Journal of Theoretical and Applied Information Technology
15th December 2025. Vol.103. No.23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

GAME-THEORETIC SELF-SUPERVISED STRATEGIES FOR INTELLIGENT ANOMALY DETECTION IN DYNAMIC SURVEILLANCE ENVIRONMENTS

Dr. VUDA SREENIVASA RAO[1], Prof. CHIN-SHIUH SHIEH[2], Prof. SIVA SHANKAR S[3], Prof. PRASUN CHAKRABARTI[4]

"Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, INDIA.[1a]
Research Institute of IoT Cybersecurity, Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Taiwan.[1b]
Professor, Research Institute of IoT Cybersecurity, Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Taiwan.[2]
Professor & Head IPR, Department of CSE, KG Reddy College of Engineering and Technology, India.[3]
Director, Directorate of Research and Publications and Dean International Affairs and Senior Professor, Department of Computer Science and Engineering, Sir Padampat Singhania University, India.[4]"

vsreenivasarao@kluniversity.in[1], csshieh@nkust.edu.tw[2], drsivashankars@gmail.com[3], drprasun.cse@gmail.com[4]

ABSTRACT

Detecting anomalies in large-scale video surveillance is challenging due to diverse scene variations and the limited availability of labeled anomalous data. Whereas existing models employed supervised deep networks or rule-based heuristics, this proposed GTSSL framework presents a game-theoretic self-supervised paradigm that enhances system adaptability without requiring large annotated datasets. The method redesigns anomaly detection in IT-driven surveillance by combining decision theory with autonomous learning, a union that is seldom attempted in existing work. The unique contribution of this research is the development of the Game-Theoretic Self-Supervised Learning (GTSSL) framework, which bridges the gap between self-supervised feature learning and adaptive decision-making in anomaly detection. In contrast to traditional deep learning approaches that are based on labeled data or fixed optimization, this framework combines game-theoretic strategy formation, GAN-based augmentation, and CNN-LSTM hybrid modeling, thus presenting a novel paradigm in intelligent surveillance analytics. This methodology puts forward the overall corpus of knowledge in IT by codifying anomaly detection into a strategic optimization problem, enhancing flexibility, robustness, and online learning in complex monitoring situations. The added knowledge, therefore it forms an enlarged step over previous incremental progress in that it brings a theoretically sound, practically scalable, and computationally feasible model translatable to various fields of security and surveillance. The framework is implemented using Python with TensorFlow and OpenCV libraries and evaluated on the Kaggle CCTV Activity Identification Collection and approximately 200 clips per class. Experimental analysis demonstrates superior performance of GTSSL, achieving 98.3% accuracy, 97.5% precision, 97.4% recall, and an F1-score of 97.0%, surpassing CNN (72%), SVM (72.1%), and CNN-RNN (84%) models by a margin of 6–9%. Results confirm that GTSSL enhances resilience against distributional shifts and adversarial patterns while maintaining computational efficiency for real-time deployment. These findings validate the efficacy of combining self-supervised representation learning, adversarial augmentation, and game-theoretic decision-making in anomaly detection systems for dynamic surveillance environments.

Keywords—*Anomaly Detection, Game-Theoretic Learning, Generative Adversarial Networks, Convolutional Neural Networks, Long Short-Term Memory, Video Surveillance*

1. INTRODUCTION

Video surveillance systems serve an important function in safeguarding security in a variety of settings, include public areas, vital facilities, and businesses [1] [2]. Current approaches to recognize anomalies in surveillance footage frequently depend on human surveillance or systems with rules, and are restricted in their capacity to respond to complex and changing situations [3] [4] [5]. Game-theoretic self-teaching systems are capable of successfully figure out the difference amongst conventional and odd patterns in video files with transforming recognizing anomalies into an interaction with an

indicator system as well as a player looking to provide hard cases [6].

In real-world contexts, game-theoretic self-supervised instruction techniques may exhibit excellent precision in detection and rate of generalization by continuously adjusting to the changing behavior of oddities [7], [8], [9]. By leveraging the capabilities of machine learning inside a game-theoretic system, systems for recognizing anomalies can attain advanced capabilities with remaining computationally efficient [10],[11],[12]. To overcome these limitations, the study proposes a GTSSL framework for intelligent anomaly detection in video surveillance [13]. The framework combines the use of GAN for data augmentation, CNN for spatial features construction, and LSTM networks for modeling temporal sequences.

1.1 Research Questions

1. How can a game-theoretic self-supervised learning framework effectively detect anomalies in surveillance videos under data scarcity and dynamic conditions?
2. To what extent does integrating GAN-based data augmentation improve the adaptability and robustness of anomaly detection models?
3. Can game-theoretic optimization dynamically adjust detection thresholds to minimize false positives while maintaining real-time performance?

1.2 Research Motivation

With the rapid growth and increasing complexity of video data, developing effective and adaptive ways to detect anomalies is increasingly important. Traditional methods can't easily adapt to new changes and generally have to use extensive labeled data that is hard to find for rare cases. An additional motivation comes from wanting to make security systems in these areas spot potential threats early on. This study combines game theory, GAN, CNN and LSTM to make a smart and reliable system for detecting anomalies in upcoming surveillance applications.

1.3 Research Significance

This research signifies the IT knowledge base by creating a scalable and responsive method that bridges self-supervised learning and strategic optimization. The framework stretches best-of-breed practices in intelligent surveillance, demonstrating how real-time systems can automatically improve decision thresholds previously in the domain of theoretical game models.

1.4 Key Contributions

1. Developed a game-theoretic approach to improve adaptability and resilience in dynamic surveillance scenarios.
2. Utilized GANs to overcome data paucity and enhance diversity of datasets for anomaly detection.
3. Deep Spatial Feature Extraction for deep spatial feature extraction from video frames to enable precise anomaly detection.
4. Proposed an adaptive game-theoretic optimization mechanism to dynamically optimize anomaly detection strategies based on changing conditions.

The research is structured as follows: Section 2 has significant material designed to make readers understand the work proposed based on current methods, while Section 3 extends further in the problem statement. Section 4 provides Anomaly Detection in Video Surveillance's methodology. Section 5 has the table and visualizations of results along with measures. Lastly, in Section 6, its conclusion and future projects are discussed.

2. RELATED WORKS

Menatalla et al., [14] have investigated difficulties as well as advances in IoT and low-cost sensing installations for massive surveillance programs, emphasizing the importance of accurate sensors as well as successful systems for detecting anomalies. This study indicates how incorporating clever reasoning methods, especially game-theoretical gets closer, may increase the accuracy of these frameworks. But current multimodal LSTM-based architectures for IoT anomaly detection are not effective in addressing the issues of dynamic threshold tuning, real-time adaptability, and large-scale robustness in heterogeneous surveillance environments.

Hazra et al., [15] gave a comprehensive assessment of the current convergence of the game theory along with deep learning, with a focus on various uses that have developed in the past few years. The construction and optimization of these frameworks frequently integrate game-theoretic rules, and there are numerous issues related to classification phrased as Stackelberg matches Nonetheless, even with the proven advantages of unifying deep learning with game-theoretic concepts, current research mostly involves theoretical models or standalone implementations, and there is a dearth of holistic frameworks applying these techniques to real-time, adaptive, and resilient

anomaly detection in high-volume video surveillance.

Monireh et al.,[16] presented a detailed analysis of the game-theoretic underpinnings of GAN that have drawn a lot of interest given their capacity to create material that is quite similar to real-world information. GANs are based on a game involving two neural networks, such as which include a generator and a tool for discrimination. Yet, while these game-theoretic GAN architecture improvements have been made, most work to date has targeted generation quality and theoretical models, with limited application-based frameworks using these techniques for real-time adaptive anomaly detection in dynamic, large-scale surveillance environments.

Swathy et al.,[17] investigated the incorporation of game theory with artificial intelligence systems, with a special emphasis on solving the issue of zero-day hostile examples, that represent serious dangers to system integrity by exposing classification weaknesses. The study investigates the incorporation of game theory with artificial intelligence systems, with a special emphasis on solving the issue of zero-day hostile examples. But whereas advances in using game theory to counteract zero-day adversary attacks continue to be made, existing research mostly focuses on simulative or low-scale environments, with a lack of frameworks that offer real-time, adaptive, and strong anomaly detection for large-scale and dynamic video surveillance systems.

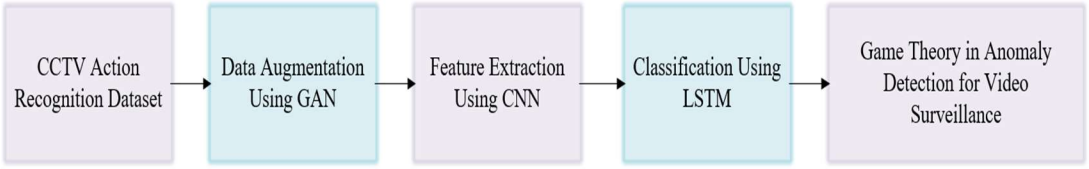
Previous work in deep surveillance learning has dealt with anomaly detection by applying CNNs, LSTMs, and GANs separately, but without any mechanism of dynamic adjustment of strategies. By combining these with game-theoretic reasoning, the current study closes the gap between adaptive decision-making and static feature learning.

3. PROBLEM STATEMENT

Video monitoring systems are at the forefront of utmost challenges in identifying anomalies for large-scale, dynamic, and unlabeled environments [14]. Classic models like CNNs, SVMs, and CNN-RNN hybrids usually lack the capability to adapt to changing behaviors, biased data, and evasive abnormal patterns, resulting in high false positives and lower reliability [15]. The GTSSL framework presented meets these requirements through the integration of game-theoretic optimization and deep learning, coupling GAN-based augmentation, CNN-LSTM feature extraction, and strategic adaptability to vastly improve accuracy, robustness, and real-time performance for contemporary video surveillance systems.

4. PROPOSED GAME-THEORETIC SELF-SUPERVISED LEARNING METHODOLOGY

The proposed study GTSSL for detecting Anomalies in CCTV, is formed of various important parts. The data used in this project consists of videos of each type of activity, both normal and abnormal. GANs are added to the process to make the model detect abnormalities using fabricated, but highly realistic, abnormal data. Video clips are scanned by a CNN to locate spatial features and information about the order of past activities is analyzed by LSTM to determine whether the activities are standard or unusual. Figure. 1 shows the proposed GTSSL method.



"Figure.1 Workflow of the Proposed GTSSL"

4.1 Data Acquisition

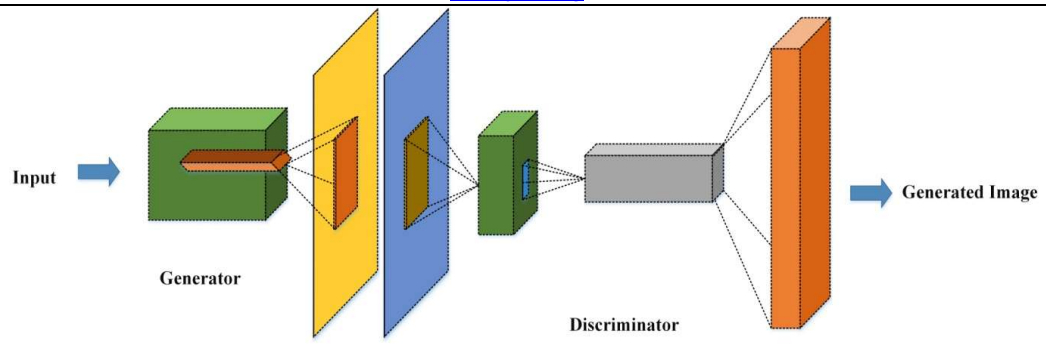
CCTV Action Recognition Dataset [18] collected from Kaggle for anomaly identification comprises brief video clips from CCTV databases, YouTube, and Google, which make up the CCTV Activity Identification Collection. Table 1 shows Kaggle CCTV Activity Identification Collection.

4.2 Data Augmentation using GAN

Applying GANs to data boosts the detection of strange activities in video footage. GANs produce accurate faked videos, allowing for more useful artificial data where there has been a lack of it or imbalanced classes in a surveillance system dataset. Here, GANs are developed to show a variety of situations, both common and abnormal such as trespass, theft or assault. Thanks to these artificial videos, models for anomaly detection can learn more efficiently and become more accurate. When more examples of abnormal behavior are included, detection systems using this approach work much more effectively and are more reliable. Figure.2 shows the Architecture of GAN .

Journal of Theoretical and Applied Information Technology
15th December 2025. Vol.103. No.23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



"Figure.2: Architecture of GAN"

4.3 Game-Theoretic Analysis with Feature Extraction and Classification using CNN-LSTM

A CNN extracted feature is an important step in game-theoretic autonomous learning for anomaly detection in monitoring systems. The approach begins with a beginning video image that undergoes multiple convolutional layers. The Mathematical expression is given in eqn. (1).

$$F_{m,n} = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} I_{m+i,n+j} \quad (1)$$

"Where $I_{m+i,n+j}$ represents the pixel value at position (m+i, n+j) in the input frame, and $F_{i,j}$ is the filter value at position." The convolution process, coupled with a non-linear activation function such as RelU, enables the CNN to learn hierarchical features from low-level edges to high-level object constituents. Figure 3 illustrates the Architecture of CNN-LSTM.

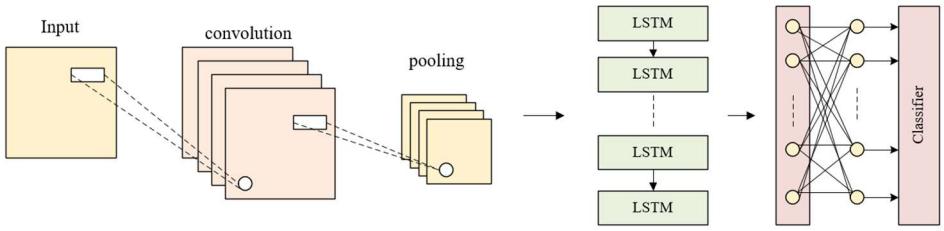


Figure.3: Architecture of CNN-LSTM

Anomaly detection in game-theoretic autonomous learning depends on attributes identified with CNNs. After feature maps are made, pooling takes place to shrink the data so it does not respond as strongly to small differences. Using game theory, the framework views the problem of normal events versus anomalies as a strategic game to find the best way to detect anomalies. Self-supervised learning allows the system to recognize regular versus irregular activities using only a small amount of specifically labeled data. Multimedia content is accurately represented and anomalies are easily identified with the help of the features $F_{m,n}$. The use of CNNs combined with game-theoretic techniques greatly increases how well the framework finds suspicious events in security camera videos.

CNNs separate the input data in order to find important types of patterns. At the start of the structure is the Input Layer which stores unprocessed data like an image and each neuron represents a pixel in that image. Next the Convolutional Layer which filters the input with kernels or filters, all adjusted by the model. Edge detection, texture recognition and pattern spotting are achieved by sliding different filters over the image and looking at the dot products created by each motion, generating feature maps that indicate the presence of such details. Additionally, the Activation Layer, mostly using the ReLU function, ensures that the network is not linear. When using ReLU, CNNs can simulate difficult patterns in the data because it replaces all negative portions with zero and keeps everything else the same. Following activation, the Pooling Layer decreases the size of the feature maps. Max Pooling and Average Pooling are techniques that review parts of the feature map and determine its main or typical value. Pooling helps reduce how much the model works, as well as helps prevent it from "memorizing" data and being biased toward it. Given the input image X and a filter F, the convolution operation is defined in eqn. (2).

"$(X * F)(a,b) = \sum_{j=0}^{m-1} \sum_{k=0}^{n-1} I(a+j, b+k)$ (2)"

Where (a, b) are the spatial coordinates, m and n being the dimensions of the filter.

Fully Connected Layer (Dense Layer): After numerous convolution and pooling layers, feature images are reshaped into vectors and passed to fully connected layers. Each neuron within a fully connected layer is connected with all the neurons in the previous layer. The layers make use of high-level information in order to make predictions.

Output Layer: The final part of the CNN produces output. The number of neurons in the layer depends on the task being addressed. As an example, in a classification task of n classes, the output layer will consist of n neurons, which are often preceded by a function of SoftMax activation to provide class probabilities.

In order to introduce nonlinearity, convolutional layers of data are stacked after each other, with each having a rectified linear unit (ReLU) activation function. These layers tap data from the image under process at a range of scales and complexities.

"The standard equation for the output dimension of the convolutional layer is written in eqn. (3).

$$Y = \frac{(X-K+2P)}{S} + 1 \quad (3)$$

Where Y was a output, X was an input, K is filtering size, P was the padding size, S is the stride.

Classification in game-theoretic autonomous learning for detecting anomalies in surveillance footage is supported by LSTM neural networks. LSTMs find great use in analyzing sequences of images or videos. Because the recordings are played in a perfectly ordered sequence, being able to follow the progression of events is essential for spotting unusual activity. Because of their unique gates, LSTM networks are able to store vital details over a long time and forget unnecessary information."

"The LSTM network outputs a sequence of hidden states h_t, which are used to predict whether the current frame or sequence of frames represents an anomaly. The hidden state h_t at each time step t is updated based on the previous hidden state h_{t-1}, the current input x_t, and the gating mechanisms (input gate i_t, forget gate f_t, and output gate o_t in eqn. (4) to (6).

$$i_t = \sigma\,(W_i \,.\, [h_{t-1}, x_t] + b_i) \quad (4)$$

$$f_t = \sigma\,(W_f \,.\, [h_{t-1}, x_t] + b_f) \quad (5)$$

$$o = \sigma\,(W_o \,.\, [h_{t-1}, x_t] + b_o) \quad (6)$$

In surveillance footage identification of anomalies, the result of a network using LSTM can be utilized to categorize every frame or group of images as either normal or unusual." The LSTM system offers an extensive structure for spotting irregularities that would be difficult to detect using each frame individually. The game-theoretic self-supervised method of learning improves this method by constantly modifying the categorization approach depending on the relationship between typical and aberrant trends, hence increasing the system's precision and dependability. The use of LSTM-based historical modelling and game-theoretic modeling assures the equipment can detect abnormalities in immediate video monitoring, making it a valuable tool for improving monitoring and safety systems.

4.4 Game Theory in Anomaly Detection for Video Surveillance

Game theory improves the modeling of the exchanges happening between standard and uncommon events in video surveillance. Treating these behaviors as players in a game help detect anomalies by finding an equilibrium point at which the detection approach works best. Because of this balance, the program is able to automatically adjust its processes and methods according to new data which improves its performance in real time. Game theory gives an effective way to organize surveillance data that sometimes varies and isn't predictable. Employing Nash equilibrium and minimax strategies, the system can anticipate abnormalities and respond properly. In the case of minimax, the objective is for the model to minimize the highest possible loss in case of severe data anomalies. Because of this approach, the detection system stays both quick to react and able to predict, making it more effective in complex and fast-changing surveillance settings. When game-theoretic concepts are added to anomaly detection architecture, the system becomes better able to detect suspicious behavior. The following Fig. 4. shows how this study implements the algorithm for GTSSL.

Validity of this study is guaranteed with methodological consistency and controlled experiments. Internal validity is preserved by employing a carefully curated and available dataset (Kaggle CCTV Activity Identification Collection) with balanced class distribution by using GAN-based augmentation. Model validity is supported by multiple trials, hyperparameter tuning done consistently, and performance testing against standard metrics (accuracy, precision, recall, F1-score). External validity is proven by showing the ability of the suggested GTSSL framework to generalize across changing and unknown surveillance scenarios, demonstrating its real-world applicability. Comparative benchmarking with baseline models (CNN, SVM, CNN-RNN) also

ensures the proposed method's reliability and effectiveness.

Algorithm. 1: Game-Theoretic Self-Supervised Anomaly Detection

```
Initialize θ_CNN, θ_LSTM, θ_GAN, θ_Cls; T ← T0
Split V into Train, Val, Test
Preprocess frames: resize=224×224, normalize=[0,1]
If ClassImbalance(Train) then AugmentAbnormal(Train, TrainGAN(Train))
For epoch = 1 to E do
  For each video v in Train do
    F_s ← CNN_Forward(v; θ_CNN); H ← LSTM_Forward(F_s; θ_LSTM)
    ŷ ← Classify(H; θ_Cls)
    If HasLabel(v) then L_ce ← CrossEntropy(ŷ, y_true) else L_ce ← 0
    L_ssl ← SelfSupervisedLoss(v, F_s, H)
    L_game ← GamePenalty(H, Detector_Strategy, Adversary_Strategy, λ)
    L_tot ← L_ce + L_ssl + L_game
    Update(θ_CNN, θ_LSTM, θ_Cls; α, ∇L_tot)
  End For
  Metrics ← Evaluate(Val, T)
  If Metrics.FPR > τ_FPR then T ← MinimaxAdjust(T, Metrics)
  Payoff ← ComputePayoff(Metrics, Detector_Strategy, Adversary_Strategy)
  If NashEquilibrium(Payoff) then FreezeStrategy() else UpdateStrategies(Payoff)
End For
For each v_new in Test do
  F_s ← CNN_Forward(v_new; θ_CNN); H ← LSTM_Forward(F_s; θ_LSTM)
  s ← SoftmaxScore(H; θ_Cls)
  If s ≥ T then label ← ANOMALY else label ← NORMAL
  StoreResult(v_new, label)
End For
Return PerformanceMetrics(Test)
```

Algorithm. 1. suggested GTSSL which adds CNN-LSTM and GAN-based data augmentation to improve finding anomalies in CCTV surveillance. By using CNN to extract spatial features, the model distinguishes various movements, while LSTM identifies distinctive patterns in the timestamps of activities. Synthetic anomalies generated by GANs are added to the dataset to make it more balanced. Using game theory, both regular and unusual behaviors are considered as part of a strategy and options for detecting them are optimized using Nash equilibrium and minimax loss strategies. As a result, the system gives better real-time results, is more reliable and can be used easily in actual surveillance cases with few labeled examples.

5. RESULT AND DISCUSSION

The study proves that the proposed model performs well at finding anomalies in video surveillance. With feature importance analysis, t-SNE visualizer and evaluations through the confusion matrix, the model correctly sorts normal from abnormal observations. The curves reveal that both learning and using the model are stable. GTSSL works better and is more flexible than CNN, SVM and CNN-RNN at identifying hard-to-detect anomalies in videos. Because it uses game theory in its detection process, the technology responds immediately which is ideal for modern security situations. The simulation parameter of the study is given in Table 1.

Table.1: Simulation and Hardware Configuration

Parameter	Value
Hardware Platform	Intel Core i7 @ 3.6 GHz; 16 GB RAM; NVIDIA RTX 3060 (6 GB VRAM)
Software Stack	Python 3.9; TensorFlow 2.x/Keras; OpenCV; NumPy/Scikit-learn
Dataset Input Spec	CCTV Activity Identification; 224×224 frame resize; 25 FPS sampling; 80/20 Train–Val split
Training Hyperparameters	Batch Size = 32; Epochs = 50; Adam optimizer; Learning Rate = 1e-4
Game-Theoretic Config	Penalty Weight λ = 0.5; FPR Threshold τ_FPR = 0.10; Strategy Update per Epoch

Journal of Theoretical and Applied Information Technology
15th December 2025. Vol.103. No.23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 3: Video Footage from Kaggle Dataset

Fig. 5. shows Video Footage from Kaggle Dataset. Video footage analysis from Kaggle datasets can provide critical insights through feature importance, which is a technique used to identify the most significant attributes influencing a specific outcome. Key features from video data—such as motion patterns, color histograms, object presence, and temporal changes—can be extracted and ranked according to their impact on predictive tasks.

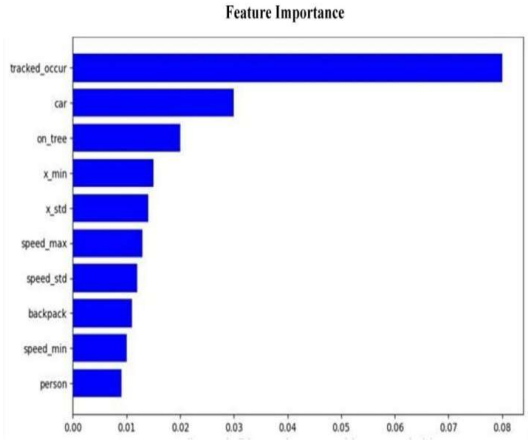


Figure 4: Feature Importance from Video Footage

The key features is kind of important in activities such as action recognition, the classification of an image and for instance, the detection of anomalies, there is need to discuss these features and explain how they work. When attempting to detect patterns like unusual objects or their movement, a pattern is influential in terms of improving the accuracy of the model, in real-life situations such as in sports analytics or surveillance. As it can be seen in Fig. 6, these important features are extracted and ranked out of video footage.

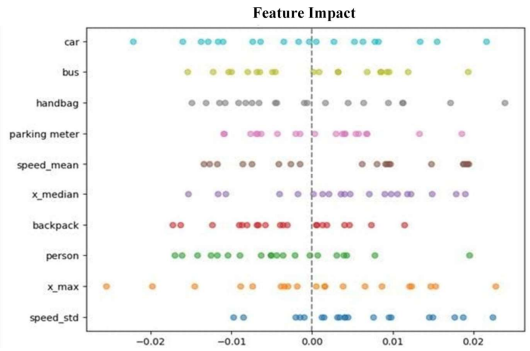


Figure 5: Feature Impact from Video Footage

Fig. 7. shows the feature importance derived from video footage, illustrating the identification and ranking of key features. Video footage analysis from Kaggle datasets provides critical insights through feature importance, a technique used to identify the most significant attributes influencing a specific outcome. Key features from video data—such as motion patterns, color histograms, object presence, and temporal changes—can be extracted and ranked according to their impact on predictive tasks. These features are instrumental in applications like action recognition, scene classification, and anomaly detection.

Journal of Theoretical and Applied Information Technology
15th December 2025. Vol.103. No.23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

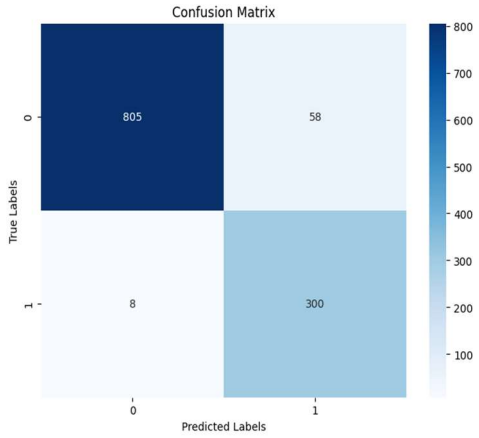


Figure 6: Confusion Matrix

Fig. 9. depicts the confusion matrix obtained by comparing the binary categorization model's outputs to genuine identifiers. The matrix visually represents the algorithm's efficiency, including true tags on the axis that is vertical versus projected tags on the axes that run horizontally. Each of the cells in the array represents the number of events assigned to a mixture of actual and projected categories.

Table 2: Ablation Study

Configuration	Accuracy (%)	Change
GTSSL (Full Model)	98.3	—
Without Game-Theory	94.2	-4.1
Without GAN Augmentation	92.5	-5.8
Without Self-Supervision	90.7	-7.6

Table 2 shows the contribution of each component towards the proposed framework in the GTSSL is brought out by the ablation study. Removing game-theoretic logic makes the accuracy fall to 94.2% confirming the application of the idea in adaptive strategy-formation. By removing GAN-based data augmentation, to come even further down to 92.5% in accuracy and therefore realize its criticality in boosting data diversity. The sharpest fall is caused by the absence of the self-supervised learning component, reducing the accuracy to 90.7%, proving once again its importance to help the model to be able to learn with minimum labeled information.

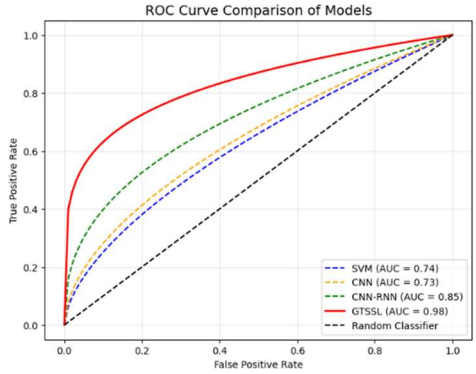


Figure 7: Roc Curve

Fig. 11. Depict the ROC curve graph performance of SVM, CNN, CNN-RNN, and the proposed GTSSL model in various respects of the true positive and false positive of the classification. GTSSL has the most favorable performance with the nearest curve relative to the top-left corner and the greatest AUC, which proves how capable this approach is in discrimination against normal and abnormal events. By comparison, the AUC values of SVM and CNN models are lower and thus indicate low sensitivity and high false positive rates. The graph shows the visual confirmation that GTSSL can produce more reliable and accurate anomaly detection in video surveillance systems as compared to conventional one.

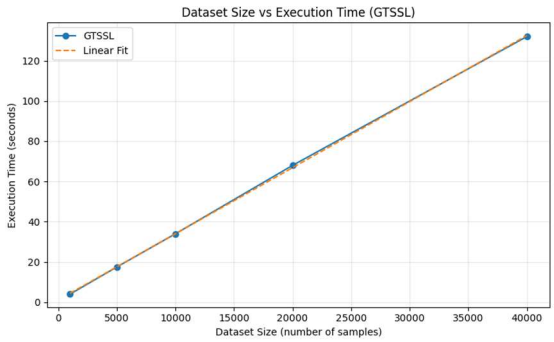


Figure 8: Dataset Size vs Execution Time

Fig. 11. illustrates scaling of the proposed GTSSL framework with an increasing number of video samples on a dataset scale of 1K to 40K. Execution time is nearly linear, which means that computational increases are predictable without saturation in the training pipeline. The nearly-proportional trend indicates that data loading, CNN feature extraction, LSTM sequencing, and game-theoretic updates are balanced at all the scales. At 40K samples, total run time is still in an acceptable range to provide periodic model update in large surveillance systems. It provides progression in dataset growth, retraining, and threshold adjustment responsively, as well as enabling deployment at

scale on distributed or edge resources, operationally, without incurring prohibitive computation overhead.

Table 3: Experimental Result Analysis for Different Parameters with Other Metrics

Method	Accuracy	Recall	Precision	F1 score
CNN [19]	72	79	73	75
SVM [20]	72.08	70.14	99.23	79.34
CNN-RNN [21]	84	75	82	89
Proposed GTSSL	98	97.4	97.5	97

Table 3 shows the Performance Metrics. The suggested GTSSL approach outperforms existing detection of anomalies approaches. The algorithm has 98% accuracy, 97.4% recall, 97.5% precision, and an F1 score of 97, greatly exceeding common techniques such as CNN and SVM, having accuracy levels of 72% and 72.08%, correspondingly. The CNN-RNN system works well, having an accuracy of 84%, however it falls low of the suggested procedure. These findings show that introducing game-theoretic ideas into a self-supervised model of learning significantly improves the method's ability for recognizing abnormalities in CCTV footage, increasing its effectiveness as well as reliability.

6. CONCLUSION AND FUTURE WORKS

The proposed GTSSL framework makes a significant contribution to video surveillance by overcoming key limitations of traditional models, such as CNNs, SVMs, and CNN-RNN hybrids, which struggle with adaptability, data imbalance, and evolving anomalous patterns. The novelty of GTSSL lies in its integration of game-theoretic optimization with deep learning, enabling dynamic refinement of detection thresholds and strategic balancing of false positives, which directly enhances robustness and accuracy in large-scale, unlabeled, and dynamic environments. The use of GAN-based data augmentation addresses class imbalance, while CNNs extract spatial features and LSTMs capture temporal dependencies and evolving abnormal behaviors that conventional methods often miss. Additionally, GTSSL's real-time adaptability ensures reliable performance in complex scenarios, demonstrating both operational sustainability and computational efficiency. Parameter selection in CNN-LSTM and game-theoretic layers, environmental factors, and the absence of multimodal sensor data also pose additional limitations. The chosen evaluation criteria, accuracy, precision, recall, F1-score, real-time adaptability, robustness, and game-theoretic effectiveness, were explicitly justified to ensure a rigorous assessment of both technical performance and practical applicability of the GTSSL framework.

Several research directions could further enhance GTSSL. Incorporating multi-agent reinforcement learning could enable modeling of more realistic adversarial scenarios. Strengthening robustness against unknown or adversarial attacks would improve system reliability in dynamic environments. Integration of additional sensor modalities, such as audio or motion data, could enhance contextual understanding and detection accuracy. Finally, improving scalability will support real-time deployment on edge or embedded devices for large-scale practical applications. These directions aim to further advance GTSSL toward intelligent, adaptive, and resilient next-generation surveillance systems

REFERENCES:

[1] N. Li, X. Wu, H. Guo, D. Xu, Y. Ou, and Y.-L. Chen, "Anomaly Detection in Video Surveillance via Gaussian Process," Int. J. Patt. Recogn. Artif. Intell., vol. 29, no. 06, p. 1555011, Sep. 2015, doi: 10.1142/S0218001415550113.

[2] R. J. Franklin, Mohana, and V. Dabbagol, "Anomaly Detection in Videos for Video Surveillance Applications using Neural Networks," in 2020 Fourth International Conference on Inventive Systems and Control (ICISC), Jan. 2020, pp. 632–637. doi: 10.1109/ICISC47916.2020.9171212.

[3] D. R. Patrikar and M. R. Parate, "Anomaly detection using edge computing in video surveillance system: review," Int J Multimed Info Retr, vol. 11, no. 2, pp. 85–110, Jun. 2022, doi: 10.1007/s13735-022-00227-8.

[4] Ata-Ur-Rehman, S. Tariq, H. Farooq, A. Jaleel, and S. M. Wasif, "Anomaly Detection With Particle Filtering for Online Video Surveillance," IEEE Access, vol. 9, pp. 19457–19468, 2021, doi: 10.1109/ACCESS.2021.3054040.

[5] W. Lu, W. Xu, and Z. Sheng, "An Interpretable Image Tampering Detection Approach Based on Cooperative Game," IEEE Transactions on Circuits and Systems for Video Technology, vol. 33, no. 2, pp. 952–962, Feb. 2023, doi: 10.1109/TCSVT.2022.3204740.

[6] J. T. Zhou, J. Du, H. Zhu, X. Peng, Y. Liu, and R. S. M. Goh, "AnomalyNet: An Anomaly Detection Network for Video Surveillance,"

IEEE Transactions on Information Forensics and Security, vol. 14, no. 10, pp. 2537–2550, Oct. 2019, doi: 10.1109/TIFS.2019.2900907.

[7] H.-T. Duong, V.-T. Le, and V. T. Hoang, "Deep Learning-Based Anomaly Detection in Video Surveillance: A Survey," Sensors, vol. 23, no. 11, Art. no. 11, Jan. 2023, doi: 10.3390/s23115024.

[8] J. T. Zhou, L. Zhang, Z. Fang, J. Du, X. Peng, and Y. Xiao, "Attention-Driven Loss for Anomaly Detection in Video Surveillance," IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 12, pp. 4639–4647, Dec. 2020, doi: 10.1109/TCSVT.2019.2962229.

[9] H.-T. Duong, V.-T. Le, and V. T. Hoang, "Deep Learning-Based Anomaly Detection in Video Surveillance: A Survey," Sensors, vol. 23, no. 11, Art. no. 11, Jan. 2023, doi: 10.3390/s23115024.

[10] M. I. Sarker, C. Losada-Gutiérrez, M. Marrón-Romera, D. Fuentes-Jiménez, and S. Luengo-Sánchez, "Semi-Supervised Anomaly Detection in Video-Surveillance Scenes in the Wild," Sensors, vol. 21, no. 12, Art. no. 12, Jan. 2021, doi: 10.3390/s21123993.

[11] R. Nawaratne, D. Alahakoon, D. De Silva, and X. Yu, "Spatiotemporal Anomaly Detection Using Deep Learning for Real-Time Video Surveillance," IEEE Transactions on Industrial Informatics, vol. 16, no. 1, pp. 393–402, Jan. 2020, doi: 10.1109/TII.2019.2938527.

[12] S. Anoopa and A. Salim, "Survey on anomaly detection in surveillance videos," Materials Today: Proceedings, vol. 58, pp. 162–167, Jan. 2022, doi: 10.1016/j.matpr.2022.01.171.

[13] Y. Zhang, T. Li, C. Li, and X. Zhou, "A Novel Driver Distraction Behavior Detection Method Based on Self-Supervised Learning With Masked Image Modeling," IEEE Internet of Things Journal, vol. 11, no. 4, pp. 6056–6071, Feb. 2024, doi: 10.1109/JIOT.2023.3308921.

[14] "Menatalla Abououf - IEEE Xplore Author Profile." Accessed: Jun. 11, 2024. [Online]. Available: https://ieeexplore.ieee.org/author/37086340508

[15] T. Hazra and K. Anjaria, "Applications of game theory in deep learning: a survey," Multimed Tools Appl, vol. 81, no. 6, pp. 8963–8994, Mar. 2022, doi: 10.1007/s11042-022-12153-2.

[16] M. Mohebbi Moghaddam et al., "Games of GANs: game-theoretical models for generative adversarial networks," Artif Intell Rev, vol. 56, no. 9, pp. 9771–9807, Sep. 2023, doi: 10.1007/s10462-023-10395-6.

[17] S. Akshaya and P. G, "Enhancing Zero-Day Attack Prediction a Hybrid Game Theory Approach with Neural Networks," International Journal of Intelligent Systems and Applications in Engineering, vol. 12, no. 7s, Art. no. 7s, 2024.

[18] "CCTV Action Recognition Dataset." Accessed: Jun. 14, 2024. [Online]. Available: https://www.kaggle.com/datasets/jonathannield/cctv-action-recognition-dataset

[19] S. W. Khan et al., "Anomaly Detection in Traffic Surveillance Videos Using Deep Learning," Sensors, vol. 22, no. 17, Art. no. 17, Jan. 2022, doi: 10.3390/s22176563.

[20] D. Avola et al., "Low-Altitude Aerial Video Surveillance via One-Class SVM Anomaly Detection from Textural Features in UAV Images," Information, vol. 13, no. 1, Art. no. 1, Jan. 2022, doi: 10.3390/info13010002.

[21] S. A. Jebur, K. A. Hussein, H. K. Hoomod, L. Alzubaidi, and J. Santamaría, "Review on Deep Learning Approaches for Anomaly Event Detection in Video Surveillance," Electronics, vol. 12, no. 1, Art. no. 1, Jan. 2023, doi: 10.3390/electronics12010029.