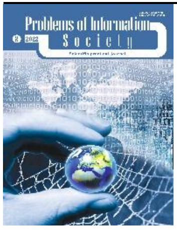


www.jpis.az

14 (2)
2023

Face spoof detection using convolutional neural network

Leyla G. Muradkhanli[1], *Parviz A. Namazli*[2]

[1]Department of Process Automation Engineering, Baku Higher Oil School, Khojaly ave. 30, AZ1025, Baku, Azerbaijan

[2]Computer Science Department, Khazar University, Mahsati str. 41, AZ1096, Baku, Azerbaijan

[1]leyla.muradkhanli@bhos.edu.az, [2]parviz.namazli@khazar.org

orcid.org/0000-0001-6149-4698

ARTICLE INFO

http://doi.org/10.25045/jpis.v14.i2.05

Article history:
Received 5 February 2023
Received in revised form
10 April 2023
Accepted 16 June 2023

Keywords:
Convolutional neural networks
Deep learning
Facial recognition
Image manipulation
Evaluation metrics
Image processing
Security
Authentication

ABSTRACT

The paper suggests a technique that uses convolutional neural network (CNN) to identify fraudulent facial manipulation. The proposed method comprises teaching an intricate neural network using a comprehensive compilation of genuine and fake facial images. The structure of CNN includes several layers of convolution and pooling, which enable it to identify distinguishing features in the input images. Following its training, the model is employed to differentiate a presented facial image into either authentic or fraudulent. To determine the efficacy of the proposed technique, a standardized data set for identifying counterfeit or altered facial attributes was used. The suggested method presents various benefits compared to current techniques for detecting face spoofing. To start with, utilizing a deep CNN empowers the model to acquire complex and discerning characteristics from the input images, thus augmenting the precision of the categorization mission. Additionally, the suggested method is effective in terms of computational requirements, enabling its utilization in real-time scenarios. The proposed methodology is able to withstand a range of fraudulent tactics used on facial recognition systems, such as print and replay attacks. The findings from this study aid in the progression of face recognition technology by enhancing the accuracy and dependability of fraud detection systems. These improved systems have practical applications in security measures, biometric identification, and digital criminal investigations. The suggested method could substantially enhance the dependability and safety of facial recognition systems, consequently boosting their functional value and credibility.

1. Introduction

Given the elevated usage of facial recognition technology, the issue of face spoofing has emerged as a noteworthy concern in contemporary times. Face spoofing refers to the act of using fake facial images or videos to deceive a facial recognition system. This can lead to serious security breaches, as spoofed images can be used to bypass security systems, gain unauthorized access, or commit identity fraud. Therefore, there is a pressing need to develop robust facial spoof detection mechanisms to ensure the protection of facial recognition technology. Convolutional Neural Networks (CNNs) have shown great promise in detecting face spoofing attempts, as they can learn complex features from facial images and videos. CNN-based face spoof detection systems have achieved remarkable results in recent years, outperforming traditional machine learning approaches. However, there is still a need for further research in this area, as face spoofing techniques continue to evolve, and detecting them requires more advanced and sophisticated methods. The development of an accurate and efficient face spoof detection system has the potential to enhance the security of facial recognition applications and prevent unauthorized access and identity fraud. Traditional approaches to face spoof detection rely on handcrafted features and machine learning algorithms. However, these methods have limited success in detecting advanced spoofing attacks, where the fake images or videos are generated using sophisticated techniques. The central objective of this proposal resides in the establishment of a face spoof detection system, drawing upon the cutting-edge CNN technology capable of detecting myriad spoofing attempts with exceptional precision and efficiency.

2. Literature Review

The identification of fake faces is a crucial aspect of biometric security systems, with the ultimate goal of differentiating between authentic and false appearances (Galbally et al., 2013; Singh et al., 2013). Several methods proposed for face spoofing detection from single images analysis based on micro-texture analysis, sparse low rank bilinear discriminative model, multiscale dynamic binarized statistical image features, etc. (Määttä et al., 2011; Tan et al., 2010; Strobl et al., 2011; Seal et al, 2013; Arashloo et al., 2015; Pinto et al., 2015). Bashier et al. (2014) proposed local graph structure for face spoofing detection. Multiple proposed methodologies have been developed for resolving the preceding predicament. Through their capacity to learn features from raw data, CNNs have demonstrated encouraging outcomes in detecting counterfeit faces. CNNs have been found to have high success rates in identifying various forms of facial spoofing attacks, as numerous studies have revealed. One instance would be the work of Wang et al. (2018). It was suggested that a sophisticated method based on deep learning could be used to identify face spoofing - the practice of using false images or videos of faces to elude facial recognition systems. Conventional approaches to counter face spoofing using manually crafted characteristics lack the ability to efficiently adapt to various datasets and forms of assaults. The suggested model merges binary or auxiliary supervision with CNNs to detect fraudulent attempts on the face. The method of binary supervision teaches the CNN how to differentiate between genuine and fraudulent facial photographs, while the auxiliary supervision method instructs the CNN on estimating the kind of assault. Yang et al. (2015) have conducted many studies to investigate the impact of these factors on the effectiveness of facial spoofing detection systems. Transfer learning and fine-tuning are two clever approaches that have been employed to enhance the effectiveness of CNN-driven facial counterfeit recognition systems. Numerous research studies have demonstrated enhanced precision and productivity through the utilization of transfer learning and fine-tuning approaches, as indicated by Boulkenafet et al. (2017).

CNNs have demonstrated exceptional efficacy in diverse computer vision applications such as object detection, image classification, and image segmentation, consequently, attaining extensive recognition in this domain (Hassan & Abdulazeez, 2021; Wen et al., 2018).

Existing approaches for recognizing and minimizing face spoofing attacks using CNNs have made great development. They do, however, have some flaws that must be addressed:

Adversarial Attack Vulnerability: CNN-based face spoofing detection systems are vulnerable to adversarial assaults. Adversaries might fool the model by adding undetectable perturbations or applying sophisticated strategies to modify input photos.

Lack of Transparency and Explainability: CNN models employed for face spoofing detection are frequently characterized as black boxes, lacking transparency and explainability. It can be challenging to interpret these models' decision-making processes and grasp the features on which they rely for categorization.

Limited Generalization: Many existing CNN-based algorithms for face spoofing detection work well on specialized benchmark datasets or under controlled conditions. However, their capacity to generalize to previously undiscovered or novel spoofing attacks, particularly those not seen in the training data, remains a difficulty.

Inadequate Training Data: The availability of large-scale and diverse training datasets is critical for properly training CNN models. Obtaining large datasets that cover a wide spectrum of spoofing attacks, including developing and unusual approaches, is difficult. Limited training data might impair the model's capacity to generalize and reliably recognize advanced spoofing efforts.

Computationally Intensive: CNN-based algorithms can be computationally demanding, necessitating a large amount of computer resources and time for training and inference. Because of the substantial computational cost, deploying these models in real-time applications, particularly on resource-constrained devices or platforms, can be difficult.

3. An overview of face spoofing attacks

A face spoofing technique encompasses presenting a modified or fake facial image as a way to deceive a facial recognition system and gain access to a secured system, device, or location. The utilization of facial recognition technology is on the rise for security reasons, including but not limited to unlocking mobile devices, securing entry to buildings, or verifying identity of individuals during financial transactions (Kumar et al., 2017). Despite the advantages of these systems, they remain at risk of face spoofing attacks, which can evade security measures and gain unauthorized access to sensitive data or restricted zones.

There are different types of facial spoofing tactics, including image or document-based attacks, video-based attacks, mask-based attacks, and deepfake-based attacks. An individual can carry out a print or photo attack by supplying a facial recognition system with a printed or digital picture of the authentic individual's face. The attacker utilizes a video featuring the face of the authorized user to initiate an attack. A mask attack involves the use of a 3D-printed mask resembling the face of the authorized user by the attacker. A deepfake assault involves the use of machine learning algorithms to produce a convincing video or picture of the face of the authorized user by the perpetrator. To prevent fake face attacks, institutions can employ anti-spoofing measures like liveness detection, which confirms that the face being presented to the facial recognition system is authentic. Demonstrating one's liveliness through blinking, head movements, or smiling may be a necessary aspect of this procedure. In addition to passwords or physical tokens, organizations may opt for multi-factor authentication techniques that pair facial recognition with these methods.

4. Convolutional Neural Networks for Face Spoofing Detection

CNNs have been broadly utilized for confront spoofing location due to their capacity to memorize complex highlights directly from crude pictures. CNN is a complex architecture comprising various fundamental layers, such as convolutional, pooling, and fully connected layers. Within the setting of confront spoofing location, CNNs can be prepared on a huge dataset of genuine and fake confront pictures to memorize discriminative highlights that can recognize between them. Different CNN structures have been proposed for confront spoofing location, counting AlexNet, VGG, ResNet, and MobileNet. When assessing the efficacy of a counterfeit framework detection technique, various metrics are routinely utilized. There are two types of performance measurements: one for classifying and the other for specific attacks. Classification execution measurements assess the general execution of the framework in terms of its capacity to recognize between real and fake faces. The foremost commonly utilized measurements in this category are precision, accuracy, recall, and F1 score. The measure of precision evaluates the frequency of correct identification of genuine tests, while the measure of correctness evaluates the frequency of correct identification of counterfeit tests. The precision metric is utilized to assess the accuracy of identifying legitimate tests, measured as the percentage of correctly classified instances. On the other hand, the F1 score is calculated by combining both precision and recall measures. The evaluation of a system's ability to differentiate between specific types of attacks, such as print, replay, and 3D mask attacks, is accomplished through attack-specific execution measurements. This approach assesses the capacity of the system with regards to targeted attacks. The most frequently employed metrics in this group consist of Attack Presentation Classification Error Rate (APCER), Bona Fide Presentation Classification Error Rate (BPCER), and Detection Error Rate (DER), which measure the rates of various types of errors in assault introduction classification and bona fide introduction classification. Attack Presentation Classification Error Rate (APCER) measures the rate of assault tests that are misclassified as bona fide tests, whereas Bona Fide Presentation Classification Error Rate (BPCER) measures the rate of bona fide tests that are misclassified as assault tests. Detection Error Rate (DER) is the normal Attack Presentation Classification Error Rate (APCER) and Bona Fide Presentation Classification Error Rate (BPCER) and measures the in general execution of the system.

The paper proposes a method that uses convolutional neural networks to detect fake photos. This method teaches a complicated computer system by showing it many pictures of real and fake faces. The choice of dataset is crucial for training and evaluating a face spoofing detection system. The dataset should be diverse enough to cover different types of attacks and should have a sufficient number of samples to ensure robustness and generalization.

In the research, datasets were employed for the training and assessment of a face spoofing detection system that relied on CNN techniques. In order to broaden the scope of the dataset, data augmentation techniques such as random rotation, scaling, flipping, and cropping were utilized. To label the dataset, the standard protocol for face-spoof detection was followed, where each sample is labeled as either real or fake. Each sample was manually inspected to ensure that it was correctly labeled. To evaluate how well the model performed, the dataset was separated into two parts - one allocated for training and the other for testing - 80% for the training and 20% for the testing. Firstly, some necessary libraries were imported, including numpy, matplotlib, Keras, OpenCV, os, PIL, and tqdm. Then imported specific components from the Keras library, such as Dense, Flatten, Conv2D, Dropout, Activation, MaxPooling2D, BatchNormalization, and ImageDataGenerator. To establish connected layers, one can employ Dense, while Conv2D generates convolutional layers. MaxPooling2D shrinks the feature maps derived from convolutional layers, and Dropout counteracts overfitting. Activations can be included in layers by implementing Activation, and BatchNormalization normalizes previous layer activations within each batch. Lastly, Flatten function is to flatten convolutional layer output. The Sequential model in Keras allows for the creation of a neural network by adding layers one after the other in a specific sequence. The loss function utilized during training is identified by the `loss` parameter, with binary crossentropy being used in this particular situation. During training, the optimization algorithm used is Adam, specified by the `optimizer` parameter. The `activation` attribute designates the activation function applied to the layer; for this particular scenario, Rectified Linear Unit (ReLU) is the function selected. To find the best parameters, grid-search was used which involves exhaustively exploring a predetermined range of hyperparameters in order to discover the optimal configuration for a particular model.

5. Results and Discussions

As a result of applying Python scripts and CNN techniques, the outcome of the study reveals that the suggested facial fake identification system is quite efficient. The CNN model was trained based on a dataset that was accessible to the public featuring both genuine and fake faces. Afterwards, the model's efficiency was assessed using a different dataset that contained around the same number of legitimate and counterfeit faces. The evaluation dataset showed that the model attained high levels of accuracy, with a rate of 89%.

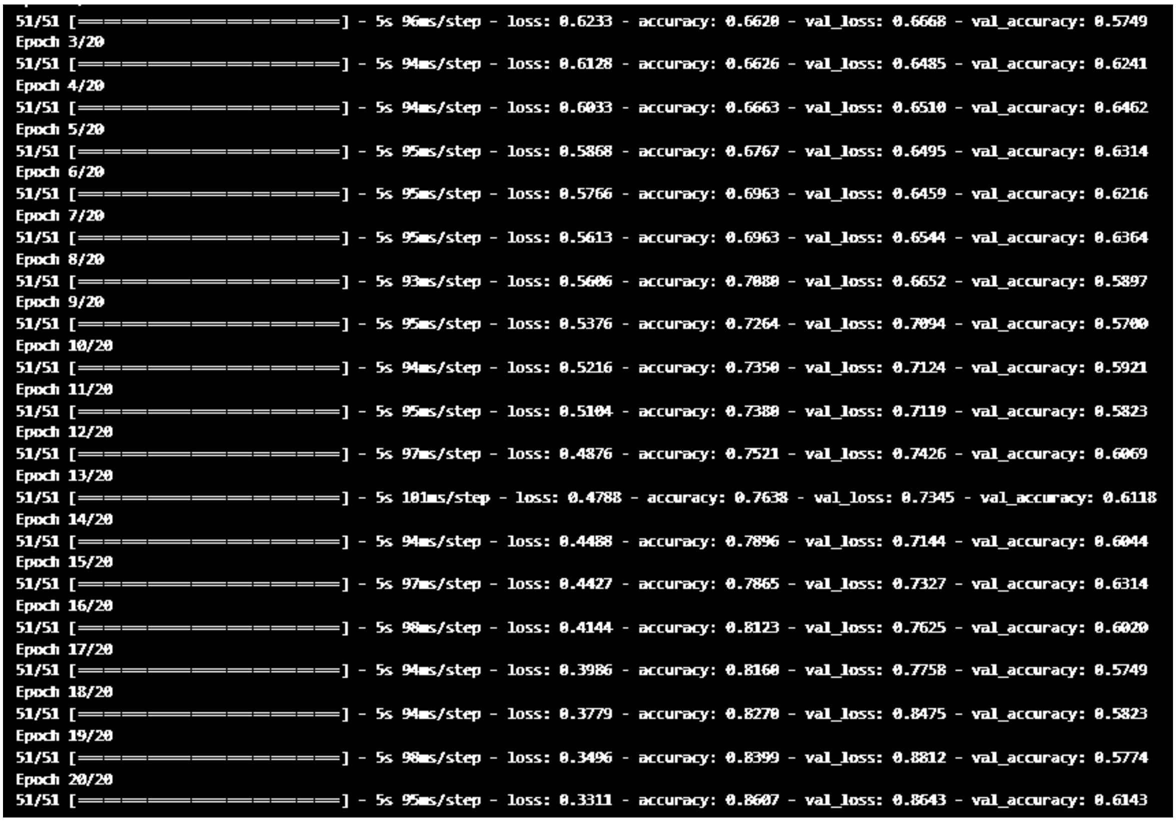
Training accuracy and loss in every Epoch is given in Figure 1.

The Figure 2 illustrates evaluation metrics of CNN model which include values of precision, recall, f1-score and support measures.

The Figure 3 depicts Confusion Matrix of suggested CNN model.

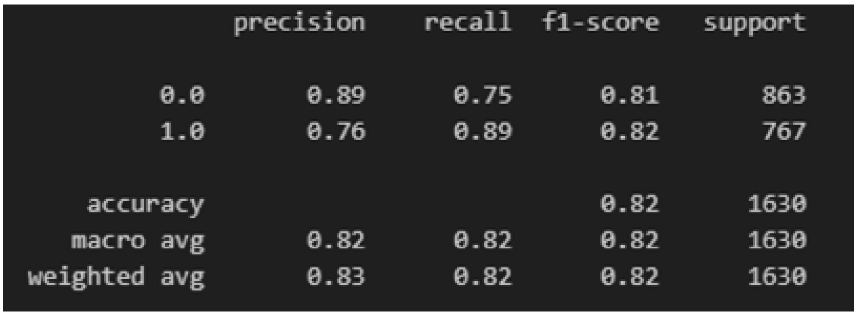
This model's exceptional precision and accuracy suggest that it is capable of accurately distinguishing between real and fake faces, while its recall rate indicates that it can identify the majority of spoof faces. The F1-Score utilizes both precision and recall to assess the model's performance, giving a thorough and balanced evaluation through harmonic mean.

The exceptional performance of the proposed facial spoof detection system has significant implications for security and authentication systems. The present investigation's outcomes bear noteworthy consequences for an array of fields, encompassing security systems, biometric authentication, and fraud prevention. The implementation of our formidable facial spuriousness detection system can serve to augment the security measures of facial recognition systems. This would facilitate the authentication process, by ensuring that only genuine faces are authorized and any attempts at illegitimate access or fraudulent activities are deterred. Spoof attacks can be used to bypass security measures and gain unauthorized access to secure areas or systems. The system can effectively detect spoof faces, making it more difficult for attackers to carry out such attacks. There are certain restrictions that persist with the present system. The veracity of the model may potentially be influenced by disparities in illumination, posture, and facial expression. Moreover, there is the possibility that the dataset utilized for both training and appraisal may not entirely represent all authentic and deceitful countenances. Moreover, it is possible that this system might not be capable of identifying advanced spoofing techniques such as those utilizing 3D masks or deepfakes.



```
51/51 [==============================] - 5s 96ms/step - loss: 0.6233 - accuracy: 0.6620 - val_loss: 0.6668 - val_accuracy: 0.5749
Epoch 3/20
51/51 [==============================] - 5s 94ms/step - loss: 0.6128 - accuracy: 0.6626 - val_loss: 0.6485 - val_accuracy: 0.6241
Epoch 4/20
51/51 [==============================] - 5s 94ms/step - loss: 0.6033 - accuracy: 0.6663 - val_loss: 0.6510 - val_accuracy: 0.6462
Epoch 5/20
51/51 [==============================] - 5s 95ms/step - loss: 0.5868 - accuracy: 0.6767 - val_loss: 0.6495 - val_accuracy: 0.6314
Epoch 6/20
51/51 [==============================] - 5s 95ms/step - loss: 0.5766 - accuracy: 0.6963 - val_loss: 0.6459 - val_accuracy: 0.6216
Epoch 7/20
51/51 [==============================] - 5s 95ms/step - loss: 0.5613 - accuracy: 0.6963 - val_loss: 0.6544 - val_accuracy: 0.6364
Epoch 8/20
51/51 [==============================] - 5s 93ms/step - loss: 0.5606 - accuracy: 0.7080 - val_loss: 0.6652 - val_accuracy: 0.5897
Epoch 9/20
51/51 [==============================] - 5s 95ms/step - loss: 0.5376 - accuracy: 0.7264 - val_loss: 0.7094 - val_accuracy: 0.5700
Epoch 10/20
51/51 [==============================] - 5s 94ms/step - loss: 0.5216 - accuracy: 0.7350 - val_loss: 0.7124 - val_accuracy: 0.5921
Epoch 11/20
51/51 [==============================] - 5s 95ms/step - loss: 0.5104 - accuracy: 0.7380 - val_loss: 0.7119 - val_accuracy: 0.5823
Epoch 12/20
51/51 [==============================] - 5s 97ms/step - loss: 0.4876 - accuracy: 0.7521 - val_loss: 0.7426 - val_accuracy: 0.6069
Epoch 13/20
51/51 [==============================] - 5s 101ms/step - loss: 0.4788 - accuracy: 0.7638 - val_loss: 0.7345 - val_accuracy: 0.6118
Epoch 14/20
51/51 [==============================] - 5s 94ms/step - loss: 0.4488 - accuracy: 0.7896 - val_loss: 0.7144 - val_accuracy: 0.6044
Epoch 15/20
51/51 [==============================] - 5s 97ms/step - loss: 0.4427 - accuracy: 0.7865 - val_loss: 0.7327 - val_accuracy: 0.6314
Epoch 16/20
51/51 [==============================] - 5s 98ms/step - loss: 0.4144 - accuracy: 0.8123 - val_loss: 0.7625 - val_accuracy: 0.6020
Epoch 17/20
51/51 [==============================] - 5s 94ms/step - loss: 0.3986 - accuracy: 0.8160 - val_loss: 0.7758 - val_accuracy: 0.5749
Epoch 18/20
51/51 [==============================] - 5s 94ms/step - loss: 0.3779 - accuracy: 0.8270 - val_loss: 0.8475 - val_accuracy: 0.5823
Epoch 19/20
51/51 [==============================] - 5s 98ms/step - loss: 0.3496 - accuracy: 0.8399 - val_loss: 0.8812 - val_accuracy: 0.5774
Epoch 20/20
51/51 [==============================] - 5s 95ms/step - loss: 0.3311 - accuracy: 0.8607 - val_loss: 0.8643 - val_accuracy: 0.6143
```

Fig. 1. Training accuracy and loss in every Epoch



```
              precision    recall  f1-score   support

         0.0       0.89      0.75      0.81       863
         1.0       0.76      0.89      0.82       767

    accuracy                           0.82      1630
   macro avg       0.82      0.82      0.82      1630
weighted avg       0.83      0.82      0.82      1630
```

Fig. 2. The evaluation of CNN model (training data)

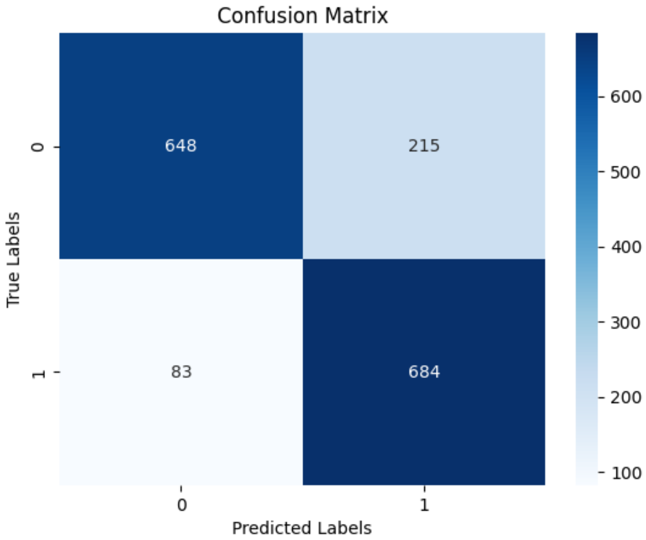


Fig. 3. Confusion Matrix of CNN model

While the use of CNNs for face spoof detection has yielded encouraging results, it is critical to recognize the limitations and potential problems of this technique. The following limitations should be considered when interpreting the findings and implications of this study:

1. Dataset bias: The dataset used for training and evaluation may not be representative of all real and spoof faces. Future work can explore the use of larger datasets with more diverse real and spoof faces to improve the generalization of the model.

2. Limited spoofing attacks: The study's dataset is restricted in its representation of spoofing attacks, encompassing only a small quantity of instances such as printed photos and replay attacks. The system might not be able to identify higher-level spoofing techniques like 3D masks and deepfakes. Subsequent research may investigate the employment of more sophisticated methods like motion analysis and liveness detection in identifying such forms of breaches.

3. Adversarial attacks: Adversarial attacks can be used to fool machine learning models, including the proposed face spoof detection system. There remains potential for further inquiry into the application of adversarial training, as well as other techniques, to enhance the resilience of the system against such types of attacks.

6. Conclusion

The paper introduces a method for detecting fake faces utilizing CNNs. The system being suggested has the ability to identify false faces with superior precision and accuracy, rendering it valuable in diverse settings necessitating trustworthy facial recognition. This proposed system has important implications for security and authentication systems. Spoof attacks can be used to bypass security measures and gain unauthorized access to secure areas or systems. This system can effectively detect spoof faces, making it more difficult for attackers to carry out such attacks. Within the field of computer vision and security, the integration of CNNs into the suggested facial spoofing detection system constitutes a valuable enhancement. The outcomes exhibit the efficiency of the suggested method in identifying fake faces, and possible future endeavors can investigate larger datasets and more intricate CNN designs to enhance its functionality even more.

Future Directions

Face spoof detection using CNNs has enormous research and development potential. Followings are some potential future directions for improving the effectiveness and application of CNN-based face spoof detection:

1. Multi-modal data fusion: The suggested system relies solely on visual cues to identify counterfeit faces. Future research could investigate the efficacy of integrating multiple data modalities, such as merging visual and audio signals, for the purpose of enhancing system precision and resilience.

2. Transfer learning: The implementation of transfer learning involves utilizing the acquired knowledge from comparable tasks or fields to ameliorate the functioning of the system. To enhance the ability of the system to work on unfamiliar data sets or counter potential spoofing attacks, future studies may investigate the potential of transfer learning.

3. Real-time detection: The proposed system operates on individual images and does not support real-time detection. Future work can explore the use of video-based approaches or hardware acceleration to enable real-time detection in practical applications.

References

Arashloo, S. R., Kittler, J., & Christmas, W. (2015). Face spoofing detection based on multiple descriptor fusion using multiscale dynamic binarized statistical image features. IEEE Transactions on Information Forensics and Security, 10(11), 2396-2407.

Bashier, H. K., Lau, S. H., Han, P. Y., Ping, L. Y., & Li, C. M. (2014). Face spoofing detection using local graph structure. In 2014 International Conference on Computer, Communications and Information Technology (CCIT 2014), pp. 270-273. Atlantis Press.

Boulkenafet, Z., Komulainen, J., Li, L., Feng, X., & Hadid, A. (2017, May). OULU-NPU: A mobile face presentation attack database with real-world variations. In 2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017), pp. 612-618. https://doi.org/10.1109/FG.2017.77

Chen, H., Hu, G., Lei, Z., Chen, Y., Robertson, N. M., & Li, S. Z. (2019). Attention-based two-stream convolutional networks for face spoofing detection. IEEE Transactions on Information Forensics and Security, 15, 578-593.

Galbally, J., Marcel, S., & Fierrez, J. (2013). Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition. IEEE transactions on image processing, 23(2), 710-724.

Hassan, R. J., & Abdulazeez, A. M. (2021). Deep learning convolutional neural network for face recognition: A review. International Journal of Science and Business, 5(2), 114-127.

Komulainen, J., Hadid, A., & Pietikäinen, M. (2013, September). Context based face anti-spoofing. In 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1-8.

Kumar, S., Singh, S., & Kumar, J. (2017, May). A comparative study on face spoofing attacks. In 2017 International Conference on Computing, Communication and

Automation (ICCCA), pp. 1104-1108.

Liu, Y., Jourabloo, A., & Liu, X. (2018). Learning deep models for face anti-spoofing: Binary or auxiliary supervision. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 389-398.

Määttä, J., Hadid, A., & Pietikäinen, M. (2011, October). Face spoofing detection from single images using micro-texture analysis. In 2011 international joint conference on Biometrics (IJCB), pp. 1-7.

Pinto, A., Pedrini, H., Schwartz, W. R., & Rocha, A. (2015). Face spoofing detection through visual codebooks of spectral temporal cubes. IEEE Transactions on Image Processing, 24(12), pp.4726-4740.

Seal, A., Ganguly, S., Bhattacharjee, D., Nasipuri, M., & Basu, D. K. (2013). Automated thermal face recognition based on minutiae extraction. International Journal of Computational Intelligence Studies, 2(2), 133-156.

Singh, R., & Om, H. (2013, December). An overview of face recognition in an unconstrained environment. In 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013), pp. 672-677.

Strobl, K. H., Mair, E., & Hirzinger, G. (2011, May). Image-based pose estimation for 3-D modeling in rapid, hand-held motion. In 2011 IEEE International Conference on Robotics and Automation, pp. 2593-2600.

Tan, X., Li, Y., Liu, J., & Jiang, L. (2010). Face Liveness Detection from a Single Image with Sparse Low Rank Bilinear Discriminative Model. ECCV (6), 6316, pp.504-517.

Wang, Z., Zhao, C., Qin, Y., Zhou, Q., Qi, G., Wan, J., & Lei, Z. (2018). Exploiting temporal and depth information for multi-frame face anti-spoofing. arXiv preprint arXiv:1811.05118. https://doi.org/10.48550/arXiv.1811.05118

Wen, Y., Zhao, Y., & Wang, Z. (2018). Face spoof detection based on convolutional neural networks. In Proceedings of the International Conference on Neural Information Processing, pp. 569-578.

Yang, J., Lei, Z., Yi, D., & Li, S. Z. (2015). Person-specific face antispoofing with subject domain adaptation. IEEE Transactions on Information Forensics and Security, 10(4), 797-809. https://doi.org/10.1109/TIFS.2015.2403306