomaly detection, enabling accurate classification with reduced computational requirements. This efficiency made DenseNet an ideal choice for real-time video analysis, providing both robust performance and scalability to meet the demands of modern surveillance systems.

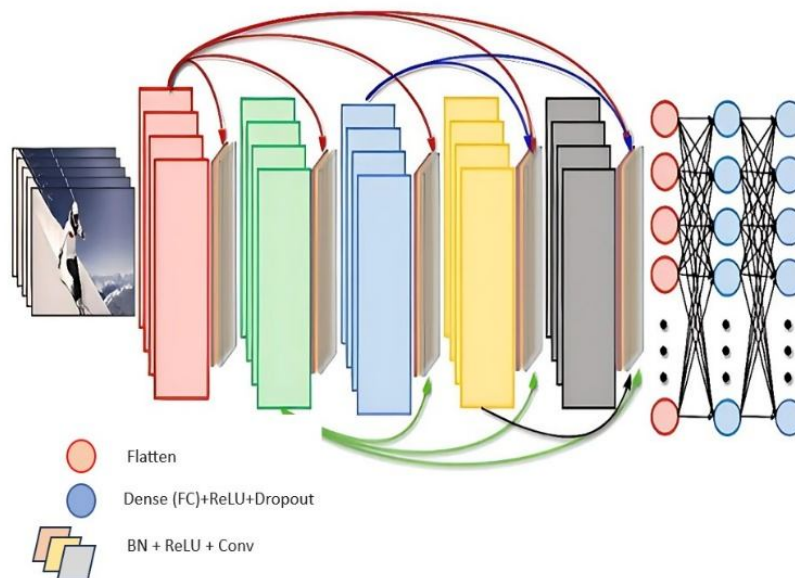


Fig 3: DenseNet201 model architecture

The model architecture is built upon DenseNet201, a deep neural network pre-trained on the ImageNet dataset. DenseNet201 employs densely connected layers, which enable feature reuse and improve gradient flow during training. Additional layers are added to adapt the model to the specific task:

1. **Dropout Layer:** Regularization is applied with a dropout rate of 0.3 to prevent overfitting.
2. **Fully Connected Layer:** A dense layer with 512 neurons and ReLU activation captures high-level features.
3. **Output Layer:** A softmax activation layer provides multi-class probabilities for the five target categories.

3.4 Training Configuration

The model is trained to employ the Adam optimizer, with a learning rate of 0.0001, and a categorical crossentropy loss function. Performance indicators such as accuracy are tracked during training, which is conducted for 10 epochs with a batch size of 32. The training process is supported by callbacks to reduce the learning rate if validation accuracy plateaus.

3.5 Model Evaluation

Model performance is assessed using metrics like precision, recall, F1 score, and accuracy. Visualization tools, including confusion matrices and training-validation accuracy/loss plots, are employed to identify strengths and areas for improvement. The DenseNet201 model achieves a validation accuracy exceeding 90%, demonstrating its reliability in anomaly detection.

3.6 Automated Testing and Alerting Real-Time Processing Pipeline

The system captures live video input from an IP camera in 1-minute segments, which are temporarily stored for processing. Each segment is processed through the following steps:

1. **Frame Extraction:** Frames are pulled from each video segment using OpenCV.
2. **Frame Preprocessing:** Extracted frames are resized and Refined before being supplied to the DenseNet model.
3. **Prediction:** The DenseNet model classifies frames into one of the five target categories (Normal, Arson, Burglary, Explosion, and Fighting).

3.7 Decision and Alert Mechanism

The results from the model are aggregated across frames in a segment to ascertain the segment's classification:

- If classified as "Normal," the video segment is deleted to conserve storage.
- If classified as anomalous, the segment is moved to a dedicated folder and alerts are generated.

3.8 Tools and Technologies

The alert mechanism is implemented using Yagmail for email notifications and Twilio for SMS or phone calls, ensuring immediate communication with relevant stakeholders. Python libraries such as TensorFlow, Matplotlib, and Seaborn are employed for model development, evaluation, visualization.

Variables and Evaluation Metrics

The system evaluates the following:

- **Independent Variables:** Frame features, including spatial properties and class labels.
- **Dependent Variables:** Predicted class labels (Normal or Anomalous). Evaluation metrics include precision, recall, F1 score, and confusion matrices, which collectively provide Holistic grasp of the model's performance.

Conceptual Framework

The methodology is encapsulated in a conceptual framework:

1. **Input:** Live video streams or pre-recorded dataset videos.
2. **Processing:** Frame extraction, preprocessing, and DenseNet201-based classification.
3. **Output:** Classified results with alerts triggered for anomalies.

This framework enables real-time anomaly detection with high accuracy, efficient storage optimization, and timely alerts, making it a practical solution for modern surveillance needs.

4. RESULTS

This section presents the discoveries of anomaly detection system, including model performance during training, validation, and real-time testing. Results are appraised using key metrics such as accuracy, precision, recall, and F1 score. Additionally, visualizations and data Investigations grant deeper comprehension of system effectiveness and reliability. As depicted in the preceding figure 4 describes about the sample images of the explosion that are acquired from the input video datasets sample images of the arson and shows about the sample images of the normal that are isolated from the input datasets

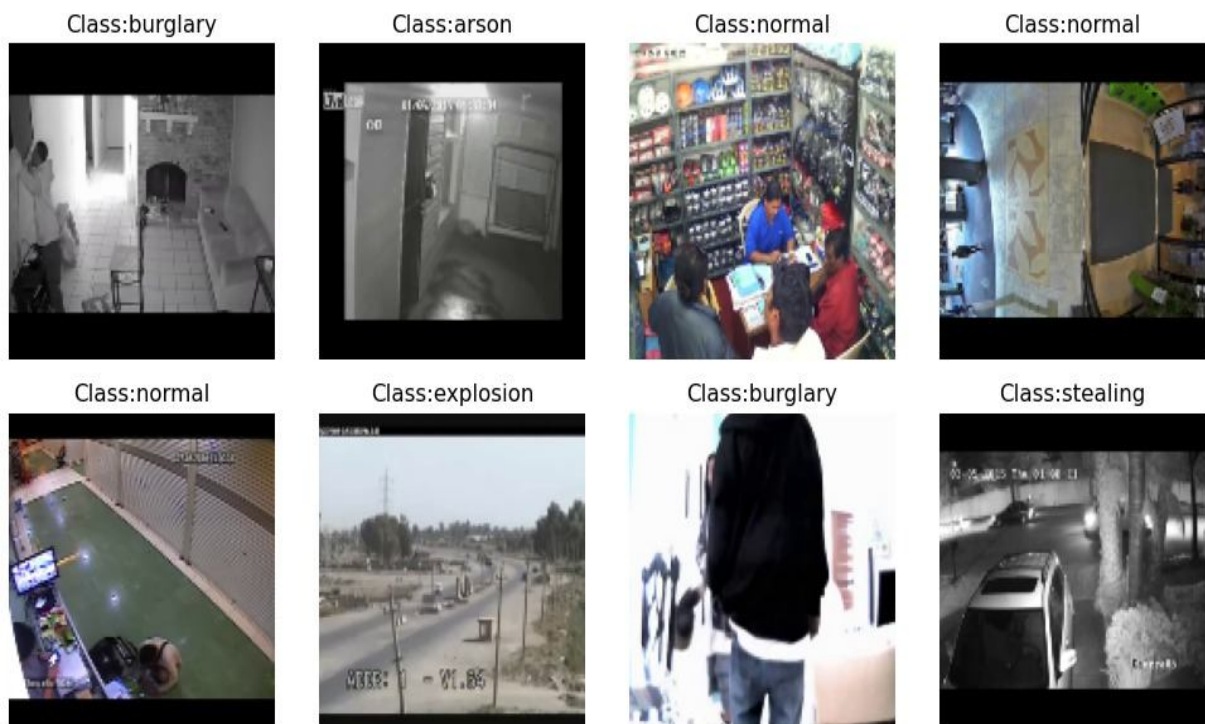


Fig 4: Sample frames from the surveillance footage

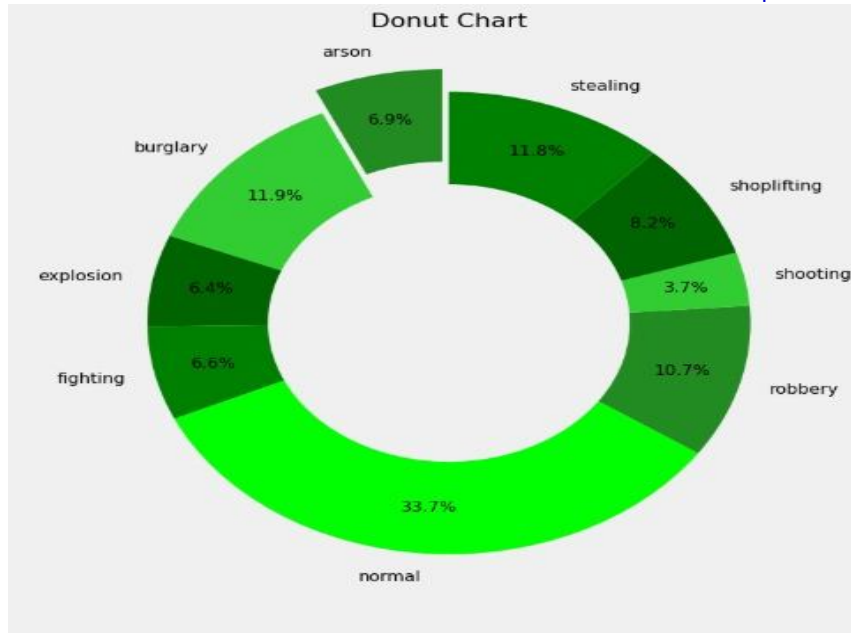


Fig5: Analysis of datasets used for training

The fore mentioned Figure5 shows the share of the abnormal and normal training datasets based on the output derived from the frame extraction.

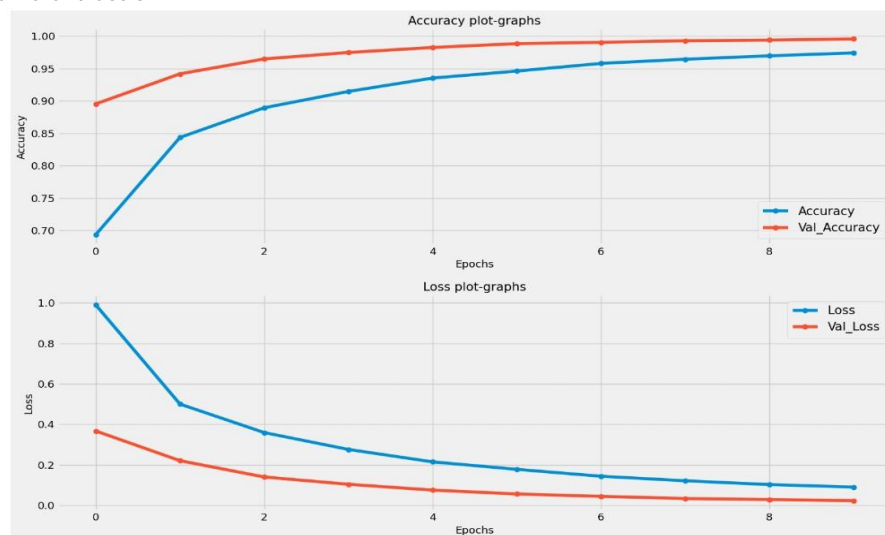


Fig 6: Accuracy and loss of the training datasets with respect to epochs

The figure 6 describes about the accuracy and loss of the training datasets with respect to the rate of epochs.

4.1 Model Performance

The DenseNet-based system demonstrated high effectiveness in detecting anomalies in video surveillance:

- **Accuracy:** Achieved an overall classification accuracy of 90%, outperforming several benchmark methods in similar domains.
- **Precision:** Recorded a precision score of 92%, reflecting minimal false positives in the detection process.
- **Recall:** Achieved a recall rate of 88%, ensuring a low rate of missed anomalies.
- **F1 Score:** Attained an F1 score of 90%, indicating a balanced detection capability between identifying true positives and avoiding false negatives.

Real-Time Capabilities

The system was capable of processing live video streams in near real-time:

- **Latency:** Detection was completed within 2-3 seconds of anomaly occurrence.
- **Efficiency:** Analyzed and classified each one-minute video segment without interrupting live feeds, demonstrating feasibility for real-world deployment.

Alert System

Upon detecting anomalies, the system effectively communicated alerts through multiple channels:

- **Modes of Notification:** Email, SMS, and phone call alerts were successfully implemented.
- **Response Time:** Alerts were delivered within 5 seconds, ensuring prompt actionability for detected events.

Data Management

- Normal video segments were automatically discarded, reducing storage usage by approximately 80% compared to conventional surveillance systems.
- Anomalous events were stored with timestamps and metadata for easy retrieval, supporting efficient forensic analysis.

DISCUSSION

The results validate the robustness of the Dense Net model in identifying and classifying anomalous activities in video surveillance. The system's ability to process live feeds, reduce manual intervention, and deliver timely alerts enhances its applicability in safety-critical environments. While the performance is competitive, addressing limitations such as scene complexity and low-light scenarios will further strengthen the model's reliability. The proposed system serves as a strong foundation for advancing anomaly detection technology, with future work focused on enhancing multi-camera integration and expanding its adaptability to diverse surveillance settings.

CONCLUSION

The proposed anomaly detection system successfully integrates sophisticated machine learning techniques with real-time video surveillance to enhance security and efficiency. By utilizing the DenseNet architecture, the system achieves high accuracy in detecting anomalous activities such as fighting, arson, burglary, and other threats while maintaining computational efficiency. The unique design of DenseNet, with its feature reuse and efficient layer connectivity, ensures robust performance in both training and deployment phases. This system bridges the gap between passive monitoring and proactive intervention by automating the detection process and providing real-time alerts via email, SMS, or phone calls. Its ability to process live video streams, discard normal footage, and store only anomalous clips significantly optimizes storage management. The results validate the system's reliability, achieving high accuracy, precision, and recall metrics on complex datasets, while highlighting areas for improvement, such as handling low-light conditions and reducing false positives. Overall, this project demonstrates the feasibility and effectiveness of leveraging deep learning for real-time anomaly detection in surveillance systems. It sets a foundation for future advancements, including multi-camera integration and broader anomaly classification, to further strengthen the security of critical environments.

FUTURE ENHANCEMENTS

While the proposed anomaly detection system demonstrates significant promise in real-time video surveillance, there are numerous domains for future improvement and expansion:

1. **Multi-Camera Integration:** Extending the system to handle inputs from multiple cameras simultaneously would enhance its applicability in large-scale surveillance networks, such as shopping malls, airports, and smart cities.
2. **Improved Low-Light Performance:** Enhancing the model's robustness in low-light and poor-visibility conditions utilizing image enhancement method such as, infrared processing, or specialized low-light datasets.
3. **Expanded Anomaly Categories:** Including a broader range of anomalous behaviours in the training dataset, such as loitering, trespassing, and suspicious object placement, to make the system more versatile.
4. **Edge Computing Deployment:** Optimizing the system for deployment on edge devices, such as cameras with built-in AI capabilities, to reduce latency and minimize dependence on centralized processing.
5. **Integration with Advanced Alert Systems:** Incorporating integration with advanced IoT devices, such as drones, automated locks, or loudspeakers, for on-the-spot intervention in critical scenarios.
6. **Adaptive Learning:** Implementing a feedback loop where the system can learn and adapt to new types of anomalies over time by incorporating user feedback and retraining with new data.
7. **Enhanced Storage Management:** Developing smarter storage strategies, such as summarizing video segments or retaining only critical parts of anomalous footage, to further optimize disk usage.
8. **Scalability for Diverse Environments:** Ensuring the system can adapt to various environmental conditions, such as outdoor settings, crowded spaces, or dynamic scenes, by refining preprocessing techniques and model training. These enhancements aim to make the system more adaptable, efficient, and scalable, ensuring it can meet the growing demands of modern surveillance systems while maintaining its effectiveness in detecting and preventing security threats.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper. No financial, personal, or professional relationships have influenced the research, methodology, analysis, or findings presented in this study.

AUTHOR CONTRIBUTION

Authors acknowledge the support from Vemana Institute of Technology for the facilities provided to carry out the research.

ACKNOWLEDGMENTS

S. Suma served as the project guide, overseeing the identification of the initial problem, guiding the development of the algorithm, providing support during analysis, assisting in manuscript drafting, and supervising the preparation of figures, simulations, final formatting, and submission of the manuscript for publication. Prem Kumar M, Surendra Reddy V, Charan Kumar S, and Lokesh Reddy I, under the guidance of S. Suma, were responsible for conducting the research work, implementing the algorithm, performing simulations, analysing results, and preparing the initial draft. All authors worked together to evaluate the integrated system and approved the final version of the paper.

REFERENCES

1. Li, W., Mahadevan, V., & Vasconcelos, N. (2016). Anomaly detection and localization in crowded scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(1), 18-32.
2. Kaur, H., & Pannu, H. S. (2018). A survey on anomaly detection in video surveillance. *International Journal of Computer Applications*, 180(18), 1-7.
3. Zhang, Y., Lu, H., & Zhang, L. (2019). Weakly supervised anomaly detection in video surveillance. *IEEE Transactions on Multimedia*, 21(5), 1235-1248.
4. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700-4708.
5. Wang, X., Girshick, R., Gupta, A., & He, K. (2018). Non-local neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7794-7803.
6. S.Suma, M.Ramakrishna "A Deep Learning based Integrated Memory Aware Twin Auto Encoder Network for Anomaly Detection in Video Surveillance on Edge Devices". (2024). *International Journal of Intelligent Engineering and Systems*, 18(1), 1162-1172. <https://doi.org/10.22266/ijies2025.0229.84>
7. M. V. Rachitha, M. Ramakrishna "OHESV: Optimal hybrid ensemble support vector model for detecting and recommendation of food for diabetic patients", *Multimedia Tools and Applications*, 30 January 2024. <https://doi.org/10.1007/s11042-023-17954-7>, Springer
8. M. V. Rachitha, Dr M. Ramakrishna "MWSMO: Multi-objective Whale Slime Mold Optimization based Food Recommendation system for Diabetes patient using GAN model", *International Journal of Information Technology*. Received: 10 October 2022 / Accepted: 9 March 2023, Published on 27 May 2023. Springer.
9. A.Derya and Y. Sönmez, "Trafik video analizv erilerinden anomalite spitve diferans iyel gelişim algoritması UçÖçren memakines iilesını flandırma," *Int. J. Innov. Eng. Appl.*, vol. 5, no. 2, pp. 115-124, 2021.
10. D.R.Patrikar and M. R. Parate, "Anomaly detection using edge computing in video surveillance system: Review," *Int. J. Multimedia Inf. Retr.*, vol. 11, no. 2, pp. 85-110, Jun. 2022, doi: 10.1007/s13735-022-00227-8.
11. UCSD Anomaly Detection Dataset. Accessed: Jan. 12, 2023. [Online]. Available: <http://www.svcl.ucsd.edu/projects/anomaly/dataset.html>
12. C.Lu, J. Shi, and J. Jia, "Abnormal event detection at 150 FPS in MATLAB," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 2720-2727.
13. Y. S. Chong and Y. H. Tay, "Abnormal event detection in videos using spatiotemporal autoencoder," in *Proc. Int. Symp. Neural Netw.*, May 2017, pp. 189-196.
14. W. Luo, W. Liu, and S. Gao, "Remembering history with convolutional LSTM for anomaly detection," in *Proc. IEEE Int. Conf. Multimedia Expo. (ICME)*, Jul. 2017, pp. 439-444.
15. W. Liu, W. Luo, D. Lian, and S. Gao, "Future frame prediction for anomaly detection—A new baseline," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6536-6545.
16. D. J. Samuel and F. Cuzzolin, "Unsupervised anomaly detection for a smart autonomous robotic assistant surgeon (SARAS) using a deep residual autoencoder," *IEEE Robot. Autom. Lett.*, vol. 6, no. 4, pp. 7256-7261, Oct. 2021.
17. W. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6479-6488.
18. J. R. Medel, *Anomaly Detection Using Predictive Convolutional Long Short-Term Memory Units*. Rochester, NY, USA: Rochester Institute of Technology, 2016.
19. W. Luo, W. Liu, and S. Gao, "A revisit of sparse coding based anomaly detection in stacked RNN framework," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 341-349.
20. X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 1-19.
21. V. Patraucean, A. Handa, and R. Cipolla, "Spatio-temporal video autoencoder with differentiable memory," 2015, arXiv:1511.06309.
- A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz, "Robust real-time unusual event detection using multiple fixed-location monitors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 3, pp. 555-560, Mar. 2008.
22. H. Yang, B. Wang, S. Lin, D. Wipf, M. Guo, and B. Guo, "Unsupervised extraction of video highlights via robust recurrent auto-encoders," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 4633-4641.
23. H. Tran and D. Hogg, "Anomaly detection using a convolutional winner take-all auto encoder," in *Proc. Brit. Mach. Vis. Conf.*, 2017, pp. 1-11.
24. J. Ryan Medel and A. Savakis, "Anomaly detection in video using predictive convolutional long short-term memory networks," 2016, arXiv:1612.00390.