MDPI

Article

Chainlet based Ear Recognition using Image Multi-Banding and Support Vector Machine

Matthew Martin Zarachoff [1,*], Akbar Sheikh-Akbari [2,*] and Dorothy Monekosso [3,*]

[1] School of Built Environment, Engineering and Computing, Leeds Beckett University, Leeds, LS6 3QR, United Kingdom;
[2] School of Built Environment, Engineering and Computing, Leeds Beckett University, Leeds, LS6 3QR, United Kingdom;
[3] Department of Computer Science, Durham University, Durham, DH1 3LE, United Kingdom
* Matthew Martin Zarachoff: M.Zarachoff4868@student.leedsbeckett.ac.uk,
Akbar Sheikh-Akbari: A.Sheikh-Akbari@leedsbeckett.ac.uk,
Dorothy Monekosso: Dorothy.Monekosso@durham.ac.uk

Abstract: This paper presents a Chainlet based Ear Recognition algorithm using Multi-Banding and Support Vector Machine (CERMB-SVM). The proposed method splits the gray input image into several bands based on the intensity of its pixels, like a hyperspectral image. It performs 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, creating a single binary edge map. This edge map is then divided into non-overlapping cells and the Freeman chain code for each group of connected edges within each cell is determined. A histogram of each group of contiguous four cells is computed, and the generated histograms are 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 obtained using the two benchmark ear image datasets demonstrates that the proposed CERMB-SVM method generates significantly higher performance in terms of accuracy than the Principal Component Analysis based techniques. Furthermore, the proposed CERMB-SVM method yields greater performance in comparison to its anchor chainlet technique and state of the art learning-based ear recognition algorithms.

Keywords: ear recognition, chainlets, support vector machine, multi-band image generation

Citation: Zarachoff, M.N.; Sheikh-Akbari, A.; Monekosso, D. Chainlet based Ear Recognition using Image Multi-Banding and Support Vector Machine. *Appl. Sci.* **2022**, *1*, 0. https://doi.org/

Received:
Accepted:
Published:

1. Introduction

Ear recognition, a field of biometrics wherein an ear image is used to identify an individual, has advanced over the last two decades. Ears are unique to an individual; even indistinguishable twins can have different ear patterns [1]. There are several challenges attached with ear recognition in comparison to face recognition. More obstruction can be found when dealing with ear images than face images due to coverings that are sometimes present in such images, e.g., hair and jewelry. Additionally, there is currently a limited number of ear image datasets available. These datasets usually contain a smaller number of images. A typical ear recognition technique consist of a feature extractor and a classification method. Some of existing feature extraction algorithms for ear recognition are Principal Component Analysis (PCA) [2–6], wavelet based [7], Curvelet based [8], local oriented patterns based [9] and neural network based methods [10–13] Over the years, researchers have proposed several machine learning based and statistical methods for ear recognition. Some of these methods include: 'Eigenfaces' [6], wavelet [7], deep learning [12], and SVM [14,15] based methods for feature extraction and classification. Both learning and statistical-based algorithms have been successfully used for ear recognition. However, more accurate results are often obtained using the learning-based techniques due to the 'width' of the data. However, promising results have been noticed with recently reported statistical based algorithms, e.g., 2D-MBPCA [16] and chainlets [17]. Most successful ear recognition techniques use a combination of statistical based feature extraction method along with a learning-based classification algorithm[2]. This has inspired the authors to investigate a new combination of multi-band image processing with chainlets and a learning-based classifier. Application of the multi-band image processing for ear recognition on non-decimated wavelet subbands of ear images using Principal Component Analysis (PCA) shows the effectiveness of multi-band image processing in recognition. In [18], authors showed that the intersection of the Eigenvector energy and number of features graphs define the optimum number of bands for recognition, where increasing the number of multi-band images changes the distribution of the energy across image Eigenvectors, consolidates most of the image Eigenvec energy into a smaller number of Eigenvectors. The result of this was an increased accuracy in recognition. In [17], authors introduced Chainlets as an efficient feature descriptor for encoding the shapes formed by the edges of an object, where the connections and orientations of the edges are more invariant to translation and rotation. They have successfully applied their method to ear recognition and reported promising results. However, to the au knowledge, the application of multi-band image processing along with Chainlets for ear recognition has not been reported in the literature. This has inspired the authors to investigate a new combination of multi-band image processing with Chainlets and a learning-based classifier for ear recognition.

This paper presents a Chainlet based Multi-Band Ear Recognition method using Support Vector Machine (CERMB-SVM). The proposed algorithm splits the input ear image into a few image bands based on the image pixel intensity. More ear features can be extracted using the created image bands rather than just the input image. The Canny edge detection is applied to each resulting image band, generating a binary image representing its edges. To suppress isolated edges, connect adjacent remaining edges, and discard inappropriate edges within each resulting binary image, morphological operators are employed. The resulting edge bands are then flattened into a single binary edge map. The generated binary edge map is divided into several cells using a windowing algorithm and the Freeman chain code for each edge group within each cell is determined. The cells are then categorized into overlapping blocks and a histogram is then computed 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. The IITD II [19] and USTB I [20] ear image datasets were used to generate experimental results. Results show that the proposed CERMB-SVM technique surpasses both the statistical and state of the art learning based ear recognition techniques. The rest of the paper is organized as follows: section 2 introduces the proposed CERMB-SVM algorithm, section 3 discussed the pairwise Support Vector Machine classifier, section 4 presents the experimental results, and 5 concludes the paper.

2. Proposed CERMB-SVM Technique

A block diagram of the proposed Chainlet based Ear Recognition algorithm using Multi-Banding and Support Vector Machine (CERMB-SVM) is shown in Fig. 1. This figure shows that the proposed algorithm contains five main stages: image pre-processing; multi-band image generation; binary edge image creation; chainlet calculation; and classification.

2.1. 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 unsigned 8-bit, grayscale image. The proposed technique performs

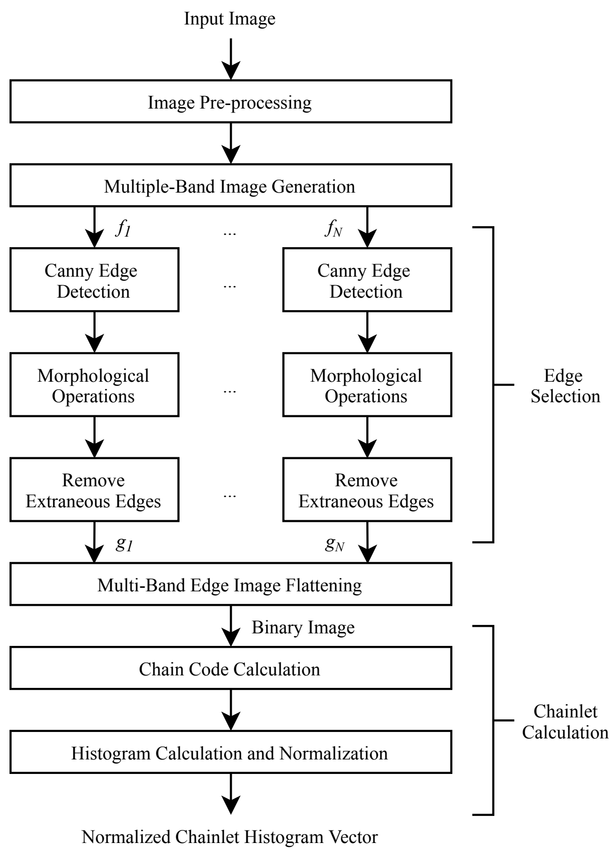


Figure 1. Block diagram of the proposed Chainlet based Multi-Band SVM (CERMB-SVM) ear recognition technique.

a histogram equalization on the input image to increasing its contrast. This is done by computing the Probability Mass Function (PMF) P_X of the input image.

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

where $x_0, x_1, ..., x_{255}$ is the pixel values and $P_X(x_k)$ indicates the probability of pixel value in bin k. The Cumulative Distribution Function (CDF) C_X of the image is then computed using the calculated PDF:

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

where $C_X(k)$ indicates 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, creating a histogram equalized image.

2.2. Multi-Band Image Generation

The proposed CERMB-SVM algorithm divides the resulting histogram equalized image into several bands based on its pixel values. Let N be the number of target bands for the input image e to be split into. The pixel value boundaries $B = \{b_1, b_2, ..., b_{N-1}\}$ are then determined using (3):

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

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

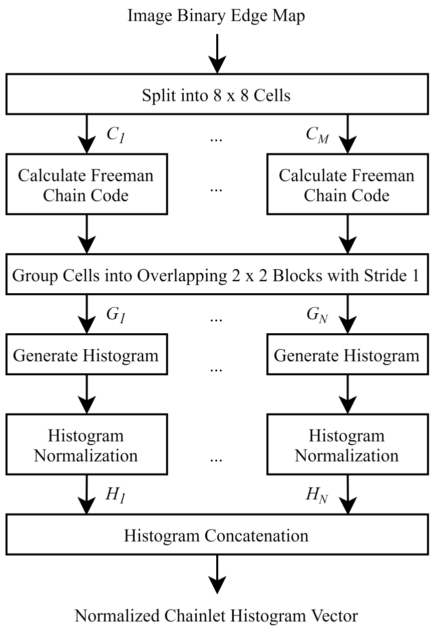


Figure 2. Block diagram of the chainlet calculation process.

2.3. Edge Selection

Input image edge selection process is as follows: the proposed algorithm first applies Canny edge detection to each resulting intensity band $f \in F$. A Gaussian filter with sigma 0.5 is applied on each resulting intensity band to smoothing the band. Then, the intensity gradient of the resulting band is computed in four directions ($0°$, $45°$, $90°$, and $135°$) using a first order derivative function. The horizontal and vertical edge gradients are first determined and used to calculate the gradients of the diagonals. The non-maximum suppression algorithm 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 discarded, 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 discarded if none of its 8-connected neighborhood pixels are strong-edge pixels. The resulting edge maps are then binarized, generating a binary edge map $g \in G$ for each band.

Each resulting binary edge map $g \in G$ is then subjected to two morphological operations. First, isolated edges are suppressed to zero. Second, a 'bridge' morphological operator is performed on the resulting edge map. If a zero-value pixel has at least two non-zero neighbors, its value is set to one, thereby decreasing the number of distinct contours in the binary image map. The resulting contours within the maps more accurately demonstrate the ear features of the original image.

One side-effect of performing edge detection on each band 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 inaccurately classified as edges. To solve this problem, these pixels are compared with their 8-connected neighborhood. If any of the neighbors has a value of zero, the edge is assumed to be induced by multi-band image generation and is then removed. After all redundant edges have been removed, the binary image bands $g \in G$ are combined, creating the final binary edge image.

(a) Freeman Chain Code (b) Edge Traversion

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

2.4. Chainlet Calculation

A block diagram of the chainlet computation procedure is shown 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 map, to each of its potential eight edge neighbors is allocated a value between zero and seven as shown in Fig. 3a.

To compute the chain codes, the resulting flattened input binary edge map is split into non-overlapping cells of size 8×8 pixels. For each edge contour within each cell, the Freeman chain code is determined starting from that edge-contour's upper leftmost pixel and traversing counter-clockwise as shown in Fig. 3b. For this 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 grouped into overlapping blocks of size 2×2 cells with a stride of 1 cell. A histogram is then created for each block's resulting chain codes, which is normalized using the L2 norm. The resulting normalized histograms of all blocks are concatenated row by row generating a normalized chainlet histogram vector.

3. Pairwise Support Vector Machine

Although various classification algorithms can be used, in this research, a pairwise Support Vector Machine (SVM) is employed for its simplicity. Pairwise SVM takes two inputs and determines if they belong to the same class, whereas standard SVM takes only one input and seeks to determine its class. Let $R = \{H_1^1, ..., H_i^j, ..., H_N^M\}$ 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:

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

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

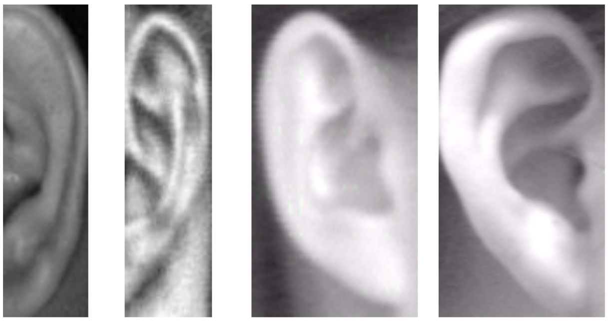
$$y_{mn} = \begin{cases} +1, m = n \\ -1, m \neq n \end{cases} \tag{5}$$

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

$$K((a,b),(c,d)) := k(a,c) + k(b,d) \tag{6}$$

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

$$k(a,c) = a^T c \tag{7}$$



(a) (b) (c) (d)

Figure 4. Sample images of two unique individuals from the IITD II dataset (a-b) [19]. Sample images of two unique individuals from the USTB I dataset (c-d) [20].

Table 1. Experimental results for the proposed Chainlet based Multi-Band Ear Recognition using Support Vector Machine (CERMB-SVM) ear recognition algorithms on the IITD II [19] 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

In this research, the parameters alpha and gamma were learned as detailed in [21]. Moreover, for each ear image dataset, the first two images of each individual were used for training and the remaining images were used for testing.

4. Experimental Results

Two benchmark ear image datasets named the Indian Institute of Technology Delhi II (IITD II) [19] and the University of Science and Technology Beijing I (USTB I) [20], which are widely used in the literature [7,14–16,22], were used to create experimental results. These two datasets were chosen due to their widespread use in recent publications and also because their images have been pre-aligned. The IITD II dataset contains 793 images of the right ear of 221 participants. Each participant was photographed between three and six times, where the images are unsigned 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 contains 180 images of the right ear of 60 participants, each of whom were photographed three times. The images in this dataset are unsigned 8-bit grayscale of size 150×80. The images in USTB I are tightly cropped; however, they demonstrate some slight rotation and shearing. Example images from both datasets are shown in Fig. 4.

The proposed CERMB-SVM algorithm 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 2.2. The number of correct matches was computed for each set of bands. A subset of the results for both the IITD II and USTB I image datasets are tabulated in Table 1 and Table 2 respectively. From these tables, the proposed technique attains its highest performance at four and seven bands when applied to the images of the IITD II and USTB I datasets, respectively. From these tables, it can be observed that the proposed CERMB-SVM technique has slightly higher performance when applied to the images of the IITD II dataset rather than those of the USTB I dataset.

To compare the performance of the proposed CERMB-SVM algorithm with the statistical PCA and anchor chainlet, and state of the art learning based techniques, the Rank-1

Table 2. Experimental results for the proposed Chainlet based Multi-Band Ear Recognition using Support Vector Machine (CERMB-SVM) ear recognition algorithm on the USTB I [20] 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

Table 3. Experimental results for the proposed Chainlet based Multi-Band Ear Recognition using Support Vector Machine (CERMB-SVM) ear recognition method on the USTB I [20] dataset.

Algorithm	Dataset	
	IITD II	*USTB I*
Statistical based Techniques		
Single Image PCA	36.35	45.00
Eigenfaces [4]	89.78	75.93
2D-MBPCA [16]	91.12	85.19
Chainlets [17]	98.54	99.02
Learning based Techniques		
BSIF and SVM[15]	97.31	-
GoogLeNet[11]	98.57	99.36
ResNet18 and SVM[12]	98.76	99.44
VGG-based Ensembles[23]	98.88	99.24
Neural Network and SVM[14]	-	98.30
Proposed CERMB-SVM Technique	**99.02**	**99.44**

experimental results of the proposed CERMB-SVM, single image PCA, 'eigenfaces' [4], 2D-MBPCA [16], 'BSIF and SVM' [15], GoogLeNet [11], 'ResNet18 and SVM' [12], VGG-based Ensembles [23] and 'neural network and SVM' based [14] methods are tabulated in Table 3. From this table, the proposed CERMB-SVM method significantly outperforms both the PCA based and learning based state of the art algorithms for the images of the IITD II dataset. Additionally, the proposed CERMB-SVM method significantly outperforms the PCA based methods and slightly outperforms the learning-based algorithms on the USTB I dataset.

A further comparison between the proposed 2D-CERMB-SVM technique and the aforementioned algorithms is demonstrated using Cumulative Match Curves (CMC). Regions of interest for the CMC curves comparing 2D-CERMB-SVM to the statistical based methods on the IITD II and USTB I datasets are shown in Fig. 5 and 6, respectively. In addition, regions of interest for the CMC curves for both the proposed 2D-CERMB-SVM and learning based techniques on both datasets are shown in Fig. 7 and 8. From Fig. 5 and 6, it can be seen that the proposed 2D-CERMB-SVM algorithm greatly outperforms the PCA based methods. In addition, the proposed technique generates superior results to its anchor chainlet method. From Fig. 7 and 8, it is evident that the 2D-CERMB-SVM algorithm generates superior results to the 'BSIF and SVM', GoogLeNet, VGG-based Ensembles, and 'Neural Network and SVM' methods. However, the proposed technique generates identical results to that of the 'ResNet18 and SVM' method on the USTB I dataset.

4.1. Justification of the Achieved Performance

From the experimental results, it is clear that the proposed CERMB-SVM technique significantly outperforms the PCA based methods. This performance can be explained by

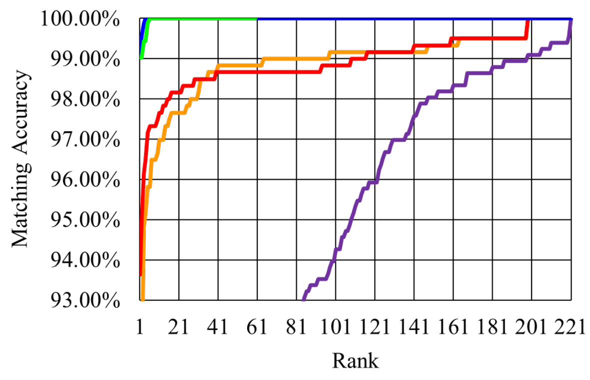


Figure 5. Region of interest of the CMC curves for Single Image PCA (purple), eigenfaces (orange), 2D-MBPCA (red), Chainlets (green) and 2D-CERMB-SVM (blue) for the IITD II dataset [19].

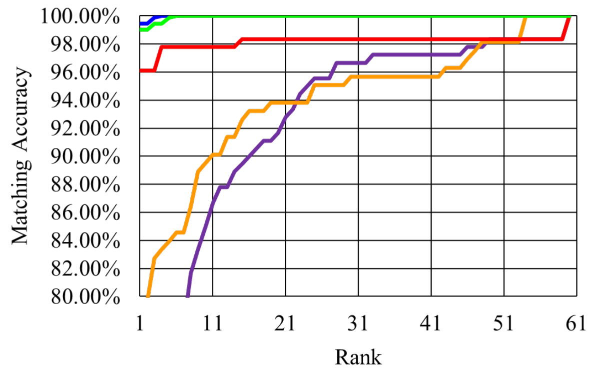


Figure 6. Region of interest of the CMC curves for Single Image PCA (purple), eigenfaces (orange), 2D-MBPCA (red), Chainlets (green) and 2D-CERMB-SVM (blue) for the USTB I dataset [20].

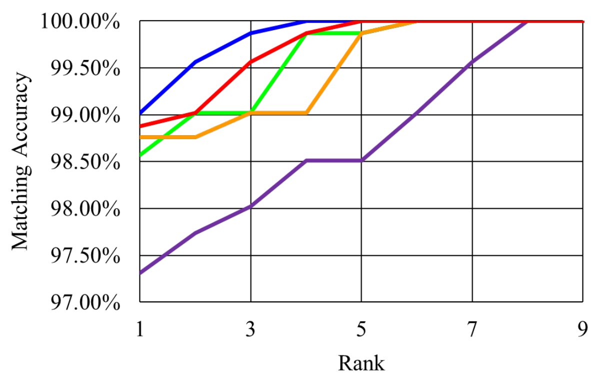


Figure 7. Region of interest of the CMC curves for 'BSIF and SVM' (purple), GoogLeNet (green), 'ResNet18 and SVM' (orange), VGG-based Ensembles (red) and 2D-CERMB-SVM (blue) for the IITD II dataset [19].

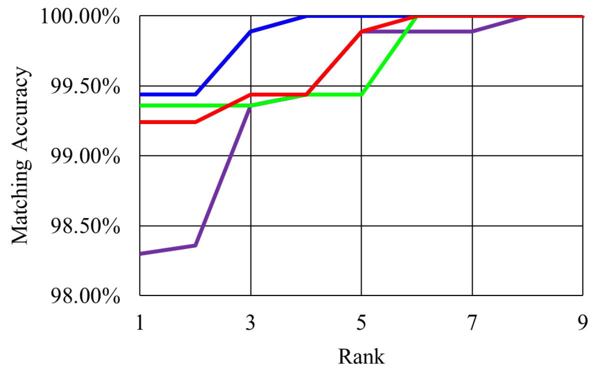


Figure 8. Region of interest of the CMC curves for 'Neural Network and SVM' (purple), GoogLeNet (green), VGG-based Ensembles (red) and 2D-CERMB-SVM (blue) for the USTB I dataset [20]. The 'ResNet18 and SVM' method produced identical results to 2D-CERMB-SVM.

the fact that the multiple band image generation process expands the ear image feature space by a factor of $b-1$, where b is the number of frames. However, due to flattening of the resulting edge maps of different bands, some edges will overlap, resulting in a slight reduction of the increased feature space. To give the reader a visualized justification of the selection of optimum number of bands, the average of the total eigenvector energy for the resulting edge maps for different bands are calculated to represent the effectiveness of the resulting features generated by the multi-banding process. In addition, the average number of features for different number of bands is also plotted on the same graph. These two calculations can be seen in Fig. 9 and Fig. 10 for the IITD II and USTB I datasets, respectively.

From these figures, it can be seen that the total eigenvector energy for the resulting edge maps decreases as the number of bands increases. At the same time, the average number of features post-flattening increases with the number of the bands in a slightly less than linear fashion due to edge overlap in the flattening process. The intersection of these two graphs represents the optimal number of frames that can be used to produce the highest matching performance. The intersection of the Eigenvector Energy and Number of Features graphs occurs at approximately four bands for the IITD II dataset and seven bands for the USTB I dataset. This is consistent with the experimental results for finding the optimal number of bands in Section 4.

4.2. Execution Time

Ear recognition techniques can generally be classified into two main categories: statistical based and learning based techniques. Statistical based techniques, including PCA, Eigenfaces, 2D-MBPCA, and the anchor Chainlet technique, extract some statistics or features directly from the image and use these features to find the best match, while learning based techniques use a range of information including image statistics, features, and other data extracted from the image dataset to train classifiers such as neural networks and support vector machines such as the proposed CERMB-SVM method. Learning based techniques then use the trained classifiers to find the best match for an input query image. Consequently, learning based ear recognition algorithms are much more computationally expensive than their statistical based counterparts.

To give the reader a sense of the computational complexity of the proposed CERMB-SVM algorithm with respect to statistical based methods, as well as the state of the art learning based techniques, Single Image PCA, eigenfaces [4], 2D-MBPCA [16], the anchor Chainlet, 'BSIF and SVM' [15], GoogLeNet [11], 'ResNet18 and SVM' [12], VGG-based Ensembles [23], 'neural network and SVM' based [14], and the proposed CERMB-SVM methods were implemented in MATLAB. The resulting algorithms were then executed

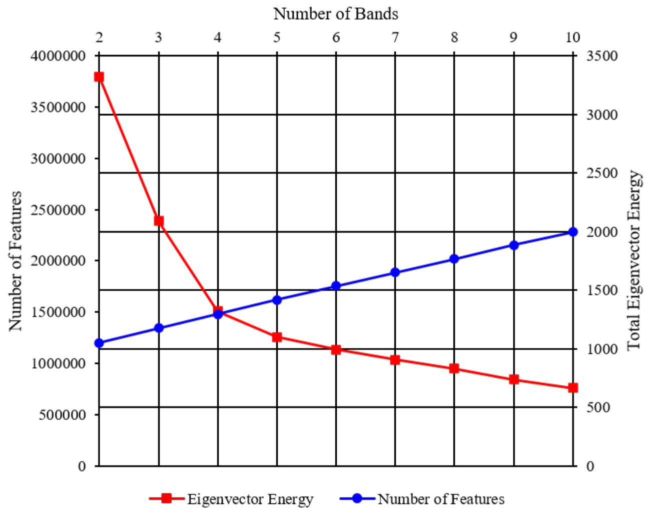


Figure 9. The number of features and total eigenvector energy versus the number of frames, where the intersection demonstrates the number of frames for maximum achievable performance, for the IITD II dataset [19].

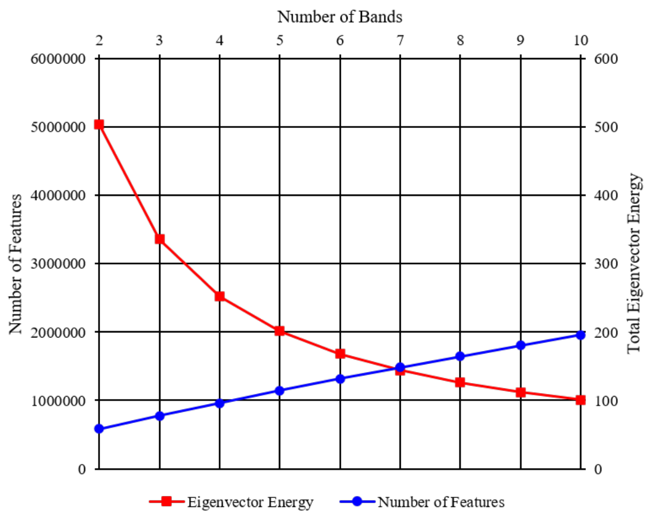


Figure 10. The number of features and total eigenvector energy versus the number of frames, where the intersection demonstrates the number of frames for maximum achievable performance, for the USTB I dataset [20].

Table 4. Average execution time (milliseconds) of the proposed CERMB-SVM and the state of the art PCA based and learning based algorithms

Algorithm	Dataset	
	IITD II	*USTB I*
PCA based Techniques		
Single Image PCA	13.55	12.16
Eigenfaces[4]	3.10	1.82
2D-MBPCA[16]	13.64	13.07
Learning based Techniques		
BSIF and SVM[15]	23.57	-
GoogLeNet[11]	22.88	21.59
ResNet18 and SVM[12]	24.24	23.88
VGG-based Ensembles[23]	23.51	22.79
Neural Network and SVM[14]	-	22.78
Proposed CERMB-SVM Technique	**22.21**	**21.85**

on a Windows 10 personal computer equipped with a 7th generation Intel core i7 processor, an Nvidia GTX 1080 graphics card, and a 512 GB Toshiba NVMe solid-state drive (no other applications, updates or background programs were running during the computation). The average computation time for processing an query image using each algorithm (learning based techniques were already trained and their training time has not been included in their measurement) was measured using 100 randomly selected query images from each dataset. The resulting measurements are tabulated in Table 4.

5. Conclusions

In this paper, application of multi-band image processing together with Chainlets and Support Vector Machine for ear recognition was investigated. This resulted in development of a Chainlet based Ear Recognition algorithm using Multi-Banding and Support Vector Machine (CERMB-SVM) algorithm, which significantly outperforms the statistical based ear recognition techniques and gives superior results to those of the learning-based methods in terms of accuracy. The proposed CERMB-SVM method splits the input ear image into several bands based on the intensity of its pixels. Canny edge detection algorithm along with morphological operators were used to generate and select edge map for each resulting bands. A single binary edge map image was created by combining the edge maps of different image bands. This resulting single binary edge map was divided into cells and the Freeman chain code for each cell was calculated. The resulting cells are then clustered into overlapping blocks, and a histogram for each block is computed. The resulting histograms are normalized and concatenated to create 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 show that the proposed CERMB-SVM technique significantly outperforms the statistical-based techniques, in terms of accuracy. The proposed CERMB-SVM technique generates 62.67%, 9.24%, 7.90% and 0.48% higher than single image PCA, Eigenfaces, 2D-MBPCA and anchor Chainlet in terms of accuracy on images of IITD II dataset, respectively. The proposed algorithm produces 54.44%, 23.51%, 14.25% and 0.42% higher accuracy than single image PCA, Eigenfaces, 2D-MBPCA and anchor Chainlet technique on images of USTM I dataset, respectively. Experimental results show that the proposed CERMB-SVM technique generates superior or the same performance in terms of accuracy than those of learning-based methods. It generates 1.71%, 0.45%, 0.26% and 0.14% higher accuracy on images of IITD II datasets compared to those of ”BSIF and SVM”, ”GoogLeNet2”, ”ResNet18 and SVM” and ”Neural Network and SVM” techniques, respectively. Moreover, the proposed technique produces 0.08%, 0.2%, 1.14% greater accuracy on images of USTM I dataset than those of "GoogLeNet", "VGG-based Ensembles" and "Neural Network and SVM" methods, respectively. The proposed algorithm generates similar results to "ResNet18 and SVM" method.

The proposed CERMB-SVM algorithm can be applied to different applications including iris and drone recognition. The MATLAB implementation of the algorithm indicates that the proposed algorithm generates competitive results compared to those of learning-based algorithms at a portion of their computation cost. However, the real-time implementation of the proposed algorithm on DSP or FPGA can be considered as the future work for this research.

Author Contributions: Conceptualization, Matthew Martin Zarachoff and Akbar Sheikh-Akbari; Funding acquisition, Akbar Sheikh-Akbari and Dorothy Monekosso; Investigation, Matthew Martin Zarachoff; Methodology, Matthew Martin Zarachoff and Akbar Sheikh-Akbari; Software, Matthew Martin Zarachoff; Supervision, Akbar Sheikh-Akbari and Dorothy Monekosso; Writing – original draft, Matthew Martin Zarachoff; Writing – review & editing, Matthew Martin Zarachoff and Akbar Sheikh-Akbari.

Funding: This research has been funded under a knowledge transfer partnership by Innovate UK (KTP 10304).

Acknowledgments: The first author would like to thank Leeds Beckett University for their support through a fully-funded studentship.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nejati, H.; Zhang, L.; Sim, T.; Martinez-Marroquin, E.; Dong, G. Wonder ears: Identification of identical twins from ear images. Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), 2012, pp. 1201–1204.
2. Emeršič, v.; Štruc, V.; Peer, P. Ear recognition: More than a survey. *Neurocomputing* **2017**, *255*, 26–39. doi:10.1016/j.neucom.2016.08.139.
3. Victor, B.; Bowyer, K.; Sarkar, S. An evaluation of face and ear biometrics. Object recognition supported by user interaction for service robots, 2002, Vol. 1, pp. 429–432 vol.1. doi:10.1109/ICPR.2002.1044746.
4. Chang, K.; Bowyer, K.W.; Sarkar, S.; Victor, B. Comparison and combination of ear and face images in appearance-based biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2003**, *25*, 1160–1165. doi:10.1109/TPAMI.2003.1227990.
5. Querencias-Uceta, D.; Ríos-Sánchez, B.; Sánchez-Ávila, C. Principal component analysis for ear-based biometric verification. 2017 International Carnahan Conference on Security Technology (ICCST), 2017, pp. 1–6. doi:10.1109/CCST.2017.8167843.
6. Turk, M.A.; Pentland, A.P. Face recognition using eigenfaces. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Proceedings, 1991, pp. 586–591. doi:10.1109/CVPR.1991.139758.
7. Nosrati, M.S.; Faez, K.; Faradji, F. Using 2D wavelet and principal component analysis for personal identification based On 2D ear structure. 2007 International Conference on Intelligent and Advanced Systems, 2007, pp. 616–620. doi:10.1109/ICIAS.2007.4658461.
8. Basit, A.; Shoaib, M. A human ear recognition method using nonlinear curvelet feature subspace. *International Journal of Computer Mathematics* **2014**, *91*, 616–624. doi:10.1080/00207160.2013.800194.
9. Hassaballah, M.; Alshazly, H.A.; Ali, A.A. Robust local oriented patterns for ear recognition. *Multimedia Tools and Applications* **2020**, *79*, 31183–31204. doi:10.1007/s11042-020-09456-7.
10. Galdámez, P.L.; Arrieta, A.G.; Ramón, M.R. Ear recognition using a hybrid approach based on neural networks. 17th International Conference on Information Fusion (FUSION), 2014, pp. 1–6.
11. Eyiokur, F.I.; Yaman, D.; Ekenel, H.K. Domain adaptation for ear recognition using deep convolutional neural networks. *IET Biometrics* **2018**, *7*, 199–206. doi:10.1049/iet-bmt.2017.0209.
12. Dodge, S.; Mounsef, J.; Karam, L. Unconstrained ear recognition using deep neural networks. *IET Biometrics* **2018**, *7*, 207–214. doi:10.1049/iet-bmt.2017.0208.
13. Alshazly, H.; Linse, C.; Barth, E.; Martinetz, T. Deep Convolutional Neural Networks for Unconstrained Ear Recognition. *IEEE Access* **2020**, *8*, 170295–170310. Conference Name: IEEE Access, doi:10.1109/ACCESS.2020.3024116.
14. Omara, I.; Wu, X.; Zhang, H.; Du, Y.; Zuo, W. Learning pairwise SVM on deep features for ear recognition. 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), 2017, pp. 341–346. doi:10.1109/ICIS.2017.7960016.
15. Benzaoui, A.; Hezil, N.; Boukrouche, A. Identity recognition based on the external shape of the human ear. 2015 International Conference on Applied Research in Computer Science and Engineering (ICAR), 2015, pp. 1–5. doi:10.1109/ARCSE.2015.7338129.
16. Zarachoff, M.; Sheikh-Akbari, A.; Monekosso, D. 2D Multi-Band PCA and its Application for Ear Recognition. 2018 IEEE International Conference on Imaging Systems and Techniques (IST), 2018, pp. 1–5. doi:10.1109/IST.2018.8577132.
17. Ahmad, A.; Lemmond, D.; Boult, T.E. Chainlets: A New Descriptor for Detection and Recognition. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 1897–1906. doi:10.1109/WACV.2018.00210.

18. Zarachoff, M.M.; Sheikh-Akbari, A.; Monekosso, D. Non-Decimated Wavelet Based Multi-Band Ear Recognition Using Principal Component Analysis. *IEEE Access* **2022**, *10*, 3949–3961. doi:10.1109/ACCESS.2021.3139684.
19. IIT Delhi Ear Database. Accessed: 2018-03-13.
20. Ear Recoginition Laboratory at USTB. Accessed: 2018-06-25.
21. Brunner, C.; Fischer, A.; Luig, K.; Thies, T. Pairwise Support Vector Machines and their Application to Large Scale Problems. p. 14.
22. Benzaoui, A.; Boukrouche, A. Ear recognition using local color texture descriptors from one sample image per person. 2017 4th International Conference on Control, Decision and Information Technologies (CoDIT), 2017, pp. 0827–0832. doi:10.1109/CoDIT.2017.8102697.
23. Alshazly, H.; Linse, C.; Barth, E.; Martinetz, T. Ensembles of Deep Learning Models and Transfer Learning for Ear Recognition. *Sensors* **2019**, *19*, 4139. Number: 19 Publisher: Multidisciplinary Digital Publishing Institute, doi:10.3390/s19194139.