**Application of CNN to counter spoofing attacks in biometric identification systems**

Bakhtiyor Abdukadirov a), Fakhriddin Toshboltaev, Shakhboz Abdullaev

Fergana state university, Fergana, Uzbekistan

a) Corresponding author: [bakxtiyor.uz@gmail.com](mailto:bakxtiyor.uz@gmail.com)

**Abstract.** The article proposes a method for detecting a situation in which, during face authentication, one person disguises himself as another by replacing his face with a photograph or a displayed video. The facial feature vector is extracted using a convolutional neural network. The classification is implemented using a support vector machine. The method can be input either one or a group of frames.

**INTRODUCTION**

Facial recognition systems are currently used in many areas of human activity, such as border control, banking payments and access to services, and mobile applications. Biometric technologies for human recognition have recently been actively developing and are finding application in access control and management systems, human traffic analysis systems, and systems for searching for intruders in crowded areas. Biometric systems for human recognition allow for the identification of a person by unique physical characteristics, such as fingerprints, facial images, and voice [1]. Biometric systems are used where a high level of security is required due to their advantages over conventional control methods: passports, special plastic cards and tokens, PIN codes, etc. [2]. Technologies based on facial images are particularly attractive. Compared to other biometric control methods, technologies based on facial recognition do not require direct contact with the equipment.

Facial recognition is an easy task for humans, but it's not so trivial for algorithms. The human brain identifies image features such as edges, lines, and movement, combines the resulting information into templates, and uses these templates to make identifications.

***Image recognition algorithms.*** Image recognition algorithms operate similarly. They extract significant characteristics from an image, present them in a convenient form, and classify the resulting data.

Image recognition quality is affected by negative factors such as changes in room lighting, the presence of multiple faces in the image, head rotation, changes in a person's emotions, and other possible facial changes (aging, makeup, etc.). Therefore, it is first necessary to reduce the image to a template. For example, by identifying each individual face, normalizing lighting, normalizing image orientation, etc. Thus, the process of personal identification can be divided into several main stages:

*- image registration and normalization,*

*- feature extraction,*

*- proximity calculation,*

*- decision making.*

***Detection of false attacks.*** Although many researchers [3, 4] have been working on facial recognition for many years, a number of challenges remain to be addressed. One such challenge is countering spoofing attacks, i.e., attempts to substitute biometric characteristics and deceive a biometric system by presenting a fake identification object to the sensor [5]. Despite their simplicity, such attacks are generally quite successful, and effective algorithms capable of countering them are still lacking. However, a positive trend is that researchers are now beginning to focus on this issue [6-8].

The large-scale implementation of biometric identification systems requires assessing their resilience to various types of attacks [9, 10].

A current area of research in this area is assessing the feasibility of constructing an attacking biometric image that allows access on behalf of the victim while, in a sense, not revealing the fact of the attack [11, 12]. That is, for a person, such an image should not be associated with the object being attacked, and the recognition system should identify it as the object being attacked.

To counter spoofing attacks, methods using additional sensors or 3D scanners were considered in [13, 14]. Despite the high efficiency achieved, the use of these methods is hampered by the need for specialized equipment. Methods that do not require additional devices or human intervention are preferable, as they can be easily integrated into existing facial recognition systems, which are typically equipped only with a camera. Such (non-invasive) methods typically utilize motion, action, and texture characteristics [15].

Motion analysis exploits the fact that the motion of flat 2D objects differs significantly from the motion of a real human face, which is a 3D object. In [16], differences in the properties of optical flow generated by 3D objects and 2D planes were analyzed. In [17], the authors estimated the motion trajectory of individual facial parts using a short image sequence and a simplified optical flow analysis. The basic idea of the method is based on the assumption that the central part of the face moves differently than its peripheral parts (for example, around the ear). It should be noted that a motion-based approach may not provide adequate performance when motion information is insufficient. For example, errors may occur when analyzing noisy and low-resolution images.

Action cues, depending on the method of interaction with the user, can be divided into two types: those requiring feedback (e.g., following the instruction "open your mouth at a certain moment," so to speak, presenting a "motion password") and those not requiring any specific activity from the user, i.e., based on actions independent of external instructions (e.g., spontaneous blinking). For example, a method based on blink analysis was proposed in [18]. In [19], the authors analyzed eye movements and trained a motion model using a support vector machine. In [20], a method was proposed in which facial vitality is determined by lips. Vitality, as applied to biometric systems, means that a "live" person is interacting with the system. Action cues are very difficult to fake using a photograph or a 3D sculpture of a human head. However, this approach may require user interaction and is highly dependent on the quality of facial feature detection. Methods based on motion analysis and open-loop action analysis may also be ineffective when using video recordings of the target.

Texture-based approaches can detect texture features not characteristic of a living person. For example, in [21], texture information was calculated using local binary patterns. Unlike the previous two approaches, texture analysis does not require user interaction and is simple to implement. This approach operates on the assumption that fake facial images are printed on paper, and the printing process and the paper's structure generate texture variations, allowing for the distinction between printed images and real faces [22].

However, these assumptions may not be true in some cases. For example, with the development of ultra-high-definition displays, the likelihood that the color intensity reproduction on the display is close to the actual color values increases. In other words, an attack can be carried out using a photo displayed on a screen rather than on paper, making it difficult to distinguish the textures of a real person's image from a fake one. Therefore, spoofing attacks using facial video played on a high-definition display will become increasingly common. Therefore, this spoofing scheme should be given closer attention when developing and testing methods for facial vitality detection.

Taking into account all of the above, this paper proposes a new method for countering spoofing attacks. The method is based on texture analysis and does not require the analysis of multiple frames. A decision can be made based on a single facial image. The approach is based on training a convolutional neural network to extract facial feature vectors and perform further classification using a support vector machine, which allows for high accuracy in distinguishing between genuine and fake facial images.

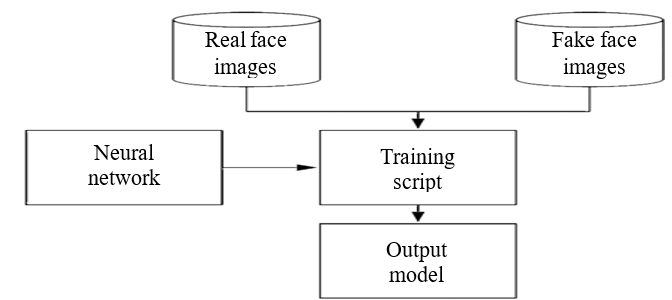
***The proposed method***. The proposed method for countering spoofing attacks determines how similar a received frame (or group of frames) is to a person in front of the camera, rather than a photograph or video. If multiple images are fed to the method as input, the resulting probability value is calculated as the average value across all images. The proposed method is based on the use of a convolutional neural network to generate a feature vector and a support vector machine for classification.

The initial data fed to the network are 300×300 pixel facial images. All facial images undergo a geometric normalization step according to the coordinates of the control points.

A convolutional neural network (AlexNet) is proposed as the neural network architecture for generating the feature vector.

The stochastic gradient descent algorithm [23] was chosen for network training. The network was trained using the NUAA Imposter Database and our own database.

The NUAA Imposter Database contains 5,105 images of real faces and 7,509 fake faces obtained from 15 respondents. Caffe [24], an open-source library developed by the Berkeley Vision and Learning Center team and designed for scientific research in the field of computer vision, is used to train the model. The input layer accepts a color image of 300×300 pixels. To improve the quality of the network, augmentation was used – adding distorted images to the sample in order to increase the diversity of the data. A transformation such as reflection was used. This procedure improves the generalization ability of the convolutional neural network and prevents overfitting [25]. The learning rate is initially set to 10-4 and gradually decreases to 10-5.

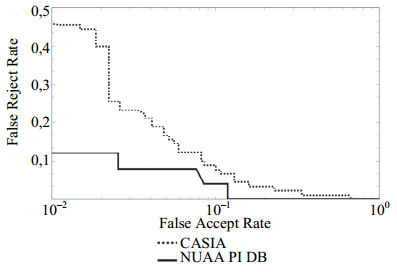


**FIGURE 1.** The structure of the learning neural network with the use of real and fake images

**EXPERIMENTAL RESEARCH**

In this paper, experiments are conducted on two datasets: NUAA PI DB and CASIA. These datasets simulate various types of spoofing attacks.

The CASIA database contains facial images of 50 individuals. For each individual, low-, medium-, and high-quality facial images were obtained. The fake faces simulate three types of attacks: a deformed photo (the attacker presents a printed photograph to the camera, intentionally distorting it in an attempt to simulate facial movement), a partially cropped photo (the eyes are cropped out of the photograph, the attacker hides behind the photograph and simulates blinking), and a video-based attack (an iPad displaying high-quality video is presented to the camera). As a result, for each individual, there are 12 video sequences (3 true faces and 9 fake ones). The total number of sequences in the database is 600, of which 240 are used for training and 360 for testing.



**FIGURE 2.** DET curves for the method of countering spoofing attacks on the NUAA PI DB and CASIA databases

The NUAA PI DB database contains images of 15 individuals and is divided into three sessions based on different lighting conditions. The data volume is unbalanced across sessions, as not all individuals participated in all three sessions. During filming, participants were asked to look frontally at the webcam, maintain a neutral facial expression, and avoid head movement or blinking, thereby achieving a close resemblance to the photo. The camera records a 25-second video at 20 frames per second (fps). Images selected from the original video sequence, rather than the original video sequence, are available for download. Images for the spoofing attack were collected for the same participants. The individual was photographed using a Canon camera so that their face occupied one-third of the image. The photographs were then printed and presented to the camera, moving during frame capture.

Based on the results of the experiments, DET curves (Define Error Tradeoff Curves) were constructed and are shown in Figure 2. Two types of errors were evaluated:

1. False Accept (FA) error of falsely accepting a photograph (video) placed in front of the screen as a live user;

2. False Reject (FR) error of falsely accepting a live user as their image (video).

**RESEARCH RESULTS**

Figures 3-4 show the accuracy and loss when training the model on this dataset.

|  |  |
| --- | --- |
|  |  |
| **FIGURE 3.** Plot of training accuracy and loss on the NUAA PI dataset | **FIGURE 4.** Plot of training accuracy and loss on the CASIA dataset |

The average value of the first and second type errors (Half Total Error Rate, HTER) and accuracy (the proportion of input data classified into the correct class) for the proposed method to counter spoofing attacks, as well as a comparison of the proposed method with existing approaches are presented in Tables 1 and 2.

**TABLE 1.** Comparison of the effectiveness of countermeasures against spoofing attacks based on NUAA PI dataset

|  |  |  |
| --- | --- | --- |
| **Method** | **Performance evaluation** | |
| **HTER, %** | **Accuracy, %** |
| Sparse logistic regression + DoG filter | – | 87,5 [26]; 93 [27] |
| LBP + SVM [28] | 19,03 | – |
| LBP + DoG filter [29] | 11,97 | – |
| LBP + Linear discriminant analysis [30] | 18,32 | – |
| Proposed method | 5,27 | 96,5 |

**TABLE 2.** Comparison of the effectiveness of countermeasures against spoofing attacks based on CASIA dataset

|  |  |  |
| --- | --- | --- |
| **Method** | **performance evaluation** | |
| **HTER, %** | **Accuracy, %** |
| LBP + Linear discriminant analysis [30] | 21,01 | – |
| LBP + SVM [28] | 18,17 | – |
| LBP [30] | 18,21 | – |
| LBP-TOP | 10 | – |
| proposed method | 8,27 | 91,3 |

As a result, it can be concluded that the proposed convolutional network architecture and the selected training data allow for the detection of spoofing attacks based on facial images with high accuracy.

**CONCLUSIONS**

This paper examines the problem of countering spoofing attacks in facial biometric systems. Existing methods for substituting biometric characteristics, as well as methods aimed at countering spoofing in facial biometric systems, are reviewed, and their shortcomings are identified. Based on an analysis of these shortcomings, a method is proposed whose results are comparable to the best results achieved using other known spoofing detection methods.

The proposed solution to countering spoofing attacks consists of sequentially performing two stages of analysis: facial feature extraction and classification to determine the degree to which a presented frame (or group of frames) resembles a real person, rather than a photograph or video.

To address the first subproblem, related to facial feature extraction, the paper proposes the use of a convolutional neural network containing five convolutional and three fully connected layers. The training algorithm is described, along with the data used for training and validation. To solve the second subproblem, which involves deciding whether to assign the obtained feature vector to one of the classes, the paper proposes using a support vector machine with a radial basis function kernel.

The software implementation developed for solving the main research problem allows the proposed solution to be used for both single-frame and group-frame analysis.

Experiments conducted showed that the average Type I and Type II errors on test data do not exceed 9%, and accuracy reaches over 91%.

Further development of the neural network method for countering spoofing attacks is possible through the use of recurrent neural networks for additional analysis of relationships between multiple frames.

**REFERENCES**

1. Matveyev Yu.N. Texnologii biometricheskoy identifikatsii lichnosti po golosu i drugim modalnostyam // Injenerniy jurnal: nauka i innovatsii. 2012. №3. S. 46–61.
2. Physical access control biometrics [Electronic resource]. Access mode: http://www.findbiometrics.com/physicalaccess/.
3. Li S.Z., Jain A.K. Handbook of Face Recognition. London: Springer-Verlag, 2011. 724 p. doi: 10.1007/978-0-85729-932-1
4. Datta A.K., Datta M., Banerjee P.K. Face Detection and Recognition: Theory and Practice. Chapman and Hall/CRC, 2015. 326 p. doi: 10.1201/b19349.
5. Mamatov N. S., Abdukadirov B. A., Samijonov A. N., Samijonov B. N. Method for false attack detection in face identification system // IEEE International Conference on Information Science and Communications Technologies. – 2021. –Pp. 1–4.
6. Mamatov N., Abdukadirov B., Kakharov Sh., Orifjonov B., Abdukadirova G. Peculiarities of face detection and recognition // IEEE International Conference on Information Science and Communications Technologies. – 2021. –Pp. 1-5.
7. Marcel S., Nixon M.S., Li S.Z. Handbook of Biometric AntiSpoofing: Trusted Biometrics under Spoofing Attacks. Springer, 2014. 281 p. doi: 10.1007/978-1-4471-6524-8
8. Galbally J., Marcel S., Fierrez J. Biometric antispoofing methods: a survey in face recognition // IEEE Access. 2014. V.2. P. 1530–1552. doi: 10.1109/ACCESS.2014.2381273
9. Marshalko G.B., Nikiforova L.O. Spufing ataka na biometricheskuyu sistemu identifikatsii, ispolzuyuщuyu algoritm raspoznavaniya Eigenfaces // Problemi informatsionnoy bezopasnosti. Kompyuterniye sistemi. 2018, №3, s. 37-44.
10. Volkova S.S., Matveev Yu.N. Convolutional neural networks for face anti-spoofing. // Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2017, vol. 17, no. 4, pp. 702–710 (in Russian). doi: 10.17586/2226-1494-2017-17-4-702-710
11. Marshalko G.B., Nikiforova L.O. Spufing ataka na biometricheskuyu sistemu identifikatsii, ispolzuyuщuyu algoritm raspoznavaniya Eigenfaces // Problemi informatsionnoy bezopasnosti. Kompyuterniye sistemi. 2018, №3, s.37-44.
12. Sharif M., Bhagavatula S., Bauer L., Reiter M. K., Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition // Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (2016), ACM, pp. 1528 – 1540.
13. Kostilev N.M., Gorevoy A. The liveness detection module based on spectral reflection characteristics of facial skin // Engineering Journal: Science and Innovation. 2013. №9 (21). p. 47–60.
14. Lagorio A., Tisterelli M., Cadoni M. et. al. Liveness detection based on 3D face shape analysis // Proc. International Workshop on Biometrics and Forensics, IWBF. Lisbon, Portugal, 2013. Art. 657310. doi: 10.1109/IWBF.2013.6547310
15. Chakarborty S., Das D. An overview of face liveness detection // International Journal on Information Theory. 2014. V. 3. №2. P. 11–25.
16. Bao, Wei & Li, Hong & Li, Nan & Jiang, Wei. A liveness detection method for face recognition based on optical flow field. Proceedings of 2009 International Conference on Image Analysis and Signal Processing, IASP 2009. 233-236.
17. Kollreider K., Fronthaler M., Bigun J. Evaluating liveness by face images and structure tensor // Proc. 4 th IEEE Workshop on Automatic Identification Advanced Technologies. Washington, USA, 2005. P. 75–80. doi: 10.1109/AUTOID.2005.20
18. Jee H.-K., Jung S.-U., Yoo J.-H. Liveness detection for embedded face recognition system // International Journal of Biomedical Sciences. 2006. V1. №4. P. 235.
19. Deng G., Coo B., Miao J. et al. Liveness check algorithm based on eye movement model using SVM // Journal of Computer Aided Design and Computer Graphics. 2003. V15. №7. P. 853–857.
20. Kollreider K., Fronthaler M., Faraj M.I., Bigun J. Real-time face detection and motion analysis with application in liveness assessment // IEEE Transactions on Information Forensics and Security. 2007. V2. №3. P. 548–558.
21. Kim G., Eum S., Suhr J.K. et al. Face liveness detection based on texture and frequency analyses // Proc. 5th IAPR Int. Conf. on Biometrics (ICB). New Deli, India, 2012. P. 67–72.
22. Yang J., Lei Z., Liao S., Li S.Z. Face liveness detection with component dependent descriptor // Proc. Int. Conf. on Biometrics (ICB). Madrid, Spain, 2013.
23. LeCun Y., Bottou L., Orr G.B., Muller K.R. Efficient Backprop // Neural Networks Tricks of the Trade. 1998. V. 1524. P. 9–50. doi: 10.1007/3-540-49430-8\_2
24. Jia, Y. Caﬀe: Convolutional architecture for fast feature embedding [Electronic resource]. — 2014. — URL: https://arxiv.org/abs/1408.5093.
25. Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classiﬁcation with deep convolutional neural networks // Advances in Neural Information Processiny system. –2012. – P. 1097–1105.
26. X. Tan, Y. Li, Liu J., L. Jiang. Face liveness detection from a single image with sparse low rank bilinear diskriminative model // Computer vision ECCV2010. — 2010.
27. Peixoto B., Michelassi C., Rocha A. Face liveness detection under bad illumination conditions // In.: IEEE 18th International Conference of Image Processing (ICIP), 11-14 September. -2011.
28. Maatta, J. Hadid A., Pietik M. Face spooﬁng detection from single images using micro-texture analysis // IEEE, Washington DC. — 2011. — P. 1–7.
29. Koze, N., Dugelay J.L. Classiﬁcation of captured and recaptured images to detect photograph spooﬁng // In.: APR International Conference of Informatics, Electronics and Vision, IEEE, Dhaka. - 2012. - P. 1027–1032.
30. Chingovska, I., Anjos A., Marcel S. On the eﬀectiveness of local binary patterns in face anti-spooﬁng// In processing of International Conference on Biometrics Internet Group (BIOSIG), IEEE, Darmstadt, 6-7 September 2012