2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE) | 979-8-3503-0641-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IITCEE59897.2024.10467801

# Deep Learning based Abnormal Event Detection in Pedestrian Pathways

1st Narasimha Chary. S
Research Scholar, *Department of ECE*, Bharath Institute of Higher Education and ResearchChennai, Tamilnadu,
narasimhachary435@gmail.com

2nd Ganesan V
*Professor, Department of ECE,*
*Bharath Institute of Higher Education and Research*
Chennai, Tamilnadu, India
vganesh1711@gmail.com

***Abstract*--A video surveillance system is capable of detecting any kind of motions that are unusual and hence proves to be a major contributor of the surveillance sector. The unusual or strange movements that has been recorded by the surveillance cameras have been used and their patterns have been studied. The pattern that do not match with the patterns of normal behavior are then used captured. In video sequences they can be seen in numerous ways which includes bikers, skaters, small carts in pedestrian pathways, etc. The main aim of the study is to develop a efficiently working Video Anomaly Detection(VAD) system that is capable of identifying and analyzing strange movements in videos using the techniques of deep learning and image processing.**

**Keywords: video surveillance, image processing, VAD, deep learning, patterns**

## I. INTRODUCTION

A complicated field of study in computer vision and machine learning systems, intelligent video surveillance has drawn more attention recently as a result of worries about international security. It's critical to keep an eye out for odd activity in public areas like bus stops, train stations, and retail centers. Due to technological advancements and decreased costs, the usage of surveillance cameras has expanded in both public and private settings. The surveillance of undesirable occurrences is often carried out by human operators. They examine simultaneous movies taken by several cameras.

Their capacity to recognize anomalous occurrences in real time is compromised when they are exposed to extended periods of time without visual stimulation. As a result, the existing system is reduced to a recording system with little utility beyond forensics. Therefore, real-time automated anomaly detection is necessary to identify and take quick action. Therefore, the goal of our study is to create an anomaly system that can identify odd movement in video footage. Here, image processing methods have been combined with deep learning approaches such as convolutional and recurrent neural networks to analyze the input image dataset.

## II. LITEARTURE SURVEY

Anomalies are abnormal events that differs from the pattern of normal behavior. Anomaly identification is the key factor for surveillance. But this often found to be a challenging task because anomaly may be falsely detected in certain areas. For example, a firing a gun is an abnormal event in regular basis but it is a normal event in a gun shooting club. Hence certain events are termed as anomalous despite being a normal behavior according to the place.[1][2]

The WEKA dataset (Waikato Environment for Knowledge Analysis) was used for the investigation of violent crimes and this analysis shows a trend between real criminal data and community data. Additive regression, Linear regression and Decision stump techniques were used in this study and among the three techniques linear regression was able to find the unpredictability of occurrence of the event and shows the ability of deep learning techniques in anomaly detection. [3][4]

In a study by Kim S et al , the anomaly detection in Philadelphia is analyzed and a trend has been identified. The machine learning model is trained for anomaly detection from massive data sources using machine learning techniques as logistic regression, ordinal regression, decision trees, and k-nearest neighbor. The models have achieved a 69% accuracy.[5][6]

In another study Elharrouss et al have used old crime locations dataset to predict the places where crimes are like to happen. The levenerg-marquardt technique was used to analyze and understand the data. In addition to this, scaled technique was also used in the study for data examination and understanding. The scaled technique used was found to be the best performer and showed an accuracy of 78%. It also results in crime reduction up to 78%.[7][8]

Sultani et al. conducted a thorough investigation on anomaly detection in metropolitan areas where data was integrated into a 200x250 m grid. It was clearly analyzed with hindsight. They proposed a model using techniques such as ensemble logistic regression and neural network for anomaly detection. The results conclude that prediction of anomaly is more accurate when performed once in every 14 days compared to performance on monthly basis.[9]

In a study by Rummens et al, the anomaly activities have been thoroughly observed and analyzed with the help of the anomaly data of the past 15 years from right before 2017. The techniques such as k-nearest neighbor and decision tree have been used for the detection. It was able to achieve an accuracy level of 39-44% when used against a dataset consisting so 5,60,000 anomaly activities.[10][11]

## III. METHODOLOGY

Deep learning, a form of machine learning, comprises additional layers called hidden layers in addition to input and output layers. Deep learning is able to replicate any action a human takes on its own. The model is more accurate the more hidden layers there are. A single layer

979-8-3503-0641-5/24/$31.00 ©2024 IEEE

can ultimately be used with a deep learning model, but it won't perform to expectations. [12]

The deep learning techniques of RNN and CNN were applied to build the model in this instance. A massive amount of picture dataset is supplied to the model as input. These photos were taken from a variety of security cameras in different places. The photos in the collection are analyzed using image processing techniques. Because of the way the model is built, it can analyze and recognize patterns. The image dataset is split into train and test data in the initial phase. The model is trained using the train data, and its performance is assessed using the test data. The deep learning techniques such as convolutional neural network (CNN) have been used for model building. The algorithms have been individually tested for their accuracy, performance , etc. and the algorithm with highest accuracy is used in model building.[13][14]

## IV. SYSTEM ARCHITECTURE

The system architecture is presented in fig.1.

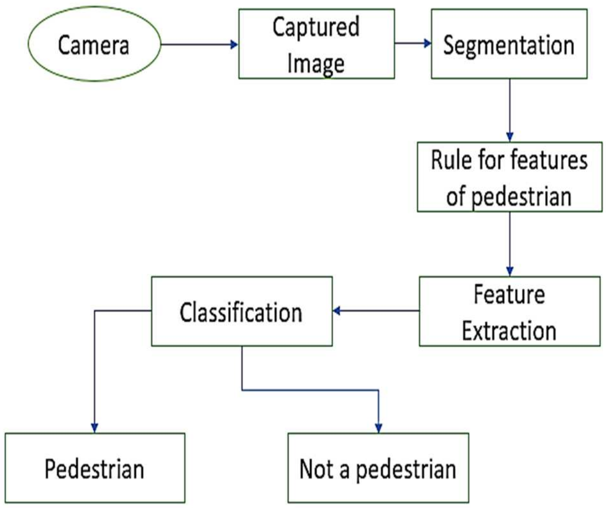


Fig.1. System architecture

The input data come in a variety of forms, including picture, audio, and video. Here, we used preprocessed pedestrian picture data to get rid of any duplicates. After that, the Feature Extraction approach is used. The necessary characteristics are retrieved from the picture at this step. The Feature Selection is then put into practice. Using this method, we pick the most important characteristics from the features that were retrieved in the previous stage. You may now separate the dataset into a test set and a train set. The Deep Learning model is then adjusted to forecast whether or not the input picture contains an anomalous item.[15][16]

## V. IMPLEMENTATION

### A. DATA COLLECTION

Data are of different types and in this method, we use videos for the detection of anomaly in pedestrian pathway and create a better safe journey of the pedestrians for doing this job the dataset used here is from UCSD called UCSD anomaly detection which is collected by mounting cameras in the pedestrian pathways and capturing the video of the pedestrian pathwayBoth ped1 and ped2 folders are part of the dataset. In ped1 and ped2, which comprise 34 training video samples, 36 testing video samples, 16 training video samples, and 12 testing films, respectively, people are shown walking in both directions. Some snippets from the video dataset are shown in fig. 2 and fig. 3.[17]

Fig.2. Captured image

Fig. 3. Captured image

### B. DATA PREPROCESSING

We may utilize object detection and optical flow to identify the anomalies based on their motion, which is the object's speed and appearance. These methods can be used with the Object Detection dataset's video frames.: Object detection helps us to identify the objects from the frames of the video and can detect the objects in the video frame such as pedestrians or an anomaly such as bikes, cars, truck, skaters etc... and it is fundamental to recognize the object in the video frame. Optical Flow: The motion of the objects in the video is determined using optical flow by comparing the consecutive frames it uses brightness constancy assumption this is based on the this is based on the brightness of the pixel, and it can be able to identify fast moving objects in the pedestrian pathway which are the anomalies.[18][19]

### BATCH NORMALIZATION

In this project, we employ the batch normalization approach to speed up the model's learning. In simple terms, by using this layer, we may speed up the input picture training process and make our model simpler to understand. 32 image samples are transmitted through the network at a time when we utilize the batch size of 32 in this model. To boost the network's pace of learning from the training data, we add a batch normalization layer.[20]

C. TRAINING AND TESTING

The dataset is divided into train and test folders, with ped1 and ped2 based on different camera locations. Pred1 has 34 training and 36 testing video samples, while Pred2 has 16 training and 12 testing video samples. Both pred1 and pred2 have binary flags to identify normal and anomalous, and each video clip is converted to 200 frames for model training. Masks are also provided to evaluate the metrics. we select the sets based on the size the training data should be higher than the testing data for better performance of the model.[21]

Any model that is trained on input data must be able to predict the output. The model is now prepared to begin training, during which the retrieved characteristics from the input image of a pedestrian are practiced using the Epoch approach. The term "epoch" refers to how many iterations of the training set should take place in a single cycle. This model was created using 20 epochs, which means that the training data was fed into the model 20 times as shown in fig. 4. Alternatively said, we 20 times trained the data using deep learning techniques. The precision is examined each time. Next, a test is permitted for the trained model. In order to forecast the outcome, the hidden picture is loaded in this case.[22]

```
Epoch: 1/20..  Training Loss: 0.468..  Validation Loss: 0.597..  Validation Accuracy: 0.675
Epoch: 2/20..  Training Loss: 0.378..  Validation Loss: 0.529..  Validation Accuracy: 0.700
Epoch: 2/20..  Training Loss: 0.349..  Validation Loss: 0.543..  Validation Accuracy: 0.738
Epoch: 3/20..  Training Loss: 0.335..  Validation Loss: 0.481..  Validation Accuracy: 0.750
Epoch: 4/20..  Training Loss: 0.328..  Validation Loss: 0.534..  Validation Accuracy: 0.744
Epoch: 4/20..  Training Loss: 0.295..  Validation Loss: 0.545..  Validation Accuracy: 0.744
Epoch: 5/20..  Training Loss: 0.283..  Validation Loss: 0.679..  Validation Accuracy: 0.700
Epoch: 6/20..  Training Loss: 0.326..  Validation Loss: 0.528..  Validation Accuracy: 0.731
Epoch: 6/20..  Training Loss: 0.290..  Validation Loss: 0.513..  Validation Accuracy: 0.769
Epoch: 7/20..  Training Loss: 0.292..  Validation Loss: 0.503..  Validation Accuracy: 0.781
Epoch: 8/20..  Training Loss: 0.277..  Validation Loss: 0.534..  Validation Accuracy: 0.769
Epoch: 8/20..  Training Loss: 0.285..  Validation Loss: 0.447..  Validation Accuracy: 0.769
Epoch: 9/20..  Training Loss: 0.284..  Validation Loss: 0.551..  Validation Accuracy: 0.725
Epoch: 10/20..  Training Loss: 0.270..  Validation Loss: 0.520..  Validation Accuracy: 0.800
Epoch: 10/20..  Training Loss: 0.276..  Validation Loss: 0.407..  Validation Accuracy: 0.806
Epoch: 11/20..  Training Loss: 0.253..  Validation Loss: 0.484..  Validation Accuracy: 0.812
Epoch: 12/20..  Training Loss: 0.265..  Validation Loss: 0.592..  Validation Accuracy: 0.750
Epoch: 12/20..  Training Loss: 0.254..  Validation Loss: 0.596..  Validation Accuracy: 0.787
Epoch: 13/20..  Training Loss: 0.252..  Validation Loss: 0.462..  Validation Accuracy: 0.794
Epoch: 14/20..  Training Loss: 0.238..  Validation Loss: 0.519..  Validation Accuracy: 0.787
Epoch: 14/20..  Training Loss: 0.244..  Validation Loss: 0.432..  Validation Accuracy: 0.806
Epoch: 15/20..  Training Loss: 0.222..  Validation Loss: 0.503..  Validation Accuracy: 0.806
Epoch: 16/20..  Training Loss: 0.238..  Validation Loss: 0.475..  Validation Accuracy: 0.787
Epoch: 16/20..  Training Loss: 0.271..  Validation Loss: 0.419..  Validation Accuracy: 0.806
```

Fig.4. Epoch

Fig. 5 and fig 6 show that the accuracy and loss graph for the Convolution neural network (CNN). In this project, we train 20 epoch in the training phase of the model.

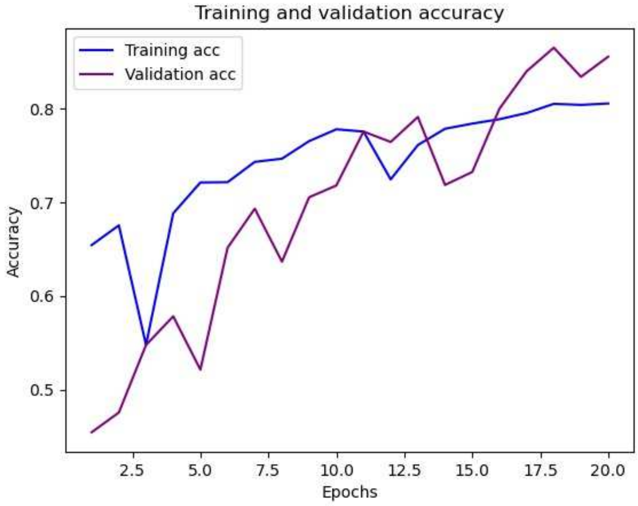


Fig.5. Accuracy

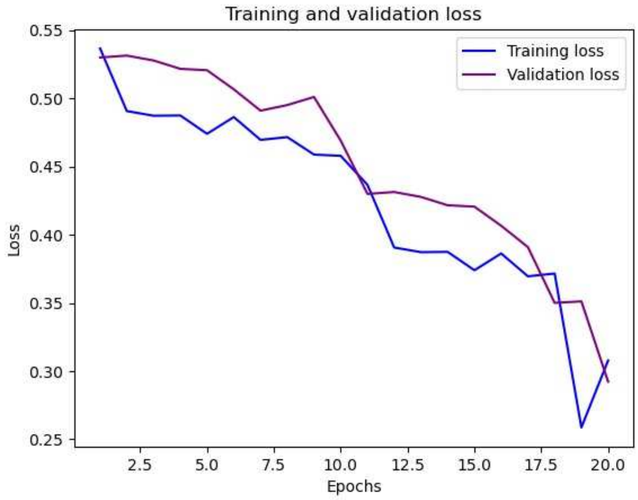


Fig.6. Loss

D. MODEL SELECTION

These models, which are a kind of recurrent neural network, are the convolution neural network and the many to many LSTM. The deep learning approaches are used here because of the ability to extract features and train them with multiple layers and get a better performance the backpropagation of the neural network helps in making a better model and the learning rates can be used for better training of the model

CONVOLUTIONAL NEURAL NETWORK (CNN)

In this work, we use a deep learning method akin to the convolutional neural network (CNN). Constructing a CNN pedestrian anomaly detection model is a challenging task. The following layers are used by convolutional neural networks (CNNs) to more efficiently build deep learning models

KERNAL

The information or characteristics from the input pictures which is taken by the camera are extracted using the Kernal approach and are represented as pixels in a matrix form. In this project, the kernel size is 3x3, which means that the size of the 3x3 matrix is derived from the input pixel matrix of the picture and compared with the kernel matrix, that is also a 3x3 matrix. The two matrices are then multiplied, and the output is a feature extracted matrix. The Kernal size is 3x3 since we used to extend the model.[23]

FLATTEN LAYER

The multi-dimensional array of picture characteristics in this project may be reduced to a single dimension by adding a flattening layer after the convolution layer. The fully connected network receives this flattened array of picture attributes as an input layer, which it uses to determine if the supplied image is cancerous or not. As a consequence, by reducing the size of the features map, adding this layer after the convolution layer would improve the network's overall performance and computing speed.[24]

On the other hand, we utilize the Sigmoid function to determine whether or not the image is an abnormality. The range of the sigmoid function is 0 to 1. This sigmoid function causes the network to become non-linear, enabling it to learn complicated models. Therefore, in this project, we utilize both the Sigmoid function and the rectified Linear unit function. By estimating the values from the preceding layer of neurons in the network, the Adaptive Moment Estimator (Adam) is utilized in this model to automatically change the parameter values of the neural network. With the aid of this optimizer, the learning rate during sample training is improved, and it also improves the model's accuracy.[25][26]

```
Layer (type)                Output Shape              Param #
=================================================================
conv2d (Conv2D)             (None, 254, 254, 16)      448

max_pooling2d (MaxPooling2D  (None, 127, 127, 16)     0
)

conv2d_1 (Conv2D)           (None, 125, 125, 32)      4640

max_pooling2d_1 (MaxPooling  (None, 62, 62, 32)       0
2D)

conv2d_2 (Conv2D)           (None, 60, 60, 16)        4624

max_pooling2d_2 (MaxPooling  (None, 30, 30, 16)       0
2D)

flatten (Flatten)           (None, 14400)             0

dense (Dense)               (None, 256)               3686656

dense_1 (Dense)             (None, 1)                 257
```

Fig.7. Output of CNN

```
for batch in test.as_numpy_iterator():
    X, y = batch
    yhat = model.predict(X)
    pre.update_state(y, yhat)
    re.update_state(y, yhat)
    acc.update_state(y, yhat)
print(pre.result(), re.result(), acc.result())
```

Fig.8. Predicting model

With the aid of training samples, the model in this project predicts if the loaded pedestrian image has any anomalies or not. The highest accuracy score out of all of them is used as the final accuracy rating. Convolution Neural Network (CNN) provides 80% accuracy in this project.

## VI. CONCLUSION

The safety of the people is very important hence anything that could cause any discomforts to the people should need to be checked hence the anomaly is checked These anomalies can be anything that is abnormal in the pedestrian pathway usually these anomalies are bikes, cycles, truck, skatters etc... They can cause harm to the pedestrians in the pathway therefore this code detects the anomalies with the help of Deep learning and image processing techniques by converting video into frames then this frames are preprocessed to extract efficient features from the dataset and the noise were removed from the frames and using motion detection and the object detection it classifies based on the motion of the object travelling and the size of the object other than the, the pedestrians. This model learns from the frames of the videos, and its learning rate is adjusted to make it learn more quickly. The hyperparameters are also adjusted for optimal model performance. Deep learning techniques, such as convolution neural network and recurrent neural network, are used to have high accuracy and to have multiple hidden layers to extract better features and high accuracy. As a result, this model performs in prediction of the anomalies in the pathway with an accuracy of 80%, making it a very good model.,By this we can make better protection for the people using the pathway by detecting the anomalies. This model is learnt based on motion and object detection. This can be extended to more parameters for the betterment of the model. By using other parameters this model can be better and the results would be in high accuracy

## REFERENCES


[1]"Evaluating the use of public surveillance cameras for crime control and prevention." by Vigne, N.G.L., Lowry, S.S., Markman, J.A., Dwyer, A.M. in US Department of Justice, Office of Community Oriented Policing Services. Urban Institute, Justice Policy Center, Washington, DC (2011)

[2] "A world with a billion cameras watching you is just around the corner." by Lin, L., Purnell, N. in Wall Str. J. (2019)

[3] "spatio-temporal adversarial networks for abnormal event detection." by Lee, S., Kim, H. G., Ro, Y. M.: STAN in IEEE international conference on acoustics, speech and signal processing (ICASSP)(2018)

[4] "Observe locally, infer globally a space-time MRF for detecting abnormal activities with incremental updates." by

Kim, J., Grauman, K. in IEEE Conference on Computer Vision and Pattern Recognition (2009)

[5] “A review of video surveillance systems” by O. Elharrouss, N. Almaadeed, S. Al-Maadeed in J. Vis. Commun. Image Represent. (2021)

[6] “E2E-VSDL: end-to-end video surveillance-based deep learning model to detect and prevent criminal activities” by M.Q. Gandapur in Image Vis Comput. (2022)

[7] “Survey on contemporary remote surveillance systems for public safety” by T.D. Raty in IEEE Transactions on Systems, Man and Cybernetics. (2010)

[8] ”Abnormal event detection at 150 FPS in MATLAB.” by Lu C., Shi, J., Jia, J. in IEEE International Conference on Computer Vision (2013)

[9] “A unified approach to interpreting model predictions.”by Lundberg, S.M., Lee, S. In Proceedings of the 31st International Conference on Neural Information Processing Systems(2017)

[10] “Robust real-time unusual event detection using multiple fixed-location monitors.” by Adam, A., Rivlin, E., Shimshoni, I., Reinitz, D. in IEEE Trans. Pattern Anal. Mach. Intell. (2008)

[11] “Video anomaly detection with compact feature sets for online performance.” by Leyva, R., Sanchez, V., Li, C.-T. in IEEE Trans. Image Process.(2017)

[12] “Real-world anomaly detection in surveillance videos” by W. Sultani, C. Chen, M. Shah in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2018)

[13] Ravanbakhsh, M., Nabi, M., Sangineto, E., Marcenaro, L., Regazzoni, C., Sebe, N.: Abnormal event detection in videos using generative adversarial nets. In: 2017 IEEE international conference on image processing (ICIP), pp. 1577–1581 (2017)

[14] “Crime analysis through machine learning” by S. Kim, P. Joshi, P.S. Kalsi, P. Taheri in IEEE 9th Annual Informaation Technology, Electronics and Mobile Communication Conference(IEMCON),IEEE (2018)

[15] “Decision-theoretic saliency: computational principles, biological plausibility, and implications for neurophysiology and psychophysics.” by Gao, D., Vasconcelos, N. in Neural Comput.(2009)

[16] “A causal framework to determine the effectiveness of dynamic quarantine policy to mitigate COVID-19.” by Kristjanpoller, W., Michell, K., Minutolo, M.C. in Appl. Soft Comput. (2021)

[17] “Hierarchical context modeling for video event recognition.” by Wang, X., Ji, Q. in IEEE Trans. Pattern Anal. Mach. Intell.(2017)

[18] “Detecting anomalies in image classification by means of semantic relationships.” by Pasini, A., Baralis, E. in IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Sardinia, Italy.(2019)

[19] “The use of predictive analysis in spatiotemporal crime forecasting: building and testing a model in an urban context” by A. Rummens, W. Hardyns, L. Pauwels in Appl. Geogr. (2017)

[20] “Segmentation of COVID-19 pneumonia lesions: a deep learning approach.” by Ghomi, Z., Mirshahi, R., Bagheri, A.K., Fattahpour, A., Mohammadiun, S., Gharahbagh, A.A., Djavadifar, A., Arabalibeik, H., Sadiq, R., Hewage, K. in Med. J. Islam. Repub. Iran(2020)

[21] “SiTGRU: single-tunnelled gated recurrent unit for abnormality detection” by H. Fanta, Z. Shao, L. Ma. in Inf. Sci. (2020)

[22] Comparative Analysis on Sediment Classification using Convolutional Neural Network: 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS).

[23] Cooperative Sensing with Heterogeneous Spectrum Availability in Cognitive Radio: 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)

[24] “Detection of brain tumor from MRI images by using segmentation and SVM” by S.R.Telrandhe, A.Pimpalkar, A.Kendhe in IEEE WCTFTR World Conf. Futur. Trends. Res. Innov. Soc. Welf. (2016)

[25] Study and analysis of Intrusion Detection System using Random Forest and Linear Regression: Periodicals of Engineering and Natural Sciences ISSN 2303-4521 Vol.6 No.1, pp, 197-200 December 2018.

[26] Detection of cardiac amyloidosis on electrocardiogram images using machine learning and deep learning techniques: Computational.Intelligence.2022;1-23. Wileyonlinelibrary.com/journal/coin 2022 Wiley Periodicals LLC.