



Ear Recognition process using Hybrid System

Jeel Doshi¹, Priyanka Patel²

¹Smt. Kundaben Dinsha Patel Department of Information Technology, Chandubhai S. Patel Institute of Technology, Charotar University of Science and Technology, Changa-388421, Gujarat, India
17IT025@charusat.edu.in

²Smt. Kundaben Dinsha Patel Department of Information Technology, Chandubhai S. Patel Institute of Technology, Charotar University of Science and Technology, Changa-388421, Gujarat, India
priyankapatel.it@charusat.ac.in

ABSTRACT

The objective of this paper is to remember a half breed approach for the field of biometric [Jain et al. 2004] investigation, coordinating a portion of the more notable methodologies utilized by ears to distinguish individuals. This look into employments Hausdorff separation as a pre-preparing stage by acquainting strength with improve sifting proficiency for tests to be utilized in the test stage. Utilizations the Image Ray Transform (IRT) and the Hair Dependent Classifier for the Detection Process. Speeded Up Robust Features (SURF) and Linear Discriminant Analysis (LDA) are then assessed by the robot as contribution from two neural systems to recognize the human by their face profile. The above calculations have been applied utilizing Microsoft C # to exhibit the applicable exploratory theory. The consequences of the examination evil presence.

Key words:Hausdorff, LDA, IRT, Ear Recognition, SURF Algorithm

1. INTRODUCTION

The investigation of deciding individual recognizable proof dependent on the physical or social attributes of a man is alluded to as biometrics. People have been using body highlights, for example, the face and discourse for a huge number of years to perceive one another. In contemporary culture, there is a developing enthusiasm for structuring applications for PC acknowledgment that can be utilized for programmed human identification. With utilizes running from crime scene investigation to national security, biometrics are progressively turning into a fundamental piece of contemporary society. The most famous biometric gadgets are those centered around highlights ordinarily used by people for verification, for example, fingerprints and facial pictures, and have the most noteworthy market share. However, a few other individual highlights are currently being recognized as potential biometric markers for singular ID. The life systems of the ear is one of those biometrical pieces of information, as the structure and plan of the ear have been found to vary significantly between individuals[Jain et al. 2004] . This is a mainstream perceptible trademark when the picture is introduced in the image and is consequently promptly acquired from video accounts or photography.[1]

While the use of human ear acknowledgment subtleties has been built up, it is as yet indistinct whether the ear ought to be esteemed exceptional or explicit adequately to be utilized as a biometric. each physical or conduct quality, be that as it may, might be utilized as a biometric distinguishing gadget on the off chance that it is general, if any person contains an identifier, is unmistakable and extraordinary to every person, is invariant in time, and can be resolved naturally or physically; the ear satisfies every one of these characteristics.

2. BACKGROUND

2.1. Anatomy of ears and growth

Between the fifth and seventh long stretches of pregnancy the ear starts appearing. Now, the substance of the incipient organism is increasingly indicated when puncturing of the head, nostrils and ear spaces are clear. In spite of the fact that there is still discussion with the specific embryology of the external ear [ArbabZavar and Nixon 2011], the general advancement of the ear during pregnancy is as per the following:[2]

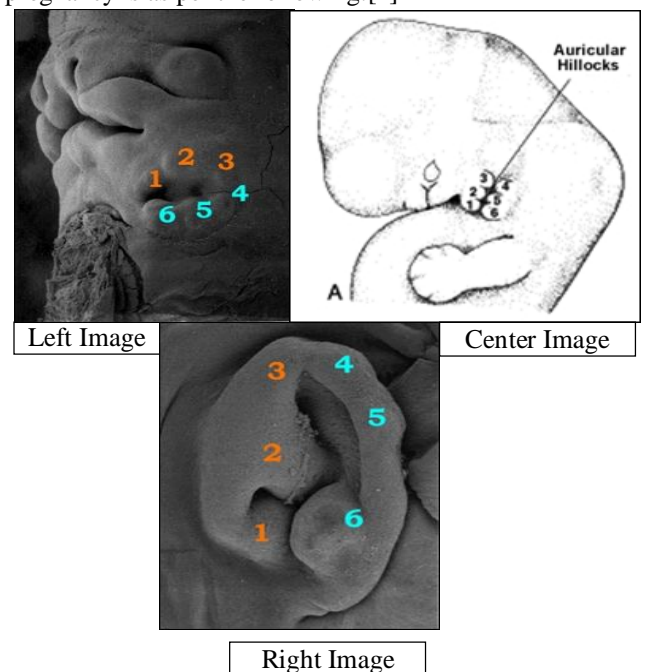


Figure 1: Initial development process of ear features

The human ear creates from the a trial hillocks (focus) that happen in the fifth seven day stretch of undeveloped turn of events. Between the sixth (left) and ninth (right) advancement weeks, the slopes augment, separate, and breaker to shape the external ear structures, Refer the figure: 1. Also, the ear transform the side of the neck to a progressively cranial and parallel site.

1. Initial incipient organism arrangement bunches of early stage cells that go about as the base on which a segment of the body or organ is made. Two of those groups, named the first and second pharyngeal curves, shape six tissue heights called auricular hillocks. Focal figure during the fifth seven day stretch of growth. Includes an undeveloped organism drawing of six auricular hillocks, numbered 1 through 6. The left figure exhibits the ascent and movement of the hillocks during the 6th seven day stretch of undeveloped turn of events.
2. Throughout the seventh week, the auricular hillocks start to grow, isolated and combine, making the last type of the head, which is gradually translocated from the side of the neck to a progressively cranial and horizontal area. All through the seventh week, the auricular hillocks start to grow, discrete and meld, making the last type of the head, which is gradually translocated from the side of the neck to a progressively cranial and parallel area.[2]

2.2. Ear Biometric System

The usage of different individual characteristics as a tool for personal identification includes the brain. As noted by Hurley *et al.*, ear biometry has received insufficient attention in the scientific community as a significant field of research, especially in comparison to the more traditional methods of recognizing and identifying head, hair or fingerprint. Nevertheless, ears have assumed a significant job in criminological science for a considerable length of time, especially in the United States, where Iannarelli (1989) built up an ear grouping plan dependent on manual estimations, which has been being used for a long time now shown in figure 2. The ear has one of a kind points of interest over the more conventional regions utilized in biometrics, particularly in that it has a rich and strong structure which doesn't modify radically as an individual age. Furthermore, while the face varies altogether relying upon voice, there is no issue with the ears.

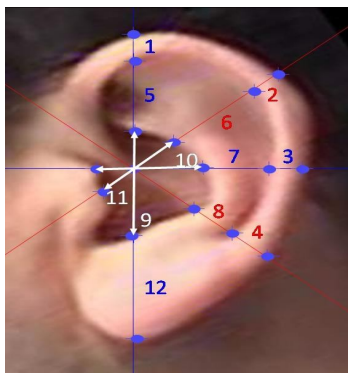


Figure 2: Iannarelli's Measurements

In differentiate, the ear's prompt importance is extremely clear (since it is constantly situated on the head), though facial acknowledgment for the most part includes a directed domain to perceive a turmoil that isn't regularly present, obviously. Like the catching of iris, hair, or unique mark (which is biometric contact), the ear may not require get close to nearness.[5]

2.3. Algorithm for whole Process

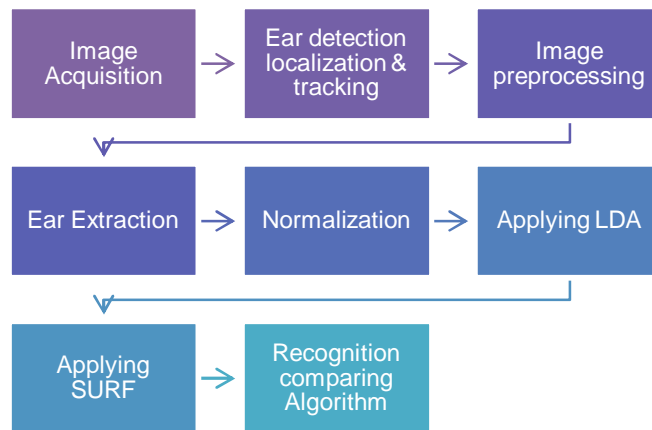


Figure 3: Pipeline Of the whole ear detection process

- Step-1: Image Acquisition
 - From Dataset [high resolution image]
 - From Web camera
 - From Video
- Step-2: extracting brightness component through color space transform
- Step-3: Localize the Ear
- Step-4: Resize the Localize area
- Step-5: Convert into binary using threshold function
- Step-6: Detect Edges.
- Step-7: Extract feature of each edges
- Step-8: Aggregate feature data in to dataset
- Step-9: Aggregated features will be used as an input to Neural Network
- Step-10: Activate each neurons
- Step-11: Continue this until the output layer is reached.
- Step-12: Compare the output to the target pattern and compute an error value.
- Step-13: Change all weights values of each weight using the formula:

$$W = \text{Weight}(\text{old}) + \text{LR} * \text{Output Error} * \text{O/P}(\text{Neuron } i) * \text{O/P}(\text{Neuron } i + 1) * (1 - \text{O/P}(\text{Neuron } i + 1))$$
- Step-14: Go to the 1st step.
- Step-15: The algorithm end, if we get all O/P pattern
- Step-16: Match target pattern.

2.4. Ear Recognition

Traditionally, recognition schemes obey a series of principles, such as the collection of images, pre-processing, Attribute collection and description and/or definition of the item in question. In the next pages, each of these tasks will be mentioned connecting important algorithms to achieve their goal. Nonetheless, it is necessary to remember that the procedure that we are going to discuss is focused on a variation of some of the current methods to create a reliable framework that enables real-time video recognition via the ear to be conducted, sensed, monitored and recognized.

The approach therefore incorporates a series of algorithms that produce reasonable outcomes on a stand-alone basis and reach a higher degree of robustness when coupled with a significant improvement in problems such as changes in picture brightness and perspective.

3. DETECTION AND PRE-PROCESSING

OpenCV and its.Net structure expansion OpenCV contains various article identifiers dependent on the Viola – Jones technique, a few of which are arranged to oblige various traits, for example, the front head, eyes, nose, and so forth. Modesto Castellón-Santana et al. built up a Haar-cascade classifier for use in OpenCV to recognize left and right hands. This classifier is the principal stage in the making of a solid ear recognizable proof and observing gadget.

The hair classifier is utilized to group the profiles of the neck, this catch is utilized to extricate the ear with a similar Haar strategy, if this technique cannot find the ear, and the gadget can gauge the IRT. The introductory photos from the document of the Ávila Police School produced for this investigation of 300 faces, 3 photographs for every point was acquired using the strategy proposed by Cummings et al.[9] where the gadget quantifies the IRT, Gaussian Smoothing is utilized to wipe out clamor and limit holes in the helix. Instead the model is changed in accordance with erase as a great part of the image as essential while endeavoring to keep the helix saved while making a curved plan to suit the model.

With the ear checked we start to play out the pre-handling step, convert the image to a dim scale and start the standardization time frame, at that point we play out the division of the image utilizing a veil to dispose of simply the ear, at that point the picture is changed over into an edge map utilizing the vigilant edge channel. On the off chance that w is the picture width in the pixel and h is the picture stature in the pixel, the shrewd edge finder can take as info the pair $w \times h$ the dim qualities and the sigma.

Information is a twofold picture with an estimation of 1 for edge pixels, for example a pixel that speaks to a limit and an estimation of 0 for every single other pixel. We measure a line between the principle and the minor y esteem in the edge picture to turn every pixel, endeavoring to situate the ear

lobule in the picture. This strategy is attempting to get all the photographs that have a structure near an obscure photo. On the off chance that the pre-preparing step is finished, we keep on figuring the Match utilizing the forms of the ear shape, where we are endeavoring to limit the quantity of contender for the recognizable proof stage by using as far as possible.

4. APPLICATION OF HAUSDORFF DISTANCE

The Hausdorff separation figuring used in this paper expands on the supposition that the ear locales are of various degrees of significance where attributes, for example, helix, antihelix, tragus, antitragus, concha, projection and ear shape assume a critical job in ear acknowledgment. The calculation applied depends on characterized in.

The investigation of edge maps works for all intents and purposes by including the Hausdorff width. The advantage of utilizing edges to fit two antiquities is that to change light, this portrayal is flexible. Represents a case of the Hausdorff separation endeavoring to arrange two pictures, the calculation attempts to quantify the separation between the articles, and with that separation we pick the kind of pictures in our bundle, this element goes about as a channel choosing and disposing of those pictures to help the characterization plot.

The strategy incorporates killing the image scenery as it was accomplished in the underlying pre-preparing, joining a few measures after pixel covering, removing the edges utilizing the Canny and Sobel channel, turning around the image to work with a white foundation, and binarizing the face, playing out a particular technique to every pixel contained in the archive. With the gathered things we contrast pixels with get how close the two figures are, as though they were conceptual figures doing a relative activity, estimating the Hausdorff separation, we contrast pixels with perceive how comparative the two figures are, coming full circle in a rundown of qualities speaking to the separation of the information picture from every component in the database.

The objective ought to be seen as an option with a littler relative separation; if the base edge esteem isn't reached and the customer is discovered, at that point the issue will be considered uncertain. In the built up approach, the Hausdorff calculation is actualized as an outside pre-handling instrument to expand the presentation of the neural system utilizing the SURF algorithm, Unless the framework methodology presume that the objective is the equivalent, it is acknowledged that the image has a place with the client input recognized by the three procedures joined.

5. FOLLOWING THE EAR

Speeded Up Robust Functionality (SURF) is a scale and revolution invariant finder and marker of focal points. The

choice of incredibly unmistakable and invariant trait focuses from the pictures was anticipated. One of the essential variables for utilizing SURF to characterize highlights is to examine how the particular highlights work in pictures and, simultaneously, to be increasingly reliable as far as change, considering heading, scale, brightening and impediment when contrasted with other invariant sort descriptors like SIFT and GLOH.

The suggestion for the SURF include vectors is that the relative estimated at the overall direction produces every vector speaking to an invariant concerning the turn of the picture. The way wherein the SURF procedure coordinating utilizations the nearest neighbor proportion coordinating to accomplish the best coordinating of the picture key-point inside another image is clarified by the recognizable proof of the nearest neighbor In the other key-purposes of the subsequent picture, where the nearest neighbor is characterized as the key-point with the least euclidean good ways from the key-point framed, their unidirectional mark networks are looked from the start. Since these SURF vectors are invariant to the turn of the sign, the ear identification framework that joins the past Viola – Jones approach with the SURF vectors is strong and powerful.

Between different invariant structure attributes in size and revolution, SURF gives momentous distinctive highlights while being vigorous to varieties in visual conditions, pivots and loads simultaneously. Utilizing a unidirectional descriptor, SURF indicates an image by first perceiving any elite characteristic focuses inside it and afterward portraying them.

6. FEATURE EXTRACTION AND RECOGNITION

This area subtleties the information extraction calculations and the parameter settings that the neural systems need to test ear acknowledgment.

7. PRINCIPLE COMPONENT ANALYSIS (PCA)

The ear acknowledgment Protocol with eigenears is characterized actually expressing that the preparation set's unique pictures are changed over into an assortment of eigenears E. For each picture, loads are then estimated on the (E) assortment, and afterward put away in the (W) collection. Observing an unclear picture X, loads for that particular picture are estimated and set in the WX framework. WX therefore comparative with picture weights. The strategy for arranging another ear in an alternate class (known ears) is the result of two stages. To begin with, the current picture is changed over into its own parts. The resultant loads structure the weight vector ΩT_{new} . [3]

$$\omega_k = u_k^T (\Gamma_{new} - \Psi) \quad k = 1, \dots, M'$$

$$\Omega_{new}^T = [\omega_1 \ \omega_2 \ \dots \ \omega_{M'}]$$

A measure of resemblance between the related images I and j is given by the euclidean distance between two vectors $d(\Omega_i, \Omega_j)$. If a certain threshold value approaches the difference between Γ_{new} and the rest of the photos on average, it may be concluded that Γ_{new} is not an identifiable ear. [8][9][10]

8. LINEAR DISCRIMINANT ANALYSIS (LDA)

LDA or a fisherman overcomes the disadvantages of the PCA method by introducing the dimensional differential criterion of a fisherman. The PCA algorithm is a linear combination of functions which maximizes information variance. It can contribute to bad images, particularly when working with picture noise, such as background, light and viewpoint shifts. The PCA would instead find the faulty parts to be graded. To prevent this issue, a fisher algorithm is used to evaluate the results of the ear recognition process. The fisher algorithm that we are implementing is exactly the same thing. [7]

We are building an picture matrix x with each column representing a picture. Each image is allocated to the member of the related data variable c . Project x in the $(N - c)$ dimensional subspace as P with the rotation matrix WP_{ca} defined by the PCA, where N is the number of samples in x . c is the specific number of classes ($\text{length}(\text{unique}(C))$) and we determine the scatter between classes of projection P as :

$$S_w = \sum_{i=1}^c \sum_{x_k \in X_i} (x_k - \text{mean}_i) * (x_k - \text{mean}_i)^T$$

Where the median is the overall value of P , the median is the sum of class I in P , and the value is the mean of class I in P .

$$S_b = \sum_{i=1}^c N_i * (\text{mean}_i - \text{mean}) * (\text{mean}_i - \text{mean})^T$$

Where x_i is a class i measure, x_k is a class x_i measure, mean is the class i mean in P . The conventional LDA is applied and the ratio of the determinant to class dispersion and unit dispersion is optimized. The solution is created by a series of generalized S_b and S_w W_{fld} downvectors which correspond to their own value. S_b 's rank is maximum $(c - 1)$, so there are just $(c - 1)$ non-zero own numbers, cut off the remainder. Finally we get the fisherears by $W = WP_{ca} * W_{fld}$.Such vectors are used as inputs for the creation of a neural network.

9. SURF ALGORITHM

Utilizing the SURF cycle, the ear picture is reproduced as an assortment of remarkable focuses, where each is combined with a vector descriptor. Each may have lengths of 64 or

128. The 128 dimensional framework of descriptors is viewed as the most testing data based structure, and is in every case best spoke to by portrayals of the most restricting highlights possible. An instrument for procuring a specific quality of a solitary individual is recommended by coordinating the attributes got by the distinctive preparing bodies.[4]

Table 1: Haar-cascade and adding SURF tracking

	#Attempts	Ear Localization Haar-Cascade (%)	SURF Tracking
2D images	308	92.53	98.70
Real time video	314	86.69	95.13

On the off chance that we have a person's ear picture for testing, A combined model is made by melding the element descriptor cluster of all the test pictures got, calling the excess descriptor exhibit just a single time. After all the photographs had been gotten, a rundown of labels showing that each picture and combination vector has been recently chosen was made. In the wake of ascertaining the Wave qualities and screening the photographs the Hausdorff scale is to utilize, the list incorporates the unidirectional mark ear lattices.[4]

Such vectors go about as system transmission inputs. In the preparation calculation, a positive return of 1 is viewed as the unidirectional network of qualities having a place with a person as the yield of neurons allotted to that client and 0 to different neurons.

In the wake of gathering the new picture, the element vectors are determined, for the obscure picture, We compute the present descriptors. Such descriptors are presented in the neural system, the yields of the individual neurons are determined and, if the joined yield recurrence approaches the predefined level, the individual with the dynamic record ear is affirmed to have a place with the neuron allocated ear.

The neural system parameters utilized in this technique are convoluted, the yield of the neurons relies upon the degree of the Hausdorff Distance Filter, where the calculation chooses numerous potential responses to the acknowledgment issue so as to decrease the quantity of competitors. The concealed layer is progressively made, gave that the entirety of shrouded neurons between the size of the info layer and the size of the yield layer ought to be 2/3 the size of the information layer in addition to the size of the yield layer; and not as much as double the size of the info layer dependent on the exploration of Jeff Heaton. than double the info layer size dependent on Jeff Heaton's work.

10. EXPERIMENTAL RESULTS

Inside this section the discoveries got during the distinguishing proof and acknowledgment process are talked about. Table 1 demonstrates the rates of exactness by utilizing only the Viola – Jones classifier utilized in OpenCV versus the capacity accomplished by including SURF checking highlights. This improvement can be exhibited in 2D photos on the grounds that when the system is performed on film, the impact isn't so clear.

When we think about the time length it works in trying to discover the article, the calculation joined with SURF observing turns out to be substantially more precise in light of the fact that these applications permit you to area the component despite the fact that it happens 180 degrees and doesn't happen through the hands.

Summarizing with viewpoint and lighting in normal natural surroundings, we accomplish 86 percent proficiency in PCA discovery, 87 percent with angling calculation, the use of the neural system with SURF descriptors, the rate expanded to 92 percent, with in excess of 300 individual test assortment. This is a typical method to include a note with exactness. The f esteem is named, so it employments: Certainly, we're

$$2 * \frac{(Precision * Recall)}{(Precision + Recall)}$$

going to guarantee from those f scores and we're going to state that SURF has the best f positioning (94 percent). LDA-NN has the second most noteworthy (84%) and PCA has the least (74%) and afterward we would choose SURF over the others on the off chance that we needed to make a beeline for the f segment.

Table 2: Table 2 Performance of conventional PCA vs LDA-NN and SURF-NN(%).

Training images	Testing images	PCA	LDA-NN	SURF-NN
20	79	72	80	81
30	70	76	82	83
50	86	77	87	83
80	103	82	87	88
100	148	83	88	92
120	185	85	90	93

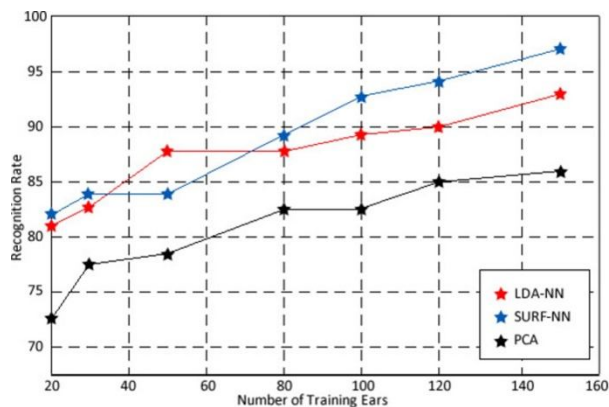


Figure 4: Recognition Rate vs Number of Training Ears

Table 3: In Normal Condition

	PCA		LDA-NN		SURF-NN	
	Positive	Negative	Positive	Negative	Positive	Negative
P	131	49	197	37	269	11
N	41	118	38	41	23	107

	PCA		LDA-NN		SURF-NN	
	Positive	Negative	Positive	Negative	Positive	Negative
P	164	69	169	58	178	27
N	89	128	49	111	32	137

Table 4: Changing illumination and perspective condition

In Table 2 the discoveries can be gotten with managed light under ordinary conditions. So as to check the legitimacy of our examination, we contrasted the outcomes got and our LDA-NN and SURF-NN here. For this situation, the outcomes are empowering, using SURF includes as input from a neural system with various test tests, we get a higher level of presentation on record than customary calculations. The relating confound lattices are appeared in Tables 3 and 4 refer figure-4 graph. [6]

The two calculations function admirably with frontal enlightenment, however the move in center has had a tremendous effect in results. The neural system utilizing the SURF descriptors as information was the calculation with less ear acknowledgment mistakes when we change the lighting and point of view. At the point when we alter the parameters, the blunder rate diminishes, leaving the PCA calculation with simply 67%, the SURF-NN with 86% and the LDA-NN with 76%.

11. CONCLUSION AND FUTURE WORK

The key consequence of this paper is the incorporation of two ear-acknowledgment calculations. The strategy utilized in this examination is to try to consolidate a portion of the more **typical techniques in the identification cycle; the venture isn't** seen as unique and excellent, yet centered around the methodologies proposed, ordered and differentiated by other researchers, So try to pick a blend of these procedures to successfully present a totally working system that perceives a human by means of their head. The main technique depends on pre-preparing with SURF followed by a Feed Forward Neural Network (SURF-NN) classifier and the second is centered around pre-handling with (LDA-NN). So try to pick a blend of these procedures to successfully present a totally working project that perceives a human through their head. The principal technique depends on preprocessing with SURF, trailed by a Feed Forward Neural Network (SURF-NN) classifier, and the second is centered on preprocessing with (LDA-NN).

The Neural system that gets contribution from SURF Descriptors gives off an impression of being more noteworthy than the difference in brightening. Changes in pre-handling forms are improving execution, particularly by utilizing Hausdorff Distance as a channel point. Tests found that with a fundamental 3-layer, about 95.03 percent of the exactness of ear recognizable proof is accomplished neural feed-forward system with back-proliferation preparing despite the fact that the photographs include some clamor.

Neural systems offer a more grounded substitution for some other customary classifiers. This product type is an adaptable device that can be educated to execute complex errands and capacities in PC vision frameworks, just as pre-preparing tasks, extraction usefulness and example acknowledgment. The most basic and sensible device for the police as future work is to accomplish the advancement of an application competent not just of proposing applicants from an ear picture yet in addition of distinguishing and perceiving a guilty party utilizing an earprint. While primer, the consequences of this investigation point to the objective, they recommend significant advancement towards the last point, affirmation fixated on these earprints.

REFERENCES

1. Maity, Sayan, and Mohamed Abdel-Mottaleb. "3D ear segmentation and classification through indexing." *IEEE Transactions on Information Forensics and Security* 10.2 (2014): 423-435.
2. Rask Andersen, Helge, et al. "Human cochlea: anatomical characteristics and their relevance for cochlear implantation." *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology* 295.11 (2012): 1791-1811. <https://doi.org/10.1002/ar.22599>
3. Skrobot, Vinicius L., et al. "Use of principal component analysis (PCA) and linear discriminant analysis (LDA)

- in gas chromatographic (GC) data in the investigation of gasoline adulteration." *Energy & Fuels* 21.6 (2007): 3394-3400.
4. Boulkenafet, Zinelabidine, JukkaKomulainen, and AbdenourHadid. "Face antispoofing using speeded-up robust features and fisher vector encoding." *IEEE Signal Processing Letters* 24.2 (2016): 141-145.
 5. Badrinath, G. S., and Phalguni Gupta. "Feature level fused ear biometric system." 2009 Seventh International Conference on Advances in Pattern Recognition. IEEE, 2009.
 6. El-Naggar, Susan, Ayman Abaza, and ThirimachosBourlai. "Ear Detection in the Wild Using Faster R-CNN Deep Learning." 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE, 2018.
 7. Fan, Zizhu, Yong Xu, and David Zhang. "Local linear discriminant analysis framework using sample neighbors." *IEEE Transactions on Neural Networks* 22.7 (2011): 1119-1132.
<https://doi.org/10.1109/TNN.2011.2152852>
 8. Patel, P., A Poster on Invariant Face Recognition(AIFR) Using Principal Component Analysis and Neural Network, poster, CSI, XXVII GSC, doi:10.13140/2.1.4613.2489, 2013
 9. Bandara, Ravimal, LochandakaRanathunga, and NorAniza Abdullah. "Nature inspired dimensional reduction technique for fast and invariant visual feature extraction." *arXiv preprint arXiv:1907.01102* (2019).
<https://doi.org/10.30534/ijatcse/2019/57832019>
 10. Patel, P., and A. Ganatra. "Investigate age invariant face recognition using PCA, LBP, Walsh Hadamard transform with neural network." *International Conference on Signal and Speech Processing (ICSSP-14)*. 2014.