
Citation:

Zarachoff, M and Sheikh Akbari, A and Monekosso, D (2021) Ear Recognition using Chainlet based Multi-Band SVM. In: IEEE International Conference on Imaging Systems and Techniques, 24 August 2021 - 26 August 2021, Virtual. (Unpublished)

Link to Leeds Beckett Repository record:

<https://eprints.leedsbeckett.ac.uk/id/eprint/7864/>

Document Version:

Conference or Workshop Item (Accepted Version)

© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

The aim of the Leeds Beckett Repository is to provide open access to our research, as required by funder policies and permitted by publishers and copyright law.

The Leeds Beckett repository holds a wide range of publications, each of which has been checked for copyright and the relevant embargo period has been applied by the Research Services team.

We operate on a standard take-down policy. If you are the author or publisher of an output and you would like it removed from the repository, please [contact us](#) and we will investigate on a case-by-case basis.

Each thesis in the repository has been cleared where necessary by the author for third party copyright. If you would like a thesis to be removed from the repository or believe there is an issue with copyright, please contact us on openaccess@leedsbeckett.ac.uk and we will investigate on a case-by-case basis.

Ear Recognition using Chainlet based Multi-Band SVM

1st Matthew Martin Zarachoff

*School of Built Environment,
Engineering and Computing
Leeds Beckett University
Leeds, United Kingdom*

m.zarachoff4868@student.leedsbeckett.ac.uk

2nd Akbar Sheikh-Akbari

*School of Built Environment,
Engineering and Computing
Leeds Beckett University
Leeds, United Kingdom*

a.sheikh-akbari@leedsbeckett.ac.uk

3rd Dorothy Monekosso

*School of Built Environment,
Engineering and Computing
Leeds Beckett University
Leeds, United Kingdom*

d.n.monekosso@leedsbeckett.ac.uk

Abstract—This paper presents a Chainlet based Multi-Band Ear Recognition using Support Vector Machine (CMBER-SVM) algorithm. The proposed method divides the gray input image into a number of bands based on the intensity of its pixels, resembling a hyperspectral image. It then applies Canny edge detection on each resulting normalized band, extracting edges that represent the ear pattern in each band. The resulting binary edge maps are then flattened, generating a single binary edge map. This edge map is then split into non-overlapping cells and the Freeman chain code for each group of connected edges within each cell is calculated. A histogram of each group of contiguous four cells is calculated, and the results histograms are then normalized and concatenated to form a chainlet for the input image. The resulting chainlet histogram vectors of the images of the dataset are then used for training and testing a pairwise Support Vector Machine (SVM). Experimental results on images of two benchmark ear image datasets show that the proposed CMBER-SVM technique outperforms both the state of the art statistical and learning based ear recognition methods.

Index Terms—ear recognition, chainlets, support vector machine, multi-band image generation

I. INTRODUCTION

Ear recognition, a field of biometrics wherein an ear image is used to identify an individual, has developed over the last two decades. Ears are unique to an individual; even identical twins can have different ear patterns [1]. In addition, ear recognition has several unique challenges when compared to face recognition. Ear images are smaller and generally contain more occlusion than images of the face due to the presence of jewelry and long hair. Furthermore, the number and size of ear image datasets is quite limited. In general, ear recognition algorithms are comprised of a feature extractor and a classification algorithm. Successful feature extraction techniques in ear recognition include Principal Component Analysis (PCA) [2]–[6], wavelet based [7], Curvelet based [8], local oriented patterns based [9] and neural network based methods [10]–[13]. Various statistical and machine learning based techniques for ear recognition have been proposed by researchers over the past two decades, including ‘Eigenfaces’ [6], wavelet [7], deep learning [12], and SVM [14], [15] based methods for feature extraction and classification. Although both statistical and learning based methods have been successfully applied to the ear recognition problem, the learning based algorithms

tend to produce the most accurate results. However, recent advances in statistical based techniques, such as 2D-MBPCA [16] and chainlets [17], have shown promising results. As many of the most successful ear recognition algorithms use a combination of statistically extracted features alongside learning based classification algorithms [2], the authors were inspired to explore a novel combination of multi-band image generation with chainlets and a learning based classifier.

In this paper, a Chainlet based Multi-Band Ear Recognition using Support Vector Machine (CMBER-SVM) is presented. The proposed method divides the input ear image into several image bands based on the image pixel intensity. This enables the extraction of more subtle ear edge features than is possible with the image as a whole. Edge selection is then performed on each resulting band using the Canny edge detection algorithm. Morphological operators are then applied to each resulting binary edge image band to suppress isolated edges, connect adjacent remaining edges, and remove extraneous edges. The resulting edge bands are then flattened into a single binary edge map. A window technique is then used to split the binary edge map into a number of cells, and the Freeman chain code for each edge group within each cell is calculated. The cells are then grouped into overlapping blocks and a histogram is then calculated from the chain codes for each block. These histograms are then normalized and concatenated to form the normalized chainlet histogram vector for the input image. A Pairwise Support Vector Machine is then trained and used to perform ear recognition. Experimental results, generated using images from the IITD II [18] and USTB I [19] datasets, demonstrate that the proposed CMBER-SVM technique outperforms both the statistical and state of the art learning based ear recognition algorithms.

II. PROPOSED TECHNIQUE

A block diagram of the proposed Chainlet based Multi-Band SVM (CMBER-SVM) ear recognition method is shown in Fig. 1. From this figure, it can be seen that the proposed technique includes five main steps: image pre-processing; multi-band image generation; edge selection; chainlet calculation; and classification.

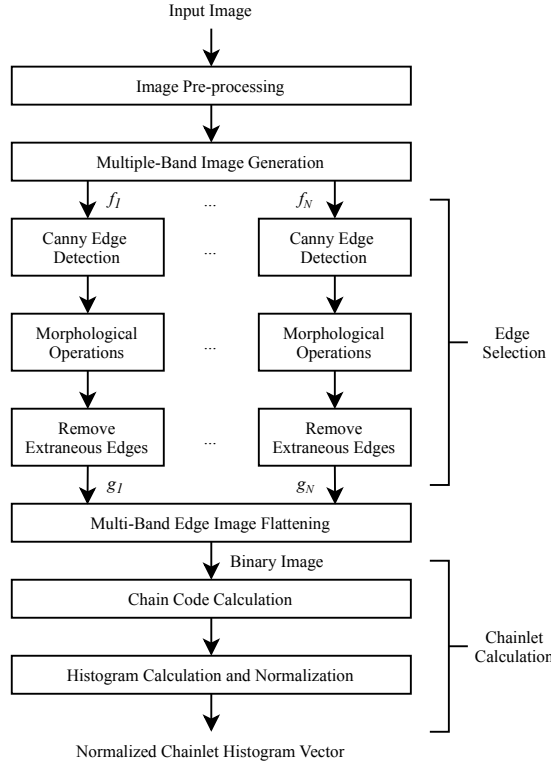


Fig. 1. Block diagram of the proposed Chainlet based Multi-Band SVM (CMBER-SVM) ear recognition technique.

A. Image Pre-processing

Let E be the set of all ear images within the image dataset. It is assumed that the input image $e \in E$ is an 8-bit, grayscale image. Histogram equalization is applied on the input image, increasing its contrast. To do so, the Probability Mass Function (PMF) P_X of the input image is first calculated:

$$P_X(x_k) = P(X = x_k) \text{ for } k = 0, 1, \dots, 255 \quad (1)$$

where x_0, x_1, \dots, x_{255} represent the pixel values and $P_X(x_k)$ is the probability of pixel value in bin k . The resulting PDF is then used to calculate the Cumulative Distribution Function (CDF) C_X of the image:

$$C_X(k) = P(X \leq x_k) \text{ for } k = 0, 1, \dots, 255 \quad (2)$$

where $C_X(k)$ is the cumulative probability of $X \leq x_k$. Finally, each pixel value within the image is mapped to a new value using its resulting CDF, generating a histogram equalized image.

B. Multi-Band Image Generation

The proposed CMBER-SVM method splits the input histogram equalized image into a number of bands based on its pixel values. Let N be the number of target bands for the input image e to be divided into. The pixel value boundaries $B = \{b_1, b_2, \dots, b_{N-1}\}$ are then calculated using (3):

$$b_n = n/N \text{ for } n = 1, 2, \dots, (N - 1) \quad (3)$$

The histogram equalized image has now been partitioned into F image bands, generating a multi-band image $F = \{f_1, f_2, \dots, f_N\}$.

C. Edge Selection

Input image edge selection procedure is as follows: Canny edge detection is first applied each resulting intensity band $f \in F$. The input intensity band is first smoothed using a Gaussian filter. Then, the intensity gradient of the resulting band is calculated in four directions (0° , 45° , 90° , and 135°) using a first order derivative function. The horizontal and vertical edge gradients are first calculated and used to determine the gradients of the diagonals. The non-maximum suppression method is then applied on the resulting gradients to preserve the edges with the largest gradients. The remaining edge pixels are then subjected to two empiric thresholds, low and high. Pixels below the low threshold are suppressed, pixels above the high threshold are classified as strong edges, and pixels between the two thresholds are considered to be weak edges. Finally, the resulting edges are subjected to edge tracking by hysteresis, wherein a weak-edge pixel is suppressed if none of its 8-connected neighborhood pixels are strong-edge pixels. This process produces a binary edge map $g \in G$ for each $f \in F$.

Each resulting binary edge image band $g \in G$ is then subjected to two morphological operations. First, isolated edges are identified by examining all single-pixel edges. If an isolated edge's 8-connected neighborhoods are zero, that edge is suppressed to zero. Second, a 'bridge' morphological operator is applied to the resulting edge band. If a zero value pixel has at least two non-zero neighbors, its value is set to one, thereby reducing the number of distinct contours in the binary image band. The resulting contours within the band more accurately represent the ear features of the original image.

One side-effect of performing edge detection on each binary image band g is that additional edges have been introduced by multi-band image generation, as pixels with values just above and below given boundary values are often adjacent. Consequently, those boundary pixels are incorrectly classified as edges. To remedy this, the boundary pixels' values are compared with their 8-connected neighborhood. If any of the neighbors has a value of zero, that pixel's corresponding edge in the map is assumed to be induced by multi-band image generation and is then discarded. After all superfluous edges have been removed, the binary image bands $g \in G$ are combined, generating the final binary edge map.

D. Chainlet Calculation

A block diagram of the chainlet calculation process can be seen in Fig. 2. Chainlets are based on the Freeman chain code of eight directions, where the chain code is commonly used to generate a vector representing the edge contour. The direction from an edge pixel, represented by a one in the binary edge

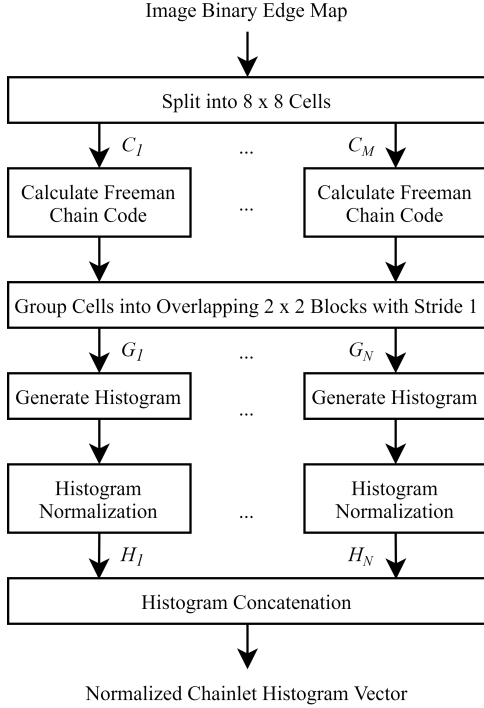
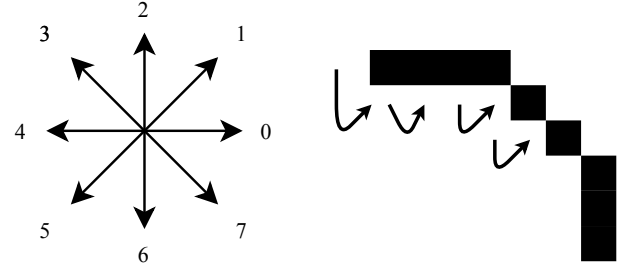


Fig. 2. Block diagram of the chainlet calculation process within the proposed CMBER-SVM ear recognition technique.

map, to each of its potential eight edge neighbors is assigned a value between zero and seven as shown in Fig. 3a. To calculate the chain codes, the resulting flattened input binary edge map is divided into non-overlapping cells of size 8×8 pixels. For each edge contour within each cell, the Freeman chain code is computed beginning from that edge-contour's upper leftmost pixel and traversing counter-clockwise as shown in Fig. 3b. For this particular edge-contour, the chain code is [0 0 0 7 7 7 6 6 2 2 3 3 3 4 4]. The resulting cell chain codes are then grouped into overlapping blocks of size 2×2 cells with a stride of 1 cell. A histogram is then generated for each block's resulting chain codes, which is then normalized using the L2 norm. The resulting normalized histograms of all blocks are then concatenated row by row to create a normalized chainlet histogram vector.

III. PAIRWISE SUPPORT VECTOR MACHINE

To perform classification, a pairwise Support Vector Machine (SVM) is employed in this research. Unlike standard SVM, which classifies new inputs to a finite number of pre-trained classes, Pairwise SVM takes two inputs and determines if they belong to the same class, making it ideal for pairwise matching. Let $R = \{H_1^1, \dots, H_i^j, \dots, H_N^M\}$ be a training set of chainlets where H_i^j corresponds to the j -th training image of the i -th individual. The pairwise decision function between H_i^j and the inquiry chainlet H_p^q can then be expressed as:



(a) Freeman Chain Code (b) Edge Traversal

Fig. 3. The Freeman Chain Code of eight directions and a traversed edge, resulting in the chain code [0 0 0 7 7 7 6 6 2 2 3 3 3 4 4].

$$D(H_i^j, H_p^q) = \sum_{(m,n)} \alpha_{mn} y_{mn} K((H_m, H_n), (H_i^j, H_p^q)) + \gamma \quad (4)$$

where α are the learned weights, K is the kernel function, γ is the learned bias, and:

$$y_{mn} = \begin{cases} +1, m = n \\ -1, m \neq n \end{cases} \quad (5)$$

In this work, the kernel K is the direct sum pairwise kernel, i.e.:

$$K((a,b), (c,d)) := k(a,c) + k(b,d) \quad (6)$$

where k is a standard kernel; in this case, the linear kernel:

$$k(a,c) = a^T c \quad (7)$$

In this research, the parameters alpha and gamma were learned as detailed in [20]. Furthermore, for each ear image dataset, the first two images of each individual were used for training and the remaining images were used for testing.

IV. EXPERIMENTAL RESULTS

Experimental results were generated using two benchmark ear image datasets called the Indian Institute of Technology Delhi II (IITD II) [18] and the University of Science and Technology Beijing I (USTB I) [19], which are widely used in the literature [7], [14]–[16], [21]. These two datasets were selected due to their widespread use in recent publications and also because their images have been pre-aligned. The IITD II dataset consists of 793 images of the right ear of 221 participants. Each participant was photographed between three and six times, where the images are 8-bit grayscale of size 180×50 pixels. The images of IITD II dataset are tightly cropped, of equal size, and are manually centered and aligned. The USTB I dataset consists of 180 images of the right ear of 60 participants, each of whom were photographed three times. The images in this dataset are 8-bit grayscale of size 150×80 . The images in USTB I are tightly cropped; however, they

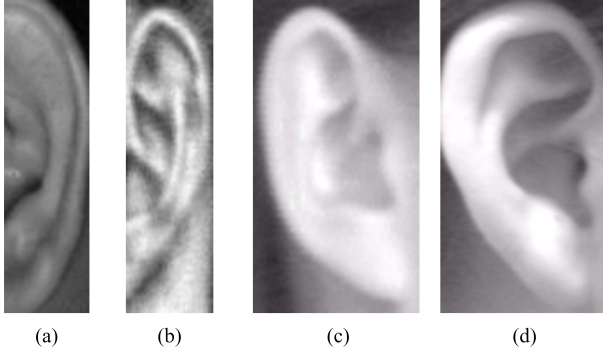


Fig. 4. Sample images of two unique individuals from the IITD II dataset (a-b) [18]. Sample images of two unique individuals from the USTB I dataset (c-d) [19].

exhibit some slight rotation and shearing. Examples from both datasets can be seen in Fig. 4.

The proposed CMBER-SVM method was applied to both images of the IITD II and USTB I datasets using two to ten bands of constant size as detailed in Section II-B. The number of correct matches was calculated for each set of bands. A subset of the results for both the IITD II and USTB I image datasets are tabulated in Table I and Table II respectively. From these tables, it can be seen that the proposed technique achieves its best performance at four and seven bands when applied to the images of the IITD II and USTB I datasets, respectively. From these tables, it is evident that the proposed CMBER-SVM technique achieves slightly higher performance when applied to the images of the IITD II dataset rather than those of the USTB I dataset.

TABLE I
EXPERIMENTAL RESULTS FOR THE PROPOSED CHAINLET BASED MULTI-BAND EAR RECOGNITION USING SUPPORT VECTOR MACHINE (CMBER-SVM) EAR RECOGNITION METHOD ON THE IITD II [18] DATASET.

Number of Bands	Rank-1	Rank-5
2	97.79	99.85
3	98.44	100.00
4	99.02	100.00
5	98.25	100.00
6	96.94	99.63

TABLE II
EXPERIMENTAL RESULTS FOR THE PROPOSED CHAINLET BASED MULTI-BAND EAR RECOGNITION USING SUPPORT VECTOR MACHINE (CMBER-SVM) EAR RECOGNITION METHOD ON THE USTB I [19] DATASET.

Number of Bands	Rank-1	Rank-5
5	99.17	100.00
6	99.17	100.00
7	99.44	100.00
8	99.44	100.00
9	98.89	100.00

To assess the performance of the proposed CMBER-SVM

technique, it, single image PCA, the ‘eigenfaces’ technique [4], 2D-MBPCA [16], ‘BSIF and SVM’ [15], and ‘Neural Network and SVM’ [14] were applied to the images of the IITD II and USTB I datasets. In single image PCA, the PCA algorithm is applied directly to the images of each dataset and the Euclidean distance between the resulting eigenvectors is used as a matching criteria. The ‘eigenfaces’ technique, first introduced by [6] and applied to ears in [4], first uses a small training set to create ‘eigeneers’. Ear images are then projected along the eigeneers, with the resulting vectors compared using Euclidean distance. 2D-MBPCA [16] applies a multi-banding technique to the input image and generates eigenvectors by applying PCA to each resulting multi-band image, with these also being compared by Euclidean distance. The two learning techniques both utilize SVM as their classifier, with [15] first extracting features using BSIF, while [14] uses a neural network to extract its features. The Rank-1 results for these techniques are tabulated in Table III. From Table III, it can be seen that the proposed CMBER-SVM method significantly outperforms both the PCA based and learning based state of the art methods for the images of both the IITD II and USTB I datasets.

TABLE III
EXPERIMENTAL RESULTS FOR THE PROPOSED CHAINLET BASED MULTI-BAND EAR RECOGNITION USING SUPPORT VECTOR MACHINE (CMBER-SVM) EAR RECOGNITION METHOD ON THE USTB I [19] DATASET.

Algorithm	Dataset	
	IITD II	USTB I
<i>PCA based Techniques</i>		
Single Image PCA	36.35	45.00
Eigenfaces [4]	89.78	75.93
2D-MBPCA [16]	91.12	85.19
<i>Learning based Techniques</i>		
BSIF and SVM [15]	97.31	-
Neural Network and SVM [14]	-	98.30
Proposed CMBER-SVM Technique	99.02	99.44

V. CONCLUSIONS AND FUTURE WORK

In this paper, a Chainlet based Multi-Band Support Vector Machine (CMBER-SVM) ear recognition method was presented. The proposed method divides the input ear image into a number of bands based on the intensity of its pixels. Edge selection is then performed on each resulting band and morphological operators are used to enhance the resulting binary edge maps. These edge maps are then flattened to generate a single binary edge map image. This edge map is then split into cells and the Freeman chain code for each cell is calculated. Cells are then grouped into overlapping blocks, and a histogram for each block is generated. The resulting histograms are normalized and concatenated to produce a normalized chainlet histogram vector for the input image. The normalized chainlet histogram vectors for different images are finally used as features for matching using pairwise SVM. Experimental results on images of two benchmark ear image datasets show that the proposed CMBER-SVM technique

outperforms the state of the art statistical and learning based ear recognition methods. An extended version of the proposed CMBER-SVM technique has great potential to be useful in other biometric fields such as face recognition. Such extensions would be applicable in the fields of security surveillance.

ACKNOWLEDGMENT

Matthew Martin Zarachoff thanks Leeds Beckett University for their support through a fully-funded studentship. This research has been funded under a knowledge transfer partnership by Innovate UK (KTP 10304).

REFERENCES

- [1] H. Nejati, L. Zhang, T. Sim, E. Martinez-Marroquin, and G. Dong, "Wonder ears: Identification of identical twins from ear images," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, Nov. 2012, pp. 1201–1204.
- [2] v. Emeršič, V. Štruc, and P. Peer, "Ear recognition: More than a survey," *Neurocomputing*, vol. 255, pp. 26–39, Sep. 2017. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S092523121730543X>
- [3] B. Victor, K. Bowyer, and S. Sarkar, "An evaluation of face and ear biometrics," in *Object recognition supported by user interaction for service robots*, vol. 1, 2002, pp. 429–432 vol.1.
- [4] K. Chang, K. W. Bowyer, S. Sarkar, and B. Victor, "Comparison and combination of ear and face images in appearance-based biometrics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1160–1165, Sep. 2003.
- [5] D. Querencias-Uceta, B. Ríos-Sánchez, and C. Sánchez-Ávila, "Principal component analysis for ear-based biometric verification," in *2017 International Carnahan Conference on Security Technology (ICCST)*, Oct. 2017, pp. 1–6.
- [6] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Proceedings*, Jun. 1991, pp. 586–591.
- [7] M. S. Nosrati, K. Faez, and F. Faradji, "Using 2D wavelet and principal component analysis for personal identification based On 2D ear structure," in *2007 International Conference on Intelligent and Advanced Systems*, Nov. 2007, pp. 616–620.
- [8] A. Basit and M. Shoaib, "A human ear recognition method using nonlinear curvelet feature subspace," *International Journal of Computer Mathematics*, vol. 91, no. 3, pp. 616–624, Mar. 2014. [Online]. Available: <https://doi.org/10.1080/00207160.2013.800194>
- [9] M. Hassaballah, H. A. Alshazly, and A. A. Ali, "Robust local oriented patterns for ear recognition," *Multimedia Tools and Applications*, vol. 79, no. 41, pp. 31 183–31 204, Nov. 2020. [Online]. Available: <https://doi.org/10.1007/s11042-020-09456-7>
- [10] P. L. Galdámez, A. G. Arrieta, and M. R. Ramón, "Ear recognition using a hybrid approach based on neural networks," in *17th International Conference on Information Fusion (FUSION)*, Jul. 2014, pp. 1–6.
- [11] F. I. Eyiokur, D. Yaman, and H. K. Ekenel, "Domain adaptation for ear recognition using deep convolutional neural networks," *IET Biometrics*, vol. 7, no. 3, pp. 199–206, 2018.
- [12] S. Dodge, J. Mounsef, and L. Karam, "Unconstrained ear recognition using deep neural networks," *IET Biometrics*, vol. 7, no. 3, pp. 207–214, 2018.
- [13] H. Alshazly, C. Linse, E. Barth, and T. Martinetz, "Deep Convolutional Neural Networks for Unconstrained Ear Recognition," *IEEE Access*, vol. 8, pp. 170 295–170 310, 2020, conference Name: IEEE Access.
- [14] I. Omara, X. Wu, H. Zhang, Y. Du, and W. Zuo, "Learning pairwise SVM on deep features for ear recognition," in *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, May 2017, pp. 341–346.
- [15] A. Benzaoui, N. Hezil, and A. Boukrouche, "Identity recognition based on the external shape of the human ear," in *2015 International Conference on Applied Research in Computer Science and Engineering (ICAR)*, Oct. 2015, pp. 1–5.
- [16] M. Zarachoff, A. Sheikh-Akbari, and D. Monekosso, "2D Multi-Band PCA and its Application for Ear Recognition," in *2018 IEEE International Conference on Imaging Systems and Techniques (IST)*, Oct. 2018, pp. 1–5.
- [17] A. Ahmad, D. Lemmond, and T. E. Boulton, "Chainlets: A New Descriptor for Detection and Recognition," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Mar. 2018, pp. 1897–1906.
- [18] "IIT Delhi Ear Database," accessed: 2018-03-13. [Online]. Available: <https://bit.ly/3rAPbWE>
- [19] "Ear Recognition Laboratory at USTB," accessed: 2018-06-25. [Online]. Available: <http://www1.ustb.edu.cn/resb/en/index.htm>
- [20] C. Brunner, A. Fischer, K. Luig, and T. Thies, "Pairwise Support Vector Machines and their Application to Large Scale Problems," p. 14.
- [21] A. Benzaoui and A. Boukrouche, "Ear recognition using local color texture descriptors from one sample image per person," in *2017 4th International Conference on Control, Decision and Information Technologies (CoDIT)*, Apr. 2017, pp. 0827–0832.