DOI: https://doi.org/10.51846/vol5iss4pp36-43 **Pakistan Journal of Engineering and Technology, PakJET**
ISSN (p): 2664-2042, ISSN (e): 2664-2050
Volume: 5, Number: 4, Pages: 36- 43, Year: 2022

Smart Surveillance and Detection Framework Using YOLOv3 Algorithm

Hassan Zaki[1], Muhammad Kashif Shaikh[1], Muhammad Tahir[1], Muhammad Naseem[1], Muzammil Ahmed Khan [2]

[1]Department of Software Engineering, Sir Syed University of Engineering & Technology, Karachi, Pakistan.
[2]Department of Computer Engineering, Sir Syed University of Engineering & Technology, Karachi, Pakistan.

Corresponding author: Muhammad Kashif Shaikh (e-mail: mkshaikh@gmail.com).

Received: 13-07-2022 Revised: 21-11-2022 Accepted:13-12-2022

Abstract- **The outlines of a video taken from an observation camera are becoming interesting these days, and much research work is to be done with fruitful results. The objective of this research is also to take care of such methods, which could be used to locate, identify and admit such activities by wasting no time, which could be done by getting the outlines of a video from an observation camera. The crucial point of this article is to explore the research more and consider such activities in real-time to have the video from a reconnaissance camera. The activity could get the message, follow the given time space and accordingly provide an activity based on that time-space in a single format. YOLO provides a better platform with better strategy and fast response than the localization within the customized dataset. The research is based on the new results, and analysis is given for insight into the proposed work. The given method proves the activity-based analysis and shows its application's suitability. The results are developed, showing a better agreement with the previous work and faster response time. The proposed work is applicable in many different ways, in shopping malls, automatic teller machines, corporate offices, residential and societies. The work is also useful in detecting ideal human actions.**

Index Terms**-- Video Analytics, Human Action Recognition, action label, deep learning, Custom Dataset, You Only Look Once (YOLO), Convolutional Neural Network.**

I. INTRODUCTION

Although, despite the later improvements in profound learning, exceptionally few profound learning-based strategies have been proposed to handle the issue of viciousness location from recordings. Convolutional Neural Systems (CNN) may be a strategy that's broadly embraced by analysts all over the world. For picture classification issues owing to the colossal victory of CNNs in examining a picture and its substance, examiners have initiated utilizing CNN for video investigation to a more noticeable degree. Profound learning methods are not application-specific, unlike the hand-crafted feature-based strategies, since a profound neural arrangement demonstrates can be effectively connected for a diverse errand without any critical changes to the design.

With moved forward execution, numerous diverse methods for tending to issue question location, following, acknowledgment activity acknowledgment, caption era, etc., have been created as a result. In this term paper, we proposed an idea that can consequently screen reconnaissance recordings and recognize the savage behaviour of people that will be of significant assistance to the law-and-order establishment. If you're utilizing these days, the rates of savage violations have extended profoundly, a fear assault includes one or numerous people with weapons and blades or it may be a battle or capturing.

That has come about in the colossal utilization of reconnaissance cameras that made a difference to the specialists in recognizing savage assaults and performing the vital steps to play down lamentable impacts. In our get-to, our appearance can notice and limit works out from fewer video diagrams (regularly in fact reasonable from disconnected diagrams). The appearance we proposed here allows an unprecedented activity title and certainty score for each diagram with an around-the-world world movement title for video gathering is gotten by finding the visit development title [1].

Occasional outlines from the video arrangement are prepared but not all the videos. Too, in most of the basics, a separate or few outlines is adequate for acknowledgment of the activity of people shown within the video [2].

In case of quick lessening of certainty grade, more outlines from the video are included to recognize. Within the display work, we have utilized the least number of outlines in this way decreasing the calculation time complexity [1]. As shown in Fig. 1.

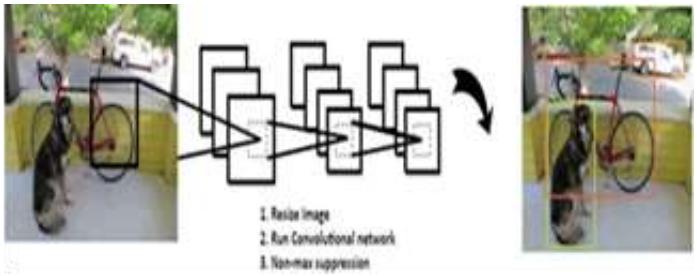


FIGURE 1. The YOLO Location Framework. Handling pictures with YOLO is basic and clear

II. BACKGROUND

In the recent past, researchers have tested a wide variety of technologies to find a breakthrough in the field of video surveillance. A review of existing literature and research was performed to support the study undertaken in this report. In modern video reconnaissance, computer vision strategies have regularly been utilized to distinguish ranges of intrigued since of the wealthy data substance contained in a picture but in some cases, video is required to efficiently detect the action performed [1].

A. VIDEO SURVEILLANCE

Investigate concepts of multistate spatiotemporal following utilizing real-time video investigation, numerous protest models, dynamic cameras, and long-term design examination to supply comprehensive circumstance awareness.[2] Kiryati et al. said that video surveillance systems make colossal wholes of data for capacity and show [3].

B. ALGORITHM HISTORY

Among the profound learning calculations utilized in computer vision, YOLO version 4 [4] is the foremost alluring for activity acknowledgment. Two truths are securing this choice. Firstly, YOLO version 4 [5] outlined its reasonability compared to other contradicting disclosure calculations [6]. It also sets an able-finding time (up to 45 takes per minute). This grants real-time certification of the supplication positions. Internal parts the taking after subsections, we show up the orchestrating of the elemental format of YOLO (YOLO adaptation 1[5]) and the unmistakable developments actualized in YOLO form 2 [7], YOLO adaptation 3 [8], and YOLO adaptation 4 [9].

a) YOLO V1

In 2016 YOLO was to start with displayed as an approach to understanding the Nonexclusive address area issues. This approach is based on arranging since it was a single CNN for the errand of the following and the assignment of investigation at the same time. YOLO version 1 accommodates 24 conditional layers for uniting induction and two related layers for making exceptional yield [10].

Into a lattice of SS * SS, YOLO separates the input picture. Each lattice cell can be related as if it were one question. Lattice cells have settled number B of boundary boxes to be expected for this dissent. For each line box, a certainty score is calculated. That comes almost as, for each cross-section cell YOLO calculates, the vector of lesson chances among class C we are centring on [4]. YOLO calculates, for each bounding box of the cell, a vector that contains five parameters: (x, y, w, h, and box certainty score). Hence, for each input picture, the YOLO organize makes a tensor of the outline:

$$SS \times SS \times (C1 + B1 * 5)$$

where:

SS x SS: Compares to numbers of grid cells.
B1: Compares the numbers of the bounding boxes.
C1: Compares the numbers of the targeted classes.

The 2nd may be a Localization Misfortune (the position of surveyed detention box compared to ground fact). 3rd can be a certainty occurrence (Box certainty rate correlated to the ground fact). We have utilized a combination of 3 misfortune capacities to prepare the YOLO organize. As shown below in the equation. [2]

$$\lambda_{coord}\sum_{i=0}^{s^2}\sum_{j=0}^{B}1_{ij}^{obj}\left[(x_i-\hat{x}_i)^2+(y_i-\hat{y}_i)^2\right]$$
$$+\lambda_{coord}\sum_{i=0}^{s^2}\sum_{j=0}^{B}1_{ij}^{obj}\left[\left(\sqrt{(w_i)}-\sqrt{(\hat{w}_i)}\right)^2+\left(\sqrt{(h_i)}-\sqrt{(\hat{h}_i)}\right)^2\right]$$
$$+\sum_{i=0}^{s^2}\sum_{j=0}^{B}1_{ij}^{obj}(c_i-\hat{c}_i)^2$$
$$+\lambda_{noobj}\sum_{i=0}^{s^2}\sum_{j=0}^{B}1_{ij}^{noobj}(c_i-\hat{c}_i)^2$$
$$+\sum_{i}^{s^2}1_{ij}^{obj}\sum\left(p_i(c_i)-\hat{p}_i(c_i)\right)^2$$

Equation: Cruel Normal Exactness [2] (1)

When YOLO was presented, it outperformed the other protest location calculations concerning speed. Its mAP (cruel Normal Exactness) was equal to or surpassed the mAP of other advanced calculations [2].

b) YOLO V2

In YOLO version 2, numerous advancements were presented to in- wrinkle the precision and diminish the handling time [6] with them, we can notice,

1. Use of Clump Normalizing (BN). In 2015 this strategy was proposed to cure the misfortune amid preparation. BN was included in all the conditional layers in YOLO, which moved forward mAP by 2.
2. The input picture from estimate 224 * 224 was supplanted by 448 * 448, which progressed the mAP by 4.

3. Choice of the conditional with the hook boxes. The course desire lifted from the level of the system cell to the level of the line box, which made several small increases in mAP by an edge of 0.3.
4. To utilize the K-means combination calculation, choose the finest stay box from the prepared set of line boxes. Euclidean removal is supplanted by the IoU (Crossing point Over Union) for the combination.
5. Forecasts are made on offsets to grapples. Rather than foreseeing (x,w,y,h,C1), YOLO version 2 forecasts (tx, ty, tw, th, tC). This makes the organization meet better.
6. To utilize fine-grained highlights. Essentially, to the personality mapping in ResNet [9], YOLO version 2 joins moo determination and tall determination highlights to progress the capacity to categorize little objects. This progresses the mAP by 1.
7. The selection of multiple scale preparation. Instead of employing a settled measure for the input picture, the YOLO version 2 arbitrarily chooses the picture to measure every 10 bunches, which moves forward the capacity to anticipate the well over the diverse capacity of the input pictures.

c) YOLO V3

As an incremental change to the past forms in April 2018, the YOLO version 3 [3] was proposed. Among the enhancements built, we can note,

1. To utilize the multi-label arrangement. Rather than the collective elite labelling within the past forms, A calculated classifier is utilized in YOLO version 3 to appraise the likeliness of a question being of a particular name. Classification misfortune is being changed to need the twofold cross entropy misfortune for each name rather than the common cruel square misfortune utilized in past forms [10].
2. To utilize another line box desire. In the arrangement, the objectless score 1 in YOLO version 3 is related to the line box remains that best covers the ground truth dissent. It is overlooked if the Crossing point Over Union (IoU) between the line box snare and the ground reality is less than control (0.7 interiors the operation). Interior the conclusion, its assistants for each ground reality address one line box snare [2].
3. To alter the yield 3D tensor. The expectation is to exhaust YOLO version 3 for one lattice cell at three diverse
4. Scales and, after that, concludes the extraordinary bounding box from those scales. This was induced by the included pyramid systems [3]. In this way, the cutting-edge estimation of the abdicate 3d sensor is at that point:

$$SS \times SS \times (3 \times (5 + C1)) \quad (3)$$

where:

SS × SS: Compares to the number of network cells.

B1: is excluded since just 1 line box mainstay is retained at the end.

C1: Compares s to the number of aimed classes.

The selection of the unused includes an extractor (Darknet-53). It has fifty-three (53) layers and employments skip connections like the ResNet. [11] It employments both the (3x3) and (1x1) convolutions. YOLO version 3 gave state-of-the-art exactness alongside more and way better speed and fewer computations.

d) YOLO V4

In April 2020, the YOLO version 4 [12] arrived with the demanding and curious updates demonstrating the finest state-of-the-art finders in precision and inference speed [10]. On those enhancements, able to note that,

1. The influence of state of art Bag of Freebies (BoF) and some Bags of specials (BoS).
2. The YOLO version 4 acquires an AP value of (43.5) percent (65.7) percent AP50 [9].
3. For batch normalization, Cross Mini-Batch Normalization (CmBN) is used with the idea that this can be run on any GPU that people use [11].
4. YOLO version 4 employs Drop Block regularization. In Drop Block, segments of the picture are covered up from the primary layer. It may be a procedure to drive the organization to memorize highlights that it may not be something else dependent upon [11].
5. On the Tesla V100, it has a real-time speed of (65 FPS). It beats the fastest and the most accurate detectors in terms of both Accuracy and Speed [9].

III. METHODOLOGY

YOLO stands for ("You Only Look Once") and is one of the broadly embraced calculations in the world of computer vision since it accomplishes tall exactness together with being able to prepare in genuine time. [4]

These charts are chosen after a cycle of a certain number of takes which builds upon the included numbers of takes interior the video. For trial purposes, we have chosen (30) charts in a video for assertion and localization of works out from the trial video charts. We arranged to illustrate utilizing diagrams containing suitable activity from the planning set of distinctive datasets (se Fig. 2) [6].

FIGURE 2. Detection of an object using an anchor box [18]

After performing movement location on each outline, the movement title is distinguished in more than five outlines and having a certainty edge of 0.6 over 30 outlines is restricted as the closing activity title inside the video. As shown in Figure 3.

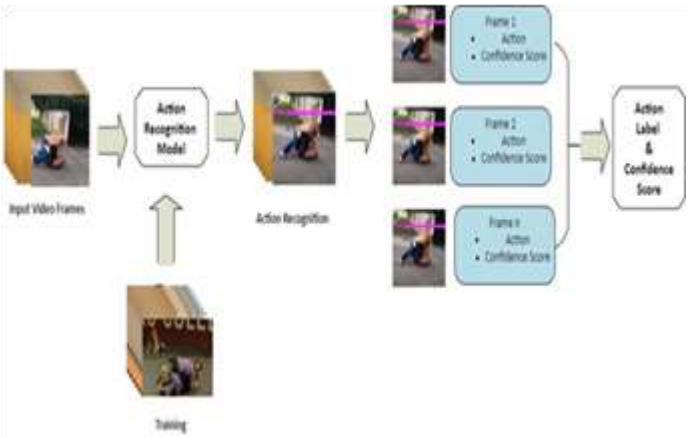


FIGURE 3. Steps to detect the YOLO architecture YOLO [8]

A. AN OBJECT DETECTION ALGORITHM

You Only Look Once (YOLO) is a modern point of view on cross-disclosure. The YOLO method considers object detection as a regression problem. The YOLO calculation exercises a separate CNN to the complete picture, which encourages partitions of the picture into lattices. YOLO specifically predicts the lesson possibilities and the line box offsets from full pictures with a single feed-forward curlicue neural. It kills the locale proposition era, includes resampling, and typifies all stages in a single arrangement to create a genuine end-to-end location framework [5].

YOLO version 3 strategy separates the input pictures into (SS × SS) little network cells. If the center of the protest falls into a network cell, the network cell can identify the protest. Each lattice cell concludes the position data of B bounding boxes and computes the question scores compared to these line boxes (see Fig. 4) [6].

YOLO engineering has 24 conditional layers and 2 completely associated layers. The YOLO takes a picture as an input and resizes it to (448 x 448) by keeping the angle proportion the same and performing cushioning [10].

Then this picture is passed within the CNN organize. The show contains twenty-four convolution layers and four max-pooling layers taken after by two completely associated layers. For the lessening of the number of layers (Channels) [13], we use a 1*1 convolution that's taken after by a 3*3 convolution [10].

Notice that the final layer of YOLO predicts a cuboidal yield from the ultimate completely associated layer, reshapes it to size and gives yield within the shape of a 7×7×30 tensor. The tensor gives data around the arranges of the bounding box's rectangle. The likelihood conveyance over all classes the framework is prepared. Holding this certainty notch (probability) disposes lesson names notching less than 30% (see Fig. 5).

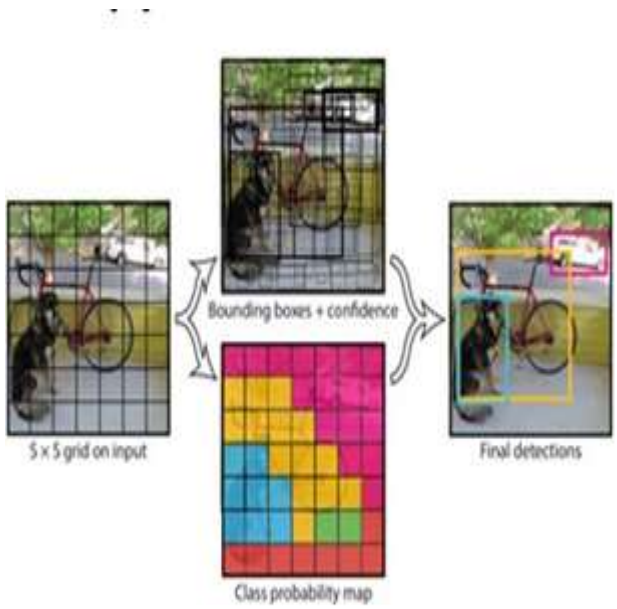
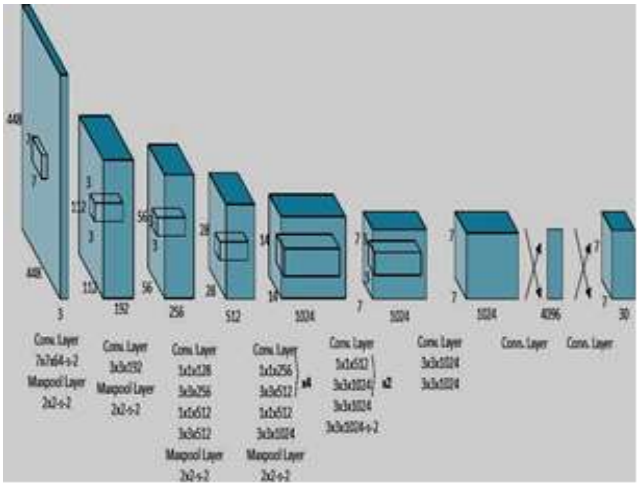


FIGURE 4. A simplified illustration of the YOLO object detection pipeline [18]



IGURE. 5. The YOLO Architecture [13]

B. YOLO ALGORITHM OVER OTHER DETECTORS

However, it isn't the most excellent in terms of precision of question location calculations. Still, it's a great choice in case we require real-time discovery without the misfortune of much exactness. YOLO (Yolo only looks one) belongs to one of the quicker protest location calculations within the field of computer vision. [14]. Flow diagram elaborates that live streaming is done through a camera from which detection and recognition of a person, objects, and activity. Each record is entered into the database, generating a voice-based message alert. No alert is generated if no harmful object or unusual activity is detected.

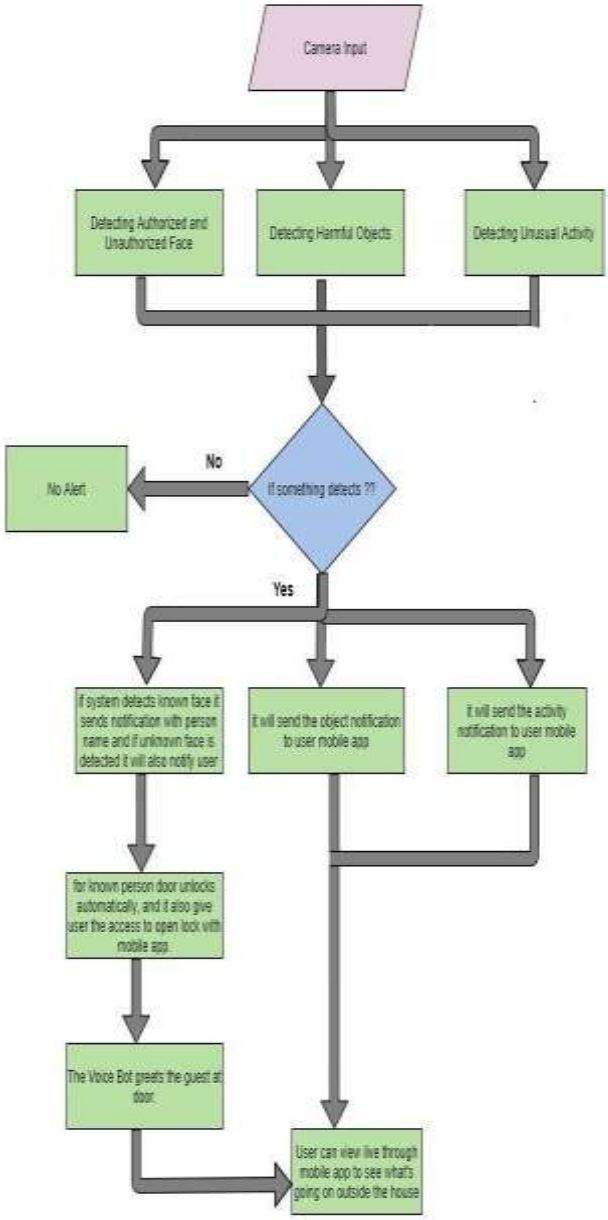


FIGURE 6. Error Analysis: YOLO vs. Fast R-CNN [16]

For localization and classification of objects instead of using the 2-steps technique. YOLO applies a single CNN for both courses of action and the following of the protest. The picture is separated into locales; it predicts bounding boxes and probabilities for each locale [15].

YOLO applies a single CNN for both courses of action. It shapes graphs at around 40 to 90 FPS, so it is especially rapid. This suggests that the gushing video can be analyzed in honest-to-goodness time, with remarkably insignificant dormancy in sensible many milliseconds. The building of YOLO makes it astonishingly quick [15]. Compared with R-CNN, it is 1000 times speedier and 100 times speedier than Expedient R-CNN (see Fig. 6).

YOLO may be a common- reason finder that learns to distinguish an assortment of objects at the same time as compared to Quick and Quicker R-CNN center on speeding up the R-CNN system by sharing computation and utilizing the Neural Systems to propose the locales rather than Particular Look. Execution [16].

The goal of object detection is to detect all objects and class their objects. It has been widely used in autonomous driving [21].

We propose a YOLOv3-based improved video processing [22] detection approach to increase robustness and eliminate the time-consuming process [23].

Nowadays, with the accessibility of huge datasets, quicker GPUs, advanced machine learning algorithms, and better calculations, we can effectively prepare PCs and develop automated computer-based systems to distinguish and identify numerous items on a site with high accuracy. Recent developments indicate that machine learning [24] and advanced image processing algorithms have dominant roles in smart surveillance and security systems [25].

These charts show the rate of localization and foundation blunders within the best N discovery for different categories (N = # objects in that category Fig. 7) [4].

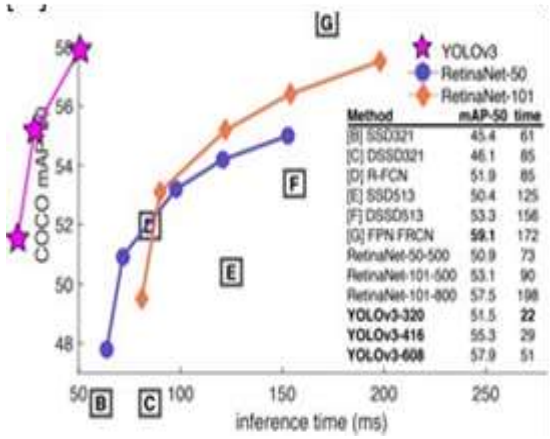


FIGURE 7. Once more, adjusted from this time show in speed/accuracy tradeoff on the mAP at .5 IOU metric. You'll be able to tell YOLO version 3 is exceptionally great because it is extraordinarily tall and distant from the cleared-out. [4]"

IV. DATASET

Surrounded by all the datasets present today, most of them belong to,

- Fight dataset (Hockey fight detection dataset [18]).
- Irregular recordings of battle scenes from the web. The whole number of collected pictures is around 500 we physically labelled them into one lesson utilizing rectangular bounding boxes. Dataset table I Number of Pictures. As shown in Figs. 8-11.

FIGURE. 8. Sample Dataset I [18]

FIGURE 9. Sample Dataset II [18]

FIGURE 10. Sample Dataset III [18]

FIGURE 11. Sample Dataset III [18]

TABLE I

TRAINING DATASET AND TESTING DATASET, RATIO

TRAINING DATASET AND TESTING DATASET

	Training Dataset	**Testing Dataset**
No. of Images	400	100
Percentage	80%	10%

Frame Extraction For the training of the YOLO model, we need frames of fighting scenes in image form. To extract those images from a video, we must do frame extraction using any available software (like VLC media player). A few test pictures of the dataset appear in Figure 2, and Table I shows up the number of pictures and the events inside the organizing and testing datasets.[20].

V. TRAINING

The stream chart for the hone of YOLO-based activity discernment show has appeared in Fig. 12.

The preparation was done utilizing Google Collaboratory so that we could get a 12GB NVIDIA Tesla K80 for speedier and more proficient network preparation. First, we made .txt records for each picture within the dataset for YOLO. Since YOLO requires a text record for each structure with a line for each activity within the structure. After preprocessing the dataset i.e., making name records for each picture, both pictures and their name records are to be kept together. Further, a few records that (YOLO) requires to begin preparing are:

- They add up to many movement classes.
- Text record with the way to all traces which we must be prepared.
- Text record with names of all movement classes.
- The way to spare organized weight records.
- A setup record (.cfg) with all layers of YOLO advancement. prepare-trained convolutional weights.

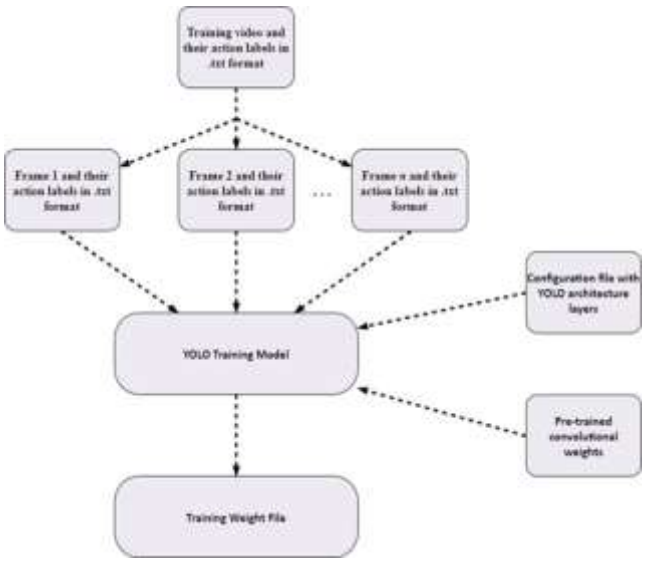


FIGURE 12. The Flow diagram for the training of YOLO

Yolo.cfg record was used to prepare the arrangements incorporating 3 YOLO layers. As a conventional strategy, each lesson must be prepared for at slightest 20emphasiszes. The bunch and the subdivision's values were set to 64 and 8 separately for ideal preparing speed [19]. The stature and width values were set at (416) each for ideal speed and superior exactness of location. The number of channels utilized within the convolution layer was set to 18 as the value is subordinate to the overall number of classes as, channels = (classes + 5) *3. For the moment

final layer, the value of channels within the setup record of YOLO (cfg record) isn't self-assertive and depends on the full number of classes. Around 7-8 hours were required to prepare the arrange with the over arrangements. The weights in this way were created after 20emphasizesses were utilized to carry out discoveries and analyze the execution. Charts and comes about are appeared in Figs. 13, 14, and 15.

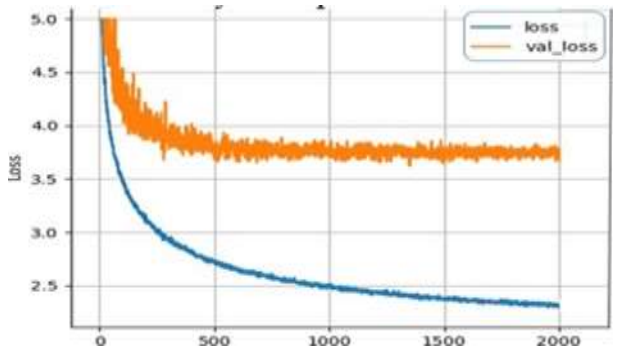


FIGURE 13. Training graph

VI. CONCLUSION & RESULTS

A. TRAINING AND TESTING

Inside the planning and organizing, each repetition took nearly 5.25s. We arranged to appear up to 500 ages and the typical misfortune was found to be 0.32 for the bunch assess of 32 with 8 segments. For more accuracy and to decrease the normal misfortune, we are ready to plan our show with a more vital number of repetitions.[17] For testing, the typical affirmation of action inside the layout takes roughly 70ms that's 13-15 FPS (Layout per minute). Hence, Genuine-time affirmation of movement is doable on off-the-rack desktop PCs with a mid-level Outlines Planning Unit (GPU).

B. QUANTITATIVE RESULT

In our research, the confusing system on the test dataset is shown in Table II, where lines talk to genuine exercises while columns talk to the expected activities.

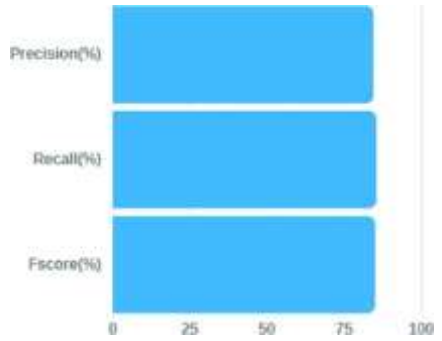


FIGURE 14. Exactness, review, F-score for activity lesson utilizing the disarray network

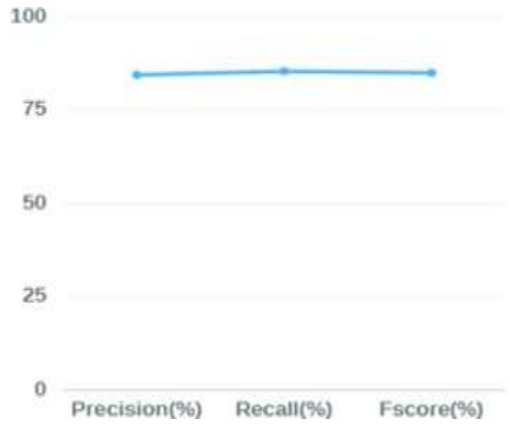


FIGURE 15. State-of-the-art calculations, Comparisons of execution

The exactness gotten by the technique proposed in this paper was found to be 85.83%. Exactness (84.4), Review (85.3), and F-score (84.9) for a lesson are displayed in figure 10. Table II shows that the proposed technique performs better when related to the F-scores of state-of-art methods. Interior the proposed approach, we recognized that the development takes put close to the camera and is less blocked.

TABLE II
ACCURACY FOR ACTION RECOGNITION AND LOCALIZATION

	Accuracy
LSTM	79.40%
ConvNet	83.90%
MLP	80.80%
KNN	65.90%
RF	78.10%
SVM	77.70%

FUNDING STATEMENT

The authors declare they have no conflicts of interest to report regarding the present study.

CONFLICT OF INTEREST

The Authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 779–788. DOI: 10.1109/CVPR.2016.91.

[2] "YOLO — You only look once, real-time object detection explained | by Manish Chablani | Towards Data Science." https://towardsdatascience.com/yolo-you-only-look-once-real-time-object-detection-explained-492dc9230006.

[3] D. Manian, "International Conference on Computer Networks and Communication Technologies Lecture Notes on Data Engineering and Communications Technologies," Springer, Jan. 2020,

[4] "Start Training YOLO with Our Own Data," Guanghan Ning's Blog, Dec. 22, 2015. http://guanghan.info/blog/en/my-works/train-yolo/.

[5] D. Manian, "International Conference on Computer Networks and Communication Technologies Lecture Notes on Data Engineering and Communications Technologies," Springer, Jan. 2020
[6] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," ArXi version 161208242 Cs, Dec. 2016.
[7] J. Redmon and A. Farhadi, "YOLO version 3: An Incremental Improvement," ArXi version 180402767 Cs, Apr. 2018,
[8] A. Koubaa et al., "Activity Monitoring of Islamic Prayer (Salat) Postures using Deep Learning," ArXi version 191104102 Cs, Nov. 2019,
[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," ArXi version 151203385 Cs, Dec. 2015
[10] F. F. Chamasemani and L. S. Affendey, "Systematic Review and Classification on Video Surveillance Systems," Int. J. Inf. Technol. Comput. Sci., vol. 5, no. 7, pp. 87–102, Jun. 2013, DOI: 10.5815/ijitcs.2013.07.11.
[11] SYNCED, "YOLO Is Back! Version 4 Boasts Improved Speed and Accuracy | Synced," Apr. 27, 2020. https://syncedreview.com/2020/04/27/yolo-is-back-version-4-boasts-improved-speed-and-accuracy/
[12] "YOLO Is Back! Version 4 Boasts Improved Speed and Accuracy | by Synced | SyncedReview | Medium." https://medium.com/syncedreview/yolo-is-back-version-4-boasts-improved-speed-and-accuracy-7e0b71bef5a9.
[13] J. S. JUN 29 and 2020 11 Min Read, "YOLOv5 New Version - Improvements And Evaluation," Roboflow Blog, Jun. 29, 2020. https://blog.roboflow.com/yolov5-improvements-and-evaluation/.
[14] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLO version 4: Optimal Speed and Accuracy of Object Detection," Apr. 2020, DOI: 10.48550/arXiv.2004.10934.
[15] "Data Augmentation in YOLO version 4." https://blog.roboflow.com/yolo version 4-data-augmentation/ [16] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," ArXi version 150602640 Cs, May 2016,
[16] L. Zhao and S. Li, "Object Detection Algorithm Based on Improved YOLO version 3," Electronics, vol. 9, no. 3, Art. no. 3, Mar. 2020, DOI: 10.3390/electronics9030537.
[17] Muhammad Kashif Shaikh, Dr. Sellappan Palaniappan, Engr. Fayyaz Ali, Engr. Dr. Muhammad Khurram. (2020). IDENTIFYING DRIVER BEHAVIOUR THROUGH OBD-II USING ANDROID APPLICATION. PalArch's Journal of Archaeology of Egypt / Egyptology, 17(7), 13636 - 13647.
[18] "Hockey Fight Detection Dataset," Academic Torrents. https://academictorrents.com/details/38d9ed996a5a75a039b84cf8a137be794e7cee89.
[19] Muhammad Kashif Shaikh; Sellappan Palaniappan; Touraj Khodadadi, An AI-driven automotive smart black box for accident and theft prevention. International Journal of Modelling, Identification, and Control (IJMIC), Vol. 39, No. 4, 2021.
[20] S. Kulkarni, and G. S. P, "Deep Learning Based Object Detection Using You Only Look Once," Int. J. Res. Advent Technol., vol. 7, no. 4, pp. 9–12, Apr. 2019, DOI: 10.32622/ijrat. 74201902.
[21] Zhao, Liquan, and Shuaiyang Li. 2020. "Object Detection Algorithm Based on Improved YOLOv3" Electronics 9, no. 3: 537. https://doi.org/10.3390/electronics9030537
[22] Diwan, T., Anirudh, G. & Tembhurne, J.V. Object detection using YOLO: challenges, architectural successors, datasets and applications. Multimed Tools Appl (2022). https://doi.org/10.1007/s11042-022-13644-y
[23] Abdusalomov, Akmalbek, Nodirbek Baratov, Alpamis Kutlimuratov, and Taeg Keun Whangbo. 2021. "An Improvement of the Fire Detection and Classification Method Using YOLOv3 for Surveillance Systems" Sensors 21, no. 19: 6519. https://doi.org/10.3390/s21196519.
[24] A. Joshi, N. Jagdale, R. Gandhi, and S. Chaudhari, "Smart surveillance system for detection of suspicious behaviour using machine learning," in Intelligent Computing, Information and Control Systems. ICICCS 2019. Advances in Intelligent Systems and Computing vol. 1039, Berlin, Germany, Springer, Cham, 2020.
[25] Narejo, Sanam & Pandey, Bishwajeet & Esenarro, Doris & Rodriguez, Ciro & Anjum, M.. (2021). Weapon Detection Using YOLO V3 for Smart Surveillance System. Mathematical Problems in Engineering. 2021. 1-9. 10.1155/2021/9975700.