International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 619-622 www.ijaem.net ISSN: 2395-5252

The Application For Facial Recognition In Big Data Security

Patel Tirth, Makawana Parth, Bhatt Shivam
Departmet of Information Technology. Parul Institute of Engginering and Technology Limda, Waghodia, India
Departmet of Information Technology. Parul Institute of Engginering and Technology Limda, Waghodia, India
Departmet of Information Technology. Parul Institute of Engginering and Technology Limda, Waghodia, India

Submitted: 01-02-2022 Revised: 07-02-2022 Accepted: 10-02-2022

ABSTRACT--A Facial Recognition System is a technology capable of matching a human face from a digital image or a video frame against a database of faces, typically employed to authenticate users through ID verification services, works by pinpointing and measuring facial features from a given image. Big Data is a field that performs, analyse and treats the information in systematic way of extracting and dealing with large data-sets.Big Data analysis includes various challenges like capturing the data, store, sharing, transfer, visualization, querying, updating the information.
Keywords— Big Data Access, Face Authentication, Face Authorization, Verification, Identification, Biometrics

I. INTRODUCTION

Face recognition system is a category of biometric security. In other forms of biometric software includes voice recognition, fingerprint recognition and eye-retina or iris recognition. Although the accuracy of face recognition systems is considered as biometric technology, it is always lower than iris recognition and fingerprint recognition. Furthermore, it is widely adopted due to its contactless process. Facial Recognition System's have been deployed in advanced human computer interaction, video surveillance and automatic indexing of images.

Big Data is not a small thing, and we can't describe it in the context of size. Volume is the main structure of big data architecture. Today almost every organization is thinking of adopting big data to oversee the potential and utilization of data. Hadoop-HDFS is the environment to process the large datasets. Big Data is a collection of the data huge in volume, yet growing exponentially with time. Data with many fields offer great statistical power, while data with higher complexity may lead to a typical false discovery rate.

In this paper, we have talked about the implications of facial recognition as a authentication system for big data framework. We have also described how beneficial is the facial recognition system in big data as a biometric authentication.

II. FACE RECOGNITION

Facial Recognition System comprises of two sub processes which are: Detection of the face and Recognition of the face on captured images which is taken as an input.

1.1 DETECTION

Face Detection is an AI based computed technology that is used in finding and identifying human faces in digital images. When it comes to face detection, it is necessary for the algorithms to know which part of an image contains face. Furthermore, it compares the images as an input with a database system which is known or unknown to find a match. It plays a vital role for the various application to know which parts of an image is used to generate the faceprints. Afterall it will be compared with previously stored faceprints to establish whether the face is matched or not.

Algorithms used in face detection are as follows:

I. Haar Cascade Classifier
II. Multi- Task Cascaded Convolutional Neural Network (MTCNN)
III. Linear Binary Pattern Histogram (LBPH)

1.1.1 RECOGNITION

Facial Recognition is a way of

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 619-622 www.ijaem.net ISSN: 2395-5252

identifying and conforming an individuals identify using their face. In these systems it can be used as a real time picture, captured images, video surveillance and raw images.

Face recognition is a category of biometric security. In other forms, biometric software includes voice recognition, fingerprint recognition and iris recognition. In mathematical context of face recognition, we must include the basic formula for calculating the designated output.

Algorithms used in Face Recognition are:

I. Eigenfaces
II. Fisher faces
III. Delaunay triangulation

III. BIG DATA

The concept of "Big Data" is defined as a software utility that is designed to analyze, process and to extract the information from a typical kind of large datasets. Basically, it is used for optimizing the parallel processing of structured and unstructured data, using negligible hardware costs.

Hadoop is an open-source software framework for storing all types of data and running applications on a specified cluster. Usually, the storage space depends on the commodity hardware. It provides massive data storage for any relatable data, enormous processing power and one ability to handle concurrent tasks or jobs in a virtual session.

1.2 NEED OF SECURITY IN BIG DATA

In a Hadoop process, there are multiple types of data which are combined and stored in a Hadoop periphery and after this the storage information is processed accordingly. Eventually, after utilizing the potential and power of big data they are using Hadoop to process these large datasets.

As of now, security is concerned protecting the data is most important aspect of the data. In several organizations usually people can't handle the database structured in single platform so dividing the setup is the only solution.

Now, as if we talked about the security, mainly it comprises various strategies like keeping out unauthorized users and intrusions with firewalls, making user authentication reliable and to give training to the end-user.

Basic data security measures in compliance consists of data loss protection and privacy mandates like the: "GENERAL DATA PROTECTION REGULATION" (GDPR). End user must use the latest antivirus package. Continuously, monitoring and auditing all of the access to sensitive data. We can also use the alert and react to intrusions and unlawful actions in real time systems.

Furthermore, the main concerns in big data security are as follows: -

- Vulnerabilities of manipulated data on end point devices may transmit the false data.
- Data Mining solutions are the heart of man big data environments. List of tools also follow the pattern of finding roots in unstructured data. For this reason, companies need to add a extra security layer.
- In a distributed architecture of big data, intrusion detection plays a vital role. It enables the security teams to protect the platform from the exploited vulnerable network.

IV. EXISTING SECURITY USED IN BIG DATA

- The centralized key management system always uses a single point access to its audit logs and term of policies. Hence, a reliable key management system is much needed to build a uniform responsive system.
- The preserve data privacy depends upon different implementing techniques to uphold the data privacy. Encouraging its practices and preserving data composition will include multiple databases of reviewing and monitoring the database infrastructure.
- The Big Data Distribution always includes the distributed programming frameworks such as: Hadoop, which plays a huge role for maintaining modern day sources.
- In present situation organizations do include the establishment of several methods like "KERBEROS AUHENTICATION" and "DE-IDENTIFY" type of approaches does ensure the conformity to predefined the structured of securing data policies.
- The authorization access to files with predefined security policy which would help in not leaking the information via several resources of system which is used by MAC (Mandatory Access Control) such as a sentry tool in Apache HBase.
- For regular maintenance and assistance of the data leakage IT department must check the worker nodes and substantial mappers in cloud environment. Afterall, they will keep an eye on the altered form of data for avoiding duplicacy of data.

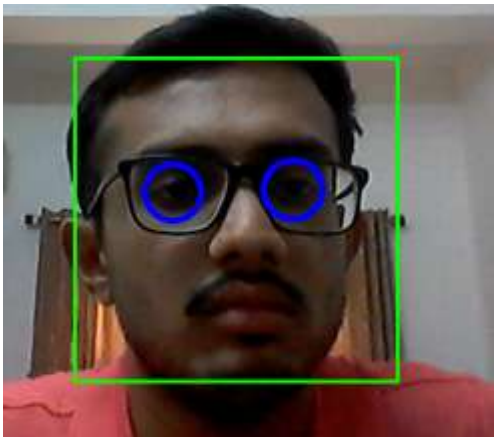
V. PROPOSED IDEOLOGY AND METHODOLOGY

To increase and improve access security in this project, we are proposing the idea to use part of

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 619-622 www.ijaem.net ISSN: 2395-5252

the biometrics with existing security system ~ KERBEROS. The part of the biometrics, we are using is face recognition system which is mentioned above.As it is a facial recognition system it is using two algorithms to fulfill the process of detection and the process of recognition which consist of detecting the face of a particular entity.

In our project, for detecting of the human faces we have specifically used the "VIOLA JONES ALGORITHM" and for recognizing the identity of the user we have implemented the Multi- Task Cascaded Convolutional Neural Network (MTCNN) Algorithm.



[Fig. 1. DETECTING A FACE]

As we have mentioned above the security concerns for accessing the big data without any security tools which can lead to wiping out, destroying and manipulating of the data. Therefore, by using face recognition for granting the access and Kerberos for Ticket Granting Server (TGS) one can improve access security and prevent any kind of unwanted data access. Furthermore, we would likely to include the authorization and authentication of the particular user in the Hadoop based big data ecosystem.

In addition, as it does not require separate maintenance of identity information, reducing complexity and risks. Ultimately, a new authentication solution is been used to cover the malicious use of the Hadoop cluster (as it would be prevented).

VI. CHALLENGES AND LIMITATIONS

The three major Big Data concerns are about data privacy, data security and data discrimination. Hence, to overcome these particular problems we would ensure that validating and accumulating data from different sources and data cleansing is the only solution.

Following the drawbacks of using face recognition which includes violating the personal freedom and right of virtues, risk of malfunctioning in data, increasing potential data theft.

Enlisting the major limitations of Face Recognition are as follows:

1) Pose variations
2) Facial expression changes
3) Ageing of the face
4) Image resolution
5) Expression

VII. FUTURE WORK

To improve the vast usage of face recognition system is to use various algorithms and using that particular algorithm increase the accuracy of it. Also, other way to improve face recognition is to collect versatile training datasets with detailed visual data.Deploying the verification and tracking of legitimate users for the avoidance of criminal activities.

VIII. CONCLUSION

To conclude that, we all know that Facial Technology has been in trend from past few years. Nowadays,machines are automatically verifying the information of the user for secure transaction, for surveillance and also for security purposes.

REFERENCES

[1]. Zhang, N., Luo, J., & Gao, W. (2020). Research on Face Detection Technology Based on MTCNN. 2020 International Conference on Computer Network, Electronic and Automation (ICCNEA). doi:10.1109/iccnea50255.2020.0004

[2]. Kaur, P., Krishan, K., Sharma, S. K., & Kanchan, T. (2020). Facial-recognition algorithms: A literature review. Medicine, Science and the Law, 002580241989316. doi:10.1177/0025802419893168

[3]. Ma, Y., Kan, M., Shan, S., & Chen, X. (2020). Learning deep face representation with long-tail data: An aggregate-and-disperse approach. Pattern Recognition Letters, 133, 48–54. doi:10.1016/j.patrec.2020.02.007

[4]. Cheng, E.-J., Chou, K.-P., Rajora, S., Jin, B.-H., Tanveer, M., Lin, C.-T., … Prasad, M. (2019). Deep Sparse Representation Classifier for Facial Recognition and Detection System. Pattern Recognition Letters. doi:10.1016/j.patrec.2019.03.006

[5]. Bah, S. M., & Ming, F. (2019). An improved face recognition algorithm and its application in the attendance management system. Array, 100014. doi:10.1016/j.array.2019.100014

[6]. Rameswari, R., Naveen Kumar, S., Abishek Aananth, M., & Deepak, C. (2020). Automated access control system using face recognition. Materials Today: Proceedings. doi:10.1016/j.matpr.2020.04.664

[7]. Kemenristekdikti (2020). New approach to the identification of the easy expression recognition system by robust techniques (SIFT, PCA-SIFT, ASIFT and SURF).DOI: 10.12928/TELKOMNIKA.v18i2.13726. ISSN: 1693-6930

[8]. Oloyede, M. O., Hancke, G. P., & Myburgh, H. C. (2020). A review on face recognition systems: recent approaches and challenges. Multimedia Tools and Applications. doi:10.1007/s11042-020-09261-2

[9]. Bhathal, G. S., & Singh, A. (2019). Big data: Hadoop framework vulnerabilities, security issues and attacks. Array, 100002. doi:10.1016/j.array.2019.100002

[10]. Hussain, Toshifa and Sanga, Anirudh and Mongia, Shweta, Big Data Hadoop Tools and Technologies: A Review (October 1, 2019). Proceedings of International Conference on Advancements in Computing & Management (ICACM) 2019, Available at SSRN: https://ssrn.com/abstract=3462554 or http://dx.doi.org/10.2139/ssrn.3462554

[11]. O. Ekundayo and S. Viriri, "Facial Expression Recognition: A Review of Methods, Performances and Limitations," 2019 Conference on Information Communications Technology and Society (ICTAS), Durban, South Africa, 2019, pp. 1-6, doi: 10.1109/ICTAS.2019.8703619.