An adaptive combination algorithm based on deep learning and genetic algorithm for anomalous events detection

Zainab K. Abbas, Ayad A. Al-Ani
Department of Information and Communication Engineering, College of Information Engineering,
Al-Nahrain University, Baghdad, Iraq

Article Info

Article history:

Received Oct 7, 2022
Revised Mar 24, 2023
Accepted Apr 2, 2023

Keywords:

BiLSTM
Deep learning
Features selection
Genetic algorithm
Video surveillance

ABSTRACT

One of the most widely used human behavior detection methods is anomaly detection, which this article covers. Ensuring a person's safety is a crucial task in every community today due to the ever-increasing actions that can be dangerous, from planned crime to harm from an accident. Classic closed-circuit television is insufficient since a person must always be awake and available to monitor the cameras, which is costly. Also, someone's attention tends to decrease after a certain period of time. Due to these reasons, a surveillance system that is automated and able to detect unusual activities in real-time and give sufferers prompt aid is necessary. It should be noted that the identification process must be completed swiftly and correctly. In this paper, we employ a model based on mixes the machine learning (ML) model, namely genetic algorithms with deep learning (DL). In this study's experimentation, the UCF-Crime dataset was employed. The detection accuracy on the testing sample dataset was equal to 89.90%, while the area under the curve (AUC) was equal to 94.58%. The developed models have demonstrated reliability and the ability to achieve the greatest accuracy when compared to models that have already been designed.

Corresponding Author:

Zainab K. Abbas
Department of Information and Communication Engineering, College of Information Engineering
Al-Nahrain University
Baghdad, Iraq
Email: zainabkudair@gmail.com

1. INTRODUCTION

Malls, roads, smart cities, hospitals, markets, banks, and educational institutions are just a few examples of places where video surveillance systems (VSS) are used to improve public security [1]. Typically, the effectiveness and quickness of video anomaly detection are the major goals of security systems [2]. Several security cameras have recently been installed in various locations across the world for public security [3]. Continually, these cameras generate huge volumes of video data [3]. Many human resources are needed for anomalous case detection and real-time video analysis. Moreover, mistakes could happen as a result of people losing focus over time [1]. Human surveillance of abnormalities is ineffective for all the aforementioned reasons, necessitating the use of autonomous anomaly detection approaches based on artificial intelligence (AI) techniques [4].

In the literature, many ways have been used to explain anomalous behavior, such as "the occurrence of variance in regular patterns" [3]. Traffic security, automated intelligent visual monitoring, and crime prevention are some uses for abnormal action recognition in security cameras [5]. Video anomaly detection was traditionally believed to be a one-class classification assignment due to the absence of real irregular events [6]-[8] i.e., The classifier is trained on regular videos, and a video is labeled as an abnormality when strange patterns are present in the test data [5]. Hence, numerous regular actions could deviate from regular happenings, resulting in false alarms [1], [9]. Many studies have been conducted on this issue, and each of these methods was applied in a specific situation. Sultani *et al*. [10] suggest a framework that can detect unusual attitudes and inform the user about the kind of behavior. Another study by Shreyas *et al*. [11] advises reducing the file size of each video before sending it to the system for detecting behavior. Anala *et al*. [12], Hao *et al*. [13], and Dubey *et al*. [14] address the detection of anomaly events as a regression problem. Another research presented a lightweight convolutional neural network (CNN) by Ullah *et al*. [5]. Ullah *et al*. [3] presented an anomaly detection system depend on merging the ResNet50 with multilayer bidirectional long short term memory (BiLSTM). A weakly supervised anomaly detection model based on video-level labels to train the model was presented by Zaheer *et al*. [15]. Another technique that used a weakly supervised learning model to handle anomalous detection and classification was offered by Majhi *et al.* [16]. Wu *et al*. [17] offered a dual-branch network utilizing the concepts of multi-detail (spatial and temporal) as the input. Cao *et al*. [18] proposed taking into account the spatial-temporal relationships between video parts to detect anomalies in the video. Abbas and Al-Ani [19] suggest compressing each video using high-efficiency video coding (H265) before feeding them into anomaly detection systems. An algorithm for decreasing the features maps size has been suggested by Abbas and Al-Ani [19]. Abbas and Al-Ani [20] suggest using principal component analysis for reducing the dimensionality of the data features.

Although the term "anomaly" is commonly used in literature, there isn't yet agreement on what it means [9]. Most existing techniques produce a lot of false alarms. However, even if these techniques perform well with small datasets, their effectiveness in real-world applications is limited. To overcome these problems, we suggested decreasing the features maps dimension using the features selection model namely genetic algorithm (GA) [21], [22], to improve the model performance and reduce the complexity of the model. Then feed our classifier model, which is a BiLSTM by the new features. We used a weakly supervised technique depending on spatio-temporal features and BiLSTM to train our classifier model. The BiLSTMs have demonstrated that they're very useful in situations when the context of the input is required [23].

2. METHOD

For the suggested model in this research, three stages were used: feature extraction using Resnet50, feature selection using GA, and BiLSTM for anomaly detection. The major goal of this research is to evaluate hybrid machine learning (ML) and deep learning (DL) algorithm on the UCF-Crime dataset for the first time. Figure 1 illustrates the proposed framework.

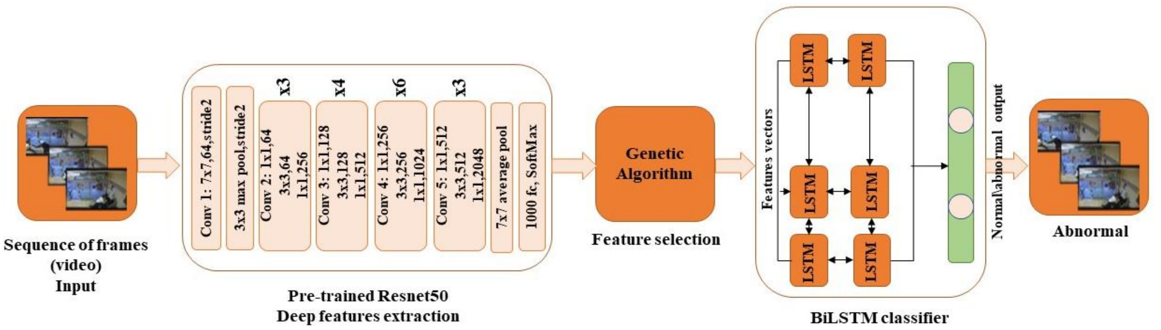


Figure 1. The suggested framework

2.1. Input database UCF-Crime

UCF-Crime is one of the most recent datasets used in the anomaly detection field used in this research [24], [25]. This dataset contains 13 diverse categories of an anomaly such as abuse, accidents, and fights, in addition to the normal category. In total it includes 1,900 videos divided into 800 normal videos and 810 anomaly videos for the training, whereas the testing had 150 normal videos and 140 anomaly videos [10], [26]. This collection includes over 129 hours of films at a resolution of 320x240 and 13 million frames [10], [24], [26]. We selected this dataset because of its vast variety of abnormal event categories and the important impact that its abnormalities have on community security [2], [24]. For this research experimentation, we selected videos with lengths less than or equal to two minutes. As a result of this condition, 1,324 videos were used in this research experimentation divided as: 1,116 videos for the training stage (in a ratio of 90:10 for training, and validation respectively) and 208 videos for the testing stage.

2.2. Machine learning (ML) and deep learning (DL)

Unstructured data was successfully processed by a DL, which is a subset of the ML. DL gradually recognizes and comprehends the information's various facets. It got more widespread as data availability expanded and powerful computers were created. The input is separated into layers by the DL method, and each level can extract features and transmit them to the layer above. The first layers collect the fundamental data, which is coupled with the explanations offered by the next layers. The effectiveness of DL classifiers greatly improves as the amount of information rises when compared to standard learning models. Several designs, such as recurrent neural networks (RNN), pre-trained networks, CNN, and others, may be utilized for DL [1], [2], [27].

Image processing typically makes use of CNN. CNN needs less setup than other categorization techniques. When it uses the appropriate filters to discover spatial and temporal relationships from an image. Several CNN kinds are ZFNet, VGGNet, GoogleNet, LeNet, AlexNet, and ResNet [2], [27]. RNN is better at understanding sequence information than CNN, despite the latter's strong understanding of spatial information. This is because RNN employs state variables to store historical information and combine it with present input to forecast present outcomes [2], [28]. RNNs can estimate time series because they memorize previous inputs. One illustration of an RNN is the LSTM [2], [27].

2.2.1. Features extraction

The 50-layer deep convolutional neural network ResNet50, which contains 23.5 million learnable parameters and returns 1,000 features for each frame, was used in this research to extract features [2], [3]. ResNet50 requires input with a resolution of 224 by 224 pixels. As a result, to match the input dimensions, the longest edges of a video were scaled and cut utilizing a center crop in this research.

2.2.2. Features selection-genetic algorithms (GA)

GA are reliable, adaptable heuristics that can be used to solve issues in any domain with small context-based adjustments. In this research, we suggest reducing the number of features using GA algorithms. Apart from its simplicity and ability to optimize, GA is used for this reason. Figure 2 illustrates the GA steps [29]. The model parameters are as follows: the maximum number of generations was equal to 20, the crossover rate and the mutation rates were 0.8 and 0.3, respectively, and the number of chromosomes was equal to 10 [30].

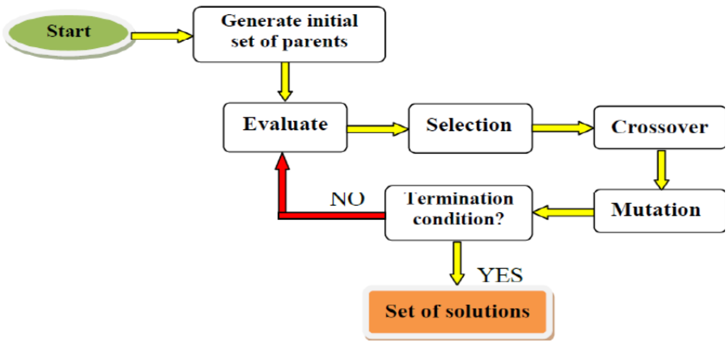


Figure 2. Genetic algorithm steps

2.2.3. Classifier

In this research, the BiLSTM model was utilized as a classifier model to detect anomalies in events, because it looks suitable for long sequential information such as videos [2]. Unlike the previous works that used the features extracted from the features' extraction stage. In this research, we utilized the features got from the features' selection stage as input to the classifier model to detect the anomaly events in the video. The confusion matrix, area under curve (AUC), and the detection accuracy were used as performance metrics in this research.

3. RESULTS AND DISCUSSION

The computer codes used in this research were implemented using the MATLAB software environment (version 2021a). The computer used for the job had the following specifications: NVIDIA GeForce MX450 graphics processing unit, Windows 10, 64-bit operating system, Intel Core i7 processor, 16 GB of RAM, and 1 TB SSD hard drive. Each stage's outcomes were as shown in:

3.1. Input dataset UCF-Crime

The UCF-Crime dataset was used in the experiments for this research. The 13 anomalous classes were used along with the regular class. As the anomalous classification in this task was performed at the video level and the duration of the video had no significant effect on how well the outlier detection operated, we only chose videos that were no longer than or equal to two minutes. We have 1,324 videos in total, which were split into the following categories based on this stipulation: 208 videos were used for testing, whereas 1,116 videos were used for training (90% for training, 10% for validation).

3.2. Machine learning (ML) and deep learning (DL)

3.2.1. Features extraction

In this research, a pre-trained ResNet50 model was employed for extracting the features. Rather than employing the SoftMax layer, the features were taken from the fc1,000 layer. This means that every frame in the movie will, after this phase, be captured by 1000 aspects, meaning that for a movie with (x) frames, the number of features will be (1,000 * x), and this is referred to as huge data. To save time for training while simultaneously improving classification accuracy, we suggest in this research passing the retrieved features to the GA model before classifying them for the first time for the UCF-Crime dataset.

3.2.2. Features selection-genetic algorithms (GA)

In this step, GA is employed to compute the new features map, which is input into the binary classifier rather than the retrieved features from ResNet50. Figures 3 and 4 illustrate the variation in the number of feature vectors between the dataset before and after applying GA for both the training and testing dataset. The red curve indicates the number of feature vectors for the video before applying GA, while the blue curve indicates the number of feature vectors after applying GA.

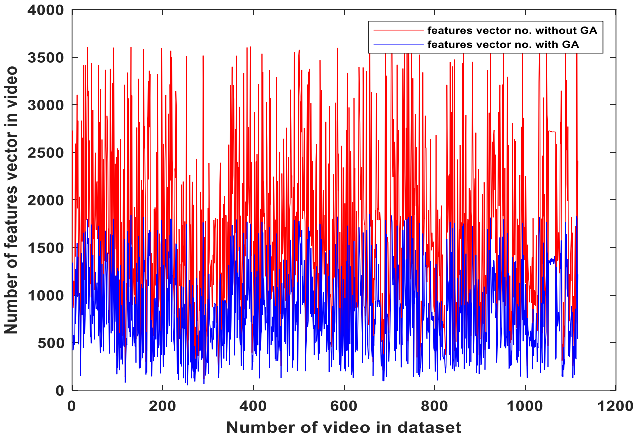


Figure 3. Video’s features vectors number without and with genetic algorithm for training data

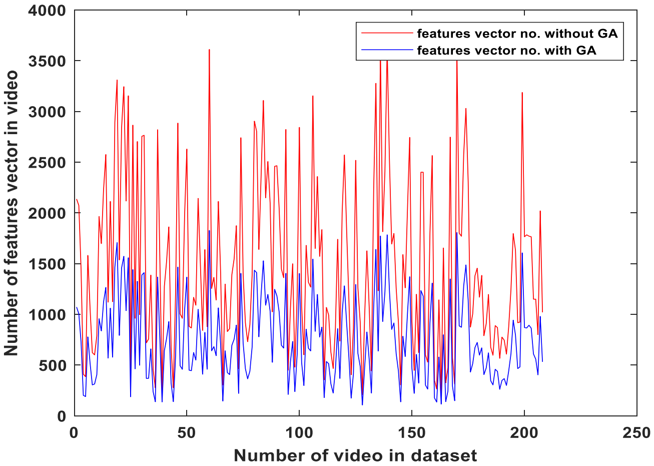


Figure 4. Video’s features vectors number without and with genetic algorithm for testing data

3.2.3. Classifier

A decision classifier application is required when all of the features are complete. This research utilized the BiLSTM method as a classifier. The Adam optimizer was employed in this study's training, with the mini-batch size equal to 8, the initial learning rate equal to 1e-4, the maximum number of epochs equal to 60, and 80 hidden layer nodes with the L2Regularization set to 0.5, and dropout set to 0.7 to reduce the model overfitting. The model's variables were chosen by trial and error. The effectiveness of the suggested model was assessed using the receiver operating characteristics (ROC) and the AUC following the earlier study. Also, we evaluated the classifier model's accuracy using the (1) [31].

$$accuracy = \frac{number\ of\ correct\ symbols}{total\ number\ of\ symbols}\ x\ 100\% \qquad (1)$$

The experiments show the efficacy of the suggested model, which more accurately recognizes unusual behavior than current works. The ROC curve of the presented design is depicted in Figure 5, where the blue curve indicates ROC with GA with AUC score equal to 94.58% while the red curve indicates ROC without GA with AUC score equal to 92.47%. Also, the time taken to train the model on features subject to GA approximately was equal to half the time for training on features not subject to GA.

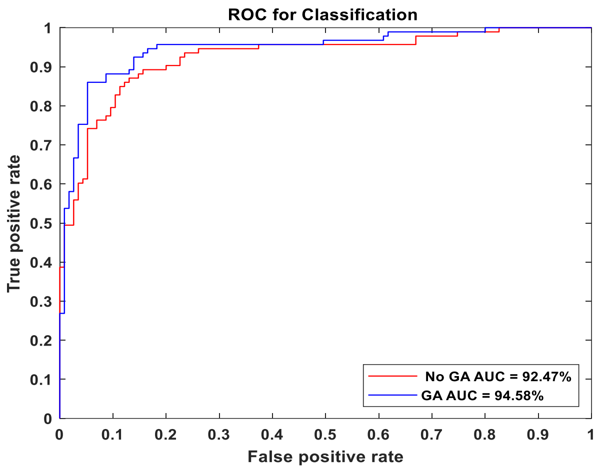
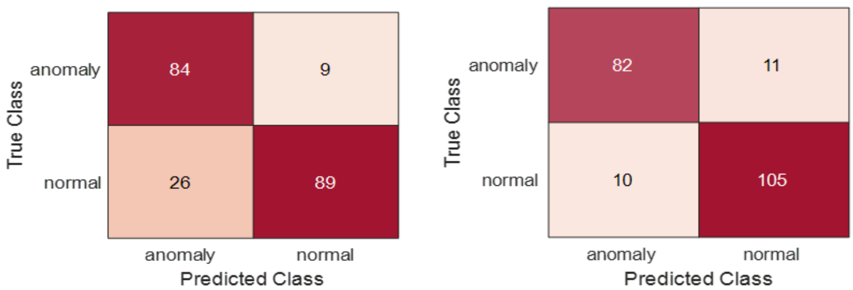


Figure 5. The ROC Curve of the model without and with GA

The confusion matrix of the classifier model can be shown in Figure 6 for both cases without and with GA. Where the false positive (FP) which means normal classified as an anomaly was equal to 10, false negative (FN) which means an anomaly classified as normal was equal to 11, true positive (TP) which means an anomaly classified as anomaly was equal to 82 and true negative (TN) which means normal classified as normal was equal to 105. The accuracy of the model has been calculated, and it is found that it was equal to 83.17% without GA, while with GA it was equal to 89.90%



Detection accuracy = 83.17% Detection accuracy = 89.90%.

Figure 6. The classifier model confusion matrix without and with GA

After describing the suggested model details, it is necessary to compare the suggested model with the earlier research. The AUC scores have been compared with the previous works in Table 1, and it is obvious that the suggested model earned the highest AUC of 94. 58%.

Table 1. AUC score comparison between the proposed work and the previous works

Method	AUC %
Waqas and colleagues, 2019 [10]	75.41
Anala and colleagues, 2019 [12]	85
Shreyas and colleagues, 2020 [11]	79.8
Hao and colleagues, 2020 [13]	81.22
Dubey and colleagues, 2021 [14]	81.91
Ullah and colleagues, 2021 [5]	78.43
Ullah and colleagues, 2021 [3]	85.53
Zaheer and colleagues, 2021 [15]	78.27
Majhi and colleagues, 2021 [16]	82.12
Wu and colleagues, 2021[17]	87.65
Cao and colleagues, 2022 [18]	83.14
Abbas and Al-Ani [2]	90.16
Abbas and Al-Ani [19]	93.61
Abbas and Al-Ani [20]	94.21
Our adaptive algorithm	94.58

6. CONCLUSION

The suggested system for anomaly detection is based on merging ML with DL in which the Resnet50 was used for extracting the features map for each video then GA was used for the first time in such work to remove particular features to generate the new feature map which will feed the BiLSTM for anomaly detection. The proposed model has been shown to have greater accuracy when compared to earlier works. The experimental results demonstrated that the AUC value of the UCF-Crime dataset improved by up to 94.58%. A measurement of the classifier's detection accuracy revealed that it was 89.90% accurate. And this demonstrates the importance of our suggestion to increase abnormal event detection accuracy by minimizing the number of false alarms for both negative and positive. To improve the indicator of the accuracy of the proposed model, future works will test other dimensionality reduction methods, feature selection methods, and feature extraction methods.

REFERENCES

[1] Z. K. Abbas and A. A. Al-Ani, "A comprehensive review for video anomaly detection on videos," *Proceedings of the 2nd 2022 International Conference on Computer Science and Software Engineering, CSASE 2022*. pp. 30–35, 2022, doi: 10.1109/CSASE51777.2022.9759598.

[2] Z. K. Abbas and A. A. Al-Ani, "Anomaly detection in surveillance videos based on H265 and deep learning," *International Journal of Advanced Technology and Engineering Exploration*, vol. 9, no. 92, pp. 910–922, 2022, doi: 10.19101/IJATEE.2021.875907.

[3] W. Ullah, A. Ullah, I. U. Haq, K. Muhammad, M. Sajjad, and S. W. Baik, "CNN features with bi-directional LSTM for real-time anomaly detection in surveillance networks," *Multimedia Tools and Applications*, vol. 80, no. 11, pp. 16979–16995, 2021, doi: 10.1007/s11042-020-09406-3.

[4] H. İ. Öztürk and A. B. Can, "ADNet: Temporal anomaly detection in surveillance videos," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12664 LNCS, pp. 88–101, 2021, doi: 10.1007/978-3-030-68799-1_7.

[5] W. Ullah, A. Ullah, T. Hussain, Z. A. Khan, and S. W. Baik, "An efficient anomaly recognition framework using an attention residual lstm in surveillance videos," *Sensors*, vol. 21, no. 8, 2021, doi: 10.3390/s21082811.

[6] R. Morais, V. Le, T. Tran, B. Saha, M. Mansour, and S. Venkatesh, "Learning regularity in skeleton trajectories for anomaly detection in videos," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2019-June, pp. 11988–11996, 2019, doi: 10.1109/CVPR.2019.01227.

[7] Y. Fan, G. Wen, D. Li, S. Qiu, M. D. Levine, and F. Xiao, "Video anomaly detection and localization via Gaussian mixture fully convolutional variational autoencoder," *Computer Vision and Image Understanding*, vol. 195, 2020, doi: 10.1016/j.cviu.2020.102920.

[8] F. P. dos Santos, L. S. F. Ribeiro, and M. A. Ponti, "Generalization of feature embeddings transferred from different video anomaly detection domains," *Journal of Visual Communication and Image Representation*, vol. 60, no. Icmc, pp. 407–416, 2019, doi: 10.1016/j.jvcir.2019.02.035.

[9] J. Ren, F. Xia, Y. Liu, and I. Lee, "Deep video anomaly detection: opportunities and challenges," *IEEE International Conference on Data Mining Workshops, ICDMW*, vol. 2021-December, pp. 959–966, 2021, doi: 10.1109/ICDMW53433.2021.00125.

[10] W. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," *arXiv* preprints, Jan. 2018, [Online]. Available: http://arxiv.org/abs/1801.04264.

[11] D. G. Shreyas, S. Raksha, and B. G. Prasad, "Implementation of an anomalous human activity recognition system," *SN Computer Science*, vol. 1, no. 3, 2020, doi: 10.1007/s42979-020-00169-0.

[12] M. R. Anala, M. Makker, and A. Ashok, "Anomaly detection in surveillance videos," *Proceedings - 26th IEEE International Conference on High Performance Computing Workshops, HiPCW 2019*, pp. 93–98, 2019, doi: 10.1109/HiPCW.2019.00031.

[13] W. Hao *et al.*, "Anomaly event detection in security surveillance using two-stream based model," *Security and Communication Networks*, vol. 2020, 2020, doi: 10.1155/2020/8876056.

[14] S. Dubey, A. Boragule, J. Gwak, and M. Jeon, "Anomalous event recognition in videos based on joint learning of motion and appearance with multiple ranking measures," *Applied Sciences (Switzerland)*, vol. 11, no. 3, pp. 1–21, 2021, doi: 10.3390/app11031344.

[15] M. Z. Zaheer, J. Lee, M. Astrid, A. Mahmood, and S.-I. Lee, "Cleaning label noise with clusters for minimally supervised anomaly detection," *arXiv*, pp. 3–6, 2021, doi: 10.48550/arXiv.2104.14770.

[16] S. Majhi, S. Das, F. Bremond, R. Dash, and P. K. Sa, "Weakly-supervised joint anomaly detection and classification," *Proceedings - 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2021*, 2021, doi: 10.1109/FG52635.2021.9667006.

[17] J. Wu *et al.*, "Weakly-supervised spatio-temporal anomaly detection in surveillance video," *IJCAI International Joint Conference on Artificial Intelligence*, pp. 1172–1178, 2021, doi: 10.24963/ijcai.2021/162.

[18] C. Cao, X. Zhang, S. Zhang, P. Wang, and Y. Zhang, "Adaptive graph convolutional networks for weakly supervised anomaly detection in videos," *IEEE Signal Processing Letters*, vol. 29, pp. 2497–2501, 2022, doi: 10.1109/LSP.2022.3226411.

[19] Z. K. Abbas and A. A. Al-Ani, "Detection of anomalous events based on deep learning-bilstm," *Iraqi Journal of Information and Communication Technology*, vol. 5, no. 3. Inpress, pp. 34–42, 2022, doi: 10.31987/ijict.5.3.207.

[20] Z. K. Abbas and A. A. Al-Ani, "An adaptive algorithm based on principal component analysis-deep learning for anomalous events detection," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 1, pp. 421–430, 2023, doi: 10.11591/ijeecs.v29.i1.pp421-430.

[21] A. Hassanat, K. Almohammadi, E. Alkafaween, E. Abunawas, A. Hammouri, and V. B. S. Prasath, "Choosing mutation and crossover ratios for genetic algorithms-a review with a new dynamic approach," *Information (Switzerland)*, vol. 10, no. 12, 2019, doi: 10.3390/info10120390.

[22] D. A. Coley, *An introduction to genetic algorithms for scientists and engineers*. World Scientific Publishing Co. Pte. Ltd., 1999, doi: 10.1142/3904.

[23] A. A. Sharfuddin, M. N. Tihami, and M. S. Islam, "A deep recurrent neural network with BiLSTM model for sentiment classification," *2018 International Conference on Bangla Speech and Language Processing, ICBSLP 2018*, 2018, doi: 10.1109/ICBSLP.2018.8554396.

[24] R. Maqsood, U. I. Bajwa, G. Saleem, R. H. Raza, and M. W. Anwar, "Anomaly recognition from surveillance videos using 3D convolution neural network," *Multimedia Tools and Applications*, vol. 80, no. 12, pp. 18693–18716, 2021, doi: 10.1007/s11042-021-10570-3.

[25] "UCF-Crime database." https://www.dropbox.com/sh/75v5ehq4cdg5g5g/AABvnJSwZI7zXb8_myBA0CLHa?dl=0. (accessed Jan. 20, 2021)

[26] P. Kumari, A. K. Bedi, and M. Saini, "Multimedia datasets for anomaly detection: a review," *arXiv preprints*, 2021, doi: 10.48550/arXiv.2112.05410.

[27] A. Mathew, P. Amudha, and S. Sivakumari, "Deep learning techniques: an overview," *Advances in Intelligent Systems and Computing*, vol. 1141, no. August 2020, pp. 599–608, 2021, doi: 10.1007/978-981-15-3383-9_54.

[28] J. M. Czum, "Dive into deep learning," *Journal of the American College of Radiology*, vol. 17, no. 5, pp. 637–638, 2020, doi: 10.1016/j.jacr.2020.02.005.

[29] M. Afzal and A. S. M. Arif, "Genetic algorithm for outlier detection," *International Journal of Computer Science and Information Technology*, vol. 7, no. 2, pp. 833–835, 2016.

[30] J. Too and A. R. Abdullah, "A new and fast rival genetic algorithm for feature selection," *Journal of Supercomputing*, vol. 77, no. 3, pp. 2844–2874, 2021, doi: 10.1007/s11227-020-03378-9.

[31] S. H. Abd, I. A. Hashim, and A. S. A. Jalal, "Automatic deception detection system based on hybrid feature extraction techniques," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 1, pp. 381–393, 2022, doi: 10.11591/ijeecs.v26.i1.pp381-393.

BIOGRAPHIES OF AUTHORS

Zainab K. Abbas she was born in Baghdad/Iraq, in 1993. She gained B.Sc. and M.Sc., from Dept. of Control and system engineering/University of technology, in 2015 and 2018, respectively. She is currently a Ph.D. student in the Department of Information and Communication Engineering, College of Information Engineering, Al-Nahrain University, Baghdad, Iraq. Her works studies are focused on digital processing with artificial intelligence. She can be contacted at email: zainabkudair@gmail.com.

Ayad A. Al-Ani he was born in Baghdad/Iraq, in 1961. He gained the B.Sc., M.Sc., and Ph.D., from Dept. of Physics/College of Science/Baghdad Univ., in 1983, 1990, and 1995, respectively. Previously, he was Univ. Vice President for Administrative affairs/Al-Nahrain Univ., Deputy Dean for Post Graduate Studies and Scientific Research/Baghdad Univ, and Head of Space and Astronomy Dept./College of Science/Baghdad Univ. Since 2006, he is a Professor of Digital Image Processing in Department of Information and Communication Engineering/College of Information Engineering/Al-Nahrain Univ. He published 74 papers and 4 books. He can be contacted at email: ayad.a@nahrainuniv.edu.iq.