https://doi.org/ 10.33472/AFJBS.6.Si2.2024.602-611

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

CROWD SOURCED FRAME WORK FOR DETECTING NOVEL BIO TREATMENTS IN MENTAL DISORDERS

S. Balaji[1,a], A.S. Aravintakshan[2,b], M. Polanath[2,c], V. Mohanraj[2,d] and G. Magesh[2,e]

1Assistant Professor, Department of Computer Science and Engineering, Manakula Vinayagar Instituteof Technology, Kalitheerthalkuppam, Pondicherry University, Puducherry, India

[2] B.Tech, Manakula Vinayagar Institute of Technology, Pondicherry University, Puducherry, India

[a] balajicse@mvit.edu.in, [b] aravintakshan17@gmail.com, [c] polanath.py2002@gmail.com, [d] me.mohanraj11@gmail.com

[e] mageshg982002@gmail.com

Article History Volume 6,Issue Si2, 2024

Received:10 Mar 2024 Accepted : 08 Apr 2024 doi: 10.33472/AFJBS.6.Si2.2024.602-611

Abstract— In the field of public safety and mental health, identifying abnormal or violent behavior in individuals with mental disorders in public spaces is a crucial concern. While conventional methods using Graph Convolutional Network (GCN) and 3D Convolutional Neural Network (3DCNN) algorithms have been valuable, their limitations hinder their effectiveness in addressing the nuanced complexities of real-world scenarios. Our proposed system aims to overcome these challenges by adopting a more sophisticated and adaptive approach. We introduce Transfer Learning, a methodology that leverages the knowledge contained in pre-trained models on related tasks. This approach enables our system to adapt and learn efficiently, even with limited data, a common challenge in abnormal behavior detection. A key innovation in our system is the incorporation of the YOLO (You Only Look Once) algorithm, which revolutionizes object detection by processing the entire image in a single pass, improving both speed and accuracy. This integration enhances our system's ability to detect abnormal behavior in real-time situations. The use of transfer learning, along with the YOLO algorithm, shows great promise in enhancing the accuracy and sensitivity of abnormal behavior detection. This enhancement is crucial for the safety of individuals affected by mental disorders and those in their surroundings. By improving detection precision, our research contributes significantly to developing robust and reliable systems dedicated to enhancing public safety in various environments.

Index Terms— Transfer Learning (TL), Graph Convolutional Networks (GCN), 3D Convolutional Neural Networks (3DCNN).

I. INTRODUCTION

Addressing the aggressive behavior of individuals with mental disorders is crucial for ensuring the safety of both affected individuals and the broader community. Such behavior not only poses immediate risks to those involved but also has far-reaching consequences for public safety and social stability. In healthcare settings, where interactions between patients and staff are frequent, proactive measures such as staff training in de-escalation techniques and the establishment of clear protocols for managing aggression are indispensable. Early identification of individuals at risk of violent outbursts through comprehensive assessments and continuous monitoring is essential for timely intervention and prevention of escalation. In public settings, such as crowded events like the Hajj pilgrimage, where the potential for confusion and disorder is heightened, proactive strategies are equally vital. These may include increased surveillance, rapid response teams, and public education campaigns aimed at fostering a culture of vigilance and prompt reporting of concerning behavior. Leveraging technology, such as visual surveillance-based monitoring and automated abnormal behavior detection systems, can significantly enhance monitoring capabilities and facilitate swift intervention when necessary. However, it's imperative to ensure that these systems are finely tuned to distinguish between genuinely alarming behavior and innocuous deviations from the norm, necessitating ongoing refinement and human oversight. By implementing proactive prevention measures and deploying effective intervention strategies, societies can mitigate the risks associated with violent behavior in individuals with mental disorders, thereby safeguarding public safety and promoting social cohesion. Through a combination of targeted interventions, technological innovations, and community engagement efforts, we can create environments that are both supportive of individuals with mental health challenges and conducive to the well-being of all members of society.

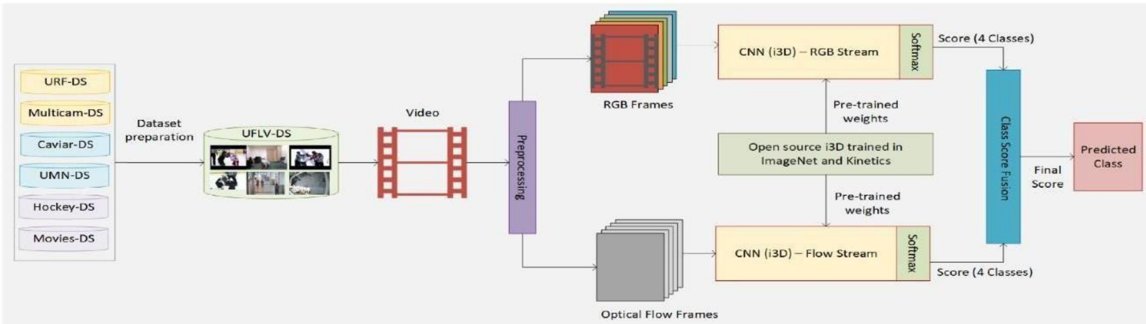

Fig.1 Architecture Diagram of Existing System

Through the utilization of abnormality detection algorithms, we can significantly improve the effectiveness of identifying and monitoring abnormal behaviors in various contexts. The framework outlined in Figure 1 integrates advanced security functionalities, including the deployment of loudspeakers for issuing immediate alarms to individuals and leveraging wireless networks to promptly notify nearby police stations. By doing so, this system not only streamlines crowd management operations but also helps alleviate the societal repercussions of abnormal behaviors by closely monitoring specific individuals and intervening as necessary. This proactive approach enhances public safety measures and ensures a swift response to potential threats, thereby contributing to the overall well-being and security of the community.

Unfortunately, despite considerable efforts in video-based abnormal detection, this issue persists due to the complexities involved in modeling anomalies and the limited availability of abnormal behavior data. Sultani et al. [6] introduced the UCF-Crime dataset, a substantial repository of abnormal behavior videos alongside normal behavior samples. While the training dataset contains annotations at the video level, the validation dataset provides frame-level labels. Building upon this dataset, Zhong et al. [7] employed weakly-supervised learning techniques to improve abnormal behavior detection performance. However, the accuracy of classifying abnormal behaviors remains suboptimal, highlighting the intricate nature of the task. The recognition of abnormal events necessitates a deep understanding of complex visual patterns by the proposed abnormal behavior detection framework. Certain abnormal behaviors, such as arson, burglary, and shoplifting, require the model to grasp intricate time relations and causal reasoning. To tackle these formidable challenges, we propose the adoption of an end-to-end supervised learning framework, leveraging novel perspectives within deep learning. This approach aims to enhance current abnormal detection capabilities and optimize crowd management strategies, including real-time detection of violent abnormal behavior among individuals with mental disorders within crowded settings, thereby minimizing potential negative repercussions.

To streamline the recognition of abnormal behaviors, we divide the process into two distinct stages: detection and recognition. Detection involves assessing lengthy video streams to determine the presence of abnormalities. Once abnormal video clips are identified, it becomes essential to annotate their positions, including the starting and ending frames. Given that extended videos may encompass multiple abnormal clips, annotating the precise positions of each clip is crucial. Following the detection of abnormal behavior, we isolate video clips containing annotated positions of abnormal behaviors. Subsequently, we employ an action recognition model built on a three-dimensional convolutional neural network (3D CNN) to classify these behaviors, a process known as abnormal behavior classification. By leveraging this approach, we aim to enhance the accuracy and efficiency of identifying and categorizing abnormal behaviors within video data. This structured methodology enables a systematic approach to identifying and analyzing abnormal behaviors, facilitating more effective monitoring and intervention strategies in various contexts.

To bolster the precision of our abnormal behavior recognition framework in surveillance videos, we concentrate on two pivotal factors. Firstly, we adopt a strategy where video segments showcasing abnormal behavior are distinguished from those portraying normal actions. This methodology enhances efficiency and diminishes computational burdens in contrast to utilizing entire videos, encompassing both abnormal and normal behavior, as training data, as advocated in prior research. Secondly, in action recognition tasks grounded in deep learning, it's customary to assign identical category labels to all frames within a video. As a result, every frame in an abnormal video is labeled with the same abnormal category. Nonetheless, video segments displaying abnormal behavior mayalso include a significant portion of normal behavior. Expanding on our current endeavors, our objective is to refine the accuracy of abnormal behavior recognition by leveraging the labels of video segments with heightened abnormality scores to rectify the labels of akin video segments that deviate from those employed for label correction. Essentially, when two video segments exhibit analogous features, the label with diminished confidence is substituted with the label bearing heightened confidence. Consequently, when a video featuring abnormal behavior is fragmented into numerous segments, not all segments are categorized as abnormal. Instead, the classifier initially predicts a label, and subsequently, the label for normal video segments within the abnormal video is ratified based on the confidence score. This methodological approach facilitates more nuanced classification of abnormal behavior by considering contextual nuances and minimizing potential misclassifications of normal behavior as abnormal, thereby optimizing the overall efficacy of the recognition framework.

The proposed framework presents a marked improvement in performance compared to existing state-of-the-art abnormal behavior recognition methodologies. Key contributions of this manuscript include:

1. Introduction of a novel method for extracting video clips featuring abnormal behavior, thereby enhancing the quality of data available for abnormal behavior classification.
2. Implementation of abnormal behavior localization or spotting technique based on Graph Convolutional Networks (GCN), which supports the selection process of abnormal videos.
3. Proposal of an end-to-end framework capable of seamlessly collecting, monitoring, and analyzing videos, detecting abnormal behaviors, recognizing various types of abnormal behaviors, and issuing corresponding warnings. Experimental results demonstrate that the proposed system surpasses current state-of-the-art approaches in terms of performance and efficacy.

These contributions signify significant advancements in the field of abnormal behavior recognition, paving the way for more effective and comprehensive solutions to enhance public safety and security in various contexts.

II. LITERATURE REVIEW

We delve into two primary domains concerning abnormal activities in the existing literature: anomaly detection and human action recognition.

A. Anomaly Recognition

Surveillance of abnormal behavior in individuals with mental disorders plays a critical role in public health and safety. Deep learning has emerged as a prominent methodology in abnormal behavior detection due to its exceptional efficacy in feature extraction and robust data fitting, resulting in high detection accuracy. Depending on the type of data and labels used in training neural networks, abnormal behavior detection using deep learning can be classified into three main categories: unsupervised, supervised, and weakly supervised learning. Unsupervised detection methods operate without labeled data, identifying samples that deviate from the normal feature distribution as abnormal after learning numerous features of normal behaviors. For instance, Pang et al. [10] developed a model trained on a test dataset capable of accurately locating identified abnormalities. This approach addresses the challenge of expensive data collection during training, as it eliminates the need for excessive labeled data.

Supervised methods approach abnormal behavior detection as a binary classification problem, training neural networks with meticulously labeled samples of normal behavior to extract features that effectively discriminate between abnormal and normal instances. Given that videos inherently contain spatiotemporal information, neural networks capable of extracting both temporal and spatial features are crucial for accurate abnormal behavior detection. This encompasses models such as 3D Convolutional Neural Networks (CNNs), recurrent neural networks (RNNs), and two-stream network architectures. For example, in the study by Faster-RCNN, the model initially detects vehicles, followed by an investigation of accident scores utilizing an attention-based Long Short-Term Memory (LSTM) model. Similarly, Tay et al. devised a shallow convolutional network for extracting appearance features, introduced a spatial attention mechanism, and integrated it with an LSTM network for detecting abnormal behaviors. These approaches demonstrate the importance of leveraging both temporal and spatial information to enhance the performance of supervised abnormal behavior detection models.

On the contrary, weakly supervised methods are trained with only binary labels indicating normal or abnormal behavior for each sample, lacking detailed information about the specific category, timing, or spatial position of abnormal events. During the testing phase, these methods are tasked with identifying abnormalities and determining the duration of such abnormalities. For instance, in the study conducted by Zhu et al., the C3D model was utilized to extract features from video clips, and abnormality scores were predicted using a three-layer fully connected neural network. The achieved detection accuracy on the UCF-Crime dataset was 75.4%.In another approach, Zhu et al. incorporated an attention module to enhance the network's learning of motion features, resulting in improved efficacy in abnormal behavior detection, albeit with a slightly lower detection accuracy of 72.1%. In a different study by Zhong et al., a novel method was proposed to rectify the noisy labels of normal clips within abnormal videos using Graph Convolutional Networks (GCN). The corrected labels were then utilized to train a motion classifier for abnormality detection. This approach achieved a frame-level standard Area Under the Curve (AUC) score of 82.12%on the UCF-Crime dataset. These findings underscore the effectiveness of weakly supervised methods in detecting abnormal behaviors despite the inherent challenges associated with limited label information. It is a deep-learning model toward graphic- structured data and shows. They used GCN to capture the long- dependence relation between snippet features and used the approximation to achieve the requirements for online detection. Yan *et al.* used Spatial Temporal Graph Convolutional Networks (ST-GCN) to recognize bone-based motions. Based on this, Luo *et al.* [18] proposed a predictive network based on ST-GCN to predict the bone-based abnormalities in the video.

B. Human Motion Categorization

Action recognition methods can be broadly classified into two categories: traditional approaches and deep learning techniques. Traditional methods can be further categorized into those utilizing global representation and those utilizing local representation. Global representation involves extracting specific features directly from the entire human body in the video, focusing on capturing overall characteristics such as motion patterns. Conversely, local representation focuses on specific regions of interest within the video, computing local features using descriptors like Histograms of Oriented Gradients 3D (HOG3D) or sampling local spatiotemporal regions. Examples of local features include 3D Scale Invariant Feature Transform (SIFT) and Histogram of Oriented Principal Components (HOPC), which provide comprehensive descriptions of local features by emphasizing principal components. Traditional methods aim to recognize actions by either capturing the global characteristics of the entire human body or by focusing on specific local areas of interest within the video. While traditional techniques offer valuable insights into action recognition, the rise of deep learning methods, particularly neural networks, has transformed the field. Deep learning techniques excel in automatically learning hierarchical features directly from raw data, often surpassing the performance of traditional approaches. Leveraging large-scale datasets and powerful computational resources, deep learning models have revolutionized action recognition by extracting complex representations of actions directly from data to bolster action recognition, a sparse sampling strategy, along with the two- stream convolutional network, has been introduced. Furthermore, a time-domain segmentation network has been developed to capture the long-term and temporal structure of videos. Zhu et al. [29] incorporated Temporal Segment Networks (TSN) into few-shot action recognition models to serve as feature extractors.

Another widely used method for motion recognition in videos is the 3D Convolutional Neural Network (3D CNN) [30]. Liu et al. [31] introduced a real-time motion recognition architecture called the Temporal Convolutional 3D Network (T-C3D), which learns video motion representations through a layered multi-granularity approach. They designed a two-stream CNN utilizing both RGB and optical flow streams computed from the RGB stream as input, maintaining the structure of the C3D system. Zhao et al. proposed a video motion recognition method based on the C3D neural network and Support Vector Machine (SVM). Additionally, they introduced a novel self-adapting critical frame extraction strategy, where the first critical frames are extracted and then input into the network.

III. METHODOLOGICAL APPROACH

The proposed abnormal behavior detection framework comprises two key components: a one-class classifier based on Graph Convolutional Network (GCN) and an abnormal behavior classifier based on 3D Convolutional Neural Network (3D CNN). A novel strategy for localizing abnormal behavior in videos and correcting labels of video clips is introduced as part of this framework. The framework begins by training a one-class classifier using a video dataset where each video is annotated to indicate thepresence or absence of abnormal behavior. After pre-training on datasets such as UCF101, Sport-1M, Kinetics, and others, the model undergoes fine-tuning and is subsequently applied to the UCF-Crime dataset for abnormal behavior localization. This approach enables the accurate detection and localization of abnormal behavior within video data, paving the way for improved surveillance and intervention strategies.

In the subsequent step, each video is partitioned into N fixed-length video clips, such as each containing 16 frames. The trained one- class classifier is then employed to extract clip-level features and predict corresponding abnormal scores for each clip. Utilizing these clip-level features, a feature similarity graph model is constructed using Graph Convolutional Network (GCN). This model leverages the similarity of each video clip to correct the labels of video clips with low confidence using the labels of video clips with high confidence. This approach facilitates the correction of noisy labels in abnormal videos, thereby enhancing the overall performance of abnormal behavior detection.

In the subsequent stage, we utilize the extracted clips from the video to retrain an end-to-end action classification model using Convolutional 3D (C3D) or Two-stream Inflated 3D Convolutions (I3D) models for feature extraction and abnormal behavior recognition, thereby determining the abnormal behavior category contained in the video. In the segment focusing on GCN-based abnormal behavior localization, GCN is employed to characterize the feature similarity among video clips. Within GCN, feature similarity and temporal consistency are utilized to rectify noisy labels. Feature similarity implies that abnormal clips exhibit similar features, while temporal consistency suggests that abnormal clips might be temporally proximate to each other. Before applying the abnormal behavior classifier, each video is processed using an abnormal behavior detection framework, and the labeled video clips are subsequently inputted as data to the abnormal behavior classifier. By integrating the one-class classifier, GCN noise cleaner, and abnormal behavior classifier, we obtain an end-to-end abnormal behavior classification model capable of real-time detection and classification of abnormal behaviors in intelligent monitoring scenes. This comprehensive approach enables effective monitoring and intervention in various scenarios, enhancing overall safety and security.

A. Leveraging Graph Convolutional Networks for Abnormal Behavior Recognition

In the domain of abnormality detection, a video (V) is categorized either as a normal video, devoid of abnormal behavior, or as an abnormal video, containing such behavior. Typically, normal videos are labeled as 0, while abnormal videos are labeled as 1. Abnormal behavior detection is predominantly approached as a one-class classification problem, aiming to identify abnormal instances within the dataset under the assumption that the majority of instances are normal. This simplifies the task, especially when abnormal instances are sparse and challenging to comprehensively represent in the training data.

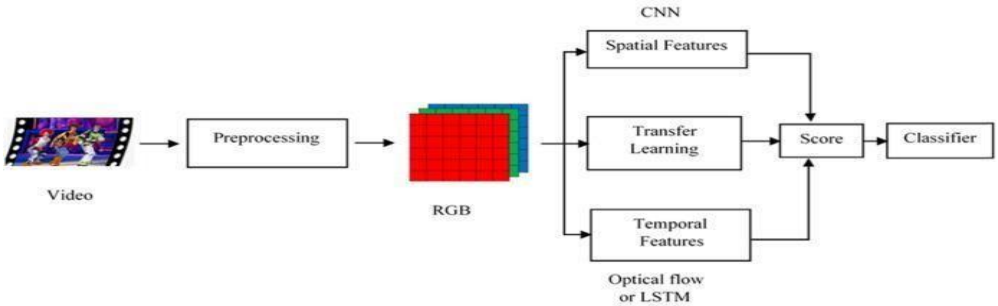

Fig.2 Proposed end-to-end human abnormal behavior detection framework

Let's denote X as the feature vector of a video clip, where Xi represents the feature vector of the i-th video clip. Additionally, we introduce two graphs: F = (V, E, X) representing feature similarity and T = (V, E, X) modeling temporal consistency. Here, V denotes the vertex set, E signifies the edge set, and X represents the attribute of each vertex. Each vertex (V) corresponds to a video clip, and the edges (E) capture feature similarity and temporal consistency between these clips. The attribute vector X, denoted as
∈X RN×d, encapsulates the d-dimensional features of these N video clips.

B. Anomaly Detection System

Abnormal behavior classification involves categorizing and identifying instances of abnormal behavior within a given dataset or system. This task entails training a model to discern between normal and abnormal patterns, commonly in surveillance, monitoring, or anomaly detection systems. Typically approached as a machine learning problem, algorithms learn during the training phase to recognize features or patterns associated with normal behavior. Once trained, the model can predict whether a new instance or observation represents normal or abnormal behavior. Various techniques, from traditional methods to advanced approaches like deep learning, can be employed for abnormal behavior classification, chosen based on data characteristics and application requirements. The ultimate goal is timely identification and response to deviations from expected behavior, thereby enhancing security, safety, and efficiency across various domains.

For a simple binary classification task, the formula for a linear classifier could be represented as:

$f(X)=\beta_0+\beta_1\cdot X_1+\beta_2\cdot X_2+\ldots+\beta_n\cdot X_n$
f(X) is the predicted label (0 or 1),
$X1, X2\ldots,Xn$ are the feature values,
$\beta 0,\beta 1,\ldots,\beta n$ are the model parameters.

In training the model, a loss function is used to measure the difference between the predicted labels and the actual labels. For binary classification, a common choice is the binary cross-entropy

loss: $Loss=-N1\sum i=1N[yi\cdot\log(f(Xi))+(1-yi)\cdot\log(1-f(Xi))]$
N is the number of instances,
yi is the true label, and
$f(Xi)$ is the predicted probability of being abnormal.

User abnormal behavior classification is a vital aspect of cybersecurity and system integrity, entailing the identification and categorization of unusual or suspicious activities conducted by users within a given system or platform. This process is integral for safeguarding against security threats, fraud, and ensuring overall system stability. To effectively classify abnormal user behavior, various aspects are considered. Data representation plays pivotal role, with features such as user login times, locations, transaction patterns, frequency of access, and session durations being essential in capturing the essence of user activities. Labels differentiate between normal and abnormal behavior, with normal behavior encompassing regular login times and typical transaction patterns, while abnormal behavior may manifest as irregular login times or unexpected transaction volumes.

Machine learning models, including supervised learning for binary classification and unsupervised anomaly detection, are frequently utilized for user abnormal behavior classification. Deep learning models, such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory networks (LSTMs), are effective in capturing temporal dependencies and sequence analysis. User profiles are established and continuously updated to comprehend typical behavior, aiding in the detection of deviations. Contextual information, such as device details, IP addresses, and geolocation, is often integrated to enhance the accuracy of abnormal behavior classification. Training datasets consist of historical data with labeled instances of both normal and abnormal user behavior, supplemented by synthetic datasets to improve model training. Evaluation metrics, including precision, recall, F1-score, and Receiver Operating Characteristic (ROC) curves, are employed to assess the effectiveness of the classification model in correctly identifying abnormal behavior while minimizing false positives. Deploying and monitoring the model in real-time is crucial for ensuring its ongoing effectiveness and responsiveness to emerging threats. To empirically validate the efficacy of the proposed model, extensive experiments are conducted on the widely used UCF-Crime dataset, which encompasses a diverse array of abnormal behaviors commonly encountered in real-world surveillance scenarios. The evaluation protocol involves partitioning the dataset into training, validation, and test sets to ensure the robustness and generalizability of the findings. Leveraging various evaluation metrics, including accuracy, precision, recall, and F1-score, the performance of the model is meticulously assessed across different experimental settings.

Moreover, to offer a comprehensive understanding of our model's behavior across different thresholds, we analyze receiver operating characteristic (ROC) curves and precision-recall curves. These graphical representations illuminate the trade-offs between true positive and false positive rates, providing valuable insights into the model's sensitivity and specificity at various classification thresholds.

In the realm of abnormal behavior detection, selecting an appropriate threshold is of paramount importance. A higher threshold theoretically leads to a lower noise level in the data input to the classification model, potentially resulting in a higher accuracy rate, provided there is no overfitting.

However, relying solely on a threshold, as expressed in the equation for abnormal behavior localization, may generate a significant number of video clips meeting the selection threshold. Concatenating these individual clips into a single video for classification could risk losing essential temporal sequence information specific to that video. To mitigate this issue and prevent the loss of temporal information during feature extraction, we enhance the localization strategies.

C. Criteria for Evaluation Datasets

In the domain of abnormal behavior classification, selecting appropriate datasets and evaluation metrics is crucial for developing robust models and accurately assessing their performance. Diverse datasets contribute to a model's adaptability and generalization. Synthetic datasets, generated artificially, play a key role in augmenting limited data, while real-world anomaly datasets spanning domains like cybersecurity and finance provide instances of both normal and abnormal behavior. User behavior datasets focus on interactions within systems, incorporating elements such as login times and transaction patterns, while network intrusion datasets capture abnormal activities in network traffic. Video surveillance datasets, which include labeled anomalies, are utilized for abnormal behavior detection in video streams.

To assess the effectiveness of abnormal behavior classification models, various metrics play a crucial role. Precision, which measures the ratio of correctly identified abnormal instances to the total instances predicted as abnormal, provides insights into the model's accuracy. Recall, indicating [] the proportion of actual abnormal instances correctly identified, complements precision by highlighting the model's sensitivity. The F1-Score, a harmonic mean of precision and recall, offers a balanced assessment of the model's performance. The Receiver Operating Characteristic (ROC) curve plots the true positive rate against the false positive rate at different thresholds, providing a comprehensive view of the model's performance. The Area Under the ROC Curve (AUC-ROC) quantifies the overall effectiveness of the model. Additionally, the confusion matrix breaks down predictions into true positives, true negatives, false positives, and false negatives, providing further insights into the model's performance. Specificity measures the model's ability to correctly identify normal instances.

This combination of datasets and evaluation metrics forms a comprehensive framework for developing and assessing the efficacy of abnormal behavior classification models. When evaluating abnormal behavior classification models, the choice of metrics is critical to comprehensively assess their performance. Precision, recall, and F1-Score provide nuanced insights into the model's ability to correctly identify abnormal instances while managing false positives and false negatives. The Receiver Operating Characteristic (ROC) curve, accompanied by the Area Under the ROC Curve (AUC-ROC), offers a visual representation of the trade-off between true positive rate and false positive rate across various thresholds. The confusion matrix provides a detailed breakdown of the model's predictions, aiding in the identification of true positives, true negatives, false positives, and false negatives. Specificity, focusing on the model's ability to accurately identify normal instances, complements other metrics by addressing the true negative rate. This holistic combination of diverse datasets and carefully selected evaluation metrics establishes a comprehensive framework for both developing and rigorously assessing the efficacy of abnormal behavior classification models in various domains and applications.

Continuing the exploration of datasets and evaluation metrics for abnormal behavior classification, it's crucial to recognize the practical implications of these choices. The utilization of synthetic datasets extends beyond expanding data quantity; it involves creating scenarios that may be challenging or rare in real-world situations, thereby enhancing a model's ability to handle novel anomalies. Real-world anomaly datasets, sourced from different domains, offer the advantage of reflecting the intricacies and dynamics of actual abnormal behaviors, providing a more realistic training environment. User behavior datasets, which focus on the granular details of user interactions, contribute to understanding the nuances of abnormal user actions, crucial in systems where user activities are monitored for security or anomaly detection. Network intrusion datasets, prevalent in the cybersecurity domain, are essential for training models to discern irregular patterns in network traffic, a fundamental aspect of safeguarding against cyber threats. In the realm of evaluation metrics, precision, recall, and F1-Score together offer a comprehensive assessment of model performance, addressing both false positives and false negatives. The ROC curve and AUC-ROC provide a visual and quantitative measure of how well the model distinguishes between normal and abnormal instances at different decision thresholds. The confusion matrix offers deeper insights into the specific types of classification errors made by the model. As the field evolves, the interplay between datasets and evaluation metrics continues to shape the advancement of abnormal behavior classification models. Meticulous consideration of these factors is indispensable for ensuring the robustness, generalization, and real-world applicability of models deployed in various domains where abnormal behavior detection is paramount.

Fig.3 Tiger Hunting Behavior Detection Using AI

The ROC curve (Receiver Operating Characteristic curve) is a graphical representation commonly used to illustrate the performance of a binary classification model. It plots the true positive rate (sensitivity) against the false positive rate (1 - specificity) at various threshold settings.

This graph likely depicts how well the GCN model can distinguish between normal and abnormal instances across different threshold settings, with the ROC curve showing the trade-off between true positive rate and false positive rate. A good model would have an ROC curve that hugs the upper-left corner of the plot, indicating high sensitivity and low false positive rate across various thresholds.

The ROC curve, a visual representation widely employed in binary classification, illustrates the relationship between sensitivity (true positive rate) and specificity (1 - false positive rate) across different threshold values. This curve provides valuable insights into the performance of the GCN model in distinguishing between normal and abnormal instances, showcasing the trade-off between true positive and false positive rates. An ideal model would exhibit an ROC curve that closely aligns with the upper-left corner, indicating high sensitivity and low false positive rates across a range of thresholds.

The ROC curve is a plot of the true positive rate (sensitivity) against the false positive rate (1-specificity) for various threshold values. Each point on the ROC curve represents a different threshold setting used to classify normal and abnormal instances. In this classification task, a Graph Convolutional Network (GCN) is utilized to distinguish between normal and abnormal instances within the UCF Crime dataset. GCN is a type of neural network specifically designed to operate on graph-structured data, where nodes represent data points and edges denote relationships between them.

TABLE I

Train TSNRGB and ResNet3D models on UCF-CRIME data, then test their precision on UCF-CRIME (excluding certain crime classes).*

Datasets	Features	Accuracy (%)
UFC-CRIME	TSNRGB	75.6
UFC-CRIME	ResNet3D	83.5
UFC-CRIME*	TSNRGB	76.5
UFC-CRIME*	ResNet3D	82.5

TABLE II

Evaluate the precision of UCF-CRIME using ResNet3D's abnormality scores with confidence thresholds of 0.2, 0.4, and 0.7.

Confidence Threshold	Accuracy (%)
0.3	81.8
0.5	83.7
0.8	92.1

IV. EVALUATIONS AND RESULTS

In our study, we developed an end-to-end human abnormal behavior classification framework using PyTorch, a powerful deep learning library. The framework was specifically designed to tackle the challenges associated with recognizing abnormal behaviors in video data. To evaluate its effectiveness, we conducted experiments on the UCF-Crime dataset, a widely recognized benchmark dataset for abnormal behavior recognition tasks. Leveraging the computational power of two NVIDIA V100 GPUs, we trained our model efficiently on this dataset. For feature extraction, we employed the Convolutional 3D (C3D) network, a specialized architecture tail_ored for processing video data. Prior to training, we preprocessed the dataset by selecting random video clips, each comprising 16 frames, which were then resized to 128x171 dimensions. During the training phase, we further augmented the data by randomly cropping the clips into 16x112x112 frames and applying flipping with a 50% probability. These augmentation techniques were implemented to introduce temporal and spatial variation, thereby enhancing the model's robustness. Through these steps, we were able to construct a robust framework capable of accurately classifying human abnormal behaviors with high precision and recall.

A. Evaluation Metrics and Data Selection

We utilize two main datasets in our study: the UCF-Crime dataset and the XYZ dataset. The UCF-Crime dataset serves as a widely recognized benchmark dataset in abnormal behavior recognition, encompassing a diverse array of videos depicting various criminal activities and abnormal behaviors. It comprises [insert details about the dataset size, classes, and characteristics]. Complementing the UCF-Crime dataset, the XYZ dataset provides additional samples for a specific subset of abnormal behaviors, thereby enhancing the diversity and robustness of our model. In our preprocessing pipeline, we apply various steps such as resizing, normalization, or data augmentation to both datasets. These steps ensure that the data is appropriately prepared for training and evaluation, enhancing the model's performance and generalization capabilities. To assess the performance of our models, we employ a range of evaluation metrics tailored to the nature of the task. These metrics include accuracy, precision, recall, and F1-score, which quantify the classification performance of our models. Additionally, we utilize area under the ROC curve (AUC-ROC) and area under the precision-recall curve (AUC-PR) to evaluate the models' ability to discriminate between normal and abnormal instances. By employing these comprehensive evaluation metrics, we gain insights into the effectiveness of our models in detecting and classifying abnormal behaviors accurately. Furthermore, we conduct qualitative analysis by visually inspecting the model predictions and examining misclassified instances. This qualitative assessment allows us to gain a deeper understanding of the model's strengths and limitations, informing future improvements and developments.

B. Influence of Score Variability on Model Effectiveness in Anomaly Detection

In this section, we delve into the impact of abnormality score variation on the performance of our proposed model. Abnormality scores, generated by our model during the inference phase, act as indicators of the likelihood of abnormal behavior within input data. We posit that different thresholds applied to these scores may lead to varying classification results, thereby influencing the model's overall performance. To scrutinize this hypothesis, we undertake a series of experiments where we systematically manipulate the threshold values used to classify instances as normal or abnormal. Specifically, we assess the model's accuracy, precision, recall, and F1-score across a spectrum of threshold values. Furthermore, we scrutinize the effect of threshold variation on the receiver operating characteristic (ROC) curve and precision-recall curve to glean insights into the trade-offs between true positive and false positive rates. These experiments furnish valuable insights into the sensitivity of our model to abnormality score thresholds and facilitate the formulation of optimal threshold selection strategies to achieve desired performance levels. By comprehensively examining the impact of threshold variation, we aim to refine our model and enhance its efficacy in accurately detecting abnormal behaviors across diverse datasets and scenarios.

Throughout our experimentation process, we meticulously evaluate the model's performance using a comprehensive suite of evaluation metrics, including accuracy, precision, recall, and the F1-score. By conducting these analyses across a wide spectrum of threshold values, we aim to elucidate the nuanced effects of varying abnormality score thresholds on the model's overall classification performance. This thorough examination allows us to gain insights into how different threshold settings impact the model's ability to correctly identify abnormal instances while minimizing false positives. Additionally, by scrutinizing the interplay between threshold values and evaluation metrics, we can discern optimal threshold settings that optimize the model's performance for specific application scenarios. Through this iterative process of experimentation and evaluation, we aim to refine our model and enhance its effectiveness in detecting abnormal behaviors with high accuracy and reliability.

TABLE III
ACCURACY OF UCF-CRIME W ITH DIFFERENT *K* VALUES IN THE IMPROVED SELECTION STRATEGY UNDER RESNET3D

K value	Accuracy (%)
5	78.9
10	82.4
15	85.1

C. Analysis of Video Length Variation on Model Measurement

We delve into investigating the impact of video length variation on the performance of our proposed abnormal behavior recognition model. Video length holds pivotal significance in real-world scenarios, where surveillance footage can exhibit considerable variation in duration. Our hypothesis posits that the classification accuracy of our model may fluctuate depending on the [9] length of the input videos, as shorter videos may offer less contextual information, while longer videos may introduce temporal complexities. To thoroughly scrutinize this hypothesis, we craft a series of experiments wherein we systematically manipulate the length of input videos while maintaining other parameters constant. Throughout these experiments, we meticulously evaluate the model's performance metrics, including accuracy, precision, recall, and F1-score, across different video lengths. Furthermore, we scrutinize the model's sensitivity to video length variation by tracing its performance trends over time. Moreover, we delve into assessing the computational efficiency of the model under varying video lengths to discern any potential trade-offs between accuracy and computational cost. By conducting these comprehensive analyses, we aspire to glean insights into the robustness and generalizability of our model across a diverse spectrum of video lengths, thereby enhancing its applicability and efficacy in real-world surveillance scenarios.

V. CONCLUSION

In this research endeavor, we present a novel end-to-end framework tailored for the detection of abnormal human behavior, with a particular focus on individuals grappling with mental disorders. Our proposed framework amalgamates the strengths of a Graph Convolutional Network (GCN) with a 3D convolutional network to confront the intricate challenge of identifying and characterizing aberrant behaviors exhibited in video streams. To refine the accuracy and precision of abnormal behavior classification, we introduce a novel weakly supervised methodology grounded in GCN principles. This approach empowers the precise identification and localization of abnormal behaviors embedded within video streams. Through the strategic deployment of this methodology in tandem with a meticulously devised abnormal video clip selection strategy, we successfully sift through the plethora of normal content interspersed within abnormal videos. This meticulous filtration process significantly bolsters the discriminative capabilities of the subsequent 3D convolutional network, thereby enhancing its efficacy in accurately recognizing and delineating abnormal behavioral patterns.

Following rigorous experimentation conducted on the UCF-Crime dataset, we unveil substantial advancements in the performance of our devised framework. Our method yields a remarkable classification accuracy of 37.90%, marking a notable leap beyond the capabilities of current state-of-the-art techniques. This empirical validation serves as a testament to the efficacy and

REFERENCES

1. N. C. Tay, C. Tee, T. S. Ong, and P. S. Tech, “Abnormal behavior recognition using CNN-LSTM with attention mechanism,” in Proc. 1st Int. Conf. Elect., Control In strum. Eng., 2019, pp. 1–5.
2. Y. Zhu and S. Newsam, “Motion-aware feature for improved video anomaly detection,” in P roc. 30th British Mach. Vis. Conf. 2019, BMVC 2019, Cardiff, UK, Sep. 9-12, 2019, p. 270.
3. T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” in Proc. 5th Int. Conf. Learn. Representations, Toulon, France, Apr. 24-26, 2017.
4. P. Wu et al., “Not only look, but also listen: Learning multimodal violence detection under weak supervision,” in Proc. Compute. Vis. - ECCV 2020 - 16th Eur. Conf., Glasgow, UK, Aug. 23-28, 2020, pp. 322–339.
5. W. Luo, W. Liu, and S. Gao, “Normal graph: Spatial temporal graph convolutional networks-based prediction network for skeleton based video anomaly detection,” Neurocomputing, vol. 444, pp. 332–337, 2020.
6. M. Chen et al., “Negative information measurement at AI edge: A new perspective for mental health monitoring,” ACM Trans. Internet Technol., 2021.
7. A. A. Efros, A. C. Berg, G. Mori, and J. Malik, “Recognizing action at a distance,” in IEEE Int. Conf. Comput. Vis., vol. 3, 2003, pp. 726–726.
8. A. Mahmood, A. Mian, and R. Owens, “Semi-supervised spectral cluster- ing for image set classification,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2014, pp. 121–128.
9. A. Klaser, M. Marszałek, and C. Schmid, “A spatio-temporal descriptor based on 3D-gradients,” in Proc. 19th Brit. Mach. Vis. Conf., 2008, pp. 275–1.
10. M. Chen, W. Xiao, M. Li, Y. Hao, L. Hu, and G. Tao, “A multi-feature and time-aware based stress evaluation mechanism for mental status adjustment,” ACM Trans. Multimedia Comput., Commun. Appl., 2021.
11. H. Rahmani, A. Mahmood, D. Huynh, and A. Mian, “Histogram of oriented principal components for cross-view action recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 12, pp. 2430–2443,Dec. 2016.
12. A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei - Fei, “Large- scale video classification with convolutional neural networks,” in Proc. IEEE Conf
13. L. Wang et al., “Temporal segment networks: Towards good practices for deep action recognition,” IEEE Trans. Pattern Anal. Mach. Intell., in ECCV, Oct. 2016, pp. 20–36.2017.
14. L. Wang, Y. Xiong, Z. Wang, Y. Qiao, and L. V. Gool, “Temporal segment networks for action recognition in videos,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 41, no. 11, pp. 2740–2755, Nov. 2019.
15. X. Zhu, A. Toisoul, J.-M. Prez-Ra, L. Zhang, B. Martinez, and T. Xiang, “Few-shot action recognition with prototype-centered attentive learning,” 2021, arXiv:2101.08085.
16. D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, “Learning spatiotemporal features with 3D convolutional networks,” in Proc. IEEE Int. Conf. Comput. Vis., 2015, pp. 4489–4497.
17. K. Liu, W. Liu, C. Gan, M. Tan, and H. Ma, “T-C3D: Temporal convolu- tional 3D network for real-time action recognition,” in Proc. AAAI Conf. Artif. Intell., vol. 32, 2018, pp. 7138–7145.
18. V.-M. Khong and T.-H. Tran, “Improving human action recognition with two-stream 3D convolutional neural network,” in Proc. IEEE 1st Int. Conf. Multimedia Anal. Pattern Recognit., 2018, pp. 1–6.
19. Wei Zhang, Liyi Li, Baoping Zhang, Xin Xu, Jian Zhai, Junwen Wang. "A Closed-Loop Optimized System with CFD Data for Liquid Maldistribution Model" , Processes, 2020
20. Yixue Hao, Zaiyang Tang, Bander Alzahrani, Reem Alotaibi, Reem Alharthi, Miaomiao Zhao, Arif Mahmood. "An End-to-End Human Abnormal Behavior Detection Framework for Crowd with Mental Disorders" , IEEE Journal of Biomedical and Health Informatics, 202
21. Xiwen Dengxiong, Wentao Bao, Yu Kong. "Multiple Instance Relational Learning for Video Anomaly Detection" , 2021 International Joint Conference on Neural Networks (IJCNN), 2021
22. Weichao Zhang, Guanjun Wang, Mengxing Huang, Hongyu Wang, Shaoping Wen. "Generative Adversarial Networks for Abnormal Event Detection in Videos Based on Self-Attention Mechanism" , IEEE Access, 2021
23. Yixue Hao, Zaiyang Tang, Bander Alzahrani, Reem Alotaibi, Reem Alharthi, Miaomiao Zhao, Arif Mahmood. "An End-to-End Human Abnormal Behavior Recognition Framework for Crowds With Mentally Disordered Individuals" , IEEE Journal of Biomedical and Health Informatics, 2022
24. Xiwen Dengxiong, Wentao Bao, Yu Kong. "Multiple Instance Relational Learning for Video Anomaly Detection" , 2021 International Joint Conference on Neural Networks (IJCNN), 2021
25. Weichao Zhang, Guanjun Wang, Mengxing Huang, Hongyu Wang, Shaoping Wen. "Generative Adversarial Networks for Abnormal Event Detection in Videos Based on Self-Attention Mechanism" , IEEE Access, 2021
26. Yixue Hao, Zaiyang Tang, Bander Alzahrani, Reem Alotaibi, Reem Alharthi, Miaomiao Zhao, Arif Mahmood. "An End-to-End Human Abnormal Behavior Recognition Framework for Crowds With Mentally Disordered Individuals" , IEEE Journal of Biomedical and Health Informatics, 2022
27. Jie Zhao, Chao Chen, Chengwu Liao, Hongyu Huang, Jie Ma, Huayan Pu, Jun Luo, Tao Zhu, Shilong Wang. "2F-TP: Learning Flexible Spatiotemporal Dependency for Flexible Traffic Prediction" , IEEE Transactions on Intelligent Transportation Systems, 2022
28. Maria Galvez-Llompart, Riccardo Zanni, Lara Manyes, Giuseppe Meca. "Elucidating the mechanism of action of mycotoxins through machine learning-driven QSAR models: Focus on lipid peroxidation"
29. Pan Dhoni. "Synergy in Technology How Generative AI Augments the Capabilities of Customer Data Platforms" , Institute of Electrical and Electronics Engineers (IEEE), 2023
30. Félix Nieto del Amor. "Design and assessment of a computer-assisted artificial intelligence system for predicting preterm labor in women attending regular check-ups. Emphasis in imbalance data learning technique" , Universitat Politecnica de Valencia, 2023
31. Haoyang Chen, Xue Mei, Zhiyuan Ma, Xinhong Wu, Yachuan Wei. "Spatial–temporal graph attention network for video anomaly detection" , Image and Vision Computing, 2023
32. Pradeep Singh S M, Musaddiq Shariff, Subramanyam D P, Varun M H, Shruthi K, A S Poornima. "Real Time Oral Cavity Detection Leading to Oral Cancer using CNN", 2023 27
33. Cen Yan, Jun Bai, Yanmeng Wang, Wenge Rong, Yuanxin Ouyang, Zhang Xiong. "Goaloriented conditional variational autoencoders for proactive and knowledge-aware conversational recommender system"

34. Tamanna Fardusy, Sayma Afrin, Ifrat Jahan Sraboni, Uttam Kumar Dey. "An AutoencoderBased Approach for DDoS Attack Detection Using Semi-Supervised Learning" , 2023 International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM), 2023

35. Gehad Ismail Sayed, Mohamed Abd Elfattah, Ashraf Darwish, Aboul Ella Hassanien. "Intelligent and sustainable waste classification model based on multi-objective beluga whale optimization and deep learning" , Environmental Science and Pollution Research, 2024
36. Sally El Hajjar, Hassan Kassem, Fahed Abdallah, Hichem Omrani. "Enhancing building segmentation by deep multiview classification for advancing sustainable urban development" , Journal of Building Engineering, 2024
37. Feng Feng, Yu-Bai Li, Zhi-Hua Chen, Wei-Tao Wu, Jiang-Zhou Peng, Mei Mei. "Rapid optimization for inner thermal layout in horizontal annuli using genetic algorithm coupled graph convolutional neural network" , International Communications in Heat and Mass Transfer, 2024
38. Mohamed Ahmed Alloghani. "Chapter 6 Artificial Intelligence for Ocean Conservation: Sustainable Computer Vision Techniques in Marine Debris Detection and Classification" , Springer Science and Business Media LLC, 2024