Kufa Journal of Engineering
Vol. 15, No. 4, October 2024, P.P. 125-140
Article history: Received 27 April 2024, last revised 2 July 2024, accepted 3 July 2024

CLOUD-SMART SURVEILLANCE: ENHANCING ANOMALY DETECTION IN VIDEO STREAMS WITH DF-CONVLSTM-BASED VAE-GAN

Sivalingan H

Assistant Professor, PG and Research Department of Computer Science, Providence College for Women (Autonomous), Coonoor, The Nilgiris. Email: drsivah23@gmail.com

https://doi.org/10.30572/2018/KJE/150409

ABSTRACT

Anomaly detection in computer vision is crucial, and manual identification of irregularities in videos is resource-intensive. Autonomous systems are essential for efficiently analysing and detecting anomalies in diverse video datasets. Video surveillance relies heavily on anomaly detection for monitoring equipment states through time-series data. Presently, deep learning methods, particularly those based on Generative Adversarial Networks (GAN), have gained prominence in time-series anomaly detection. This paper proposes a novel solution: the double-flow convolutional Long Short-Term Memory (DF-ConvLSTM) - based Variational Autoencoder- Generative Adversarial Network (VAE-GAN) method. By co-training the encoder, generator, and discriminator, this approach leverages the encoder's mapping skills and the discriminator's discrimination capabilities simultaneously. The proposed strategy is compared with LSTM-VAE, LSTM-VAE-Attention, and VAE. The proposed method is evaluated using metrics for recall, accuracy, precision, and F1 score. With classification accuracies of 91% on the University of Central Florida (UCF) crime dataset, the experimental results outperformed alternative techniques. Furthermore, the analysis of the ROC curve revealed that the suggested method performed better than the others, as evidenced by its higher ROC (Receiver Operating Characteristic) values. Experimental results demonstrate the proposed method's ability to rapidly and accurately detect anomalies in surveillance videos, ensuring efficient and reliable anomaly detection. Experimental results show the method's rapid, accurate anomaly detection in surveillance videos, ensuring efficiency and reliability. However, challenges include high computational costs, affecting the practicality of implementation for real-time anomaly detection.

KEYWORDS

Anomaly Detection, Video surveillance, Time series, LSTM, VAE, GAN.

1. INTRODUCTION

The proliferation of surveillance systems in public spaces has escalated, necessitating real-time analysis of video surveillance streams. A surge in research efforts has focused on enhancing detection algorithm performance, addressing challenges arising from factors like data volume, dynamic environments, and diverse anomaly types. Video anomaly detection involves identifying frames within a video sequence that deviate significantly from the expected norm, facilitating the pinpointing of unusual occurrences such as fires, altercations, . . .etc. Researchers like (Pang et al. ,2020) and (Mahmood et al. ,2021) have explored this area, showcasing its significant practical utility. Beyond security applications, anomaly detection in video surveillance proves valuable in industrial settings, enabling the detection of equipment malfunctions or defects. However, it takes a lot of time and labour for humans to manually identify anomalies in surveillance footage. This is because manual analysis is an impractical solution due to the volume of data generated by critical systems in security applications. The need for automated systems to identify anomalies in videos has grown significantly in the last few years. Automated systems are more productive and economical when it comes to identifying irregularities in surveillance footage because they drastically cut down on human labour and time. Finding and following anomalies in the captured footage is one of the surveillance field's expanding research challenges. To mitigate this problem and save a great deal of time and human labour, researchers have suggested automated systems for identifying anomalies in videos. Even with the progress that has been made, there is always space for improvement in terms of precision, dependability, and expandability when creating an ideal video surveillance system. With the development of new models and improvements in hardware performance in recent years, smart learning techniques have gained popularity in the field of video anomaly detection.

However, there are a number of limitations to using these strategies. The creation of big data, which demands a significant amount of processing power, is one of the main obstacles. This study has addressed the difficulties involved in intelligent anomaly detection while concentrating on creating new, more effective, and efficient anomaly detection techniques. In general, the continued development of video surveillance depends on ability to comprehend the drawbacks of existing unsupervised anomaly detection techniques and investigate novel approaches by (Pereira and Silveira ,2018). Soft computing techniques, including machine learning and deep learning classifiers, are pivotal for early prediction of anomalous activities in videos (lin et al. ,2020) and (xu et al ,2018). The recent proliferation of deep learning techniques has brought about automation with higher accuracy compared to traditional machine learning classifiers, marking a notable advancement in video anomaly detection. This paper's main goal is to investigate this gap for Anomaly detection using deep learning methods. Using University of Central Florida (UCF) crime dataset, the suggested method makes use of explanatory power to comprehend AI decisions.

This paper's main contributions can be summed up as follows:

- The study presents DF-ConvLSTM -based VAE-GAN, a ground-breaking method for anomaly detection in video surveillance.
- To demonstrate the effectiveness of the model, compare the performance of the suggested framework with the current approaches using the UCF-Crime dataset.
- The model's performance was assessed using a variety of metrics, including area under the ROC curve, recall, accuracy, precision, and F1-score.
- The experiment's outcomes showed that the suggested model performed better than the others.

The structure of the paper is as follows: A review of recent developments in the field is given in Section 2, followed by an explanation of the proposed technique in Section 3, testing and comparison of the suggested method with alternatives in Section 4, and a conclusion detailing future research possibilities in Section 5.

2. RELATED WORKS

One important advancement that improves security and efficiency in cloud computing is anomaly detection through the use of video monitoring. This device continuously analyses camera feeds to spot odd or suspicious activity by utilizing cloud resources. Graphical representations of these anomaly categories are shown in Fig.1.

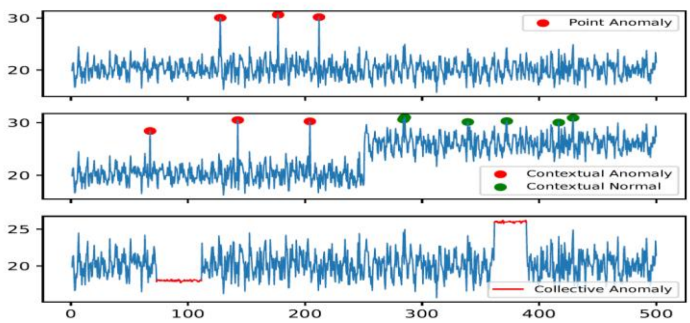

Fig.1. Types of anomalies.

According to a recent study by (Baur et al. ,2018) developed VAE-GAN, a method for anomaly detection that incorporates GANs for adversarial training. During testing, VAE-GAN utilizes pixel-level L1-distance to evaluate abnormality. (Sultani et al. ,2018) introduced a method employing deep multi-instance ranking, training on ambiguously labelled videos with video-level labels to enhance anomaly detection by learning from both normal and anomalous events. (Schlegl et al. ,2019) used a similar methodology, concentrating on anomaly identification in optical coherence tomography images and replacing variational autoencoders with convolutional ones. Deep generative models can include reconstruction probability or likelihood scores as extra anomaly indicators in their detection processes, in addition to the standard pixel distance.

(Meena et al. ,2019) highlighted the difficulties in anomaly detection, especially when employing video surveillance in traffic situations. Their suggested technique made use of spatiotemporal elements to identify a number of anomalies, such as cars driving on pedestrian walkways and on the wrong side of the road. This joint-based method integrated ConvLSTM with k-means clustering and combined reconstruction and clustering loss. Through the utilization of convolutional neural networks (CNN) with long short-term memory (LSTM) and previous frame data, the model effectively recognized and detected abnormalities in the traffic scene. The real-time implementation outperformed existing approaches in anomaly identification inside traffic scenes, achieving an exceptional accuracy rate of 93.02%. Testing and training phases were carried out to discern between normal and abnormal events.

A unique hybrid autoencoder for anomaly detection in video data was suggested by (Zhou et al. ,2020). They discovered a drawback of conventional LSTM autoencoders: because of fixed dimension representations, they had trouble handling global context anomalies. They developed a hybrid autoencoder that extracted both temporal and spatial information in order to get around this restriction. In addition, they added shortcut connections to improve the decoder's functionality. Reconstruction error was the basis for anomaly identification, and this method performed better when managing global context anomalies. (Shi et al. ,2022), cloud computing is recognized as a major development in distributed computing that offers exciting business opportunities. Effective resource management and use are made possible by this centralization. Virtual machines (VMs) are the means via which all resources inside the IaaS architecture can be accessed. Dynamic resource allocation approaches are essential for maximizing resource utilization, cost savings, and computing efficiency. This is especially true when using cutting-edge multi-objective optimization methodologies that guarantee stable resource distribution among several virtual machines.

These traditional methods for surveillance video anomaly detection are often ineffective and time-consuming. This paper introduces a novel anomaly detection system for video surveillance leveraging DF-ConvLSTM-based VAE-GAN. Experimental results showcase its effectiveness through impressive F1 Score, accuracy, recall, and precision metrics. Addressing existing research gaps, the system offers enhanced accuracy, real-time performance, and robustness across diverse surveillance videos. These findings represent a significant advancement in anomaly detection technology, emphasizing the critical role of reliable detection systems in achieving early outcomes.

3. MATERIALS AND METHODS

By utilizing the discriminator's discriminating skills and the encoder's mapping prowess, this novel anomaly detection technique trains the discriminator, generator, and encoder all at once. It allows the model to discover abnormalities quickly by cutting down on the amount of time required for anomaly detection. Furthermore, the combined optimization of these modules improves the model's ability to identify anomalies with a significant degree of accuracy.

The suggested method's process is shown in Fig. 2, where videos taken with a security camera are sent to the cloud platform as input data at predetermined intervals. After being transformed into a timeseries format, this raw visual data is pre-processed and feature extracted. After that, the input is passed into a deep learning system, which uses a learning technique to model the behaviour of surveillance targets and evaluate abnormality based on anomaly scores. Five tasks have to be performed for prediction: i) Data collection ii) Pre-processing the data; iii) Feature selection; iv) Data partition and v) Detection methodology.

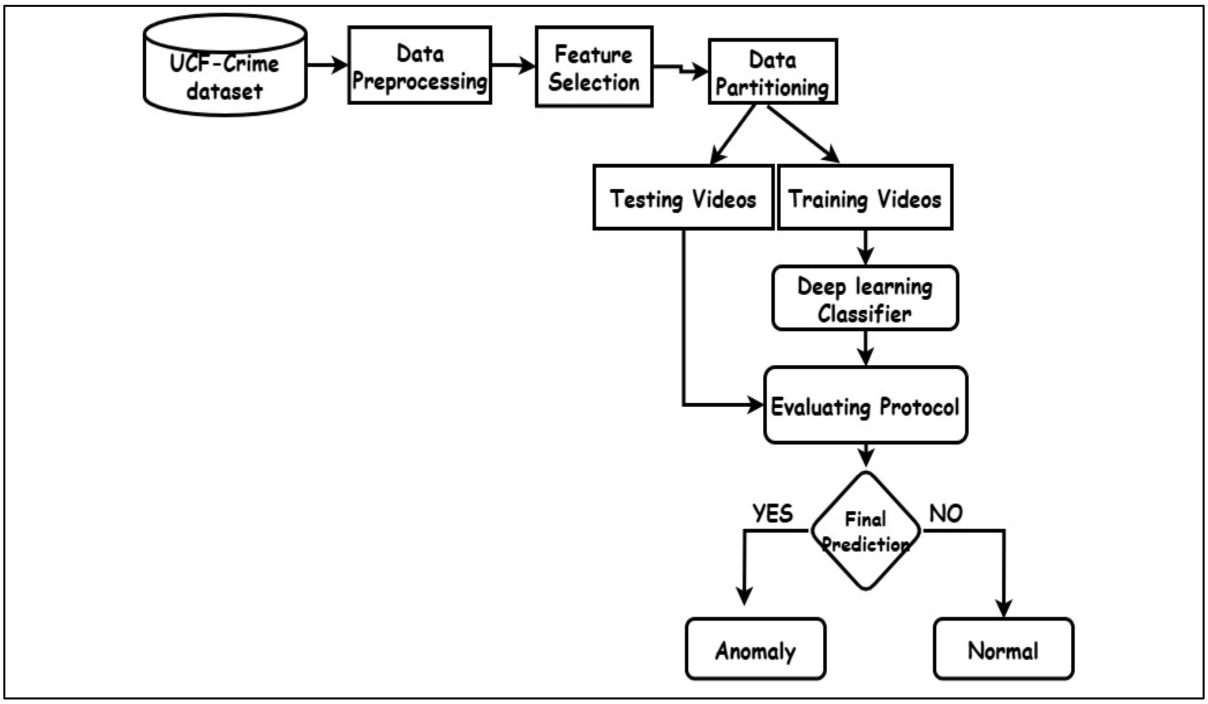

Fig.2. Flow diagram of proposed work

3.1. Data Collection

The University of Central Florida (UCF) dataset, introduced by Sultani, (Chen, and Shah ,2018), is a comprehensive resource for tackling anomaly detection challenges shown in Table 1. This dataset encompasses 128 hours of video data, featuring 1,900 real-world street and indoor surveillance cameras. These cameras, utilizing various RGB technologies, capture diverse locations, providing a multi-scene perspective crucial for robust anomaly detection models. The dataset comprises 13 anomaly classes, covering a spectrum of activities. With 1,610 training videos (800 normal, 810 abnormal) and 290 test videos (150 normal, 140 deviant), the dataset is meticulously annotated with video-level and frame-level labels, respectively. Notably, the balanced representation of normal and abnormal samples in both training and testing sets enhances the dataset's utility for developing and evaluating anomaly detection algorithms.

Table 1. UCF-Crime dataset

Type of Anomaly	No of videos	Training	Test set
Abuse	50	48	2
Arrest	50	45	5
Arson	50	41	9
Assault	50	47	3
Burglary	100	87	13
Explosion	50	29	21
Fighting	50	45	5
Road Accidents	150	127	23
Robbery	150	145	5
Shooting	50	27	23
Shoplifting	50	29	21
Stealing	100	95	5
Vandalism	50	45	5
Normal Events	950	810	140

3.2. Preprocessing

Data cleaning in anomaly detection involves identifying and rectifying errors or inconsistencies in data to enhance the accuracy of anomaly detection models. This process includes handling missing values, removing duplicates, and correcting erroneous data entries. One key aspect is outlier detection and removal, which can be performed using statistical methods like Z-score. The Z-score measures how many standard deviations a data point is away from the mean of the dataset. For instance, using the Z-score method, anomalies can be identified if the absolute Z-score value of a data point exceeds a threshold (usually 3):

$$\mathrm{Z-score} = \frac{(X-\mu)}{\sigma} \quad (1)$$

Where X is the data point, μ is the mean, and σ is the standard deviation of the dataset. Points with Z-scores outside the range of [-3, 3] are considered outliers.

Data normalization is crucial to ensure that features from different scales contribute equally to the detection process. Min-Max normalization, also known as feature scaling, it transforms features to a fixed range, typically [0, 1]. This process prevents features with larger numeric ranges from dominating those with smaller ranges, which is essential for many anomaly detection algorithms that rely on distance metrics or statistical thresholds. The formula for Min-Max normalization is:

$$X' = \frac{X-X_{min}}{X_{max}-X_{min}} \quad (2)$$

where X is the original value, X_{min} and X_{max} are the minimum and maximum values of the feature, and X′ is the normalized value. In anomaly detection, normalized data ensures that anomalies can be detected more effectively by eliminating the bias introduced by varying scales of features. This enhances the performance of algorithms as they rely on the relative distances between data points. Normalization makes it easier to identify outliers that deviate significantly from the normalized range.

3.3. Feature selection

Feature selection in anomaly detection using Double-Flow ConvLSTM involves identifying the most relevant features that contribute to detecting anomalies in spatiotemporal data. This model is particularly effective for handling structured data, such as videos or time series and it has two parallel ConvLSTM flows: one for the spatial domain and another for the temporal domain. This ensures that the ConvLSTM model focuses on the most critical aspects of the data, leading to more reliable and timely anomaly identification. The feature selection process can be formulated as:

$$X_t = F(W_f * X_{t-1} + W_i * I_t + b) \quad (3)$$

Where X_t is the selected feature set at time t, F is the activation function, W_f and W_i are weight matrices for the forget and input gates, $*$ denotes convolution, X_{t-1} is the feature set at the previous time step, I_t is the input at time t, b is the bias term.

3.4. Data Partition

Dataset is partitioned into training and testing dataset. First, the dataset is split into two parts: a larger portion for training and a smaller one for testing. The training dataset should primarily contain normal data to allow the model to learn typical patterns without bias from anomalies. This helps the DF-ConvLSTM-VAE model, integrated with GAN, to effectively encode normal behaviour and generate realistic data. During testing, the model evaluates both normal and anomalous data. Anomalies are detected based on deviations from the learned normal patterns. Proper data partitioning ensures the model is trained on representative normal data, enhancing its ability to accurately detect anomalies during testing.

3.5. Detection using DF-ConvLSTM-Based VAE-GAN

The DF-ConvLSTM-based VAE-GAN method leverages both the encoder's mapping abilities and the discriminator's capabilities to achieve high-quality image synthesis and reconstruction. In this hybrid framework, the encoder, part of the Variational Autoencoder (VAE) structure, is responsible for mapping input data to a latent space. The ConvLSTM layers enhance this mapping by capturing temporal dependencies and spatial features, which is particularly beneficial for sequential data or dynamic textures in videos. The encoder transforms the input image into a compressed latent representation, capturing the essential features while reducing dimensionality. This latent code is then fed into the decoder, which attempts to reconstruct the original video, ensuring that the key characteristics of the input are preserved. Simultaneously, GAN component plays a crucial role. The generator, which is integrated with the VAE's decoder, produces images from the latent code. The discriminator, a vital element of the GAN, evaluates these generated images against real images, providing feedback to improve the generator's outputs. The discriminator's capabilities in distinguishing real from fake video drive the generator to produce more realistic and high-fidelity images over time. The two networks work in parallel form for efficient process in anomaly detection.

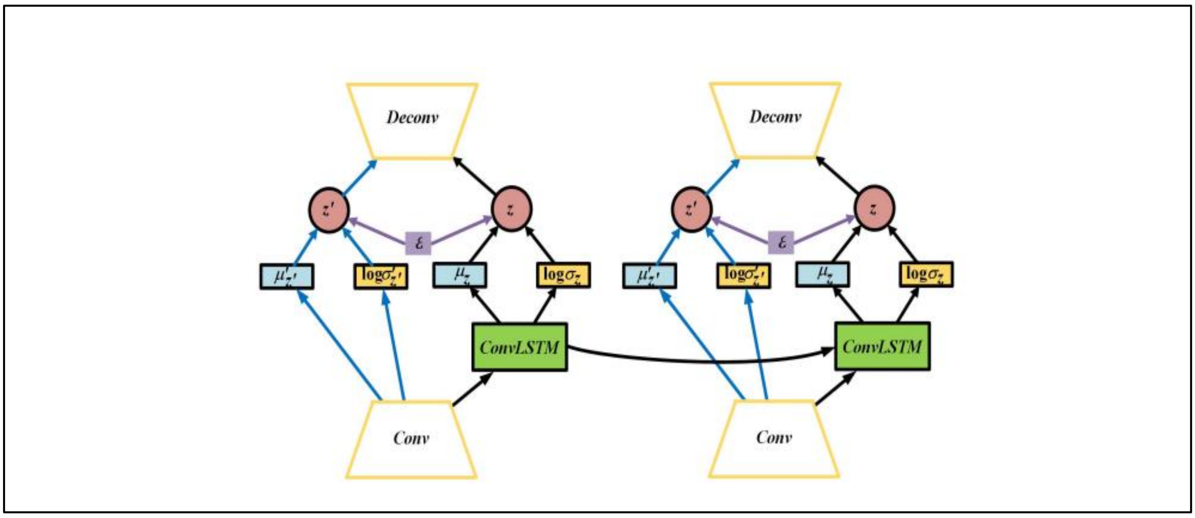

Fig. 3. Architecture of proposed method

The proposed model's structure, which consists of the left and right flows, is shown in Fig. 3. Figure 3's left and right flows are indicated by blue and black arrows, respectively. The three components of the left flow model are the encoder (Conv), sample, and decoder (Deconv). The encoder comprises Conv and ConvLSTM modules. ConvLSTM captures temporal video patterns from spatial features, while Conv utilizes convolutional layers to extract spatial features from each frame. Sampling process is made up of two sample processes: N (μ, σ^2) is used to sample data z of right flow, N (μ', σ'^2) is used to sample data z' of the left flow. Lastly, the Deconv module serves as the decoder.

$$L_{MSE}(\tilde{x}, x) = ||\tilde{x} - x||^2 \qquad (4)$$

$$L_{KLD}(\mu, \sigma) = \mathrm{KL}(\mathrm{N}(\mu, \sigma^2)||N(0, I)) \qquad (5)$$

Where L_{MSE} is the reconstruction error (MSE, the mean squared error) between the inputs x and their reconstructions $\tilde{x}$. μ, σ denotes mean and standard deviation. Kullback - Leibler divergence of left flow is represented by third term L'_{KLD} and the right flow by the second term L_{KLD}. Equation (6) represents the suggested model's objective function.

$$L = L_{MSE}(\tilde{x}, x) + L_{KLD}(\mu, \sigma) + L'_{KLD}(\mu', \sigma') \qquad (6)$$

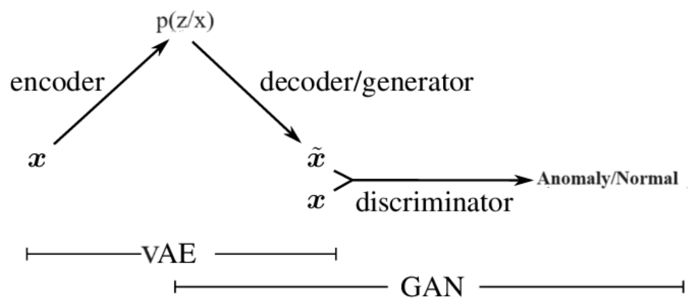

Fig. 4. VAE – GAN

The GAN component enhances the VAE by introducing a generator and a discriminator shown in Fig. 4. The generator synthesizes data samples from the learned latent space, aiming to trick the discriminator, which distinguishes between real (normal) and generated (potentially anomalous) data. Training and deploying the DF-ConvLSTM-based VAE-GAN model incurs significant computational costs and resource demands. Training involves intensive GPU (Graphics Processing Units) usage due to complex convolutional and LSTM layers, requiring substantial memory and processing power. Deployment requires ongoing GPU resources for real-time inference, potentially limited by hardware capabilities and operational costs, necessitating efficient resource management and optimization strategies to balance performance and affordability in production environments.

3.6. Proposed Method Algorithm

Input: Normal training dataset X for every frame x_t, t = 1, . . . , T.
Output: probabilistic encoder E_φ, $\grave{E}_{\varphi}$, probabilistic decoder G_θ, $\grave{G}_\theta$
(E_φ = Conv + ConvLSTM, G_θ = Deconv, $\grave{E}_{\varphi`} = conv$, $G`_\theta$ = Deconv)
φ, θ, C_0, h_0, φ`, θ`, $\grave{C}_0$, $\grave{h}_0$ ← Initialize parameters
(φ, θ - denote the hidden parameters of encoder E and decoder G, C_0, h_0 - denotes the ConvLSTM unit cell, hidden state)

repeat
for t = 1 to T do
μ`, σ`, F_t = Conv (x_t)
μ, σ, C_t, h_t = ConvLSTM (F_t, C_{t-1}, h_{t-1})
z← samples from N (μ, $σ^2$)
z`← samples from N (μ`, $\sigma^{`2}$)
$\tilde{x}_t$= Deconv(z`, z)
Calculate $L_t = L_{MSE}(\tilde{x}, x) + L_{KLD}(\mu, \sigma) + L'_{KLD}(\mu', \sigma')$
end for
φ, θ, φ`, θ` update parameters using gradients of L =$\sum_{t=1}^{T} L_t$
until convergence of parameters

3.7. Anomaly Detection

Reconstruction error produced by VAE - GAN, discriminator's results are combined to form the anomaly score, which is derived from reconstruction error probability (REP) in this video anomaly detection model. The encoder uses test video clip's frame x_t as input to estimate parameters of the latent gaussian variables μ, σ, and uses this information as the output. Then, using the reparameterization trick, z is sampled for L times in accordance with latent distribution N (μ, $σ^2$), i.e. $z^{(l)} = \mu + \sigma \cdot \in^{(l)}$, where $\in \sim N(0, I)$ and l = 1, . . ., L. Generative network outputs the reconstructed frame $\tilde{x}_t^{(l)}$ after receiving z $^{(l)}$ as input data. Probability of reconstruction error for an intensity value (I) of a pixel at location (u, v) in frame x_t of video sequence using Equation (7).

$$REP_{(u,v,t)} = \frac{1}{L}\sum_{l=1}^{L}\left\|\tilde{I}^{(l)}_{(u,v,t)} - I_{(u,v,t)}\right\| 2 \quad (7)$$

where intensity in reconstructed frame $\tilde{x}_t^{(l)}$ is indicated by the symbol $\tilde{I}^{(l)}_{(u,v,t)}$. REP of frame x_t is calculated from each frame by adding up all of the probability of pixel-wise errors:
$REP_{(t)} = \sum_{(u,v)} REP_{(u,v,t)}$. A video sequence's regularity scores, or s(t), are determined by Equation (8):

$$s(t) = 1 - \frac{REP(t) - min_t REP(t)}{maxREP(t)} \quad (8)$$

Additionally, to determine the number of abnormal events in a video, analyse the noisy local minima in the regularity score time series to detect these anomalies. Anomalies are most likely to be present in video frames if there are distinct local minima. In this stage, two local minima are considered to be a component of the same abnormal event if their separation is smaller than 50 frames. For every point, the anomaly detector calculates an average anomaly score. Furthermore, anomalies exist in a subset of the test data, which facilitates the choice of an ideal threshold. This threshold allows it to find anomalies in the full test set by effectively identifying those in this subset and allowing it to generalize.

4. EXPERIMENTAL RESULTS

To assess the efficacy of the proposed DF-ConvLSTM -based VAE-GAN anomaly detection model, three baseline methods—VAE (Xu et al. ,2018), LSTM-VAE (Lin et al. ,2020), and LSTM-VAE-Attention (Pereira et al. ,2018)—were employed. Evaluation metrics include accuracy, precision, recall, and F1-score. The assessment utilizes the UCF-Crime dataset, providing a robust benchmark for comparing the model's performance against established methods in anomaly detection. In task classification accuracy, "true positive" (TP) denotes correctly identified anomalous events, "true negative" (TN) represents accurately identified normal events, "false positive" (FP) signifies incorrectly identified anomalous events, and "false negative" (FN) indicates wrongly identified normal events. These metrics are crucial for evaluating classification performance. Fig. 5 illustrates the accuracy of the proposed method alongside existing systems like VAE, LSTM-VAE, and LSTM-VAE-Attention, providing a comparative analysis of their performance in anomaly detection.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (9)$$

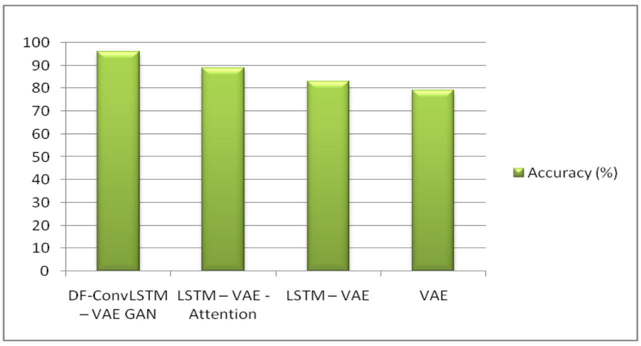

Fig.5. Comparison of accuracy with other models

Recall quantifies the percentage of successfully anticipated positive instances to all actual positive instances, whereas precision shows the ratio of correctly predicted positive cases to all projected positive cases. The F1 score offers a fair performance metric since it is a weighted harmonic mean of recall and precision. The thorough study based on these assessment measures in anomaly detection is shown in Fig. 6. The following is the calculating formula:

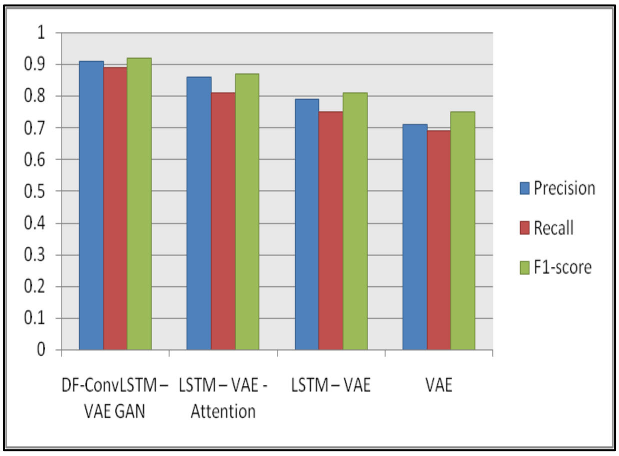

Fig.6. Comparison of evaluation metrics with other models

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{10}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{11}$$

$$\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{12}$$

When evaluating various threshold settings, the Receiver Operating Characteristic (ROC) curve serves as a valuable tool. The ROC curve incorporates the false positive rate (FPR) and the true positive rate (TPR). The FPR measures the ratio of false-positive outcomes to the total negative samples in the test, while the TPR evaluates the classifier's ability to correctly identify positive instances among all positive samples in the test.

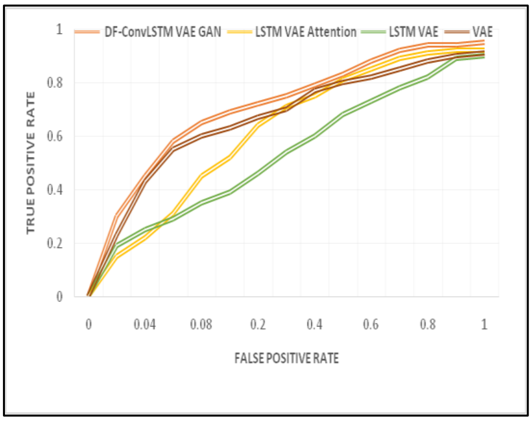

Fig. 7. ROC Curve for proposed and baseline method

Fig. 7 displays the ROC curve, illustrating model performance, where the proposed method outperforms other models. It adeptly balances actual positive and false positive rates, making it suitable for anomaly detection in surveillance systems. The Area Under the Curve (AUC), a classification metric, is computed using ground truth and frame-level anomaly scores, providing insight into the model's effectiveness in distinguishing anomalies from normal instances.

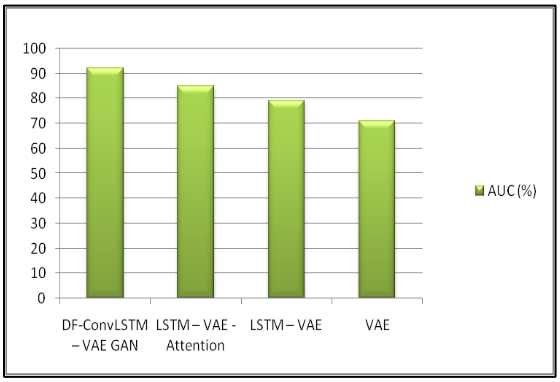

Fig.8. AUC results of different models

Fig. 8 highlights that proposed method achieves the highest AUC value of 97.2% among the models, showcasing superior performance. A larger AUC signifies the model's enhanced ability to distinguish between anomalies and normal instances. Fig. 9 shows the relationship between the computational time t and M corresponding to its methods (in graph red colour denote proposed system).

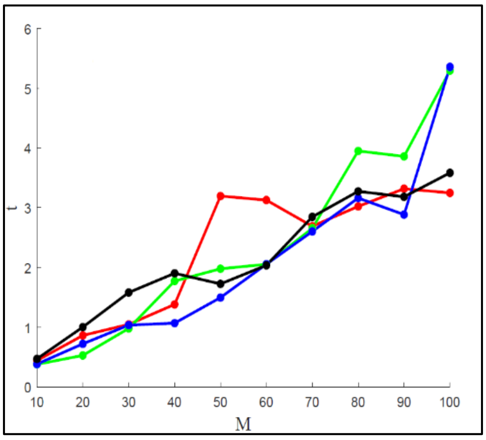

Fig. 9. The graph of computational time with method

This sophisticated method is utilized in various real-world applications due to its capabilities in handling sequential data and generating high-quality videos and images. It is particularly effective in video prediction and generation, where it can predict future frames in a video sequence. This method is also used in weather forecasting, helping to predict future weather conditions based on historical data. This method also utilized in enhancing medical imaging techniques. For instance, it improves the quality of MRI scans and CT images by generating high-resolution images from low-resolution inputs. The ConvLSTM component helps in understanding the temporal dependencies in sequential medical data, providing better diagnostic tools. Additionally, it finds applications in anomaly detection in surveillance systems, where it can identify unusual activities by learning normal patterns and detecting deviations. The method's versatility makes it valuable in diverse fields requiring advanced data modelling and prediction.

5. CONCLUSION AND FUTURE WORK

Traditional methods for surveillance video anomaly detection are often ineffective and time-consuming. In this research, a novel approach is introduced using DF-Conv LSTM-based VAE-GAN for precise and efficient detection. The primary objective is to monitor equipment operation through the analysis of collected data. The reconstruction process involves two phases: model training, where normal data patterns are learned, and anomaly detection, which calculates scores for each time series for precise identification. The proposed model simultaneously trains the encoder, generator, and discriminator, enabling swift and accurate anomaly detection without additional optimization. Experimental results demonstrate its superiority over other models, including LSTM-VAE-Attention, LSTM-VAE, and VAE, across various metrics, the current approach achieves.0.92 F-Measure, 96% Accuracy, 0.89% Recall, and 0.91% Precision. The proposed method achieves a noteworthy accuracy of 96%, surpassing other models by margins of 7%, 6%, and 4%. This research marks a substantial 9.8% improvement in overall classification accuracy, positioning it as a promising solution for real-world anomaly detection.

The experiments validate the proposed model's validity and competitiveness on several publicly available benchmark data sets when compared to other conventional methods. Because the suggested model's simple Gaussian model is unable to capture the complex structure of actual data. Recognizing this limitation, attempt to build a new model of the probability graph in the future to achieve this task by making the representation z obey a more sophisticated model. To refine the anomaly scoring module for scalability, integrate it seamlessly with existing systems by optimizing algorithms, ensuring compatibility, and leveraging robust APIs for efficient data processing and feedback mechanisms. Additionally, an adaptive threshold adjustment technique is proposed to enhance flexibility and applicability, reflecting the commitment to advancing the model's capabilities for broader practical use.

Acknowledgment

The author wants to thank the editors and reviewers for their efforts. This research did not receive funding from any public, commercial, or not-for-profit organizations.

6. REFERENCES

Baur, C., Wiestler, B., Albarqouni, S., & Navab, N. (2018) “Deep autoencoding models for unsupervised anomaly segmentation in brain MR images,” arXiv [cs.CV]. Available at: http://arxiv.org/abs/1804.04488.

Lin, S., Clark, R., Birke, R., Schönborn, S., Trigoni, N., & Roberts, S. (2020, May). Anomaly detection for time series using vae-lstm hybrid model. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 4322-4326). IEEE.

Mahmood, S. A., Abid, A. M. and Lafta, S. H. (2021) “Anomaly event detection and localization of video clips using global and local outliers,” Indonesian journal of electrical engineering and computer science, 24(2), p. 1063. doi: 10.11591/ijeecs.v24.i2.pp1063-1073.

Meena, K., Viji, A., Athanesious, J. J., & Vaidehi, V. (2019, March). Detecting abnormal event in traffic scenes using unsupervised deep learning approach. In 2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET) (pp. 355-362). IEEE.

Pang, G., Yan, C., Shen, C., Hengel, A. V. D., & Bai, X. (2020). Self-trained deep ordinal regression for end-to-end video anomaly detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 12173-12182).

Pereira, J. and Silveira, M. (2018) “Unsupervised anomaly detection in energy time series data using variational recurrent autoencoders with attention,” in 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE.

Schlegl, T., Seeböck, P., Waldstein, S. M., Langs, G., & Schmidt-Erfurth, U. (2019). f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. Medical image analysis, 54, 30-44. doi: 10.1016/j.media.2019.01.010

Shi, Feng & Lin, Jingna. (2022). Virtual Machine Resource Allocation Optimization in Cloud Computing Based on Multi objective Genetic Algorithm. Computational Intelligence and Neuroscience. 2022. 1-10. 10.1155/2022/7873131.

Sultani, W., Chen, C., & Shah, M. (2018). Real-world anomaly detection in surveillance videos. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 6479-6488).

Xu, H., Feng, Y., Chen, J., Wang, Z., Qiao, H., Chen, W., Zhao, N., Li, Z., Bu, J., Li, Z., Liu, Y., Zhao, Y., & Pei, D. (2018) “Unsupervised anomaly detection via variational auto-encoder for seasonal KPIs in web applications,” in Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW ’18. New York, New York, USA: ACM Press.

Zhou, F., Wang, L., Li, Z., Zuo, W., & Tan, H. (2020) “Unsupervised learning approach for abnormal event detection in surveillance video by hybrid autoencoder,” Neural processing letters, 52(2), pp. 961–975. doi: 10.1007/s11063-019-10113-w.