iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XII **Month of publication:** December 2024

DOI: https://doi.org/10.22214/ijraset.2024.65732

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XII Dec 2024- Available at www.ijraset.com

Attention-Oriented Two-Stream Convolutional Networks for the Identification of Face Spoofing

Dugyala Shivani[1], Karnala Vyshnavi[2], Kotapati Govind Sai[3], Pallam Venkatapathi[4]
ECE Department, CMR Institute of Technology, Medchal, Hyderabad, Telangana, India

Abstract: ***Face recognition has been extensively researched and has seen tremendous success in a variety of applications over the past few decades because the human face retains the most information for identifying individuals. Modern face recognition systems are still vulnerable to face spoofing attacks, such as the face video replay attack. Despite the fact that numerous efficient antispoofing techniques have been put forth, we discover that illuminations impair the effectiveness of many of the current techniques. It encourages us to create illumination-invariant anti-spoofing techniques. In this work, we propose a two stream convolutional neural network (TSCNN) that operates on two complementary spaces: multi-scale retinex (MSR) space (illumination invariant space) and RGB space (original imaging space). In particular, MSR is invariant to illumination but contains less detailed facial information than RGB space, which contains detailed facial textures but is sensitive to illumination. Furthermore, the high-frequency information that is discriminative for face spoofing detection can be efficiently captured by MSR images. To learn the discriminative features for anti-spoofing, the TSCNN is fed images from two spaces. We suggest an attention-based fusion technique that can successfully capture the complementarity of two features in order to smoothly fuse the features from two sources (RGB and MSR). We test the suggested framework on a number of databases, including OULU, CASIA-FASD, and REPLAY-ATTACK, and we obtain very competitive results. We perform cross-database experiments to further confirm the suggested strategies' capacity for generalization, and the outcomes demonstrate the high efficacy of our approach.***

Keywords: Face spoofing, convolutional neural networks, attention-based fusion, softmax

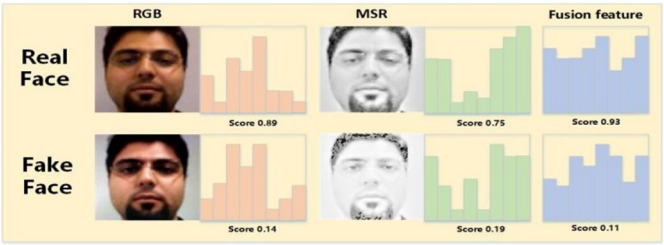
I. INTRODUCTION

Because face recognition technology is more convenient and user-friendly than more conventional techniques like passwords or PINs, it is frequently used for user authentication. However, security concerns have been raised by the growing use of face recognition systems, particularly in light of face spoofing attacks like photo and video replay attacks. These attacks can get around conventional recognition systems by using recaptured images or videos that try to look like a real person. These types of attacks frequently result in a reduction in image quality, which makes the spoof images display traits like moire effects, image banding, or loss of high-frequency information. Face recognition systems need to be resilient to these difficulties in order to avoid spoofing, even though these image degradation factors can be helpful in differentiating between real and fake faces.

Using a deep learning technique known as the Two-Stream Convolutional Neural Network (TSCNN), this paper suggests a novel approach to face spoofing detection in order to address this problem. The multi-scale retinex (MSR) space and the conventional RGB space are the two complementary image spaces that this network operates with. Despite having rich and detailed facial information, the RGB space is extremely sensitive to illumination, which can lead to issues in different lighting scenarios. However, because the MSR space is illumination-invariant, it can withstand changes in illumination while still capturing discriminative features, such as high-frequency data. Because MSR images can capture important high-frequency details, they are useful for differentiating spoof images from real ones, even though they lack some of the finer details found in RGB images. Both RGB and MSR images are used by the TSCNN by feeding them into different network branches.

Every branch uses its own image space to learn discriminative features for anti-spoofing. The authors present an attention-based fusion technique to integrate the complementary information from both image spaces. This technique maximizes the network's capacity to discriminate between real and fake faces by dynamically weighting and adapting the features from both streams. By allowing the model to concentrate on the most pertinent features from both RGB and MSR images, the attention mechanism improves the model's performance. The network is better able to generalize under various lighting conditions thanks to this adaptive fusion, which increases its efficacy in spotting spoofing attacks in practical situations. A number of popular face spoofing databases, such as CASIA-FASD, REPLAY-ATTACK, and OULU, are used to assess the suggested framework.

In both intra-database and cross-database evaluations, the results show that the TSCNN with the attention-based fusion method outperforms other existing methods, achieving highly competitive performance. These results imply that the suggested approach is reliable and has good generalization capabilities across various datasets, including those taken in various lighting scenarios.



II. LITERATUREREVIEW

1. **Wolpert (1992)**: Introduced stacked generalization, an ensemble learning technique that combines predictions from multiple models to improve accuracy. This method has become fundamental in machine learning for tasks like classification and regression.
2. **Itti, Koch, Niebur (1998)**: Proposed a saliency-based visual attention model, using bottom-up processes driven by visual features to simulate human attention. This model has influenced fields like computer vision and robotics.
3. **Gupta et al. (2018)**: Developed an attention model for recognizing emotions in group settings. The model focuses on individual contributions within a group to improve accuracy in social emotion recognition applications.
4. **Nogueira, Lotufo, Machado (2016)**: Presented a CNN-based approach for detecting fake fingerprints, enhancing biometric security by distinguishing live from spoofed fingerprints with high accuracy.
5. **He et al. (2016)**: Introduced Residual Networks (ResNets), addressing vanishing gradient issues in deep networks. ResNets have significantly improved image recognition tasks and become a foundational architecture in deep learning.

L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 11, pp. 1254–1259, 1998.

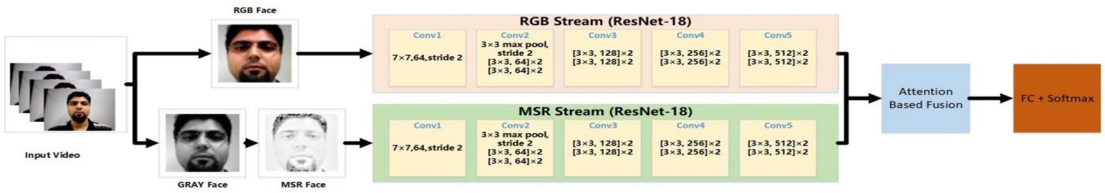
In the 1998 paper "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis," L. Itti, C. Koch, and E. Niebur proposed a computational model that simulates human visual attention mechanisms. The model is based on saliency maps, which highlight the most prominent features of a scene to guide attention toward areas that are most likely to be relevant for processing. By using bottom-up processes driven by low-level visual features like color, intensity, and orientation, the model mimics the way humans selectively focus on important elements of an image.

III. METHODOLOGY

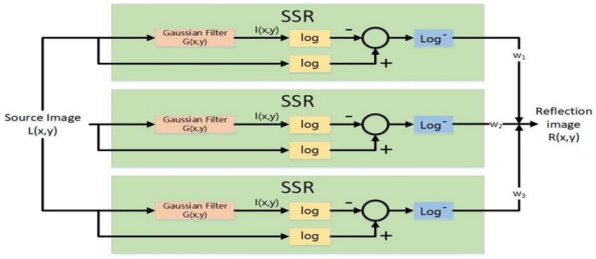
Spoofing detection is actually a binary (real vs. fake face) classification problem. In deep learning era, a natural solution of this task is to feed the input RGB images to a carefully designed CNN with classification loss (softmax and cross entropy loss) for end-to-end training. This CNN-based frame- work has been widely investigated by [25], [26], [47]–[50]. Despite the strong nonlinear feature learning capacity of deep learning, the performance of anti-spoofing degrades when the input images are captured by different devices, under dif- ferent lighting, etc. In this work, we aim to train a CNN which generalizes better to various environments, mainly various lightings.

The RGB images are sensitive to illumination variations yet cover very detailed facial texture information. Motivated by extensive research of (single-scale and multi-scale) Retinex image, we find the Retinex (we use Multi-Scale Retinex - MSR in this work) image is invariant to illumination yet loses minor facial texture. Thus, in this work, we propose a two-stream CNN (TSCNN) which trains two separate CNNs accepting RGB images and MSR images as input respectively. To effectively fuse RGB feature and MSR feature, we propose an attention based fusion method.

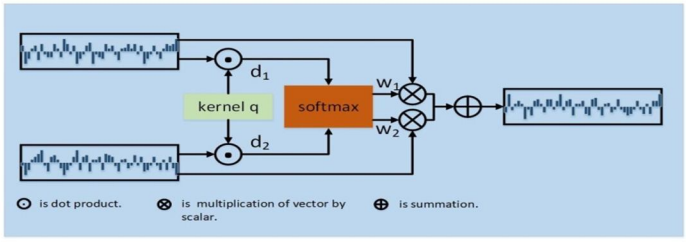
In this section, firstly, we introduce the theory of the Retinex to explain the reason why MSR image is discriminative for anti-spoofing. After that, the complementarity of the RGB and MSR features is analyzed and the proposed TSCNN is detailed. Last, we introduce our attention-based feature fusion method. Retinex (MSR in our work) is used for face spoofing detection with two reasons.(1) The MSR can separate illumination and reflectance. In this work, we use there flectance images (MSR image) to train a CNN for illumination- invariant face spoofing detection. (2) Since the fake face image is regarded as there captured image in many cases, which may lose some high frequency informationcompared to genuine ones. Thus, high frequency in formation can work as a discriminative clue for anti-spoofing. MSR algorithm can be viewed as an optimized high pass filter to capture the high frequency information for spoofing detection.



(A) Attention Based Two-stream Architecture for Face Spoofing Detection



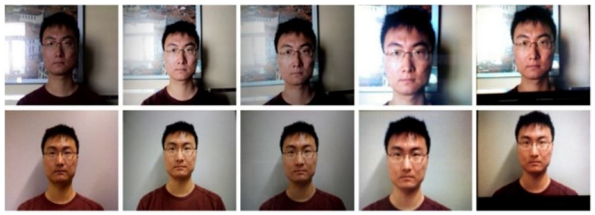
(B) MSR Algorithm



(C) Attention Based Fusion

The CASIA Face Anti-Spoofing Database is divided in to the training set cons is ted of 20 subjects and the test set containing 30 individuals (see, Fig.3). The fake faces weremade by capturing the genuine faces. Three different cameras are used in this database to collect the videos with various imaging qualities :low, normal, and high. In addition, the individuals were asked to blink and not to keep still in the videos to collect abundant frames for detection. Three types of face attacks were designed as follows: 1) Warped Photo Attack: Ahigh resolution (19201080) image, which is recorded by a SonyNEX-5camera,was used to print a photo.

1) *REPLAY-ATTACK Data base:* The REPLAY-ATTACK Data base consists of video recordings of real accesses and attack attempts to 50clients (see,Fig.4). The reare1200 videos taken by the web cam on a Mac Book with the resolution 320240 under two illumination conditions: 1) controlled condition with a uniform back ground and light supplied by a fluorescent lamp, 2) adverse condition with non-uniform back ground and the day-light. For performance evaluation, the data set Is divided in to three sub sets of training (360videos), development (360 videos), and testing (480 videos). To generate the fake faces, a high resolution videos were taken for each person using a Canon Power Shot camera and a ni Phone 3GS camera, under the same illumination conditions.Three types of attacks were designed: (1) Print Attacks: High resolution pictures were printed on A4 paper and recaptured by cameras; (2)MobileAttacks:



2) *OULU-NPU Database:* OULU-NPU face presentation attack database consists of 4950 real access and attack videos that were recorded using front facing cameras of six different mobile phones (see, Fig.5). The real videos and attack materials were collected in three sessions with different illumination condition. The attack types considered in the OULU-NPU data base are print *and* video-replay.

IV. RESULTS

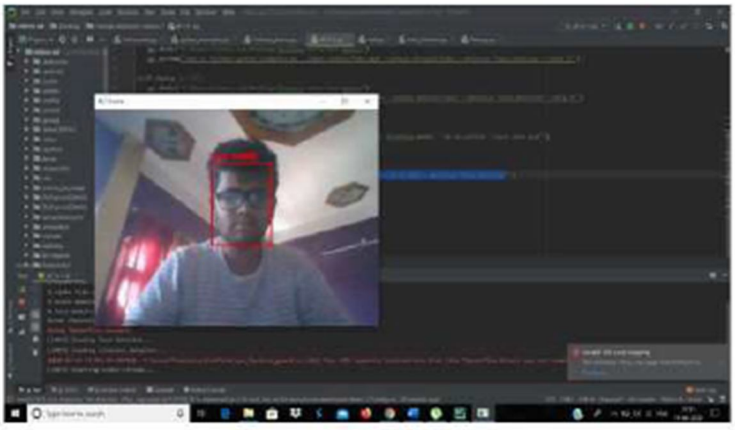


Figure : Detection of real image

As you can see in the above figure a real face is detected

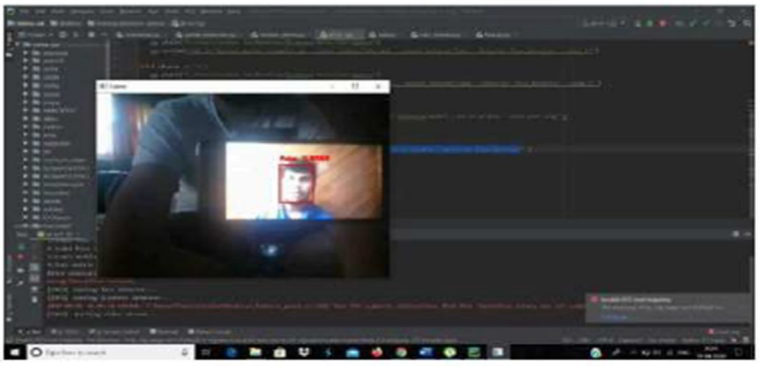


Figure: Detection of fake image

As you can see in the above figure a fake face image is detected

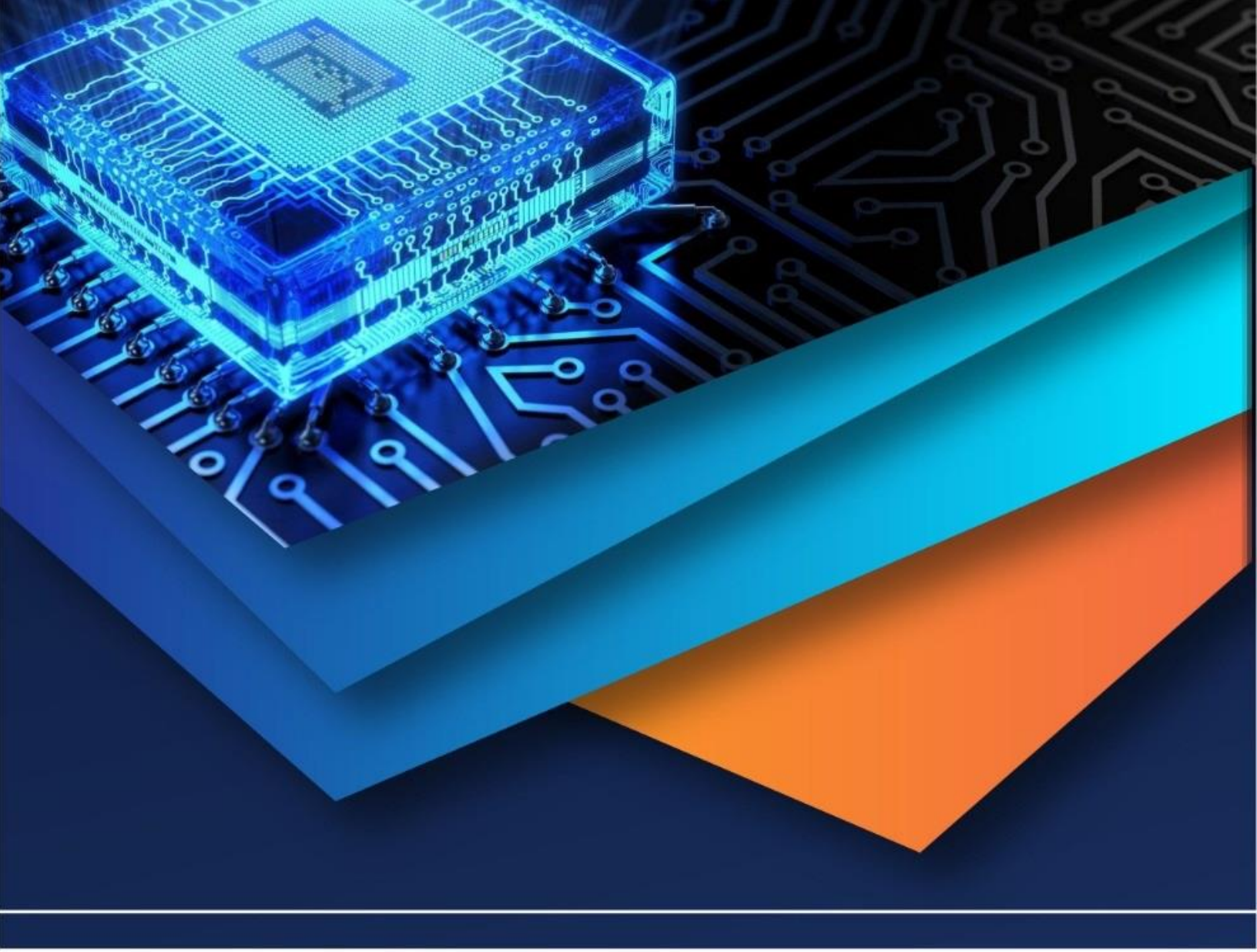
V. CONCLUSIONS

For face spoofing detection, the authors suggest a novel attention-based two-stream convolutional network (TSCNN) that makes use of RGB and Multi-Scale Retinex (MSR) data that are taken from deep learning models like MobileNet and ResNet-18. The attention-based fusion methodology, which adaptively integrates the RGB and MSR data to enhance the model's discriminative power, particularly under different lighting circumstances, is the method's main contribution. The method performs competitively in both intra- and inter-database testing situations when tested on three well-known face spoofing databases: CASIA-FASD, REPLAY-ATTACK, and OULU-NPU. The findings demonstrate how the attention fusion mechanism successfully integrates complimentary features to improve the model's capacity to differentiate between real and spoof faces. The suggested model's robustness is further demonstrated by the cross-database evaluations, which show notable gains in generalization to novel, untested datasets. In a variety of illumination conditions, the combination of RGB and MSR characteristics works very well to counteract the difficulties presented by face spoofing attempts. Overall, the study demonstrates how well the TSCNN architecture works in practical face anti-spoofing applications and offers insightful information about how to combine several feature domains to enhance model performance in security-sensitive situations.

REFERENCES

[1] J. Li, Y. Wang, and A. K. Jain, “Live face detection based on the analysis of fourier spectra,” Proc Spie, vol. 5404, pp. 296–303, 2004.
[2] X. Tan, Y. Li, J. Liu, and L. Jiang, “Face liveness detection from a single image with sparse low rank bilinear discriminative model,” in Computer Vision - ECCV 2010 - 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part VI, 2010, pp. 504–517.
[3] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, “A face antispoofing database with diverse attacks,” in Iapr International Conference on Biometrics, 2012, pp. 26–31.
[4] J. Galbally and S. Marcel, “Face anti-spoofing based on general image quality assessment,” in 22nd International Conference on Pattern Recognition, ICPR 2014, Stockholm, Sweden, August 24-28, 2014, 2014, pp. 1173–1178.
[5] D. Wen, H. Han, and A. K. Jain, “Face spoof detection with image distortion analysis,” IEEE Trans. Information Forensics and Security, vol. 10, no. 4, pp. 746–761, 2015.
[6] J. Yan, Z. Zhang, Z. Lei, D. Yi, and S. Z. Li, “Face liveness detection by exploring multiple scenic clues,” in 12th International Conference on Control Automation Robotics & Vision, ICARCV 2012, Guangzhou, China, December 5-7, 2012, 2012, pp. 188–193.
[7] K. Patel, H. Han, A. K. Jain, and G. Ott, “Live face video vs. spoof face video: Use of moiré patterns to detect replay video attacks,” in International Conference on Biometrics, ICB 2015, Phuket, Thailand, 19-22 May, 2015, 2015, pp. 98–105.
[8] Z. Boulkenafet, J. Komulainen, and A. Hadid, “Face spoofing detection using colour texture analysis,” IEEE Trans. Information Forensics and Security, vol. 11, no. 8, pp. 1818–1830, 2016.
[9] Z. Boulkenafet, J. Komulainen, X. Feng, and A. Hadid, “Scale space texture analysis for face anti-spoofing,” in International Conference on Biometrics, ICB 2016, Halmstad, Sweden, June 13-16, 2016, 2016, pp. 1–6.
[10] Z. Boulkenafet, J. Komulainen, and A. Hadid, “Face anti-spoofing based on color texture analysis,” in 2015 IEEE International Conference on Image Processing, ICIP 2015, Quebec City, QC, Canada, September 27-30, 2015, 2015, pp. 2636–2640.
[11] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971–987, 2002.
[12] J. Määttä, A. Hadid, and M. Pietikäinen, “Face spoofing detection from single images using micro-texture analysis,” in 2011 IEEE International Joint Conference on Biometrics, IJCB 2011, Washington, DC, USA, October 11-13, 2011, 2011, pp. 1–7.
[13] D. Menotti, G. Chiachia, A. da Silva Pinto, W. R. Schwartz, H. Pedrini, A. X. Falcão, and A. Rocha, “Deep representations for iris, face, and fingerprint spoofing detection,” IEEE Trans. Information Forensics and Security, vol. 10, no. 4, pp. 864–879, 2015.
[14] Chinnaiah, M. C., Sanjay Dubey, N. Janardhan, Venkata Pathi, K. Nandan, and M. Anusha. "Analysis of pitta imbalance in young indian adult using machine learning algorithm." In 2022 2nd International conference on intelligent technologies (CONIT), pp. 1-5. IEEE, 2022.
[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States., 2012, pp. 1106–1114.
[16] G. Pan, L. Sun, Z. Wu, and S. Lao, “Eyeblink-based anti-spoofing in face recognition from a generic webcamera,” in IEEE 11th International Conference on Computer Vision, ICCV 2007, Rio de Janeiro, Brazil, October 14-20, 2007, 2007, pp. 1–8.
[17] G. Pan, L. Sun, Z. Wu, and Y. Wang, “Monocular camera-based face liveness detection by combining eyeblink and scene context,” Telecommunication Systems, vol. 47, no. 3-4, pp. 215–225, 2011.
[18] L. Sun, G. Pan, Z. Wu, and S. Lao, “Blinking-based live face detection using conditional random fields,” in Advances in Biometrics, International Conference, ICB 2007, Seoul, Korea, August 27-29, 2007, Proceedings, 2007, pp. 252–260.
[19] A. Anjos, M. M. Chakka, and S. Marcel, “Motion-based countermeasures to photo attacks in face recognition,” Iet Biometrics, vol. 3, no. 3, pp. 147–158, 2014.
[20] Gudipelly Mamatha, B.Manjula and P.Venkatapathi “Intend Innovative Technology For Recognition Of Seat Vacancy In Bus” International Journal of Research and Analytical Reviews, Volume 6, Issue 02, April-June.-2019, ISSN: 2349-5138
[21] Y. Kim, J. Na, S. Yoon, and J. Yi, “Masked fake face detection using radiance measurements,” J Opt Soc Am A Opt Image Sci Vis, vol. 26, no. 4, pp. 760–766, 2009.
[22] A. da Silva Pinto, H. Pedrini, W. R. Schwartz, and A. Rocha, “Face spoofing detection through visual codebooks of spectral temporal cubes,” IEEE Trans. Image Processing, vol. 24, no. 12, pp. 4726–4740, 2015.

[23] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," IEEE Trans. Image Processing, vol. 6, no. 7, pp. 965–976, 1997.
[24] Sudhakar Alluri, Karnati Mahidhar, Kalluru Kavya, Dulam Srija, P.Venkatapathi "High Performance Of Smartcard With Iris Recognition For High Security Access Environment In Python Tool" Industrial Engineering Journal ISSN: 0970-2555; Volume : 52, Issue 10, No. 2, October : 2023
[25] Y. Atoum, Y. Liu, A. Jourabloo, and X. Liu, "Face anti-spoofing using patch and depth-based cnns," in 2017 IEEE International Joint Conference on Biometrics, IJCB 2017, Denver, CO, USA, October 1-4, 2017, 2017, pp. 319–328.
[26] Z. Xu, S. Li, and W. Deng, "Learning temporal features using LSTMCNN architecture for face anti-spoofing," in 3rd IAPR Asian Conference on Pattern Recognition, ACPR 2015, Kuala Lumpur, Malaysia, November 3-6, 2015, 2015, pp. 141–145.
[27] S Venkatapathi Pallam, Vasudev Biyyala, Chandra Shekar Jadapally, Ramsai Nalla, Dr. Sudhakar Alluri "Doctors Assistive System Using Augmented Reality Glass Critical Analysis" International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; Volume 11 Issue X Oct 2023
[28] Z. Zhang, D. Yi, Z. Lei, and S. Z. Li, "Face liveness detection by learning multispectral reflectance distributions," in Ninth IEEE International Conference on Automatic Face and Gesture Recognition (FG 2011), Santa Barbara, CA, USA, 21-25 March 2011, 2011, pp. 436–441. [Online]. Available: https://doi.org/10.1109/FG.2011.5771438
[29] J. Komulainen, A. Hadid, M. Pietikäinen, A. Anjos, and S. Marcel, "Complementary countermeasures for detecting scenic face spoofing attacks," in International Conference on Biometrics, ICB 2013, 4-7 June, 2013, Madrid, Spain, 2013, pp. 1–7.
[30] Venkatapathi, Pallam, Habibulla Khan, S. Srinivasa Rao, and Govardhani Immadi. "Cooperative spectrum sensing performance assessment using machine learning in cognitive radio sensor networks." Engineering, Technology & Applied Science Research 14, no. 1 (2024): 12875-12879.
[31] MARKING, N.V.W., 2014. MULTI-WAVELET BASED ON NON-VISIBLE WATER MARKING.
[32] Sudhakar Alluri, Komireddy Shreyas, Lingampally Ganesh, Mangali Vamshi, Venkatapath Pallam "A System Based in Virtual Reality to Manage Flood Damage" International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; Volume 11 Issue XI Nov 2023



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY