Data Mining using face detection with Logic

[1]Bhavishya Rathore, [2]Dr Neelesh Jain, [3]Prof Jaya Malviya

[1]Student, [2]Professor, [3]Assistant Professor
Department of Computer Science & Engineering
SAM Collage of Engineering, Bhopal (M.P.)

***Abstract*: The growth of massive data stores has led to the development of a number of automated processors that work to discover relationships in and between the data in those stores. These processors are often referred to by a number of names including data mining, knowledge discovery, pattern recognition, artificial and machine learning. Data mining is the nontrivial extraction of implicit, previously unknown, and potentially useful information from data. It is used to automatically extract structured knowledge from large datasets. The application of logic with data mining makes information understandable to human. Data mining can have many methods like association rules, classification, clustering. One of the methods of implementing association rules is Apriori algorithm.**

In this thesis, A Apriori System is built; it uses Apriori algorithm alone, then Apriori algorithm with the application of logic to find association rules. It will find the relationships among items stored in a supermarket to present knowledge about what are the most soled items and the relations among items. From the experimental results, it was found that functions filters the results and make number of extracted rules less than the number of rules extracted by applying Apriori algorithm only.

***Keywords*: Biometrics, LBP, OpenCV, Python, Surveillance.**

Introduction

Progress in digital data acquisition and storage technology has resulted in the growth of huge databases. This has occurred in all areas of human endeavor, from the mundane (such as supermarket transaction data, credit card usage records, telephone call details, and government statistics) to the more exotic (such as images of astronomical bodies, molecular databases, and medical records). After that interest was grown toward tapping these data, and extracting from them information that might be of value to the owner of the database. The discipline concerned with this task has become known as data mining. Data mining is the analysis of large observational data sets to find unsuspected relationships, and to summarize the data in novel ways that are both understandable and useful to the data owner [Dav01].

The amount of data stored in databases continues to grow fast. Intuitively, this large amount of stored data contains valuable hidden knowledge, which could be used to improve the decision-making process of an organization. For instance, data about previous sales might contain interesting relationships between products and customers. The discovery of such relationships can be very useful to increase the sales of a company. However, the number of human data analysts grows at a much smaller rate than the amount of stored data. Thus, there is a clear need for semi automatic methods for extracting knowledge from data. This need has led to the emergence of a field called data mining and knowledge discovery. This is an interdisciplinary field, using methods of several research areas (specially machine learning and statistics) to extract high level knowledge from real-world data sets. Data mining is the core step of a broader process, called knowledge discovery in databases, called knowledge discovery in databases, or knowledge discovery, for short [Ale01].

Data Mining

Data mining is the process of extracting meaningful information from large quantities of data. It involves uncovering patterns in the data and is often tied to data warehousing because it makes such large amounts of data usable. Data elements are grouped into distinct categories so that predictions can be made about other pieces of data. For example, a bank may wish to ascertain the characteristics that typify customers who pay back loans. Although this could be done with database queries, the bank would first have to know what customer attributes to query for. Data mining can be used to identify what those attributes are and then make predictions about future customer behavior[Sar05]..

Data Mining and Logic

Knowledge discovery, whose objective is to obtain useful knowledge from data stored in large repositories, is recognized as a basic necessity in many areas, especially those related to business. Since data represent a certain real-world domain, patterns that hold in data show interesting relations that can be used to improve human understanding of that domain. Data mining is the step in the knowledge discovery process that attempts to discover novel and meaningful patterns in data. The theory of sets can certainly help data mining to reach this goal. It is widely recognized that many real world relations are intrinsically. For instance, clustering generally provides a more suitable partition of a set of objects than crisp clustering do. Moreover, sets are an optimal tool to model imprecise terms and relations as commonly employed by humans in communication and understanding. As a consequence, the theory of sets is an excellent basis to provide knowledge expressed in a meaningful way.

PROBLEM DEFINITION

Face recognition has many challenges due to illumination variations, large dimensionality, uncontrolled environments, pose

variations and aging. In the recent years, Face recognition get remarkable improvement and accuracy to overcome these challenges, but illumination change is still changing. It proposed an NIR imaging system that gives satisfactory results for face recognition in illumination variance conditions but it does not give good results when matching NIR image to visible images. Unfortunately, all face images in the database are store in the visible spectrum.

A general articulation of the face acknowledgment issue (in PC vision) can be figured as follows: given still or video pictures of a scene, distinguish or confirm at least one person in the scene utilizing a put-away database of appearances. Facial acknowledgment for the most part includes two phases: Face Detection where a photograph is looked to discover a face, at that point the picture is prepared to harvest and concentrate the individual's face for simpler acknowledgment. Face Recognition where that identified and prepared face is contrasted with a database of known appearances, to choose who that individual is. Since 2002, face identification can be performed decently effectively and dependably with Intel's open-source system called OpenCV. This system has an inbuilt Face Detector that works in around 90-95% of clear photographs of an individual looking forward at the camera. In any case, recognizing an individual's face when that individual is seen from an edge is generally harder, once in a while requiring 3D Head Pose Estimation. Additionally, the absence of appropriate brilliance of a picture can significantly expand the trouble of recognizing a face, or expanded differentiation in shadows on the face, or perhaps the image is hazy, or the individual is wearing glasses, and so forth.

Face acknowledgment anyway is considerably less solid than face recognition, with a precision of 30-70% when all is said in done. Face acknowledgment has been a solid field of research since the 1990s, however, it is as yet a far route away from a dependable strategy for client validation. An ever-increasing number of procedures are being built up every year. The Eigenface strategy is viewed as the least complex technique for precise face acknowledgment, however numerous other (considerably more confused) strategies or mixes of different techniques are marginally progressively exact.

Intel's open-source PC vision library can incredibly disentangle computer vision programming. It incorporates propelled abilities - face identification, face following, face acknowledgment, Kalman sifting, and an assortment of artificial intelligence (AI) techniques- in prepared touse structure. What's more, it gives numerous essential PC vision calculations using its lower-level APIs**.**

LITERATURE REVIEW

Asas During the most recent decade, a few promising face identification calculations have been created and distributed. Among these three stands apart because they are regularly alluded to when execution figures and so on are looked at. This segment quickly presents the framework and primary concerns of every one of these calculations.

Existing Methods for Face Detection

***Robust Real-Time Object Detection*:** By Paul Viola and Michael J. Jones [8-12] . This is by all accounts the primary article where Viola-Jones presents the lucid arrangement of thoughts that comprise the basics of their face location calculation. This calculation just discovers frontal upstanding countenances yet is in 2003 introduced in a variation that likewise recognizes profile and pivoted sees.

***Neural Network-Based Face Detection*:** An image pyramid is calculated to detect faces at multiple scales. In the pyramid, a fixed size sub-window is moved through each image. The substance of a sub-window is remedied for non-uniform lightning and exposed to histogram equalization. The prepared substance is taken care of to a few equal neural systems [10] that do the real face detection. The outputs are joined utilizing logical AND, along these lines lessening the quantity of bogus location. In its first structure this algorithm additionally just recognizes frontal upstanding faces.

A Statistical method for 3d Object Detection [11] applied to Faces and Cars: The fundamental mechanics of this calculation is additionally to ascertain a picture pyramid and sweep a fixed size sub-window through each layer of this pyramid. The substance of the sub-window is exposed to a wavelet investigation [13-20] and histograms are made for the distinctive wavelet coefficients. These coefficients are taken care of to diversely prepare equal locators that are delicate to different directions of the item. The direction of the article is dictated by the locator that yields the most noteworthy yield. Restricted to the essential Viola-Jones calculation and the calculation introduced by Rowley et al. this calculation likewise recognizes profile sees.

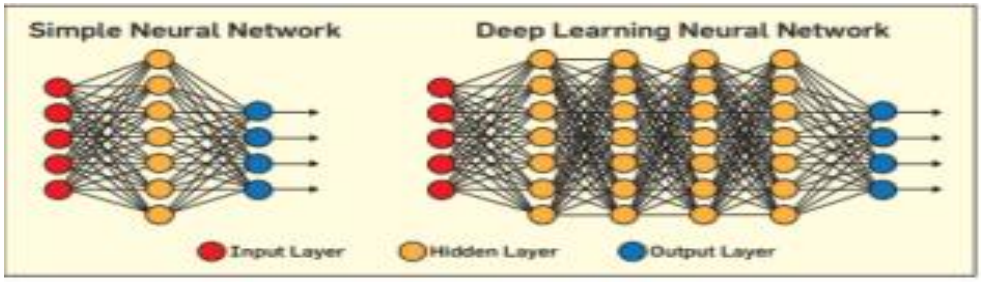
Existing Methods for Recognition: [21-22]

Introduction to Biometrics: Biometrics is a robotized technique for distinguishing an individual or confirming the personality of an individual dependent on a physiological or conduct trademark. Instances of physiological qualities incorporate hand or finger pictures, facial attributes. Biometric confirmation requires looking at an enlisted or selected biometric test (biometric format or identifier) against a recently caught biometric test (for instance, caught picture during a login). During Enrollment, as appeared in the image beneath, an example of the biometric quality is caught, handled by a PC, and store for later examination.

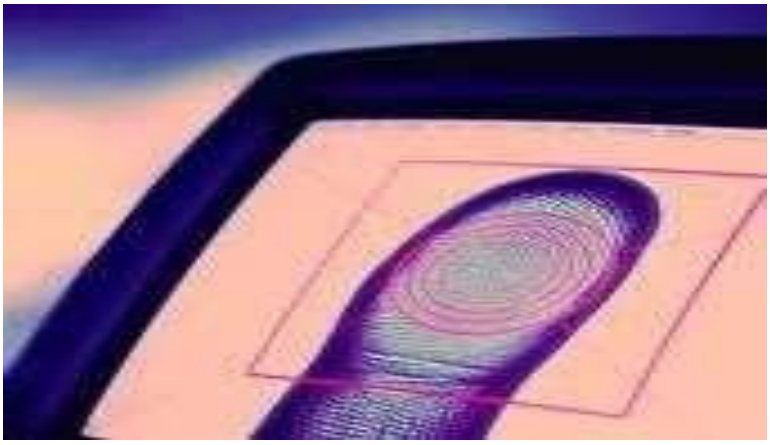
Biometric Authentication Technology: Biometric acknowledgment can be utilized in Identification mode, where the biometric framework recognizes an individual [23] from the whole selected populace via looking at a database for a match dependent on the biometric. Now and then distinguishing proof is called one-to-many coordinating. A framework can likewise be utilized in Verification mode, where the biometric framework [24-26] verifies an individual's asserted personality from their recently enlisted design. This is additionally called one-to-one coordination. In most PC access or system get to situations, confirmation mode would be utilized.

Types of Biometrics: There are various types of biometric techniques [27] that we observe in our daily life. Some of them are shown below

Face Recognition: The recognizable proof of an individual by their facial picture should be possible from various perspectives, for example, by catching a picture of the face in the obvious range utilizing a cheap camera or by utilizing the infrared examples of facial warmth emanation. Facial acknowledgment [28-31] in noticeable light commonly model key highlights from the focal bit of a facial picture. Utilizing a wide grouping of cameras, the obvious light frameworks remove highlights from the caught pictures that don't change after some time while evading shallow highlights, for example, outward appearances or hair. A few ways to deal with demonstrating facial pictures in the noticeable range are Principal Component Analysis, Local Feature Analysis, neural systems, versatile diagram hypothesis, and multi-goals investigation. Figure 1 speaks to the face acknowledgment framework for various preparing faces.



Fingerprint Recognition: Fingerprints are extraordinary for each finger of an individual including indistinguishable twins. One of the most economically accessible biometric advances, unique mark acknowledgment gadgets for work area, and PC get to are currently broadly accessible from a wide range of merchants effortlessly. With these device clients no longer need to type passwords – rather, just a touch gives a moment to get to. Unique mark frameworks can likewise be utilized in distinguishing proof mode. A few states check fingerprints for new candidates to social administrations' advantages to guarantee beneficiaries don't deceitfully get benefits under phony names. Figure 2 speaks to the unique finger impression acknowledgment framework.



Palm Recognition: Palm print acknowledgment is a biometric validation technique dependent on the interesting examples of different qualities in the palms of individuals' hands. Palm prints and fingerprints are regularly utilized together to upgrade the exactness of distinguishing proof. The structure of the palm appears in fig. 3.

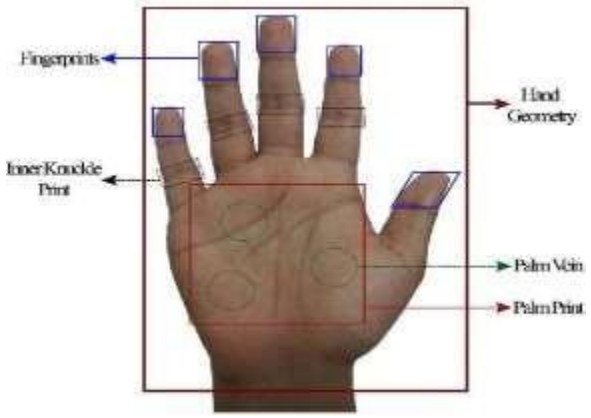


Figure 3: *Palm Recognition*

Gait Recognition: The Machine Visions way to deal with step acknowledgment involves the securing of stride signals utilizing at least one camcorders from a separation. In this manner, it requires a surrounding set-up. As a first regular advance, frameworks in this classification use strategy for video and image handling to distinguish the client's image in a scene, to follow the client's walk, and to separate step highlights for client acknowledgment Figure4 speaks to the style of stride acknowledgment.

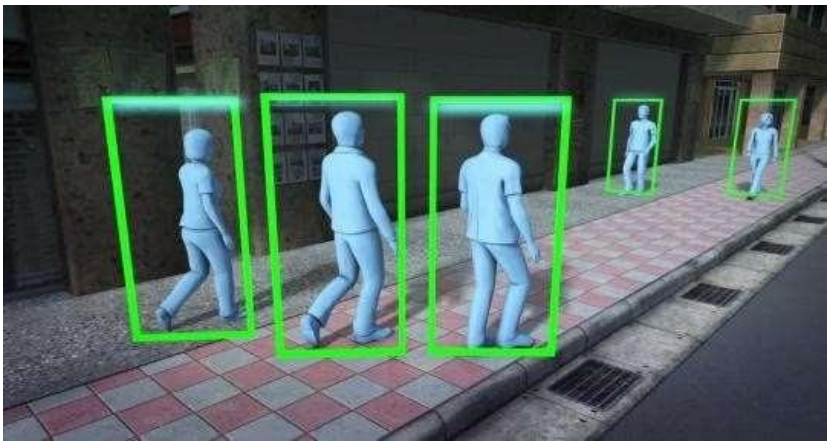


Figure 4: *Gait Recognition.*

Voice Recognition: Voice acknowledgment is the innovation by which sounds, words, or expressions verbally expressed by people are changed over into electrical signs, and these signs are changed into coding examples to which importance has been assigned‖. While the idea could all the more, for the most part, be called―sound recognition‖, we center here around the human voice since we frequently and most normally utilize our voices to convey our plans to others in our quick environmental factors. The trouble in utilizing voice as a contribution to a PC reproduction lies in the major contrasts between human discourse and the more conventional types of PC input. While PC programs are usually intended to deliver an exact and all around characterized reaction after accepting the correct (and similarly exact) input, the human voice and verbally expressed words are not exact. Every human voice is unique, and indistinguishable words can have various implications whenever spoken with various intonations or in various settings.

Iris Recognition: This acknowledgment strategy utilizes the iris of the eye, which is the shaded territory that encompasses the student. Iris designs are thought one of a kind. The iris designs are gotten through a video-based image procurement framework. Iris filtering gadgets have been utilized in close to home confirmation applications for quite a while. Frameworks dependent on iris acknowledgment have considerably diminished in cost and this pattern is relied upon to proceed. The innovation functions admirably in both confirmation and distinguishing proof modes. Figure 5 speaks to the state of the iris that appeared in the piece of eye.

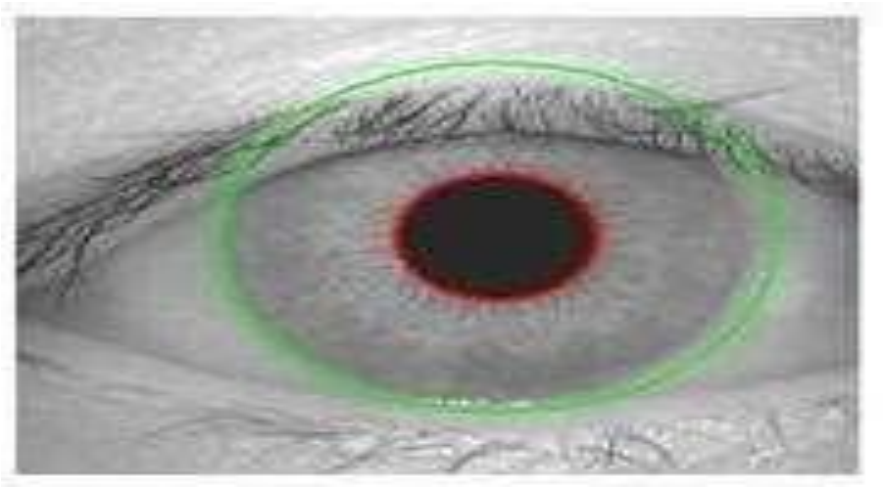


Figure 5: *Iris Recognition.*

DNA Recognition: Deoxyribonucleic acid (DNA) is the genetic material found in most organisms, including humans. DNA serves as a hereditary code that is one of a kind to each living being, no two being similar; just indistinguishable twins are an accurate DNA coordinate. On account of individuals, there are around 3 million bases, 99% of which are the equivalent from individual to individual. The varieties found in the last 1% are the way DNA gets exceptional to every person. The last 1%also fills in as the establishment for DNA biometrics

OpenCV*:* OpenCV has the upside of being a multi-stage structure; it bolsters the two Windows and Linux, and all the more as of late, Mac OS X. OpenCV has such huge numbers of capacities it can appear to be overpowering from the start. A decent comprehension of how these techniques work is the way to getting great outcomes when utilizing OpenCV. Luckily, just a chosen few should be known in advance to begin. OpenCV's usefulness that will be utilized for facial Acknowledgment is contained inside a few modules.

PROPOSED METHOD

Figure 6 speaks to the schematic of the proposed technique for face identification utilizing python programming. Face location and acknowledgment from an image or a video is a famous subject in biometrics examination. Face acknowledgment innovation has generally stood out because of its tremendous application worth and market potential. We structure a continuous face

acknowledgment framework dependent on IP camera and image set calculation by the method of OpenCV and Python programming improvement. The framework incorporates three sections: Detection module, preparing module, and acknowledgment module.

The distinguished facial image is then handled to address the direction and increment the difference, in this manner, it keeps up high facial acknowledgment precision. Huge databases with faces and non-faces images are utilized to prepare and approve face location and facial acknowledgment algorithms. The calculations accomplish a general genuine positive pace of 98.8% for face identification and 99.2% for right facial acknowledgment.

Image Acquisition

The image is acquired from a high definition camera and is processed.

Histogram Normalization

The caught images now and again have brilliance or murkiness in it which ought to be evacuated for acceptable outcomes. To begin with, the RGB image is changed over to the grayscale image for an upgrade.

Face Detection

An image of the separate individual is looked to discover a face, at that point the image is handled to harvest and concentrate the individual's face for simpler acknowledgment.

Face Recognition

The recognized and handled face is contrasted with a database of known appearances, to choose whom than an individual is.

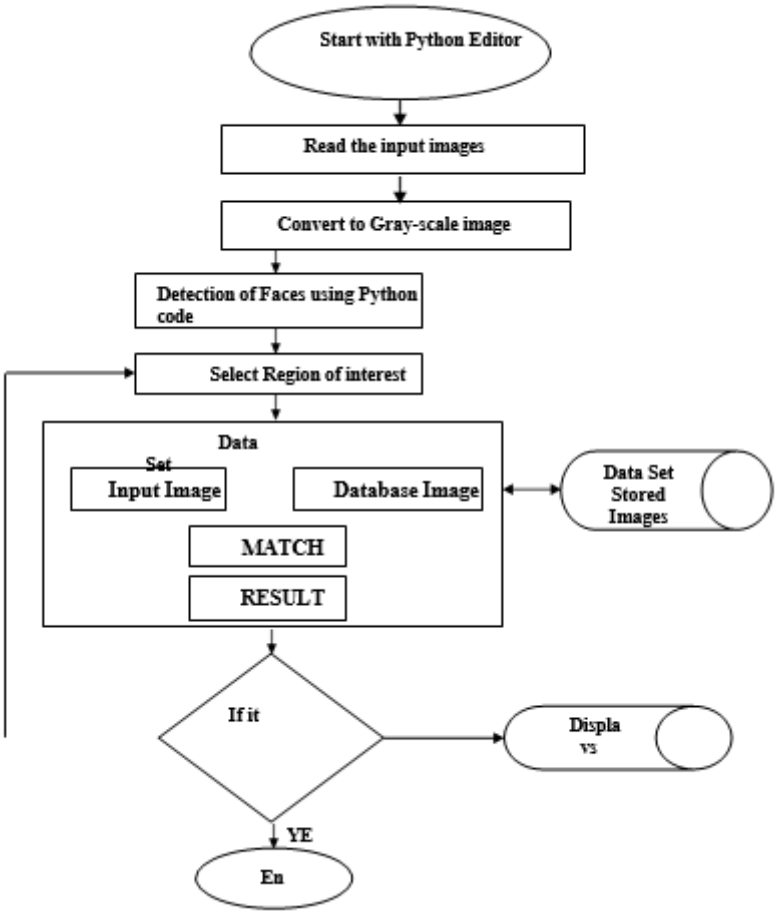
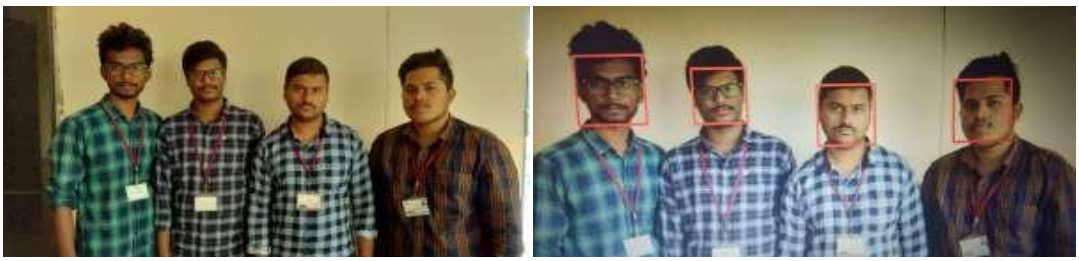


Figure 6: *Schematic Diagram of the Proposed Method.*

EXPERIMENTAL RESULTS

Figure 7 speaks to the trial consequences of the proposed strategy for face location framework utilizing python-based programming. The primary image appeared in Fig. 7a is an information image with the gathering of individuals. The subsequent image appeared in Fig. 7b is a face recognized image identifying all faces present in the gathering image. Figure 7d is the prepared images that are prepared and put away in a database for additional handling. Figure 7e speaks to the recognized individual countenances of individuals

and showed with name os the individual to which that face has a place.



(a) (b)

(c) (d)

(e)

Figure 7: *Output Images for Proposed Method a) Input image b) Face Detected image c) Grayscale Image d) Trained Images e) Output Images.*

CONCLUSION

There are several challenges in the cross spectral face recognition model as follows: Visible and near infrared spectrum has deferent• wavelength, visible spectrum has wavelength in between .4µm to .7µm and NIR spectrum has wavelength from.7µm to 1.4µm. Human face Images of same person in NIR spectrum and Visible spectrum are look very differently so even is very difficult task for human to• recognize these images.

Illumination changes and differences in facial expressions. Some practical challenges due to the oriental and• misalignment of the face in the different images. Illumination variances, facial expressions, surrounding environment and lighting condition also• effect the matching.

Utilizing Haar-falls for face recognition worked amazingly well in any event when subjects wore exhibitions. Ongoing video speed was agreeable also without observable casing slack. Thinking about all elements, LBPH joined with Haar-falls can be actualized as a cost effective face acknowledgment stage. A model is a framework to distinguish known troublemakers in a shopping center or a market to give the proprietor an admonition to keep him alert or for programmed participation taking in a class.

REFERENCES

1. J. Daugman (1999) Recognizing persons by their Iris patterns," in Biometrics: Personal Identification in a Networked Society, A. K. Jain, R. Bolle, and S. Pankanti, Eds. Norwell, M. A.: Kluwer, pp. 103–121.
2. M. P. Potadar, V. V. Marathe, A. S. Khose, and L. A. Kotkar (2015), "Biometric Attendance Recording and Communication System," *International Journal of Innovations in engineering and technology (IJIET)*, Vol. 5, No. 2, pp.230-234, Available at: http://ijiet.com/wp-content/uploads/2015/04/33.pdf.
3. C. Sanderson (2008) Biometric Person Recognition: Face, Speech and Fusion. VDM Publishing
4. M. B. Srinidhi and R. Roy (2015), "A Web-Enabled Secured System for Attendance Monitoring and Real-Time Location Tracking Using Biometric and Radio Frequency Identification (RFID) Technology," *IEEE International Conference on Computer Communication and Informatics (ICCCI), Coimbatore,* DOI: 10.1109/ICCCI.2015.7218103.
5. A. Jain, L. Hong, S. Pankanti, and R. Bolle (1997), "An Identity Authentication System Using Fingerprints," *Proceedings of the IEEE*, Vol. 85, No. 9, pp. 1365-1388, DOI: 10.1109/5.628674.
6. Syed Mohsin Saif et al . (205), "Effectiveness Of Gaussian And Average Noise Reduction Filters On Ideal Fingerprint Image In Biometric Fingerprint Identification System," *International Journal of Recent Scientific Research*, Vol. 6, No. 8, pp. 5731-5736, Available at: http://www.recentscientific.com/effectiveness-gaussian-and-average-noise-reduction-filters-ideal-fingerprint-image-biometric-fingerp.
7. S. Ouyang, T. Hospedales, Y. Song, and X. Li. A survey on heterogeneous face recognition: Sketch, infra-red, 3d and low-resolution. In arXiv preprint arXiv: 1409.5114, 2014
8. P. Viola, and M. J. Jones (2004), "Robust Real-Time Face Detection," *International Journal of Computer Vision*, Vol. 57, No. 2, pp. 137-154, Available at: https://link.springer.com/article/10.1023%2F B%3AVISI.0000013087.49260.fb.
9. S. Z. Li, and Z. Q. Zhang (2004), "Float Boost Learning and Statistical Face Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 9, pp. 1112-1123, Available at: http://www.cbsr.ia.ac.cn/Li%20Group/papers/ Li-PAMI-04.pdf.
10. Henry A. Rowley, Shumeet Baluja, Takeo Kanade, (2005) "Neural Network-Based Face Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 53, no. 2, pp. 105-115, DOI: 10.1109/34.655647.
11. J. Kittler, P. Koppen, et. al. (2018), "Conformal Mapping of a 3D Face Representation onto a 2D Image for CNN Based Face Recognition," *2018 International Conference on Biometrics (ICB)*, Gold Coast, QLD, pp. 124 -131. DOI: 10.1109/ICB2018.2018.00029.
12. T. Tirupal and B. Chandra Mohan (2012), "Pixel-Level Multifocus Image Fusion based on Wavelet Transform & Principal Component Analysis," *Journal of Innovation in Electronics & Communication*, Vol. 2, No. 2, pp. 60–64, Available at: https://scholar.google.co.in/citations?user=cQ HsneIAAAAJ&hl=en.
13. E. Jagadeesh and T. Tirupal (2015), "A Multimodal Image Fusion based on Dual-Tree Complex 2D-Discrete Wavelet Transform Method," *International Journal for Technological Research in Engineering*, Vol. 2, No. 12, pp. 3151-3156, Available at: https://www.researchgate.net/publication/304 024714_A_MULTIMODAL_IMAGE_FUSI
14. [Ale01] Alex Freitas, "A Survey of Evolutionary Algorithms for Data Mining and Knowledge Discovery", Postgraduate Program in Computer Science, Pontificia Universidade Catolica do Parana Rua Imaculada Conceicao, 1155. Curitiba - PR. 80215-901. Brazil, 2001.
15. [Dav01] David Hand, et al., "Principles of Data Mining", MIT Press 2001.