

Hand-Sign To Text Using CNN

Anupam Gupta, Arshad Khan, Ayush Bansal, Kanish Verma, Jagbeer Singh

Department of Computer Science and engineering.

Meerut Institute of Engineering and Technology, Meerut, U.P., India

{ anupamgupta2607@gmail.com, arshadkhancse71@gmail.com, Aayushbansal35@gmail.com, kanish.chirag@gmail.com, jagbeer.singh@miet.ac.in }

DOI: 10.47750/pnr.2023.14.S01.158

Abstract

In day-to-day life, people with hearing impairment feel discomfort in public areas. They can't explain their words to others. So, Sign language recognition comes. Sign Language recognition has many applications and increasing continuously. This paper shows the CNN model of a hand sign-detection system converting it into a message. From the existing designs which are based on image processing, showing a simple single letter, it can make a complete message by joining the all-single letter. The predicted message is displayed on a side screen.

Introduction

Signs are the expressive way to tell our thoughts or views, and Sign Language is developed for people with hearing and speaking impairments [1][2].

There are three types of Sign language:-

- 1.ASL(American Sign Language)
- 2.BSL(British Sign Language)
3. Australian Sign Language

In this paper, the CNN model to understand ASL(American Sign language) and transform signs or letters into words.

They can classify into Following Types: -

- Face sign: There are many examples of face signs, some of them are: shaking of the head, movement of the eyebrows, the movement of the eyeball, opening and closing the mouth to speak, and different gestures of a face when we look of happiness fear, surprise, anger, contempt, sadness, etc.
- Hand gestures: Showing the different signs from fingers to tell the letter or thing to another person.
- Body Movement: It involve movement of different body parts such as legs, stomach and other body part.

The essential aim of developing a hand sign detection system is to establish a connection between computer and human, where the machine can easily understand the sign that they are trying to say and convert it into text [3][4].

Keywords: model, ASL (American sign Language), Python, CNN, OpenCV

1. Related Theory

1.1 CNN

In machine learning different neural network for different purposes. E.g., RNN (Recurrent Neural Network) is used for predicting the sequences of words. Same as, for image classification or image operation we use CNN (Convolutional Neural Network).

In Neural Network there are 3 types of layers: -

Input layer

Hidden Layer (Convolutional layer)

Output Layer [5][6].

1.1.1 Input Layer: - In this layer we provide input to our model.

1.1.2 Hidden Layer: - From input layer the input is passed to the Hidden Layer. Hidden layer can vary in number depend upon the layer added by us. Activation function is used in this layer.

1.1.3 Output Layer: - The data is passed from hidden layer to output layer. Different logistic functions are used in this layer such as Sigmoid or SoftMax. It converts the output of each layer into the probability score of each class.

The construction of CNN is a multi-layered feed forward neural network. In CNN hidden layers are grouped with activation layers. The pre-processing of image is needed before feeding it into the input layer.

Different CNN Model are: -

ResNet

AlexNet

VGG

MobileNet

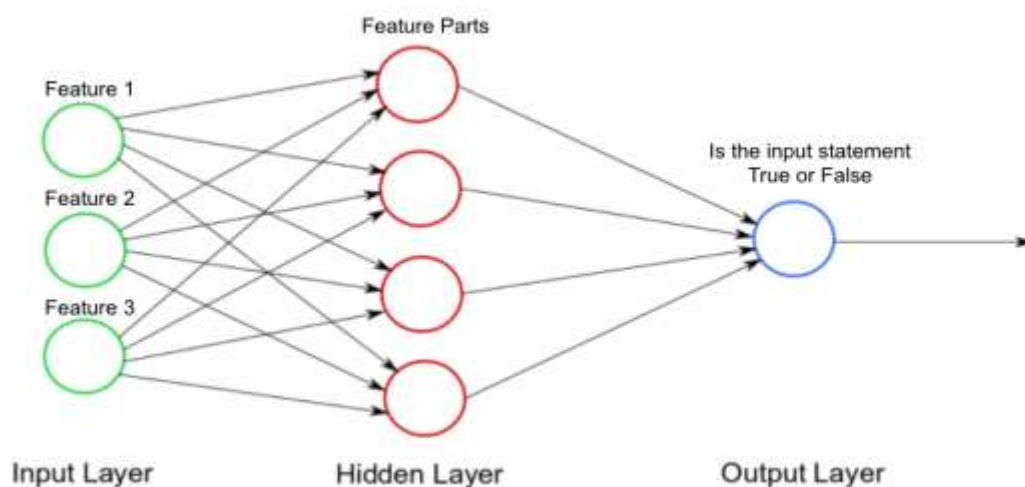
1.2 Application of CNN

1. Analysing image of different body part.
2. Advertising
3. Understanding Climate

1.3 Hand Sign Recognition Using CNN: In this paper we presented a scheme of detecting the different sign of hand to get a text from the person. Initially, the hand sign is captured on the basis of ASL (American Sign Language). A CNN model which is trained on our dataset which contain alphabet A to Y (not including J) of AS [7][8].

1.4 Hand Sign to text System for the people who have difficulty in speaking: In this paper a model is created for the DUMB people to understand their different sign similar to ASL. For the hand sign detection CNN technology is used. It is a dynamic system for recognizing the American Sign using the web camera. We use GRAY_SCALING and python thresholding technique for making simple image. It helps dumb people to get interaction with other [9][10].

Figure 1



2.

Methodology

The overview of the hand sign detection system is made by following the under given stages.

- 1) The first stage is the hand sign image are captured where the images are taken using web camera with different position of same letter.
- 2) the next stage is a pre-processing stage in which image is pre-processed by Grey Scaling using `cv2.COLOR_BGR2GRAY` method.
- 3) After pre-processing of images we set the path in model file i.e., `model.py`.
- 4) It will create model. Now choose the file for the `Gesture_recognize.py` [11][12].

2.1 Hand sign image capture

The construction of a dataset for hand signs is done by web camera. American Sign language is used here. Hand image is captured in the green rectangular box.

Image is captured by pressing the key “c”.

The web cam keep run until the key “q” is pressed.

OpenCV library of python is used to capture the hand images.



Figure 2

2.2 Pre-processing stage

Image processing is an art of converting an image into digital format such that it can be manipulate for data extraction [13][14].

2.2.1 GRAY SCALLING

Each image which is captured by camera is preprocessed first. In preprocessing each image is first converted into Gray scale from BGR (Blue, Green, Red). This is done for simplicity, as operation can easily performed on simple 2D array rather than BGR 3D array image. cv2.COLOR_BGR2GRAY parameter is passed in the cv2.cvtColor function for this purpose.



Figure 3

2.2.2 IMAGE THRESHOLDING

In OpenCV, Thresholding technique is used for image preprocessing. In this technique each image pixel value is matched with the provided threshold value.[2] The pixel value is set to zero(0) if it is smaller than the threshold value. The image pixel value is set to max(i.e., 255) if it is higher than the threshold value. Thresholding technique helps in image segmentation, in this it changes the image pixel to differentiate object on foreground from background. In CV thresholding technique is done on grayscale images.[3] There are many types of thresholding technique which are: -[15][16].

- 1)cv.THRESH_TRUNC
- 2)cv.THRESH_TOZERO_INV
- 3)cv.THRESH_TOZERO
- 4)cv2.THRESH_BINARY
- 5)cv2.THRESH_BINARY_INV

In this model, “cv2.THRESH_BINARY_INV” is used. In this technique, if the image pixel intensity is greater than the provided threshold it is set to zero(black). If it less than the threshold value then the pixel value is set to max(i.e., 255). In this model (127,255) is used as threshold value.

2.3 Training/Testing Phase

2.3.1 Sequential model

For training or testing of the images, images are read as NumPy array using imread method. To predict captured image as letter a label of alphabet is provided. American Sign language is used. For training Sequential API model from keras module is used. Keras model is an aggregate of multiple training and inferential layers. The Sequential API are less refined API for modelling simple neural network. [10] “relu (rectified linear activation unit)” activation function for training is used. Many layers are added into the model by using add() method. Convolutional layer is also added i.e. Conv2D from keras. Layer module. It creates a convolutional kernel which bind with the input layers. It helps to produce tensor output. Different layers are added by changing kernel size(it is

integer/list/tuple of 2 integers defining the height and width of convolutional window). Here 3 layers having 3 different kernel size are applied[17][18].

Now this model is loaded to our code for the testing of the data image.

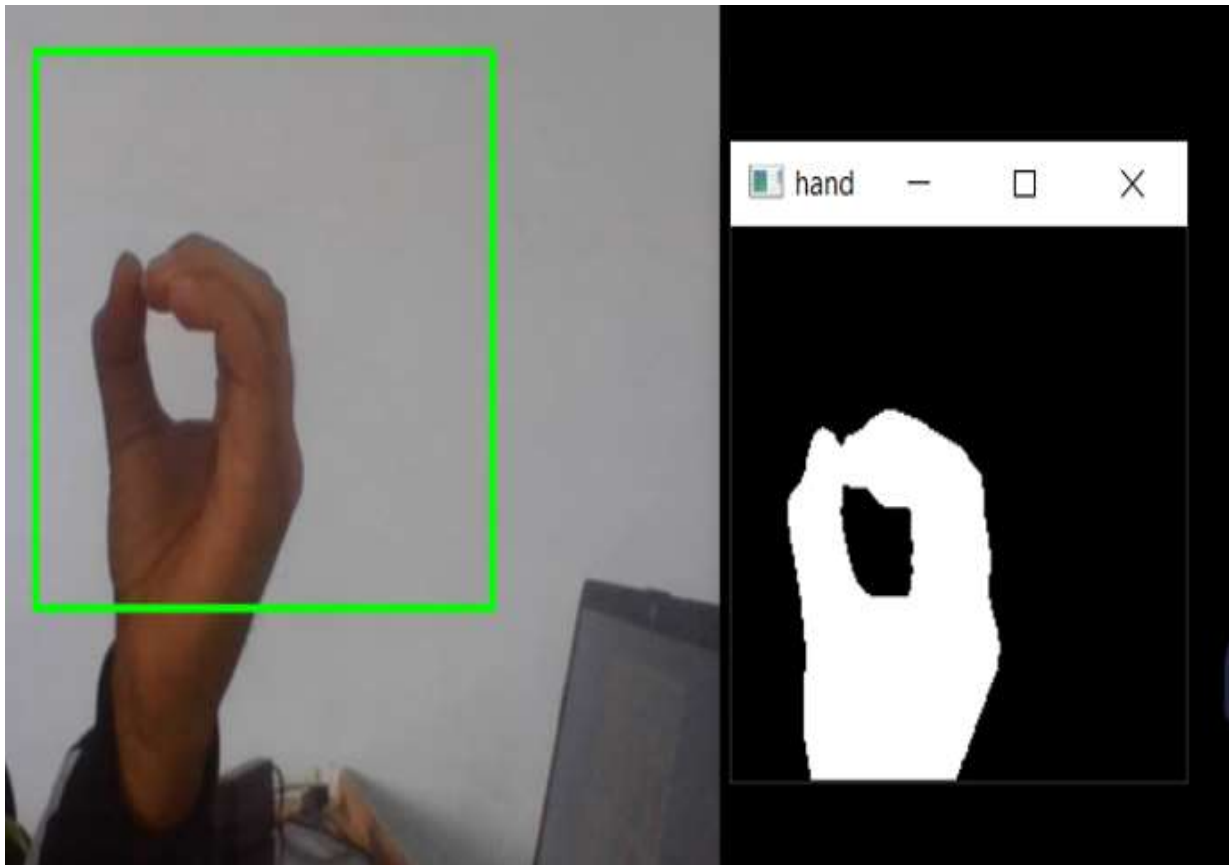


Figure 4

Diagram that shows the flow of Hand-Sign Detection

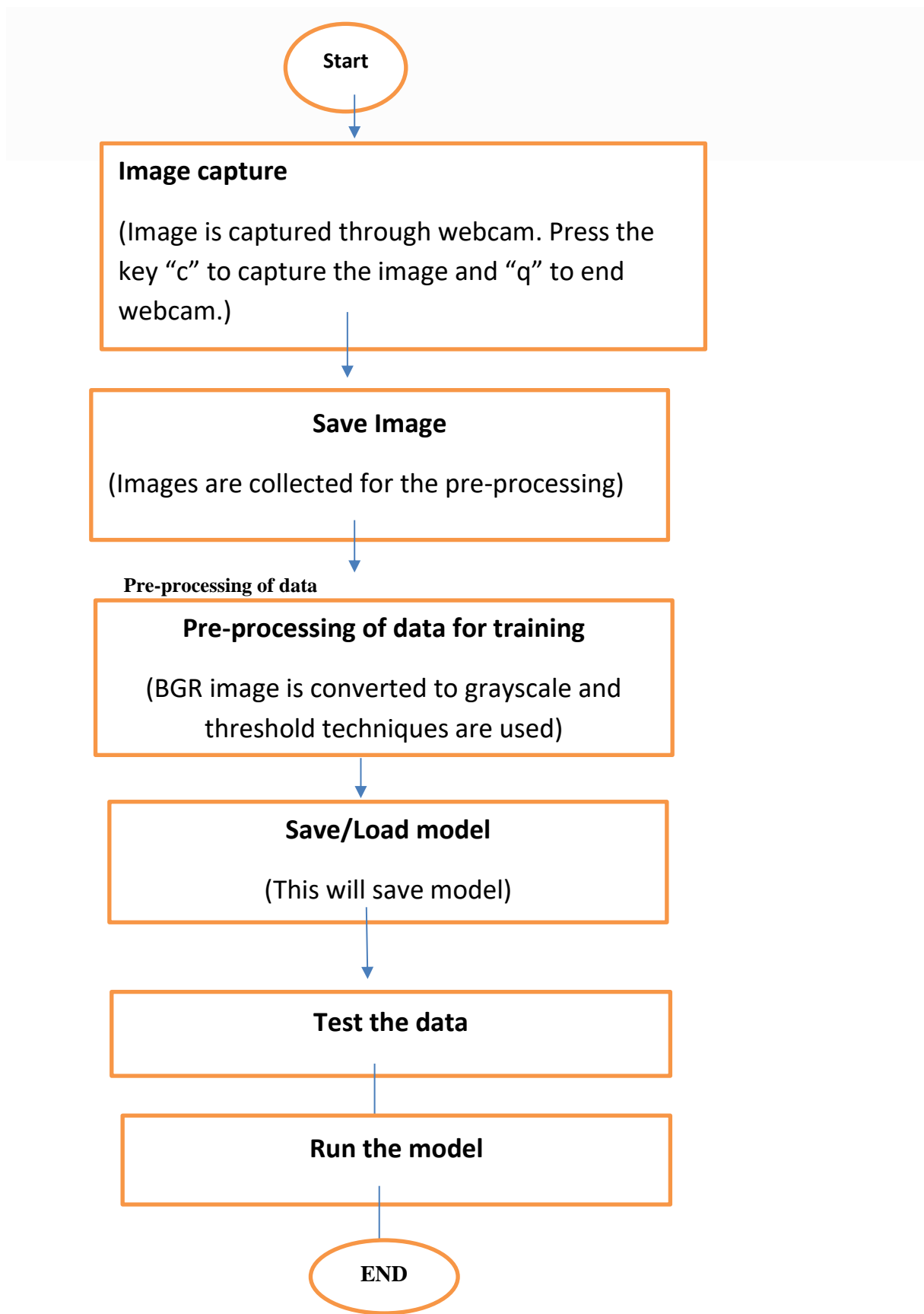


Figure 5

APPLICATION

Our paper is beneficial at different places such as Restaurant/Airports/Hospital/ and many other places. We can install a camera at the front of reception of these places, so that a person can show their signs in the camera and through our model it can predict the text and show to the screen of the other side. So, it creates an interaction between the both [19][20][21].

OUTPUT SCREEN

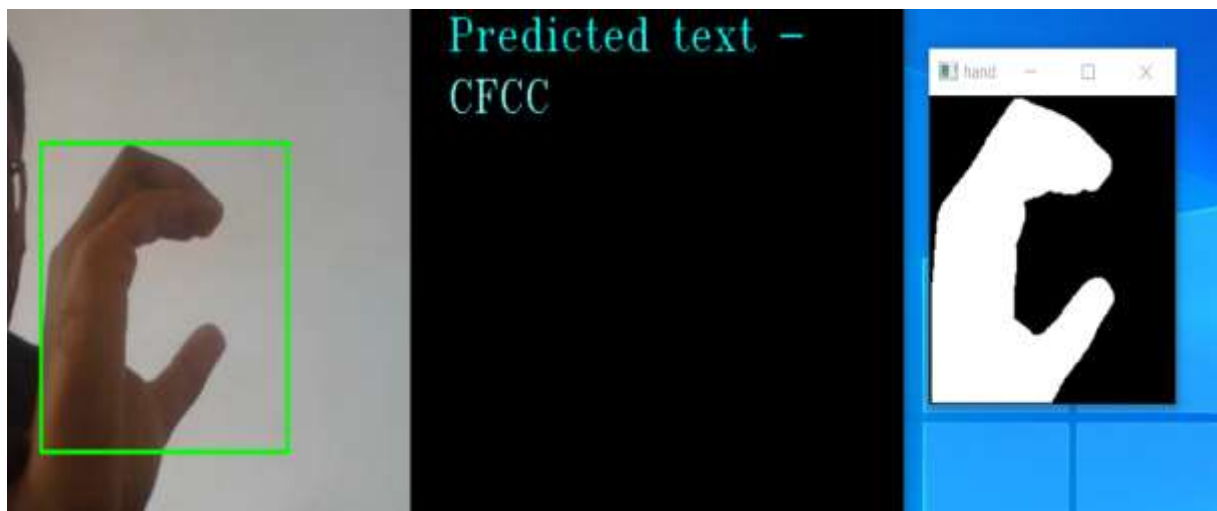


Figure 6

Conclusion

This paper demonstrates how we can use CNN and OpenCV to detect the sign of a hand and convert it into text. Here we use ASL for the training of model. Hand sign Detector is obtained with an accuracy of 75%. The CNN model is trained on a dataset containing alphabets of ASL (American Sign Language). This project can be extended to further to increase its accuracy. Sign language is spoken more in circumstances rather than as finger spelling language. The main goal has been achieved. A few finer points need to be considered when we run into this project. Bad gesture postures will not yield correct predictions. This project can be extended in a few ways in the future, and it can be expanded to work for other domestic sign languages with necessary datasets and training. This project implements a hand sign recognition system. However, sign languages are also spoken on a contextual basis, where each sign could represent an object or verb.

References

- [1] Narayan, Vipul, and A. K. Daniel. "Design consideration and issues in wireless sensor network deployment." (2020): 101-109.
- [2] T. Ahmad, C. J. Taylor, "Tracking and recognizing hand gestures, using statistical shape models", The University of Manchester.
- [3] Jannatun Magna, Mohammad Shahriar, "Hand gesture- Based Character Recognition OpenCV and Deep Learning", 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI).
- [4] Ankit Chaudhary, Karen Das, Sonia Raheja "Intelligent Approaches to interact with Machines using Hand Gesture Recognition in Natural way: A Survey", International Journal of Computer Science & Engineering Survey (IJCSES) Vol.2, No.1, Feb 2011.
- [5] Pratibha Pandey, Vinay Jain, Hand Gesture Recognition for Sign Language Recognition: A Review, International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 3, March 2015.
- [6] 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.
- [7] D. Bhavana, K. Kishore Kumar, Medasani Bipin Chandra, P.V. Sai Krishna Bhargav, D. Joy Sanjanaa, and G. Mohan Gopi. Hand Sign Recognition using CNN [J]. Int J Performability Eng, 2021, 17(3): 314-321.
- [8] Narayan, Vipul, and A. K. Daniel. "RBCHS: Region-based cluster head selection protocol in wireless sensor network." Proceedings of Integrated Intelligence Enable Networks and Computing: IIENC 2020. Springer Singapore, 2021.
- [9] Narayan, Vipul, and A. K. Daniel. "CHOP: Maximum coverage optimization and resolve hole healing problem using sleep and wake-up technique for WSN." ADCAI: Advances in Distributed Computing and Artificial Intelligence Journal 11.2 (2022): 159-178.

- [10] Narayan, Vipul, and A. K. Daniel. "CHHP: coverage optimization and hole healing protocol using sleep and wake-up concept for wireless sensor network." *International Journal of System Assurance Engineering and Management* 13.Suppl 1 (2022): 546-556.
- [11] Narayan, Vipul, and A. K. Daniel. "IOT based sensor monitoring system for smart complex and shopping malls." *Mobile Networks and Management: 11th EAI International Conference, MONAMI 2021, Virtual Event, October 27-29, 2021, Proceedings*. Cham: Springer International Publishing, 2022.
- [12] Narayan, Vipul, and A. K. Daniel. "Energy Efficient Protocol for Lifetime Prediction of Wireless Sensor Network using Multivariate Polynomial Regression Model." *Journal of Scientific & Industrial Research* 81.12 (2022): 1297-1309.
- [13] Awasthi, Shashank, et al. "A Comparative Study of Various CAPTCHA Methods for Securing Web Pages." *2019 International Conference on Automation, Computational and Technology Management (ICACTM)*. IEEE, 2019.
- [14] Narayan, Vipul, and A. K. Daniel. "FBCHS: Fuzzy Based Cluster Head Selection Protocol to Enhance Network Lifetime of WSN." *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal* 11.3 (2022): 285-307.
- [15] Narayan, Vipul, et al. "E-Commerce recommendation method based on collaborative filtering technology." *International Journal of Current Engineering and Technology* 7.3 (2017): 974-982.
- [16] Narayan, Vipul, et al. "To Implement a Web Page using Thread in Java." (2017).
- [17] 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).
- [18] 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).
- [19] Smiti, Pujja, 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.
- [20] 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.
- [21] Narayan, Vipul, and A. K. Daniel. "Multi-tier cluster based smart farming using wireless sensor network." *2020 5th international conference on computing, communication and security (ICCCS)*. IEEE, 2020