

---

# Multimodal Biometrics Data Analysis for Gender Estimation Using Deep Learning

Shivanand Sharanappa Gornale<sup>1</sup>, Abhijit Patil<sup>1, \*</sup>, Kruti Ramchandra<sup>2</sup>

<sup>1</sup>Department of Computer Science, Rani Channamma University, Belagavi, Karnataka, India

<sup>2</sup>Department of Computer Science, Jain University, Bangalore, India

## Email address:

shivanand\_gornale@yahoo.com (S. S. Gornale), abhijitpatil05@gmail.com (A. Patil), krutisrinivas85@gmail.com (K. Ramchandra)

\*Corresponding author

## To cite this article:

Shivanand Sharanappa Gornale, Abhijit Patil, Kruti Ramchandra. Multimodal Biometrics Data Analysis for Gender Estimation Using Deep Learning. *International Journal of Data Science and Analysis*. Special Issue: *Multimodal Biometric Data Analysis*.

Vol. 6, No. 2, 2020, pp. 64-68. doi: 10.11648/j.ijdsa.20200602.11

**Received:** December 9, 2019; **Accepted:** December 16, 2019; **Published:** May 29, 2020

---

**Abstract:** In the recent past with the rapid growing technology security problem is ubiquitous to our daily life pertinent to it, now a day the usage of biometrics is becoming inevitable. Correspondingly, the field of biometrics has gained tremendous acceptance because of its individualistic and authentication capabilities. In many practical scenario the multimodal-based gender estimation will helps to increase the security and efficiency of other biometrics system. Likewise, in contrast to it uni-modal biometric, the multimodal biometrics system would be very difficult to spoof because of its multiple distinct biometrics features. Gender identification using biometrics traits are mainly used for reducing the search space list, indexing and generating statistical reports etc In this paper, a robust multimodal gender identification method based on the deep features are computed using the off-the-shelf pre-trained deep convolution neural network architecture based on AlexNet. The proposed model consists of 20 subsequent layers which contain different window size of convolutional layers following with fully connected layers for feature extraction and classification. Extensive experiments have been conducted on a homologous SDUMLA-HMT (Shandong University Group of Machine Learning and Applications) multimodal database with 15052 images. The proposed method achieved the accuracy of 99.9% which outperforms the results noticed in the literature.

**Keywords:** AlexNet, Biometrics, Convolutional Neural Network, Deep Neural Network, Gender Estimation, Multimodal, SDUMLA-HMT

---

## 1. Introduction

Gender estimation is one such auxiliary information for determining gender of a person based upon the biometrics features. A very few researchers have worked on unimodalities for gender classification like fingerprints [1-5], iris [6-9, 32], palmprint [10-14], face [15-19], speech [20-23] and gait [27] and have obtained the competitive results.

Moreover, any uni-modal biometric systems suffers from a variety of problems including [29] distorted data, intra-class variations, a constrained level of freedom, non-comprehensiveness, spoofs, and intolerable error rates and so forth. However; these confines can be mitigated by fusing different uni-modalities. Intuitively multimodal systems are more reliable and viable solution as multiple independent characteristics of modalities are fused together

which interns can provides higher accuracy rather than any uni-modal system. Multi-modal biometrics based gender identification represents an emerging trend and practically have many real-world applications like human-computer interface, security, gender-based advertisement, forensics and surveillance etc [30].

In this work, we present a gender identification system using deep neural network architecture from multi-modalities like Face, Iris, and Fingerprints. The proposed work divulges effective and real time method for multimodal gender estimation by network fine tuning strategy which interns provides competitive result. The paper is organized as follows section 2 contains the literature review; section 3 is focused on proposed methodology. In section 4, database is described and in section 5 results are discussed and finally, the conclusions are drawn in Section 6.

## 2. Related Work

Preceding research has shown that the possibility of authenticating an individual from their respective modalities, However from the related literature, it is observed that a limited work has been carried out for identifying a person's gender using their multi-modal biometrics. This section reviews the studies which have been reported on gender identification using multi-modal biometric, Xiong Li *et al.* [24] have performed multimodal based gender identification by combining local binary pattern and Bag of words features based on decision level fusion on the face and fingerprint traits of the internal database of 397 volunteers from Han nationality and obtains an accuracy of 94% using Bayesian Hierarchical model. Mohamed A *et al.* [25] have performed multimodal based gender identification by combining features, Eigenvalue, Syntactic Complexity, Response length, shallow and deep syntax and mean, heart rate max-min difference features are combined on five different traits i.e. visual linguistic, physiological, thermal and acoustic traits. A database of 51 males and 53 females were utilized and obtained an overall accuracy of 80.6% using decision tree classifier. Abdenour H *et al.*[26] have utilized three different unimodal databases namely CRIM, vidTimit and Cohn-Kanade dataset of face and video which are further combined to form a multimodal dataset (face and video) by using local binary pattern obtained an overall accuracy of 96.3%by support vector machine. Caifeng Shan *et al.* [27] performed multimodal gender identification by combining the face and gait modalities from CASIA Gait b dataset from which frontal face image is extracted from 119 subjects which are further combined with gait videos. AdaBoost based face detector and Background subtraction feature extraction techniques are implemented respectively and using the support vector machine classifier an overall accuracy of 97.2% is achieved. S S Gornale *et al.*[28] have utilized the binarized statistical image features along with multi-block local binary pattern features on face, iris, and fingerprint traits of SDUMLA-HMT dataset. Then by Support Vector Machine classifier the highest accuracy of 99.8% accuracy is yielded.

Further, S. S. Gornale *et al.*[29] have performed synthesization of features along with fusion of classifiers on Face, Fingerprints and Iris images of SDMULA-HMT and KVK-Multimodal datasets an highest accuracy of 99.9% and 99.8% is noted respectively.

From the literature it is observed that, major focus has been given on the handcrafted features and limited work is presented on deep learning algorithm there is still scope to develop a multi-modal system for gender classification based on deep convolution method.

## 3. Proposed Methodology

Deep learning is well known and special kind of neural network which processes the data in grid like topology. The CNN accepts the whole image as input and automatically extracts the meaningful features which are required for efficient image classification. Deep learning is a branch of machine learning algorithm which have been used in solving many classical artificial intelligence problems, such as image classification, medical imaging for skin disease identification, cancer cell detection and natural language processing [31] etc.

In this work, we have trained AlexNet from the scratch on SDUMLA multimodal images for gender identification. AlexNet was initially designed by Alex Krizhevsky [32] in 2012 to solve the 1000 class image classification problem in (ILSVRC-2012). We have fine tune the AlexNet because AlexNet is a good model to be trained with limited amount of data to train the net directly and AlexNet is a deep network that can extract many abstract and high scale multimodal features.

Gender classification using multimodal biometrics generally involves three steps namely pre-processing, feature extraction and classification. The initial task is to do pre-processing, as each image is resized to 224\*224 grey scale images. Since the input requirement for the AlexNet is a fixed size of 224\*224 color image, we empirically converted grey image to color by duplicating the grey channels. Feature computation step deals with extraction of features information using meaning full feature map. Lastly, is evaluated with different binary classifiers.

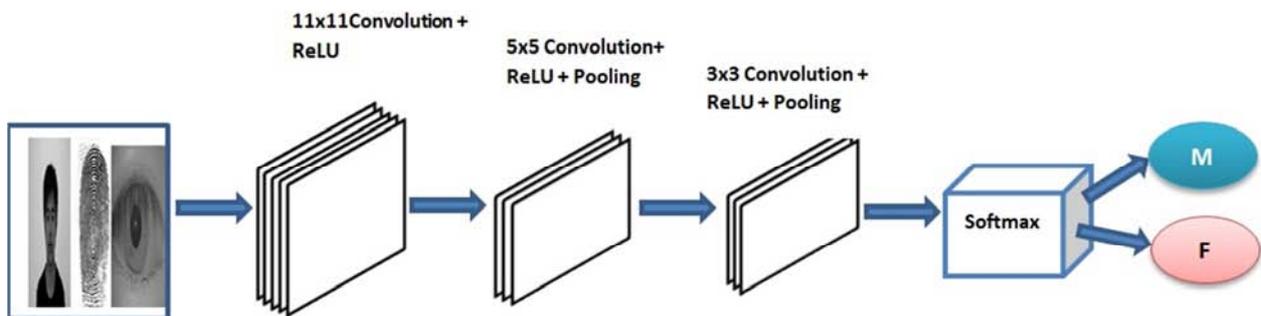


Figure 1. Proposed Methodology.

AlexNet is simple feed forward neural network architecture with multiple layers in subsequent connected denser neurons with successive hidden layers. In this work, we performs the fine tuning of AlexNet architecture by configuring the network with activation function is rectified linear unit (ReLU)

whose operation is non-linear in nature. With ReLU activation function the training time is seen significantly decreased with the network hyperbolic non-linearly tangent. Where  $k$  indicates the current layer and  $u^k$  is the input of ReLU.

$$\text{ReLU}(u^k) = \text{Max}(U^k, 0) \quad (1)$$

Generally, there are 3 fully connected layers in the network namely FC6, FC7 and FC8, where FC6 and FC7 layers have 4096 neurons and FC8 has 1000 neurons. An image can be presented to any connected layer which results in a feature map. During training we have initialized with FC7, and then with propagation the resultant feature map is convolved with set of weights called filters. Likewise different filters sets were applied over different features map however same kind of filters has been shared among all neurons. A pooling layer reduces this spatial content of the representation output by a convolution layer resulting in a reducing number of the parameters and the number of computations within the network. Pooling work independent on every depth slice of its input usually by applying max operation. Here  $w_k$  is a weight of the layer,  $x_{k-1}$  represents the output of the  $k-1$  layer and  $b$  is the bias which is initialized to 0.

$$Y = (W^k \cdot X^{k-1} + b^k) \quad (2)$$

Overfitting is one of the most common problems in neural networks which can be prevented by dropout and Batch Normalization. The Dropout is based on the idea of leaving a certain neurons in each iteration to train the network and the batch normalization is applied which intern reduces covariate shifts of the over fitted data and makes training faster. Gradient descent (GD) is the learning process of neural network is based on the searching for the combination of learning parameters that assures the lowest errors for the loss functions. The gradient of the loss function is used to find the direction towards the weighted vector

$$W_i : W_i - \alpha \frac{\delta}{\delta W_i} L \quad (3)$$

Where  $\alpha$  represents the learning rate and 'Wi' are weights of  $\delta$  layers, during the training the weights are tends to increases in each such iterations. The stochastic gradient descent (SGD) is considered to be as optimization method which automatically updates its corresponding weight at each iterations, The learning rate was 0.003 is initialized in all cases, it may seems to small but accordingly to the dataset size, the weight updates were to control the better training, Basically here weights are been transferred instead of sparse sparsefying the weight which results in higher accuracy.

## 4. Database

In this work, we have used publicly available SDUMLA-HMT standard dataset which is collected and maintained by machine learning and application lab of Shandong University [33]. The dataset includes real multimodal data from 106 individuals out of which 59 are male volunteers and 47 are female volunteers. The dataset contains face images which were collected from different poses, expressions and accessories. An Iris images were captured giving proper direction to volunteers, images from the both eyes are collected. Likewise, fingerprint dataset

images are acquired with FT-2BU sensors, from each such subject images of both-hand thumb, index and middle finger were collected by giving prerequisite directions. Thus in total the database consists of 15052 multimodal images. A sample dataset of male and female images are shown in figure 2.

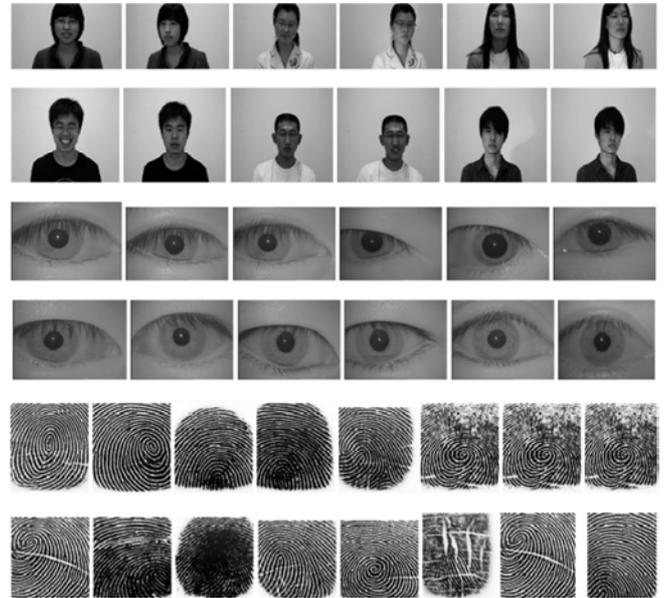


Figure 2. Sample of Database.

## 5. Results and Discussion

In this study, we have trained the AlexNet from the scratch on multimodal biometrics images which are further been validation over different binary classifiers i.e. Naive Bayes, K-Nearest Neighbour, Support Vector Machine and Decision Tree classifier on publically available SDUMLA-HMT database.

The algorithmic steps are as follows:

Input: Multimodal images.

Output: Gender Identification.

Step 1: Input the images from database.

Step 2: Image pre-processing is carried out.

Step 3: Fine tuning of pre-trained AlexNet model is performed.

Step 4: Resultant feature map is trained and tested.

Step 5: Gender Classification is performed by different binary classifiers.

The proposed architecture consists of 20 layers, which includes five convolutions layers with different convolution window sizes i.e. of  $11 \times 11$ ,  $5 \times 5$  and  $3 \times 3$  pixels, followed by two fully connected layers and a softmax layer. The first layer is the input layer which accepts the image with  $250 \times 250 \times 3$  pixel image dimension. The subsequent layers are convolution convlayer1 which have a window size of the  $11 \times 11$  pixels with 96 filters and then next layer is a ReLU, which have the non-linear activation function shadowed by cross normalization layer and Max-pooling layers

Likewise, then ReLU is tailed with a convolution i.e. convlayer2 which have a window size of the  $5 \times 5$  pixels with 256 filters and an intermediate pooling is sub-sampled with

another cross normalization layer and Max-pooling layers, followed with a convolution convlayer3 which have a window size of the 3×3 pixels with 384 filters and ReLU function. Follow by convolution convlayer4 which have a window size of the 3×3 pixels with 384 filters and another ReLU function. Further, with convolution convlayer5 which have a window size of the 3×3 pixels with 256 filters and another ReLU function an intermediate pooling is down-sampled

Next layer is a fully connected layer1 (FC7) with 4096 neurons, by ReLU activation function followed by the dropout layer is used to reduce over fitting in neural network, fully connected layer2 (FC8) compel with 1000 class classifier (neurons), The last softmax layer has 2 neurons that identify the class membership which predicts whether the multimodal biometrics images belong to male or female.

**Table 1.** Confusion matrix of Different Features on SDUMLA HMT Database.

Classifier	Female	Male	Average accuracy of Male (%)	Average accuracy of female (%)	Accuracy in %
Naïve Bayes	5467	1628	81.9	80.6	81.2
	1207	6750			
KNN Cityblock	6554	83	98.2	99	98.7
	120	8295			
KNN Euclidean	6556	89	98.2	98.9	98.6
	118	8289			
SVM	6673	0	100	99	99.9
	1	8378			
Decision. Tree	6570	101	98.5	98.8	98.6
	104	8277			

The results are depicted in the form of confusion matrix in Table 1. With the Table 1 it is witnessed that the highest accuracy of 99.9% is observed using the SVM classifier and the lowest accuracy of 81.2% is noted by using Naive Bayes classifier. Further, by KNN classifier K=3 which is empirically fixed, the City-Block distance an accuracy of 98.7% is noted. With Decision Tree classifier and K-NN classifier Euclidean distance a similar accuracy of 98.6% is observed respectively.

## 6. Conclusion

Gender identification has been explored by several researchers, in the past decades major focus is been towards handcrafted features based unimodal methods and its implementation. In this paper a robust algorithm has been deliberated which automatically perform multimodal based gender identification using fine-tuned AlexNet architecture. In this work, deep convolution method is considered to get robustness in gender determination and achieved an accuracy of 99.9% on SDUMLA-HMT multimodal biometric database.

## References

- [1] S. S. Gornale, Abhijit Patil and Veersheety C. "Fingerprint Based Gender Identification Using DWT and Gabor Filters", International Journal of Computer Applications Vol. 152 Issue 4, page. 34-37 Oct-2016.
- [2] Kruti R, Abhijit Patil and S. S. Gornale, "Fusion of Features and Synthesis Classifiers for Gender Classification using Fingerprints", International Journal of Computer Sciences and Engineering (IJCSE) ISSN: 2347-2693, Vol-7 Issue-5, May 2019.
- [3] Kruti R, Abhijit Patil and S. S. Gornale,"Fusion of Local Binary Pattern and Local Phase Quantization features set for Gender Classification using Fingerprints", International Journal of Computer Sciences and Engineering, vol. 7, issue 1, pp. 22-29 Feb -2019.
- [4] S. S Gornale, Malikarjun Hangarge, Rajmohan P, Kruthi R, "Haralick Feature Descriptors for Gender Classification Using Fingerprints: A Machine Learning Approach", International Journal of Advanced Research in Computer Science and Software Engineering", Volume 5, Issue 9, ISSN: 2277 128X, PP-72-78 September 2015.
- [5] K. S. Arun and K. S. Sarath," A machine learning approach for fingerprint based gender identification, 2011 IEEE Recent Advances in Intelligent Computational Systems, ISBN: 978-1-4244-9478-1, pp: 22-24 Sept. 2011.
- [6] Zhang H., Sun Z., Tan T., Wang J. (2011) "Ethnic Classification Based on Iris Images", In: Sun Z., Lai J., Chen X., Tan T. (eds) Biometric Recognition. CCBR 2011. Lecture Notes in Computer Science, Vol 7098. Springer, Berlin, Heidelberg 2011.
- [7] S. Aryanmehr, Fanny Dufoss, Farsad Zamani Boroujeni. "CVBL IRIS Gender Classification Database Image Processing and Biometric Research, Computer Vision and Biometric Laboratory (CVBL)", 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC). 2018.
- [8] Shruti Nagpal, Mayank Vatsa; Richa Singh; Afzel Noore; Angshul Majumdar Maneet Singh, "Gender and ethnicity classification of Iris images using deep class-encoder", 2017 IEEE International Joint Conference on Biometrics (IJCB), DOI: 10.1109/BTAS.2017.8272755, ISSN: 2474-9699, Denver, CO, USA. 2017.
- [9] Tapia J. E., Perez C. A., Bowyer K. W. (2015) "Gender Classification from Iris Images Using Fusion of Uniform Local Binary Patterns. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science, vol. 8926. Springer, Cham.
- [10] W Ming and Y Yuan, "Gender Classification Based on Geometrical Features of Palmprint Images", The Scientific World Journal vol. 2014. Article Id: 734564, pp.: 7 2014.

- [11] S. S Gornale, Abhijit Patil and Kruti R, "Fusion Of Gabor Wavelet And Local Binary Patterns (LBP) Features Sets For Gender Identification Using Palmprints", *International journal of Imaging Science and Engineering*, Vol. 10, Issue No. 2, April 2018.
- [12] Zhihui Xie, Zhenhua G and Chengshan Q, "Palmprint gender classification by Convolution neural network", *IET Computer Vision*, Vol 12, Issue 4 pp: 476-483, 2018.
- [13] Shivanand S Gornale, Abhijit Patil and Mallikarjun Hangarge, "Binarized Statistical Image Feature Set for Palmprint based Gender Identification", *Book of Abstract of First International Conference on Machine Learning, Image Processing, Network Security and Data Sciences (MIND-2019)* pp.: 61, held on 3rd-4th march 2019 at National Institute of Technology, Kurukshetra, India.
- [14] Shivanand S Gornale, Abhijit Patil, Mallikarjun Hangarge and Rajmohan P, "Automatic Human Gender Identification using Palmprint". *Smart Computational Strategies: Theoretical and Practical Aspects*. Springer, Singapore, Online ISBN 978-981-13-6295-8, Print ISBN 978-981-13-6294-1, 22-march-2019.
- [15] Paul Viola and MichaelJ. Jones. 2004. Robust Real-Time Face Detection. *International Journal of Computer Vision* Vol. 57, Issue 2 pp. 137–154 (2004).
- [16] Len Bui; Dat Tran; Xu Huang; Girija Chetty, "Classification of gender and face based on gradient faces", 3rd European Workshop on Visual Information Processing, 10.1109/EuVIP.2011.6045544 Paris, France -July 2011.
- [17] Arnulf B. A. Graf and Felix A. Wichmann, "Gender Classification of Human Faces", *Biologically Motivated Computer Vision*, eds. H. H. B'ulthoff, S.-W. Lee, T. A. Poggio and C. Wallraven, LNCS 2525, pp.: 491-501, 2002, Springer Verlag, Heidelberg.
- [18] S. Ravi, S. Wilson, "Face Detection with Facial Features and Gender Classification Based On Support Vector Machine", 2010 Special Issue - *International Journal of Imaging Science and Engineering*.
- [19] Abul Hasnat; Santanu Haider; Debotosh Bhattacharjee; Mita Nasipuri, "A proposed system for gender classification using lower part of face image.", 2015 International Conference on Information Processing (ICIP), DOI: 10.1109/INFOP.2015.7489451, ISBN: 978-1-4673-7758-4, Pune, India.
- [20] Hadi Harb, Liming Chen, "Voice-Based Gender Identification in Multimedia Applications", *Journal of Intelligent Information Systems*, vol. 24: pp.: 179, 2005.
- [21] Rafik Djemili; Hocine Bourouba; Mohamed Cherif Amara Korba, "A speech signal based gender identification system using four classifiers", 2012 International Conference on Multimedia Computing and Systems, ISBN no. 978-1-4673-1520-3, 2012.
- [22] Daniel Reid, Sina Samangoeei, Cunjian Chen, Mark Nixon, and Arun Ross. 2013. So. biometrics for surveillance: an overview. *Machine learning: theory and applications*. Elsevier pp: 327–352, 2013.
- [23] L Walawalkar, M. Yeasin, A. Narasimhamurthy, and R. Sharma. Support vector learning for gender classification using audio and visual cues. *International Journal of Pattern Recognition and Artificial Intelligence*, 17 (3): 417–439, 2003.
- [24] X. Li, X. Zhao, H. Liu, Y. Fu and Y. Liu, "Multimodality gender estimation using Bayesian hierarchical model," 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX, 2010, pp.: 5590-5593. doi: 10.1109/ICASSP.2010.5495242.2010.
- [25] Mohamed Abouelenien, Veronica Perez-Rosas, Rada Mihalcea, and Mihai Burzo, "Multimodal Gender Detection", in *Proceedings of ICMI '17*, Glasgow, UK, November, pp.: 302-311. ISBN no. 978-1-4503-5543-8.2017. 2017.
- [26] A. Hadid and M. Pietikainen, "Combining motion and appearance for gender classification from video sequences", 2008 19th International Conference on Pattern Recognition, Tampa, Florida, 2008, pp: 1-4. doi: 10.1109/ICPR.2008.4760995.
- [27] Caifeng Shan and Shaogang Gong and McOwan, Peter W, "Learning Gender from Human Gaits and Faces", 2007 IEEE Conference on Advanced Video and Signal Based Surveillance Washington, DC, USA. pp: 505-510, isbn 978-1-4244-1695-0.
- [28] Abhijit Patil, Kruthi R and Shivanand Gornale, " Analysis of Multi-modal Biometrics System for Gender Classification Using Face, Iris and Fingerprint Images", *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, Vol. 11, No. 5, pp. 34-43, 2019. DOI: 10.5815/ijigsp.2019.
- [29] Shivanand S Gornale, Abhijit Patil and Kruthi R, "Multimodal Biometrics Data Based Gender Classification using Machine Vision", *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Vol-8 Issue-11, September 2019.
- [30] Maneet Singh, Richa Singh, Arun Ross, A comprehensive overview of biometric fusion, *Information Fusion*, Volume 52, 2019, pp. 187-205, ISSN 1566-2535, <https://doi.org/10.1016/j.inffus.2018.12.003>.
- [31] Nour Eldeen M. Khalifa, Mohamed Hamed N. Taha and Hamed Nasr Eldin T. Mohamed, "Deep Iris: Deep Learning for Gender Classification Through Iris Patterns", *Acta Informatica Medica*, Vol. 22 Issue 2, pp-96-102, June 2019, doi: 10.5455/aim.2019.27.96-102.
- [32] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems*. Curran Associates Inc.; 2012: 1097-1105.
- [33] Yilong Yin, Lili Liu, and Xiwei Sun, "SDUMLA-HMT: A Multimodal Biometric Database", *The 6th Chinese Conference on Biometric Recognition (CCBR 2011)*, LNCS 7098, pp.: 260-268, Beijing, China, 2011.