Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 3 (2023)

Utilizing Neural Network Models for Detecting Anti-Social Activities in Surveillance Monitoring within the Internet of Things (IoT)

Tuannurisan Sukitjanan[1], Sakesan Chana[2], Thaphat Chaichochok[3] ,Surachet Channgam[4] and Keattisak Chankaew*

[1,2,3] Department of Information Technology and IoT, Faculty of Industrial Technology, Songkhla Rajabhat University, Thailand. E-mail: tuannurisan.su@skru.ac.th[1] , sakesan.ch@skru.ac.th[2], thaphat.ch@skru.ac.th[3]
[4] Program in Digital Business Technology, Faculty of Management Science, Chandrakasem Rajabhat University, Thailand. E-mail: surachet.c@chandra.ac.th
*Coresponding Author.

Abstract
The use of sensors and electronic devices for remote monitoring has become increasingly prevalent across various sectors, including traffic management, forestry, military operations, commerce, and medical applications. These monitoring systems are designed to detect and respond to abnormalities in real-time, providing valuable insights and enhancing safety and security. However, in the realm of surveillance videos, the computational complexity of computer vision-based video processing systems has posed a significant challenge. To address this, recent research has developed a novel approach known as the Slow-Fast Convolutional Neural Network (SF-CNN). This innovative CNN architecture is designed to automatically learn from video frames and effectively classify anomalous behavior in surveillance videos. What sets SF-CNN apart is its adaptability, as it can learn at varying speeds depending on the frame rate, enabling it to capture both spatial and temporal information. By identifying humans, vehicles, and animals and handling normal and aberrant activities in different scenarios, SF-CNN offers a robust solution to address the limitations associated with detecting anomalous motions end-to-end. In testing on benchmark datasets, this proposed method achieved a remarkable accuracy rate of 99.6%, surpassing the performance of previous approaches.

Keywords: Sensors, Artifical Neural Network, Video and Image Processing, Internet of Thing (IoT)

INTRODUCTION

The proliferation of surveillance systems, driven by advancements in technology, has transformed the way we monitor public spaces, critical infrastructure, and private premises. In an era characterized by an increasingly interconnected world, the Internet of Things (IoT) has emerged as a transformative paradigm, enhancing the capabilities of surveillance systems through the integration of smart sensors and devices. This integration not only facilitates real-time data collection but also opens doors to more sophisticated data analysis techniques. This research delves into the intersection of IoT and artificial intelligence, particularly neural network models, to address a critical societal concern: the identification of anti-social activities in surveillance monitoring. Surveillance monitoring plays an indispensable role in modern society, encompassing a wide array of applications, from security and crime prevention to traffic management and environmental monitoring. However, the sheer volume of data generated by surveillance systems has outpaced the capacity of human operators to effectively analyze and respond to incidents in real time. Consequently, there exists a compelling need to harness the power of advanced technologies, such as IoT and neural network models, to automate the detection of anti-social activities. These activities may include criminal behavior, vandalism, harassment, or any actions that

disrupt public safety and harmony. The IoT revolution has ushered in an era where everyday objects and devices are connected to the internet, enabling them to collect and transmit data autonomously. In the context of surveillance monitoring, this translates to an extensive network of cameras, sensors, and data sources, all providing a constant stream of information. Leveraging IoT technologies, this research aims to enhance the efficiency and accuracy of surveillance systems by automating the identification of anti-social activities. This not only reduces the burden on human operators but also enables proactive responses to potential threats or incidents.

Neural network models, particularly deep learning algorithms, have demonstrated remarkable success in various fields, including computer vision and pattern recognition. Their ability to learn complex patterns and features from vast datasets makes them a natural choice for the task of identifying anti-social activities within surveillance footage. By combining the capabilities of IoT-enabled devices with neural network models, this research seeks to create a robust framework that can distinguish between normal and anti-social behavior, providing actionable insights for surveillance operators and law enforcement agencies.[1],[2]. Surveillance systems play a pivotal role in ensuring security and safety in a multitude of sectors, including healthcare, forests, research centers, aerospace, transportation hubs, malls, and expansive buildings. Employing advanced CCTV cameras, these systems continuously record activities, generating video footage that is subsequently processed to identify abnormal behavior using object detection and recognition methods. This classification categorizes objects' activities as either normal or abnormal, facilitating proactive security measures.

Businesses today find myriad reasons to invest in video surveillance systems, with ten prominent motivations at the forefront. These include conflict resolution, productivity enhancement, theft reduction, customer experience improvement, real-time monitoring capabilities, safety enhancement, secure digital data storage, evidence generation, access control, and substantial cost savings. While the core surveillance monitoring process remains similar, the configuration and capacity of cameras and surveillance systems are tailored to suit specific applications. Surveillance applications span a wide spectrum, ranging from offices and roadways to official buildings and residential areas. One persistent challenge in the realm of surveillance is the delay in detecting abnormal activities, often occurring after the event has taken place. Addressing this limitation requires the development of methodologies that can automatically or manually preempt strange activities, making early detection paramount. To proactively control and prevent abnormal movements, location data combined with a comprehensive understanding of the geographic region is indispensable. This combination not only aids in early abnormal activity detection but also augments security within the specific surveillance area. To enhance security further, this research employs a deep learning-based convolutional neural network (CNN). Deep learning, with its capacity to mimic human behavior and process a range of data types, has revolutionized various fields, from recognizing street signs to distinguishing objects. Its accuracy, grounded in multi-layered neural networks and extensive labeled data, has propelled it to the forefront of technology, with applications extending to robotics, autonomous vehicles, natural language processing, and image recognition. Leveraging enormous labeled data and significant computational power, deep learning continues to redefine the landscape of technological progress, meeting and exceeding user expectations in various domains. This scientific breakthrough, powered by GPUs' parallel processing capabilities, has significantly cut training time and enhanced performance, solidifying deep learning as an invaluable tool in today's technology-driven world. Its capacity to surpass human capabilities is a driving force in fields such as robotics, AI-driven automobiles, language processing, and image captioning, underscoring its pivotal role in reshaping our technological landscape.

Office [1], Road [2], Building [3], House [4]

Figure-1. Various Applications of Surveillance System

Deep neural networks, often called deep learning models, stand out from conventional neural networks due to their remarkable depth, characterized by numerous hidden layers that can extend to over 150 layers. This depth enables them to perform complex and hierarchical feature extraction, making them highly effective in tackling intricate tasks. Unlike traditional networks, deep neural networks learn features directly from data through extensive training on large datasets, without the need for explicit human guidance. Convolutional Neural Networks (CNNs) are a prominent example of deep neural networks, particularly well-suited for image-related tasks, thanks to their 2D convolution layers. They excel in capturing intricate patterns and details within images, revolutionizing fields like computer vision and image analysis. Deep neural networks play a pivotal role in the future of artificial intelligence and data analysis.

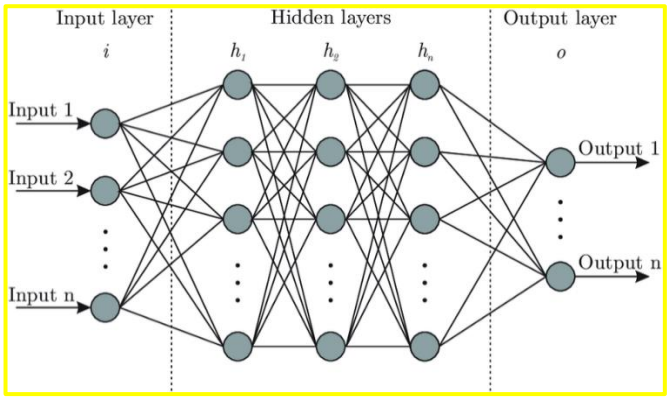


Figure-2. Structure of a Neural Networks

Convolutional Neural Networks (CNNs) represent a significant breakthrough in computer vision and image analysis due to their remarkable ability to classify photos without the need for manual identification of attributes. CNNs excel at automating feature extraction directly from visual data, making them exceptionally well-suited for image-related tasks. What sets CNNs apart is their capacity to learn and extract the necessary features from images during training without the requirement for pre-training on specific attributes. This simultaneous learning of features and classification leads to more efficient and accurate object classification within the field of computer vision. The complexity of visual features is effectively managed by employing multiple hidden layers within CNNs, with the initial layers specializing in recognizing edges and simple forms, gradually progressing to discern complex object attributes.

CNNs leverage their deep architecture, comprising numerous hidden layers, to progressively recognize intricate visual features within images. The initial hidden layers serve as the foundation, detecting fundamental features like edges and basic shapes, which are essential building blocks for identifying more complex objects. These early layers act as feature extractors, transforming raw pixel data into hierarchical representations that capture increasingly abstract and nuanced visual attributes. As the network advances through deeper hidden layers, it gains the capability to distinguish intricate patterns and attributes, ultimately culminating in the accurate classification of objects within images. CNNs' ability to automate this feature recognition process has revolutionized the field of computer vision, enabling advancements in tasks such as image classification, object detection, and facial recognition.

Machine Learning Versus Deep Learning

Deep learning, which uses human-like artificial intelligence, is more efficient than machine learning. Machine learning algorithms parse data manually. It learns from data to create a model that categorizes the thing. Machine learning models use data to improve.

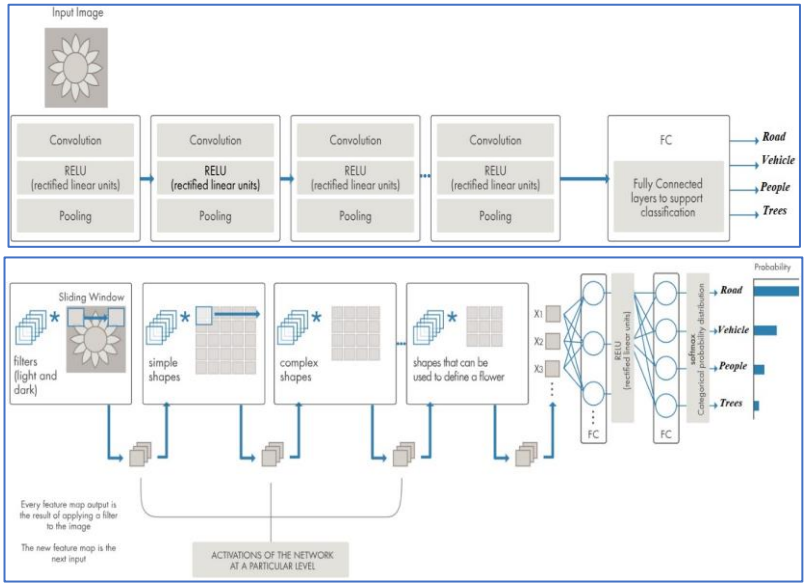


Figure-3. A Sample Layer Architecture of CNN - Training and Testing process

Deep learning models, drawing inspiration from the human brain's neural system, harness the power of Artificial Neural Networks (ANNs). ANNs are designed to continuously analyze data, akin to the decision-making capabilities of the human brain. The efficacy of deep learning models significantly improves with larger datasets, reflecting the importance of data size in this domain. While machine learning offers a diverse array of models and approaches for various applications, deep learning stands out for its ability to learn intricate patterns and make accurate decisions. Training a deep learning model, however, demands access to extensive data, often involving millions of images and hours of videos. To process such data efficiently, Graphics Processing Units (GPUs) play a crucial role, enabling rapid computation. When GPUs are unavailable, traditional machine learning algorithms may offer more practical solutions compared to deep learning.

The paper presents an innovative approach to enhancing object recognition by leveraging two distinct learning algorithms, each characterized by its speed of learning. What sets this research apart is the introduction of frame rate-based video data analysis within the realm of deep learning models. When the frame rate is low, the model focuses on analyzing frames slowly, extracting spatial semantics to understand the static elements in the video. Conversely, at high frame rates, the model shifts its focus to the analysis of structures and temporal semantics, which capture dynamic changes over time. By integrating spatial and temporal information, this deep learning model elevates video processing and object recognition to a new level, enabling it to understand and categorize objects more comprehensively within dynamic visual contexts. This innovative approach holds promise for applications requiring nuanced video analysis and object recognition.

Deep Learning-Based Object Detection and Recognition

The deep learning [2, 13-15] approach is ideal for video processing, object detection and recognition, pattern recognition, and speech recognition. This research suggested Deep Learning's Convolution Neural Network for object and anomaly detection. The application's wired or wireless CCTV cameras feed the system/PC's input footage. If hooked, the PC stores the video file. Routers send the video to the PC. This paper makes assumptions to clarify the method. This paper detects abnormalities in surveillance videos using deep learning. Deep learning is inspired by neural network architectures that learn features and represent data. A neural network model has input, output, hidden, and more hidden layers. Deep learning has several massive, multilayer networks. Convolutional Neural Networks [16-18]are popular deep learning networks. (CNN). This research uses CNN architecture to detect aberrant activity. CNN automatically classifies video/image features. This study analyzes and assesses the suggested CNN architecture's performance on human, vehicle, and animal behaviors in varied backgrounds, a novel data processing method.

Proposed Approach

The proposed CNN architecture is explained in this section. The video V is divided into frames (images) F, in which various objects and their activities are normal and abnormal. Some of the specific abnormal activities are different activities than the usual activities. For improving the efficiency of video/image processing, the images are initially applied for preprocessing using a moving $\mathbf{3} \times \mathbf{3}$ average filter, which removes the noises occurring in the images. It can be represented as,

$$\mathbf{y_{ij}} = \sum_{\mathbf{k=-m}}^{\mathbf{m}} \sum_{\mathbf{l=-m}}^{\mathbf{m}} \mathbf{w_{kl}} \mathbf{x_{i+k,j+l}}$$

Where the input image is represented as $\mathbf{x_{ij}}$, $(\mathbf{i}, \mathbf{j})$ represents the pixels in the image, and $\mathbf{y_{ij}}$ represents the output image. Similarly, a linear filter with the size $\mathbf{3} \times \mathbf{3}$, is used on

$$(\mathbf{2m} + \mathbf{1}) \mathbf{x} (\mathbf{2m} + \mathbf{1})$$

having the weights $\mathbf{w_{kl}}$ for every $\mathbf{k}$ and $\mathbf{l}$ from $-\mathbf{m}$ to $\mathbf{m}$, equal to $\mathbf{1}$.

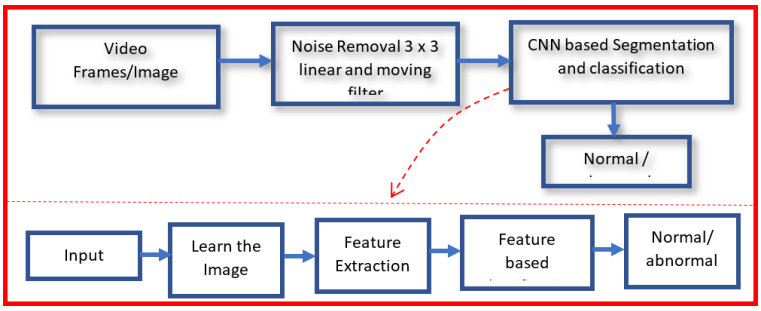
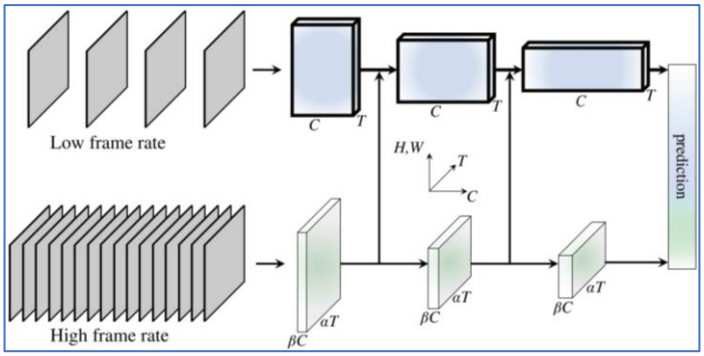


Figure-4. Neural Network-Based for Video Processing

In video processing, the selection of key frames is crucial to streamline the process and enhance classification accuracy. Approximately 30% of video frames are meticulously labeled as "normal" or "abnormal" for training, with a focus on annotating anomalous behaviors to conserve computing resources. A dedicated database stores aberrant frames for reference during training. The Convolutional Neural Network (CNN) architecture is central to this methodology, consisting of input, middle, and output layers for classification. Frames are resized to 32x32 dimensions for consistency, and key operations, including convolution, pooling, and Rectified Linear Unit (ReLU) activation, are applied in the middle layers. The proposed SF-CNN architecture adapts to varying frame rates, offering adaptability and comprehensive capabilities for video analysis.



(a).

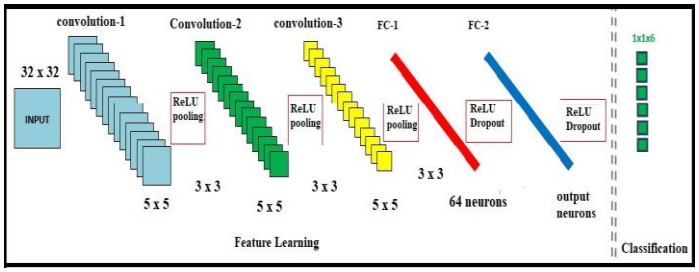


Figure-5. (a). SF-CNN framework, (b). Functionality of CNN

In image processing and feature extraction, convolution filters are instrumental in highlighting crucial information like edges and textures. They play a vital role in recognizing abnormal actions in video frames by detecting discrepancies in feature values. This deep learning model employs 5x5x3 convolution filters on input color images, with 2-pixel padding to preserve edge data. Rectified Linear Unit (ReLU) activation function is used for efficient training. Pooling layers further downsample the data to 15x15 to enhance feature extraction, with deliberate avoidance of excessive downsampling. The CNN architecture transitions to classification with fully connected layers and SoftMax activation, determining class probabilities for accurate categorization. Random weight initialization aids network robustness and effectiveness.

SF-CNN

This paper introduces a novel approach that integrates a deep Convolutional Neural Network (CNN) with a Slow-Fast learning method to analyze video segments effectively. The proposed method employs two parallel CNN models, referred to as the Slow and Fast learners, to process the same input video. Video content typically comprises two distinct types of data: static and dynamic. Static data remains relatively unchanged or changes slowly, while dynamic data represents continuously changing elements such as moving objects. Figure-5(a) illustrates the flow of video frames obtained from fast streaming to the slow frame rate learner. The rationale behind this approach is that the slow learner can effectively learn from the output of the fast learner. In this context, the data format used in the Slow-Fast (SF) learner is defined as follows:

Fast: {αT, S^2, βC}

Slow: {T, S^2, αβC}

These two sets of data are fused together to facilitate the SF learning process, allowing for the efficient handling of both static and dynamic information within the video. The SF-CNN approach introduces a distinct methodology for transforming the data, with T-2-C (Time-2-Channel) playing a pivotal role. This involves reshaping and transposing the data from {αT, S^2, βC} to {T, S^2, αβC}, effectively consolidating all α frames into one frame per channel. Another technique, TSS (Time-strided-sampling), treats each α frame as a sample, resulting in data transformation: {αT, S^2, βC} = {T, S^2, βC}. Additionally, TSC (Time-strided-convolution) employs a three-dimensional convolution with a 5 x 12 kernel, generating 2βC output channels with α as the stride. To further streamline the SF learning process, global pooling operators are applied to both slow and fast learners, effectively reducing dimensionality. The fully connected (FC) layer then identifies output from both learners, with subsequent classification of object behavior using SoftMax. The SF-CNN methodology is implemented via an algorithm, compatible with various programming languages, ensuring its adaptability and practicality for video analysis and object recognition tasks.

RESULTS AND DISCUSSION

This research paper's algorithm is meticulously implemented and rigorously verified using the MATLAB software, which features a built-in CNN module within the Image Processing toolbox. MATLAB not only provides a robust platform for algorithm implementation but also offers an array of functionalities and the flexibility to leverage various inbuilt algorithms. This versatility extends to regression methods and decision-making techniques, enhancing the algorithm's performance verification capabilities. To comprehensively evaluate the proposed framework, the study incorporates seven distinct datasets. Table-1 outlines the datasets, encompassing human, vehicle, and animal activities, each comprising frames captured as RGB images with varying pixel sizes. While some datasets feature photos with dimensions such as 276x236, 352x240, and 360x288, all frames are standardized to 32x32 pixels due to frame constraints. These datasets exhibit a balanced distribution of both positive and negative frames, totaling 12,000 frames generated from 100 videos. Among these, 5,000 frames represent normal activities, while the remainder showcase problematic or aberrant behaviors. The proposed framework undergoes a rigorous evaluation in a two-stage experimental setup. In Experiment-1, the algorithm is tasked with the classification of normal and abnormal activities. In Experiment-2, the focus shifts to classifying the various aberrant classes within the dataset. The training of networks is facilitated through stochastic gradient descent with momentum. To fine-tune network hyperparameters, the experiments encompass a range of settings, including 10-100 epochs and learning rates spanning from 0.001 to 0.1. Each epoch involves the forward and backward transmission of training samples, accompanied by the learning rate adjustment. This

comprehensive experimentation strategy ensures the robustness Achieving high-accuracy results in deep learning demands an abundance of input data. This necessitates the utilization of powerful hardware, particularly Graphics Processing Units (GPUs) capable of handling the computational demands of deep learning tasks. In this context, the experiments involving Convolutional Neural Networks (CNNs) were conducted using an Intel Core i7 processor, running MATLAB-2017. This setup provided the computational resources necessary for the CNN experiments. The experiments revolved around the binary classification of photos, with various datasets offering a collection of normal and abnormal images that were systematically categorized. These datasets comprised human-related videos capturing activities such as walking, pointing, hugging, and shaking, as well as inappropriate actions including kicking, shoving, and punching. The CNN model proposed in this study was employed to classify images into distinct classes, which were then compared to the labeled classes within the datasets to assess the model's performance. The results, illustrated in Figure-7, showcased the CNN's proficiency in identifying abnormalities within individual frames through a binary classification model. The images presented in Figure-7 served as visual evidence, with yellow bounding boxes highlighting instances of aberrant activity. Notably, the CNN effectively recognized and classified various abnormal activities, ranging from pushing, kicking, and fighting to pedestrians on roads designated for vehicles, contributing to the evaluation of the experiment's learning rate for detecting abnormalities.

(a). Abnormality Detection in First three dataset

(b). Abnormality Detection in Four dataset

Figure-6. Abnormality Detection using proposed CNN

In both experiments, normal and abnormal, including all abnormal classes, are more accurate using the proposed CNN, which is understood from Figure-7. The performance of the proposed CNN is high and is evident by comparing both results. The abnormal detection accuracy is increased since the testing process is always compared with the training process. Hence, for human interacted classification, the training classes are highly accurate and are used in the testing process. The experiment further evaluates object classification performance

by testing 70% of the frames and calculating categorization accuracy. Table-2 provides a summary of the experiment's results, encompassing total frames and accurately classified frames. Notably, among the 12,134 frames derived from all seven datasets, 6,035 frames represent typical activities, while the remaining 6,099 frames depict predefined aberrant actions verified by previous studies. These results offer a comprehensive perspective on the CNN technique's efficacy in object classification within various real-world scenarios.

Table-2. Performance Calculation

Data	Total before the error	Total after error	Normal frames	Abnormal frames
DB	**13,158**	**11,122**	**5,923**	**6,532**
Proposed CNN	13,170	11,134	5,934	6,544

Table-2 provides a detailed breakdown of the suggested CNN architecture's performance in classifying frames as normal and abnormal. Before the error: "DB" dataset had a total of 13,158 frames, with 5,923 being normal and 6,532 being abnormal. "Proposed CNN" dataset had a total of 13,170 frames, with 5,934 being normal and 6,544 being abnormal. After the error: "DB" dataset had a total of 11,122 frames, with 6,532 being normal and 4,590 being abnormal. "Proposed CNN" dataset had a total of 11,134 frames, with 6,544 being normal and 4,590 being abnormal.highlighting its proficiency in detecting and categorizing aberrant behaviors. The evaluation encompasses several performance metrics, including True Positives (TP), True Negatives (TN), False Positives (FP), False Negatives (FN), Sensitivity, Specificity, and overall accuracy.

Showcasing the outstanding performance of the proposed CNN architecture. Comparative analysis against alternative methods reveals that the suggested CNN consistently outperforms these approaches. Impressively, the proposed CNN achieved an accuracy score of 99.6%, surpassing the performance of techniques identified in the literature survey. The choice of epochs and learning rate significantly influences classification accuracy, with CNN learning rates spanning 0.001, 0.01, and 0.1 for epochs ranging from 10 to 50. Diverse datasets, learning rates, and epoch combinations were systematically examined to calculate accuracy, as detailed in Table-3.

The findings underscore that the suggested CNN model exhibits superior time complexity, object detection capabilities, and classification accuracy compared to alternative techniques. Notably, the incorporation of 100 epochs contributed to the enhancement of accuracy, reaffirming the proposed approach's efficacy in real-world scenarios involving video analysis and object recognition.

Table-3. Accuracy Based on Learning Rate

Learning Rate	Max. No. of. Epochs	Dataset Accuracy (%)					
		CMU	UTI	PEL	HOF	WED	UCS-AD
0.001	10	99.63	64.77	91.40	58.47	89.18	99.87
	20	98.97	56.52	91.40	98.97	98.97	98.97
	30	98.97	57.71	91.40	98.97	98.97	98.97
	40	98.97	70.15	91.40	98.97	98.97	98.97
	50	98.97	99.12	91.40	98.97	98.97	98.97
0.01	10	98.97	54.87	91.40	98.97	98.97	98.97
	20	98.97	99.57	91.40	98.97	98.97	98.97
	30	98.97	99.84	87.95	98.97	98.97	98.97
	40	98.97	98.81	98.97	98.97	98.97	98.97
	50	98.97	99.77	98.97	98.97	98.97	98.97
0.1	10	46.61	54.87	8.09	98.97	98.97	98.97
	20	0	0	8.10	98.97	98.97	98.97

Learning Rate	Max. No. of. Epochs	Dataset Accuracy (%)					
		CMU	UTI	PEL	HOF	WED	UCS-AD
	30	0	54.90	0	98.97	98.97	98.97
	40	0	54.90	91.37	98.97	98.97	98.97
	50	0	0	9.62	98.97	98.97	98.97

CONCLUSION

This research represents a significant advancement in the field of surveillance system anomaly detection by introducing a novel deep learning-based technique. The cornerstone of this work is the development of a Convolutional Neural Network (CNN) architecture designed to excel in education, information extraction, and the classification of abnormal video frames in surveillance scenarios. The primary objective of this article is to identify and categorize anomalies present in diverse datasets, with the overarching aim of creating a universal system for the detection of anomalies across human, animal, and vehicle surveillance domains. The utilization of deep learning models contributes to a substantial improvement in accuracy, enhancing the effectiveness of surveillance systems.

The performance of the proposed deep learning model is rigorously tested, with particular focus on the impact of varying learning rates and epochs. It is observed that increasing the number of epochs has a positive effect on accuracy, resulting in improved anomaly categorization. The experimental results unequivocally demonstrate that the suggested CNN architecture outperforms competing methods, achieving an exceptional anomaly categorization accuracy of 99.6%. Notably, this technology offers a completely autonomous solution, making it an ideal choice for integration into any monitoring system, as it excels in the precise classification of abnormalities across a wide spectrum of surveillance scenarios.

REFERENCES

1. Zhu, Z., Z. Xu, and J. Liu, Flipped classroom supported by music combined with deep learning applied in physical education. Applied Soft Computing, 2023. 137.
2. Yu, C., X. Bi, and Y. Fan, Deep learning for fluid velocity field estimation: A review. Ocean Engineering, 2023. 271.
3. Zhao, W., et al., A learnable sampling method for scalable graph neural networks. Neural Netw, 2023. 162: p. 412-424.
4. Rajeshkumar, G., et al., Smart office automation via faster R-CNN based face recognition and internet of things. Measurement: Sensors, 2023. 27.
5. Jiang, H., et al., A review of deep learning-based multiple-lesion recognition from medical images: classification, detection and segmentation. Comput Biol Med, 2023. 157: p. 106726.
6. You, K., C. Zhou, and L. Ding, Deep learning technology for construction machinery and robotics. Automation in Construction, 2023. 150.
7. Zhu, G., et al., Various solutions of the (2+1)-dimensional Hirota-Satsuma-Ito equation using the bilinear neural network method. Chinese Journal of Physics, 2023.
8. Xie, G., et al., A life prediction method of mechanical structures based on the phase field method and neural network. Applied Mathematical Modelling, 2023. 119: p. 782-802.
9. Mohammad-Rahimi, H., et al., Deep learning: A primer for dentists and dental researchers. J Dent, 2023. 130: p. 104430.
10. Rayadurgam, V.C. and J. Mangalagiri, Does inclusion of GARCH variance in deep learning models improve financial contagion prediction? Finance Research Letters, 2023.
11. Ravikumar, K.C., et al., Challenges in internet of things towards the security using deep learning techniques. Measurement: Sensors, 2022. 24.
12. Mohammed, A. and R. Kora, A comprehensive review on ensemble deep learning: Opportunities and challenges. Journal of King Saud University - Computer and Information Sciences, 2023. 35(2): p. 757-774.

13. Fernando, K.R.M. and C.P. Tsokos, Deep and statistical learning in biomedical imaging: State of the art in 3D MRI brain tumor segmentation. Information Fusion, 2023. 92: p. 450-465.
14. khelili, M.A., et al., Deep learning and metaheuristics application in internet of things: A literature review. Microprocessors and Microsystems, 2023. 98.
15. Tanveer, M., et al., Deep learning for brain age estimation: A systematic review. Information Fusion, 2023. 96: p. 130-143.
16. Alshehri, D.M., Blockchain-assisted internet of things framework in smart livestock farming. Internet of Things, 2023. 22.
17. Huang, R., X. Yang, and P. Ajay, Consensus mechanism for software-defined blockchain in internet of things. Internet of Things and Cyber-Physical Systems, 2023. 3: p. 52-60.
18. Si, L.F., M. Li, and L. He, Farmland monitoring and livestock management based on internet of things. Internet of Things, 2022. 19.
19. Hemmati, A. and A.M. Rahmani, The Internet of Autonomous Things applications: A taxonomy, technologies, and future directions. Internet of Things, 2022. 20.
20. Alshammari, H.H., The internet of things healthcare monitoring system based on MQTT protocol. Alexandria Engineering Journal, 2023. 69: p. 275-287.
21. Siegel, J.W., et al., Greedy training algorithms for neural networks and applications to PDEs. Journal of Computational Physics, 2023. 484.
22. Seo, M. and S. Min, Graph neural networks and implicit neural representation for near-optimal topology prediction over irregular design domains. Engineering Applications of Artificial Intelligence, 2023. 123.