iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** XI **Month of publication:** November 2021

DOI: https://doi.org/10.22214/ijraset.2021.38777

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: 7.429
Volume 9 Issue XI Nov 2021- Available at www.ijraset.com

Voice and Face Recognition for Web Browser Security

Reddy Phanidhar Reddy[1], Y. Ashok Reddy[2], C. Gowtham[3], Santhosh Reddy Chilaka[4]
[1, 2, 3, 4]*School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India*

***Abstract:** This paper analysis about browser privacy multimodal authentication mechanisms. Face and voice recognition will be used as authentication methods in this process. The OpenCV library is used in the framework's face recognition section. It detects and recognizes faces from a database using basic eigen face recognition approaches. The MFCC (Mel Frequency Cepstrum Coefficients) and Gaussian Mixture Model are used to recognize voices. Following successful authentication, the cookies on the local hard disc are decrypted, allowing us access to the browser cookies. Initially, after a user registers, we will encrypt the browser cookies with AES, one of the most secure encryption methods available.*
***keywords:** MFCC, Gaussian Mixture Model, Browser cookies, authentication, AES, encryption, decryption, Open CV, Eigen.*

I. INTRODUCTION

Our goal was to create a framework that could protect a user's session information in a shared computer environment, such as a family computer or a University or School computer, where any unaffiliated person may access sessions that were not closed when the browser was closed. Private messaging programmes, such as Instagram, whatsapp or mail sessions, such as gmail or outlook, are examples of such sessions that may contain very sensitive and critical information.

II. MOTIVATION

Many innovative implementations of Biometrics have been researched and published to the public and are widely applied in many use cases. But the sort of use case that we are going to discuss in our project implementation hasn't been explored. In this paper we are going to develop computer image processing and train the system to detect the face and secure browsers from false authentications.

III. METHODOLOGY

Biometrics are one of the very few things that can be used safely in such environments without being forged. Traditional methods like passwords can be leaked by shoulder surfing or keyloggers that are usually installed on computers in unsafe environments like University computers. Biometrics don't have such disadvantages and can be used safely in public spaces. We have implemented face recognition based on the OpenCV library that has been developed specifically for computer image processing. This library also has many other supplementary functions that we have used to recognize faces from a few images of users that have been captured and used to train various methods that are available.

IV. PROJECT DESCRIPTION AND GOALS

A. To identify the user correctly and give access to any session cookies we have implemented a multimodal authentication system that uses Face recognition along with a Voice identification to recognize the user from a database of multiple users.

B. To Identify a person using their face we have used OpenCV in python. This enables the script to use a camera that is available to capture images of the person and then use those images to extract features from and then train the machine learning algorithm using those extracted feature dataset.

C. To Identify users, A Gaussian Mixture model has been used with MFCC for feature extraction. Multiple recordings of user speech are used to create training data. These audio clips are then trimmed and filtered by using MFCC to remove unwanted frequency ranges. This data is passed to a Gaussian Mixture model to classify the frequency ranges and identify a user.

And finally our main goal is to correctly identify a user using their Face and Voice biometrics to grant access to their sensitive session information

V. MODULES

1) *Open CV:* OpenCV (Open Source Computer Vision Library) is a free software library for computer vision and machine learning. OpenCV was developed to offer a standard infrastructure for computer vision applications and to let commercial products incorporate machine perception more quickly. Because OpenCV is a BSD-licensed product, it is simple for companies to use and change the code. More than 2500 optimised algorithms are included in the library, which contains a comprehensive range of both traditional and cutting-edge computer vision and machine learning techniques.

The following features we used which comprises of OpenCv are:

a) Image Processing(imgproc): this module includes linear and non-linear image filtering, geometric image transformations (resizing, affine and perspective warping, remapping), colour space conversion, histograms.
b) Video Analysis(video) : In this module consists of motion estimation and object tracking algorithm with background subtraction.
c) High-level GUI (highgui) : An interface for simple UI(user-interface) capabilities.
d) Video I/O(videoio) : An interface for capturing videos.

2) *MFCCs:* The MFCCs (mel-frequency cepstral coefficients) are the coefficients that make up an MFC. They're mainly composed of a cepstral representation of the audio sample (a nonlinear "spectrum-of-a-spectrum"). The distinction between the cepstrum and the mel-frequency cepstrum is that the MFC's frequency bands are evenly separated on the mel scale, which more closely approximates the human auditory system's response than the normal spectrum's linearly spread frequency bands. This frequency warping can help with sound representation such as audio compression.

a) Take (a windowed clip of) a signal and perform the Fourier transform on it.
b) Apply triangular overlapping windows or cosine overlapping windows to map the powers of the spectrum obtained above onto the mel scale.
c) Take the logs of the powers at each of the mel frequencies.
d) compute the discrete cosine transform of the list of mel log powers.
e) The amplitudes of the resultant spectrum are the MFCCs.

3) *GMM:* Gaussian mixture models are a probabilistic representation of normally distributed subpopulations within a larger population. In general, mixture models don't require knowing which subpopulation a data item belongs to, enabling the algorithm to automatically learn the subpopulations. This is a kind of unsupervised learning because the subpopulation assignment is unknown. GMMs have been used to extract features from voice data, and that they've also been used to predict object positions at each frame in a video sequence, where the number of mixture components and their means predict entity locations.
4) *PyAesCrypt:* The Advanced Encryption Standard (AES) is perhaps the most popular and widely adopted symmetric encryption algorithm available today (AES).pyAesCrypt is a Python 3 file encryption module and script that encrypts and decrypts files and binary streams using AES256-CBC. The AES Crypt file format is supported by pyAesCrypt (version 2). It's Free Software, licenced under the Apache License 2.0.

VI. DESIGN APPROACH AND DETAILS

Our project source directory has 4 main files:

1) *Main_function.py:* This script supplies all the common functions that are used in the subsequent scripts.
2) *Add_user.py:* This script uses the functions in main_functions script to create a face and voice model for a particular user. This script also encrypts any Cookie file found in the users google chrome profile directory.
3) *Recognize.py:* This script uses the functions defined in the main_funcctions script and uses the face and voice model generated for a user by the add_user script to recognize a user. Face and Voice samples are captured by this script when it runs, this data is compared with the models that were stored on the disk by add_user script.
4) *Delete_user.py:* This script is used to delete any registered users data from the disk by supplying a username to the script. This also decrypts the encrypted cookie file and reverts the state of files to as they were before running the add_user script.

First we start by launching add_user.py script:

- This script asks the user for the input of username. This name is then used to store the files and data models that will be created later .
- The script then opens the camera and takes about 300 pictures of the user.
- The images are cropped by opencv model using a haar cascade; these images are then passed to the facenet model (weights are stored in a folder called weights) that trains on user images.
- The face model of the user that is generated is stored in a folder called face_database.
- User is then asked to speak their name out loud 3 times to generate a voice model.
- Generated model of the voice is then stored in a gmm_models folder.
- The cookies file in linux is at a location called $HOME/.config/google-chrome/Default/Cookies This file is then encrypted using AES encryption algorithm and a hardcoded password. This password is stored in the file in a variable.

User cookies are now secure and can't be read without the password or biometrics.

For recognizing the user and decrypting the cookie for just during the session, we use recognize.py script:

- This script first asks the user to speak their name or something very similar.
- That audio clip is then compared against the generated GMMs.
- If the voice of the speaker matches with any of the GMM, face recognition starts. Or the script stops running.
- Since the name of GMMs is username, name can be extracted from this and passed to face recognition model.
- Face recognition module then starts and takes multiple pictures of the user.
- These images are compared to the face model of the name that has been supplied by the voice recognition module.
- If the face matches, the cookie file on the disk is decrypted and a subprocess in python that launches google-chrome is started.
- Users can normally browse the web using their old session cookies and close the browser.
- When Browser is closed the subprocess ends and the main process is notified.
- Main process then re-encrypts the cookie file and deletes the plaintext file and exits.

When the user wants to delete their account .i.e Delete_user.py script:

- Users voice GMM is deleted from the gmm_models folder
- The face model is deleted from the face_database folder.
- Cookie file is decrypted by using the hardcoded password. This can be modified a little bit to entirely delete the user cookies but this will be like resetting the browser.

VII. RESULTS

Figure 1: User registration has started and the script has asked for username

Figure 2: Images of Users face are being captured

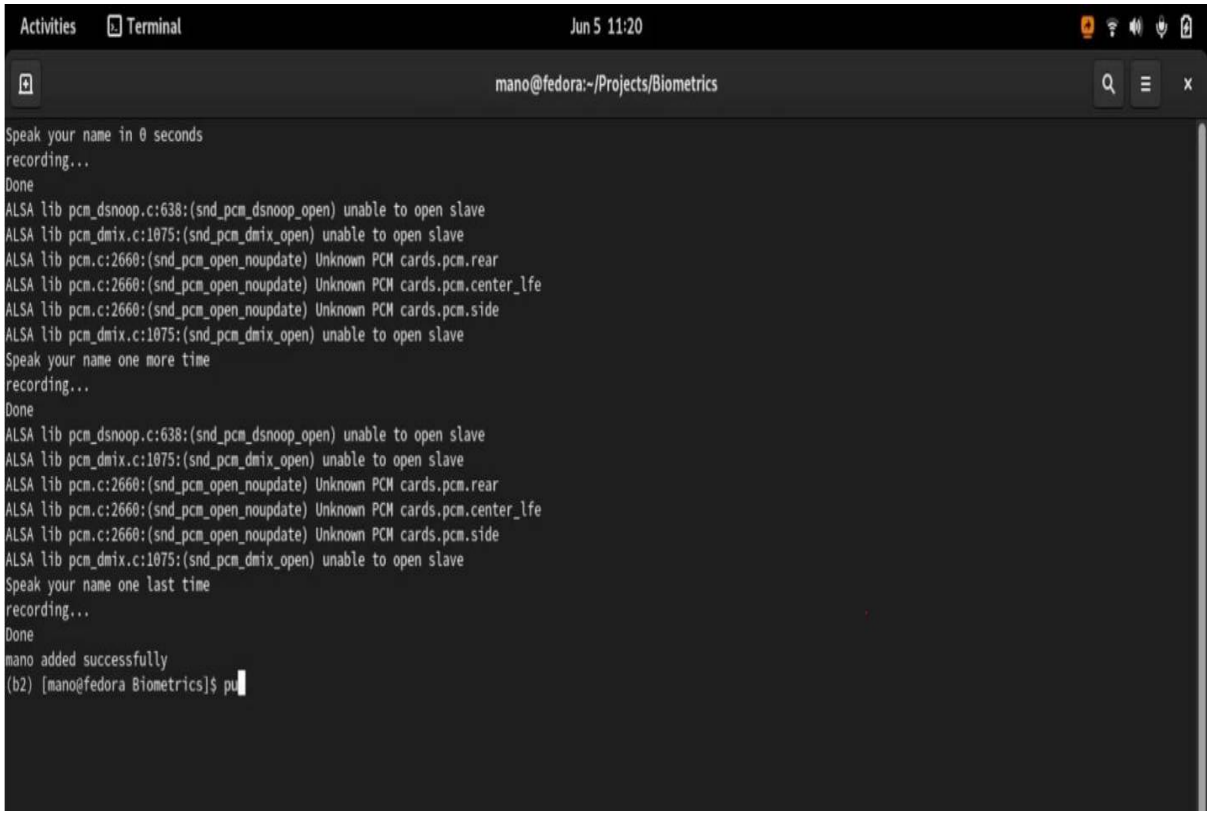


Figure 3: Users voice samples are captured and User is registered after encrypting cookies

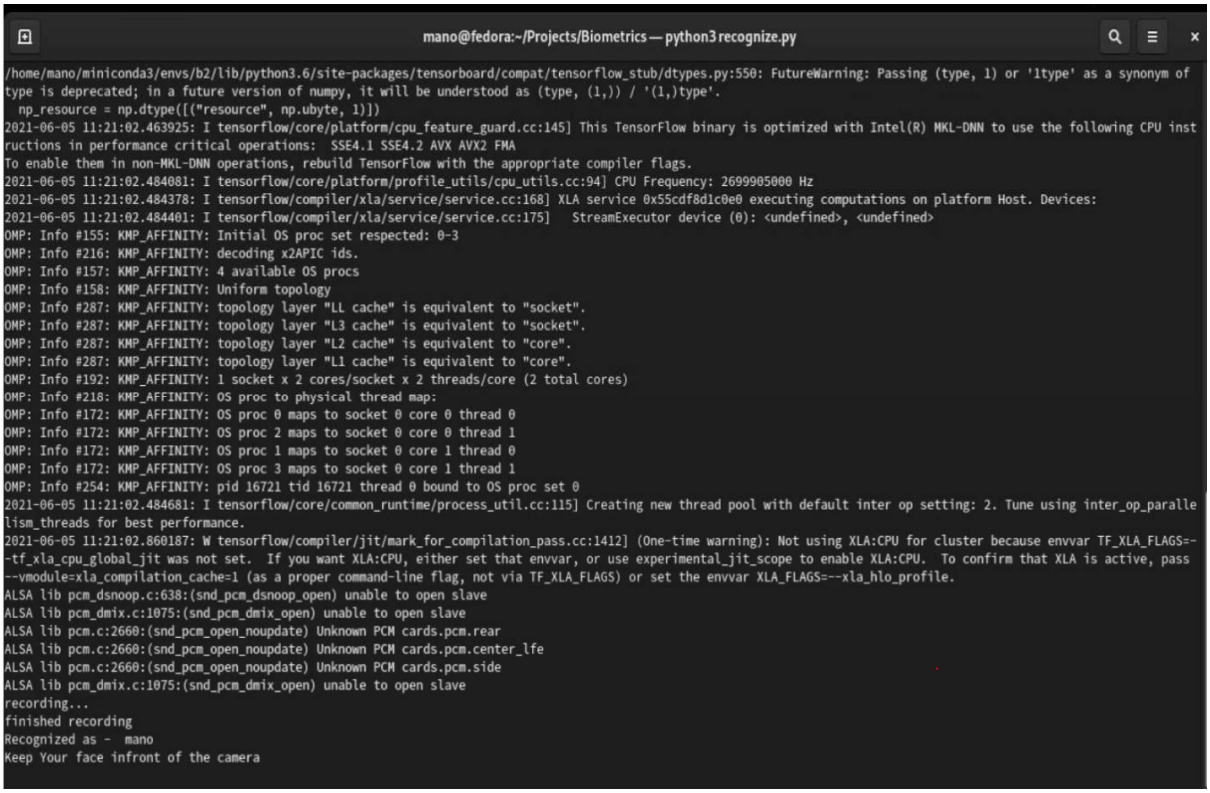


Figure 4: Recognition script has been started by the user and it is listening for users voice

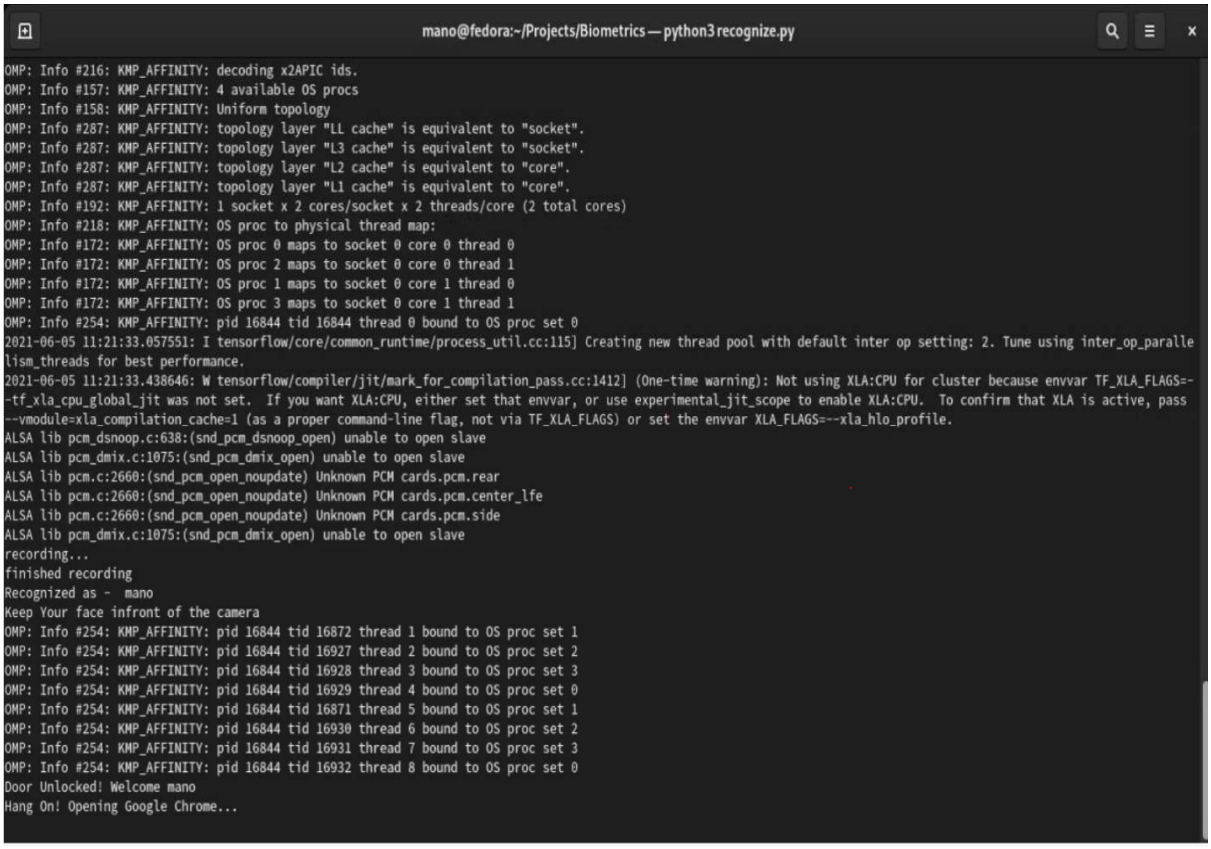


Figure 5: Users Face images were taken and User has been identified as legitimate user

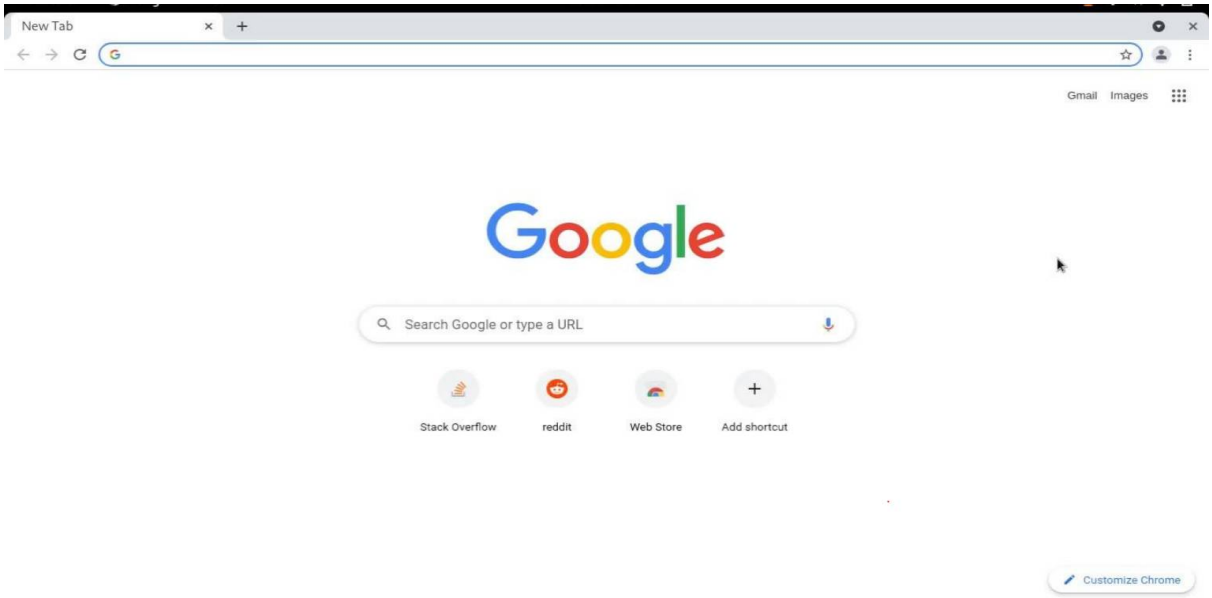


Figure 6: A subprocess opened Google Chrome Browser with Decrypted Cookies

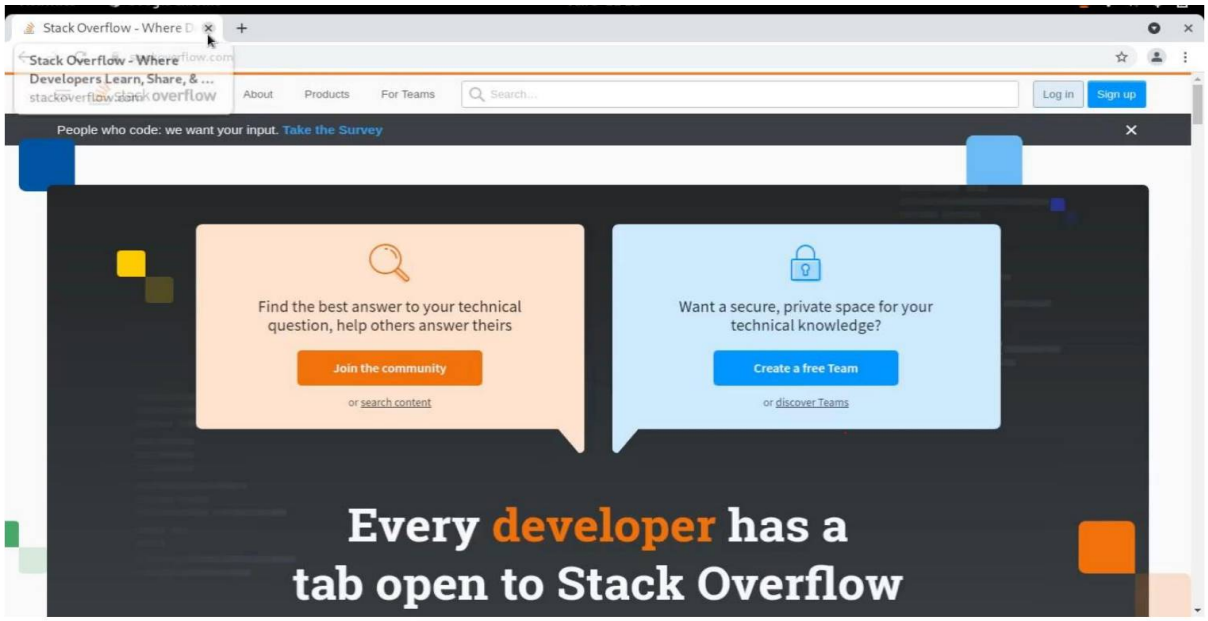


Figure 7: There is no cookie prompt after decrypting whereas a cookie consent will be prompted if you are visiting the website without any previous cookies related to that site.

This framework works really well. Browser cookies can't be read without running the script even if an administrator account tries to access them. Accuracy of the face recognition system is very high in normal and bright lighting conditions. Voice recognition works well in optimal conditions.

However, there are a few drawbacks to this system.

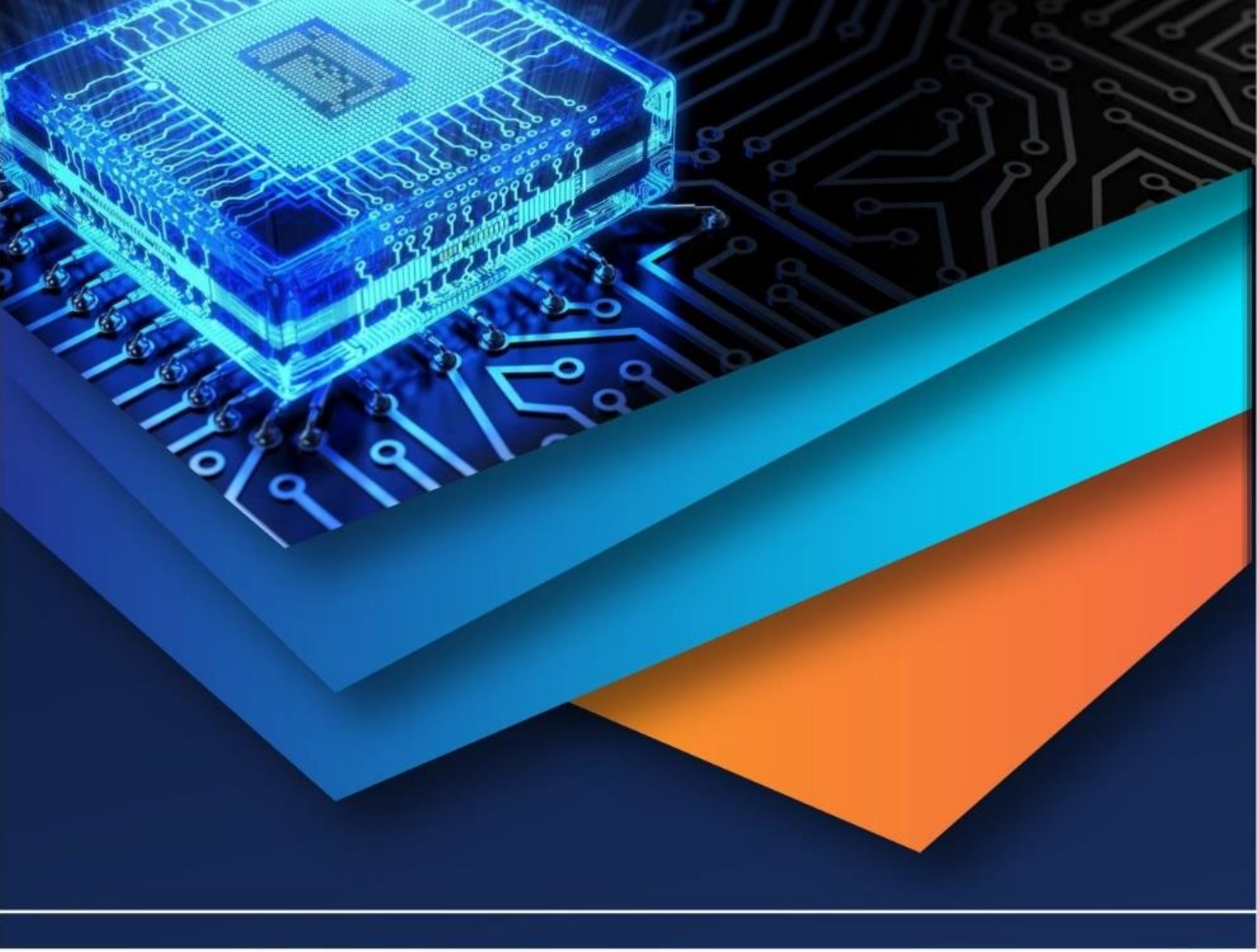
1) The face recognition algorithm doesn't work well with low resolution cameras and poor lighting conditions like backlight and dark conditions.
2) Voice recognition has to be performed in really good surroundings without any loud noises and disturbances. Or the voice sample will not be identified.

VIII. CONCLUSION

This framework works really well. Browser cookies can't be read without running the script even if an administrator account tries to access them. Accuracy of the face recognition system is very high in normal and bright lighting conditions. Voice recognition works well in optimal conditions. By following the process we can protect our browsers from unauthorized third parties.

REFERENCES

[1] Joseph, T.,Kalaiselvan, S.A., Aswathy, S.U. et al. A multimodal biometric authentication scheme based on feature fusion for improving security in cloud environments. https://doi.org/10.1007/s12652-020-02184-8
[2] O. Tymchenko, B. Havrysh, O. O. Tymchenko, O. Khamula, B. Kovalskyi and K. Havrysh,"Person Voice Recognition Methods," 2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, 2020, pp. 287-290, doi: 10.1109/DSMP47368.2020.9204023.
[3] Prasad, Puja & Pathak, Rashmi & Gunjan, Vinit & Rao, H.. (2020). Deep Learning Based Representation for Face Recognition. 10.1007/978-981-13-8715-9_50
[4] H. Cevikalp, H. S. Yavuz and B. Triggs, "Face Recognition Based on Videos by UsingConvex Hulls," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no.12, pp. 4481-4495, Dec. 2020, doi: 10.1109/TCSVT.2019.2926165.
[5] Y. Huang and H. Hu, "A Parallel Architecture of Age Adversarial Convolutional Neural Network for Cross-Age Face Recognition," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 1, pp. 148-159, Jan. 2021, doi: 10.1109/TCSVT. 2020.2965739.
[6] A. Sepas-Moghadda m, A. Etemad, F. Pereira and P. L. Correia, "Long Short-Term Memory With Gate and State Level Fusion for Light Field-Based Face Recognition," in IEEE Transactions on Information Forensics and Security, vol. 16, pp. 1365-1379, 2021, doi: 10.1109/TIFS.202 0.3036242.
[7] N. H. Tandel, H. B. Prajapati and V. K. Dabhi, "Voice Recognition and Voice Comparison using Machine Learning Techniques: A Survey," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020, pp. 459-465, doi: 10.1109/ICACCS48 705.2020.9074184.
[8] Feng Zhao, Jing Li, Lu Zhang, Zhe Li, Sang-Gyun Na, Multi-view face recognition using deep neural networks, Future Generation Computer Systems, Volume 111, 2020, Pages 375-380, ISSN 0167-739X
[9] M. Luo, J. Cao, X. Ma, X. Zhang and R. He, "FA-GAN: Face Augmentation GAN for Deformation-Invariant Face Recognition," in IEEE Transactions on Information Forensics and Security, vol. 16, pp. 2341-2355, 2021, doi: 10.1109/TIFS.2021.3 053460.
[10] Madarkar, Jitendra and Sharma, Poonam. 'Occluded Face Recognition Using NonCoherent Dictionary'. 1 Jan. 2020: 6423 – 6435.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY