

Video Anomaly Detection in Crime Analysis using Deep learning Architecture- A survey

G. Sivakumar¹, G. Mogesh², N. Pragatheeswaran³,

T. Sambathkumar⁴

¹Professor and Head, Department of Computer Science and Engineering, Perundurai, Erode, Tamilnadu, India.

^{2,3,4}Student, Department of Computer Science and Engineering, Perundurai, Erode, Tamilnadu, India.

Email: ¹sivakumarganapathi@gmail.com, ²gmogesh2003@gmail.com, ³pragathees1208@gmail.com,

⁴sambakumart95@gmail.com.

Abstract

The importance of developing automated video surveillance systems for public safety and security, particularly in crime analysis, has witnessed significant growth in recent years. This survey delves into the current landscape of automated video surveillance systems, emphasizing advancements in crime analysis and exploring existing methodologies and technologies. The study underscores the significance of employing deep learning models in video analysis. Furthermore, the study suggests a deep learning architecture to address the challenges of the existing methods. The goal of the suggested approach is to help security and law enforcement organizations quickly react to any dangers by precisely identifying unusual occurrences or actions in video sequences. The DenseNet-121 architecture is used for efficient spatial and temporal data acquisition from the video frames. This architecture is characterized by a dense connection structure in which all levels get feature mappings from all layers before them. The characteristics of DenseNet-121 can help in the accurate identification of anomalies in video streams and differentiate between normal and abnormal actions. In addition, the study also delves into the topic of using a cell structure with varied sizes to effectively split video

sequences. This allows for flexible analysis and can accommodate different sorts of abnormalities. Anomaly detection accuracy can be further improved by adding size, motion, and location information to prediction and measurement models. This study serves as a foundation for the future research that aims to develop a more robust and efficient automated video surveillance solutions.

Keywords: Video Anomaly Detection, Public Safety, Security, Crime Analysis, Densenet-121.

1. Introduction

In an era characterized by the widespread presence of surveillance cameras and the continuous growth of video data, the necessity for effective video analysis has never been more significant. One crucial application within the field of video analytics is known as "Video Anomaly Detection." This technology plays a vital role in ensuring the protection of critical infrastructure, public spaces, and various environments where constant monitoring and early threat detection are of utmost importance. Video anomaly detection represents a sophisticated branch of computer vision and machine learning, intended to automatically identify unusual or anomalous events within a video stream.



Figure 1. Video Anomaly Detection [11]

In recent times, the development of automated video surveillance systems (VSSs) has become imperative in ensuring the security and safety of the people, particularly during events involving large gatherings, such as conferences, talent shows, street party, exhibition etc. While artificial intelligence (AI) facilitates computers to emulate human thinking, machine learning (ML) and deep learning (DL) play a more significant role by incorporating training and

learning components. DL algorithms necessitate data labeling and high-performance computers to effectively analyze and comprehend surveillance data obtained from fixed or mobile cameras installed in both indoor and outdoor environments. However, these algorithms may not perform as expected, require extensive training time, or lack sufficient input data for effective generalization. To address these challenges, deep transfer learning (DTL) and deep domain adaptation (DDA) have recently emerged as promising solutions.

So, the proposed study presents a review on the existing deep learning and machine learning models and suggests a video anomaly detection framework using DenseNet-121 transfer learning. The following Figure.2 shows the general block diagram of the DenseNet-121 architecture.

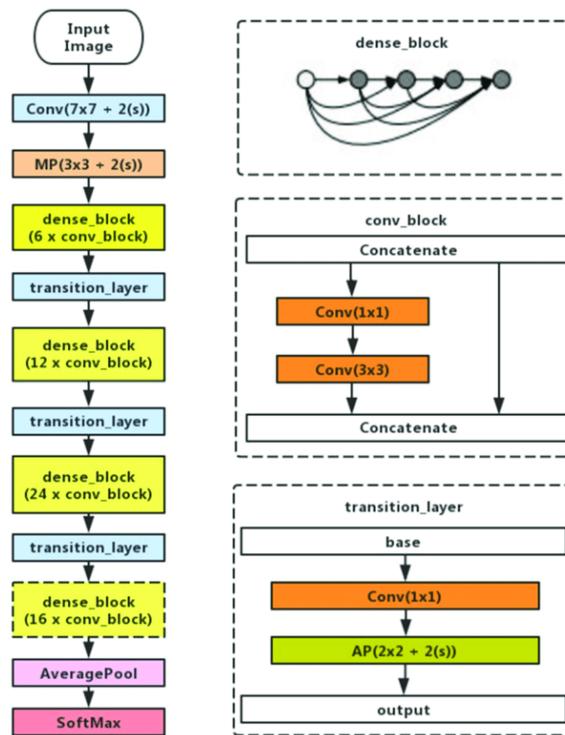


Figure 2. Block Diagram of DenseNet-121 [12]

2. Literature Review

One approach proposed by Nina Tuluptceva et.al [1]. introduces a comprehensive method for image anomaly detection that combines the power of GANs, perceptual loss functions, and a novel similarity metric. Their approach demonstrates significant

improvements in detecting anomalies in images and holds promise for various real-world applications where anomaly detection is critical.

The framework proposed by Davide Abati et.al [2]. addresses the inherent challenges of novelty detection by formalizing a model that captures the twofold nature of novelties. The model acknowledges the incapability to remember unseen data and the surprise aroused by the observation of their latent representations. This dual perspective is crucial for robust novelty detection, as it allows the model to generalize well to unseen data while also detecting deviations from learned patterns. The core of the proposed framework lies in a deep generative autoencoder, augmented with an autoregressive density estimator, which learns the distribution of latent vectors using maximum likelihood principles. By incorporating masked layers tailored for image and video data, the framework demonstrates adaptability to different types of input modalities, further enhancing its versatility. A key contribution of the proposed framework is the introduction of an auxiliary module operating in latent space, aimed at minimizing the encoder's differential entropy. This regularization technique proves effective in improving the model's ability to discern novel instances by encouraging the learned latent representations to capture essential characteristics of the data distribution. Experimental evaluations demonstrate the effectiveness of the proposed framework in various settings, including one-class and anomaly detection scenarios.

The work proposed by R. Tibor Schirrmeister et.al [3]. presents a novel contribution to the field of anomaly detection by introducing two log-likelihood-based metrics tailored for this purpose. In their study, the authors demonstrate the effectiveness of their approach, particularly in the unsupervised setting, where it outperforms existing state-of-the-art methods. In supervised contexts, the method performs competitively but somewhat lower than classifier-based systems. The work focuses on outlier loss to improve anomaly detection by penalizing the model for incorrectly identifying out-of-distribution samples. Hierarchical viewpoints help understand the structure of data, and future research efforts to bring together all of them for better efficiency. In summary, the proposed approach demonstrates strong anomaly detection performance, particularly in challenging unsupervised settings, and provides valuable insights into the interplay between hierarchical distributions and features in generative-model-based anomaly detection. This work contributes to advancing the state-of-the-art in anomaly detection and lays the foundation for future research in this domain.

The proposed method for automatic abnormal event detection in videos, based on particle filtering and optical flow analysis, represents a novel approach in the field of anomaly detection. While traditional methods often rely on reconstructing normal patterns and flagging deviations as anomalies, this approach leverages the predictability of normal events versus the unpredictability of abnormal ones. One such study by Xinwen Gao et.al.[4] explored the predictability of normal events in complex systems, laying the foundation for prediction-based anomaly detection methods. introduced a prediction-based approach to anomaly detection, emphasizing the advantages of leveraging historical data and real-time observations for predicting normal behaviors. the proposed method integrates particle filtering with optical flow analysis to track the L2-norm series of crowd motion, enabling the detection of abnormal events based on significant deviations from predicted values. Furthermore, the proposed method aligns with the growing interest in incorporating spatial feature extractors for more precise anomaly localization. Studies has explored the integration of spatial features to enhance anomaly detection accuracy, laying the groundwork for future advancements in anomaly localization techniques.

WenLiuto et.al [5]. have proposed a system for anomaly detection in videos, which involves identifying events that deviate from expected behavior. However, existing methods primarily focus on minimizing reconstruction errors of training data, which may not necessarily result in a larger reconstruction error for an abnormal event. The study proposes a novel approach to address the anomaly detection problem within a video prediction framework. To the best of our knowledge, this is the first work that utilizes the difference between a predicted future frame and its ground truth to detect an abnormal event. To predict a future frame with higher quality for normal events, we introduce a motion (temporal) constraint in video prediction by enforcing the consistency of optical flow between predicted frames and ground truth frames, in addition to the commonly used appearance (spatial) constraints on intensity and gradient.

Abnormal event detection in video surveillance has gathered a significant attention due to its applications in various domains such as health monitoring, public security, and pedestrian surveillance. The primary challenges in this field revolve around effectively localizing anomalies in both space and time while also managing computational time costs. Furthermore, many existing methods focus solely on utilizing normal event data for training, thereby

overlooking valuable abnormal event information. In response to these challenges, C. He, J. Shao, and J. Sun [6] have proposed novel approaches to improve abnormal event detection. One such approach is anomaly-introduced learning (AL), which integrates abnormal event information into the learning process. This method leverages a graph-based multi-instance learning (MIL) model that incorporates both normal and abnormal video data. By generating a set of potentially abnormal instances and a coarse classifier through the MIL model, AL aims to enhance the detection of abnormal events. An essential component of the AL method is anchor dictionary learning (ADL), an improved dictionary learning technique. ADL utilizes the generated instances to refine the dictionary, which plays a crucial role in detecting abnormalities. The sparse reconstruction cost (SRC) is then employed to quantify the abnormality of detected instances. Compared to existing methods, the proposed AL method offers several advantages. Firstly, it effectively incorporates abnormal event information into the learning process, thereby improving detection accuracy. Secondly, by employing a coarse filter to prune testing instances, the method reduces the computational time required for SRC computation. Experimental results demonstrate the competitive performance of the AL method, highlighting its potential to enhance abnormal event detection in video surveillance applications.

The work by Waqas Sultani et.al [7]. presents a novel approach to anomaly detection in surveillance videos using a deep Multiple Instance Learning (MIL) framework with weakly labeled data. By exploiting both normal and anomalous videos, the proposed method aims to learn a general model of anomaly detection that can effectively identify diverse real-world anomalies. The authors introduce a new large-scale anomaly dataset, which encompasses a wide range of anomalies commonly encountered in surveillance scenarios, thereby providing a comprehensive evaluation platform for their proposed approach. Experimental results demonstrate the superiority of the proposed anomaly detection approach over baseline methods, highlighting its efficacy in detecting various types of anomalies in surveillance videos. Furthermore, the authors showcase the utility of their dataset for the task of anomalous activity recognition, emphasizing its potential to advance research in this domain. The proposed deep learning approach presents a significant advancement in the field of anomaly detection in surveillance videos. By leveraging weakly labeled data and a deep MIL

framework, the method achieves robust performance across a diverse range of real-world anomalies, underscoring its practical applicability in security and surveillance applications.

Several studies in the field have focused on improving anomaly detection in video surveillance through various techniques. Some researchers have explored deep learning approaches to enhance the accuracy and robustness of anomaly detection systems. For example, convolutional neural networks (CNNs) have been widely utilized for feature extraction and classification tasks in video data, showing promising results in detecting anomalous events. Additionally, recurrent neural networks (RNNs) have been employed to capture temporal dependencies in video sequences, enabling more effective anomaly detection over time. However, existing approaches often struggle to adapt to dynamic environments where anomalous behaviors may evolve continuously. This limitation has prompted researchers to explore novel methods that can better handle the changing nature of anomalies in real-time video surveillance. One such approach is the Incremental Spatio-Temporal Learner (ISTL), proposed in the abstract, which aims to address the challenges and limitations of current anomaly detection techniques. ISTL is described as an unsupervised deep learning approach that leverages active learning with fuzzy aggregation to continuously update and distinguish between new anomalies and normality as they evolve. By incorporating incremental learning mechanisms, ISTL aims to provide a more adaptive and robust solution for real-time video surveillance applications. To evaluate the effectiveness of ISTL, the authors Rashmika Nawaratneet.al [8]. conducted experiments using three benchmark datasets, focusing on accuracy, robustness, computational overhead, and contextual indicators. The results of these experiments validate the proposed approach and suggest its suitability for real-time video surveillance applications.

Sharmila Chackravarhi, and Steven Schmitt [9] have proposed a method of securing homes with the rapid and exact detection of unlawful behavior. To achieve this, conventional video monitoring has long been extensively depended upon. This often results in a backlog of video material that has to be glanced over by an official in charge. This leads to an increasing strain on supervisory authorities in major metropolitan areas, which enhances the mistake rate. Workload reduction strategies have been put in place. Though they have a lot of disadvantages, auto-regressive models have been lately applied to forecast criminal occurrences more

effectively. It offers employing neural networks as a cure coupled with a Hybrid Deep Learning system to investigate data from video feeds.

The proposed method, temporally coherent Sparse Coding (TSC) coupled with a stacked Recurrent Neural Network (SRNN), represents a novel approach to anomaly detection in video data. This literature review will explore related works in sparse coding, anomaly detection, and recurrent neural networks, highlighting their contributions and shortcomings, and paving the way for the introduction of the TSC-SRNN model. Sparse coding has been widely employed in anomaly detection due to its ability to effectively represent data using a small number of non-zero coefficients. One prominent approach is the use of dictionary learning techniques, such as dictionary-based sparse coding, where anomalies are detected based on the reconstruction error or sparsity of the representation. For instance, Weixin Luo et al [10]. proposed a dictionary learning-based anomaly detection method that achieved promising results on various datasets. However, traditional sparse coding methods often struggle with capturing temporal dependencies in sequential data like videos. To address this limitation, researchers have explored incorporating temporal coherence into sparse coding frameworks. For example, a Temporally Coherent Sparse Subspace Clustering (TCSSC) method for video anomaly detection, considers both spatial and temporal coherence in the learned representations. While TCSSC showed improvements over traditional sparse coding approaches, it still faced challenges in parameter tuning and computational efficiency. Recurrent Neural Networks (RNNs) have also gained popularity in anomaly detection tasks, particularly for their ability to model sequential data and capture long-term dependencies. For instance, introduced an RNN-based anomaly detection method that achieved state-of-the-art performance on various benchmarks. However, traditional RNN architectures often suffer from vanishing or exploding gradient problems during training, making them challenging to optimize. In light of these existing approaches, the proposed TSC-SRNN method presents several innovations. Firstly, by enforcing temporal coherence in sparse coding through the TSC framework, the model can effectively capture both spatial and temporal information in video data. This enables more robust anomaly detection by considering the contextual information across neighboring frames. Secondly, the integration of TSC with a stacked RNN architecture offers several advantages. By leveraging the capabilities of SRNN to learn parameters simultaneously, the model can avoid the need for manual hyperparameter tuning, thereby

improving ease of use and scalability. Additionally, the shallow structure of SRNN enables efficient inference of reconstruction coefficients within a single forward pass, reducing computational costs compared to traditional RNN architectures.

The Table .1 below presents the key findings of the research.

Table 1. Comparative Table

Ref. No	Methodology	Merits	Demerits
[1]	Combination of Encoder and Generator Networks, Utilization of Generative Adversarial Networks (GANs), Perceptual Loss Function.	State-of-the-art performance, Robustness to Image Variations, Adaptability to Hyperparameter Tuning.	Complexity and Computational Cost, Dependency on Training Data Quality, Interpretability.
[2]	Model Formalization, Deep Generative Autoencoder, Masked Layers for Image and Video Data.	Performance in Novelty Detection Tasks, Comparison with Baseline Methods.	Complexity and Computational Overhead, Generalization to Unseen Data, Data-Dependent Performance.
[3]	Log-Likelihood Based Metrics, Outlier Loss, Hierarchical Views of Distributions and Features.	Outlier Loss, Hierarchical Views.	Performance Trade-offs, Dependency on Training Data, Generalization Issues.
[4]	Construction of L2-norm Series, Particle Filtering (PF), Anomaly Detection.	Accuracy, Precision and Recall, False Alarm Rate.	Computational Complexity, Dependency on Optical Flow.
[5]	Video Prediction Framework, Difference Analysis, Integration of Motion Constraint.	Novel Approach, Enhanced Accuracy, Potential for Real-Time Detection.	Computational Complexity, Dependency on Optical Flow, Generalization.

[6]	Graph-Based Multi-Instance Learning (MIL) Model, Anchor Dictionary Learning (ADL), Sparse Reconstruction Cost (SRC).	Incorporation of Abnormal Event Information, Improved Dictionary Learning, and Reduction of Computational Time.	Sensitivity to Model Parameters, Limited Abnormal Event Representation, Generalization to Diverse Datasets.
[7]	Data Collection and Preprocessing, Weakly Supervised Learning with Multiple Instance Learning (MIL), Deep Learning Model Architecture.	Utilization of Weakly Labeled Data, Exploitation of Normal and Anomalous Videos, Deep MIL Framework.	Scalability, Model Complexity, Dataset Bias.
[8]	Incremental Spatio-Temporal Learner (ISTL), Active Learning with Fuzzy Aggregation, Benchmark Datasets.	Accuracy, Robustness, Contextual Indicators.	Dependence on Training Dataset, Evaluation Metrics.
[9]	Auto-Regressive Models, Hybrid Deep Learning Framework, Deep Convolutional Neural Network (DCNN).	Detection Accuracy, False Positive Rate, Processing Speed.	Complexity and Computational Cost, Data Privacy and Ethical Concerns, Generalization and Adaptability
[10]	Temporally-coherent Sparse Coding (TSC), SRNN Mapping, Dataset Construction.	Enhanced Temporal Coherence, Parameter Optimization, Large-Scale Dataset.	Complexity, Data Dependency.

Although machine learning and deep learning play a more important role by adding training and learning components. To properly evaluate and interpret surveillance data acquired from fixed or mobile cameras positioned in both indoor and outdoor areas, DL algorithms require data labeling and high-performance computing. However, these algorithms may not perform as predicted, necessitate extensive training, or lack adequate input data for effective

generalization. Deep transfer learning (DTL) and deep domain adaptation (DDA) are emerging as possible answers to these difficulties. In real-time video analysis, deep transfer learning outperforms classic deep learning and machine learning methods. Deep transfer learning, as demonstrated by models such as DenseNet-121, outperforms its competitors due to its capacity to apply pre-trained information to massive datasets. This pre-training improves accuracy, allowing for efficient feature extraction and pattern recognition, which is critical for real-time applications. DenseNet-121's densely connected design, combined with transfer learning capabilities, ensures improved accuracy, scalability, and flexibility in reacting to changing visual surroundings. In contrast to standard machine learning, which may require significant labeled data, deep transfer learning performs in settings with low labeled data, which is a common difficulty in real-time video analysis. The model's optimized structure enables real-time detection, extensive analysis, and stable performance, making it a top choice for security and law enforcement applications. The Table.2 below presents the advantages of the suggested method.

Table 2. Advantages of using Densenet-121

Advantages	Real-Time Analysis
Advantages of Using Densenet-121	Enhanced Accuracy
	Real-time Detection
	Comprehensive Analysis
	Scalability and Flexibility
	Robustness to Environmental Variability
	Transfer Learning Capabilities
	Optimized Detection Performance

3. Proposed System

We will present a novel approach for real-time anomaly detection in video sequences, utilizing particle filtering and leveraging posterior probabilities to enhance accuracy. Our proposed method, DenseNet121, which stands for "Densely Connected Convolutional Networks," is distinguished by its dense connectivity pattern. In DenseNet121, each layer

receives feature maps from all preceding layers and transmits its feature maps to all subsequent layers. The objective is to develop a system that not only identifies anomalous video frames but also precisely identifies specific anomalous regions within these frames, enabling meticulous anomaly analysis. To achieve this, we will implement a variable-sized cell structure that efficiently divides scenes in the video, facilitating flexible analysis and accommodating various types of anomalies. Furthermore, we will incorporate size, motion, and location features to enable effective prediction and measurement models, thereby improving the accuracy of anomaly detection. The Figure.3 shows the flow diagram of the proposed.

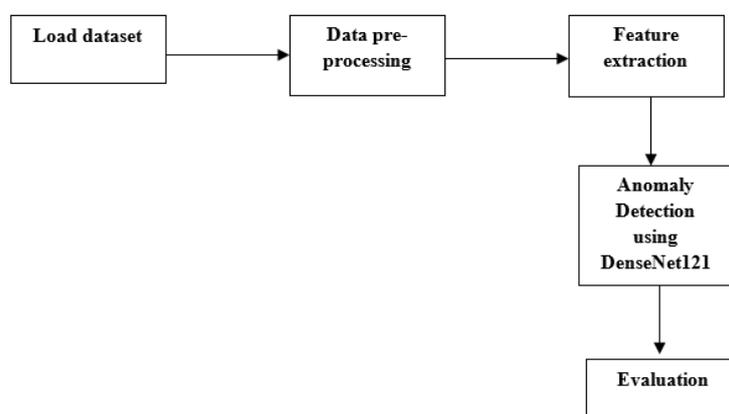


Figure 3. Proposed Flow Diagram

3.1 General Working Procedure

The initial procedure entails the loading of the video dataset. This dataset ought to encompass a diverse range of videos, encompassing both conventional and atypical occurrences. The conventional videos should aptly depict the customary activities anticipated within a surveillance setting. Conversely, the atypical videos should accurately portray the categories of incidents one aims to identify, such as criminal activities, accidents, or any other extraordinary happenings. Upon loading the dataset, it is imperative to undertake pre-processing procedures. These may encompass resizing the frames, standardizing the colors, and eliminating any noise or artifacts. The primary objective of data pre-processing is to guarantee that the frames are in a uniform format that can be effortlessly processed by the feature extraction and anomaly detection modules. The feature extraction module is responsible

for extracting features from the pre-processed frames. These features can be generated through either manual design or acquired through a deep learning model. Hand-crafted features usually rely on the fundamental characteristics of the images, such as color, texture, and motion. The anomaly detection module utilizes the features derived from the feature extraction module to ascertain their normality or abnormality. The ROC curve and evaluation module serves the purpose of assessing the efficacy of the anomaly detection module. The ROC curve illustrates the relationship between the true positive rate (TPR) and the false positive rate (FPR) of the detector. The TPR represents the proportion of anomalous events that are accurately detected, while the FPR represents the proportion of normal events that are erroneously classified as anomalous.

3.2 Particle Filtering

Particle filtering is an effective technique for probabilistic estimates and state tracking, particularly in dynamic systems where conventional models fall short. It excels at dealing with uncertainty and complex probability distributions, as seen in robotics, computer vision, finance, and other applications. In video anomaly detection, Particle Filtering, when paired with DenseNet-121, uses Bayesian estimation to effectively describe temporal dynamics and exploit discriminative deep features for robust solutions.

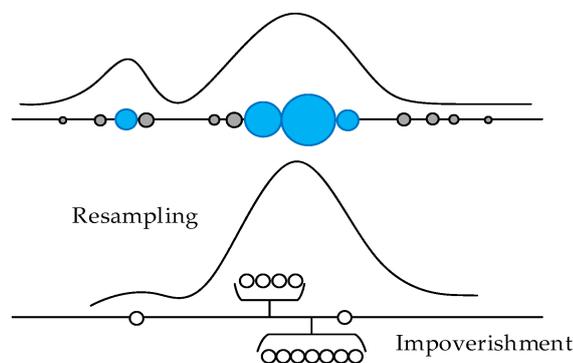


Figure 4. Particle Filtering

3.3 Inference Mechanism

Inference techniques, which are necessary for deduction in AI and human reasoning, play an important role in video anomaly detection with DenseNet-121. This convolutional neural network gathers detailed data from each frame, capturing a variety of visual patterns.

Aggregating these features using approaches such as average pooling reduces dimensionality while keeping important information. The aggregated features are then analyzed by anomaly detection algorithms, which use methods like thresholding and clustering to discover unexpected patterns. Post-processing techniques, such as temporal smoothing and spatial filtering, refine observed anomalies and increase accuracy. Using DenseNet-121's deep features ensures accurate identification of abnormal incidents in video sequences, demonstrating the model's efficiency in anomaly detection. The Table.3 below shows the layer details in the DenseNet-121 architecture.

Table 3. Details of Layers used in Densenet-121

Layer	Type	Number of Images	Size	Convolution Kernel Size
Input Layer	-	1	224x224	-
DenseNet121 Blocks	-	-	-	-
DenseBlock1	Dense Block	-	-	-
	Convolution Layer	64	112x112	7x7
	Pooling Layer	64	56x56	2x2
TransitionLayer1	Transition Layer	64	56x56	-
	Convolution Layer	128	56x56	1x1
	Pooling Layer	128	28x28	2x2
DenseBlock2	Dense Block	-	-	-
	Convolution Layer	256	28x28	7x7
	Pooling Layer	256	14x14	2x2
TransitionLayer2	Transition Layer	256	14x14	-
	Convolution Layer	512	14x14	1x1
	Pooling Layer	512	7x7	2x2
DenseBlock3	Dense Block	-	-	-
	Convolution Layer	1024	7x7	7x7
	Pooling Layer	1024	1x1	2x2
Classifier Layer	Classifier Layer	1024	1x1	-
	Fully Connected Layer	1000	-	-
	Output Layer	11	-	-

4. Discussion

It provides an efficient solution to the complexities and computational costs of deep learning. Its dense connections promote feature reuse, reducing parameters and computational overhead, making it ideal for resource constraints and real-time applications. The model excels at learning from minimal data, adjusting effectively to various data characteristics thanks to its free-flowing information flow. While interpretability remains an issue, tools such as visualization can provide insights into its hierarchical representations. DenseNet-121's high generalization capabilities assure stability over a wide range of datasets. Its consistent performance across diverse datasets, as well as its flexibility in parameter modification, makes it a powerful tool that offers a promising blend of efficiency, performance, and interpretability in anomaly detection.

4.1 Future Work

Future work will include obtaining a broad dataset of both normal and abnormal activity videos to evaluate the suggested method's efficiency. In addition, the model will be tested to determine its predictive accuracy. A comparison with state of art methodologies will be performed, followed by the model's deployment for real-time use.

5. Conclusion

The study gives a brief overview of machine learning and deep learning models for detecting abnormal activity in real-time, as well as a discussion of the challenges that each existing method faces. The survey suggests that deep transfer models are better than deep learning and machine learning models in real-time video analysis because the latter require a high computational resource. The study suggests utilizing the deep transfer learning model DenseNet-121 for real-time video analysis because of its scalability and adaptability to shifting visual environments. The study's future activities will involve dataset collection, design implementation, evaluation, and deployment.

Reference

- [1] Tuluptceva, Nina, Bart Bakker, Irina Fedulova, and Anton Konushin. "Perceptual image anomaly detection." In Asian Conference on Pattern Recognition, pp. 164-178. Cham: Springer International Publishing, 2019.
- [2] Abati, Davide, Angelo Porrello, Simone Calderara, and Rita Cucchiara. "Latent space autoregression for novelty detection." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 481-490. 2019.
- [3] Schirrmester, Robin, Yuxuan Zhou, Tonio Ball, and Dan Zhang. "Understanding anomaly detection with deep invertible networks through hierarchies of distributions and features." *Advances in Neural Information Processing Systems* 33 (2020): 21038-21049.
- [4] Gao, Xinwen, Guoyao Xu, Shuaiqing Li, Yufan Wu, Edvins Dancigs, and Juan Du. "Particle filter-based prediction for anomaly detection in automatic surveillance." *IEEE Access* 7 (2019): 107550-107559.
- [5] Liu, Wen, Weixin Luo, Dongze Lian, and Shenghua Gao. "Future frame prediction for anomaly detection—a new baseline." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6536-6545. 2018."
- [6] He, Chengkun, Jie Shao, and Jiayu Sun. "An anomaly-introduced learning method for abnormal event detection." *Multimedia Tools and Applications* 77 (2018): 29573-29588.
- [7] Sultani, Waqas, Chen Chen, and Mubarak Shah. "Real-world anomaly detection in surveillance videos." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6479-6488. 2018.
- [8] Nawaratne, Rashmika, Damminda Alahakoon, Daswin De Silva, and Xinghuo Yu. "Spatiotemporal anomaly detection using deep learning for real-time video surveillance." *IEEE Transactions on Industrial Informatics* 16, no. 1 (2019): 393-402.

- [9] Chackravathy, Sharmila, Steven Schmitt, and Li Yang. "Intelligent crime anomaly detection in smart cities using deep learning." In 2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC), pp. 399-404. IEEE, 2018.
- [10] Luo, Weixin, Wen Liu, and Shenghua Gao. "A revisit of sparse coding based anomaly detection in stacked rnn framework." In Proceedings of the IEEE international conference on computer vision, pp. 341-349. 2017.
- [11] Khan, Arfat Ahmad, Muhammad Asif Nauman, Muhammad Shoaib, Rashid Jahangir, Roobaea Alroobaea, Majed Alsafyani, Ahmed Binmahfoudh, and Chitapong Wechtaisong. "Crowd Anomaly Detection in Video Frames Using Fine-Tuned AlexNet Model." *Electronics* 11, no. 19 (2022): 3105.
- [12] Ji, Qingge, Jie Huang, Wenjie He, and Yankui Sun. "Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images." *Algorithms* 12, no. 3 (2019): 51.