INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Internet of Things with Deep Learning Enabled Fraud Detection in Surveillance Video Processing

B Pushpa[1]*, V Narmatha[1], P Anandababu[1], C Senthilkumar[1]

1 Assistant Professor, Department of Computer and Information Science, Annamalai University

OPEN ACCESS

Received: 18-06-2022

Accepted: 19-08-2022

Published: 21-09-2022

Citation: Pushpa B, Narmatha V, Anandababu P, Senthilkumar C (2022) Internet of Things with Deep Learning Enabled Fraud Detection in Surveillance Video Processing. Indian Journal of Science and Technology 15(36): 1769-1778. https://doi.org/10.17485/IJST/v15i36.550

* **Corresponding author**.

pushpasidhu@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2022 Pushpa et al. 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.

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Background: In the present digital era, fraud detection using surveillance video has become a mandatory tool to determine the occurrence of abnormal events in an automated way. Since the traditional visual inspection of surveillance videos for fraud detection is time-consuming and labour-intensive, intelligent fraud detection approaches based on Deep Learning (DL) concepts have been presented in the literature. **Methods:** This paper designs a DL with optimal classification based on fraud detection in a video surveillance system, named the DLOC-FDVS technique. The proposed DLOC-FDVS technique aims to examine the surveillance videos for the existence of frauds (i.e., robbery) in the IoT environment. At the initial stage, IoT enables the data acquisition and frame conversion process to be carried out. For fraud detection, densely connected networks (DenseNet-169) feature extractor and optimal Long Short Term Memory (LSTM) classifiers are applied. Finally, the Grasshopper Optimization Algorithm (GOA) is utilised to alter the LSTM model's hyperparameters. It is frequently employed in a number of industrial settings and achieves suitable answers due to its ease of deployment and excellent precision. **Findings:** The DLOC-FDVS model is experimentally validated using a benchmark anomaly detection dataset from the Kaggle repository, which comprises 211 frames of normal video and 100 frames of fraud video. The experimental results show that the suggested model is an excellent fraud detection tool in the IoT context, achieving maximum precision, recall, F1score, and AUC of 96.56%, 96.56%, 96.56%, and 96.56% respectively. **Novelty:** The use of GOA for hyperparameter adjustment of the LSTM model for fraud detection demonstrates the work's uniqueness. As a result, the DLOC-FDVS model may be used to identify fraud in real-time surveillance recordings.

Keywords: Video surveillance; Deep learning; Video processing; Internet of Things; Fraud detection; Hyperparameter tuning

1 Introduction

Owing to the widespread availability of CC cameras, the number of videos acquired by security cameras has expanded significantly, making human video processing impractical in certain real-time settings.[1,2]. Anomaly detection defines the issue of determining unwanted patterns which deviate from normal behaviour. Anomalies in video comprises of theft, fight, traffic accident, vehicles in pedestrian walkways, etc.,[3]. Although it appears that the recognition of abnormal objects or occurrences is an important component to examine, the context of a video is equally vital for detection.[4,5]. Due to the subjective depiction of anomaly, insufficient quantity of annotated data due to the unusual frequency of anomalous events, low-quality surveillance video, and various intra/inter-class differences, anomaly identification in surveillance films is a time-consuming and critical procedure.[6]. Previous works of anomaly detection in videos are based on features like object trajectory and Spatio-temporal context. They were manually chosen and computed for classification[7,8]. It requires manual annotation of every individual frame. It consumes more time in attaining a video dataset. The recently presented deep learning (DL) models become familiar with the area of computer vision and pattern recognition[9]. In contrast to conventional approaches, the features involved in the DL models are chosen and optimized automatically. The DL models show a promising performance in video anomaly classification with large-scale datasets[10]. It can perform well in complex and harsh environments.

Pustokhina et al.[11] Established an automated DL-enabled anomaly identification model in pedestrian walkways. The purpose of the presented technique is to detect and classify several anomalies which occur in the PW namely jeep, cars, skating, and so on. The DLADT-PW approach contains pre-processed as an initial stage that is executed to remove the noise and enhance the quality of images. Also, the mask region, CNN (Mask-RCNN) with the DenseNet technique was utilized.

Ullah et al.[12] Projected an effective and robust infrastructure for recognizing anomalies in the surveillance. Big Video Data (BVD) utilizing Artificial Intelligence of Things (AIoT). A primary stream contains instant AD, which is useful over energy-limited IoT devices, but the second step is a 2-stream DNN allowing for detailed anomaly analysis, appropriate that utilized as cloud computing services. The authors in[13] exhibited a background deduction technique with a continuous DL framework of MLP-RNN which is appropriate to decide on many objects of distinct sizes by pixel-wise forefront examining infrastructure.

Li et al.[14] Presented a new 2-stream Spatiotemporal structure named as Two-Stream DSTAE. Primarily, the spatial stream removes presence features in which the temporal stream, extracts the motion pattern correspondingly. Huang et al.[15] Presented a TAC-Net for addressing the issues of AD to intelligent video surveillance. The TAC-Net is an unsupervised approach that employs deep contrastive self-supervised learning for capturing the maximum level semantic feature and tackling AD with several self-supervised tasks. During the inference step, the anomaly score was calculated using a large number of task losses and contrastive similarities. The authors of[16] created a novel CNN-RNN combination model to identify violence in real-time. Furthermore, the authors of[17] employed an attention-based residual LSTM model to detect abnormalities in surveillance films.

There are already several DL-based strategies available in the literature for learning the high-level representation of data without the requirement for expert knowledge. Occlusion, illumination change, motion blur, and other environmental modifications are difficult to detect in video surveillance. When detecting anomalies in videos, the widely used strategy is to extract the features at the initial level and then use an individual classification model to detect anomalies at the latter stage. Another technique is to use the DL model to optimise the error during the training data reconstruction and then employ those flaws for abnormality recognition in testing data. Because of the complicated architecture with multiple models or layers, training these systems is challenging. At the same time, most studies have not focused on the process of hyperparameter tuning, which is critical for the successful performance of DL models. This research proposes an intelligent DL-based anomaly detection approach with a hyperparameter tuning strategy to address this issue.

This study has focused on DL with optimum classification-based fraud detection in video surveillance systems, named the DLOC-FDVS technique. The proposed DLOC-FDVS technique majorly intends to categorize the occurrence of events into normal and fraud (i.e., robbery) in the IoT environment. Initially, IoT devices collect data, and then pre-processing occurs, which includes the video-to-frame conversion procedure. Furthermore, the densely connected networks (DenseNet-169) model is used as a feature extractor, and the best long-term memory (LSTM) model is used as a classification model. Finally, the grasshopper optimization algorithm (GOA) is used to alter the LSTM model's hyperparameters. The use of GOA for hyperparameter adjustment of the LSTM model for fraud detection demonstrates the work's uniqueness. The DLOC-FDVS model simulation is performed using a benchmark anomaly detection dataset from the Kaggle repository. In summary, the significant contributions are as follows.

- For fraud detection in video surveillance, an intelligent DLOC-FDVS model comprises of data pre-processing, DenseNet-169 feature extraction, LSTM, and the GOA parameter optimization. As far as we are aware, the DLOC-FDVS model has never been discussed in the literature.
- A unique GAO-based hyperparameter approach for LSTM-based fraud detection is proposed. The parameter optimization of the LSTM model using GOA with cross-validation helps to boost the predictive outcome of the DLOC-FDVS model for

unseen data.

2 Methodology

2.1 Dataset Used

A benchmark anomaly detection dataset from the Kaggle repository is used in this investigation [18]. In this study, we have taken two videos in each class on normal and fraud. Figure 1 depicts some sample images of fraud detection.

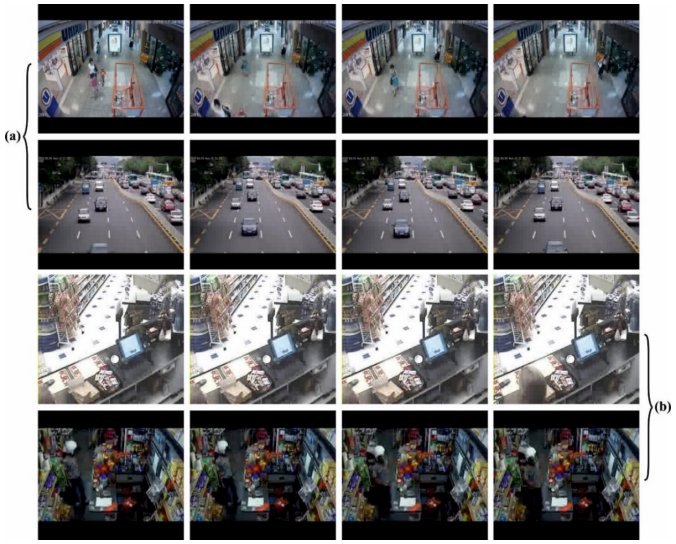

Fig 1. a) Normal Frames b) Fraud Identified Frames

2.2 Materials and Methods

In this work, a new DLOC-FDVS approach was created for the detection and classification of frauds in surveillance videos acquired by IoT devices. The proposed DLOC-FDVS technique primarily converts surveillance videos into a series of frames, after which pre-processing is conducted. Secondly, the DenseNet-169 model generates a useful set of features, and then the LSTM model is utilized to classify the videos into normal or frauds. Furthermore, the GOA is applied to optimally modify the hyperparameters involved in the LSTM model. Figure 2 illustrates the overall process of the DLOC-FDVS technique.

2.3 Data Pre-processing

At the initial stage, the IoT devices are used to collect the surveillance videos. Then, the frame conversion process takes place. Here, frames can be obtained from a video and converted into images. To convert a video frame into an image, frame rate is used. In this work, two sets of videos namely normal and robbery are considered into account. After the frame conversion process, the normal videos hold a set of 211 frames and 100 frames exist under fraud videos.

2.4 DenseNet-169 feature extractor

Once the input frames are pre-processed, they are fed into the DenseNet-169 model to derive feature vectors [19]. DenseNets [20] is a popular deep learning model that reduces the relation between input and output, allowing it to overcome vanishing gradient difficulties. All of the DenseNet layers have a lower feature map size, which is necessary for training the CNN on smaller datasets to reduce the likelihood of overfitting and to ensure that no data is lost during the transfer. Furthermore, all layers are supervised in the loss function and regularised outcome with a shorter connection, resulting in a simpler training technique. The DenseNet is made up of three main components, which are listed below.

Dense Block: The DenseNet is made up of N Dense Blocks. Within each Dense Block, there is a M layer in which all of the layers are linked to every subsequent layer using the feedforward technique. Once xm is represented as the resultant in the mth

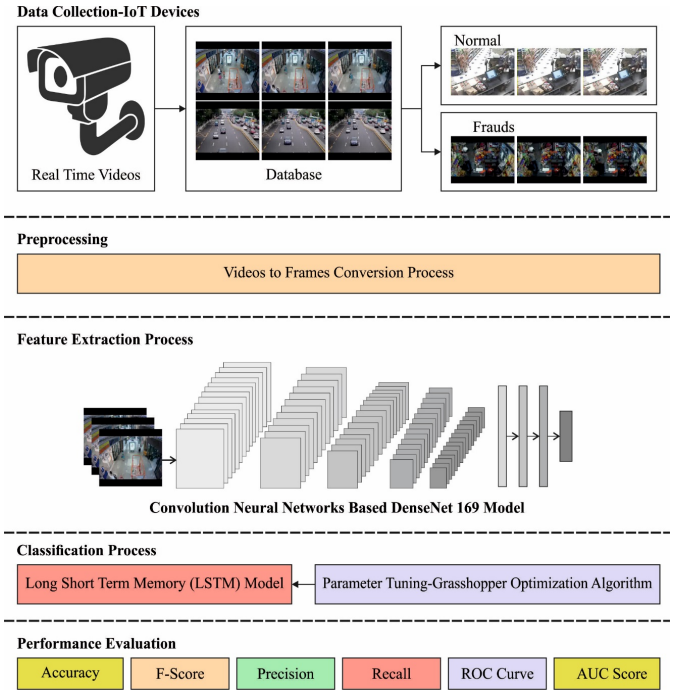

Fig 2. The overall process of the DLOC-FDVS technique

layer afterwards, it can be calculated as:

$$x_m = H_m\left(\left(x_1, x_2, \ldots, x_{m-1}\right]\right) \qquad (1)$$

while Hm defines the composite operation function beneath this layer, and the concatenation function was controlled across all feature levels underneath it. The concatenated feature deal with a combined function that contains Convolutional (3×3), Batch Normalization (BN), and ReLU. The sample of the internal infrastructure of a dense block is passed on to the Transition layer.

Transition Layer: Among all the Dense Blocks, a layer was established for decreasing the spatial dimensional of feature map and it is named as transition layer. It is comprised of Convolutional (1×1) and Average Pooling (2×2).

Growth Rate: The resultant in all concatenation functions in (1) is feature map f. The size of the Mth layer is a ref(m-l)+f0, whereas f0 refers to the number of channels of the original input images. This variable supports control of the count of novel data all the layer holds.

2.5 Fraud Detection and Classification

During the fraud detection and classification process, the LSTM model has been employed. LSTM is a kind of RNN model which manages the memory details of time-series data by the addition of memory cells to the hidden layers. The data is fed among the cells in the hidden layer using a sequence of programmable gates[17]. It maintains the cell state via a gate mechanism that resolves the short-term as well as long-term memory-dependent issues, thereby eliminating vanishing gradient and explosion problems. The fundamental LSTM unit includes three gates in a memory cell. The purpose of the input gate is to trail recent data in the memory cell whereas the output gate function is involved form controlling the data dissemination over the network. The third one (forget gate) is applied to compute whether the data needs to be removed depending upon the previous cell status. The update functions and LSTM output can be defined as follows.

$$F_t = \sigma\left(W_x f X_t + W_h f H_{t-1} + B_f\right) \qquad (2)$$

$$I_t = \sigma\left(W_{xi} X_t + W_{hi} H_{t-1} + B_i\right) \qquad (3)$$

$$C_t = \sigma\left(W_{xc}X_t + W_{hc}H_{t-1} + B_c\right) \quad (4)$$

$$C_t = F_t * C_{t-1} + I_t * C_t \quad (5)$$

$$O_t = \sigma\left(W_{xo}X_t + W_{ho}H_{t-1} + B_0\right) \quad (6)$$

$$H_t = o_t\ tanh\ \left(C_t\right) \quad (7)$$

$$Y_t = \sigma\left(W_{hy}H_t + B_y\right) \quad (8)$$

$$\sigma\left(x\right) = \frac{1}{1 + exp^{-x}} \quad (9)$$

where Xt=input vector; Yt=output vector; It=input gate outputs; Ft=forget gate outputs; Ot=output gate outputs; Ct=finishing state in memory block; Ct=temporary, σ=sigmoid function; Wxf,Wxi,Wxc, and Wxo are input weight matrix; Wh,Whi,Whc/f and Who are recurrent weight matrices; Why is output weight matrix; and Bf,Bi,Bc,B0, and By are the related bias vectors. Figure 3 demonstrates the framework of LSTM.

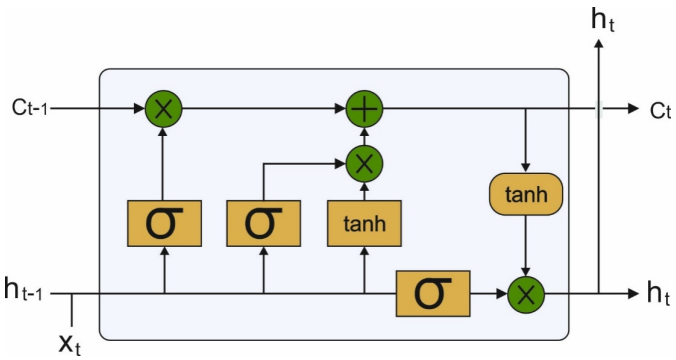

Fig 3. LSTM structure

2.6 Hyperparameter Optimization

The GOA is used to optimally tune the hyperparameters in the LSTM model[21]. GOA is a population-based approach that naturally displayed the performance of a grasshopper swarm. Exploration and exploitation of seeking space are two crucial elements in optimising; the grasshopper provides these two stages under the food hunt with these social interactions[21]. As illustrated in three evolutionary operators are accessible from the place upgrading of individuals from swarms, the social interaction operator (Si), gravity force operator (Gi), and wind advection operator (Ai) as demonstrated in Eq. (10).

$$X_i = S_i + G_i + A_i \quad (10)$$

whereas Xi implies the place of ith grasshopper. All the behaviours are mathematically processed here. The interaction operator was computed in Eq. (11).

$$S_i = \sum_{j=1, j\neq i}^{N} S\left(\left(X_j - X_i\right|\right)\frac{X_j - X_i}{d_{ij}} \quad (11)$$

In which, N represents the number of grasshoppers from the swarm, dij stands for the distance amongst ith and jth grasshoppers, and S defines the function of that power of social force and is computed as in Eq. (12).

$$S(r) = fe^{\frac{-r}{l}} - e^{-r} \tag{12}$$

whereas f and l are 2 constant values that specify correspondingly the intensity of attractions and the attraction length scale, and r refers to the real value[22].

But the gravity operator is not assumed by the authors, and it is considered that the wind direction is continuously near a target. Afterwards, Eq. (13) is developed as given below:

$$X_i^d = c\left(\sum_{\substack{j=1 \\ j\neq i}}^{N} c\frac{ub_d - lb_d}{2} S\left(\left|{}_j^d - X_i^d\right|\right)\frac{X_j - X_i}{d_{ij}}\right) + T_d \tag{13}$$

In which ubd represents the upper bound from the dth dimensional, lbd denotes the lower bound from the dth dimensional. Td refers to the value of dth dimensional from the target (optimum solution establish already), and the co-efficient c decreases the comfort region related to several rounds of operation and is computed below.

$$c = Cmax - l\frac{C\,max\, - C\,min}{L} \tag{14}$$

whereas Cmax signifies the maximal value, Cmin demonstrates the minimal value, l denotes the present iteration, and L defines the maximal count of iterations. It is utilize Cmax=1 and Cmin=0.00001.

Eq. (13) illustrates that the next place of the grasshopper was determined according to their existing place and the place of every other grasshopper and target.

The GOA derived an objective function with an intention to attain maximum classifier outcomes. The GOA, defined a fitness function whose value is a positive integer representing improved outcomes of the candidate solutions. Here, the reduction of classification error can be treated as the fitness function. The optimum solutions include lower error rates and maximum accuracy, as provided below.

$$fitness(x\_i) = Classifier\ Error\ Rate(x\_i) = \frac{number\ of\ misclassified\ events}{Total\ number\ of\ events} * 100 \tag{15}$$

3 Results and Discussion

In this section, the fraud detection and classification outcomes of the DLOC-FDVS model are examined using a dataset from the Kaggle repository. The proposed model is simulated using Python 3.6.5 tool. Also, it is experimented on PC i5-8600k, GeForce 1050Ti 4GB, 16GB RAM, 250GB SSD, and 1TB HDD. The parameter settings are given as follows: learning rate: 0.01, dropout: 0.5, batch size: 5, epoch count: 50, and activation: ReLU.

Figure 4 highlights the confusion matrices produced by the DLOC-FDVS model on test datasets. The figure indicated that the DLOC-FDVS model can identify frauds effectively on videos. On the entire dataset, the DLOC-FDVS model has recognized 208 frames as normal and 87 frames as fraud. Moreover, on 70% of the training dataset, the DLOC-FDVS approach has recognized 133 frames as normal and 70 frames as fraud. Furthermore, on 30% of the testing dataset, the DLOC-FDVS technique has recognized 75 frames as normal and 17 frames as fraud.

emonstrates detailed fraud detection outcomes of the DLOC-FDVS methodology on the input dataset. The result indicated that the DLOC-FDVS model has accomplished effectual detection efficiency on all datasets. For instance, with the entire dataset, the DLOC-FDVS model has reachedaccuracy, precision, recall, F1score, and AUC of 94.86%, 95.39%, 92.79%, 93.94%, and 92.79% respectively. Simultaneously, with 70% of the training dataset, the DLOC-FDVS algorithm has attained accuracy, precision, recall, F1score, and AUC of 95.5%, 94.47%, 91.94%, 92.95%, and 91.94% correspondingly. Concurrently, with 30% of the testing dataset, the DLOC-FDVS methodology has reached accuracy, precision, recall, F1score, and AUC of 97.87%, 96.56%, 96.56%, 96.56%, and 96.56% correspondingly.

Figure 5 demonstrates the precision-recall curve examination of the DLOC-FDVS method on test data. The figure indicated that the DLOC-FDVS model has obtained maximal precision-recall values on the classification of normal and fraud class labels.

A detailed ROC investigation of the DLOC-FDVS model on the test dataset is represented in Figure 6. The results indicated that the DLOC-FDVS model has exhibited its ability in categorizing two different classes as normal and fraud on the test dataset.

Figure 7 illustrates the training/validation accuracy inspection of the DLOC-FDVS method on an applied dataset. The figure conveyed that the DLOC-FDVS approach has offered higher training/validation accuracy in the classification process.

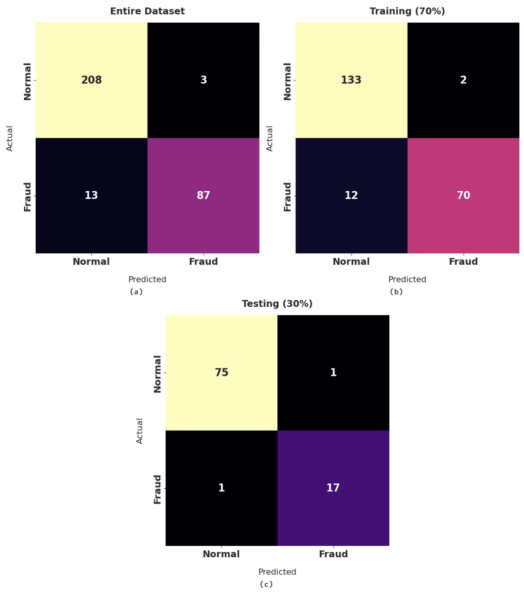

Fig 4. Confusion matrix of DLOC-FDVS technique

Table 1. Result analysis of DLOC-FDVS technique distinct measures

Class Labels	Accuracy	Precision	Recall	F-Score	AUC Score
Entire Dataset					
Normal	94.86	94.12	98.58	96.30	92.79
Fraud	94.86	96.67	87.00	91.58	92.79
Average	**94.86**	**95.39**	**92.79**	**93.94**	**92.79**
Training (70%)					
Normal	93.55	91.72	98.52	95.00	91.94
Fraud	93.55	97.22	85.37	90.91	91.94
Average	**93.55**	**94.47**	**91.94**	**92.95**	**91.94**
Testing (30%)					
Normal	97.87	98.68	98.68	98.68	96.56
Fraud	97.87	94.44	94.44	94.44	96.56
Average	**97.87**	**96.56**	**96.56**	**96.56**	**96.56**

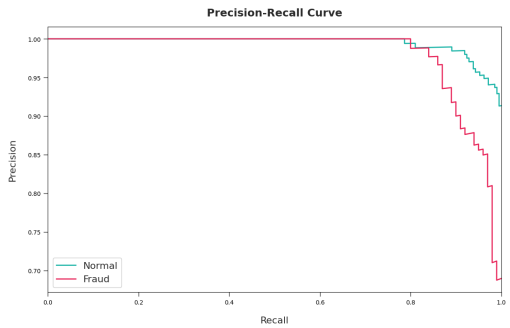

Fig 5. Precision-recall curve analysis of DLOC-FDVS technique

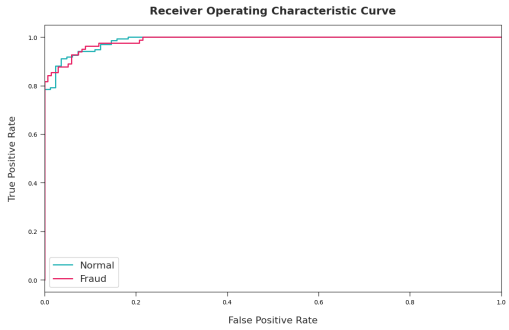

Fig 6. ROC curve analysis of DLOC-FDVS technique

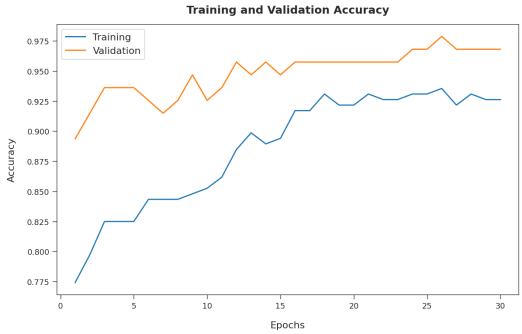

Fig 7. Accuracy graph analysis of DLOC-FDVS technique

Figure 8 reports the overall classification outcomes of the DLOC-FDVS model with other DL models. The experimental results indicated that the MobileNetV2-LSTM model has obtained lower precision, recall, F1score, and AUC of 73.21%, 86.53%, 76.88%, and 87.64% respectively. Followed by, the MobileNetV2-BDLSTM model has offered slightly increased precision, recall, F1score, and AUC of 84.25%, 80.39%, 76.78%, and 87.94% respectively. Moreover, the MobileNetV2-RLSTM model has gained moderate precision, recall, F1score, and AUC of 79.24%, 90.9%, 83.37%, and 95.78% respectively. Though the EARF-ARLSTM model received reasonable outcomes, the DLOC-FDVS model has surpassed the other methods with maximum precision, recall, F1score, and AUC of 96.56%, 96.56%, 96.56%, and 96.56% respectively.

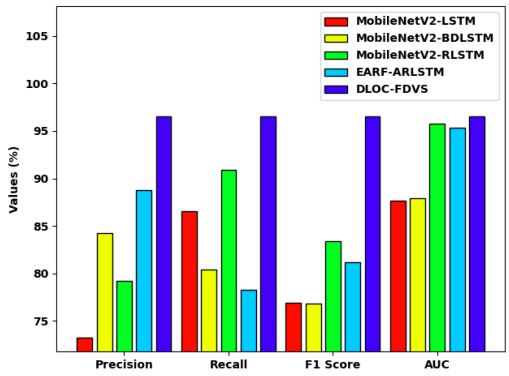

Fig 8. Overall classification result analysis of DLOC-FDVS technique with other DL approaches

To demonstrate the enhanced outcomes of the DLOC-FDVS model, a comparative accuracy examination [16,17] is made in Table 2 and Figure 9. The results indicated that VGG19-LSTM and ResNet152-LSTM methodologies resulted in lower accuracy values of 60.82% and 57.74% respectively. In addition, the ResNet50-LSTM, InceptionV3-LSTM, and ResNet101-

LSTM approaches have gained closer accuracy values of 63.77%, 64.09%, and 65.11% respectively.

Table 2. Comparison study of DLOC-FDVS technique model

Method	$Accu_y$
ResNet50-LSTM	63.77
InceptionV3-LSTM	64.09
VGG19-LSTM	60.82
ResNet101-LSTM	65.11
ResNet152-LSTM	57.74
DLOC-FDVS	97.87

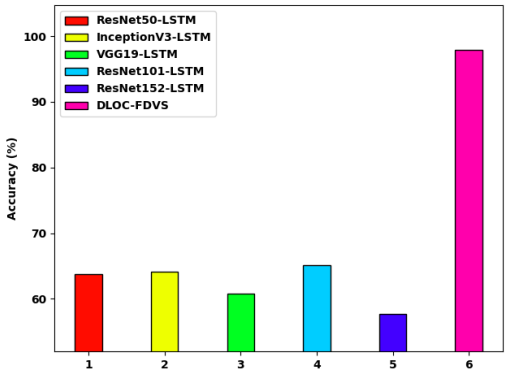

Fig 9. Overall result analysis of DLOC-FDVS technique with other DL approaches

However, the DLOC-FDVS model has accomplished a superior outcome with a maximum accuracy of 97.87%. Therefore, the DLOC-FDVS model has resulted in enhanced performance over the other methods.

4 Conclusion

In this study, a new DLOC-FDVS technique has been devised for the recognition and categorization of the frauds that exist in the surveillance videos captured by IoT devices. The proposed DLOC-FDVS approach first converts surveillance recordings into a series of frames and performs pre-processing. The DenseNet-169 model then creates a meaningful collection of features, and the LSTM model is used to identify the vents as normal or fraudulent. Furthermore, the GOA is exploited to modify the hyperparameters involved in the LSTM model. The experimental validation of the DLOC-FDVS model takes place using a benchmark anomaly detection dataset from the Kaggle repository. The experimentation values indicate the capable performances of the DLOC-FDVS model with maximum precision, recall, F1score, and AUC of 96.56%, 96.56%, 96.56%, and 96.56% respectively. Thus, the DLOC-FDVS model can be employed as an effective tool for fraud detection in real-time surveillance videos. In future, the DLOC-FDVS model can be extended to the design of anomaly detection in pedestrian walkways.

References

1) Shadroo S, Rahmani AM, Rezaee A. Survey On The application of Deep Learning in Internet of Things (IoT). *Telecommunication Systems*. 2022;79(4):601–627.
2) Datta D, Sarkar NI. Deep Learning Frameworks for Internet of Things. In: Internet of Things. Springer International Publishing. 2022;p. 137–161.
3) Ileberi E, Sun Y, Wang Z. A machine learning based credit card fraud detection using the GA algorithm for feature selection. *Journal of Big Data*. 2022;9(1):1–17.
4) Öğrek M, Öğrek E, Bahtiyar Ş. A deep learning method for fraud detection in financial systems: Poster. *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks*. 2019;p. 298–299.
5) Rai AK, Dwivedi RK. Fraud Detection in Credit Card Data using Unsupervised Machine Learning Based Scheme. *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*. 2020;p. 421–426.
6) Soleymanzadeh R, Aljasim M, Qadeer MW, Kashef R. Cyberattack and Fraud Detection Using Ensemble Stacking. *AI*. 2022;3(1):22–36.

7) Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*. 2021;2(3):1–21.
8) Le Q, Miralles-Pechuán L, Sayakkara A, Le-Khac NAA, Scanlon M. Identifying Internet of Things software activities using deep learning-based electromagnetic side-channel analysis. *Forensic Science International: Digital Investigation*. 2021;39:301308–301308.
9) Chen JIZIZ, Lai KLL. Deep Convolution Neural Network Model for Credit-Card Fraud Detection and Alert. *June 2021*. 2021;3(2):101–112.
10) Supriya M, Deepa AJ. Machine learning approach on healthcare big data: a review. *Big Data and Information Analytics*. 2020;5(1):58–75.
11) Pustokhina IV, Pustokhin DA, Vaiyapuri T, Gupta D, Kumar S, Shankar K. An automated deep learning based anomaly detection in pedestrian walkways for vulnerable road users safety. *Safety Science*. 2021;142:105356–105356.
12) Ullah W, Ullah A, Hussain T, Muhammad K, Heidari AA, Ser JD, et al. Artificial Intelligence of Things-assisted two-stream neural network for anomaly detection in surveillance Big Video Data. *Future Generation Computer Systems*. 2022;129:286–297.
13) Murugesan M, Thilagamani S. Efficient anomaly detection in surveillance videos based on multi layer perception recurrent neural network. *Microprocessors and Microsystems*. 2020;79:103303–103303.
14) Murugesan M, Thilagamani S. Efficient anomaly detection in surveillance videos based on multi layer perception recurrent neural network. *Microprocessors and Microsystems*. 2020;79:103303–103303.
15) Huang C, Wu Z, Wen J, Xu Y, Jiang Q, Wang Y. Abnormal Event Detection Using Deep Contrastive Learning for Intelligent Video Surveillance System. *IEEE Transactions on Industrial Informatics*. 2022;18(8):5171–5179.
16) Vosta S, Yow KCC. A CNN-RNN Combined Structure for Real-World Violence Detection in Surveillance Cameras. *Applied Sciences*. 2022;12(3):1021–1021.
17) Ullah W, Ullah A, Hussain T, Khan ZA, Baik SW. An Efficient Anomaly Recognition Framework Using an Attention Residual LSTM in Surveillance Videos. *Sensors*. 2021;21(8):2811–2811.
18) Anomaly-Detection-Dataset-UCF. . Available from: https://www.kaggle.com/minhajuddinmeraj/anomalydetectiondatasetucf.
19) Ghatwary N, Ye X, Zolgharni M. Esophageal Abnormality Detection Using DenseNet Based Faster R-CNN With Gabor Features. *IEEE Access*. 2019;7:84374–84385.
20) Islam MZ, Islam MM, Asraf A. A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. *Informatics in Medicine Unlocked*. 2020;20:100412–100412.
21) Mirjalili SZ, Mirjalili SZ, Saremi S, Faris H, Aljarah I. Grasshopper optimization algorithm for multi-objective optimization problems. *Applied Intelligence*. 2018;48(4):805–820.
22) Ewees AA, Elaziz MA, Houssein EH. Improved grasshopper optimization algorithm using opposition-based learning. *Expert Systems with Applications*. 2018;112:156–172.