iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52283

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Predict, Identify and Alert on Suspicious Activity by Multiple Zone

Pratik Yadav[1], Mayuri Ghodke[2], Vivek Dhokane[3], Sanjana Chavan[4]
[1, 2, 3, 4] *Zeal College of Engineering and Research, Pune*

Abstract: ***An active topic of image processing and computer vision research is the detection of suspicious human behaviour in surveillance footage. In order to stop terrorism, theft, accidents, illegal parking, vandalism, fighting, chain snatching, crime, and other suspicious activities, human activity can be observed through visual surveillance in sensitive and public places like bus stations, railway stations, airports, banks, shopping malls, schools, and colleges. It is highly challenging to continually monitor public spaces, thus an intelligent video surveillance system that can track people's movements in real-time, classify them as routine or odd, and send out an alert is needed. The field of visual surveillance to identify aberrant actions has seen a significant amount of publications in the last ten years. Furthermore There are a few surveys that can be found in the literature for the recognition of various abnormal activities, but none of them have reviewed various abnormal activities. This study presents the state-of-the-art in the field of recognising suspicious behaviour from surveillance recordings during the past ten years. We give a quick overview of the concerns and difficulties associated in recognising suspicious human activity. This article examines six aberrant behaviours, including the identification of abandoned objects, theft, falls, traffic accidents, and unlawful parking, as well as the detection of violence and fire. Generally speaking, we have covered all the processes that have been [1]***
Foreground object extraction, object identification based on tracking or non-tracking approaches, feature extraction, classification, activity analysis, and recognition are some of the techniques that have been used to identify human activity from surveillance movies in the literature. This paper's goal is to give field researchers a literature assessment of six different suspicious activity identification systems together with its broad framework.[1]

I. INTRODUCTION

Unusual Human Activity An active field of image processing and computer vision research, identification from video surveillance includes identifying human behaviour and classifying it into normal and abnormal activities. The term "abnormal activity" refers to any odd or suspicious behaviour that is seldom displayed by a person in a public setting. Examples include leaving explosives in luggage, stealing, avoiding crowds, fighting and attacking people, vandalism, and border crossing. Normal activities are the typical human behaviours that take occur in public spaces, such as hand- waving, applauding, jogging, boxing, and walking. Video surveillance is used more and more often these days to keep an eye on people's movements and stop any suspicious behaviour.[9]

II. LITERATURE SURVEY

Members Ting Yao, Jiajun Deng, Yingwei Pan, Wengang Zhou, and [1] The authors are assessing Abstract— Compared to two-stage detectors, single shot detectors have the potential to be quicker and simpler, making them more useful for object recognition in movies. However, extending these object detectors from images to videos is not simple, especially when appearance degradation, such as motion blur or occlusion, occurs in movies. How to investigate temporal coherence across frames for improving detection is a legitimate topic. In this study, we suggest that the issue may be solved by aggregating surrounding frames to improve the per-frame characteristics. We provide Single Shot Video Object Detector (SSVD) in particular, a An innovative new architecture for object recognition in movies incorporates feature aggregation into a one-stage detector.[5] To create multi-scale features, SSVD technically uses the Feature Pyramid Network (FPN) as a backbone network. SSVD differs from previous feature aggregation techniques in that it directly samples data from consecutive frames in a two-stream structure while simultaneously estimating motion and aggregating neighbouring features along the motion route. On the ImageNet VID dataset, extensive tests are carried out, and competitive results are provided when compared to cutting-edge methodologies. More astonishingly, SSVD processes a frame on an Nvidia Titan X Pascal GPU in 85 milliseconds with 448 448 input, achieving 79.2mAP on Image Net VID. Object detection algorithms are used in a variety of industries, including defence, security, and healthcare, according to Apoorva Raghunandan, Mohana, Pakala Raghav, and H. V. Ravish Aradhya [2].

In order to more accurately identify different sorts of objects for video surveillance applications, a number of object recognition algorithms, including face detection, skin detection, colour detection, shape detection, and target detection, are simulated and implemented in this work. The difficulties and uses of different Object Detection methods are also discussed.[1]

Wenguan Wang, Member, and Jianbing Shen, Senior Member The purpose of visual attention in understanding object patterns in films is thoroughly investigated in this study. By thoroughly annotating three well-known video segmentation datasets (DAVIS16, Youtube- Objects, and SegTrackV2) with dynamic eye- tracking data in the unsupervised video object segmentation (UVOS) setting, we quantitatively verified for the first time the high consistency of visual attention behaviour among human observers. During dynamic, task-driven viewing, we also found a significant link between human attention and explicit primary object judgements. These ground-breaking discoveries provide a full grasp of the reasoning thatunderlies the patterns of video objects. The UVOS-driven Dynamic Visual Attention Prediction (DVAP) in the spatiotemporal domain and Attention-Guided Object Segmentation are the two sub-tasks that we separate from UVOS in light of these findings (AGOS) in the spatial domain. Our UVOS system has three key benefits: 1) modular training without the use of expensive video segmentation annotations; instead, the initial video attention module is trained using more affordable dynamic fixation data, and the subsequent segmentation module is trained using existing fixation-segmentation paired static/image data; 2) thorough foreground understanding through multi-source learning; and 3) additional interpretability from the biologically-inspired and assessable attention. Experiments on four well-known benchmarks demonstrate that, even without using pricey video object mask annotations, our model outperforms state-of-the-arts and processes data quickly.[5]

Object detection became one of the main fields in computer vision, according to Jung Uk Kim and Yong Man Ro [4]. The duties of object categorization and object localization arecarried out during object detection. With feature maps produced by entirely shared networks, previous deep learning-based object identification networks function well. Object classification, though, focuses on the feature map's most discriminative object region. Contrarily, object localisation necessitates afeature map that is concentrated on the full object's region. In this study, we analyse the distinction between the two objectives and present an unique object detection network. The two main components of the proposed deep learning-based network are 1) theattention network part, which generates task- specific attention maps, and 2) the layer separation part, which separates the layers for estimating two tasks. The proposed object detection network outperformed state-of-the- art techniques, according to extensive experimental results based on the PASCAL VOC dataset and MS COCO dataset.

Song Wang, Xingyuan Zhang, Qi Zou, Yanting Pei, Yaping Huang, and Image categorization has advanced significantly recently thanks to deep learning neural networks, particularly the Convolutional Neural Networks, just like many other computer vision-related fields (CNNs). As shown by thefrequently used image databases Caltech-256, PASCAL VOCs, and ImageNet, the majority ofthe existing works concentrated on classifying extremely clear natural images. However, theacquired images may have some degradations in many actual applications that result in different types of blurring, noise, and distortions. The impact of such degradations to the performance of CNN-based networks is one significant and intriguing issue. the ability of degradation removal to improve CNN-based picture categorization. More specifically, we question whether image classification performance decreases with each type of degradation, whether this decline can be prevented by training on degraded images, and whether the performance of existing computer vision algorithms that try to remove such degradations can be enhanced. In this research, we investigate such issues empirically for nine different types of damaged photos. Images that are hazy, motion-blurred, fish-eye, underwater, poor resolution, salt-and-peppered, have white Gaussian noise, are blurred, and are out of focus. We anticipate that more people will be interested in studying the categorization of damaged photos as a result of this work.[3]

III. METHODOLOGY

We used two label of first webpage.Label 1 ---- 1st zone Label2-----2nd zone

In each zone separately run YOLOV3 algorithm. 1st zone = SUSPICIOUS human activity 2nd Zone= = SUSPICIOUS vehicle activity They choose a zone and Run the software.They start working.

A. YOLOv3 Algorithm

You Only Look Once, Version 3 (YOLOv3), a real-time object recognition system, can identify specific objects in moving pictures, live feeds, or still images.[7] The YOLO machine learning system makes use of characteristics that a deep convolutional neural network has learnt tolocate an object. The third iteration of the YOLOmachine learning algorithm is a more accurate rendition of the first ML approach. Ali Farhadi and Joseph Redmond created YOLO versions 1- [2] 3.

YOLO's first version was released in 2016, while the most recent, version 3, which is the one this article focuses on extensively, was released in 2018. YOLOv3 is an improved variant of YOLO and YOLOv2. YOLO is implemented using the Keras or Open CV deep learning frameworks.

Installation of YOLOv3 is quite simple. After some dependencies and libraries have been installed, using it to train models is simple. YOLOv3 can be set up either through a notebook or directly on a computer (such asGoogle Collaborator or Jupyter). The commandsare the same for both implementations. The command to install YOLOv3 is pip install YOLOv3, assuming all libraries have been installed. I'll give you a quick tutorial on installing YOLOv3 along with the necessary libraries. Visit Viso.AI/DeepLearning/Yolov3- Overview to learn more.[7]

Choosing a specific object detection project would be the first step in using YOLOv3. Forbeginners to get started with YOLOv3, it is best to choose a simple project with an easy premise, such as detecting a specific sort of animal or car in a video. YOLOv3 does real-time detections. We will go over the necessaryprocedures and information in this section so that you may successfully use the YOLO machine learning method.[2]

B. Model Weights

Weights and cfg (or configuration) files are available for download on the website of YOLOv3's original developer, https://pjreddie.com/darknet/yolo. After downloading the model weights, save them as "yolov3.weights" in your current directory. You may alternatively (easier) utilise YOLO's COCO pertained weights by initialising the model with model = YOLOv3 (). Only if you use the pre-trained weights from COCO can you use YOLO for object detection with any of the 80 pertained classes that are included with the COCO dataset. This is a good option for beginners because it requires the least amount of new code and customization.[2]

IV. SYSTEM ARCHITECTURE

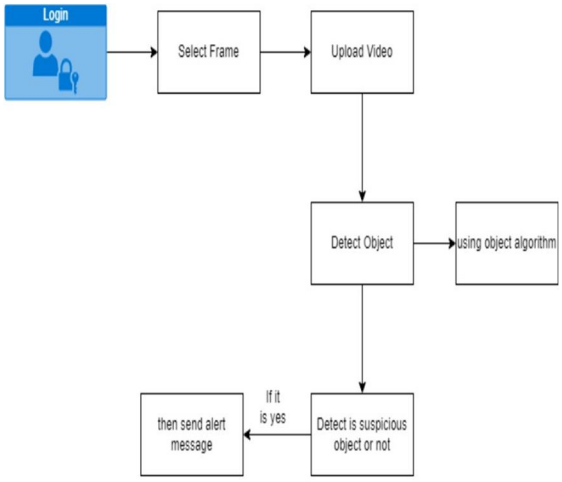


Figure 4.1: system Architecture

V. DISADVANTAGES

One disadvantage crash or damage camera then system is not working because system is also depend on camera. APPLICATION

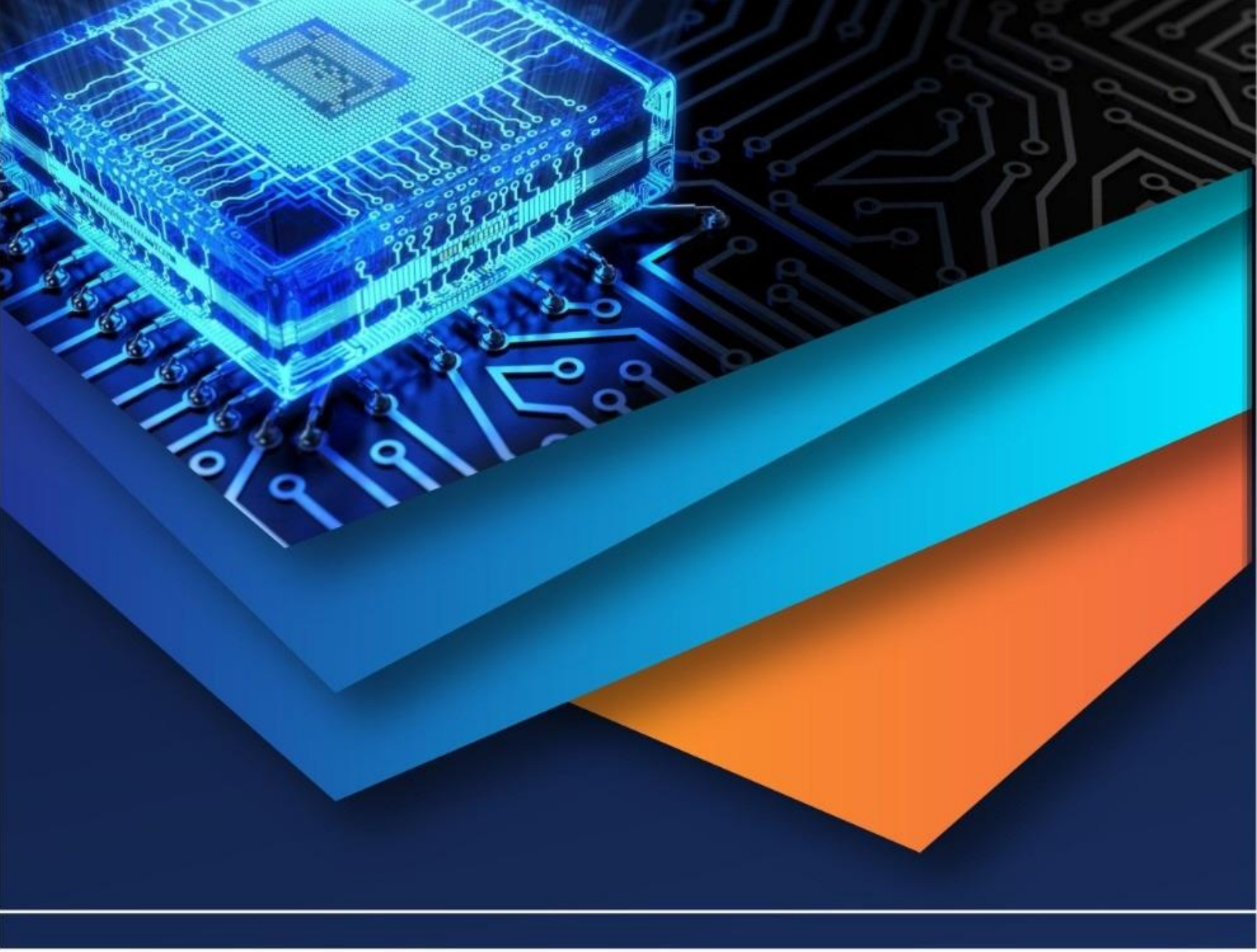
1) Forest
2) Banking
3) Society

VI. CONCLUSION

In the modern world, practically everyone is aware of the value of CCTV video, yet in the majority of situations, these footages are only used for investigation after a crime or event has occurred. The benefit of the suggested paradigm is that it prevents crime before it occurs. Real-time CCTV footage is being monitored and examined. If the analysis's conclusion predicts an unfortunate occurrence will occur, the appropriate authority is instructed to take action. So, it is possible to stop this. work with various datasets. And establish the zone For the real-time detection of deformed objects after motion in the videos, our method can offer useful suggestions. Using a live system for human interaction in the forest

REFERENCES

[1] MUHAMMAD SUALEH AND GON-WOO KIM, "Visual-LiDAR Based 3D Object Detection andTracking for Embedded Systems", IEEE Access (August 24,2020,)10.1109/ACCESS.2020.3019187
[2] Jiajun Deng, Yingwei Pan, Ting Yao, Member,IEEE, Wengang Zhou, Member, IEEE, Houqiang Li, Senior Member, IEEE, and Tao Mei, Fellow, IEEE , "Single Shot Video Object Detector", information: DOI 10.1109/TMM.2020.2990070, IEEE Transactions on Multimedia.
[3] Deng-Ping Fan, Ge-Peng Ji, Ming-Ming Cheng, and Ling Shao , "Concealed ObjectDetection", information: DOI 10.1109/TPAMI.2021.3085766, IEEE, Transactions on Pattern Analysis and Machine Intelligence
[4] Wenguan Wang, Member, IEEE, Jianbing Shen, Senior Member, IEEE, Xiankai Lu, Member, IEEE, Steven C. H. Hoi, Fellow IEEE,Haibin Ling , "Paying Attention to Video Object Pattern Understanding", information: DOI 10.1109/TPAMI.2020.2966453, IEEETransactions on Pattern Analysis and Machine Intelligence
[5] Runsheng Zhang, Yaping Huang, Mengyang Pu, Jian Zhang, Qingji Guan, Qi Zou, Haibin Ling Member, IEEE, "Object Discovery From a Single Unlabeled Image by Mining Frequent ItemsetWith Multi-scale Features", DOI10.1109/TIP.2020.3015543, IEEE Transactionson Image Processing.
[6] Yanting Pei, Yaping Huang, Qi Zou, Xingyuan Zhang, Song Wang , "Effects of Image Degradation and Degradation Removal to CNN- based Image Classification", information: DOI 10.1109/TPAMI.2019.2950923, IEEE Transactions on Pattern Analysis and Machine Intelligence.
[7] Jung Uk Kim and Yong Man Ro, "ATTENTIVE LAYER SEPARATION FOR OBJECT CLASSIFICATION AND OBJECT LOCALIZATION IN OBJECT DETECTION", 978-1-5386-6249- 6/19/$31.00 ©2019 IEEE
[8] Apoorva Raghunandan, Mohana, Pakala Raghav and H. V. Ravish Aradhya , "Object Detection Algorithms for Video Surveillance Applications", 978-1-5386-3521-6/18/$31.00 ©2018 IEEE
[9] H.-K. Chu, W.-H. Hsu, N. J. Mitra, D. Cohen- Or, T.-T. Wong, and T.-Y. Lee, "Camouflage images." ACM Trans. Graph., vol. 29, no. 4, pp.51–1, 2010.
[10] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review," IEEE T. Neural Netw. Learn. Syst., vol. 30, no. 11,pp. 3212–3232, 2019.
[11] D.-P. Fan, J. Zhang, G. Xu, M.-M. Cheng, and L. Shao, "Salient objects in clutter," arXiv preprint arXiv:, 2021.
[12] J.-X. Zhao, J.-J. Liu, D.-P. Fan, Y. Cao, J. Yang, and M.-M. Cheng, "Egnet:edge guidance network for salient object detection," in Int. Conf.Comput. Vis., 2019.
[13] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisser-man, "The pascal visual object classes (voc) challenge," Int. J. Comput. Vis., vol. 88, no. 2, pp. 303–338, 2010.
[14] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet:A large-scale hierarchical image database," in IEEE Conf. Comput. Vis.Pattern Recog., 2009, pp. 248–255.
[15] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan,P. Dollar, and C. L. Zitnick, "Microsoft coco: Common objects in´context," in Eur. Conf. Comput. Vis., 2014, pp. 740–755.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY