



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.524

Volume 14, Issue 1, January 2025

A Review Paper on Face Anti-Spoofing System using Deep Learning Technique

Shivani Kapse¹, Ram Kumar Sahu²

M. Tech. Scholar, Department of Computer Science and Engineering, TIT-Excellence, Bhopal, India¹

Assistant Professor, Department of Computer Science and Engineering, TIT-Excellence, Bhopal, India²

ABSTRACT: Face spoofing detection is one of the well-studied problems in computer vision. Face recognition has become a widely adopted technique in biometric authentication systems. In face recognition based authentication techniques, the system first recognized the person to verify the legitimacy of the user before granting access to the system resources. The system must be able to determine the liveness of the person in front of the camera, for example, by recognizing the face and denying the types of face presentation attacks related to photographs, videos and the 3D mask of the targeted person. Attackers try to directly or indirectly, masquerade the biometric system as another person by forging biometric traits and get unauthorized access. This work studies computer vision-based feature extraction techniques for real and spoof face imaging and combines different features in the area of face anti-spoofing.

KEYWORDS: Face Spoofing, Face Recognition, Machine Learning (ML)

I. INTRODUCTION

Every day we come across biometric systems for authentication and many other purposes. The biometrics includes the face, iris, and fingerprint and palm recognition. But the criminals very accurately stimulate the biometrics of the genuine users to access their private systems and this process is called a spoofing attack. It is comparatively easy for the criminals to spoof iris, fingerprint and palm to face spoofing [1, 2]. Face recognition is universally used to ensure authenticity. Mobile unlocking, to safe deposit, locker unlocking exercised face recognition technology. However, hackers or the criminals pretend to pose as a genuine user in three peculiar ways for instance photo attacks, replay video attacks and 3-D mask attack. Presenting themselves as a real user by the way of using printed photo face is entitled as photo attacks [3].

To prohibit impostors from this attack, highlight removal method, User interaction method had been formalized. However, impostors find a way to avail the resources. They exploit motion picture and video replay attacks, 3-D mask attacks to beneficially acquire the user's possessions. Hence, various states-of-art methods have been evolved to retard the imitators. The existing methodologies include Face liveness detection, Texture analysis, User Interaction (Eye and head movements), Image Quality Assessment, Image Distortion Analysis, Depth sensing, Convolutional neural networks [4, 5].

The success of machine learning in computer vision provides a strong hope of its use in biometrics for the face antispoofing and other similar applications. With these neural architectures, it is easy for us to solve complex solutions with better performance and features. They work with high-level features and are self-trained from the training dataset. The machine learning is able to deal with digital images with better performances. Many kinds of research are done with neural layers to achieve more accuracy and increase its performance [6].

II. LITERATURE REVIEW

Enoch Solomon et al. [1], face acknowledgment innovation have been broadly utilized because of the comfort it gives. Notwithstanding, face acknowledgment is helpless against caricaturing assaults which restricts its use in touchy application regions. This work presents an original face hostile to ridiculing framework, FASS, that breakers result of two classifiers. One, arbitrary timberland, utilizes the distinguished by us seven no-reference picture quality highlights got from face pictures and its outcomes are intertwined with a profound learning classifier results that utilizes whole face pictures as information. Broad investigations were performed to contrast FASS and cutting edge enemy of mocking frameworks on five benchmark datasets: Replay-Assault, CASIA-MFSD, MSU-MFSD, OULU-NPU and

SiW. The outcomes show that FASS outflanks all face hostile to ridiculing frameworks in light of picture quality highlights and is additionally more precise than large numbers of the best in class frameworks in light of profound learning.

Woubie et al. [2], replay assaults are endeavors to get fake admittance to a programmed speaker confirmation framework. In this paper, we explore the helpfulness of voice quality highlights to distinguish replay assaults. The voice quality elements are utilized along with the cutting edge steady Q cepstral coefficients (CQCC) highlights. The two capabilities are intertwined at the score level. Subsequently, the log-probability scores assessed from the two capabilities are directly weighted to get a solitary combined score. The melded score is utilized to characterize whether a given discourse test is real or mock. Our analyses with the ASVspoof 2017 dataset exhibit that the combination of log-probability scores removed from the CQCC and voice quality elements move along the Equivalent Mistake Rate (EER) contrasted with the standard framework which depends just on CQCC highlights.

Fang et al. [3], facial coverings have become one of the principal strategies for diminishing the transmission of Coronavirus. This makes face acknowledgment (FR) a difficult undertaking since covers conceal a few discriminative highlights of countenances. Additionally, face show assault location (Cushion) is vital to guarantee the security of FR frameworks. As opposed to the developing number of concealed FR studies, the effect of face covered assaults on Cushion has not been investigated. Accordingly, we present novel assaults with genuine facial coverings put on introductions and assaults with subjects wearing veils to mirror what is happening. Moreover, this study researches the impact of concealed assaults on Cushion execution by utilizing seven cutting edge Cushion calculations under various exploratory settings. We additionally assess the weakness of FR frameworks to covered assaults. The trials show that truly concealed assaults represent a serious danger to the activity and security of FR frameworks.

Muhammad et al. [4], customary 2D face biometric frameworks are helpless against show assaults performed with various face antiquities, e.g., printouts, video-replays and wearable 3D covers. The exploration center in face show assault recognition (Cushion) has been as of late moving towards start to finish gaining of profound portrayals straightforwardly from commented on information as opposed to planning hand-created (low-level) highlights. In any case, even the cutting edge profound learning based face Cushion models have shown unacceptable speculation execution while confronting obscure assaults or procurement conditions because of absence of agent preparing and tuning information accessible in the current public benchmarks. To reduce this issue, we propose a video pre-handling strategy called Fleeting Grouping Examining (TSS) for 2D face Cushion by eliminating the assessed between outline 2D relative movement in the view and encoding the appearance and elements of the subsequent smoothed video succession into a solitary RGB picture. Moreover, we influence the highlights of a Convolutional Brain Organization (CNN) by presenting a self-managed portrayal learning plan, where the marks are consequently created by the TSS technique as the balanced out outlines gathered over video clasps of various worldly lengths give the management. The learnt highlight portrayals are then adjusted for the downstream undertaking utilizing marked face Cushion information. Our broad tests on four public benchmarks, to be specific Replay-Assault, MSU-MFSD, CASIA-FASD and OULU-NPU, exhibit that the proposed system gives promising speculation ability and support further concentrate in this area.

Wanling Zhang et al. [5], a face against mocking procedure by utilizing DWT (Discrete Wavelet Transform), LBP (Local Binary Pattern) and DCT (Discrete Cosine Transform) with a SVM classifier to assess whether a video is legitimate. First and foremost, the DWT highlights are delivered by deteriorating some chosen outlines into various recurrence parts at the multi-goal blocks. Besides, the DWT-LBP highlights are produced to address spatial data of the squares by interfacing LBP histograms of the DWT blocks in each edge on a level plane. Then, at that point, the DWT-LBP-DCT highlights with the fleeting data of a video document are accomplished by performing DCT procedure on the DWT-LBP elements of those chose outlines upward. Accordingly, these took advantage of DWT-LBP-DCT highlights have the ability to address the recurrence spatial–fleeting data of a video. At last, the SVM classifier with RBF piece is prepared for face against mocking. Contrasted and past phenomenal works, exploratory outcomes on two benchmark information bases (REPLAY-ATTACK and CASIA-FASD) have shown the proposed approach has better discovery execution.

Z. Wang et al. [6], face against ridiculing is basic to the security of face acknowledgment frameworks. Profundity directed learning has been demonstrated as one of the best techniques for face against mocking. Notwithstanding the extraordinary achievement, most past works actually define the issue as a solitary casing perform multiple tasks one by just enlarging the misfortune with profundity, while dismissing the nitty gritty fine-grained data and the exchange

between facial profundities and moving examples. Conversely, we plan another way to deal with recognize show assaults from numerous edges dependent on two experiences: 1) nitty gritty discriminative pieces of information (e.g., spatial angle extent) among living and mocking face might be disposed of through stacked vanilla convolutions, and 2) the elements of 3D moving appearances give significant insights in distinguishing the ridiculing faces. The proposed technique can catch discriminative subtleties through Residual Spatial Gradient Block (RSGB) and encode spatio-worldly data from Spatio-Temporal Propagation Module (STPM) effectively. Also, an original Contrastive Depth Loss is introduced for more precise profundity oversight. To survey the adequacy of our strategy, we likewise gather a Double-modular Anti-ridiculing Dataset (DMAD) which gives genuine profundity to each example. The investigations show that the proposed approach accomplishes best in class results on five benchmark datasets including OULU-NPU, SiW, CASIA-MFSD, Replay-Attack, and the new DMAD.

S. Zhang et al. [7], face against mocking is fundamental to keep face acknowledgment frameworks from a security break. A significant part of the advances have been made by the accessibility of face against caricaturing benchmark datasets lately. Notwithstanding, existing face against mocking benchmarks have set number of subjects (≤ 170) and modalities (≤ 2), which thwart the further improvement of the scholarly local area. To work with face against caricaturing research, we present an enormous scope multi-modular dataset, in particular CASIA-SURF, which is the biggest openly accessible dataset for face hostile to satirizing as far as the two subjects and visual modalities. In particular, it comprises of 1,000 subjects with 21,000 recordings and each example has 3 modalities (i.e., RGB, Depth and IR). We additionally give an estimation set, assessment convention and preparing/approval/testing subsets, fostering another benchmark for face hostile to satirizing. Also, we present a new multi-modular combination strategy as standard, which performs highlight re-weighting to choose the more enlightening channel highlights while smothering the less valuable ones for every modular. Broad investigations have been led on the proposed dataset to check its importance and speculation ability.

X. Yang et al. [8], Face against mocking is a significant errand in full-stack face applications including face discovery, check, and acknowledgment. Past approaches construct models on datasets which don't recreate this present reality information well (e.g., little scale, immaterial difference, and so on) Existing models may depend on helper data, which keeps these antispoofing arrangements from summing up well by and by. In this paper, we present an information assortment arrangement alongside an information amalgamation strategy to reenact advanced medium-based face satirizing assaults, which can undoubtedly assist us with getting a enormous measure of preparing information well mirroring this present reality situations. Through taking advantage of a clever Spatio-Temporal Against Spoof Network (STASN), we can push the exhibition on open face hostile to mocking datasets over state-of-the-workmanship strategies by a huge degree. Since the proposed model can naturally take care of discriminative locales, it makes breaking down the practices of the organization conceivable. We direct broad investigations and show that the proposed model can recognize parody faces by extricating highlights from an assortment of locales to search out unobtrusive confirmations like boundaries, moire examples, reflection antiques, and so on.

J. Zhang et al. [9], face biometrics have accomplished exceptional execution over the previous many years, yet unforeseen caricaturing of the static faces represents a danger to data security. There is an expanding interest for steady and discriminative organic modalities which are difficult to be copied and misled. Discourse driven 3D facial movement is an unmistakable and quantifiable conduct signature that is promising for biometrics. In this paper, we propose a clever 3D behaviometrics structure dependent on a "3D visual password" got from discourse driven 3D facial elements. The 3D facial elements are mutually addressed by 3D-keypoint-based estimations and 3D shape fix highlights, extricated from both static and discourse driven powerful districts. A group of subject-explicit classifiers are then prepared over chosen discriminative highlights, which considers a discriminant discourse driven 3D facial elements portrayal. We build the primary freely accessible Speech-driven 3D Facial Motion dataset (S3DFM) that incorporates 2D-3D face video in addition to sound examples from 77 members. The exploratory outcomes on the S3DFM show that the proposed pipeline accomplishes a face distinguishing proof pace of 96.1%. Definite conversations are introduced, concerning hostile to satirizing, head present variety, video outline rate, and pertinence cases. We additionally give examination with different baselines on "profound" and "shallow" 2D face highlights.

X. Li et al. [10], face biometric frameworks are helpless against spoofing attacks. Such assaults can be acted in numerous ways, including presenting a falsified picture, video or 3D cover of a substantial user. A broadly involved methodology for separating certified appearances from fake ones has been to catch their intrinsic contrasts in (2D or 3D) surface utilizing neighborhood descriptors. One impediment of these methods is that they might fall flat assuming an

inconspicuous assault type, for example a highly reasonable 3D veil which looks like genuine skin surface, is used in spoofing. Here we propose a strong enemy of spoofing method by recognizing beat from face recordings. In view of the way that a pulse signal exists in a genuine living face however not in any mask or print material, the strategy could be a summed up solution for face liveness recognition. The proposed strategy is evaluated first on a 3D cover spoofing information base 3DMAD to demonstrate its viability in recognizing 3D veil assaults. More importantly, our cross-data set analysis with great REAL-F masks shows that the beat based technique can recognize even the previously inconspicuous veil type while surface based strategies fail to sum up past the improvement information. At long last, we propose a powerful course framework joining two corresponding assault specific parody indicators, for example use beat location against print attacks and shading surface investigation against video assaults.

L. Feng et al. [11], numerous characteristic explicit countermeasures to confront satirizing assaults have been created for security of face validation. In any case, there is no better face against satirizing strategy than manage each sort of caricaturing assault in shifting situations. To further develop the speculation capacity of face hostile to ridiculing approaches, an extendable multi-signs incorporation structure for face against satirizing utilizing a progressive neural organization is proposed, which can intertwine picture quality signals and movement prompts for liveness discovery. Shearlet is used to foster a picture quality-based liveness highlight. Thick optical stream is used to remove movement based liveness highlights. A bottleneck include combination technique can coordinate diverse liveness includes adequately. The proposed approach was assessed on three public face hostile to caricaturing information bases. A half all out blunder rate (HTER) of 0% and an equivalent mistake rate (EER) of 0% were accomplished on both REPLAY-ATTACK information base and 3D-MAD data set. An EER of 5.83% was accomplished on CASIA-FASD information base.

Problem Identification:-

The facial recognition frameworks are highly vulnerable to spoofing attacks and this vulnerability generates an effective security concerned issues in biometric domain.

- Moreover, some of the earlier proposed approaches have attained attractive results with intra test evaluation done to detect the face spoofing attack.
- However, this impact is considered as a major difficulty in the highly focused biometric anti-spoofing research domain.
- A biometric recognition system that is used to recognize the genuine entries, extracts the features among a set of users input image [7, 8].
- The person's identity is established by comparing the extracted feature set against a set of feature sets in the database.
- The object is to increase accuracy and to decrease the false positives occurred in the preceding papers.
- To decrease the complexity in the cases of high dimensional data. Also make data sets available for face spoofing attack while suppressing the features with increase detection [9].

III. CHALLENGES OF FACE RECOGNITION

Pose/Viewpoint: - It is a particular attitude or instance, especially with the hope of impressing others and to present oneself insincerely. It is a movement in which the dancer steps, in any desired position, from one foot to the other and it is a straight knee onto the flat foot.

Facial Expression: - A facial expression is one or more motions or positions of the muscles beneath the skin of the face. According to one set of controversial theories, these movements convey the emotional state of an individual to observers. Facial expressions are a form of nonverbal communication.

Occlusion: - Occlusion occurs when a portion of the picture visible on one image is occluded in the other by the scene itself or, a section of the scene near the image boundary moves out of the field of persuasion on the other picture.

Image Orientation: - The image-orientation property allows us to rotate an image in multiples of 90 degree. Values like 47 degree are rounded to the nearest multiple of 90.

Spoofing: - The word "spoof" means to trick, or deceive. Therefore, spoofing refers to tricking or deceiving computer systems or other computer users. This is typically done by hiding one's identity or faking the identity of another user on the Internet.

Spoofing Attack: It is a situation in which one person (or) program successfully masquerades as another by falsifying data, thereby gaining an illegitimate advantage. Some of the methods that are used for spoofing facial recognition systems are by presenting fake faces with masks, Printed photos, Videos etc.

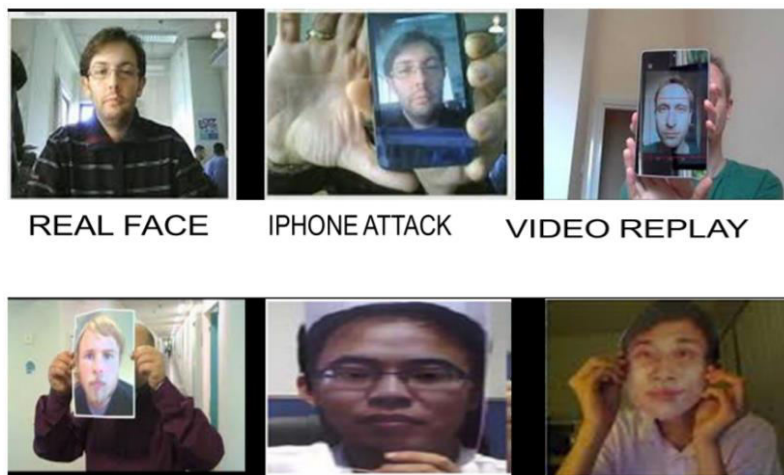


Figure 1: Types of Spoofing Attack [10]

IV. FACE ANTI-SPOOFING DETECTION USING DEEP LEARNING

Spoof detection is able to provide generalized performance and be able to be utilized in improving the effectiveness of the system particularly when working with noisy images. The anti-spoofing community should also consider engaging in new fundamental research regarding the biological dimension of biometric traits, in order to break with the current popular trend embraced by many of the latest research where some well-known sets of features are extracted from images in public databases and passed through a classifier. Classification plays again a vital role in recognition, some of the optimization techniques can be used for enhancing the performance of the system. Spoof detection measures appear to play an important role in human vision and could be incorporated into feature descriptors in a more general form than the single spatial frequency used by the current descriptors [10, 11].

DEEP LEARNING

Deep learning (DL) is a subset of machine learning (ML) procedures zeroed in on order undertakings and transformative calculations [14]. There are three sorts of learning: supervised learning, semi-supervised and unsupervised. DL structures consolidating DL models, completely associated networks, repetitive brain organizations, and fake brain networks were utilized in fields including AI, man-made reasoning, PC vision, information examination, understood, virtual entertainment site separating, computational semantics, computational science, drug plan, data recovery, and clear outline, among others [15]. Information obtaining and decentralized authoritative foundation in natural frameworks impacted fake ANNs. ANNs change from the human mind in more ways than one. Specifically, brain networks are steady and emblematic, though most working substances' natural minds are dynamic and simple.

Profound learning gets its name from the way that it utilizes many layers in the organization. Early exploration exhibited that a direct perceptron can't be utilized as an all inclusive classifier yet that an organization with a non-polynomial information layer and one unreasonable width stowed away layer may [15]. Profound learning is a later variation including many layers of limited size, considering practical application and improvement while keeping up with hypothetical subjectivity under gentle circumstances. For execution, teachability, and clarity, profound learning structures are additionally permitted to be different and wander away generally from deductively informed connectionist models, consequently the "coordinated" segment [16].

Most of new profound learning procedures center around AI, particularly CNNs. They may likewise incorporate propositional recipes or dormant factors organized layer-wise in profound generative models like profound conviction organizations and profound Boltzmann machines. Each degree of profound learning figures out how to transform the information it gets into a somewhat more conceptual and composite portrayal. The crude contribution to a picture acknowledgment program could be a grid of pixels; the main delegate layer could digest the pixels and encode edges; the subsequent layer could form and encode edge courses of action; the third layer could encode a nose and eyes, and the fourth layer could perceive that the picture contains a face. Significantly, a profound learning calculation might sort out which highlights have a place with which level all alone.

More layers, then again, don't work on the organization's capacity to inexact capabilities. Additional layers help in learning the highlights successfully in light of the fact that profound models can separate preferable elements over shallow models. Profound convolutional layers can build profound learning models in CNN. The DL can support the deconstruction of these reflections and the ID of which elements further develop results. Profound learning strategies wipe out include designing for managed learning undertakings by changing over information into minimal element vectors closely resembling factor stacking and creating layered structures that lessen overt repetitiveness. Solo gaining errands might profit from profound learning calculations. This is a critical benefit since unlabeled information is more copious than named information. ANN and profound conviction networks are the two fundamental brain network that works like solo learning approach. There are one or two sorts of profound learning calculations, which are referenced underneath [8].

V. CONCLUSION

In this paper, a framework is review with machine learning to detect the spoofed face. The image is given as an input to the machine learning framework. Since we used the strategy of training the fake and real face dataset separately it is easy and efficient for the system to classify the image between real and fake when an input image is given. The flexibility of the neural layers is the main reason for the better performance in detecting the spoofed images.

REFERENCES

1. Enoch Solomon and Krzysztof J. Cios, "FASS: Face Anti-Spoofing System Using Image Quality Features and Deep Learning", *Modern Computer Vision and Image Analysis*, Vol. 12, Issue 10, Electronics 2023.
2. Woubie, A.; Bäckström, T. Voice Quality Features for Replay Attack Detection. In *Proceedings of the 2022 30th European Signal Processing Conference (EUSIPCO)*, Belgrade, Serbia, 29 August–2 September 2022; pp. 384–388.
3. Fang, M.; Damer, N.; Kirchbuchner, F.; Kuijper, A. Real masks and spoof faces: On the masked face presentation attack detection. *Pattern Recognit.* 2022, 123, 108398.
4. Muhammad, U.; Yu, Z.; Komulainen, J. Self-supervised 2D face presentation attack detection via temporal sequence sampling. *Pattern Recognit. Lett.* 2022, 156, 15–22.
5. Wanling Zhang and Shijun Xiang, "Face anti-spoofing detection based on DWT-LBP-DCT features", *Signal Processing: Image Communication*, Elsevier, 2020.
6. Z. Wang, Z. Yu, C. Zhao, X. Zhu, Y. Qin, Q. Zhou, F. Zhou, Z. Lei, Deep spatial gradient and temporal depth learning for face anti-spoofing, in: *The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
7. S. Zhang, X. Wang, A. Liu, C. Zhao, J. Wan, S. Escalera, H. Shi, Z. Wang, S.Z. Li, A dataset and benchmark for large-scale multi-modal face anti-spoofing, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 919–928.
8. X. Yang, W. Luo, L. Bao, Y. Gao, D. Gong, S. Zheng, Z. Li, W. Liu, Face antispoofing: Model matters, so does data, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 3507–3516.
9. J. Zhang, R.B. Fisher, 3D Visual passcode: Speech-driven 3D facial dynamics for biometrics, *Signal Process.* 160 (2019) 164–177.
10. X. Li, J. Komulainen, G. Zhao, P.C. Yuen, M. Pietikainen, Generalized face antispoofing by detecting pulse from face videos, in: *International Conference on Pattern Recognition*, 2017.
11. L. Feng, L.M. Po, Y. Li, X. Xu, Y. Fang, C.H. Cheung, K.W. Cheung, Integration of image quality and motion cues for face anti-spoofing, *J. Vis. Commun. Image Represent.* 38 (C) (2016) 451–460.

12. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
13. M. Asim, Z. Ming, M.Y. Javed, CNN Based spatio-temporal feature extraction for face anti-spoofing, in: 2017 2nd International Conference on Image, Vision and Computing (ICIVC), IEEE, 2017, pp. 234–238.
14. E. Marasco, A. Ross, A survey on antispoofing schemes for fingerprint recognition systems, *Acm Comput. Surv.* 47 (2) (2015) 1–36.
15. S. Tirunagari, N. Poh, D. Windridge, A. Iorliam, N. Suki, A.T. Ho, Detection of face spoofing using visual dynamics, *IEEE Trans. Inf. Forensics Secur.* 10 (4) (2015) 762–777.
16. P. Allan, P. Helio, S. William Robson, R. Anderson, Face spoofing detection through visual codebooks of spectral temporal cubes, *IEEE Trans. Image Process.* 24 (12) (2015) 4726–4740.
17. S. Tirunagari, N. Poh, D. Windridge, A. Iorliam, N. Suki, A.T. Ho, Detection of face spoofing using visual dynamics, *IEEE Trans. Inf. Forensics Secur.* 10 (4) (2015) 762–777.
18. P.L.D. Leon, M. Pucher, J. Yamagishi, I. Hernaez, I. Saratxaga, Evaluation of speaker verification security and detection of HMM-based synthetic speech, *IEEE Trans. Audio Speech Lang. Process.* 20 (8) (2012) 2280–2290.
19. H. Wendt, S.G. Roux, S. Jaffard, P. Abry, Wavelet leaders and bootstrap for multifractal analysis of images, *Signal Process.* 89 (6) (2009) 1100–1114.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details