

# A Secured Surveillance System Employing Dense Optical Flow Estimation

Dheeraj D<sup>1\*</sup>, Abhishek V Tatachar<sup>2</sup>, Vishwas K V<sup>2</sup>, Shivaneer Kondur<sup>2</sup>, Sumukh B K<sup>2</sup>

<sup>1</sup>Assistant Professor, Department of ISE, Global Academy of Technology, Bangalore.

<sup>2</sup>Undergraduate Student, Department of ISE, Global Academy of Technology, Bangalore.

**Abstract** – With the growing amount of threat to humankind, security concerns arise. Security of oneself as well as tangible assets of a person has become an overly critical and important issue. It becomes important to include some measures that keeps the surrounding environment safe and secure. One simple and most common approach to security is surveillance. Today, we can see surveillance cameras being deployed in many places, including hospitals, banks, apartments and so on. Majority of such deployments are manual, that is they require human interference for alerting the concerned personal in case of any anomaly event. With the amount of technology at one's disposal today, it is possible to automate the process to an extent. Object detection, object tracking, motion estimation and alerting can be combined to produce a system that automates the process of surveillance. In this literature, we propose a surveillance system that makes use of object detection and optical flow estimates to identify potential criminal components and determine whether a given scene or event qualifies as an anomaly or not.

**Keywords** – Object Detection, YOLOv3, Optical Flow Estimation, Dense optical flow, Gunnar Farneback method.

## 1. INTRODUCTION

Security is a pressing concern in the modern world. Threat not only applies to people but also to the property and assets owned by people. Today, people look for security in every scenario. Banks are installed with security cameras as well as security personal to ensure safety. Similarly, security installations are also found in hospitals, housing societies and apartments, and other areas where people tend to congregate.

There are several approaches or measures to ensure security. One such measure is surveillance. Surveillance refers to the process of constant monitoring of activities in an environment for the purpose of information aggregation or supervision. Historically surveillance has been used for the purpose of spying or gaining information on the enemy. As technology advanced, surveillance methods also changed. Modern day surveillance technologies include CCTV and GPS.

Recently, Computer Vision and Image Processing have become extremely popular and widely used. Several research and optimization efforts has been carried out in this field over the last decade. Computer vision is a part of a larger field of study called Artificial Intelligence (AI), that takes a visual input such as an image or a video and derives useful information from them. Computer Vision finds its applications in multiple areas such as Optical Character Recognition (OCR) [1], Surgical Operating Rooms [2],

metrology [3], military applications [4] and, surveillance and security [5].

Computer vision is becoming increasingly popular, due to its potential in applications that today need human intervention. Computer vision is increasing efficiencies of several processes and reducing the burden on humans. Surveillance is one such field where computer vision is being increasingly used. If we have a closer look, it makes sense that the ideas and techniques from the field of computer vision benefits the surveillance process by being able to detect elements in a scene (object detection), track them (object tracking) and gather useful information from what the machine saw. To an extent the process of surveillance can be automated to ensure security even in the absence of a security personal.

This is not the first time that a proposal is being made for the use of computer vision and its associated technologies for surveillance. Multiple approaches have been proposed and devised earlier to enable intelligent surveillance solutions [6] [7] [8] [9] [10]. The main difficulties in designing such systems are occlusions, noisy video streams, and variations in lighting. In this literature, we put forward a system that makes use of optical flow estimation and object identification to assess the safety of the surroundings.

## 2. LITERATURE REVIEW

Over the last several decades technology has advanced and so has the methods used for surveillance. People have advanced from physical surveillance to use of Artificial Intelligence (AI) to increase efficiencies. In [11] we have briefly studied the background required for developing a surveillance system that makes use of optical flow estimation. But before we look into such systems, we need to understand the various other implementations of surveillance systems used for securing the environment.

With the advancements of IoT-based technologies, traditional security installations are improving. Surveillance systems are utilised in addition to existing technologies to improve security. To identify security holes and mitigate safety issues, [12] proposes a Smart Surveillance System with machine learning capabilities. Intruders and questionable activities are recognised using machine learning techniques. Energy efficiency is the main issue with continuous surveillance systems. In order to function on the absolute least amount of power while still producing the appropriate output, the suggested system places a strong emphasis on power consumption. A fire sensor has even been added for safety concerns. Upcoming smart surveillance systems for a safe future can be established by

expanding the infrastructure. The built system includes the tools required to spot intruders using facial recognition. Using ambient sensors, a safe atmosphere is also provided throughout working and non-working hours. The technology detects humans and flames with excellent accuracy. With the advancement of this study, a more dependable security system may be established.

Many sites are equipped with CCTV cameras, but manual security is still used to monitor them, leaving plenty of room for human mistake, whether via neglect or a dangerous circumstance. Numerous bizarre events, like hostage situations, hostility, or flames, might result from this. An automation system that makes an effort to avoid this using deep learning methods was created. Three datasets, including UCF crime, Real life violent events, and UBI-fights, are used to train the DL models based on CNNs, particularly the compact MobileNetV2 and ResNet50V2 models [13]. The video's effectiveness in detecting violence is assessed using frames. The technology recognises a threat in the frame and generates an alert for the situation. In severe circumstances, notifications are transmitted immediately to neighbouring police stations and emergency services, including exact location information and a description of the suspicious activity at the moment.

Closed-circuit television (CCTV) systems are essential for avoiding issues with public safety. The usage of weapons in public places nowadays creates a serious security concern and adds to the high level of violence. In order to reduce or restrict risks, it is essential to rapidly identify guns in public places. Even in places where guns are forbidden, pistols are used in numerous crimes all throughout the world. It has long been common practise to watch these situations using closed-circuit video (CCTV); now it's time to monitor the visuals. Humans frequently execute this job, making them more prone to forgetting important details out of fatigue or being preoccupied [14]. They used deep learning methods in their research to identify pistols in an unsupervised mode using surveillance footage data.

Motion estimation is a crucial component for processing of surveillance video, including video filtering and compression of video frame. [15] proposes a straightforward and effective surveillance system that uses surveillance video frames for motion detection and motion vector estimation. Motion detection is accelerated by the application of a revolutionary method called edge region determination. The Horn-Schunck method is then used to analyse the surveillance film for motion estimates due to its trustworthy performance and simplicity. A detection accuracy of 98.6 percent is what they see. The Horn-Schunck optical flow method performs the motion estimation approach. High motion vector density and ease of implementation are provided by this. The computing cost is therefore lower. Additionally, their technique is quick since no morphological procedures are carried out. It is not affected by shifting camera conditions or variations in lighting such as indoor, outdoor, sunny, foggy, night, or day. There are a myriad number of algorithms available for detecting object motion. One of the well-known techniques the optical flow that involves determining the apparent filed velocity between two consecutive images in the same scene.

[16]. In this literature the authors have developed a system for accurate estimation of motion based on the calculation of optical flow using the Lucas Kanade algorithm. They observe that, their proposed method has a lower EPE and the angular errors attain results with an average which is as low as 0.2.

In [17] the authors have proposed a way to detect some kind of violent behaviour in videos based on optical flow estimation. The method proposed by them involves extracting frames from the video, convert the frames to HSV format, process the frame through an optical flow estimation and finally use behaviour analysis to classify the event as an anomaly event or a normal event. They have made use of the SDHA 2000 dataset to evaluate their work. They concluded that their work is suitable for detecting anomalous activities in unguarded environments or some kind of public places.

In [18] the authors propose an effective surveillance system that works on motion estimation, by calculating the motion vectors from the frames of the surveillance videos. In their proposal, the surveillance video is passed through the Horn-Schunck algorithm for the determination of the motion vectors. They observe that their proposed method is computationally fast without the need for a special hardware for the purpose of image processing.

### 3. OBJECT DETECTION WITH YOLOV3

One of the most difficult issues in the field of computer vision is thought to be object detection. Object detection is a subfield of computer vision that aims at detecting instances of particular objects in images or videos. Object detection has become very relevant to today's world, and it is being used in multiple areas. We can find tonnes of examples in real-life where object detection is being used. One simple example that employs object detection is the system that counts the number of vehicles that enter a parking lot, captures the car model, colour and the vehicle number.

While defining object detection, we can also define two more important terminologies that is very important for the task at hand – object localization and image classification. In very simple words object localization is the process of defining or identifying where the object is located in the image and drawing a bounding box (anchor box) around the object. Object localization is mainly considered as a regression problem, and sometimes it is also referred to as image localization. Image classification is the process of identifying to which class the image belongs. To better understand these terms let us consider an example of a system that is responsible for classifying a given animal's image as a cat or dog. Here, the process of identifying where the animal is in the image and drawing a bounding box around the image is object localization while classifying the identified animal as cat or dog is called image classification. These two terminologies are extremely important in the further course of this paper.

There are a number of algorithms that are being used for object detection. Some common algorithms include Fast R-CNNs [19], Histogram of Oriented Gradients [20], Single Shot Detector [21] and You Only Look Once (YOLO) [22]. YOLO stands for You Only Look once, it is a real-time, state of the art system for object detection. YOLO has become

popular for its accuracy and speed. YOLO deals with object detection as a regression problem and it makes use of CNNs (Convolution Neural Networks) for performing object detection. As the name “You Only Look Once” suggests, this algorithm requires only one forward propagation through the network for detecting objects. Primarily YOLO stands apart from other object detection for three reasons – it is quick, it is very accurate and its outstanding learning capabilities.

There are several variants of YOLO such as YOLO, YOLOv2, YOLOv3, YOLO9000 and so on. YOLO had a major drawback when it came to detection of small objects that occur in groups and detection of small objects. YOLOv2 fixes this problem. YOLOv2 increases the mAP (mean Average Precision) of the network by making use of batch normalization. YOLOv2 was a major upgrade to the YOLO algorithm. YOLOv2 is quite a fast network but more systems like the SSD (Single Shot Detector) also came into picture, which although slow, provides better accuracy than YOLOv2. YOLOv3 was proposed as an improvement. DarkNet-19 is used in YOLOv2, whereas DarkNet-53, which is far more complicated, serves as the core in YOLOv3. YOLOv3 overcomes the drawbacks of YOLOv2 and YOLO, especially in the detection of smaller objects. While YOLOv2 used to predict 5 bounding boxes per cell, YOLOv3 predicts only 3. YOLOv3 detects features at 3 different scales therefore bring up the bounding box count up to 9 [23].

YOLOv3 makes use of a single neural network for the entire image. The algorithm works on the basis of the scoring different regions of the image based on its resemblance with the predefined classes. The image is divided into grids and then a number of anchor boxes (bounding boxes) are created for each cell around the objects that score highly with the predefined classes.

#### 4. DENSE OPTICAL FLOW ESTIMATION AND GUNNAR FARNEBACK APPROACH

Optical flow, in simple terms can be defined as the motion of the individual pixels on the image plane [24]. It is a primary way of estimating the movement of picture intensities, which can be attributed to the movement of objects in the scene.

Optical flow estimation work on an assumption that the brightness of a particular point or a pixel in the pattern will be constant [25]. Let us consider the intensity at a point  $(x, y)$  at time  $t$  be  $I(x, y, t)$ . After a small time period has elapsed, say  $dt$ , the point would have moved to the position  $(x + dx, y + dy)$  and the intensity at this point can be given as  $I(x + dx, y + dy, t + dt)$ . Figure 1 provides illustrates the two cases discussed here.

According to the assumption, we have:

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad \dots \text{eq}(1)$$

Applying Tylor series approximation to the RHS we have:

$$\begin{aligned} I(x + dx, y + dy, t + dt) &= I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt \\ &+ \dots \quad \dots \text{eq}(2) \end{aligned}$$

When we substitute eq(2) in eq(1) we have:

$$0 = \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt$$

or,

$$\frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt = 0$$

Dividing LHS and RHS by dt, we have:

$$\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} = 0 \quad \dots \text{eq}(3)$$

where,

$$u = \frac{dx}{dt} \text{ and } v = \frac{dy}{dt}$$

The equation obtained in eq(3) is called the optical flow constraint equation. This equation simply defines the rate of change of intensity, or in other words how fast the intensity changes moving across the image.

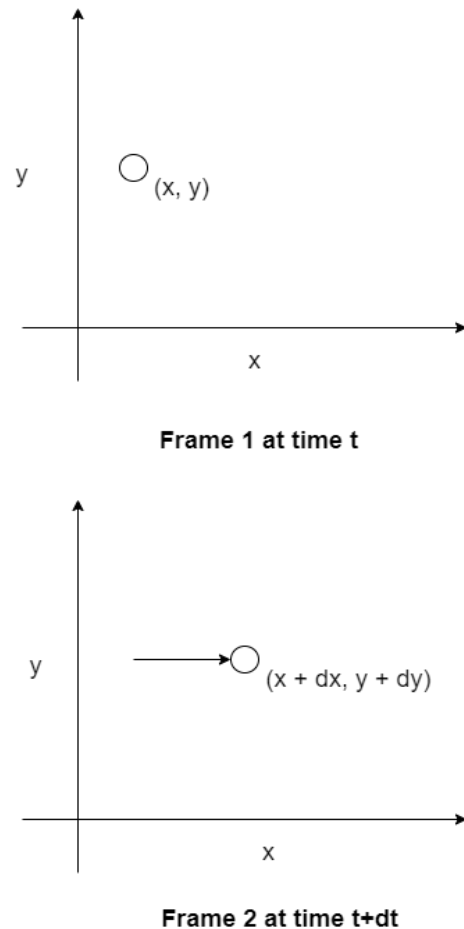


Figure 1. Optical flow estimation problem

There are two methods to solve this equation, namely the sparse optical flow and the dense optical flow. As the name suggests the sparse optical flow implementation will consider only the sparse set of pixels from the image for tracking the motion vectors. The sparse optical flow is often implemented using the Lucas Kanade method, one such implementation is provided in [26].

Dense Optical Flow, on the other hand, finds the optical flow vector for each pixel in a given frame, this accounts for the slow pace but at the same time the results are more precise. Dense optical flow estimation may be utilized for determining movement in the videos and learn structure

from motion. There are myriad ways for the implementation of the dense optical flow estimation. The dense optical flow is implemented using Gunnar Farneback method. In [27] a two-frame motion estimate based on polynomial expansion is suggested by Gunnar Farneback. According to this literature, it is possible to determine a way to estimate the displacement fields from observing how an absolute polynomial transforms under translation and this leads to a resilient algorithm post a series of refinement. Figure 2 illustrates the use of the dense optical flow method. To look into dense optical flow, OpenCV provides a method called `calcOpticalFlowFarneback()`

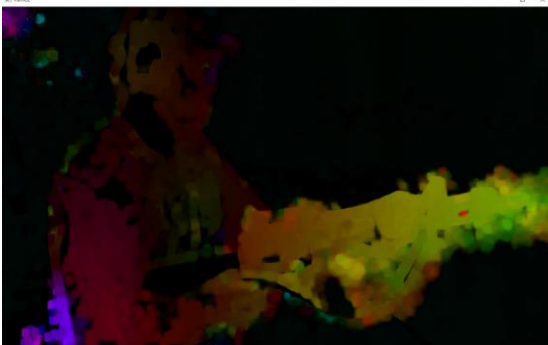


Figure 2. Dense optical flow

### 5. PROPOSED METHODOLOGY

Figure 3 shows the proposed methodology and the flow of events. The process flow starts at the camera that is installed in a particular location. The camera module provides a visual feed that is provided as input to the YOLOv3 module. YOLOv3 is used for the purpose of object detection, in this case we have trained the model for detecting custom objects. We have trained the model with 9 classes from the OIDv4 dataset, these classes correspond to the items, or more accurately the weapons that are used in most crimes. The nine classes are handgun, rifle, axe, scissors, kitchen knife, drill, hammer, dagger, knife. YOLOv3 model will detect for these objects from the live stream and when any of these objects are detected, a call to the dense optical flow model is made. This call is made for each frame in which the anomalous object is detected. In the call, the current captured frame is sent as a parameter for further processing. The dense optical flow model captures the next frame, and a processing is made to determine the optical flow. The function returns back the HSV converted frame to the YOLOv3 module for display. A counter is maintained externally to count the number of calls made to the dense optical flow function. This is achieved by making use of file handling and updating the count in a file for each call. Every time a call is made to the dense optical flow function, the counter in the dense optical flow function is updated and written into an external file.

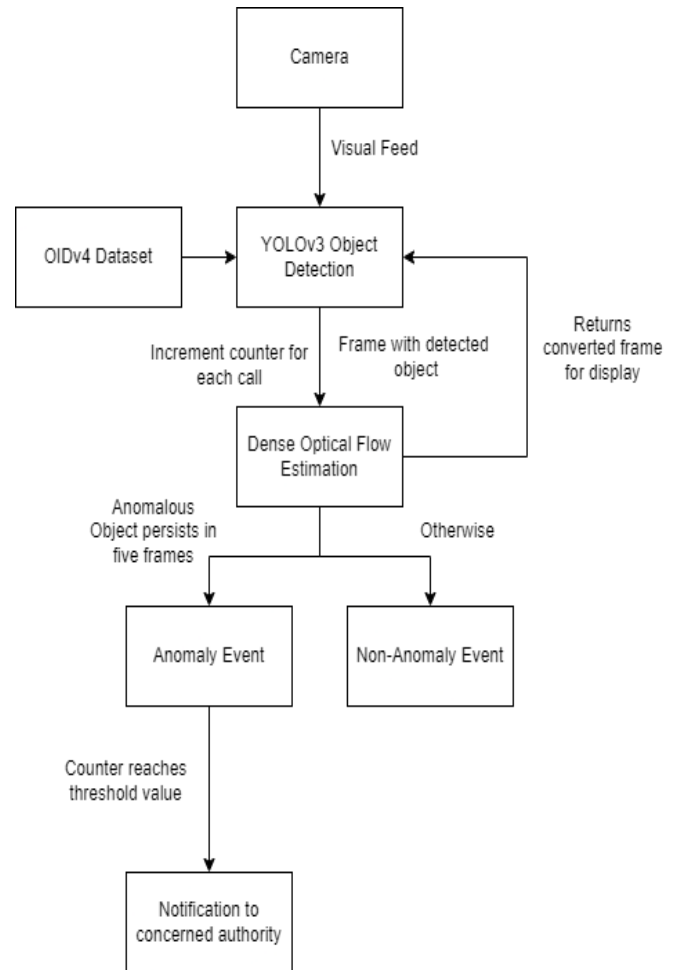


Figure 3. Proposed methodology and flow of events

Each call corresponds to a frame and therefore if X calls are made to the function, it means that X frames have elapsed. Further it is important to note that this call is only made when an anomalous object has been detected. When the counter reaches a specific value, say three, it means that the anomalous object has been spotted for three frames. We define a threshold value for the counter, once the threshold value is reached a notification is sent to the concerned authority regarding the anomalous activity. This notification is sent using the Twilio REST API, to the mobile of the concerned authority as an SMS.

### 6. RESULTS

The results can be broken down into two components – the training loss component and the accuracy component. These results are purely for the trained YOLOv3 module for the dataset containing nine classes. The dataset comprised of a total of 3852 images belonging to these nine classes. The model has been trained for 260 epochs. We define some common loss terminologies that are used for the task of segmentation and object detection. These include GIoU, objectness loss and classification loss. IoU or the Intersection over Union error is the most common evaluation metric used for the purpose of segmentation and object detection tasks. The Intersection over Union is calculated as follows:

$$IoU = \frac{|A \cap B|}{|A \cup B|}$$

For the smallest enclosing convex object C, the generalized Intersection over Union is given as:

$$GIoU = IoU - \frac{|C \setminus (A \cap B)|}{|C|}$$

Objectness loss occurs due to wrong box-object IoU prediction. Classification loss is caused by errors in predicting '1' for the proper classes and '0' for all other classes for the item in that box.

Table 1 summarises the error values observed during the training.

Table 1. Error values observed for the last training epoch

GIoU	Obj	cls	total
2.72	1.71	4.33	8.77

Accuracy is measured in terms of Precision, Recall and F1 score. Precision (P) answers the question as to what proportion of the positive identification are found to be correct? Recall on the other hand answers the question as to what proportion of actual positives was identified correctly?

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

In simple terms F1 score can be defined as the harmonic mean of the precision and the recall values.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Table 2 summarises the results observed during when the system was tested against inputs. The table holds the results for 25 sample results that the system was faced against.

Table 2. Results of event classification

Object Detected	Expected Classification	Obtained Classification	Result Type
Handgun	Anomaly	Anomaly	TP
Handgun	Anomaly	Anomaly	TP
Handgun	Anomaly	Not Anomaly	FN
No Weapon	Not Anomaly	Not Anomaly	TN
No Weapon	Not Anomaly	Anomaly	FP
No weapon	Not Anomaly	Not Anomaly	TN
Rifle	Anomaly	Anomaly	TP
Rifle	Anomaly	Anomaly	TP
Rifle	Anomaly	Anomaly	TP
Kitchen Knife	Anomaly	Not Anomaly	FN
Kitchen Knife	Anomaly	Anomaly	TP
Kitchen Knife	Anomaly	Not Anomaly	FN
Knife	Anomaly	Anomaly	TP
Knife	Anomaly	Not Anomaly	FN
Knife	Anomaly	Not Anomaly	FN
Dagger	Anomaly	Anomaly	TP
Dagger	Anomaly	Anomaly	TP
Dagger	Anomaly	Anomaly	TP
Scissors	Anomaly	Anomaly	TP
Scissors	Anomaly	Not Anomaly	FN
Scissors	Anomaly	Anomaly	TP
Hammer	Anomaly	Anomaly	TP
Hammer	Anomaly	Anomaly	TP
Hammer	Anomaly	Anomaly	TP

Axe	Anomaly	Anomaly	TP
Axe	Anomaly	Anomaly	TP
Axe	Anomaly	Not Anomaly	FN
Drill	Anomaly	Anomaly	TP
Drill	Anomaly	Anomaly	TP
Drill	Anomaly	Not Anomaly	FN

Abbreviations used in Table 2:

- TN – True Negative
- TP – True Positive
- FN – False Negative
- FP – False Positive

From the results table (Table 2) we make two important observations. One, the False Positive values, When there was no weapon of the listed category, there was a wrong classification made of a mobile phone as a handgun, which lead to this False Positive Value. Second, the False Negative values, although there was a presence of an anomalous object it was not detected.

Table 3 provides the confusion matrix for the classification. A confusion matrix is a matrix that helps in understanding how a particular classification tasks has performed. It helps us in determining the True Positive, True Negative, False Positive and False Negative responses. We plot a confusion matrix for anomaly versus non anomaly classification events.

Table 3. Confusion Matrix

	Anomaly	Non-Anomaly
Anomaly	19	8
Non-Anomaly	1	2

Calculating the values for precision, recall and the F1score:

$$Precision = \frac{19}{19 + 1} = 0.95$$

$$Recall = \frac{19}{19 + 8} = 0.70$$

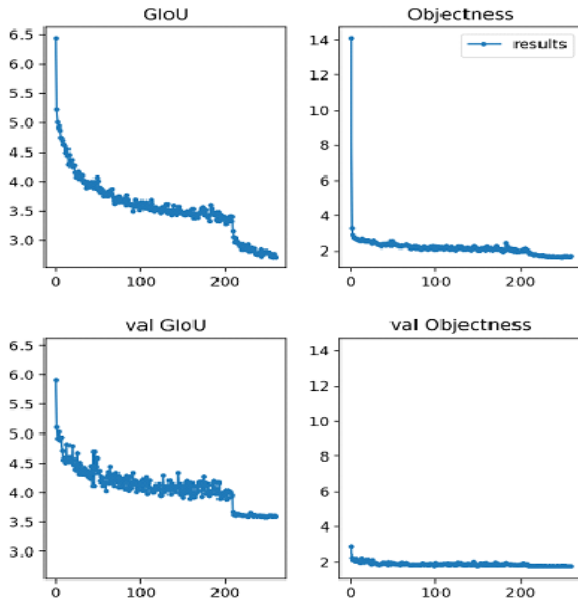
$$F1\ score = 2 * \frac{0.95 * 0.70}{0.95 + 0.70} = 2 * \frac{0.665}{1.65} = 0.806 \approx 0.81$$

Table 4 summarises the accuracy values for the predictions made.

Table 3. Accuracy Values

Precision	Recall	F1 Score
0.95	0.70	0.81

Figure 4 shows the error and accuracy trends that were observed during the training.



Figures 5 to 9 illustrates the outputs obtained on deploying the system in identification of the anomalous events.

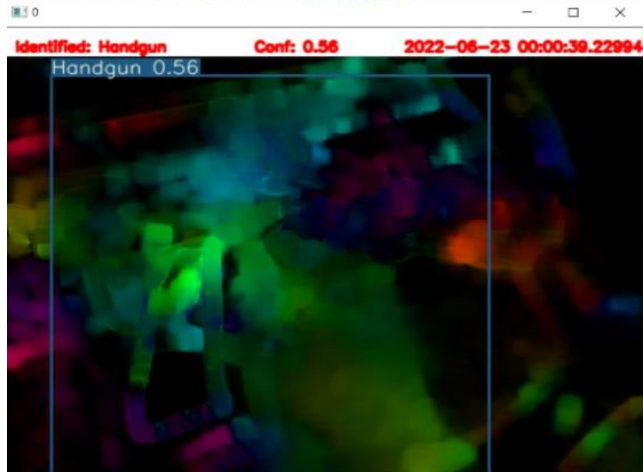


Figure 5. Detection of object followed by application of dense optical flow estimation.



Figure 6. Detection of object followed by application of dense optical flow estimation.



Figure 7. Detection of object followed by application of dense optical flow estimation.

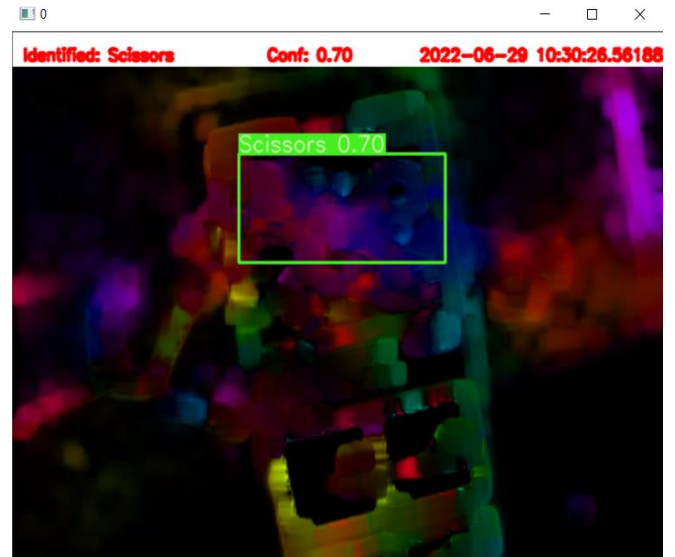


Figure 7. Detection of object followed by application of dense optical flow estimation.



Figure 8. Detection of object followed by application of dense optical flow estimation.



Figure 9. Anomalous object not detected

## 7. CONCLUSION

Security always remains an important concern for humankind. Despite various efforts, there is some kind of threat that is ever-present in the environment against which the person or the asset owned by the person has to be protected. Surveillance is an effective way to ensure security in any given scenario. Although CCTV cameras have been present in multiple areas, there is still some room for human errors. Computer Vision and its associated technologies can be incorporated to develop an intelligent solution for automating the process of surveillance. In the past several attempts have been made at developing systems that automate the process of surveillance, using myriad ways.

In this paper we have proposed and developed one such system that makes use of YOLOv3 object detection and dense optical flow estimation (Gunnar Farneback method) to detect anomalous events in the scenario. As mentioned earlier, we have made use of nine classes during the model training. These nine classes correspond to the nine most commonly used weapons in crimes. Detection of these weapons becomes important to keep the environment safe. Especially, gun crimes are a serious concern in many places. A system deployed to detect the presence of a gun can be very helpful in alerting the concerned persons. This can be just one example. The applications of such systems are vast, covering many important domains right from healthcare institutions to banking institutions. If we investigate closely, the weapons used in these different domains can vary, or additional classes may get added upon the current set of items. Here a transfer learning solution can be a great enhancement.

We have observed a need for storing the crime footages after detection of anomalous events, suppose, the police or any concerned authority requires any evidence in proving the crime the system and its output can come handy. In realising this, several concerns could be raised regarding the requirement of a high-capacity storage array and space complexity. A linked storage system and the mechanisms for handling the storage can be one of the major future scope of improvements in this perspective.

The dense optical flow estimation runs at an advantage for us to understand the movement of the entity in the scene. This will make analysis quite easy.

## Data Availability

The examples to the dataset used in the literature can be accessed here. The items in the dataset can be accessed using this toolkit.

## REFERENCES

- [1] Y. He, "Research on Text Detection and Recognition Based on OCR Recognition Technology," 2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE), 2020, pp. 132-140, doi: 10.1109/ICISCAE51034.2020.9236870..
- [2] F. Chadebecq, F. Vasconcelos, E. Mazomenos, and D. Stoyanov, "Computer vision in the surgical operating room," *Visceral Medicine*, vol. 36, no. 6, pp. 456-462, 2020.
- [3] L. Yeh and R. Chen, "Virtual Metrology of Visualizing Copper Microstructure Featured with Computer Vision and Artificial Neural Network," 2021 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), 2021, pp. 1-5.
- [4] W. Budiharto, V. Andreas, J. S. Suroso, A. A. S. Gunawan and E. Irwansyah, "Development of Tank-Based Military Robot and Object Tracker," 2019 4th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS), 2019, pp. 221-224, doi: 10.1109/ACIRS.2019.893.
- [5] J. Harikrishnan, A. Sudarsan, A. Sadashiv and R. A. S. Ajai, "Vision-Face Recognition Attendance Monitoring System for Surveillance using Deep Learning Technology and Computer Vision," 2019 International Conference on Vision Towards Emerging Trends in Com.
- [6] S. Ray, S. Das and A. Sen, "An intelligent vision system for monitoring security and surveillance of ATM," 2015 Annual IEEE India Conference (INDICON), 2015, pp. 1-5, doi: 10.1109/INDICON.2015.7443827.
- [7] W. Dai and J. Ge, "Research on Vision-based Intelligent Vehicle Safety Inspection and Visual Surveillance," 2012 Eighth International Conference on Computational Intelligence and Security, 2012, pp. 219-222, doi: 10.1109/CIS.2012.56.
- [8] W. Dai and J. Ge, "Research on Vision-based Intelligent Vehicle Safety Inspection and Visual Surveillance," 2012 Eighth International Conference on Computational Intelligence and Security, 2012, pp. 219-222, doi: 10.1109/CIS.2012.56.
- [9] A. A. Abdulhusein, H. K. Kuba and A. N. A. Alanssari, "Computer Vision to Improve Security Surveillance through the Identification of Digital Patterns," 2020 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), 2020, pp. 1-5, doi: 10.1109/ICIEAM48468.2020.9112022.
- [10] M. H. Rohit, "An IoT based System for Public Transport Surveillance using real-time Data Analysis and Computer Vision," 2020 Third International Conference on Advances in Electronics, Computers and Communications (ICAEC), 2020, pp. 1-6, doi: 10.1109/ICAEC50550.2020.9339485.
- [11] Tatachar, Abhishek & Kondur, Shivane & D, Dheeraj & K V, Vishwas & K, Sumukh. (2022). The Background Study for An Optical Flow Analysis Based Real Time Intelligent Video Surveillance System for People Safety. *International Journal of Innovative Research & Growth*. 8. 621-627.
- [12] F. Ahmed, T. A. M. R. Shahriar, R. Paul and A. Ahammad, "Design and Development of a Smart Surveillance System for Security of an Institution," 2021 International Conference on Electronics, Communications and Information Technology (ICECIT), 2021, pp. 1-4, doi: 10.1109/ICECIT54077.2021.9641422.
- [13] N. Suba, A. Verma, P. Baviskar and S. Varma, "Violence detection for surveillance systems using lightweight CNN

- models," 7th International Conference on Computing in Engineering & Technology (ICCET 2022), 2022, pp. 23-29, doi: 10.1049/icp.2022.0587.
- [14] P. T. R. Thangaraj, P. P. U. R. M and B. Vadivelu, "Real-Time Handgun Detection in Surveillance Videos based on Deep Learning Approach," 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2022, pp. 689-693, doi: 10.1109/ICAAIC53929.2022.9793288.
- [15] M. K. Hossen and S. H. Tuli, "A surveillance system based on motion detection and motion estimation using optical flow," 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV), 2016, pp. 646-651, doi: 10.1109/ICIEV.2016.7760081.
- [16] Ammar, A.; Fredj, H.B.; Souani, C. Accurate Realtime Motion Estimation Using Optical Flow on an Embedded System. *Electronics* 2021, 10, 2164. <https://doi.org/10.3390/electronics10172164>
- [17] P. D. Garje, M. S. Nagmode and K. C. Davakhar, "Optical Flow Based Violence Detection in Video Surveillance," 2018 International Conference On Advances in Communication and Computing Technology (ICACCT), 2018, pp. 208-212, doi: 10.1109/ICACCT.2018.8529501.
- [18] M. K. Hossen and S. H. Tuli, "A surveillance system based on motion detection and motion estimation using optical flow," 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV), 2016, pp. 646-651, doi: 10.1109/ICIEV.2016.7760081.
- [19] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster R-CNN: towards real-time object detection with region proposal networks. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'15)*. MIT Press, Cambridge, MA, USA, 91–99.
- [20] H. Ren and Z. Li, "Object detection using edge histogram of oriented gradient," 2014 IEEE International Conference on Image Processing (ICIP), 2014, pp. 4057-4061, doi: 10.1109/ICIP.2014.7025824.
- [21] V. R. A G, M. N and D. G, "Helmet Detection using Single Shot Detector (SSD)," 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), 2021, pp. 1241-1244, doi: 10.1109/ICESC51422.2021.9532985.
- [22] M. Ahmadi, Z. Xu, X. Wang, L. Wang, M. Shao and Y. Yu, "Fast Multi Object Detection and Counting by YOLO V3," 2021 China Automation Congress (CAC), 2021, pp. 7401-7404, doi: 10.1109/CAC53003.2021.9727949.
- [23] "Yolo: Real-time object detection explained," V7. [Online]. Available: <https://www.v7labs.com/blog/yolo-object-detection#h3>. [Accessed: 24-Jun-2022].
- [24] Pavan Turaga, Rama Chellappa, Ashok Veeraraghavan, *Advances in Video-Based Human Activity Analysis: Challenges and Approaches*, Editor(s): Marvin V. Zelkowitz, *Advances in Computers*, Elsevier, Volume 80, 2010, Pages 237-290, ISSN 0065-2458, ISBN 9780123810250, [https://doi.org/10.1016/S0065-2458\(10\)80007-5](https://doi.org/10.1016/S0065-2458(10)80007-5).
- [25] Horn, Berthold & Schunck, Brian. (1981). Determining Optical Flow. *Artificial Intelligence*. 17. 185-203. 10.1016/0004-3702(81)90024-2.
- [26] D. Patel and S. Upadhyay, "Optical flow measurement using Lucas Kanade method," *International Journal of Computer Applications*, vol. 61, no. 10, pp. 6–10, 2013
- [27] Farneback, Gunnar. (2003). Two-Frame Motion Estimation Based on Polynomial Expansion. In: *Image analysis*. 2749. 363-370. 10.1007/3-540-45103-X\_50