IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

SUSPICIOUS ACTIVITY DETECTION AND ALERT SYSTEM

[1]K DINESH BABU, [2]R NATHASHA

[1]Associate Professor, [2]UG Student
Department of Electronics and Communication Engineering,
Adhi College of Engineering and Technology, Kancheepuram-631 605

Abstract: Video Surveillance plays a pivotal role in today's world. The technologies have been advanced too much when artificial intelligence, machine learning and deep learning pitched into the system. Using above combinations, different systems are in place which helps to differentiate various suspicious behaviours from the live tracking of footages. The most unpredictable one is human behaviour and it is very difficult to find whether it is suspicious or normal. Deep learning approach is used to detect suspicious or normal activity in an academic environment, and which sends an alert message to the corresponding authority, in case of predicting a suspicious activity. Monitoring is often performed through consecutive frames which are extracted from the video. The entire framework is divided into two parts. In the first part, the features are computed from video frames and in second part, based on the obtained features classifier predict the class as suspicious or normal.

IndexTerms **- Activity detection, Segmentation, Alert system.**

I. INTRODUCTION

In recent years, ever-increasing technological advances have made automated human activity recognition a common research subject. Video surveillance has a wide range of applications. These applications include normal and suspicious activities such as gaming, human-computer interaction, exam invigilation, detecting chaos, analysing sports, predicting crowd behaviour, etc. It is an important safety aspect for indoor and outdoor environments.

Innovations are occurring rapidly, and since there is a large amount of video data to process, manual intervention is not feasible and is error-prone. Additionally, it is exceedingly challenging to monitor public spaces constantly. Hence, it is necessary to install intelligent video surveillance that can track people's movements in real time, classify them as routine or exceptional, and provide alerts.

Human activity detection relies on sensors like radar, cameras, and cell phones to identify abnormalities in human behaviour. They are being used for human-computer interaction, surveillance, monitoring suspicious activities, and other security purposes. The majority of today's systems rely on video gathered from CCTV cameras. If a crime or act of violence occurs, this footage will be utilized in the investigation. It would be preferable, however, to build a system that might identify an anomalous or unexpected circumstance beforehand and notify the authorities.

II. EXISTING SYSTEM

Object Detection is the process of finding and recognizing real-world object instances such as car, bike, TV, flowers, and humans out of an images or videos. An object detection technique lets you understand the details of an image or a video as it allows for the recognition, localization, and detection of multiple objects within an image.

It is usually utilized in applications like image retrieval, security, surveillance, and advanced driver assistance systems (ADAS). Object Detection is done through many ways

III. PROPOSED SYSTEM

Activity Recognition is divided into two categories: sensor-based activity recognition and vision-based activity recognition, depending on the system's components.

Using a camera-based system, vision-based activity recognition may identify the activities present in an environment by processing and analyzing video. These systems often employ digital image processing to draw out relevant data from video, which is regarded as a series of images. In this project, we used a vision-based system.

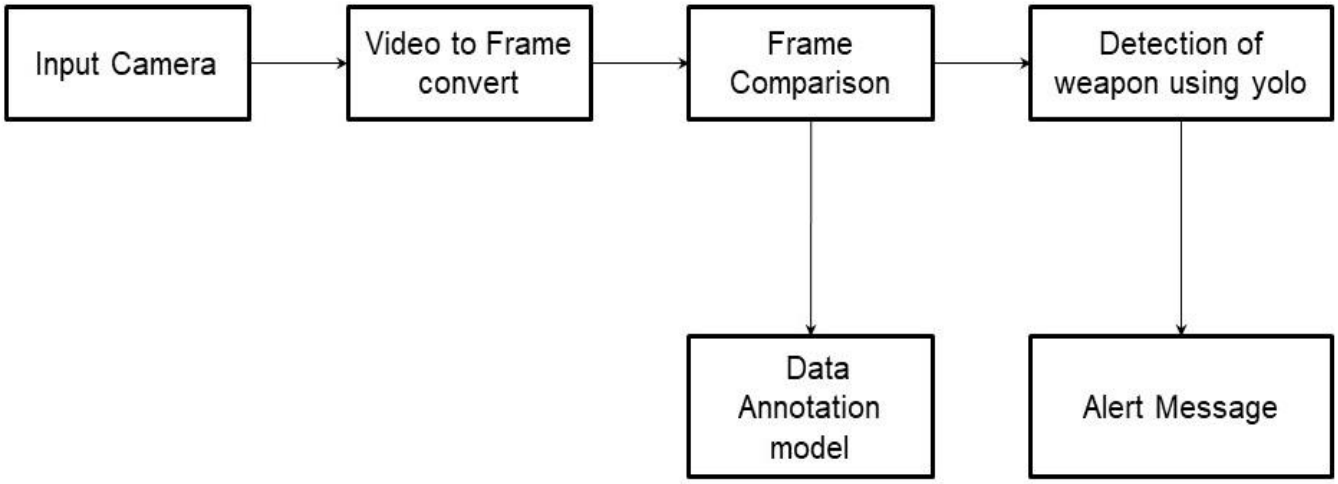

Figure :1 Block diagram

A. Input Camera

A camera is an optical instrument that captures images. Most cameras can capture 2D images, while some more advanced models can capture 3D images. At a basic level, most cameras consist of a sealed box (the camera body), with a small hole (the aperture) that allows light to pass through and capture an image on a light-sensitive surface (usually a digital sensor or photographic film). Cameras have various mechanisms to control how light falls onto the light-sensitive surface, including lenses that focus the light and a shutter that determines the amount of time the photosensitive surface is exposed to the light.

B. Video to Frame Convert

The concept of converting a video into frames involves extracting individual frames from the video file and saving them as image files. Here's an overview of the steps involved in a video to frame conversion process:

- Open the video file:
  You need to open the video file using a video capturing library, such as OpenCV in Python. This allows you to read the frames from the video.
- Read frames from the video:
  Use the video capturing library to read frames from the video file one by one. Each frame is a single image that represents a snapshot of the video at a particular time.

C. Frame Comparison

Capturing a video on camera then it was converting to frames involves extracting individual frames from the video file and saving them as image files. Then it was comparing an image files with reference image.

D. Detection of weapon using yolo

YOLOv5 is a state-of-the-art real-time object detection system, developed by Ultralytics LLC. It is the latest version of the YOLO (You Only Look Once) series, which uses a single neural network to detect and classify objects in an image. Compared to previous versions of YOLO, YOLOv5 has several improvements such as a smaller model size, faster inference time, and higher accuracy. It is based on a "Scaled-YOLOv4" architecture, which uses a combination of CSP (cross-stage partial) and PAN (path aggregation network) modules to improve the accuracy of object detection. YOLOv5 has been trained on several large-scale datasets such as COCO, VOC, and Open Images, and it can detect more than 80 different object categories. It has become increasingly popular in the computer vision community due to its excellent performance on both desktop and mobile platforms.

E. Data Annotation model

Data annotation is the categorization and labelling of data for AI applications. Training data must be properly categorized and annotated for a specific use case. With high-quality, human-powered data annotation, companies can build and improve AI implementations.

Image Annotation

Image annotation is vital for a wide range of applications, including computer vision, robotic vision, facial recognition, and solutions that rely on machine learning to interpret images. To train these solutions, metadata must be assigned to the images in the form of identifiers, captions, or keywords.

From computer vision systems used by self-driving vehicles and machines that pick and sort produce, to healthcare applications that auto-identify medical conditions, there are many use cases that require high volumes of annotated images. Image annotation increases precision and accuracy by effectively training these systems.

IV. RESULTS AND DISCUSSION

The outcomes of this model, if motion is detected on the camera, a green frame covering the detected motion is created, as shown in the figure below. making it easy for the user to identify suspicious activity.

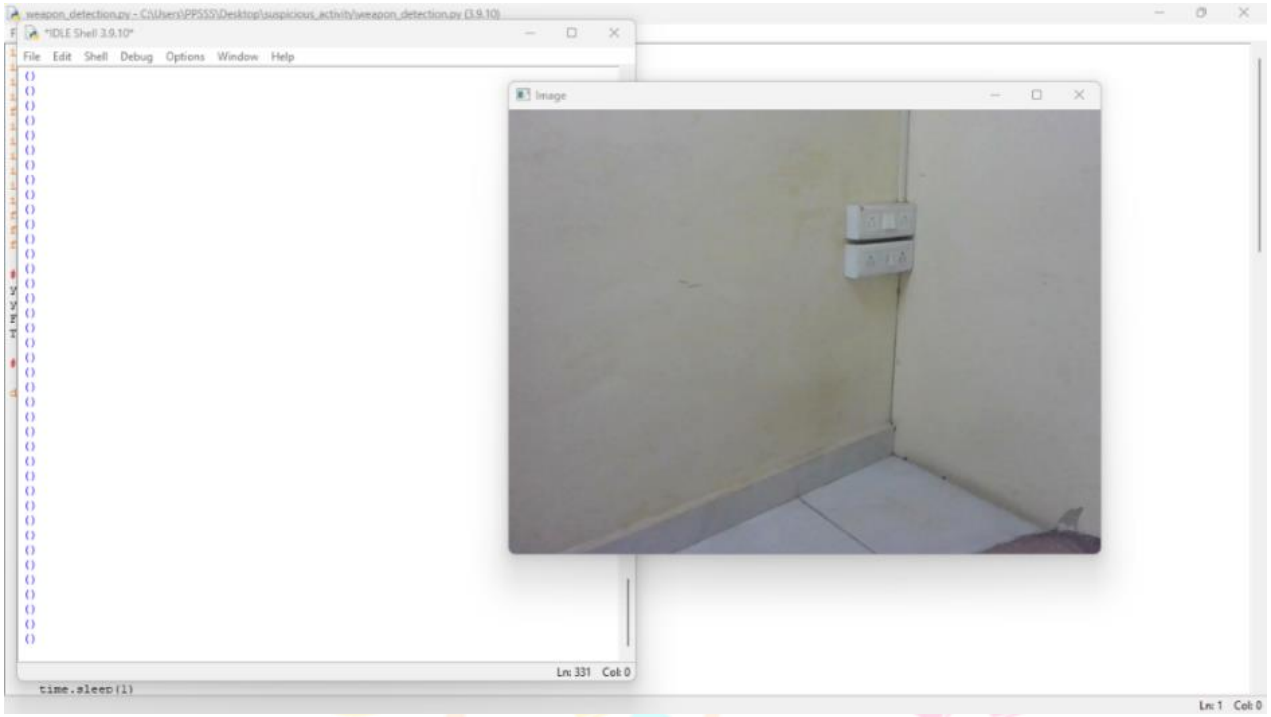

Figure 2. Software Running on Specific place

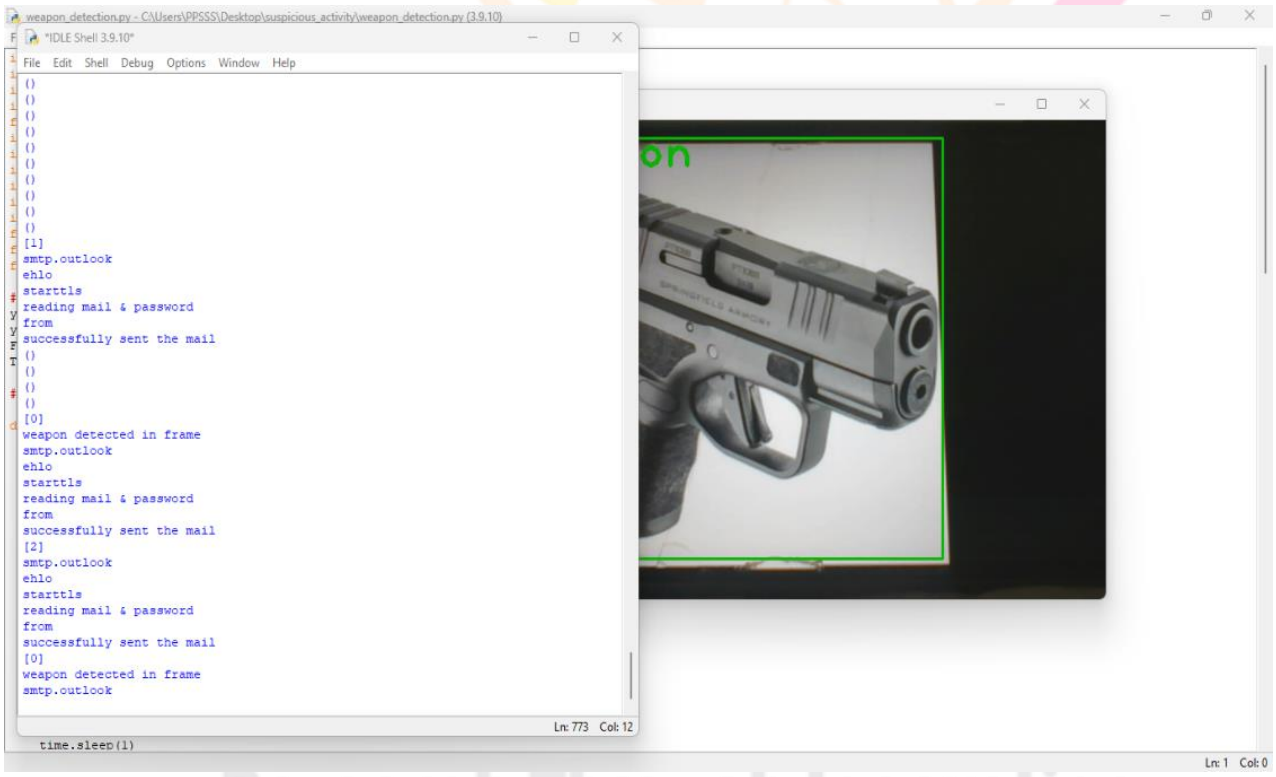

Figure 3. Weapon Detection

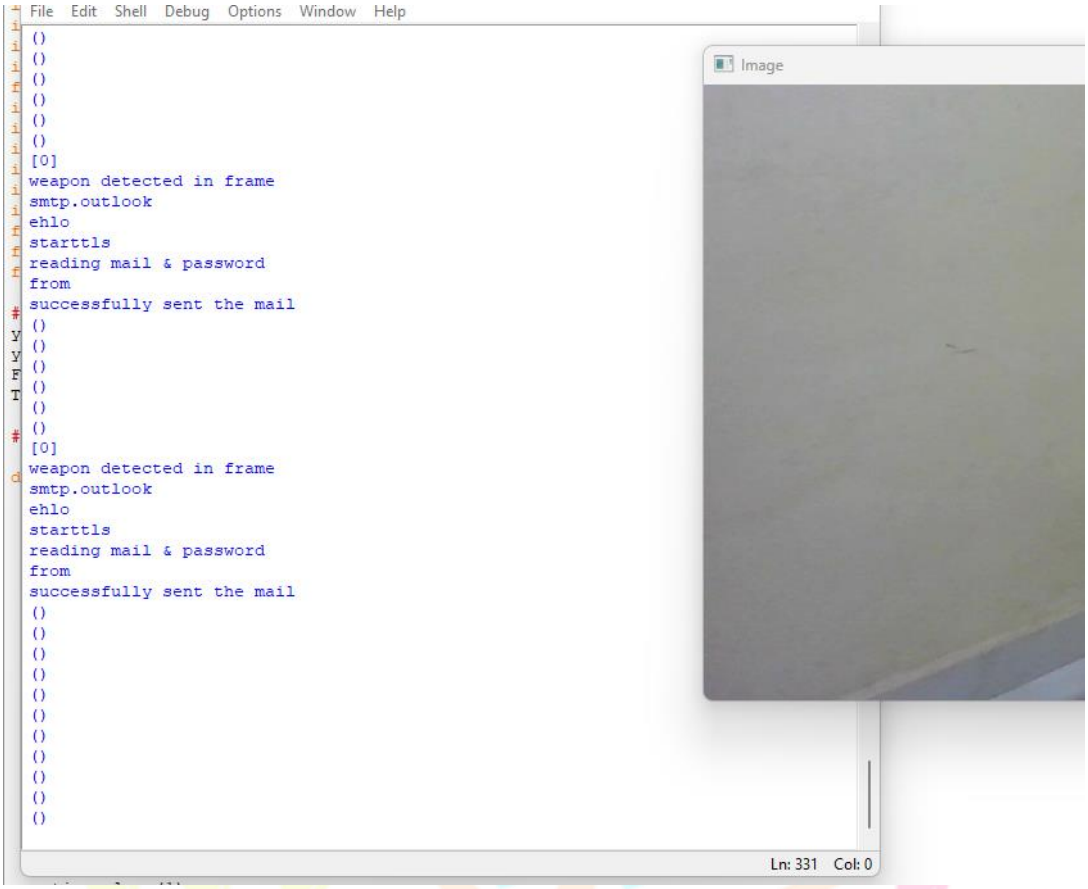

Figure 4. Detected Message to send Mail

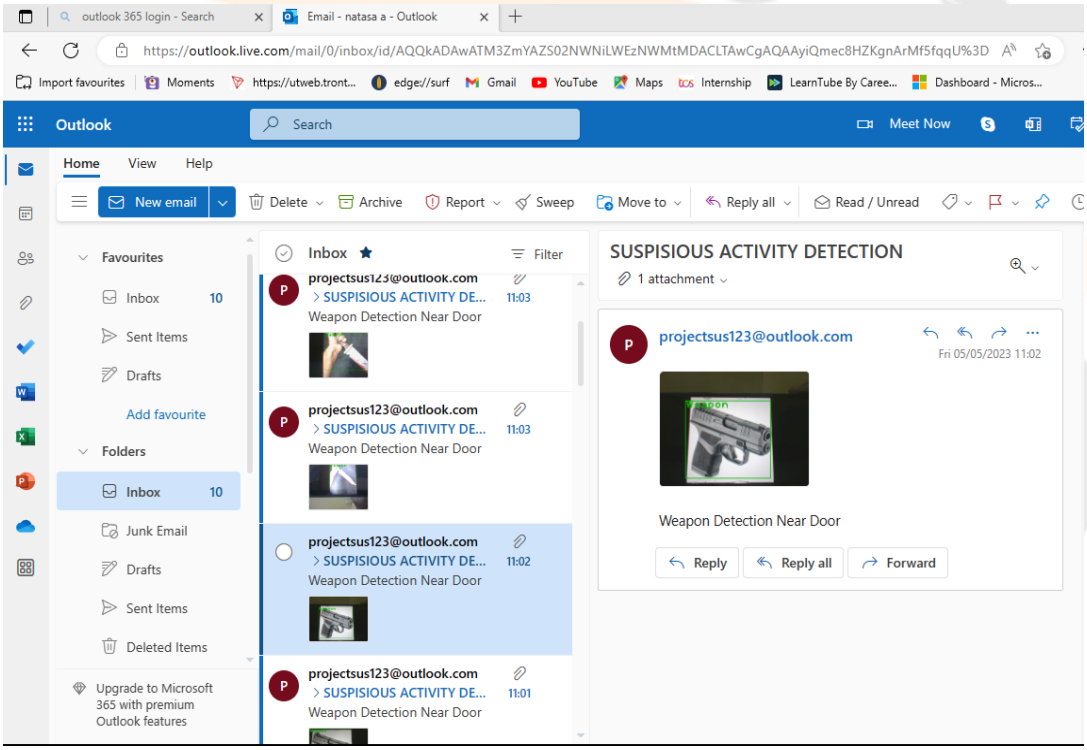

Figure 5. Alert Message Send To Mail

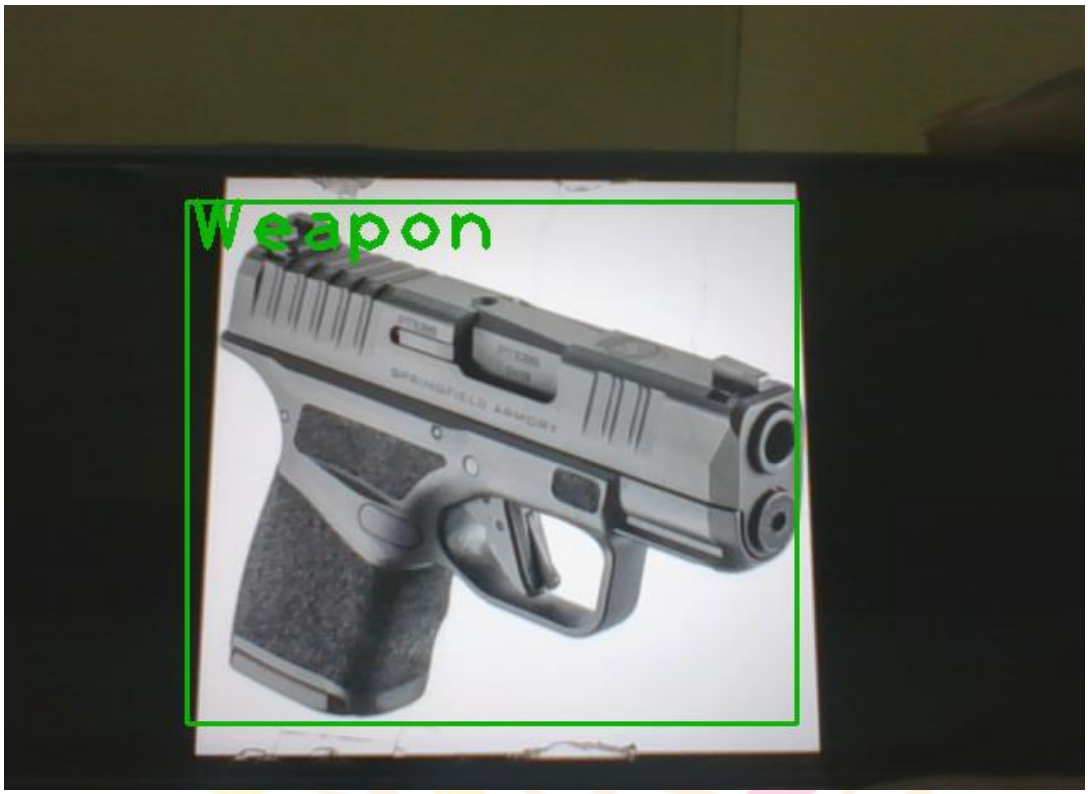

Figure 6. Received Suspicious activity in mail

In Shown figure 2, Detecting a specific place where we need to found a suspicious activity on camera. In Shown figure 3, Detecting a suspicious things like guns, knife, etc. then captured by camera. In Shown figure 4 and 5, Then captured image send to the mail and received on the alert message to mail In Shown figure 6, Then found a suspicious activity via mail to the owner.

V. FUTURE SCOPE

A major challenge for automatic image analysis is that the sheer complexity of the visual task which has been mostly ignored by the current approaches. New technological breakthrough in the areas of digital computation and telecommunication has relevance for future applications of image processing1. The satellite imaging and remote sensing applications programs of the future will feature a variety of sensors orbiting the earth. This technology is required for military and other types of surveillance, statistical data collection in the fields of forestry, agriculture, disaster prediction, weather prediction. In order to extract scientifically useful information, it will be necessary to develop techniques to register real-time data recorded by a variety of sensors for various applications.

VI. CONCLUSION

The concept of Suspicious Activity Detection and Alert involves leveraging computer vision algorithms to analyze video feeds from surveillance cameras in real-time. YOLO, as a popular object detection algorithm, can accurately detect objects of interest, including people and other potential threats, in video frames with high speed and accuracy. By training YOLO on labeled video frames, the system can learn to recognize various types of suspicious activities, enabling it to detect potential security threats, unauthorized access, or abnormal movements.

REFERENCES

[1] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, ``An image is worth 16x16 words: Transformers for image recognition at scale," Jun. 2021, arXiv:2010.11929. Accessed: Apr. 28, 2022.

[2] A. Vora and V. Chilaka, ``FCHD: Fast and accurate head detection in crowded scenes," May 2019, arXiv:1809.08766. Accessed: Apr. 29, 2022.

[3] C. Patel, D. Bhatt, U. Sharma, R. Patel, S. Pandya, K. Modi, N. Cholli, A. Patel, U. Bhatt, M. A. Khan, S. Majumdar, M. Zuhair, K. Patel, S. A. Shah, and H. Ghayvat, ``DBGC: Dimension-based generic convolution block for object recognition," Sensors, vol. 22, no. 5, p. 1780, Feb. 2022.

[4] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in arXiv, vol. abs/1512.00567, 2015

[5] C. Wang, S. Zhao, and R. Zhang, ``Occlusion-aware discriminative networks for visual object tracking," in Proc. 2nd Int. Conf. Robot., Intell. Control Artif. Intell., New York, NY, USA, Oct. 2020

[6] CRCV | Center for Research in Computer Vision at the University of Central Florida. Accessed: Apr. 28, 2022. [Online]. Available: https://www.crcv.ucf.edu/data/UCF-ARG.php

[7] Claire-Hlne Demarty, et. al The MediaEval 2012 affect task: violent scenes detection Working Notes Proceedings of the MediaEval 2012 Workshop. 2012.
[8] G. Chevalier, "LARNN: Linear attention recurrent neural network,"Aug. 2018, arXiv:1808.05578. Accessed: Apr. 28, 2022.
[9] H. Wang and L. Wang, ``Modeling temporal dynamics and spatial configurations of actions using two-stream recurrent neural networks," Apr. 2017, arXiv:1704.02581. Accessed: Apr. 29, 2022.
[10] J. Chorowski, D. Bahdanau, D. Serdyuk, K. Cho, and Y. Bengio,"Attention-based models for speech recognition," Jun. 2015, arXiv:1506.07503. Accessed: Apr. 28, 2022.
[11] J. Liu, A. Shahroudy, D. Xu, and G. Wang, ``Spatio-temporal LSTM with trust gates for 3D human action recognition," Jul. 2016, arXiv:1607.07043. Accessed: Apr. 29, 2022.
[12] J. Liu, N. Akhtar, and A. Mian, ``Skepxels: Spatio-temporal image representation of human skeleton joints for action recognition," Aug. 2018, arXiv:1711.05941. Accessed: Apr. 28, 2022.
[13] K. He, X. Zhang, S. Ren, and J. Sun, ``Deep residual learning for image recognition," Dec. 2015, arXiv:1512.03385. Accessed: Apr. 28, 2022.
[14] Lisa Torrey and Jude Shavlik, "Transfer Learning", University of Wisconsin, Madison WI, USA, Available: http://ftp.cs.wisc.edu/machine-learning/shavlikgroup/torrey.handbook09.pdf [Accessed Mar 15, 2019]
[15] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," Sep. 2020, arXiv:1905.11946. Accessed: Apr. 28, 2022.
[16] N. AlDahoul, H. A. Karim, M. H. L. Abdullah, M. F. A. Fauzi, A. S. B. Wazir, S. Mansor, and J. See, ``Transfer detection of YOLO to focus CNN's attention on nude regions for adult content detection," Symmetry, vol. 13, no. 1, p. 26, Dec. 2020.
[17] New System Performs Persistent Wide-Area Aerial Surveillance. Accessed: Apr. 28, 2022.
[18] Pierre Sermanet David Eigen Xiang Zhang Michael Mathieu Rob Fergus Yann LeCun "OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks" Courant Institute of Mathematical Sciences, New York University.
[19] S. Ren, K. He, R. Girshick, and J. Sun, ``Faster R-CNN: Towards real-time object detection with region proposal networks," Jan. 2016, arXiv:1506.01497. Accessed: Apr. 28, 2022.
[20] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards RealTime Object Detection with Region Proposal Networks," in NIPS 2015 - Advances in Neural Information Processing Systems 28. Neural Information Processing Systems Foundation, 2015
[21] S. Song, C. Lan, J. Xing, W. Zeng, and J. Liu, ``An end-to-end spatiotemporal attention model for human action recognition from skeleton data," in Proc. AAAI Conf. Artif. Intell., Feb. 2017, vol. 31, no. 1, Art. no. 1. Accessed: Apr. 28, 2022.
[22] S. Y. Nikouei, Y. Chen, S. Song, R. Xu, B.-Y. Choi, and T. R. Faughnan, "Real-time human detection as an edge service enabled by a lightweight CNN," Apr. 2018, arXiv:1805.00330. Accessed: Apr. 29, 2022.
[23] Samir K. Bandyopadhyay, Biswajita Datta, and Sudipta Roy Identifications of concealed weapon in a Human Body Department of Computer Science and Engineer, University of Calcutta, 2012
[24] Soft Computing and Intelligence Information Systems, [Online] Available: https://sci2s.ugr.es/weapons-detection [Accessed Mar 27, 2019]
[25] Roberto Olmos, Siham Tabik, and Francisco Herrera Automatic Handgun Detection Alarm in Videos Using Deep Learning Soft Computing and Intelligent [26] Information Systems research group, Department of Computer Science and Artificial Intelligence, University of Granada, 2017
[27] Rohit Kumar Tiwari and Gyanendra K. Verma, "A Computer Vision based Framework for Visual Gun Detection using Harris Interest Point Detector", Procedia Computer Science, vol 54, p. 703 - 712, 2015.
[28] T. Li, J. Liu, W. Zhang, Y. Ni, W. Wang, and Z. Li, ``UAV-human: A large benchmark for human behavior understanding with unmanned aerial vehicles," Aug. 2021, arXiv:2104.00946. Accessed: Apr. 28, 2022.
[29] The IMFDB Internet Movie Firearms Database, [Online] Available: http://www.imfdb.org/wiki/Main_Page [Accessed Mar 20, 2019]
[30] W. Li, H. Li, Q. Wu, F. Meng, L. Xu, and K. N. Ngan, ``HeadNet: An end-to-end adaptive relational network for head detection," IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 2, pp. 482-494, Feb. 2020.
[31] W. Zhu, C. Lan, J. Xing, W. Zeng, Y. Li, L. Shen, and X. Xie, "Co-occurrence feature learning for skeleton based action recognition using regularized deep LSTM networks," Mar. 2016, arXiv:1603.07772. Accessed: Apr. 29, 2022.
[32] Z. Shao, G. Cheng, J. Ma, Z. Wang, J. Wang, and D. Li, ``Real-time and accurate UAV pedestrian detection for social distancing monitoring in COVID-19 pandemic," IEEE Trans. Multimedia, vol. 24, pp. 20692083,2022.