iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50729

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 IV Apr 2023- Available at www.ijraset.com

Suspicious Activity Detection Using Different Models

Dr. Vaqar Ansari[1], Aditya Ghadge[2], Prasham shah[3], Sumon Ghosh[4], Vaibhav Sanghavi[5]
[1, 2, 3, 4, 5]*Under-graduating Electronics and Telecommunication Engineering at St Francis Institute Of Technology Mumbai*

Abstract: ***In today's insecure world, video surveillance systems play a significant role in keeping both indoors and outdoors secure. Real-time applications can utilize video surveillance components, such as behavior recognition, understanding and classifying activities as normal or suspicious. People are at risk from suspicious activities when it comes to the potential danger they pose. Detecting criminal activities in urban and suburban areas is necessary to minimize such incidents as criminal activity increases. The early days of surveillance were carried out manually by humans and involved a lot of fatigue, since suspicious activities were rare compared to everyday activities. Various surveillance approaches were introduced with the advent of intelligent surveillance systems. This paper analyzes two cases that could pose a threat to human lives if ignored, namely the detection of gun-related crimes, the detection of abandoned luggage, the detection of human violence, the detection of lock hammering, the theft of wallets, and the tempering of ATMs on surveillance video frames. In these papers they have used a neural network model that is Faster R-CNN and YOLOv3 technique to detect these activities.***

I. INTRODUCTION

Video surveillance systems are the only way to detect crimes such as stealing bags, abandoning bags on stations, stabbing with knives, and using guns, which are on the rise every day. However, video surveillance systems have the disadvantage of requiring continuous human attention, reducing their efficiency. Video surveillance has been automated to solve this problem. It is impossible to manually monitor all events on CCTV cameras today. A manual search in the recorded video would waste a lot of time, even if the event had already occurred. Automated video surveillance systems are investigating abnormal events from video footage. Video surveillance can be automated to solve this problem. Automated systems give indications in the form of alarms or other forms when predefined abnormal activities occur. As stated in the papers, they used a semantic based approach which involves defining suspicious activities, background subtraction, object detection, tracking & classification of suspicious activities within the framework of a system

II. LITERATURE SURVEY

In [1] paper, unavailability of datasets, generalizing ability of the classifier is overcomed using a semantic based approach. But, object detection accuracy is just 57%. In [2] paper, gun based crime detection and abandoned luggage detection is performed. But, detecting abandoned luggage does not address issues like identification of objects in sudden changes of illumination. In [3] paper, YOLOv3 technique outperforms R-CNN and accuracy achieved is around 95%. But, due to the small amount of data in training, there were still some mismatches in comparison between the test results and the ground truth.

III. RELATED WORK

Any suspicious movement can be detected through video surveillance, which acquires and processes the data. Many research studies have been conducted on the detection of anomalies in video data. Most researchers deal with the problem of abandoned bag detection. A framework for detecting abandoned objects in a scene with multiple interacting objects was described by James David Hogg et al. The datasets they use are the standard ones. Using Gaussian Mixture Models (GMM), the object (bag) is detected by the dual background approach. A modified multi hypothesis tracker is used for tracking extended objects. Based on the relationship between bags and people, a situation analysis is conducted. An approach based on logic is then used to assess threats. According to Fuentes & Velastin, a video surveillance algorithm is based on trajectories for detecting events.The position, trajectory, and split/merge events can be used to describe any event. Tracking is then done using the matching matrices. Through a single camera, Kim et al detect and track multiple moving objects. To extract moving regions, they use RGB color background modeling. Moving objects are grouped using the blob labeling. The foreground image is typically obtained by background subtraction in anomaly detection. The background subtraction technique is employed in our system because it does not require any prior training.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

In order to detect objects, most researchers use a machine learning approach. For training, a standard reliable dataset is required, which is difficult to obtain. The machine learning approach becomes less reliable as a result. The hierarchical semantic approach is used in our system. There will be a focus on areas such as early detection and recognition of activities. A method for predicting human activity is presented in the research paper. Their primary concern is recognizing events early (for instance, a man picking up a gun with his hand). A probabilistic activity prediction problem is formulated, and new methodologies are introduced to solve it. Spatio-temporal features are analyzed using an integral histogram. As a result of considering the sequential nature of human activities and handling noisy data, they named their new recognition methodology dynamic bag-of-words.

IV. METHODOLOGY

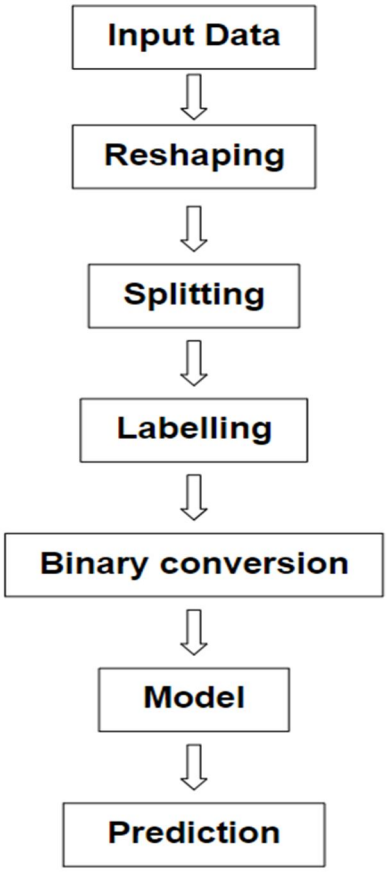


Figure 1: Block Diagram of Image Processing

A. Input Data

The dataset of images is given as an input which contains the images of three activities that are gun detection, knife detection and fight detection. This data is raw data which has images of different shapes, pixels. We were having the 55,000 images data, but due to less computing power we were able to use only 19,000 images. We converted fight videos to frames for the fight dataset from the following link: [https://github.com/seymanurakti/fight-detection-surv-dataset] The images for the gun and knife detection were collected from different sources on the internet. The link to it is given below: [https://drive.google.com/drive/folders/1bC3BmrRsx-_papIyvUbBvXpnR3Dsxr-s?usp=sharing]

B. Reshaping

We were having the images of different pixels in our dataset. Hence, we converted all the images to a constant pixels ratio that is 300x200.

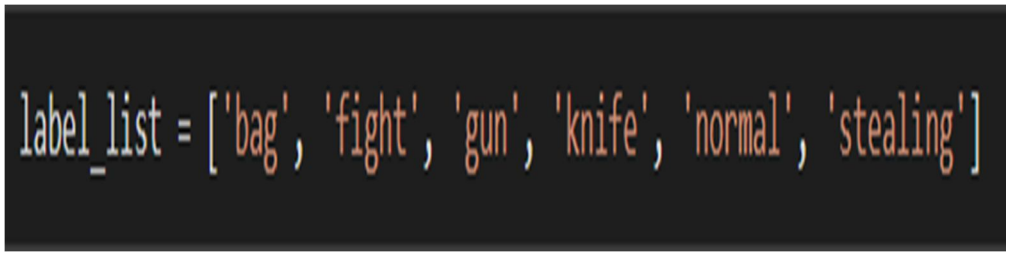
C. Splitting

The dataset is splitted in three subfolders: training (70%), testing (15%) and validation (15%).

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

D. Labelling



bag : 0
fight : 1
gun : 2
knife : 3
normal : 4
stealing : 5
Converted the input data to labelled input data from which the model will learn to predict the activity.

E. Binary Conversion

Read the images from training, testing and validation folders. Converted these images to the color images and then converted into binary format. After this step, our .npy binary converted file will be ready to provide as an input to the model.

F. Model

For the classification of these activities we have used three models that are simple CNN, ResNet50v2 and VGG19. These models will take the binary images as an input, extract the features from it and learn from it. Models will provide accuracy, loss, validation accuracy and validation loss. Based on this we can judge the performances of these models and find out the best model for prediction. These parameters are explained below:

1) *Accuracy:* Accuracy is the fraction of correct predictions made by the model on a given set of data.

Accuracy = (Number of Correct Predictions) / (Total Number of Predictions)

2) *Loss:* Loss is a measure of how well the model is able to predict the target variable. It is the error between the predicted value and the actual value of the target variable. The loss is usually computed using a loss function, such as mean squared error (MSE) or cross-entropy loss.

Loss = Loss Function(Predicted Value, Actual Value)

For example, the mean squared error (MSE) loss function can be defined as:

$MSE = (1 / n) * \Sigma(y_pred - y_actual)^2$

where ‘n’ is the number of examples, ‘y_pred’ is the predicted value, and ‘y_actual’ is the actual value.

3) *Validation Accuracy:* Validation accuracy is the accuracy of the model on a validation set, which is a set of data that is not used for training but is used to evaluate the model's performance. The validation accuracy is a measure of how well the model is able to generalize to new, unseen data.

Validation Accuracy = (Number of Correct Predictions on Validation Set) / (Total Number of Predictions on Validation Set)

4) *Validation Loss:* Validation loss is the loss of the model on a validation set. It is a measure of how well the model is able to predict the target variable on new, unseen data. The validation loss is usually used to monitor the model's performance during training and to prevent overfitting.

Validation Loss = Loss Function(Predicted Value on Validation Set, Actual Value on Validation Set)

For example, the cross-entropy loss function can be defined as:

$Cross\text{-}entropy\ loss = -\Sigma(y_actual * \log(y_pred) + (1 - y_actual) * \log(1 - y_pred))$

where ‘y_pred’ is the predicted value and ‘y_actual’ is the actual value, and the sum is taken over all examples in the validation set.

G. Prediction

Model will predict the activity in the image with some accuracy label. For the following image, this image is detected as a fight with 99.84% accuracy.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



Figure 2: Prediction of the activity

To follow the above steps, code is given below: [https://github.com/aditya423/Suspicious-Activity-Detection/tree/main/project]

V. SIMULATION

From the confusion matrix, we can calculate various metrics such as accuracy, precision, recall, and F1- score, which provide insight into the performance of the classification model. Hence, first we will take a look at the confusion matrices of these models.

1) Simple CNN:

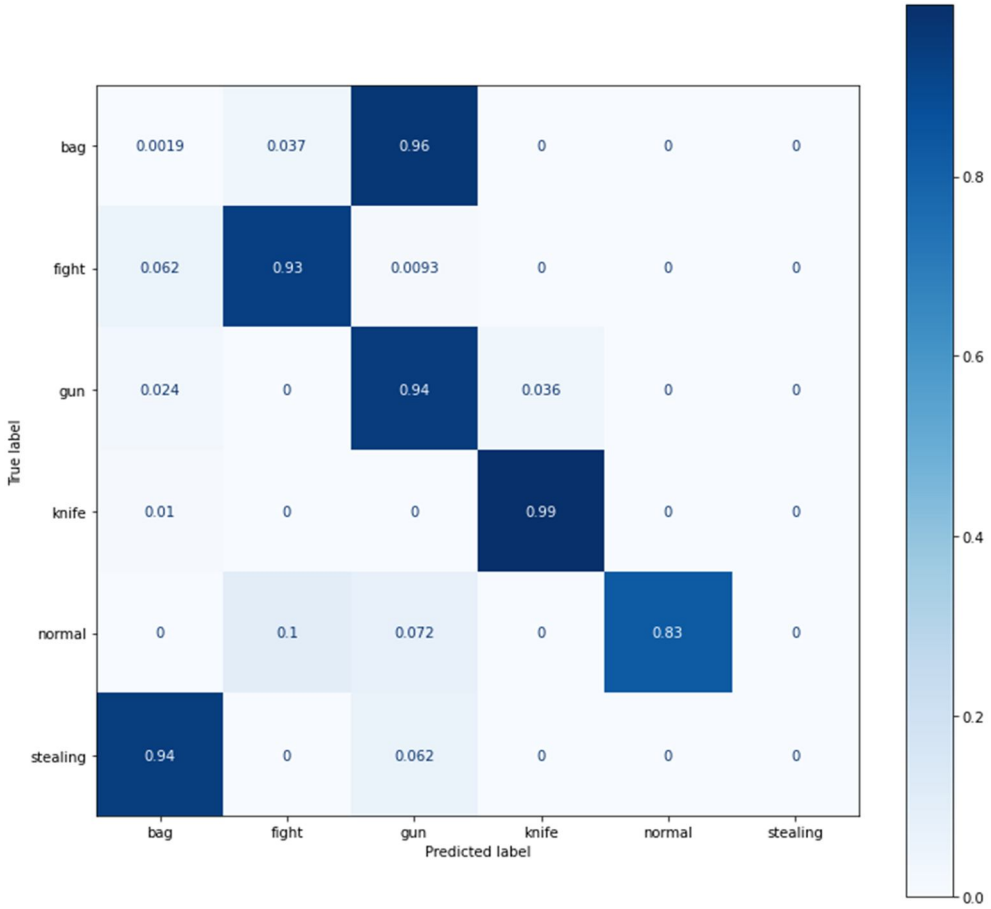


Figure 3: Confusion matrix for simple CNN model

Validation Accuracy: 77.2232

From the observation of this matrix, we get that bag and stealing activity is not detected correctly and hence we removed these categories to increase the overall accuracy of the model. We are able to detect only three activities correctly with good accuracy and that are gun detection, knife detection and fight detection.

Confusion matrices after excluding bag and stealing activities

To improve more, we decreased the learning rate from 10^{-3} to 10^{-5} and also increased the dropout rate from 0.2 to 0.4.

2) Simple CNN:

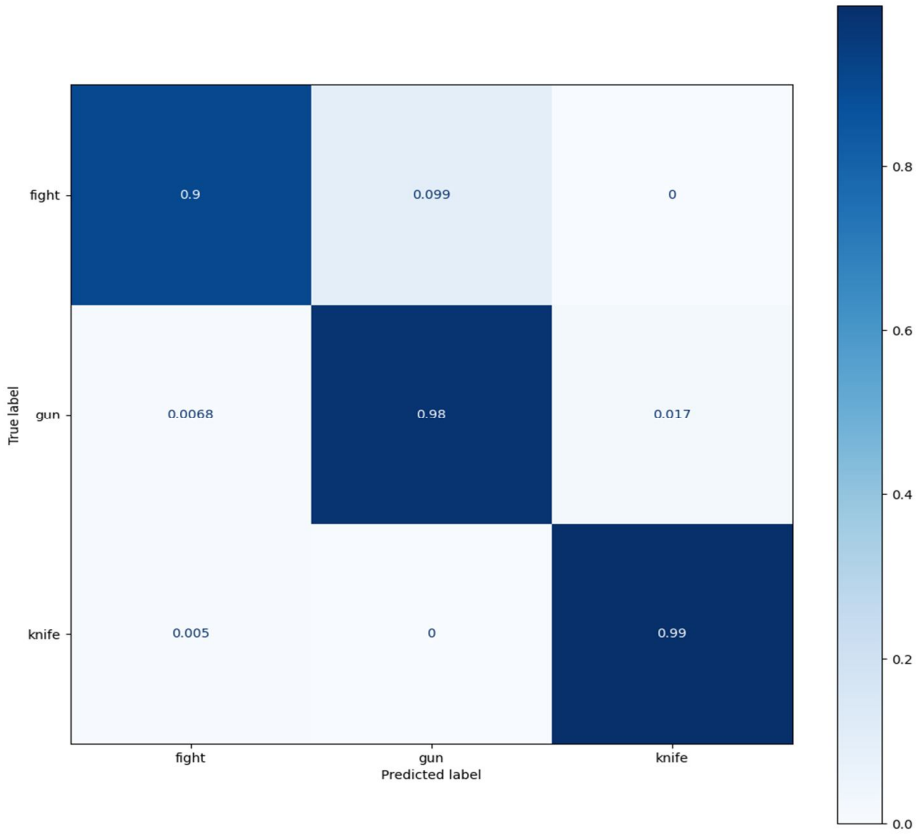


Figure 4: Confusion matrix for improved simple CNN model

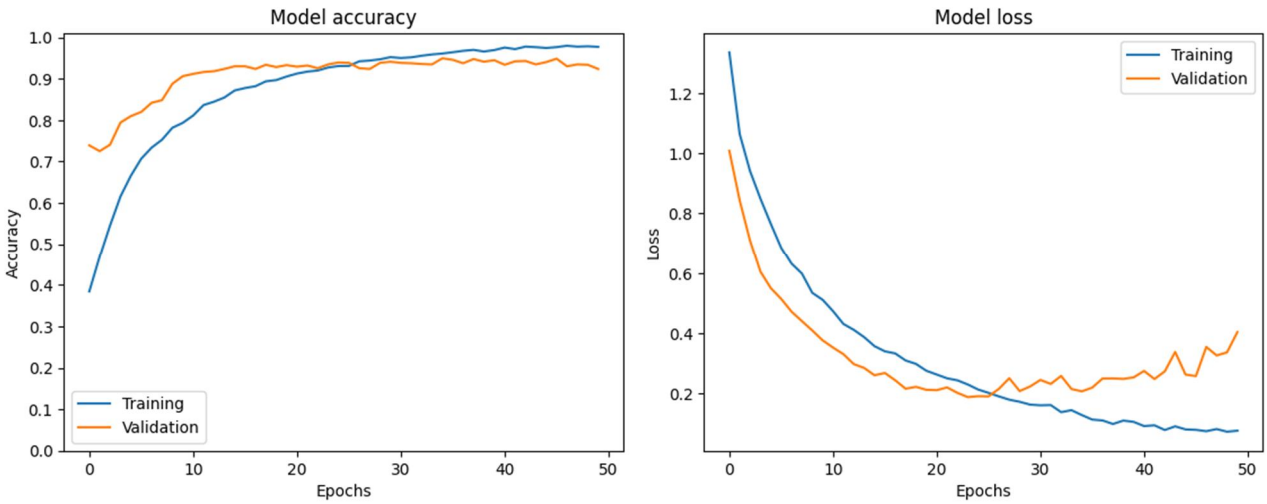


Figure 5: Model accuracy and loss graph for simple CNN model

	precision	recall	f1-score	support
0	0.99	0.90	0.94	433
1	0.87	0.98	0.92	296
2	0.99	0.99	0.99	400
accuracy			0.95	1129
macro avg	0.95	0.96	0.95	1129
weighted avg	0.96	0.95	0.95	1129

Figure 6: Classification report for simple CNN model

Validation Accuracy: 95.3942

We got good accuracy using this model and it is much improved from around 77% to 95%. But, still the two models that are ResNet50v2 and VGG19 we used to check if we get more accuracy.

3) ResNet50v2:

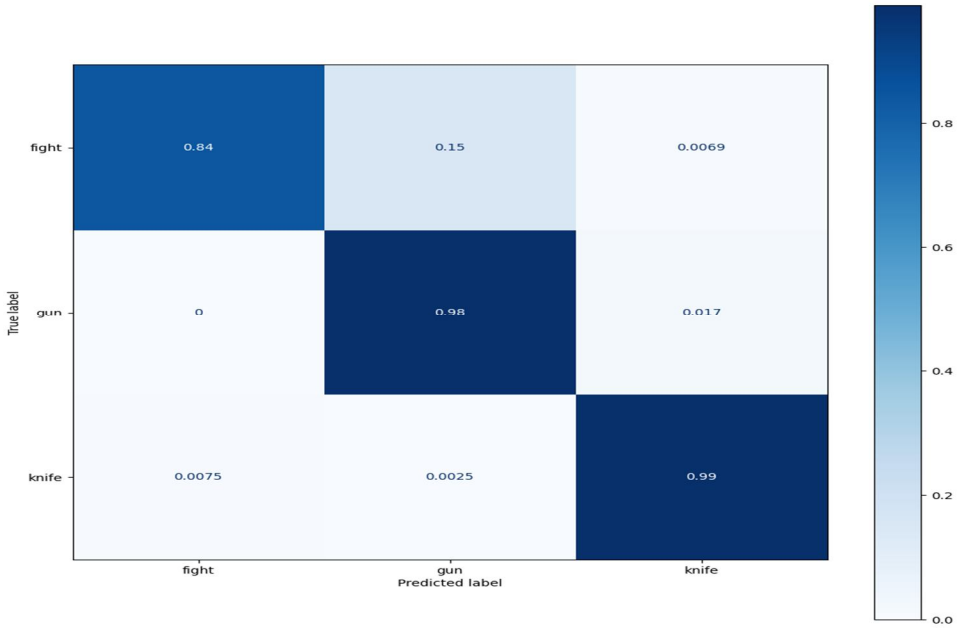


Figure 7: Confusion matrix for ResNet50v2 model

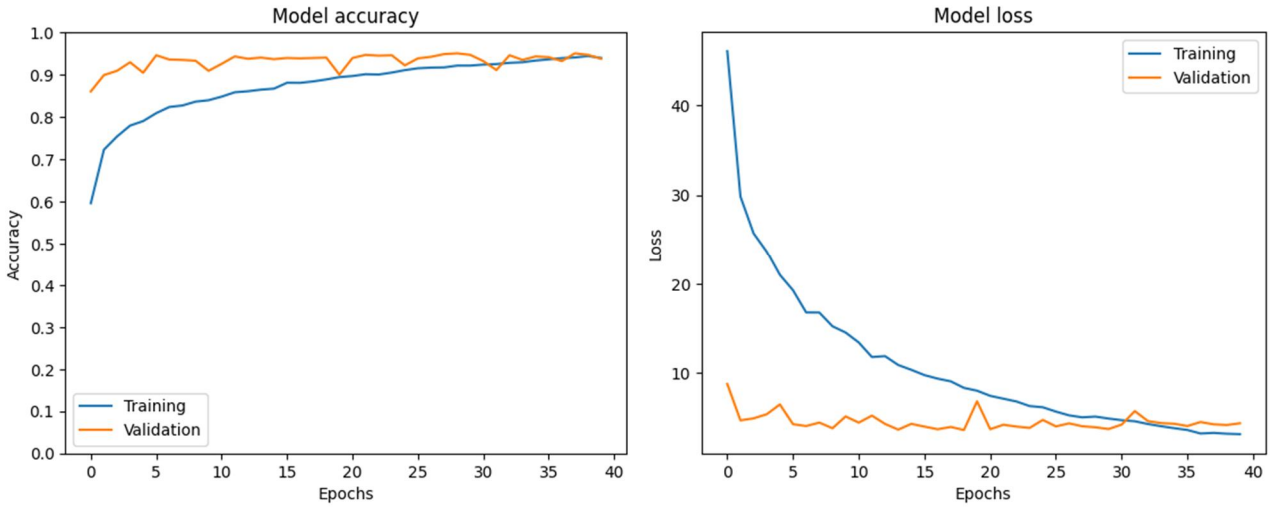


Figure 8: Model accuracy and loss graph for ResNet50v2 model

	precision	recall	f1-score	support
0	0.99	0.84	0.91	433
1	0.82	0.98	0.89	296
2	0.98	0.99	0.99	400
accuracy			0.93	1129
macro avg	0.93	0.94	0.93	1129
weighted avg	0.94	0.93	0.93	1129

Figure 9: Classification report for ResNet50v2 model

Validation Accuracy: 93.1798

Observe that simple CNN and ResNet50v2 gave the same accuracy for gun and knife detection. But, ResNet50v2 has lesser accuracy for fight detection than simple CNN. It might be possible that pre-trained model features of fight for simple CNN are more similar to the features that we extracted as compared to ResNet50v2.

4) VGG19:

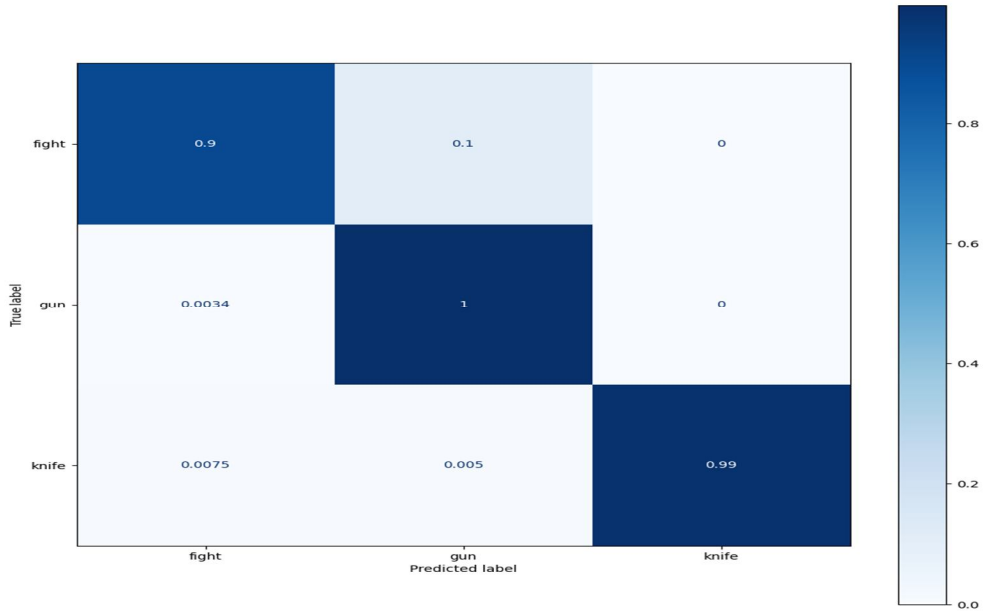


Figure 10: Confusion matrix for VGG19 model

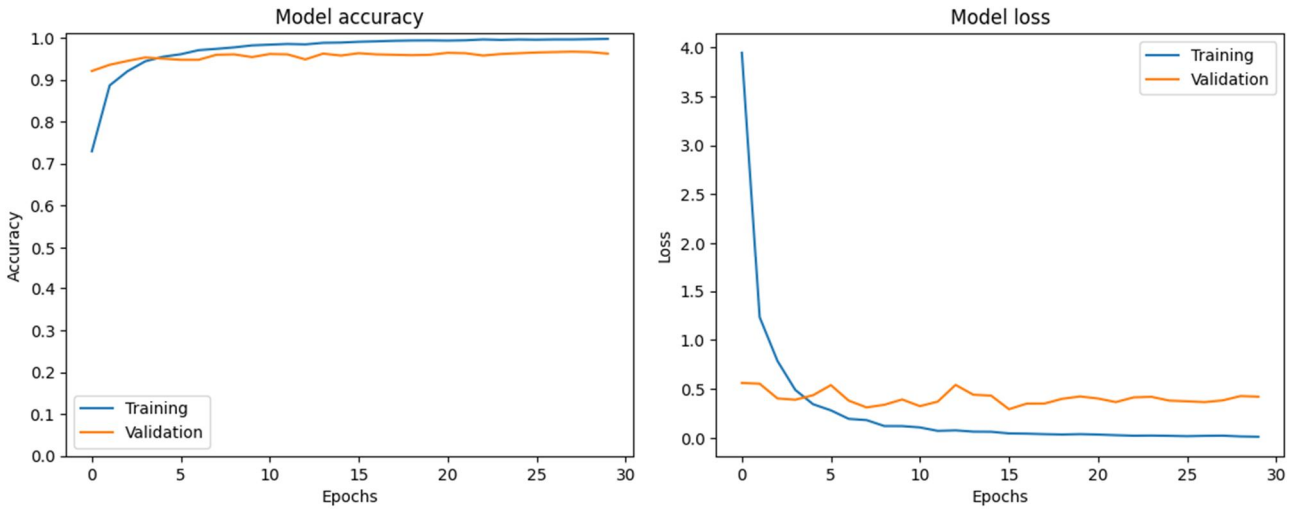


Figure 11: Model accuracy and loss graph for VGG19 model

	precision	recall	f1-score	support
0	0.99	0.90	0.94	433
1	0.86	1.00	0.92	296
2	1.00	0.99	0.99	400
accuracy			0.95	1129
macro avg	0.95	0.96	0.95	1129
weighted avg	0.96	0.95	0.96	1129

Figure 12: Classification report for VGG19 model

Validation Accuracy: 95.4827

This is also similar to simple CNN but it has a bit higher accuracy for gun detection as compared to simple CNN. We got the highest accuracy using this model.

VI. RESULTS

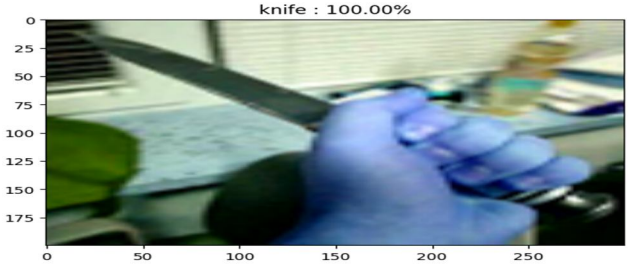
1) Gun Detection:



Figure 13: VGG19 detecting guns

The VGG19 model is used because of its highest accuracy. It detects guns accurately as we can see in the above images which are the output given by the model with accuracy percentage label.

2) Knife Detection:



ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

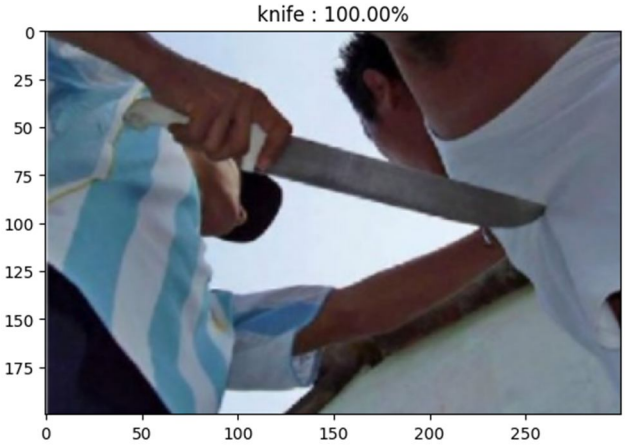


Figure 14: VGG19 detecting knives

The simple CNN will also provide the same results since both the models (simple CNN and VGG19) has same accuracy for knife detection.

3) Fight Detection:

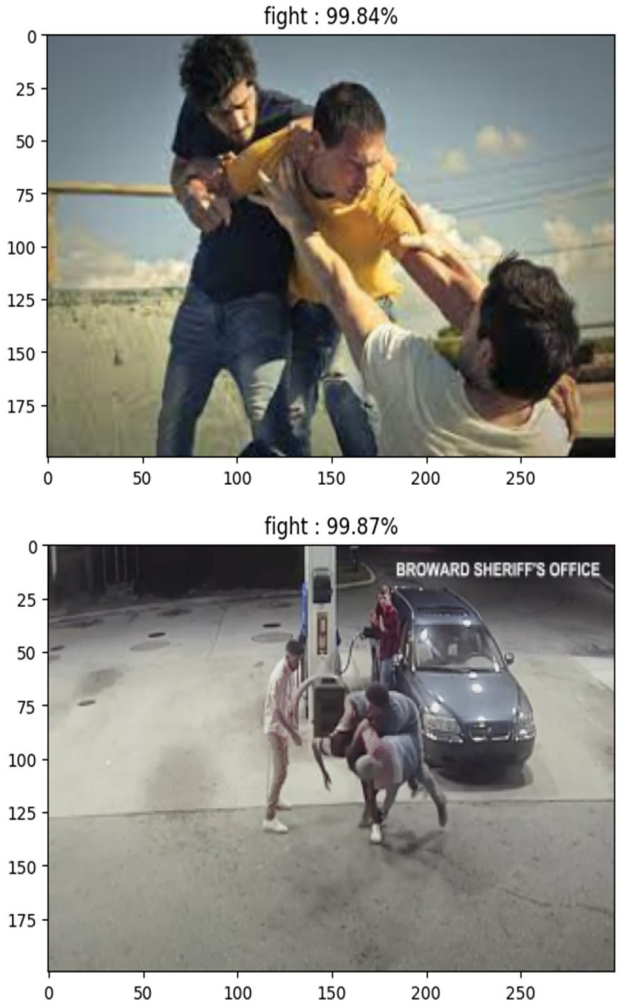


Figure 15: VGG19 detecting fights

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VII. RESEARCH GAPS

1) *Limited Availability Of Annotated Datasets:* The availability of annotated datasets for training and evaluating suspicious activity detection algorithms is limited, which can hinder the development of more accurate and robust models.
2) *The Trade-Off Between Accuracy And Computational Efficiency:* Most state-of-the-art suspicious activity detection algorithms require significant computational resources, which can limit their real-world applicability. Developing more efficient algorithms that can maintain high accuracy is an active research area.
3) *Difficulty In Detecting New Types Of Suspicious Activities:* Existing suspicious activity detection algorithms are often limited to detecting predefined types of suspicious activities. Detecting novel and previously unseen types of suspicious activities is still a challenge.

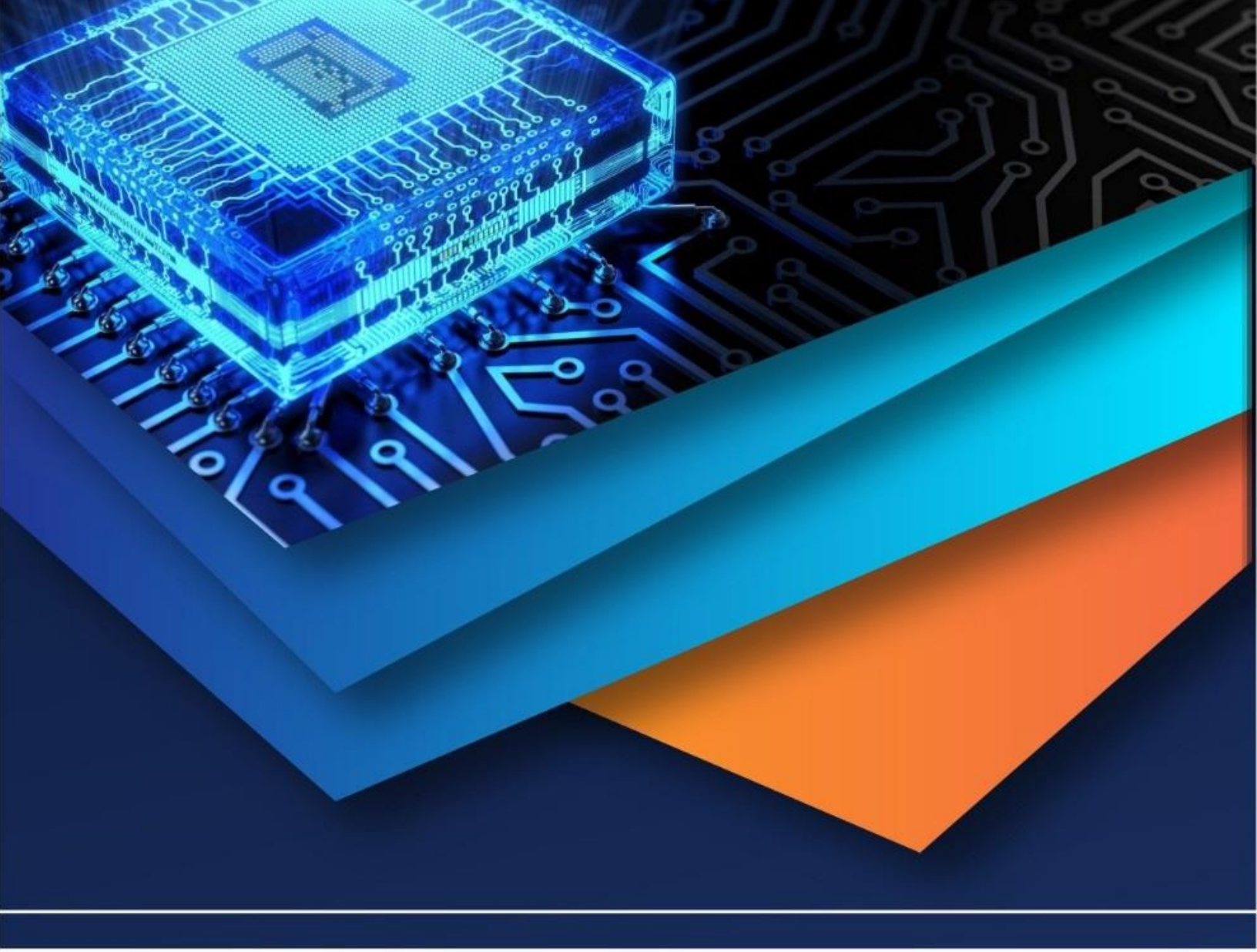
VIII. CONCLUSION

In this work, we have performed gun detection, knife detection and fight detection. We used three models that are simple CNN, ResNet50v2 and VGG19. From these three, we got the highest accuracy by VGG19 model, which is **95.4827**. But, rather than having this much validation accuracy, sometimes this model wrongly detects the activities from unseen image data. This issue is known as an overfitting issue which generally occurs in machine learning models. We can solve this issue by increasing the training data if we get more computing power. Also, we are not able to perform live detection of these activities because we didn't get the annotated dataset and this works only for a classification purpose.

REFERENCES

[1] U. M. Kamthe and C. G. Patil, "Suspicious Activity Recognition in Video Surveillance System," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-6, doi: 10.1109/ICCUBEA.2018.8697408.
[2] S. Loganathan, G. Kariyawasam and P. Sumathipala, "Suspicious Activity Detection in Surveillance Footage," 2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA), 2019, pp. 1-4, doi: 10.1109/ICECTA48151.2019.8959600.
[3] N. Bordoloi, A. K. Talukdar and K. K. Sarma, "Suspicious Activity Detection from Videos using YOLOv3," 2020 IEEE 17th India Council International Conference (INDICON), 2020, pp. 1-5, doi: 10.1109/INDICON49873.2020.9342230.
[4] Aditya G, Prasham S, Sumon G, Vaibhav S, Dr. Vaqar A, "Suspicious Activity Detection", Volume: 10, Issue: XI, Month of publication: November 2022, pp: 113-116 (ISSN no. 2321-9653, IC Value: 45.98, Impact Factor: 7.538), UGC Approved, doi: https://doi.org/10.22214/ijraset.2022.47186

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)