International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Violence Detection System using AWS Cloud

Shashank Srivastav[1], Vinod Jain[2], Akshay Varkale[3], Sushil Chhabra[4], Gurvinder Singh[5]

Submitted: 26/01/2023 **Revised:** 14/03/2023 **Accepted:** 06/04/2023

Abstract. With the rapid growth of surveillance cameras to monitor human activities, it has become really important to develop such systems that can detect violent activities and harmful people that are already banned from entering the premises by any organization before they cause any trouble. In this work, an effective architecture is proposed which is based on deep learning and real-time processing to alert the authorities about any violent activity in real-time without giving any false alarm. The architecture used in this work uses several parameters to conclude. Firstly, captured faces are compared with the faces of those individuals who were previously involved in violent activitiesin or around the institution. And for new cases, the system performs Facial gesture detection, violent object detection, violent action detection, weapon detection, and blood detection to draw any conclusion to avoid any false alarm. Then, faces involved are compared with the faces of all the members of the institution to discover if the person involved in violence belongs to the institution itself. An automatic email is sent to the discipline authority describing the person and the location of violence. For better understanding, all the steps of the research approach are presented with the help of architectural diagrams.

Keywords: *Violence detection system, weapons, and blood detection, facecomparison, AWS, cloud*

1 Introduction

Video Surveillance systems are widely used to monitor human activities to detect, prevent, and reduce crime. Due to various terrorist activities around the globe, proper detection of suspicious and violent human activities has become the need of the hour. However, as the number of surveillance cameras increases the need for operators to monitor the videos also increases. And soon there becomes a continuously increasing gap between the data captured from the video and human tendency to intelligently analyze the visual information [23],[24]. As a result, some of the suspicious activities left unnoticed giving rise to the need for a system having an automatic understanding of human actions. However, recognition of human activities is a difficult task because of various challenges like different body shapes and dressing styles of the individuals, the difference in the execution rate of activity by different individuals, camera viewpoint, etc. In the violence detection system, the first step is to divide the video into some segments [1]. In the second step, the objects in the various frames are detected. Then, features are extracted from the frames according to the applied method. Lastly, violent activity is detected. In early attempts, most of the studies on violence detection systems used some violence-based characteristics such as gunshots, blood strains, and explosives [2-3]. However, these methods have a disadvantage that they are dependent on the ability of surveillance systems to detect audio and their low detection rate can produce false alarms. Clarin et al. [4] tried to detect blood pixels in each image to conclude any violence incidence. Datta et al. [5] detected violence by using the trajectory of motion information and limb orientation of any individual during the fight scene. To detect violence Mahadevan et al [6] researched how to recognize violence by detecting blood and flames. Zixuan Wang [7] proposed facial recognition based on the location adaptive robustness. Wang used the system for mobile devices for real-time authentication. Wang tried to improve the overall

1Buddha Institute of Technology, Gorakhpur
shashank253@bit.ac.in
2 GLAUniversity Mathura, vinod.jain@gla.ac.in
3IES College of Technology , Bhopal
varkale.akshay@gmail.om
4Echelon Institute of Technology, Faridabad
sushilchhabra@eitfaridabad.co.in
5Asia Pacific Institute of Information Technology SD India Panipat
gurvinder@apiit.edu.in

accuracy of the system. More recently Chen [8] used action and face detection to recognize violence [25]. L. Nayak et al. [9] again tried to detect blood and flames in a video and degree of motion to detect violence. In recent times, the research area of violence detection has attracted the researchers when they noticed that there is a fine difference between violence and abnormal behaviour. The activities that are different from normal routine activities are called abnormal activities and activities which contain fighting, stealing is termed as violent activities.[10][11][12]. All the previous works are based on their definition of violence. Thus, there is a need to resolve these ambiguities to make the violence detection system more reliable in the real world [26],[27].

2 Proposed Work

There were few problems with earlier applications. They were focused on any one or two parameters to detect violence, as a result, some of the violence might get undetected and others might produce a false alarm. Secondly, applications were not designed to fetch the details of the people involved in violence and send them to authorities [28],[29]. To resolve these problems in this work a strong architecture is designed which can produce more accurate results based on different parameters involved in it. And with the help of a face comparison node, the application identifies the person involved in violence with the help of a college database. Since the entire infrastructure is implemented on the cloud, the latency involved in the process is also decreased. This research implemented this solution in the College to prevent the violence in College Campus.

2.1 Steps for Creating Model

Steps in creating architecture have been represented by some steps for betterunderstanding.

Step 1: First created a setup of the camera which is sending the live data to the local system server storage.

Step 2: Created two S3 Buckets A and B to store images in AWS cloud

Step 3: Wrote the bash script in the local system server which will break the live video frames in multiple images and send that images to the destination S3 Bucket A. **Step 4:** Created an IAM role, permitting AWS Lambda to access AWS S3, AWS Rekognition, AWS SNS AWS CloudWatch services, and MariaDB server.

Step 5: Launched the EC2 server on AWS.

Step 6: Installed MariaDB on EC2 server

Step 7: Created 2 databases on it. One contains the rusticated or banned student data (Photo, roll number, class, section, parents phone number) called R(d) database and the other is the whole college student database which contains the details of all students in college called S(d) database.

Step 8: Created Lambda(X) function: Configured S3 Bucket A for triggering Lambda(X), Choose the created IAM role and permission.

Step 9: Created Lambda(Y) function: Lambda(Y) will be triggered by lambda(X). So, Lambda(X) generates an event that will trigger Lambda(Y). Choose the created IAM role and permission.

Step 10: Added code in lambda for SNS functionality to send email to authority Email-id example (**authority@gmail.com**) [30],[31],[32]

2.2 Algorithm

This algorithm will alert the discipline authority of the organization if any rusticatedor banned person enters the organization or any violence happens or the violence possibility is detected [33].

Step 1: Live image is sent to the S3 bucket A by written script in a local systemserver. This process will trigger the Lambda(X) and call the rekognition functionality, Lambda(X) will take the decision according to following decision nodes criteria parameters. These nodes generate output 1 or 0.

1 Face comparison nodes:

```
{ S3 bucket A data compared with Rd database
(Database containingrusticated/banned students):
if (match found):return 1
else:
return 0
}
```

2 Violence detection node (This node undertakes several parameters):

```
{
Logics: {if (weapon_detection>60%) or
(blood_detection>60%)return 1
else:
return 0}
Logics: {if(violent_facial_gesture_detection>90%):
            return 1
else:
return 0 }
Logics: {if (Violent_object_detection>75%)):
            return 1
                  else
                return 0}
```

Logics: {if (Violent_action_detection>75%)}return
1
else
return 0}
}

Step 2: Next lambda(X) applies the following logic between nodes.

Face comparison nodes OR Violence detection node Here FC- Face comparison, VD - Violence detection, inputs= FC and VD are inputs to the OR logic gate and c= Output of OR**.**

Scenario 1: If neither any rusticated or banned student enter the college nor any violence happens, i.e.

if FC=0, VD=0, then
c= {FC=0 OR VD=0} = 0

Scenario 2: If any rusticated or banned student enters the collegethen Face comparison node will be triggered, i.e.

if FC=1, VD=0, then
c= {FC=1 OR VD=0} = 1

Scenario 3: If any new person who is not part of the rusticated list does any violence. Then the violence node will not be triggered i.e.

if FC=0, VD=1, then
c= {FC=0 OR VD=1} = 1

*Scenario 4***:** If any rusticated or banned student does any violence thenboth the face and violence node will be triggered i.e.

if FC=1, VD=1, then
c= {FC=1 OR VD=1} = 1

Step 3: Next output c in Lambda(X) will decide the further process.

If (c==1):
then { "Invoke SNS Functionality in Lambda and Send mail toauthority"
"Put that person image into S3 Bucket B"
"Invoke another lambda function Lambda(Y)"
"Move to Step 4"}
else { "Violence not detected" "Process will be stop here""End Process**"**}

Step 4: -- Lambda(Y) invoked due to output c = 1 of Lambda(X). Lambda(Y) has only Face comparison node which compares S3 Bucket B data and Student database (Sd). Lambda(Y) is used to determine whether the student involved in violence is a college student and is used to fetch his information from the college database (Sd).

Face comparison node of Lambda(Y){ if (match found):
then: {"Send the involved Student detail from S(d) database to authority by using SNS function in Lambda(Y)
"Put the matched Student data details in the R(d) database"."END".}
else: {"Put the unmatched new data detail to R(d) database."END"}
}

2.3 Designed Architecture

The designed architecture has been represented by some images for better understanding [34],[35],[36].

Images are processed by a script in the local system . They are sent to s3 bucket A by the same script. This event triggers the lambda function X.

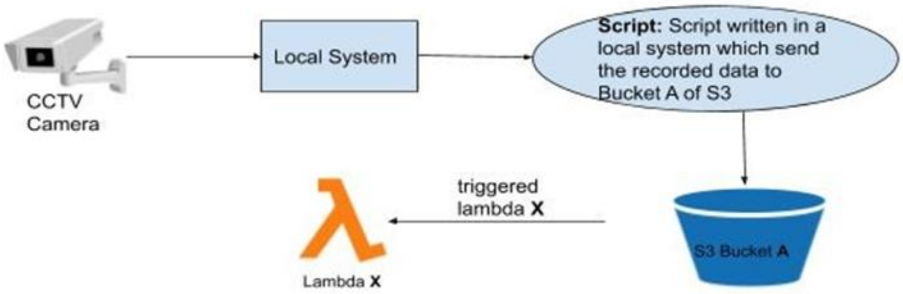

CCTV Camera
Local System
Script: Script written in a local system which send the recorded data to Bucket A of S3
triggered lambda X
S3 Bucket A
Lambda X

Fig. 1. Workflow in local system

Lambda X Compares image in the Bucket data A with rusticated student database R(d) and also detects violence in the image. If the output of these nodes is 1. Then putthe image in S3 bucket B. Calls SNS and Triggers Lambda Y [37],[38],[39].

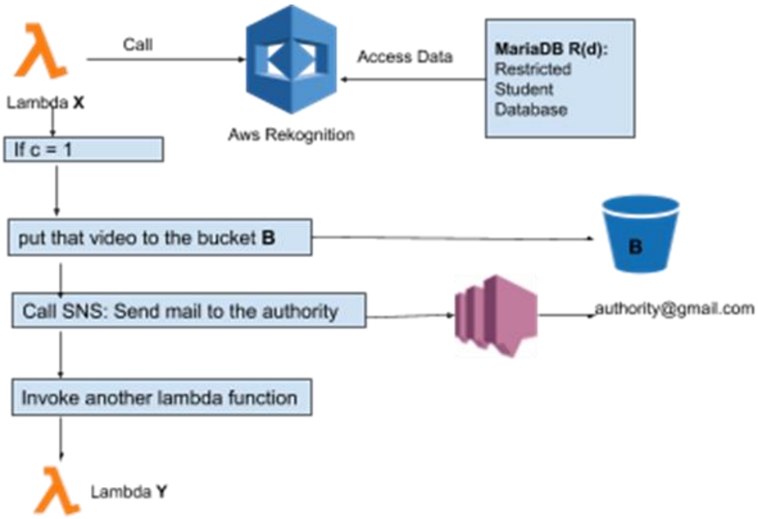

Call
Lambda X
Aws Rekognition
Access Data
MariaDB R(d): Restricted Student Database
If c = 1
put that video to the bucket B
B
Call SNS: Send mail to the authority
authority@gmail.com
Invoke another lambda function
Lambda Y

Fig. 2. Triggering of Lambda X

Lambda Y compares the Bucket B data with database Sd. Call SNS and Put the datainto Rd.

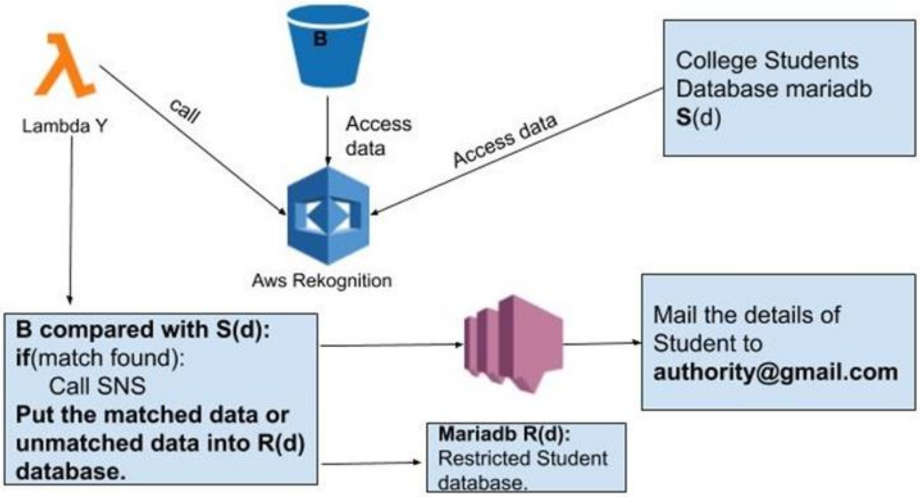

Lambda Y
call
B
Access data
Access data
College Students Database mariadb S(d)
Aws Rekognition
B compared with S(d):
if(match found):
Call SNS
Put the matched data or unmatched data into R(d) database.
Mail the details of Student to authority@gmail.com
Mariadb R(d): Restricted Student database.

Fig. 3. Working of Lambda Y

3 Experimental Results

When an image is pushed to bucket A. Lambda X is triggered. For testing purpose, a violence scene from Spiderman movie, 2002 is taken and uploaded in S3 bucket A

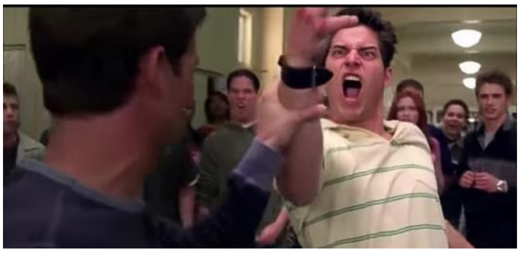
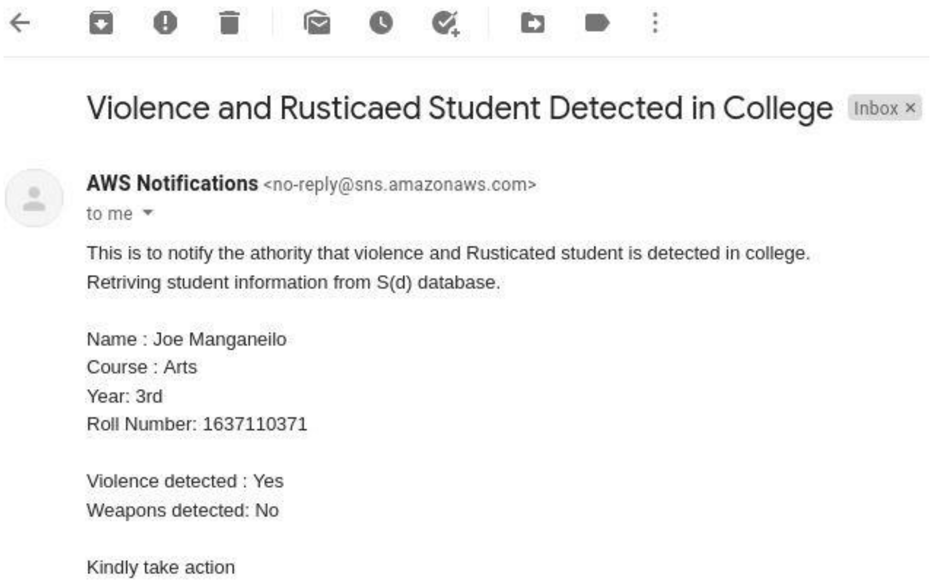


Fig. 4. Image uploaded in S3 bucket A (Source: Spiderman movie,2002)

As soon as the image is pushed to Bucket A, Lambda(X) gets triggered and executes the algorithm to detect violence and rusticated students. Bucket A image is compared with images in rusticated database R(d). The student involved in violence is detectedas a rusticated student.

Fig. 5. Rusticated student image in Rusticated database R(d) (Source: Spiderman movie,2002)

As soon as violence and rusticated students are detected in college premises an automated email is sent to the authorities regarding violence and student details.



Violence and Rusticaed Student Detected in College

AWS Notifications <no-reply@sns.amazonaws.com>

to me

This is to notify the athority that violence and Rusticated student is detected in college.
Retriving student information from S(d) database.

Name : Joe Manganeilo
Course : Arts
Year: 3rd
Roll Number: 1637110371

Violence detected : Yes
Weapons detected: No

Kindly take action

Fig. 6. Screenshot of alert email

Table. Comparison of the proposed method with previous studies on various parameters

Studies and researches	**Blood detection**	**Violent action detection**	**Face comparison**	**Weapon Detection**	**violent object detection**	**Violent gesture detection**
Clarin et al.[4]	✔	✖	✖	✖	✖	✖
Datta et al.[5]	✖	✔	✖	✖	✖	✖
Mahadevanet al[6]	✔	✖	✖	✖	✖	✖
Zixuan Wang **[7]**	✖	✖	✔	✖	✖	✖
Chen [8]	✖	✔	✖	✖	✖	✔
Proposed method	✔	✔	✔	✔	✔	✔

4 Conclusion and Future Work

With the rapid increase in the number of surveillance cameras to monitor human activities, it has become very important to develop applications that can detect violent activities automatically without the need for any human intervention. In computer vision, violence detection has become a hot topic to attract new researchers. In recent times, there has been much research work to detect human activities however automatically detecting violence is comparatively less studied. Few problems arise in automatic detection of violence due to subjective nature in defining what can be termed as violence. Also, different human activities might be misclassified as violence leading to a false alarm. All the previous works are based on their definition of violence. Thus there is a need to resolve these ambiguities to make the violence detection system more reliable in the real world. However, the demand for better violence detection has increased in areas ranging from public safety to military intelligence. This work includes various parameters to detect violence which are face recognition, facial gesture detection, violent action detection, violent object detection, blood, and weapon detection. Previous works focused on any one or two parameters only and thus were not as reliable and produced false alarms in many cases

In future, Alexa skills can be integrated in the architecture to alert authorities through voice calls instead of emails . This will prevent any important alert being missed.

References

[1] S. Chaudhary, M. A. Khan, and C. Bhatnagar, ‘‘Multiple anomalous activity detection in videos,’’ Procedia Comput. Sci., vol. 125, pp. 336–345, Jan. 2018.

[2] Chen LH, Hsu HW, Wang LY, Su CW. Violence detection in movies. In: Computer Graphics, Imaging, and Visualization (CGIV), 2011 Eighth International Conference on. IEEE; 2011. p. 119–124.

[3] Chen LH, Hsu HW, Wang LY, Su CW. Violence detection in movies. In: Computer Graphics, Imaging, and Visualization (CGIV), 2011 Eighth International Conference on. IEEE; 2011. p. 119–124

[4] Clarin C, Dionisio J, Echavez M, Naval P, DOVE: Detection of movie violence using motion intensity Analysis on skin and blood, Technical report, University of the Philippines, 2005

[5] Datta, A.; Shah, M.; Lobo, N.D.V. Person-on-person violence detection in video data. In Proceedings of the 16th International Conference on Pattern Recognition, Quebec, QC, Canada, 11–15 August 2002; pp. 433–438.

[6] Mahadevan, V.; Li, W.; Bhalodia, V.; Vasconcelos, N. Anomaly detection in crowded scenes. In Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA, 13–18 June 2010; pp. 1975–1981.

[7] Zixuan Wang, "Who is Here: Location-Aware Face Recognition", PhoneSense’12, November 6, 2012, Toronto, ON, Canada. ACM 978-1-4503-1778-8.

[8] Juan, I.E.; Juan, M.; Barco, R. A low-complexity vision-based system for real-time traffic monitoring. IEEE Trans. Intell. Transp. Syst. 2017, 18, 1279–1288.

[9] Nayak, L. Audio-Visual Content-Based Violent Scene Characterisation. Ph.D. Thesis, National Institute of Technology, Rourkela Odisha, India, 2015.

[10] A. B. Mabrouk and E. Zagrouba, ‘‘Abnormal behavior recognition for intelligent video surveillance systems: A review,’’ Expert Syst. Appl., vol. 91, pp. 480–491, Jan. 2018

[11] T. Zhang, Z. Yang, W. Jia, B. Yang, J. Yang, and X. He, ‘‘A new method for violence detection in surveillance scenes,’’ Multimedia Tools Appl., vol. 75, no. 12, pp. 7327–7349,2016.

[12] M. Alvar, A. Torsello, A. Sanchez-Miralles, and J. M. Armengol, ‘‘Abnormal behavior detection using dominant sets,’’ Mach. Vis. Appl., vol. 25, no. 5, pp. 1351–1368, Jul. 2014.

[13] A. A. Mishra and G. Srinivasa, ‘‘Automated detection of fighting styles using localized action features,’’ in Proc. 2nd Int. Conf. Inventive Syst. Control (ICISC), Jan. 2018, pp. 1385–1389.

[14] T. Agrawal, A. Kumar, and S. K. Saraswat, ‘‘Comparative analysis of convolutional codes based on ML decoding,’’ in Proc. 2nd Int. Conf. Commun.Control Intel. Syst. (CCIS), Nov. 2016, pp. 41–45.

[15] C. Ding, S. Fan, M. Zhu, W. Feng, and B. Jia, ‘‘Violence detection in a video by using 3D convolutional neural networks,’’ in Proc. Int. Symp. Visual Comput., 2014, pp. 551–558.

[16] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, ‘‘NetVLAD: CNN Architecture for weakly supervised place recognition,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.(CVPR), Jun. 2016, pp. 5297–5307.

[17] G. Mu, H. Cao, and Q. Jin, ‘‘Violent scene detection using convolutional neural networks and deep audio features,’’ in Proc. Chin. Conf. Pattern Recognit., 2016, pp. 451–463.

[18] S. Sudhakaran and O. Lanz, ‘‘Learning to detect violent videos using convolutional long short-term memory,’’ in Proc. 14th IEEE Int. Conf. Adv. Video Signal-Based Surveill. (AVSS), Aug./Sep. 2017, pp. 1–6.

[19] Z. Meng, J. Yuan, and Z. Li, ‘‘Trajectory-pooled deep convolutional networks for violencedetection in videos,’’ in Proc. Int. Conf. Comput. Vis. Syst., 2017, pp. 437–447.

[20] I. Serrano, O. Deniz, J. L. Espinosa-Aranda, and G. Bueno, ‘‘Fight recognition in video using Hough forests and 2D convolutional neural network,’’ IEEE Trans. Image Process., vol. 27, no. 10, pp. 4787–4797, Oct. 2018.

[21] F. U. M. Ullah, A. Ullah, K. Muhammad, I. U. Haq, and S. W. Baik, ‘‘Violence detection using spatiotemporal features with 3D convolutional neural network,’’ Sensors, vol. 19,no. 11, p. 2472, May 2019.

[22] W. Lejmi, A. B. Khalifa, and M. A. Mahjoub, ‘‘Fusion strategies for recognition of violence actions,’’ in Proc. IEEE/ACS 14th Int. Conf. Comput. Syst. Appl. (AICCSA), Oct./Nov. 2017, pp. 178–183. D. Maniry, E. Acar, F. Hopfgartner, and S. Albayrak, ‘‘A visualization tool for violent scenes detection,’’ in Proc. Int. Conf. Multimedia Retr., Apr. 2014, p. 522.

[23] Kumar, V. and Kumar, R., 2015. An adaptive approach for detection of blackhole attack in mobile ad hoc network. Procedia Computer Science, 48, pp.472-479.

[24] Kumar, V. and Kumar, R., 2015, April. Detection of phishing attack using visual cryptography in ad hoc network. In 2015 International Conference on Communications and Signal Processing (ICCSP) (pp. 1021-1025). IEEE.

[25] Kumar, V. and Kumar, R., 2015. An optimal authentication protocol using certificateless ID-based signature in MANET. In Security in Computing and Communications: Third International Symposium, SSCC 2015, Kochi, India, August 10-13, 2015. Proceedings 3 (pp. 110-121). Springer International Publishing.

[26] Kumar, Vimal, and Rakesh Kumar. "A cooperative black hole node detection and mitigation approach for MANETs." In Innovative Security Solutions for Information Technology and Communications: 8th International Conference, SECITC 2015, Bucharest, Romania, June 11-12, 2015. Revised Selected Papers 8, pp. 171-183. Springer International Publishing, 2015.

[27] Kumar, V., Shankar, M., Tripathi, A.M., Yadav, V., Rai, A.K., Khan, U. and Rahul, M., 2022. Prevention of Blackhole Attack in MANET using Certificateless Signature Scheme. Journal of Scientific & Industrial Research, 81(10), pp.1061-1072.

[28] Pentyala, S., Liu, M., & Dreyer, M. (2019). Multi-task networks with universe, group, and task feature learning. arXiv preprint arXiv:1907.01791.

[29] Srivastava, Swapnita, and P. K. Singh. "Proof of Optimality based on Greedy Algorithm for Offline Cache Replacement Algorithm." International Journal of Next-Generation Computing 13.3 (2022).

[30] Smiti, Puja, Swapnita Srivastava, and Nitin Rakesh. "Video and audio streaming issues in multimedia application." 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2018.

[31] Srivastava, Swapnita, and P. K. Singh. "HCIP: Hybrid Short Long History Table-based Cache Instruction Prefetcher." International Journal of Next-Generation Computing 13.3 (2022).

[32] Srivastava, Swapnita, and Shilpi Sharma. "Analysis of cyber related issues by implementing data mining Algorithm." 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2019.

[33] P Mall and P. Singh, “Credence-Net: a semi-supervised deep learning approach for medical images,” Int. J. Nanotechnol., vol. 20, 2022.

[34] Narayan, Vipul, et al. "Deep Learning Approaches for Human Gait Recognition: A Review." 2023 International Conference on Artificial Intelligence and Smart Communication (AISC). IEEE, 2023.

[35] Narayan, Vipul, et al. "FuzzyNet: Medical Image Classification based on GLCM Texture Feature." 2023 International Conference on Artificial Intelligence and Smart Communication (AISC). IEEE, 2023.

[36] "Keyboard invariant biometric authentication." 2018 4th International Conference on Computational Intelligence & Communication Technology (CICT). IEEE, 2018.

[37] Mall, Pawan Kumar, et al. "Early Warning Signs Of Parkinson’s Disease Prediction Using Machine Learning Technique." Journal of Pharmaceutical Negative Results (2023): 2607-2615.

[38] Pentyala, S., Liu, M., & Dreyer, M. (2019). Multi-task networks with universe, group, and task feature learning. arXiv preprint arXiv:1907.01791.

[39] Choudhary, Shubham, et al. "Fuzzy approach-based stable energy-efficient AODV routing protocol in mobile ad hoc networks." Software Defined Networking for Ad Hoc Networks. Cham: Springer International Publishing, 2022. 125-139.