Convolutional Neural Network-Based Human Identification Using Outer Ear Images

Harsh Sinha, Raunak Manekar, Yash Sinha and Pawan K. Ajmera

Abstract This paper presents a deep learning approach for ear localization and recognition. The comparable complexity between human outer ear and face in terms of its uniqueness and permanence has increased interest in the use of ear as a biometric. But similar to face recognition, it poses challenges such as illumination, contrast, rotation, scale, and pose variation. Most of the techniques used for ear biometric authentication are based on traditional image processing techniques or handcrafted ensemble features. Owing to extensive work in the field of computer vision using convolutional neural networks (CNNs) and histogram of oriented gradients (HOG), the feasibility of deep neural networks (DNNs) in the field of ear biometrics has been explored in this research paper. A framework for ear localization and recognition is proposed that aims to reduce the pipeline for a biometric recognition system. The proposed framework uses HOG with support vector machines (SVMs) for ear localization and CNN for ear recognition. CNNs combine feature extraction and ear recognition tasks into one network with an aim to resolve issues such as variations in illumination, contrast, rotation, scale, and pose. The feasibility of the proposed technique has been evaluated on USTB III database. This work demonstrates 97.9% average recognition accuracy using CNNs without any image preprocessing, which shows that the proposed approach is promising in the field of biometric recognition.

H. Sinha · R. Manekar · P. K. Ajmera (✉)
Department of Electrical and Electronics Engineering, BITS, Pilani, Pilani Campus, Pilani, Rajasthan, India
e-mail: pawan.ajmera@pilani.bits-pilani.ac.in

Y. Sinha
Department of Computer Science and Information Systems, BITS Pilani, Pilani Campus, Pilani, Rajasthan, India
e-mail: mail.yash.sinha@gmail.com

© Springer Nature Singapore Pte Ltd. 2019
J. C. Bansal et al. (eds.), *Soft Computing for Problem Solving*, Advances in Intelligent Systems and Computing 817, https://doi.org/10.1007/978-981-13-1595-4_56

1 Introduction

As a form of security enhancement, human traits are being used extensively for biometric authentication. The most prominent biometric used today are face, iris, and fingerprint. However, there are some major drawbacks with them such as illumination, intrusiveness, facial makeup, expressions, surgical alterations, birth defects, and physical changes with aging. Hence, researchers are finding different recognition techniques for better security. In recent years, using human ear as a biometric trait has become a popular choice. Descartes Biometrics developed an app for Android that uses ear as a password [5]. An iOS application was developed that allows a medical practitioner to search the medical data on using patient's ear as the key [3].

The motivation behind the present study derives from extensive work done in the field of image recognition using convolutional neural networks (CNNs) and histogram of oriented gradients (HOG) feature extraction technique [9, 23, 24]. This paper demonstrates the use of CNNs for the task of ear recognition and achieves very high scale-invariant ear localization accuracy using a simple (HOG+SVM) framework on a multi-scale image pyramid.

It is stated that most of recognition techniques give very poor accuracy when applied directly on acquired images [12]. Reported works have suggested addition of new image enhancement and feature extraction stages in the image preprocessing pipeline for making recognition invariant to scale, rotation, illumination, etc. [2, 27]. Hence, LeCun et al. [23] have suggested a traditional recognition framework which is preceded by a feature extraction step to extract relevant information from images (training samples). These features (or vectors) are then fed into a generic classifier, which classifies them into one-hot-encoded classes. Extracting features on the basis of performance of classifiers is found to be a much better approach [19, 23]. Hence, CNNs are suitable for such a scenario as they extract relevant features and train the end classifier pivoting on backpropagation algorithm. As a result, any explicit feature extraction pipeline was eliminated.

1.1 Detection

Ear detection is an important preliminary step for ear biometrics. In the literature, researchers achieve detection of ear in a given profile face using image processing pipeline similar to skin detection [4], skin segmentation [29], edge computation [41], ear-candidates generation by frequent subgraph mining [28], and template matching [6]. However, some researchers preferred to achieve it manually.

Prakash and Gupta [31] achieved a ear localization accuracy of 94% over 150 individuals which involved skin segmentation followed by template matching. Yan and Bowyer [41] used contour information while Yuan and Mu [43] used skin color and contour information to localize the ear and edges to detect ear pit. Chen and Bhanu [4] presented a template matching approach which used canny edge detector for segregating convex and concave images. Researchers have also focused on assuming a shape for the ear, e.g., elliptical [4], wavelet based [32] or average of known ear images [31].

Clearly, the detection pipeline in recent literature depends upon texture and edge information for localization of ear. However, the proposed HOG+SVM technique involves only two steps, calculation of gradients on the image and classification using pre-trained SVM on given dataset. There are no pre-assumptions regarding the background or shape of the ear. Moreover, HOG feature extraction has advantage of lower complexity in terms of computational time and greater accuracy as compared to popular feature extraction like Gabor filters and Haar features [9, 34, 37].

1.2 Recognition

In the past decade, most of the work in ear recognition have focused on identifying new features and representations from ear images.

Apart from holistic techniques such as linear discriminant analysis (LDA), wavelet transforms [32, 39], Gabor wavelets [22], log-Gabor filters [18, 21], 2-D quadrature filters [7] which are applied on the entire image, another common way of feature extraction is extracting local image information by dividing an image into many small windows and applying transforms on them. Sparse representation of localized Radon transform created a robust ear representation [20] while an extension of principal components analysis (PCA) was applied to create new features from the features calculated from input image to make recognition less sensitive to environmental variations [15]. Calculating local information for forming new feature set was also suggested [42]. Force-field transformations proved efficient as they achieved 99.2% accuracy on XM2VTS dataset [16]. Moreover, a modular neural network achieved a recognition accuracy of 75.44% using USTB dataset [33]. Image contrast enhancement techniques such as scale-invariant feature transform (SIFT) with artificial bee colony (ABC) algorithm, histogram equalization, and contrast-limited adaptive histogram equalization (CLAHE) have also been explored [12].

The performance of approaches based on 2D images rely on pose and illumination variation. Hence, 3D ear images have also been proposed in order to overcome these problems [25, 28, 30] as 3D models contain depth information which further enhances the accuracy of a biometric system. However, using 3D models as a biometric would lead to bulky databases and a slow verification system.

In this paper, a HOG-based descriptor for localization of ear and CNN for recognition is proposed. HOG computes gradients to classify images. Localized ear is cropped and fed into CNN for recognition which combine the feature extraction and ear recognition. CNNs have the ability to identify which features are most suitable for recognition on the particular training set and learn specific features resulting in most optimized recognition. Hence, it eliminates the need for handcrafted features. The proposed methodology is summarized in Fig. 1.

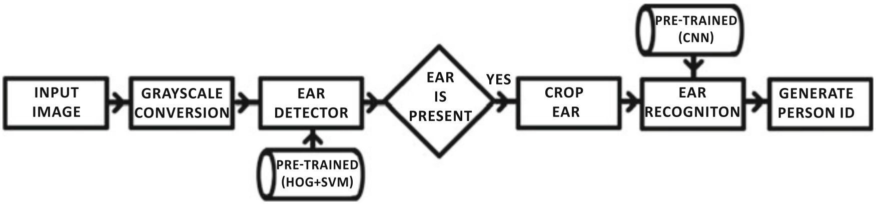

Fig. 1 Proposed pipeline

The paper is organized as follows, the proposed methodology along with HOG and CNN is described in Sect. 2. The dataset used and performance evaluation is presented in Sect. 3. Conclusions based on the experimental results are presented in Sect. 4.

2 Proposed Method

A biometric system mainly consists of of image preprocessing, detection, and recognition. In subsequent sections, detection and recognition methodology used is discussed. Preprocessing is minimal; that is, the acquired images were normalized and grayscaled.

2.1 Ear Localization

Object localization consists of three basic steps, viz. choosing a feature descriptor, training a classifier on positive and negative samples of the object, and classification of pixels inside scanning window using the trained classifier. Different approaches varies in types of feature descriptor and the generic classifier chosen.

Two major approaches which are used for ear localization are Haar feature-based cascade classifier [37] and HOG with SVM [8, 26]. Both techniques have been applied on a multi-scale image pyramid representation of the side profile image for scale-invariant ear detection.

In HOG, an object appearance and shape is characterized by intensity gradients or edge directions. The image window is divided into many small regions. The gradients of an image is calculated. Moreover, the computed gradient is converted to polar coordinates to accommodate gradient directions.

The flowchart in Fig. 2 describes the procedure for ear detection. The first step is to generate an image pyramid for each image in the training dataset which ensures that the object detector is scale invariant. For each region, a histogram of gradient directions is calculated over the pixels. Then, the cells are grouped to perform nor-

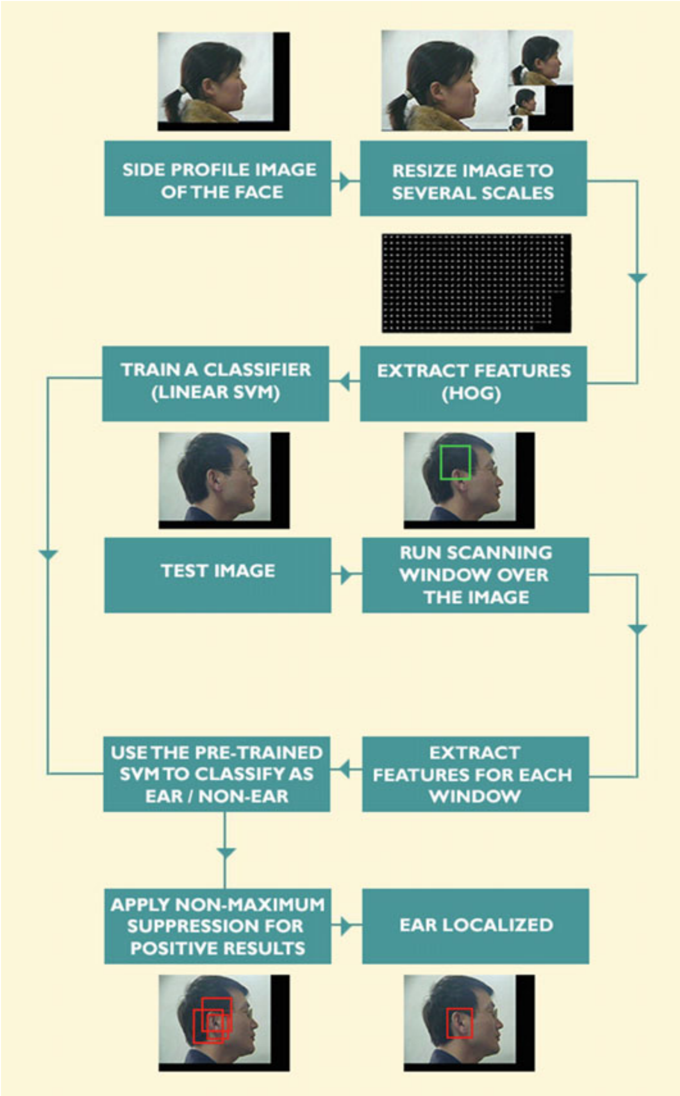

Fig. 2 Flowchart of ear localization methodology

malization across 2 $\times$ 2 cells. The normalized features are concatenated as a single HOG feature descriptor.

A classifier (linear SVM) is trained on extracted features using large number of positive and negative samples of the object (ear). When a test image is fed as input, pixels inside a scanning window are classified using the trained classifier. Generally, it is found that the detector generates several overlapping bounding boxes for a detected image. Hence, non-maximum suppression is used to generate final bounding box. Figures 4 and 5 shows the performance of (HOG+SVM) on USTB III dataset.

2.2 Ear Recognition

The problem of ear recognition is very similar to that of face recognition. It faces problem of bad illumination, pose variation, and occlusion. Recognition models are generally unable to overcome problems such as local deformation, rotation, or translation.

CNNs possess certain characteristic concepts such as local receptive fields, shared weights, and subsampling which makes it capable to overcome aforementioned problems. In this paper, a CNN structure depicted in Fig. 3 is considered for feature extraction and classification.

The CNN architecture chosen is a standard network used for image recognition. However, the architecture is much smaller as compared to very recent architectures used for large-scale general image recognition. As in our case, the number of objects to be learned is much smaller, a moderate-sized network turned out to be fast and sufficient enough to classify less than 100 individuals.

The applied CNN architecture can be described as $[INPUT - CONV - RELU - POOL - CONV - RELU - POOL - FC1 - FC2 - OUTPUT]$ where $CONV$: convolutional layer, $RELU$: rectified linear unit, $POOL$: max pooling layer, FC: fully connected.

The parameters of CNN were optimized using Adagrad optimizer [10] in which different learning rates are calculated for different parameters θ. Assuming, for tth time step and ith parameter, the gradient of the cost function is

$$g_{t,i} = \nabla_\theta J(\theta_i) \tag{1}$$

According to Adagrad update,

$$\theta_{t+1,i} = \theta_{t,i} - \frac{\eta}{\sqrt{G_{t,ii} + \epsilon}} g_{t,i} \tag{2}$$

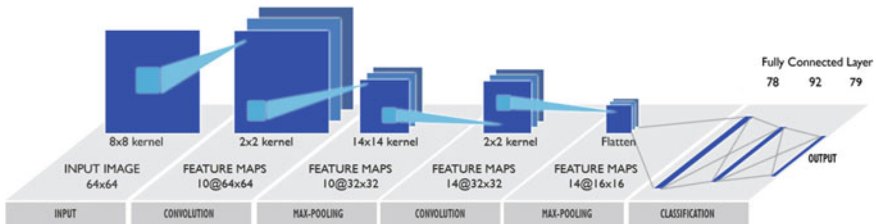

Fig. 3 Proposed convolutional neural network for ear recognition

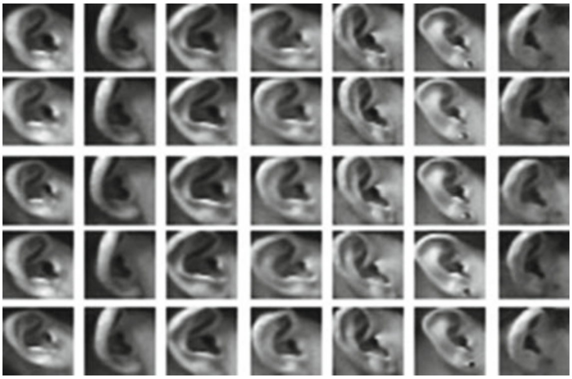

Fig. 4 Cropped ears: localization performance of HOG + SVM

Fig. 5 Test image and cropped ear after localization

Here, G_t is a $d \times d$ diagonal matrix. Every element ii is summation of squares of gradients w.r.t θ_i upto time step t.

However, in stochastic gradient descent (SGD), the algorithm has a fixed learning rate.

$$\theta_{t+1,i} = \theta_{t,i} - \eta g_{t,i} \quad (3)$$

Hence, Adagrad optimizer automatically manipulates the learning rate to achieve a better accuracy.

Hyperparameters such as kernel size or number of layers are used to tune the performance of a machine learning model for unseen data. SigOpt was used to identify best hyperparameters [35, 38]. SigOpt creates a feedback loop between model performance and different values for hyperparameters.

To further improve the performance on the validation set, dropout and L2 regularization were employed [36]. Dropout prevents co-adaptation by nullifying nodes in the network. L2 regularization or weight decay selects the smallest values for different parameters by penalizing for large weights. The weights or parameters were initialized with a defined variance, which ensured that the problem of vanishing gradients is eliminated. Moreover, weight initialization improved consistency of results [1, 13]. Batch normalization was employed which resolves the problem of internal covariate shift [17]. Applying these techniques together during training had a compounding effect on performance of model.

3 Performance Evaluation

In this section, the proposed algorithm has been analyzed and evaluated by performing various experiments on the ear recognition task using the USTB III dataset. A comparative performance has been carried out against some of the existing ear recognition schemes such as wavelet transforms [14], log-Gabor filters [21] and image enhancement techniques [12].

The recognition rates in all the experiments are computed for the correct number of matches out of the total ears used for testing. The various performance measures used are defined as follows:

$$Precision_{micro} = \frac{tp_1 + ... + tp_k}{tp_1 + ... + tp_k + fp_1 + ... + fp_k} \tag{4}$$

$$Recall_{micro} = \frac{tp_1 + ... + tp_k}{tp_1 + ... + tp_k + fn_1 + ... + fn_k} \tag{5}$$

$$Precision_{macro} = \frac{Precision_1 + ... + Precision_k}{k} \tag{6}$$

$$Recall_{macro} = \frac{Recall_1 + ... + Recall_k}{k} \tag{7}$$

where, tp = true positives, fp = false positives, fn = false negatives, tn = true negatives.

3.1 Dataset

The University of Science and Technology in Beijing (USTB) provides three databases for public use. The USTB III ear dataset contains side profile images of face. The database contains images from 79 subjects, each image having a resolution of 768 × 576. Each subject has approximately 10 uncropped images (a total of 785 images).

3.2 Detection

For ear detection, the dataset was divided into 60:40 as training data and test data. For Haar cascade classifier, the standard pre-trained model of ear was used. But, the classifier was unable to detect ears in images with pose variations. The detector could detect ears with 74.5% accuracy. For HOG+SVM detector, while training the detector achieved an accuracy of 97.02% (457 out of 471 images). The trained detector when used over all images was able to locate an ear in 775 images (Out of a total of 785 images), i.e., an accuracy of 98.72%.

Haar features or other wavelets are good for texture-based characterization as it relies on difference in intensities in their vicinity. But, in object detection, edge orientation is very important to identify objects. HOG descriptors depend on the shape of the object using gradient orientations. Hence, HOG has a better performance than Haar for ear localization.

3.3 Recognition

The dataset contains $8 - 10$ uncropped images per person. However, the dataset was unbalanced (as the number of images per person was unequal). So, gamma correction was used for data augmentation using following equation. Gamma correction also simulates the effects of uneven lighting. After data augmentation, there were 14 images per person. These images were then divided in the ratio of 8:4:2 as training, validation, and test dataset, respectively.

$$I_2 = A \times I_1^{\gamma} \tag{8}$$

In our case, $A = 1$, $\gamma \epsilon (0.5, 1.5)$ was used.

The convolutional neural network optimized using Adagrad optimizer (incorporating adaptive learning rate) obtained an average recognition accuracy of 97.9% on the test dataset (Fig. 6). The CNN also used dropout with a keep probability of 0.5 for incorporating regularization into the network.

For recognition, several variations in parameters for CNN were used before finally arriving at the aforementioned architecture. It was found that applying image enhancement techniques just before CNN allows it to converge much quickly. However, the recognition accuracy saturates around 75%.

To asses the performance of the trained recognition network, confusion matrix was computed which is visualized as a heat map in Fig. 7. The different columns represent the predicted label by trained network, while the rows represent the ground truth. The diagonal shows number of correct predictions, which makes it is easy to identify mis-classified classes.

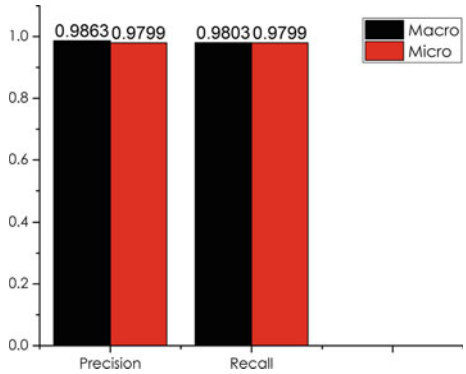

Fig. 6 Precision and recall of proposed CNN

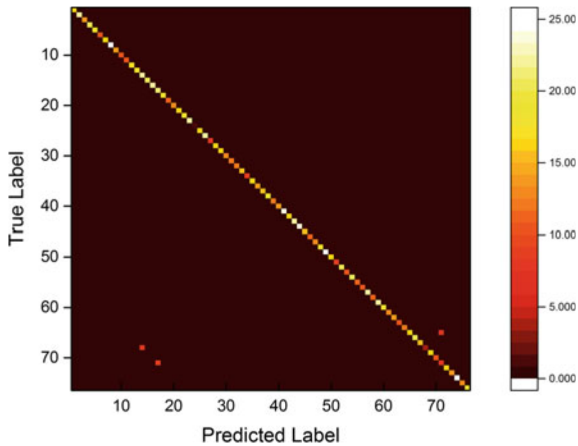

Fig. 7 Confusion matrix of CNN

3.4 *Comparative Performance*

This section compares performance of proposed method with some of the other prominent techniques which have used the similar dataset.

Yuan and Mu [43] achieved an accuracy of 90% used active shape model for detecting ears. Further, they used full-space linear discriminant Analysis (FSLDA) for recognizing ears. Xie and Mu [40] used local linear embedding (LLE) for nonlinear dimensionality reduction. Initially, they obtained an accuracy of 60.75% accounting for all poses. However, they developed a better LLE model which had a recognition rate of 80% for ear poses in the range [−10, 20] and 90% for ear poses in the range [0, 10]. Hai-Long and Zhi-Chun [14] used orthogonal centroid algorithm over low-frequency sub images and achieved an accuracy of 97.2%.

Galdmez et al. [11] used CNN over Avila's police school dataset and Bisite videos dataset and reported an accuracy of 94.79% and 40.13%, respectively. Kumar and Wu [21] achieved a recognition rate of 96.27% over IITD dataset while Ghoualmi et al.

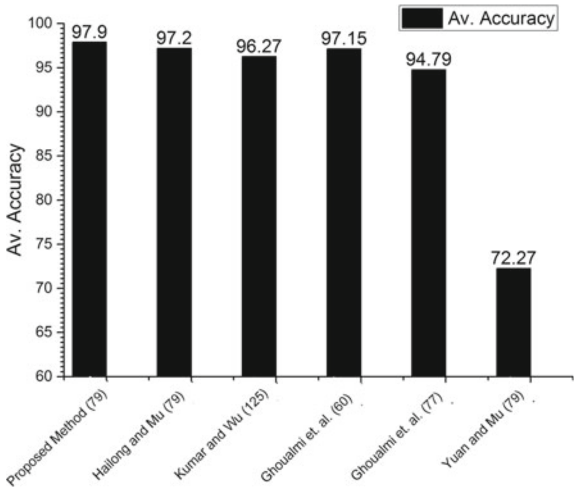

Fig. 8 Comparative performance of the proposed methodology (numbers in parenthesis denote no. of subjects)

[12] achieved 97.15% and 94.79% over USTB I and USTB II datasets, respectively. In Fig. 8, recent methodology is compared with a comparable subject count. Proposed method achieves highest recognition rate of 97.9% on USTB III dataset.

4 Conclusions

In this paper, a methodology for ear localization and recognition is proposed which reduces the pipeline for a biometric recognition system. The proposed framework uses HOG with SVM for ear localization and CNN for ear recognition. HOG descriptors depend on image gradient orientations which describe the shape of an object. Hence, HOG features provide accurate and faster ear localization. Localization accuracy of 98.72% was achieved using HOG with SVM framework. CNN combines feature extraction and ear recognition into one network. It has the ability to learn most suitable features on a given training set resulting in optimized recognition. The feasibility of the proposed technique has been evaluated on USTB III dataset. This work demonstrates 97.9% average recognition accuracy using CNN without any image preprocessing, which shows that the proposed approach is a promising in the field of biometric recognition.

References

1. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X.: TensorFlow: large-scale machine learning on heterogeneous systems (2015). http://tensorflow.org/. Software available from tensorflow.org
2. Anwar, A.S., Ghany, K.K.A., Elmahdy, H.: Human ear recognition using geometrical features extraction. Procedia Comput. Sci. **65**, 529–537 (2015)
3. Bargal, S.A., Welles, A., Chan, C.R., Howes, S., Sclaroff, S., Ragan, E., Johnson, C., Gill, C.: Image-based ear biometric smartphone app for patient identification in field settings. VISAPP **3**, 171–179 (2015)
4. Bhanu, B., Chen, H.: Human ear recognition by computer. Springer Science & Business Media (2008)
5. Biometrics, D.: Ergo Ear Biometric App: Unlock Your Phone with Your Ear. http://www.descartesbiometrics.com/ergo-app/ (2017). Accessed February 24, 2017
6. Burge, M., Burger, W.: Ear biometrics in computer vision. In: 15th International Conference of Pattern Recognition (ICPR) pp. 826–830 (2000)
7. Chan, T.S., Kumar, A.: Reliable ear identification using 2-d quadrature filters. Patt. Recogn. Lett. **33**(14), 1870–1881 (2012)
8. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)
9. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, vol. 1, pp. 886–893. IEEE (2005)
10. Duchi, J., Hazan, E., Singer, Y.: Adaptive subgradient methods for online learning and stochastic optimization. J. Mach. Learn. Res. **12**, 2121–2159 (2011)
11. Galdmez, P.L., Raveane, W., Arrieta, A.G.: A brief review of the ear recognition process using deep neural networks. J. Appl. Logic (2016). https://doi.org/10.1016/j.jal.2016.11.014. http://www.sciencedirect.com/science/article/pii/S1570868316300684
12. Ghoualmi, L., Draa, A., Chikhi, S.: An ear biometric system based on artificial bees and the scale invariant feature transform. Exp. Syst. Appl. **57**, 49–61 (2016)
13. Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. Aistats **9**, 249–256 (2010)
14. Hai-Long, Z., Zhi-Chun, M.: Combining wavelet transform and orthogonal centroid algorithm for ear recognition. In: 2nd IEEE International Conference on Computer Science and Information Technology, 2009. ICCSIT 2009, pp. 228–231. IEEE (2009)
15. Hanmandlu, M., et al.: Robust ear based authentication using local principal independent components. Exp. Syst. Appl. **40**(16), 6478–6490 (2013)
16. Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field feature extraction for ear biometrics. Comput. Vis. Image Understand. **98**(3), 491–512 (2005)
17. Ioffe, S., Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv:1502.03167 (2015)
18. Jamil, N., AlMisreb, A., Halin, A.A.: Illumination-invariant ear authentication. Procedia Comput. Sci. **42**, 271–278 (2014)
19. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
20. Kumar, A., Chan, T.S.T.: Robust ear identification using sparse representation of local texture descriptors. Pattern Recogn. **46**(1), 73–85 (2013)
21. Kumar, A., Wu, C.: Automated human identification using ear imaging. Pattern Recogn. **45**(3), 956–968 (2012)

22. Kumar, A., Zhang, D.: Ear authentication using log-gabor wavelets. In: Defense and Security Symposium on International Society for Optics and Photonics, p. 65,390A (2007)
23. LeCun, Y., Bengio, Y., et al.: Convolutional networks for images, speech, and time series. Handbook Brain Theor. Neural Netw. **3361**(10), 1995 (1995)
24. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. Neural Computat. **1**(4), 541–551 (1989)
25. Liu, Y., Zhang, B., Zhang, D.: Ear-parotic face angle: a unique feature for 3d ear recognition. Pattern Recogn. Lett. **53**, 9–15 (2015)
26. Malisiewicz, T., Gupta, A., Efros, A.A.: Ensemble of exemplar-svms for object detection and beyond. In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 89–96. IEEE (2011)
27. Omara, I., Li, F., Zhang, H., Zuo, W.: A novel geometric feature extraction method for ear recognition. Exp. Syst. Appl. **65**, 127–135 (2016)
28. Prakash, S., Gupta, P.: A rotation and scale invariant technique for ear detection in 3d. Pattern Recogn. Lett. **33**(14), 1924–1931 (2012)
29. Prakash, S., Gupta, P.: An efficient ear recognition technique invariant to illumination and pose. Telecommun. Syst. 1–14 (2013)
30. Prakash, S., Gupta, P.: Human recognition using 3d ear images. Neurocomputing **140**, 317–325 (2014)
31. Prakash, S., Jayaraman, U., Gupta, P.: A skin-color and template based technique for automatic ear detection. In: Seventh International Conference on Advances in Pattern Recognition, 2009. ICAPR'09, pp. 213–216. IEEE (2009)
32. Sana, A., Gupta, P., Purkait, R.: Ear biometrics: A new approach. In: Advances in Pattern Recognition, pp. 46–50. World Scientific (2007)
33. Sánchez, D., Melin, P.: Modular neural network with fuzzy integration and its optimization using genetic algorithms for human recognition based on iris, ear and voice biometrics. In: Soft Computing for Recognition Based on Biometrics, pp. 85–102. Springer (2010)
34. Shu, C., Ding, X., Fang, C.: Histogram of the oriented gradient for face recognition. Tsinghua Sci. Technol. **16**(2), 216–224 (2011)
35. SigOpt: Sigopt—Amplifies your Research. https://www.sigopt.com/ (2017)
36. Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**(1), 1929–1958 (2014)
37. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001, vol. 1, p. I. IEEE (2001)
38. Wang, J., Clark, S.C., Liu, E., Frazier, P.I.: Parallel Bayesian Global Optimization of Expensive Functions. arXiv:1602.05149 (2016)
39. Wang, Y., Mu, Z.c., Zeng, H.: Block-based and multi-resolution methods for ear recognition using wavelet transform and uniform local binary patterns. In: 19th International Conference on Pattern Recognition, 2008. ICPR 2008, pp. 1–4. IEEE (2008)
40. Xie, Z., Mu, Z.: Ear recognition using lle and idlle algorithm. In: 19th International Conference on Pattern Recognition, 2008. ICPR 2008. pp. 1–4. IEEE (2008)
41. Yan, P., Bowyer, K.: Empirical evaluation of advanced ear biometrics. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops, pp. 41–41. IEEE (2005)
42. Yuan, L., Chun Mu, Z.: Ear recognition based on local information fusion. Pattern Recogn. Lett. **33**(2), 182–190 (2012)
43. Yuan, L., Chun Mu, Z.: Ear recognition based on 2d images. In: First IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007, pp. 1–5. IEEE (2007)