RESEARCH & INNOVATION DIVISION
IJIST

Unusual Activity Categorization using Machine Learning

Waqar Ishaq[1], Syed Waqar Shah[2], Bilal Ur Rehman[2]*, Mir Muhammad Wasiq[1]
[1]Telecommunication Department, Hazara University, Mansehra, KPK, Pakistan, 21120
[2]Department of Electrical Engineering, University of Engineering and Technology, Peshawar, KPK, Pakistan.
***Correspondence**: bur@uetpeshawar.edu.pk

Citation | Ishaq. W, Shah. S. W, Rehman. B. U, Wasiq. M. M, "Unusual Activity Categorization using Machine Learning", IJIST, Vol. 8 Issue. 11 pp 4188-4211, November 2026

Received | Sep 22, 2026 **Revised |** Oct 24, 2026 **Accepted |** Nov 10, 2026 **Published |** Nov 16, 2026.

Recognizing unusual human activities is a major challenge in intelligent video-surveillance systems because variations in illumination, complex backgrounds, shadows, object appearance, image noise, and human motion can reduce recognition performance. This study proposes a deep-learning-based framework for unusual-activity categorization using a Multi-Domain Deep Unrolling Network (MudNet) and pretrained ResNet101 for deep feature extraction. The extracted features are fused, while the Whale Optimization Algorithm (WOA) is employed to select the most discriminative features before classification using different variants of K-nearest neighbor (KNN) and support vector machine (SVM) classifiers. The proposed framework is evaluated on three publicly available datasets: IXMAS Action, Anomalous Activity, and KTH. For the IXMAS dataset, WOA selected 1,946 features, and Fine KNN achieved 99.2% accuracy, 0.99 precision, 0.99 recall/sensitivity, and an F1-score of 0.99 using 10-fold cross-validation. For the Anomalous Activity dataset, 1,091 WOA-selected features with Cubic SVM achieved 100% accuracy, 1.00 precision, 1.00 recall/sensitivity, and an F1-score of 1.00 using 5-fold cross-validation. For the KTH dataset, 885 WOA-selected features with Cubic SVM achieved 96.6% accuracy, 0.96 precision, 0.96 recall/sensitivity, and an F1-score of 0.96 using 10-fold cross-validation. These results demonstrate that the proposed MudNet–ResNet101 feature-fusion framework combined with WOA-based feature selection can effectively categorize human activities across different surveillance datasets while maintaining high classification performance.

Keywords: Machine Learning, Deep Learning, MudNet, ResNet101, Whale Optimization Algorithm, Human Activity Recognition, Video Surveillance System.

IPIndexing Indexing Portal
CiteFactor Academic Scientific Journals
RESEARCHBIB ACADEMIC RESOURCE INDEX
IDEAS
Scientific Journal Impact Factor TOGETHER WE REACH THE GOAL
JOURNALS MASTER LIST
ROOTINDEXING JOURNAL ABSTRACTING AND INDEXING SERVICE
Scilit
INFOBASE INDEX
DRJI
Crossref
ResearchGate
WIKIDATA
HEC

Introduction:

Technology has strong relation with the human in the current age. All social, official and personal activities have strong connections with digital tech. Such interactions have transformed human behavior as well. These activities have system log files which are categorized by the researchers as usual and unusual activities. Detecting unusual activities in a surveillance system is the most prevalent research area in the computer vision. Computer vision and many more fields work to detect or recognize suspicious activities. In video surveillance, both living and non-living actions/ behaviors are recorded. Moreover, time and environmental changes are monitored and analyzed to describe and recognize human behavior. Currently, the increase in the crime rates, threats, violence, terrorist attack[1], and hazardous conditions create a chaotic environment around us. Many applications have been devised for security maintenance such as human-computer interaction, nursing care facilities, theft detection, violence detection[2] fire & smoke detection[3], and many more. Furthermore, these applications are used for indoor and outdoor activity monitoring such as university campus, academic institutions, airport, railway, bus stations, shopping malls, illegal parking detection[4] etc.

Researchers focus on capturing normal and abnormal human behaviors such as running, fighting, jogging, jumping, and playing using advanced video monitoring system. Such systems play an imperative role to separate suspicious behavior from normal behavior without any manual operation. It helps the officers to act or know what is happening on the spot without their physical presence. The aim of this work is to provide a safe and secure environment, reduce crime rates via the video surveillance scheme. Video surveillance systems offer automated monitoring system to provide safe environment for indoor or outdoor activities. These monitoring systems are utilized in medical stores, banks, border security [5], traffic signals, and many more areas. However, the researchers not only focused on the analysis of static (spatial) representation but also consider dynamic (representation)[6] to capture the camera eye in the surveillance system. The Closed-circuit television (CCTV) is an example of video surveillance monitoring system to detect suspicious actions. In CCTV footage, various activities are recognized in public areas such as traffic signal monitoring anomalous activities, detecting terrorist attacks[7] accidental conditions[8], criminal activity[9], child-abusing, etc[10]. It can generate a message alarm and send video/image recordings to the security monitoring system in case of any emergency, accident, or threat and many more [11] Likewise, video surveillance plays a vital role in the design and construction of the smart city and homes. It deploys various Internet of Things (IoT) devices that remotely monitor security systems and detect anomalous activities automatically.

The purpose of this work is to address the challenges that occurred during the recognition of suspicious or unusual activities in a surveillance system. These challenges include poor contrast, illumination, complex background, lighting effects, similar texture, size, and shape variations of objects, which hamper examining anomalous activities in images/videos. Different methods and algorithms are employed to overcome the challenges that come across in achieving the optimal solution such as classification and accuracy.

Research Objectives:

In this study, deep convolution neural network architecture (DCNN)[12] is proposed to execute a cubic KNN classifiers for analyzing deep features. The proposed method consists of following steps to predict unusual behavior in video surveillance data: The pre-processing techniques are used to enhance the image quality. This phase includes de-noising, normalization, image sharpening**,** histogram equalization, image enhancement, etc. Feature extraction techniques are employed to acquire valuable information from the selected dataset images. Deep networks are used to extract convenient features automatically. Whale optimization algorithms are used to select an optimized set of features. Feature fusion, a

combination of selected features from a different layer of the CNN model are serially concatenated after Feature selection. Then different classifiers are applied for classification to assess the accuracy and performance.

Novelty Statement:

The novelty of this study lies in the development of an integrated unusual-activity-recognition framework that combines the proposed Multi-Domain Deep Unrolling Network (MudNet) with pretrained ResNet101, deep-feature fusion, Whale Optimization Algorithm (WOA)-based feature selection, and comparative KNN/SVM classification. Rather than relying on a single feature extractor or directly classifying the complete deep-feature representation, the proposed framework obtains complementary feature representations from MudNet and ResNet101 and serially fuses them to form a richer feature space. WOA is subsequently applied to the fused representation to reduce redundant features and retain a compact set of discriminative features before classification. A further distinguishing aspect is the systematic comparison of multiple KNN and SVM variants under 5-fold and 10-fold validation rather than relying on a single classifier. The complete framework is evaluated across three different benchmark datasets IXMAS Action, Anomalous Activity, and KTH allowing its behavior to be examined across different activity classes and surveillance conditions. hus, the principal contribution is the combined MudNet–ResNet101 feature representation, feature-level fusion, WOA-driven feature optimization, and multi-classifier evaluation within a single unusual-activity-recognition pipeline, rather than claiming novelty for any individual component in isolation.

Research Gap:

Existing unusual-activity-recognition methods often rely on a single feature extractor, a limited classification strategy, or separate optimization stages. This can reduce the ability to capture complementary deep features and remove redundant information effectively. The proposed framework addresses this gap by combining MudNet and ResNet101 feature extraction, feature fusion, WOA-based feature selection, and comparative KNN/SVM classification within one unified pipeline evaluated on the IXMAS, Anomalous Activity, and KTH datasets.

Related Work:

This section deliberates the instigation of suspicious activities recognition and the utilization of numerous progressive approaches. Human activity recognition (HAR) is trending in various areas, including smartphones[13] healthcare monitoring, bus tracking systems, self-management systems, Pickpocket Suspects[14] and daycare or infant monitoring systems[15][16]finds miscellaneous applications in monitoring systems such as video surveillance systems, crime prevention, automated intelligent visual monitoring, human-computer interaction, and traffic security systems[17]. The video surveillance system captures real-time static and dynamic data[18]. It is also employed in border areas to detect abnormal activities, suspicious movements, and generate automatic warning alarms in control rooms. An entropy Markov model [19] proposes successful Automatic Human Posture estimation (A-HPE) [20] and a tracking system to detect critical behaviors. Human activity recognition tracks human body parts using the Pose-based Convolutional Neural Network (P-CNN) approach, which collects gesture and appearance-based data [21] The Deep Full-Body-Human-Pose-Estimate (DFB-HPE) approach analyses RGB images using Convolutional Neural Network (ConvNets) providing information on humans' daily life activities within unrestricted locations. It predicts the positions or angles of fifteen human body joints and classifies the features using a support vector machine (SVM).

These approaches are used in various applications such as human-robot interaction[22] gesture recognition [23], and sports performance analysis [24][25]. The study [26] employs regularization and thresholding techniques to denoise pictures, eliminating salt and pepper

noises. In [27] study presents a denoising framework for video surveillance systems, utilizing learning-based techniques to remove noise from images. Additionally, a machine learning-based transfer learning technique addresses criminal activity detection in densely populated areas[28]. Moreover, a spatiotemporal stability approach is proposed to enhance prediction accuracy via deep CNN approaches[29]. In [30], research utilizes CNN techniques to develop a 2D posture prediction model to improve detection accuracy. By employing algorithm, i.e., You Only Look Once (YOLO) and deep learning techniques in the surveillance system, this research study devises a model for predicting robbery actions [31].Hybrid whale-and-aquila optimizer method selects the best features before applying SVM for classification to achieve optimal results [32]. Median filters are utilized by researchers to enhance image quality and more accurately identify suspicious activity. Additionally, KNN and SVM-based classifiers are employed [33][34] to improve prediction performance.

More recently, advanced deep-learning approaches have focused on learning both spatial and temporal characteristics of human activities and abnormal events. Transformer and Vision Transformer architectures have demonstrated their ability to capture long-range dependencies and relevant contextual information from video sequences [35].

Spatiotemporal deep-learning models have also been investigated for video anomaly detection and human activity recognition by jointly modeling appearance and motion information across consecutive frames [36]. Furthermore, recent optimization-based feature-selection approaches have been employed to eliminate redundant features and retain more discriminative information, thereby improving classification efficiency and recognition performance [37]. These developments highlight the growing importance of combining robust deep feature extraction, spatiotemporal representation, and optimized feature selection in intelligent surveillance systems.

Proposed Methodology:

This section discusses traditional and advance deep learning approaches to recognize anomalous behavior during surveillance. The architecture of the proposed methodology is shown in the figure 1, which consists of different.

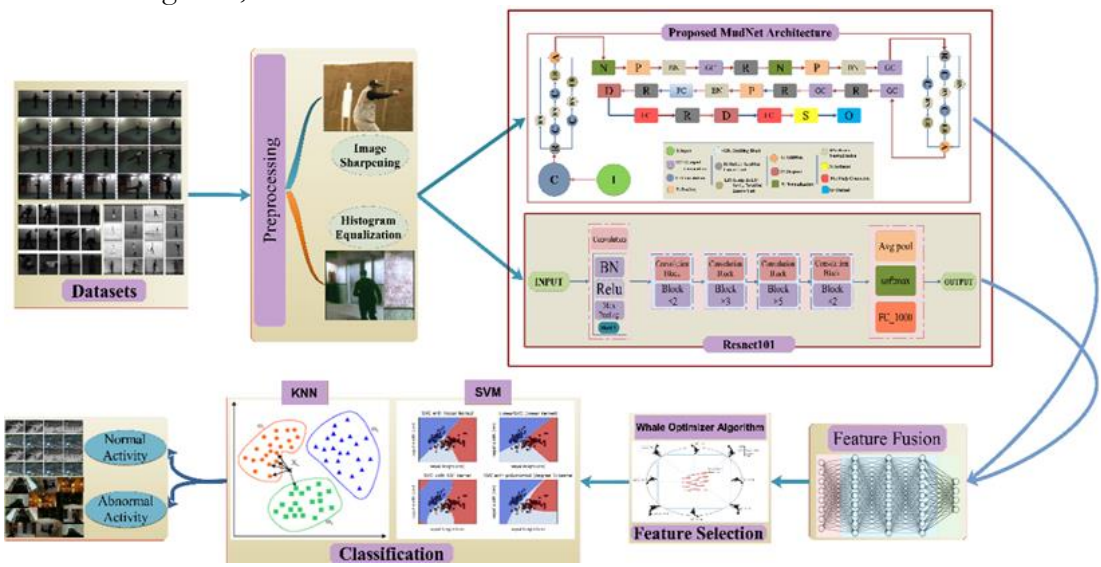


Figure 1. Block Diagram for Proposed Architecture

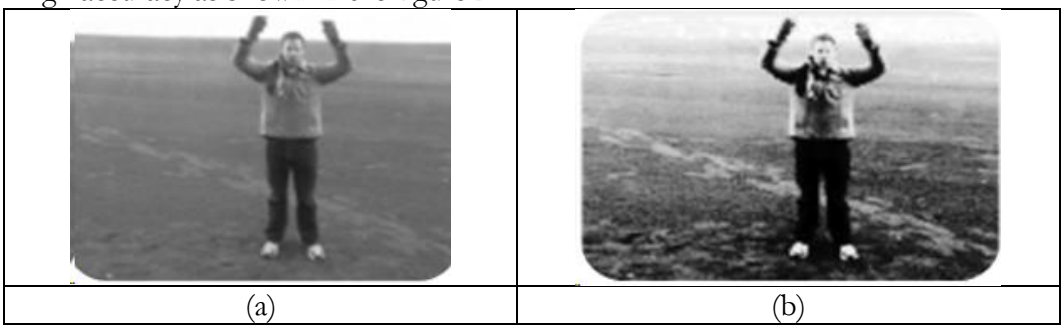
Pre-processing:

Image pre-processing is the first step of the proposed method to transforms the image into format that fits the model requirement. Such processing improves the image quality to achieve better results and recognize unusual behaviour(s) with high accuracy. Various image pre-processing techniques such as image enhancement, colour correction, noise removal, data

normalization, data augmentation, image sharpening, and many more. The proposed method uses the following two techniques in the image pre-processing step:

Image Sharpening:

In the proposed method, image sharpening in pre-processing step, increases the distinction along the edges of the image to make it comprehensible and simple to perceive. It enhances the quality of the images which are dull image and poor at the edges. Two types of filters, namely spatial and Laplacian filters are used in image sharpening. It emphasizes the boundaries and fine details of the edges. Figure 2. Shows the image sharpening output using the spatial filter i.e. the first derivation removes the blur effect in the image, and a Laplacian filter using the second derivation for fine details that cleans the texture and detects activities with high accuracy as shown in the figure 2.



(a) (b)

Figure 2. (a) Original Image (b) Image Sharpened

The image sharpening equation is shown below:

$$\nabla^2 \mathrm{f} = \frac{\partial^2 f}{\partial a^2} + \frac{\partial^2 f}{\partial b^2} (2)$$

Where, 'f' represents the function value, '∇^2' is the symbol for the Laplacian filter, and '∂^2' is a partial derivative sign. The second derivative of 'a' is added to the second derivative of 'b' which takes row and column values in a 2D image for the image sharpening. The 'a' axis represents a vertical line and 'b' axis shows a horizontal line. '$\nabla^2 \mathrm{f}$' is getting an image that highlights the edges and fine details.

Histogram Equalization:

Histogram equalization is lossless technique of improving the image intensities by adjusting the contrast in the pre-processing step of the proposed method. Here the number of pixels is equally distributed in each grey level from maximum to minimum to improve image quality or contrast. Equations (3) and (4) transform the image into discrete form.

$$H_m = T(d_m) = (L-1) \sum_{i=0}^{m} Q_r(d_i) (3)$$

$$Q_r(d_i) = \frac{n_m}{\mathrm{n}} \cdots\cdots let\ m = 0,1,\cdots\cdots, L-1 (4)$$

Where, 'd_m' is a gray level, 'n_m' is the number of pixels, '$Q_r(d_i)$' is normalized values, 'H_m' is the value of cumulative addition. This equation is equalized by the grey level to the total number of pixels. Histogram equalization enhances the image as shown in the figure 3.



(a) (b)

Figure 3. (a) The original image and (b) histogram equalization image

Feature Extraction:

Feature extraction is the second step in the proposed method followed by the pre-processing step. Here the data is subjected to feature extraction. To overcome the high computational cost due to large data size, feature extraction performs dimensionality reduction to capture useful information. The multi-domain deep unrolling network (MudNet) and residual network (Resnet) are the two models used to extract suitable features and the labels from the pre-processed data.

Multi-Domain Deep Unrolling Network (MudNet) Model:

The multi-domain deep unrolling network (MudNet) model of convolutional neural network (CNN) is trained over the extracted features. The MudNet model categorizes unusual activity and enhances the effectiveness of security systems. The framework of this model has input of size 227×227×3, consists of CNN pipeline which includes 23 convolution (conv) layers, 3 group convolution (GC) layers, 3 max-pooling (P) layers, 18 batch normalization (BN) layers, 3 layers' normalization (N), 7 dropout (D) layers, 17 Rectified linear unit (Relu) layers, 3 fully connected (FC) layers, one softmax layer, and one classification output layer are used. The convolutional layer is a key component in a CCN pipeline; in fact, most networks require at least one convolutional layer. The pooling layer is used for down-sampling. In the network, BN and ReLU layers are the most used activation functions for model stability and learning complex patterns respectively. The FC_2 layer has 2048 active features in the MudNet model. Following is the proposed equation of the MudNet model:

$$f_{mud} = \{f_{11} + f_{12} + \ldots\ldots\ldots\ldots + f_{nm}\} \quad (5)$$

Whereas " f_{mud}" and "f_{nm}" represents MudNet model and the extracted features. The proposed architecture of the MudNet model in the Figure.1 shows various layers that are used sequentially to perform their required functions. The first layer of the network is an input image, with a size of 227×227×3. Then, a convolutional layer is used with a filter window of size 11×11. Furthermore, the procedure is divided into three blocks i.e. Block 1 (B1) and block 3 (B3) which perform identical functions, with ReLU, BN, Conv, LR (Leaky ReLU). Then the last layer of block 1 is known as Addition layer, is used for joining. After B1, the layers used are N, P, BN, GC (Group Convolution), R (ReLU), N, P, BN, GC, and then use B3. Additionally, the layers of block 3 are used in reverse order depending on the network requirements, and then additional layers are added with GC, R, GC, R, P, BN, FC, R, D, FC, R and D layers for making the network performance better. At the end of this framework, some finishing layers are utilized that are fully connected, including softmax and output layers. The most suitable features are collected by the fully connected layer and then the softmax function is applied to them. The output layer generates accurate detection results, while improving security performance. The model is trained by using trained part of the data and accuracy is evaluated by test part of the dataset. Each iteration improves the network performance to recognize unusual activities or actions in the video surveillance system. Different set of filters and pooling layers are used in each iteration. The proposed approach forms a framework that extracts the strongest and most suitable features. Various activation layers are employed to improve the performance of abnormal behavior categorization. In this framework, active layer FC 2, collects the best features in match to other FC layers that are employed in the network. The selected features are used in the video surveillance system to identify people with unusual behavior. The architecture of the proposed MudNet model is shown in figure 4.

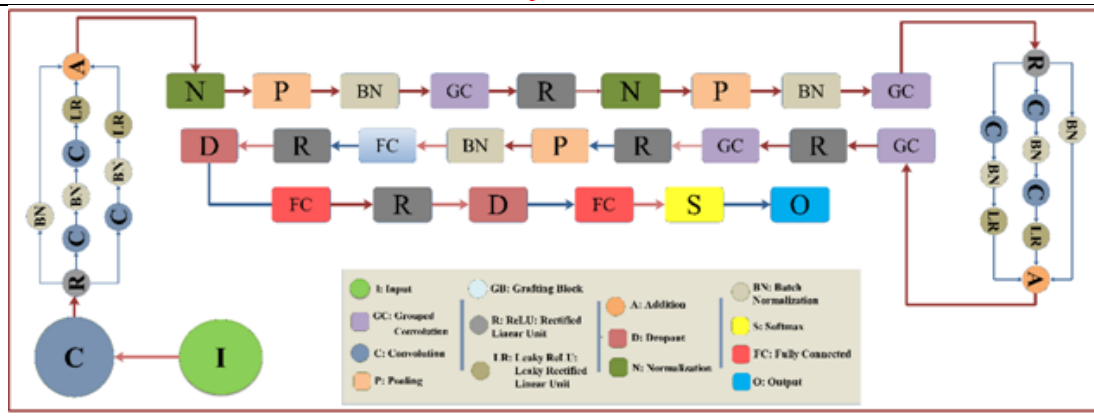


Figure 4. The Architecture of the Proposed MudNetModel

Resilient Network (Resnet):

The resilient network i.e. Resnet101, is a pre-trained network of convolutional neural networks, consists of 101 deep layers. In total, 347 layers are used to build this architecture. The input image size is 224×224×3 and the convolution, pooling, batch normalization (BN), and Relu layer are used in the network. The convolution layer is very important component of the CCN pipeline to detect specific patterns. The pooling layer is used for subsampling to control overfitting. In the network, BN and Relu layers are the most common layers. The average pooling layer has 2048 features in resnet101, followed by a fully connected layer known as FC_1000. The equation 5 representing the MudNet model is shown below, where " f_{Res}" and "f_{sr}"denotes Resnet model and selected features. The architecture of the Resnet101 model is shown in Figure 5.

$$f_{Res} = \{f_{11} + f_{12} + \ldots\ldots\ldots\ldots + f_{sr}\} (6)$$

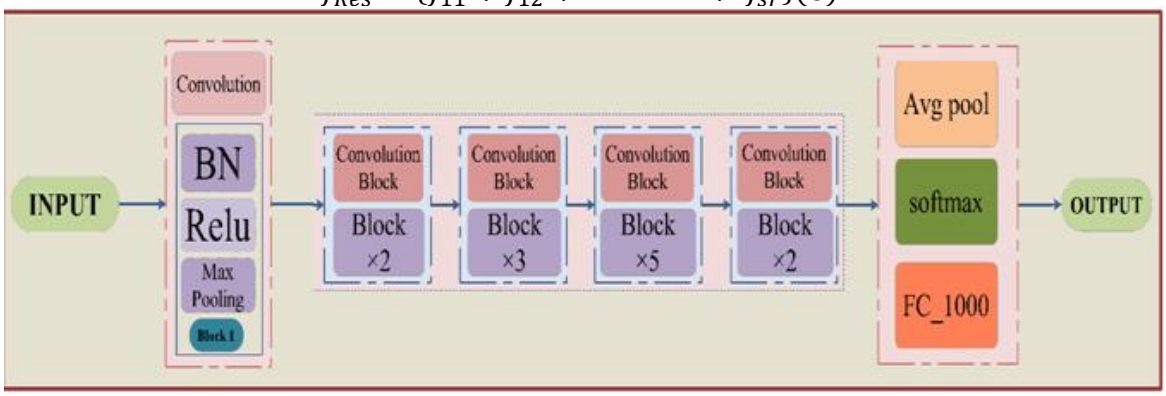


Figure 5. The Architecture of the Resnet101 Model

Feature Fusion:

In this phase, the features corresponding to the proposed models are integrated together to accurately classify and locate the objects. Such features are called feature fusion. The feature fusion summarizes all the features and serially **concatenates** them. The expression for the feature fusion (f_{fus}) is below:

$$f_{fus} = f_{mud} \cap f_{Res} \quad (7)$$

Where, 'f_{mud}' represents extracted features by MudNet architecture and 'f_{Res}' represents extracted features by the resnet101 model. Intersect '∩' is a symbolic representation to collect all optimized features.

Feature Selection:

In this stage, various optimized algorithms are used to select unique features from the fusion result. After fusion, various optimized algorithms are applied over the features which proceed using a pre-trained and proposed model. The selection entropy-based algorithm is

utilized to sort all the extracted features. After some iterations, feature selection algorithms reduced the number of features and provide only sorted features. During this phase for different datasets, features are selected e.g. 1091,1436, and 1943 for further classification process. In our proposed approach, we use Whale optimizer algorithm (WOA) for the feature selection. WOA is a technique used to address the optimization problems which selects the best features for further processing to improve the performance. The mathematical assumption of the WOA in the model is described below.

$$A = |B.\beta_t(k) - \beta(k)| \quad (8)$$

$$\beta(k+1) = \beta_t(k) - \gamma.A \quad (9)$$

Where, 'k' is a sign of the current iteration, 'β_t' represents the best point vector solution. 'B' and 'γ' are the vector coefficients and 'β' is the position of the vector.

$$\gamma = 2c.r - c(10)(10)$$

$$B = 2.c(11)(11)$$

Here, 'c' represents linearly decreased in iteration from (2 – 0) and 'r' is a binary number i.e. either 1 or 0.

Classification:

Classification is the last phase of the proposed method, where the suitable and significant features are executed. The purpose of the classification algorithms is to classify unusual behaviour in the given data. In the proposed method, we apply different variants of SVM, and KNN algorithms to achieve the desired results with high accuracy. Different variants of the KNN classifier include fine KNN[34] medium KNN[38], coarse KNN[39]cosine KNN [40] cubic KNN [41], and weighted KNN [42] algorithms. Likewise, different types of the SVM classifiers are linear SVM [43], cubic SVM [44] quadratic SVM [45] fine gaussian SVM [46], medium gaussian SVM [47] and coarse gaussian SVM [48] algorithms are used.

Data Augmentation:

Data augmentation technique deals with overfitting problem by generating modified copies of the existing data. Different machine and deep learning approaches are employed to generate such new data points. These approaches synthetically enlarge the data for better model performance with accurate detection of the suspicious activities. There are many types of data augmentation techniques, following are used in our proposed method:

Flipping Image:

Flipping image is one of the types of data augmentation, which is the mirror reversal of the original image. It has two types i.e. vertical and horizontal flipping. In our proposed approach, vertical flipping is used to increase the number of images in the given datasets as shown in the Figure 6.



(a) (b)

Figure 6. (a) The original image and (b) Flipping image

Image Noise:

Image noise is the visual distortion in the given image. Technically speaking, noise is the statistical variation in measurement generated by random processes. Image noise takes the

colour and brightness information of random pixels to create variations in the image. Salt and pepper noise is the most used technique to convert all the image pixels into white and black colours. In the Figure. 7, salt and pepper noise are shown.



(a) (b)

Figure 7. (a) Original image (b) Noise image

The mathematical form of Gaussian image noise is shown in the equation:

$$V(r) = \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-(r-\mu)^2}}{2\sigma^2} \quad (16)$$

Where, V is the Gaussian noise for the pixel value "r", depends on the standard deviation (σ) and mean "μ" which symbolizes the gray level and 'e' is the base of a logarithm. It is the transformed image to create a new form of variation.

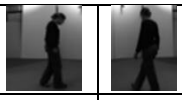
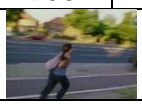
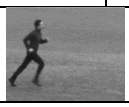
Results and Discussion:

This section provides a detailed study and description of each experiment and analysis of the results. The publicly available datasets are used in this study. Moreover, the best optimization algorithms are used to select the best features to be processed by the classifier in this study. These features are divided into small portions and their comparative analytical results details are mentioned in their respective tables and graphs. All 12 variants of the SVM and KNN classifiers are used to perform classification and highlighted the best accuracy values. The classifier performance is evaluated by using different metrics such as accuracy (acc), precision (prec), sensitivity (sens), specificity, F1 score, training cost (TC), total time (TT), prediction speed (PS) and area under the curve (AOC), ROC, graft and class confusion matrix. The proposed method is trained and tested using the Nvidia GTX 10TO graphical user interface. The results of all experiments are stored in the DESKTOP-HA3D2BC and HPLV201 hardware platforms. Intel(R) Core (TM) i3-4010U CPU @ 1.70GHz, 1700 Mhz, 2 Core(s), 4 Logical Processor(s), and 8 GB RAM power the DESKTOP-HA3D2BC. For coding and classification performance, MATLAB software is used.

Dataset:

Our proposed model is evaluated by using different datasets such as IXMAS action, ANOMOLOUS ACTIVITY and KTH, having 12, 6 and 5 classes respectively. These datasets are publicly available on Kaggle repository. Several researchers have used these datasets for their research work. These datasets are mostly used in the image processing and machine learning for the recognition of anomalous behaviors or actions. Different filter techniques are utilized for data augmentation and balancing the effects in the image to enhance the quality of datasets. The collective data is used in the proposed approach for classification purpose. The resnet101 and MudNet models concatenate their features. The features extracted are 22804×4096, 53094 × 3048 and 13250×304 from the IXMAS, KTH and other datasets respectively. Here 22804, 53094, and 13250 in the datasets IXMAS, KTH and other dataset are represented as the total images. Likewise, 4096, 3048, and 53094 are the features in the datasets IXMAS, KTH and other dataset respectively, acquired from the proposed MudNet and pre-trained resnet101 model. Some samples images are shown in Table 1.

Table 1. Datasets Results

Ref.	Datasets	Classes	Samples	Number of images	Resolution	Results
[49]	IXMAS ACTION	Check Watch	A	1425	256 × 256	99.2%
		Cross Arm	B	1610	256 × 256	
		Getup	C	2033	256 × 256	
		Kick	D	2105	256 × 256	
		Pickup	E	1328	256 × 256	
		Point	F	2140	256 × 256	
		Punch	G	1564	256 × 256	
		Scratch Head	H	2080	256 × 256	
		Sit down	I	1860	256 × 256	
		Turn Around	J	1825	256 × 256	
		Walk	K	3030	256 × 256	
		Wave	L	1747	256 × 256	
								
a	b	c	d	e	f	g	h	i	j	k	l
[50]	ANOMOLOUS ACTIVITY	Falling	A	2526	444 × 256	100%
		Fighting	B	2796	461 × 256	
		Fire	C	2897	461 × 256	
		Firing	D	2746	341 × 256	
		Running	E	2284	341 × 256	
				
a	b	c	d	e
[51]	KTH	Boxing	A	11320	160 × 120	96.6%
		Handclapping	B	10687	160 × 120	
		Handwaving	C	13106	160 × 120	
		Jogging	D	10985	160 × 120	
		Running	E	6996	160 × 120	
				
a	b	c	d	e

Experimentations Brief:

In this section, the detailed experiments are discussed related to the proposed MudNet model. Various test studies are performed for different sets of features and compared to each other. The experiments are performed using the KNN and SVM classifiers with 10- and 5-fold validation techniques. The experiments are divided into three phases and each phase shows their performance's results in Tables and Figures.

The proposed method executes three different datasets. The first experiment works on IXMAS dataset, where 1946 features and 22804 number of images are used. In the second experiment, ANOMOLOUS ACTIVITY dataset is executed, which consists of 1091 features with 13250 images. The third experiment works on KTH dataset which has 1091 features with 13250 total images. All the variants of the SVM and KNN classifiers are used for classification and accordingly unusual behaviour is detected and measured using different evaluation metrics.

Experiment-1 (IXMAS, 22804 Images, WOA):

IXMAS dataset with testing features (feat) 1946, 700, and 200 are used with 5 and 10-fold validation in the first experiment. Each case details are discussed step by step in the Table 2. The experiment is performed over three cases categorized as A, B and C using 200 features and 22804 images. Then 5-fold validation method is used to remove biasness. Additionally, experiment is performed over same dataset with 1946 features and 22804 images using 10-Fold validation. By analysing the experimental results, it is found that fine KNN outperforms all other variants of both SVM and KNN giving detection rate with accuracy of 99.2% using 10-fold validation. Furthermore, metrics evaluation graphical details of the IXMAS dataset are shown in Figure 8. Likewise, class confusion matric detail are shown in the figure 9.

Experiment-2 (Anomalous Activity, 13250 Images, WOA):

Anomalous Activity dataset with testing features (feat) 1091, 700, and 200 and 13250 images with cases A, B and C respectively, are used in the second experiment. Each case details are discussed step by step in the Table 3. The experiment is performed over case A, B and C using 200 features and 22804 images. 5- and 10-folds validation methods are used to deal with biasness. Then the SVM and KNN based classifiers are used to generate classification results. The result detail with evaluation metrics is shown in the Table 3. By analysing the experimental results, it is found that Cubic SVM outperforms all other variant of both the SVM and KNN giving detection rate with accuracy of 100%. Furthermore, metrics evaluation graphical details of the ANOMALOUS ACIVITY dataset are shown in the Figure 10. Likewise, class confusion matric details are shown in Figure.11.

Experiment-3 (KTH, 885 Features, 53094 Images, WOA):

KTH dataset is chosen for first experiment and use different set of features (feat) which are used for testing such 885, and 550 features using 5-fold and 10-fold validation. Step by step each cases of feat are discussed. The experiment performs for case A take 1946 Feat, case B take 550 Feat and total number of images in dataset are 53094. Then 5-fold validation method is used to remove biasness for testing. The SVM and KNN based classifiers are used to gather classification results. Now experiment perform for 10-Fold validation, by choosing 885 features and 53094 Images. Then perform all SVM and KNN based classifiers for classification.

KTH dataset with features (feat) 885, and 550 features with 5 and 10-fold validation are used in the third experiment. Case A and B utilize 1946 and 550 features respectively with 53094 images. The SVM and KNN based classifiers are used to generate classification results. The result detail with evaluation metrics is shown in the Table 4. By analyzing the experimental results, it is found that Cubic SVM outperforms all other variant of both the SVM and KNN giving detection rate with accuracy of 96.6 %. Furthermore, metrics evaluation graphical details of the KTH dataset are shown in the Figure 12. Likewise, class confusion matric details are shown in the Figure 13.

Table 2. IXMAS dataset Results

Dataset	Fold	Case	Feat	Images	Sr no	Classifiers	Acc (%)	Prec	Sens	F1 score
IXMAS	Experiment-1									
	5	A	1946	22804	1	Fine KNN	99.0	0.99	0.99	0.99
					2	Medium KNN	95.0	0.95	0.94	0.95
					3	Cubic KNN	94.5	0.94	0.94	0.93
					4	Cosine KNN	95.2	0.95	0.95	0.94
					5	Weight KNN	96.5	0.96	0.95	0.96
					6	Coarse KNN	74.4	0.72	0.75	0.74
					7	Cubic SVM	98.7	98	0.98	0.99

					8	Linear SVM	96.2	0.96	0.96	0.95
					9	Coarse Gaussian SVM	90.7	0.90	0.91	0.90
					10	Quadratic SVM	98.2	0.98	0.97	0.98
					11	Fine Gaussian SVM	58.3	0.55	0.58	0.58
					12	Medium Gaussian SVM	97.7	0.97	0.98	0.97
		B	700	22804	1	Fine KNN	98.9	0.99	0.98	0.98
					2	Medium KNN	91.9	0.91	0.92	0.91
					3	CubicKNN	91.9	0.91	0.92	0.91
					4	Cosine KNN	92.2	0.92	0.92	0.92
					5	Weight KNN	97.1	0.96	0.97	0.97
					6	Coarse KNN	58.6	0.56	0.58	0.58
					7	Cubic SVM	98.1	0.98	0.98	0.98
					8	Linear SVM	83.3	0.83	0.83	0.83
					9	Coarse Gaussian SVM	63.0	0.63	0.63	0.63
					10	Quadratic SVM	96.6	0.96	0.96	0.96
					11	Fine Gaussian SVM	86.4	0.71	0.94	0.89
					12	Medium Gaussian SVM	94.7	0.95	0.94	0.94
		C	200	22804	**1**	Fine KNN	99.0	0.99	0.99	0.99
					2	Medium KNN	91.6	0.92	0.91	0.91
					3	Cubic KNN	91.7	0.91	0.91	0.91
					4	Cosine KNN	91.9	0.91	0.92	0.91
					5	Weight KNN	97.1	0.96	0.97	0.97
					6	Coarse KNN	58.6	0.56	0.58	0.56
					7	Cubic SVM	98.0	0.98	0.98	0.98
					8	Linear SVM	75.0	0.75	0.75	0.75
					9	Coarse Gaussian SVM	60.5	0.58	0.60	0.58
					10	Quadratic SVM	96.1	0.96	0.96	0.96
					11	Fine Gaussian SVM	86.2	0.78	0.86	0.86
					12	Medium Gaussian SVM	94.2	0.94	0.94	0.94
	Experiment-2									
	10	A	1946	22804	**1**	**Fine KNN**	**99.2**	**0.99**	**0.99**	**0.99**
					2	Medium KNN	95.9	0.95	0.96	0.95
					3	Cubic NN	95.4	0.95	0.95	0.94
					4	Cosine KNN	96.1	0.96	0.96	0.95
					5	Weight KNN	98.0	0.98	0.98	0.98
					6	Coarse KNN	75.7	0.75	0.75	0.75
					7	Cubic SVM	98.8	0.98	0.99	0.98
					8	Linear SVM	96.6	0.96	0.95	0.96

					9	Coarse Gaussian SVM	91.1	0.91	0.91	0.91
					10	Quadratic SVM	98.2	0.98	0.98	0.98
					11	Fine Gaussian SVM	60.8	0.60	0.61	0.60
					12	Medium Gaussian SVM	97.8	0.97	0.98	0.97

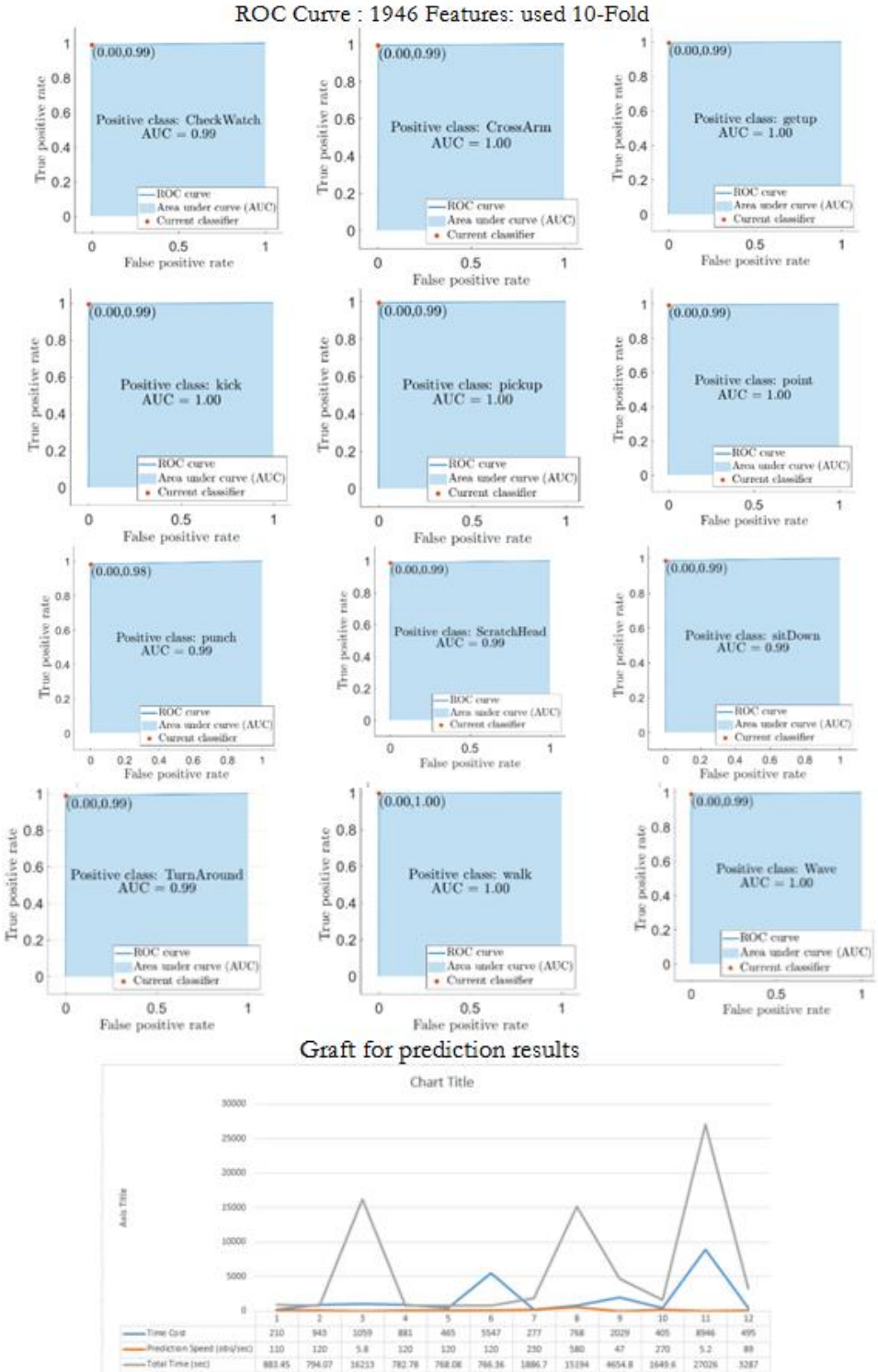


Figure 8. ROC curve and graft results for Fine KNN

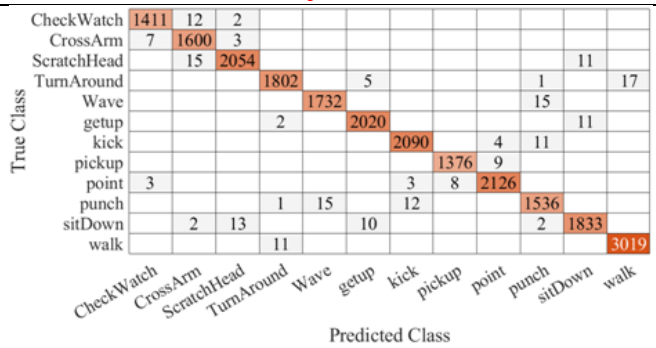


Figure 9. Confusion Matrix for IXMAS action Dataset

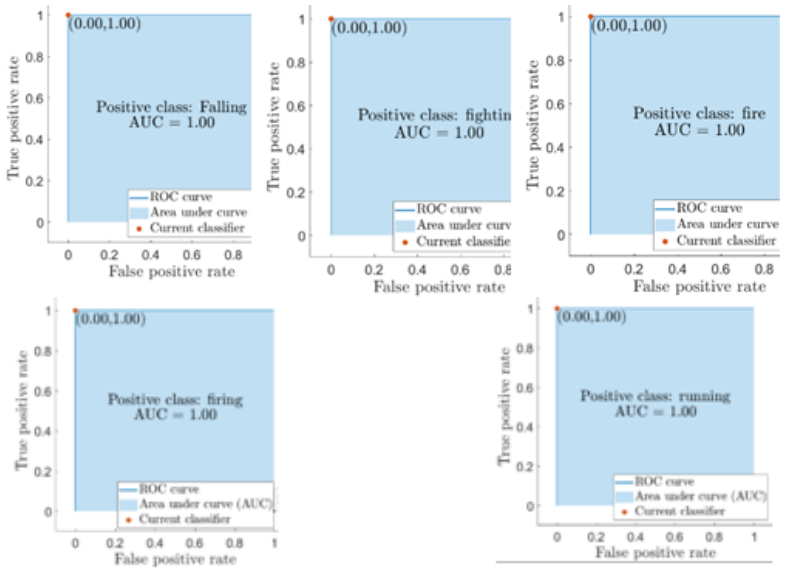
Table 3. Anomalous Activity dataset Results

Dataset	Fold	Case	Feat	Images	Sr no	Classifiers	Acc (%)	Prec	Sens	F1-Score
Anomalous Activity	Experiment-1									
					1	Fine KNN	99.9	0.99	0.99	0.99
					2	Medium KNN	99.4	0.99	0.99	0.99
					3	Cubic KNN	99.4	0.99	0.99	0.99
					4	Cosine KNN	99.6	0.99	0.99	0.99
					5	Weight KNN	99.6	0.99	0.99	0.99
					6	Coarse KNN	93.6	0.93	0.94	0.93
					7	**Cubic SVM**	**100**	**1.0**	**1.0**	**1.0**
					8	Linear SVM	99.1	0.99	0.99	0.99
					9	Coarse Gaussian SVM	97.1	0.97	0.97	0.97
					10	Quadratic SVM	99.9	0.99	0.99	0.99
					11	Fine Gaussian SVM	81.2	0.81	0.81	0.83
					12	Medium Gaussian SVM	100	1.0	1.0	1.0
		B	700	13250	1	Fine KNN	99.2	0.99	0.99	0.99
					2	Medium KNN	96.9	0.97	0.97	0.97
					3	Cubic KNN	97.0	0.97	0.96	0.97
					4	Cosine KNN	97.9	0.98	0.97	0.98
					5	Weight KNN	97.5	0.98	0.97	0.98

					6	Coarse KNN	86.4	0.86	0.87	0.87
					7	Cubic SVM	99.5	0.99	0.99	0.99
					8	Linear SVM	96.3	0.94	0.94	0.94
					9	Coarse Gaussian SVM	89.1	0.83	0.83	0.83
					10	Quadratic SVM	99.2	0.99	0.99	0.99
					11	Fine Gaussian SVM	83.8	0.72	1.0	0.83
					12	Medium Gaussian SVM	99.3	0.99	0.99	0.99
		C	200	13250	1	Fine KNN	99.1	0.99	0.99	0.99
					2	Medium KNN	97.1	0.97	0.98	0.97
					3	Cubic KNN	97.0	0.97	0.97	0.97
					4	Cosine KNN	97.7	0.97	0.97	0.98
					5	Weight KNN	97.8	0.97	0.97	0.98
					6	Coarse KNN	85.8	0.85	0.86	0.85
					7	Cubic SVM	99.2	0.99	0.99	0.99
					8	Linear SVM	92.0	0.92	0.92	0.92
					9	Coarse Gaussian SVM	86.4	0.85	0.86	0.86
					10	Quadratic SVM	98.8	0.98	0.99	0.98
					11	Fine Gaussian SVM	84.0	0.84	0.84	0.84
					12	Medium Gaussian SVM	99.1	0.99	0.99	0.99
	Experiment-2									
	10	A	1946	13250	1	Fine KNN	100	1.0	1.0	1.0
					2	Medium KNN	99.5	0.99	0.99	0.99
					3	Cubic KNN	99.5	0.99	0.99	0.99

					4	Cosine KNN	99.6	0.99	0.99	0.99
					5	Weight KNN	99.7	0.99	0.99	0.99
					6	Coarse KNN	94.3	0.94	0.94	0.94
					7	Cubic SVM	100	1.0	1.0	1.0
					8	Linear SVM	99.3	0.99	0.99	0.99
					9	Coarse Gaussian SVM	97.5	0.98	0.97	0.98
					10	Quadratic SVM	99.	1.0	0.99	0.99
					11	Fine Gaussian SVM	83.0	0.82	0.83	0.84
					12	Medium Gaussian SVM	100	1.0	1.0	1.0

ROC results of Cubic SVM: Case A: 1091 Features: used 5-Fold



Graft for prediction results

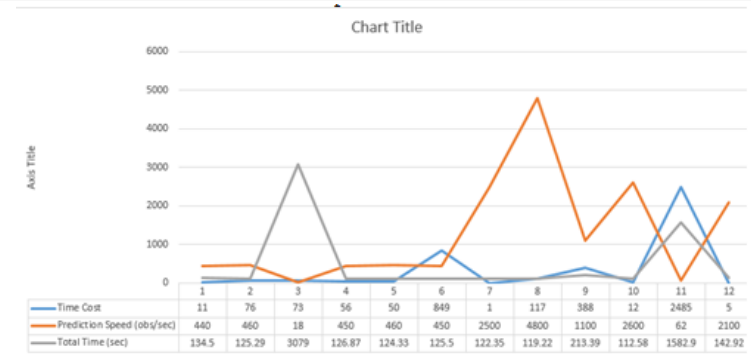


Figure 10. ROC curve and graft results for Cubic SVM

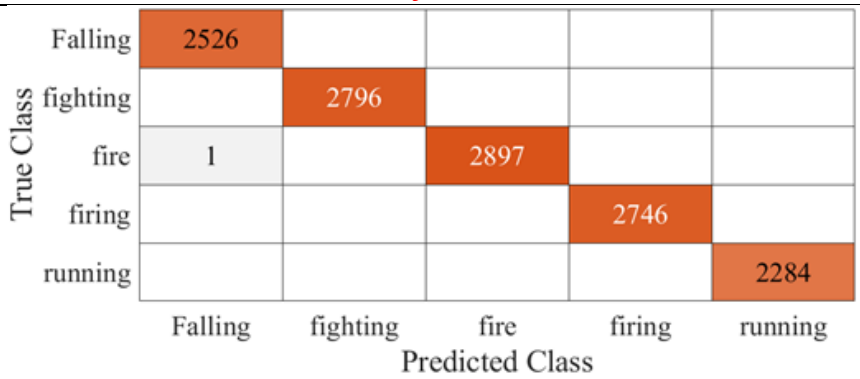


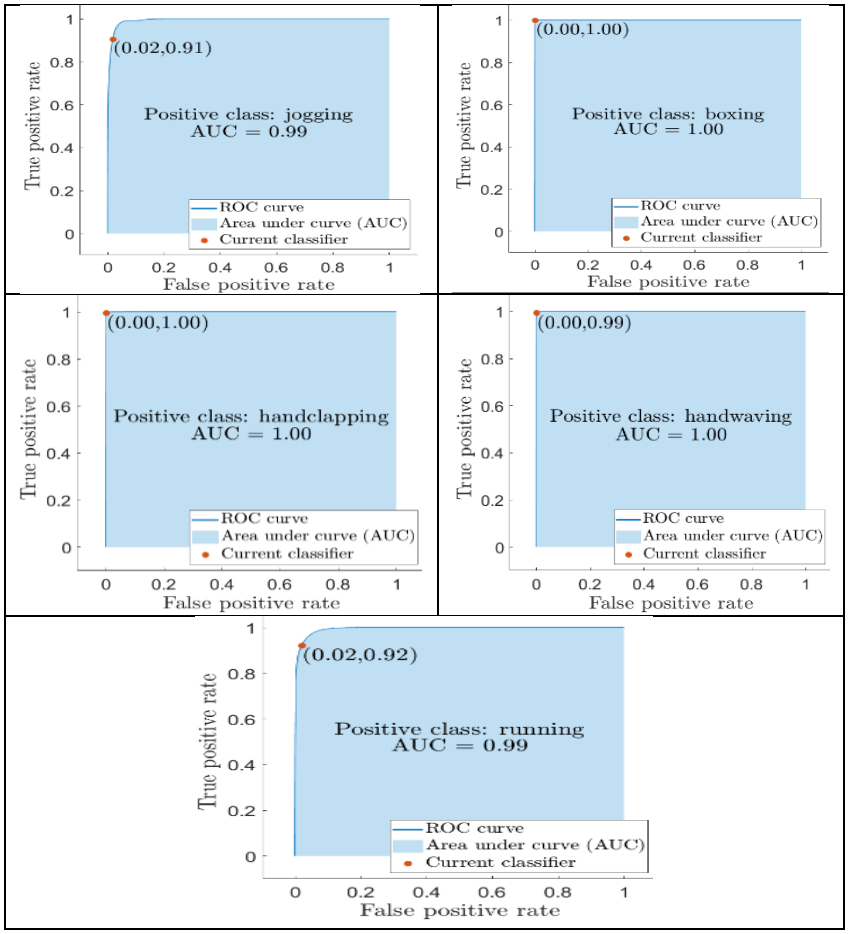
Figure 11. Confusion matrix for ANOMOLOUS ACTIVITY Dataset

Table 4. KTH dataset experiments Results.

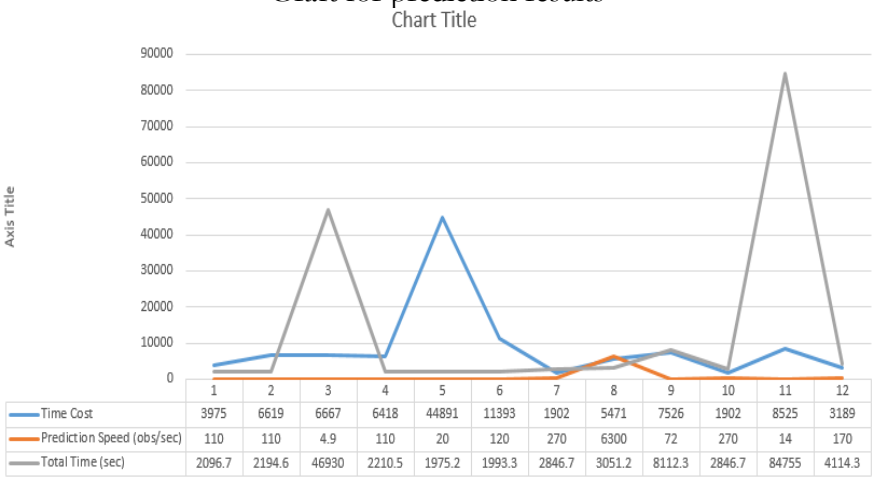
Dataset	Fold	Case	Feat	Images	Sr no	Classifiers	Acc(%)	Prec	Sens	F1 score
KTH	Experiment Phase 1									
					1	Fine KNN	92.2	0.92	0.92	0.91
					2	Medium KNN	87.8	0.87	0.87	0.85
					3	Cubic KNN	87.8	0.87	0.85	0.86
					4	Cosine KNN	88.4	0.88	0.87	0.88
					5	Weight KNN	91.3	0.91	0.92	0.91
					6	Coarse KNN	78.9	0.78	0.78	0.79
					7	Cubic SVM	96.4	0.96	0.96	0.95
					8	Linear SVM	90.1	0.90	0.89	0.90
					9	Coarse Gaussian SVM	86.1	0.85	0.86	0.85
					10	Quadratic SVM	95.1	0.95	0.96	0.95
					11	Fine Gaussian SVM	83.6	0.83	0.84	0.84
					12	Medium Gaussian SVM	94.1	0.93	0.94	0.94
	B	550	53094	B	1	Fine KNN	91.5	0.91	0.91	0.91
					2	Medium KNN	86.3	0.85	0.86	0.85
					3	Cubic KNN	86.4	0.86	0.86	0.85

					4	Cosine KNN	86.6	0.86	0.86	0.85
					5	Weight KNN	90.3	0.89	0.90	0.89
					6	Coarse KNN	76.7	0.85	0.75	0.76
					7	Cubic SVM	94.4	0.94	0.94	0.94
					8	Linear SVM	84.3	0.83	0.84	0.84
					9	Coarse Gaussian SVM	79.6	0.79	0.79	0.80
					10	Quadratic SVM	92.5	0.92	0.92	0.92
					11	Fine Gaussian SVM	87.3	0.87	0.90	0.81
					12	Medium Gaussian SVM	91.6	0.91	0.91	0.91
	Experiment Phase 2									
	10	A	885	53094	1	Fine KNN	92.9	0.92	0.92	0.92
					2	Medium KNN	88.1	0.88	0.88	0.88
					3	Cubic KNN	88	0.87	0.88	0.88
					4	Cosine KNN	88.5	0.88	0.88	0.88
					5	Weight KNN	91.9	0.91	0.91	0.91
					6	Coarse KNN	79.5	0.79	0.80	0.79
					7	**Cubic SVM**	**96.6**	**0.96**	**0.96**	**0.96**
					8	Linear SVM	90.2	0.89	0.90	0.90
					9	Coarse Gaussian SVM	86.5	0.86	0.86	0.86
					10	Quadratic SVM	96.6	0.95	0.96	0.96
					11	Fine Gaussian SVM	84.7	0.84	0.89	0.84
					12	Medium Gaussian SVM	94.3	0.95	0.94	0.94

ROC results of Cubic SVM: Case A: 885 Features: used 10-Fold



Graft for prediction results



	1	2	3	4	5	6	7	8	9	10	11	12
Time Cost	3975	6619	6667	6418	44891	11393	1902	5471	7526	1902	8525	3189
Prediction Speed (obs/sec)	110	110	4.9	110	20	120	270	6300	72	270	14	170
Total Time (sec)	2096.7	2194.6	46930	2210.5	1975.2	1993.3	2846.7	3051.2	8112.3	2846.7	84755	4114.3

Figure 12.ROC curve and graft results results for Cubic SVM

Confusion Matrix:

The number of true positives and true negatives is added up, and the total number of predictions is divided by that number to determine the model's accuracy. For the best outcomes, this section provides a confusion matrix (CM). For the CM displayed in Figure 13, the global topmost value has been chosen.

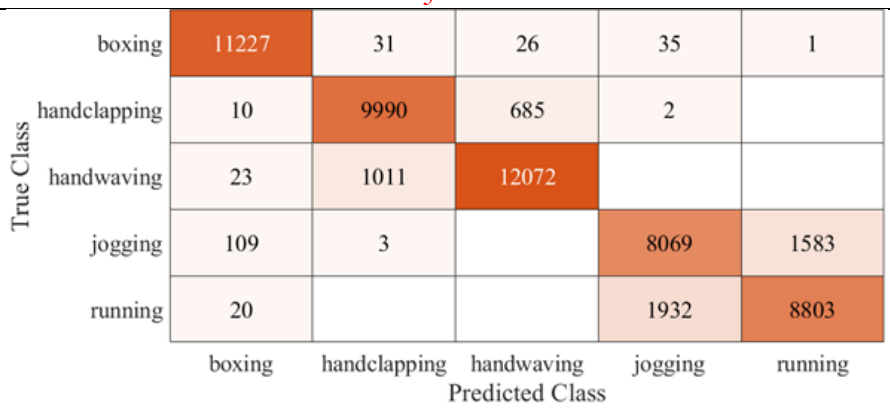


Figure 13. Confusion matrix for KTH dataset

Comparison Results:

Unusual activities have been compared and classified using machine learning. According to studies, machine learning-based abnormality detection methods are more effective than other types of sentiment analysis. The accuracy variation is insignificant, though. The results comparison of the proposed approach with the state-of-the-art approaches is shown in Table 5.

Table 5. Comparison Results

Datasets	Comparison References			Proposed Method	
	Ref.	Method	Accuracy (%)	Method	Accuracy (%)
ANOMOLOUS ACTIVITY	[49]	L4-Branched Action Net	0.99	MudNet	100
IXMAS action	[52]	3D-ResNet-LSTM	94.9	MudNet	99.2

Conclusion:

This study presented an integrated framework for unusual-activity recognition in video-surveillance environments using MudNet, pretrained ResNet101, feature fusion, Whale Optimization Algorithm (WOA)-based feature selection, and multiple KNN and SVM classifiers. The proposed approach was evaluated on three publicly available datasets, namely IXMAS Action, Anomalous Activity, and KTH. Image enhancement and preprocessing were first applied to improve input quality, after which deep features were extracted from MudNet and ResNet101, fused into a combined representation, and optimized using WOA before classification. The framework was assessed using 5-fold and 10-fold cross-validation under different feature-selection settings. The experimental results demonstrated strong classification performance across the three datasets, with the best reported accuracies reaching 99.2% for IXMAS, 100% for Anomalous Activity, and 96.6% for KTH. These results indicate that combining complementary deep features with optimization-based feature selection can support effective recognition of unusual human activities under different surveillance conditions. Future work may therefore focus on larger and more diverse datasets, real-time surveillance testing, cross-dataset evaluation, computational-efficiency improvement, and the integration of newer deep-learning architectures to further improve robustness and generalization.

References:

[1] J. A. Albarakati *et al.*, "Multi-Agent-Based Fault Location and Cyber-Attack Detection in Distribution System," *Energies 2023, Vol. 16, Page 224*, vol. 16, no. 1, p. 224, Dec. 2022, doi: 10.3390/EN16010224.

[2] S. Vosta and K. C. Yow, "A CNN-RNN Combined Structure for Real-World Violence Detection in Surveillance Cameras," *Appl. Sci. 2022, Vol. 12, Page 1021*, vol.

12, no. 3, p. 1021, Jan. 2022, doi: 10.3390/APP12031021.

[3] J. Pincott, P. W. Tien, S. Wei, and J. Kaiser Calautit, "Development and evaluation of a vision-based transfer learning approach for indoor fire and smoke detection," *Build. Serv. Eng. Res. Technol.*, vol. 43, no. 3, pp. 319–332, May 2022, doi: 10.1177/01436244221089445.

[4] X. Peng *et al.*, "Real-Time Illegal Parking Detection Algorithm in Urban Environments," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 20572–20587, Nov. 2022, doi: 10.1109/TITS.2022.3180225.

[5] S. Boukhalfa, A. Amine, and R. M. Hamou, "Border Security and Surveillance System Using IoT," *Int. J. Inf. Retr. Res.*, vol. 12, no. 1, pp. 1–21, Oct. 2022, doi: 10.4018/IJIRR.289953.

[6] S. Han, X. Dong, X. Hao, and S. Miao, "Extracting Objects' Spatial–Temporal Information Based on Surveillance Videos and the Digital Surface Model," *ISPRS Int. J. Geo-Information 2022, Vol. 11, Page 103*, vol. 11, no. 2, p. 103, Feb. 2022, doi: 10.3390/IJGI11020103.

[7] M. Q. Gandapur, "E2E-VSDL: End-to-end video surveillance-based deep learning model to detect and prevent criminal activities," *Image Vis. Comput.*, vol. 123, Jul. 2022, doi: 10.1016/J.IMAVIS.2022.104467.

[8] H. Ghahremannezhad, H. Shi, and C. Liu, "Real-Time Accident Detection in Traffic Surveillance Using Deep Learning," *IST 2022 - IEEE Int. Conf. Imaging Syst. Tech. Proc.*, Aug. 2022, doi: 10.1109/IST55454.2022.9827736.

[9] R. Mathur, T. Chintala, and D. Rajeswari, "Detecting Criminal Activities and Promoting Safety Using Deep Learning," *Proc. - IEEE Int. Conf. Adv. Comput. Commun. Appl. Informatics, ACCAI 2022*, 2022, doi: 10.1109/ACCAI53970.2022.9752619.

[10] M. C. Matheba, P. Kgadima, and K. W. Mabelane, "PARENTING A CHILD WHO IS MISUSING ILLEGAL SUBSTANCES: STRENGTHENING FATHERHOOD," *Soc. Work. W.*, vol. 57, no. 2, pp. 167–167, Jun. 2021, doi: 10.15270/57-2-927.

[11] C. S. Sung and J. Y. Park, "Design of an intelligent video surveillance system for crime prevention: applying deep learning technology," *Multimed. Tools Appl.*, vol. 80, no. 26–27, pp. 34297–34309, Nov. 2021, doi: 10.1007/S11042-021-10809-Z;PAGE:STRING:ARTICLE/CHAPTER.

[12] A. A. Tulbure, A. A. Tulbure, and E. H. Dulf, "A review on modern defect detection models using DCNNs – Deep convolutional neural networks," *J. Adv. Res.*, vol. 35, pp. 33–48, Jan. 2022, doi: 10.1016/J.JARE.2021.03.015.

[13] J. Dawes, M. Brian, H. Bialek, and M. L. Johnston, "Wireless Smartphone Control using Electromyography and Automated Gesture Recognition," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2018, pp. 5390–5393, Jul. 2018, doi: 10.1109/EMBC.2018.8513640.

[14] B. Du, C. Liu, W. Zhou, Z. Hou, and H. Xiong, "Catch me if you can: Detecting pickpocket suspects from large-scale transit records," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, vol. 13-17-August-2016, pp. 87–96, Aug. 2016, doi: 10.1145/2939672.2939687.

[15] S. Kim, M. Hirokawa, A. Funahashi, and K. Suzuki, "Parental Influence in Disengagement during Robot-Assisted Activities: A Case Study of a Parent and Child with Autism Spectrum Disorder," *Multimodal Technol. Interact. 2022, Vol. 6, Page 39*, vol. 6, no. 5, p. 39, May 2022, doi: 10.3390/MTI6050039.

[16] I. H. Sarker, "AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems," *SN Comput. Sci. 2022 32*, vol. 3, no. 2, pp. 158-, Feb. 2022, doi: 10.1007/S42979-022-01043-X.

[17] 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 2021, Vol. 21, Page 2811*, vol. 21, no. 8, p. 2811, Apr. 2021, doi: 10.3390/S21082811.

[18] D. Z. Satybaldina, N. S. Glazyrina, K. A. Kalymova, and V. S. Stepanov, "Development of an algorithm for abnormal human behavior detection in intelligent video surveillance system," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1069, no. 1, p. 012046, Mar. 2021, doi: 10.1088/1757-899X/1069/1/012046.

[19] A. Nadeem, A. Jalal, and K. Kim, "Automatic human posture estimation for sport activity recognition with robust body parts detection and entropy markov model," *Multimed. Tools Appl. 2021 8014*, vol. 80, no. 14, pp. 21465–21498, Mar. 2021, doi: 10.1007/S11042-021-10687-5.

[20] B. Shuai, A. G. Berneshawi, D. Modolo, and J. Tighe, "Multi-Object Tracking with Siamese Track-RCNN," Apr. 2020, Accessed: Sep. 23, 2026. [Online]. Available: https://arxiv.org/pdf/2004.07786

[21] Schmid, Guilhem Chéron; Ivan Laptev; Cordelia, "P-CNN: Pose-Based CNN Features for Action Recognition", doi: 10.1109/ICCV.2015.368.

[22] I. Rodomagoulakis *et al.*, "Multimodal human action recognition in assistive human-robot interaction," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2016-May, pp. 2702–2706, May 2016, doi: 10.1109/ICASSP.2016.7472168.

[23] G. Devineau, F. Moutarde, W. Xi, and J. Yang, "Deep learning for hand gesture recognition on skeletal data," *Proc. - 13th IEEE Int. Conf. Autom. Face Gesture Recognition, FG 2018*, pp. 106–113, Jun. 2018, doi: 10.1109/FG.2018.00025.

[24] E. E. Cust, A. J. Sweeting, K. Ball, and S. Robertson, "Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance," *J. Sports Sci.*, vol. 37, no. 5, pp. 568–600, Mar. 2019, doi: 10.1080/02640414.2018.1521769.

[25] S. N. Boualia and N. E. Ben Amara, "Deep Full-Body HPE for Activity Recognition from RGB Frames Only," *Informatics 2021, Vol. 8, Page 2*, vol. 8, no. 1, p. 2, Jan. 2021, doi: 10.3390/INFORMATICS8010002.

[26] N. N. Hien, D. N. H. Thanh, U. Erkan, and J. M. R. S. Tavares, "Image Noise Removal Method Based on Thresholding and Regularization Techniques," *IEEE Access*, vol. 10, pp. 71584–71597, 2022, doi: 10.1109/ACCESS.2022.3188315.

[27] D. Akiyama and T. Goto, "Improving Image Quality Using Noise Removal Based on Learning Method for Surveillance Camera Images," *LifeTech 2022 - 2022 IEEE 4th Glob. Conf. Life Sci. Technol.*, pp. 325–326, 2022, doi: 10.1109/LIFETECH53646.2022.9754743.

[28] P. Liyanage and P. Fernando, "Suspicious Human Crowd Behaviour Detection - A Transfer Learning Approach," *21st Int. Conf. Adv. ICT Emerg. Reg. ICter 2021 - Proc.*, pp. 63–68, 2021, doi: 10.1109/ICTER53630.2021.9774784.

[29] Y. Hao, J. Li, N. Wang, X. Wang, and X. Gao, "Spatiotemporal consistency-enhanced network for video anomaly detection," *Pattern Recognit.*, vol. 121, p. 108232, Jan. 2022, doi: 10.1016/J.PATCOG.2021.108232.

[30] A. S. Dileep, S. S. Nabilah, S. Sreeju, K. Farhana, and S. Surumy, "Suspicious Human Activity Recognition using 2D Pose Estimation and Convolutional Neural Network," *2022 Int. Conf. Wirel. Commun. Signal Process. Networking, WiSPNET 2022*, pp. 19–23, 2022, doi: 10.1109/WISPNET54241.2022.9767152.

[31] S. Pavithra and B. Muruganantham, "DL Based Automatic Robbery Detection using Video Surveillance in Residential Areas," *Proc. - IEEE Int. Conf. Adv. Comput. Commun. Appl. Informatics, ACCAI 2022*, 2022, doi: 10.1109/ACCAI53970.2022.9752545.

[32] R. Bousmaha, "A novel hybrid Aquila optimizer with Whale optimziation algorithm for global optimization, feature selection and optimizing SVM parameters," *4th Int. Conf. Pattern Anal. Intell. Syst. PAIS 2022 - Proc.*, 2022, doi: 10.1109/PAIS56586.2022.9946891.

[33] D. Esan, P. Owolawi, C. T.-I. I. J. of, and undefined 2022, "Anomalous detection in noisy image frames using cooperative median filtering and KNN," *Res. Esan, PA Owolawi, C TuIAENG Int. J. Comput. Sci. 2022•researchgate.net*, Accessed: Sep. 23, 2026. [Online]. Available: https://www.researchgate.net/profile/Dorcas-Oladayo-Esan/publication/381457985_Anomalous_detection_in_noisy_image_frames_using_cooperative_median_filtering_and_KNN/links/6683ae222aa57f3b8265a57c/Anomalous-detection-in-noisy-image-frames-using-cooperative-median-filtering-and-KNN.pdf

[34] L. CUI *et al.*, "A method for satellite time series anomaly detection based on fast-DTW and improved-KNN," *Chinese J. Aeronaut.*, vol. 36, no. 2, pp. 149–159, Feb. 2023, doi: 10.1016/J.CJA.2022.05.001.

[35] A. Hussain, W. Ullah, N. Khan, Z. A. Khan, M. J. Kim, and Baik, Sung Wook, "TDS-Net: Transformer enhanced dual-stream network for video Anomaly Detection", doi: 10.1016/j.eswa.2024.124846.

[36] W. Ullah, T. Hussain, F. U. M. Ullah, M. Y. Lee, and S. W. Baik, "TransCNN: Hybrid CNN and transformer mechanism for surveillance anomaly detection," *Eng. Appl. Artif. Intell.*, vol. 123, p. 106173, Aug. 2023, doi: 10.1016/J.ENGAPPAI.2023.106173.

[37] "Two-Stream Deep Learning Architecture-Based Human Action Recognition", doi: https://doi.org/10.32604/cmc.2023.028743.

[38] N. Lal, S. Kumar, and G. Kaidan, "A Deep Learning Approach for Anomalous User-Intrusion Detection in Social Media Network System," *Stud. Big Data*, vol. 113, pp. 255–275, 2022, doi: 10.1007/978-3-031-10869-3_14/SAVE-RESEARCH.

[39] R. Wan Nurazwin Syazwani, H. Muhammad Asraf, M. A. Megat Syahirul Amin, and K. A. Nur Dalila, "Automated image identification, detection and fruit counting of top-view pineapple crown using machine learning," *Alexandria Eng. J.*, vol. 61, no. 2, pp. 1265–1276, Feb. 2022, doi: 10.1016/J.AEJ.2021.06.053.

[40] G. Abosamra and H. Oqaibi, "Using Residual Networks and Cosine Distance-Based K-NN Algorithm to Recognize On-Line Signatures," *IEEE Access*, vol. 9, pp. 54962–54977, 2021, doi: 10.1109/ACCESS.2021.3071479.

[41] K. Veena, K. Meena, Y. Teekaraman, R. Kuppusamy, and A. Radhakrishnan, "C SVM Classification and KNN Techniques for Cyber Crime Detection," *Wirel. Commun. Mob. Comput.*, vol. 2022, no. 1, p. 3640017, Jan. 2022, doi: 10.1155/2022/3640017.

[42] L. G. Fahad and S. F. Tahir, "Activity recognition in a smart home using local feature weighting and variants of nearest-neighbors classifiers," *J. Ambient Intell. Humaniz. Comput. 2020 122*, vol. 12, no. 2, pp. 2355–2364, Jul. 2020, doi: 10.1007/S12652-020-02348-6.

[43] A. Kumar, S. Das, V. Tyagi, R. N. Shaw, and A. Ghosh, "Analysis of Classifier Algorithms to Detect Anti-Money Laundering," *Stud. Comput. Intell.*, vol. 950, pp. 143–152, 2021, doi: 10.1007/978-981-16-0407-2_11.

[44] S. Jamian, T. S. Gunawan, M. Kartiwi, R. A. Razak, K. Abdul Kadir, and M. N. Nordin, "Human Activity and Posture Classification using Smartphone Sensors and Matlab Mobile," *Conf. Rec. - IEEE Instrum. Meas. Technol. Conf.*, 2022, doi: 10.1109/I2MTC48687.2022.9806551.

[45] H. Masood and H. Farooq, "Utilizing Spatio Temporal Gait Pattern and Quadratic SVM for Gait Recognition," *Electron. 2022, Vol. 11, Page 2386*, vol. 11, no. 15, p. 2386,

Jul. 2022, doi: 10.3390/ELECTRONICS11152386.

[46] A. Papaleonidas, A. P. Psathas, and L. Iliadis, "High accuracy human activity recognition using machine learning and wearable devices' raw signals," *J. Inf. Telecommun.*, vol. 6, no. 3, pp. 237–253, Jul. 2022, doi: 10.1080/24751839.2021.1987706.

[47] A. Pesin, A. Louzir, and A. Haskou, "A Novel Approach for Radar-Based Human Activity Detection and Classification," *Dig. Tech. Pap. - IEEE Int. Conf. Consum. Electron.*, vol. 2021-January, Jan. 2021, doi: 10.1109/ICCE50685.2021.9427670.

[48] R. Q. Minguez, I. P. Alonso, D. Fernandez-Llorca, and M. A. Sotelo, "Pedestrian Path, Pose, and Intention Prediction Through Gaussian Process Dynamical Models and Pedestrian Activity Recognition," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 5, pp. 1803–1814, May 2019, doi: 10.1109/TITS.2018.2836305.

[49] N. ur R. Malik, S. A. R. Abu-Bakar, U. U. Sheikh, A. Channa, and N. Popescu, "Cascading Pose Features with CNN-LSTM for Multiview Human Action Recognition," *Signals 2023, Vol. 4, Pages 40-55*, vol. 4, no. 1, pp. 40–55, Jan. 2023, doi: 10.3390/SIGNALS4010002.

[50] V. T. Le and Y. G. Kim, "Attention-based residual autoencoder for video anomaly detection," *Appl. Intell. 2022 533*, vol. 53, no. 3, pp. 3240–3254, May 2022, doi: 10.1007/S10489-022-03613-1.

[51] R. Singh, A. Kumar Singh Kushwaha, Chandni, and R. Srivastava, "Recent trends in human activity recognition – A comparative study," *Cogn. Syst. Res.*, vol. 77, pp. 30–44, Jan. 2023, doi: 10.1016/J.COGSYS.2022.10.003.

[52] J. Zhang, F. Bai, J. Zhao, and Z. Song, "Multi-views Action Recognition on 3D ResNet-LSTM Framework," *2021 IEEE 2nd Int. Conf. Big Data, Artif. Intell. Internet Things Eng. ICBAIE 2021*, pp. 289–293, Mar. 2021, doi: 10.1109/ICBAIE52039.2021.9389913.

Copyright © by authors and 50Sea. This work is licensed under Creative Commons Attribution 4.0 International License.