Ear-biometrics for human identification

Shruti Nikose
Department of Electrical Engineering
Malaviya National Institute of Technology
Jaipur, India
shrutinikose2208@gmail.com

Hemant Kumar Meena
Department of Electrical Engineering (of Aff.)
Malaviya National Institute of Technology
Jaipur, India
hmeena.ee@mnit.ac.in

***Abstract*—The potential of human ear for identification was advocated long ago. It has been proved that ear of every individual is unique and can be used as a biometric to overcome the limitations of the biometrics used today. This paper presents an approach towards using ear as a biometric for identification. In this paper, a deep learning based, Convolutional neural network model is applied to the digitally processed database which gives promising results. We have used Gaussian filter and Canny edge detector for processing the image to increase recognition rate. For authentication we have used database provided by University Of Science and Technology(USTB), Bejing.**

I. Introduction

Over the years, we have witnessed a steady upward growth of biometric technology for myriad reasons but mostly due to the fact that personal identification and authentication is given more and more importance. From border and immigration control to identifying criminals to workforce management, the practical applications of biometrics are growing rapidly. In the present scenarios, identification through biometrics play a very important role in identifying humans for criminal, forensics or general identification and verification purposes so that person could be identified by self-personality rather that external id-card, pin number or password [1]. Biometrics such as facial recognition, fingerprints and iris recognition are most used due to their highly accurate results. But these biometrics have several drawbacks. Facial recognition often yield inaccurate results by the signs of aging or in an event of facial surgery, makeup or beard and in some cases due to change in expression. Iris recognition systems are not the most accurate if the subject is wearing lenses and glasses. Fingerprints recognition also doesn't quiet work in an event of burns in fingers. Another reason it can prove to be disadvantageous because it can easily be faked. Researchers have claimed that a human ear can be used as biometric and it overcomes almost all the limitations of the biometrics used today. Iannarelli etal, in [2] has shown that human ears don't age in the same proportion as that of the whole body. Rather, they age slowly [2]. Only earlobes get elongated. Furthermore, human ear is also one of our senses therefore it is mostly visible for enabling good hearing. In process of acquiring information, in contrast to facial identification system, ears cannot be disturbed by makeup, facial hair or glasses and also ears do not have emotional traits such as facial expression. However occlusion by hair, cloth or earring is possible, but it is much faster to move back hair or cloth and remove earring than removing beard or makeup. Studies have shown that ear-prints and ear-shapes of each individual is unique. Using ear as biometric is not about replacing the existing recognition systems rather it's about supplementing them to improve its efficiency.

A. DEEP LEARNING:

Deep learning, also known as deep structured learning or hierarchal learning, is a sub-part of broader family of machine learning. It is inspired by the structure and function of the human brain called artificial neural network. Andrew Ng., chief scientist from Baidu research, founded Google brain which led to the usage of deep learning techniques across several Google services. This recently explored technology is now used for object detection, handwriting recognition et. [12], advertisement, real estate and so on for its high accuracy. What deep learning methods aim for, is learning feature hierarchies from higher levels formed by composition of lower level features. Since the system automatically learns features at multiple levels of abstraction without depending on human-crafted features, it can learn complex functions mapping the input directly from the data.

B. CONVOLUTION NEURAL NETWORK:

Convolution neural network ensures some degree of shift and distortion invariance by combining three architectural ideas et. [14]. CNN image classification takes input image, process it and classify it under certain category. Computer receives an image as matrix of pixels value. Each neuron receives an input, performs dot product, and follows it with non linearity. CNN perceives images as volumes i.e. three dimensional object rather than two dimensional to be measured by only width and height. It consists of a filter which is used to extract some specific feature from an image in which it is used. CNN is composed of various layers and activation functions as follows:-

Convolution layer: It is the main building block of CNN. It merges two sets of information mathematically. Convolution is applied to an input data to produce a feature map. This is done by using a convolution

978-1-7281-6453-3/20/$31.00 ©2020 IEEE

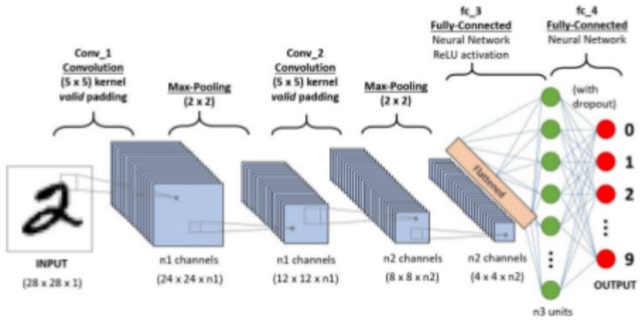


Fig. 1: CNN Framework with ReLU activation function

filter. The first layer to extract feature from an input image is convolution.

Max-pool Layers : After convolution we usually perform pooling layer (Max-pooling commonly) to reduce the size of representation, to speed up computation as well as to make some of the feature it detects a bit more robust.

Padding: Every time we apply convolution or max-pooling layer to an input image, dimensions of image shrinks. To solve this problem we add extra border to an image called padding an image before convolving.

Striding: It is the number of shifts over an input image. Usually striding of 1 or 2 is used.

Non-Linearity: ReLU is an activation function which stands for rectified Linear Unit for non-linear operation.

II. Ear Anatomy

Ear has standard parts as other biometric traits such as face, iris., etc. figure 2 shows anatomy of ears. It has several prominent features like intertragic notch, a distinctive hairpin-bend shape just above the lobe, the helix and the anti-helix which runs parallel to it and of course, the ear lobe. The central area or concha is named for its shell-like appearance[16]

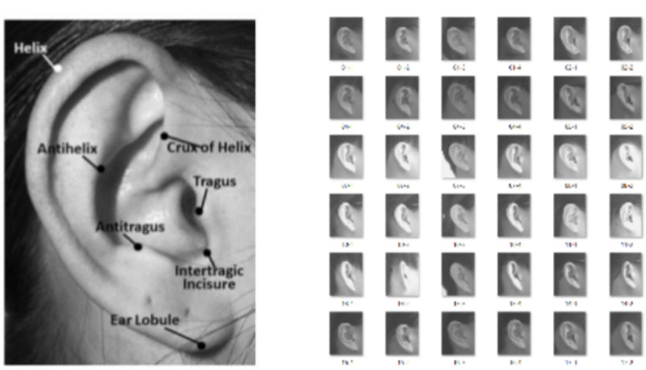


Fig. 2: Anatomy of an Ear

Using PCA(Principal component Analysis), the curve of the ear helix is found to be the most reliable anatomical structure and the basis for re-identification. Even though the individual ear yielded a high re-identification rate (88.3%), when both, the right and the left ears were paired together, amidst the pool of potential matches, our rate of re-identification was a whopping 100% [17].

III. Approaches towards Ear Biometric

The most important framework of ear print is done by Alfred Iannnarelli et.[2], who gathered 10,000 different images of ear and proved that all are unique. The classification system consists of the following steps: Firstly, a number of measurements around the ear were taken by placing a transparent compass with 8 spokes at 45 degree intervals over an enlarged photograph of the ear. For achieving the first step of registration, it is ensured that the reference line touches the crus of the helix at the top and touches the innermost point on the tragus at the bottom. Next, the enlargement mechanism is adjusted until a second reference line exactly spans the concha from top to bottom. Recent attempts have been made to automate Iannarelli's system. [18]. Burge and Burger [5] in their paper, form the Voronoi diagram using the main curve segments. They then use the Voronoi diagram to build the adjacency graph. They use adjacency graph matching based algorithm for authentication. Lightning, shadowing and occlusion are taken into account here. But, even a very small change in camera-to-ear orientation or in lighting will affect the curve segment thereby making this method unstable [5]. Hossein Nejati, Li Zhang, Terence Sim, Elisa Martinez-Marroquin, Guo Dong et.[6] proposed the uniqueness of ear even among identical twins. Even the identical twins have similar but non identical ear features. Chang, K., Bowyer, K.W., Sarker Chang, K., Bowyer, K.W., Sarker et.[4],in their paper compared the accuracy of face detection and ear print detection 71.6% for ear and 70.5 %for face .Zhichun mu et.[3] proposed an edge based recognition system including ear edge detection, feature extraction, ear description, recognition method and ear database construction. He constructed an ear database which composed of 77 subjects. Using back Propagation network as network classifier, he got recognition accuracy of 85%. Hurley et. [5] proposed an approach by extraction energy features from ear images for feature extraction. Eigen ears could provide high accuracy in closely controlled situation, otherwise have drastic performance reduction. Md. Mahbubur Rahman, Md. Rashedul Islam, Nazmul Islam Bhuiyan, Bulbul Ahmed, Md. Aminul Islam, in their paper that tracts and detects ear features simply and robustly with detection rate of 90% on same day, 88% on day variation and 87% on with light variation. If the extent of intra-individual variability is know, the designed system will be capable of differentiating between intra-individual and inter-individual variability et. [9]. Mohd Shafry Mohd Rahim1, Amjad Rehman1, Fajri Kurniawan1, Tanzila Saba, in the paper Ear biometrics for human classification based on region features mining has presented a feature extraction method using local and global features and achieved an accuracy of 86.6% using only global features and an accuracy of 93.3% with a fusion of global feature and the eigenvector feature that extracted from segmented slice regions.

IV. Database

Ear Biometrics: A Survey on Ear Image Databases and Techniques for Ear Detection and Recognition Tejas. V. Kandgaonkar1 , Rajivkumar. S. Mente2, Ashok. R. Shinde3 and Shriram. D. Raut et. [11], has presented a survey of ear biometric for any researcher who wants to work on ear biometric. In this model, we have applied the ear database collected by USTB (University of Science and Technology), Beijing. This database is composed of three sets:-

Image database 1: 60 subjects, 180 images of only their right ear . There are three images of each subject having image under normal condition, with trivial angle rotation and image under different lightening condition.[11]

Image database 2: 77 Subjects, 308 images. They are side view image(0), 2 images with difference in angle (plus 30 and minus 30) and one includes lighting dissimilarity.[11]

Image Database 3: 79 subjects: Ear and face images. Each subject requires two images to be taken at an angle of 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60 degrees. It includes ear images with partial occlusion, trivial occlusion and normal occlusion.[11]

V. Proposed Method

In this paper we propose a method which uses python libraries such as keras, matplotlib, numpy opencv and os for preprocessing database and apply neural network to extract feature and predict the correct label[22]. The steps will be as follows:-

- Digital Image Processing
- Dividing images in training and testing sets
- Using keras layers for feature extraction(deep learning)
- Predicting the label of an image.

A. Digital Image Processing

Digital image processing is a technique which analyze and manipulate the digitalized image in order to improve the quality of digital image and to help detect features of an image with much increased accuracy. In this model we have used opencv, numpy, os library from pyhton for digital image processing [22]. Steps taken to perform Digital Image Processing are:-

- Loading database into model
- Resizing images
- Converting images to grayscale
- Image smoothning and edge detection

Loading database

Loading of database is done using a function which input images from a particular database into the model and output an array of images and corresponding labels. Path of the folder is input to the function.

Fig. 3: Database 1 [11]

Fig. 4: Database 2 [11]

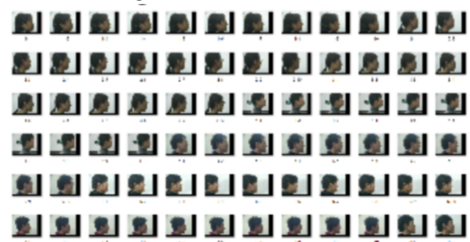


Fig. 5: Database 3 [11]

Resizing Images

Resizing of images can be done in the same function as that of loading database or one we can use different function as well. For resizing we have used an inbuilt function cv2.resize() in opencv library in python. It has required attributes such as src(source),and dsize(desired size of an output image) and optional attributes such as fx(scale factor along the horizontal axis), fy(scale factor along the vertical axis), and interpolation.[20]

Conversion to grayscale

A huge amount of information is given by edges being robust to lighting changes. Pixel intensity and position, both these factors are used to compute the pixel distance. The same thing can compute with color too[20].

Image Smoothing and Edge Detection

Edge detection is one of the most important steps in the analysis of an image as it provides structural and topological information of relevant object in an image [21]. Various tasks like region segmentation and object identification [23] heavily depend on efficiently detected edges. Like noise, edges are also known to be characterized by high frequencies [24]. Due to this, edge detection becomes quite a tedious task since the chances of mistaking noises as edges are high [25].

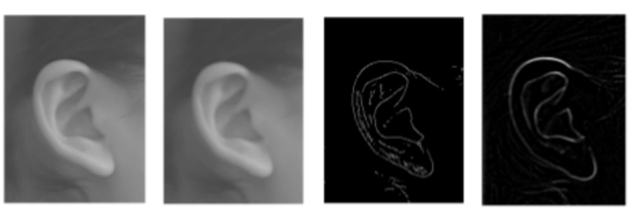
Gaussian Filter normally written as:-
$G(x_i, y_j) = 1/(2\pi\sigma^2)e^{(}(x_i^2 + y_i^2)/(2\sigma^2))$
has a standard deviation σwhich determines the width of the filter as well as the outcome of the smoothed image. The value of σ and the frequency band of the Gaussian filter are directly related.

Canny Edge Detector: Canny's three edge detection criteria are as follows[30]:

- Good detection: The probability of failing to mark true edge points and that of falsely marking non edge points show be low, whereas the signal-to-noise ratio should be as high as possible.
- The points marked as edge points by the operator should approximate to the center of the real edge as much as possible.
- False boundary responses are suppressed to the maximum and also, single edge yields fewer multiple responses.

First-order partial derivative finite difference in 2x2 neighborhoods are used for calculating the gradient amplitude and direction of smooth data array I(x, y). Px[i, j] and Py[i, j] respectively refer to array of x and y directions partial derivative.

Px[i, j]= (I[i+1, j]-I[i, j]+ I[i+1, j+1]-I[i, j+1])/2,
Py[i, j]= (I[i, j+1]-I[i, j]+ I[i+1, j+1]- I[i+1, j])/2.
Pixel gradient magnitude and gradient direction are obtained with the coordinate conversion formula involving Cartesian coordinates to polar coordinates. The gradient amplitude calculated with the second order norm is
$M(i, j) = \sqrt{f_x(i, j)^2 + f_y(i, j)^2}$
And gradient direction is
$\theta(i, j) = arctan(f_y(i, j)/f_x(i, j))$
After determining the gradient direction θ[i, j], the edge direction can be divided into eight directions with 45 degrees. The adjacent pixels of this pixel gradient direction can be found out using the gradient of direction,[30].



(a) Grayscale Image (b) Gaussian Filtered Image (c) Canny Edge Detector (d) Laplacian edge detector

B. Feature Extraction

After pre-processing, we two third of the image dataset is taken for training and one third is taken for testing. Images are then reshaped to input to our convolution neural network model. For example Google photos organizes all your photos and classify them according to date, person, time, etc. and make them searchable. But how are photos organized when they as labeled with non-descriptive labels. Image classification provides text label to images. But even with huge number of training examples, computers cannot extrapolate enough meaningful information from raw pixels alone to easily identify an object in an image (ear in this case). A model which can identify objects just like human do is convolution neural network.[33]

CONVOLUTIONAL FILTER:

A convolution filter is an n*n grid of matrix (for example 3*3) that is applied to an image in different ways to generate a feature map, a map of all feature crosses or combinations in an image. Filters form a feature hierarchy. Convents use filters t0 build representations that become progressively easier to classify. From pixels to texture, from texture to shape and from shapes to objects.[33]

CONVOLUTIONAL LAYER:

Image is matrix of pixel values. A convolutional neural network is able to successfully capture temporal and special dependencies in an image through the application of required filter. CNN perform better fitting to the image database due to reusability of weights and reduction in number of parameters. Hence CNN can be trained to understand details of image better. The filter moves right with specified stride value till it passes the complete width of the matrix of pixels. As it reaches the edge, it hopes down to the beginning and repeats the process till the entire matrix values are covered. In case of multiple channels i.e. RGB, the filter has the same depth as that of the input image. Objective of CNN is to extract features such as edges (high level features) from an image. It needs not to be limited to only one convolution layer. Conventionally one convolutional layer is responsible for extracting features such as edges, color, gradient, etc. While added multiple layers recognizes more high-level feature which gives the wholesome idea of an image. Usually dimensionality of the images input to the layers are reduced significantly thus results loosing important information from the input image. To solve this problem, we perform padding on the matrices i.e. we add extra layers to the input matrix so that the dimensionality of the matrix does not change significantly. When the dimensions of the convolved matrix is same to that of the input matrix, padding done is known as same padding. When matrix is convolved with no padding then the padding applied is known as valid padding.[33]

MAXPOOL LAYERS:

It is also known as noise suppressant layer discards the noisy activations and also performs dimsionality reduction. Convolution layer and maxpool layer together form ith layer in Convolutional Neural Network. Depending on the complexities of the image the number of convolution neural network can be increased to have better understanding of the sophistication of image at the cost of more computational cost.[33]

FULLY CONNECTED LAYER:

It converts image matrix into a column vector the flattened output is applied to feed forward neural network and back-propagation is applied at every iteration. Over the series of epochs, the model is able to distinguish between features in images and certain low-level features of images and classify them using softmax classifier.[33]

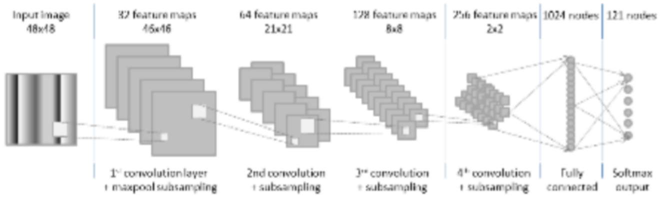


Fig. 7: Convolutional Neural Network

VI. RESULT AND DISCUSSION

The proposed identification techniques is tested on University of Science and Technology (USTB), Bejing ear database [11]. We have divided the training and testing dataset in the ratio of 2:1. In Table I, accuracies i.e recognition rate of proposed method for different databases provided by USTB is shown.

TABLE I: Recognition rate of Proposed Method

Database	No. of Subjects	Recognition rate
USTB Database1	186	93.345%
USTB Database2	308	65.234%
USTB Database3	786	91.281%

In the Table II, accuracy of proposed technique along with the accuracies of different techniques are shown and compared.

TABLE II: Recognition Rate of different Methods

Classifiers	No. of Subjects	Recognition rate
Geometrical features[3]	144	85%
Principal Component[31]	285	71.6%
Neural Network[32]	188	93%

Reasons for low accuracy are:

- Important information such as mole or scar in ear is lost by using canny edge detector.
- One of the main reasons for less accuracy is memory exhaustion. We had to reduce the size of the images considerably, because of which accuracy is decreased significantly.
- In Database 2, the different orientations of the ear have been emphasized. As a result, when the images from this database are divided into training and testing images, our algorithm is not able to map a direct relationship. This leads to the lower accuracy score in Database 2.
- However in Database 1 and Database 3, the orientation of different images have not been warped. This provides a direct relationship between training and testing images and hence improves the accuracy.

VII. CONCLUSION

In this paper, we have presented an approach towards automated human ear recognition system. We have used the basic deep learning concept of convolutional neural network. This approach consists of four stages such as digital processing of an image, dividing images into training and testing datasets, applying Keras sequential CNN model to the dataset, and finally classifying images using softmax classifier. Recognition rates of proposed method on the two database 1 and database 3 are 93.345% and 91.281% which proves that our technique is efficient and reliable also it gives higher recognition rate than similar algorithms using same databases.

ACKNOWLEDGMENT

The authors of this paper would like to thank University Of Science and Technology (USTB), Bejing for providing their ear databases for carrying out this research.

REFERENCES

[1] Rehman A, Saba T. Document skew estimation and correction: analysis of techniques, common problems and possible solutions. Appl Artificial Intel 2011; 25: 769-787.
[2] Iannarelli A. Ear identification, forensic identification series. Paramont Publ Company 1989.
[3] Mu Z, Yuan L, Xu Z, Xi D,Qi S. Shape and Structural Feature Based Ear Recognition. Lecture Notes in Computer Science (LNCS); 2004.3338, p. 663-670.
[4]] Chang, K., Bowyer, K.W., Sarker, S., Victor, B., "Comparison and Combination of Ear and Face Machine Image in appearance –Based Biometrics', IEEE Transaction on pattern Analysis and machine Intelligence, vol. 25, no, 9 (2003).
[5] Hurley DJ, Nixon MS, Carter JN. Force field energy functional for ear biometrics. Computer Vision and Image Understanding 2005; 98: 491-512.
[6] Hossein Nejati, Li Zhang,Terence Sim, Elisa Martinez-Marroquin, Guo Dong, Wonder ears: Identification of identical twins from ear images. 2012
[7] M. I. Saleh, W. C. L., and P. Schaumont. Eigen-face, Eigen-ears. using ears for human identification. Master's thesis, Virginia Polytechnic Institute and State University, 2007.
[8] Md. Mahbubur Rahman, Md. Rashedul Islam, Nazmul Islam Bhuiyan, Bulbul Ahmed, Md. Aminul Islam. Person Identification Using Ear Biometrics
[9] Lynn Meijerman, Sarah Sholl, Francesca De Conti, Marta Giacon, Cor van der Lugt, Andrea Drusini, Peter Vanezis, George Maat. Exploratory study on classification and individualisation of earprints.
[10] Mohd Shafry Mohd Rahim1*, Amjad Rehman1, Fajri Kurniawan1, Tanzila Saba. Ear biometrics for human classification based on region features mining.
[11] Tejas. V. Kandgaonkar1 , Rajivkumar. S. Mente2, Ashok. R. Shinde3 and Shriram. D. Raut4. Ear Biometrics: A Survey on Ear Image Databases and Techniques for Ear Detection and Recognition
[12] Shubham Sanjay Mor, Shivam Solanki, Saransh Gupta, Sayam Dhingra, Monika Jain, Rahul Saxena. Handwritten Text Recognition: With Deep Learning and Android
[13] Benjio, Yoshua, Learning deep architectures through AI
[14] Yann LeCun, Yoshua Bengio. Convolutional Networks for Images, Speech, and Time-Series
[15] Fares Jalled. Object Detection Using Image Processing. Moscow Institute of Physics and Technology, Department of Radio Engineering and Cybernetics Ilia Voronkov, Moscow Institute of Physics and Technology, Department of Radio Engineering and Cybernetics
[16] D. J. Hurley, B. Arbab-Zavar, and M. S. THE EAR AS A BIOMETRIC. Nixon University of Southampton.
[17] Elizabeth J. Ragan, Courtney Johnson, Jacqueline N. Milton, and Christopher J. Gil. Ear biometrics for patient identification in global health: a cross-sectional study to test the feasibility of a simplified algorithm.

[18] G.N. Rutty, A. Abbas, and D. Crossling. Could earprint identification be computerised? an illustrated proof of concept paper. International Journal of Legal Medicine, 119(6):335-343, 2005.
[19] S. Muthuselvi and P. Prabhu. Digital image processing techniques - A Survey.
[20] Bhumika Gupta, Ashish Chaube, Ashish Negi, Umang Goel. Study on Object Detection using Open CV - Python.
[21] Jonathan, M. Black Ledge. Digital Image Processing Mathematical and Computational Methods. (2005)
[22] K. R. Srinath, Pragati Mahavidyalaya. Python – The Fastest Growing Programming Language
[23] Mukesh, R.K., Jaiswal, A.K. and Rohini, S. (2013) Performance Analysis of Adaptive Canny Edge Detector Using Bilateral Filter. International Journal of Innovative Research in Computer and Communication Engineering, 1, 930-936.
[24] Lim, J.S. (1990) Two Dimensional Signal and Image Processing.
[25] Marr, D. and Hildreth, E. (1980) Theory of Edge Detection. Proceedings of the Royal Society of London B, 207, 187- 217.
[26] Amoako-Yirenkyi, P., Appati, J.K. and Dontwi, I.K. (2016) Performance Analysis of Image Smoothing Techniques on a New Fractional Convolution Mask for Image Edge Detection. Open Journal of Applied Sciences, 6, 478-488.
[27] Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. The Journal of Physiology 195(1), 215–243 (1968)
[28] S. Haykin, Neural Networks, A Comprehensive Foundation. New York: Macmillan, 1994.
[29] Steve Lawrence, C. Lee Giles, Ah Chung Tsoi, and Andrew D. Back. Face Recognition: A Convolutional Neural-Network Approach.
[30] Jun Li1, Sheng Ding. A Research on Improved Canny Edge Detection Algorithm.
[31] K.H. Pun, Y.S. Moon. Recent advances in ear biometrics
[32] 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, September 2003, vol. 25, no. 9.
[33] Simonyan, Zisserman, 2015, K. Simonyan, A. Zisserman Very Deep convolutional networks for large-scale visual recognition Proceedings of ICLR (2015)