

- [34] Wu Z, Das RK, Yang J, Li H (2020) Light convolutional neural network with feature genuinization for detection of synthetic speech attacks. arXiv preprint arXiv:2009.09637.
- [35] Bartusiak ER, Delp EJ (2021) Frequency domain-based detection of generated audio. *Electronic Imaging* 2021(4): 273-1-273-7.
- [36] Zhang Y, Jiang F, Duan Z (2021) One-class learning towards synthetic voice spoofing detection. *IEEE Signal Processing Letters* 28: 937-941.
- [37] Gomez-Alanis A, Peinado AM, Gonzalez JA, Gomez AM (2019, September) A light convolutional GRU-RNN deep feature extractor for ASV spoofing detection. In *Proc. Interspeech*, Vol 2019, pp 1068-1072.
- [38] Huang L, Pun CM (2020) Audio replay spoof attack detection by joint segment-based linear filter bank feature extraction and attention-enhanced DenseNet-BiLSTM network. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 28: 1813-1825.
- [39] Hua G, Bengjinteoh A, Zhang H (2021) Towards end-to-end synthetic speech detection. *IEEE Signal Processing Letters*.
- [40] Mittal T, Bhattacharya U, Chandra R, Bera A, Manocha D (2020, October) Emotions don't lie: an audio-visual deepfake detection method using affective cues. In *Proceedings of the 28th ACM international conference on multimedia*, pp 2823-2832.
- [41] Chugh K, Gupta P, Dhall A, Subramanian R (2020, October) Not made for each other: audio-visual dissonance-based deepfake detection and localization. In *Proceedings of the 28th ACM International Conference on Multimedia*, pp 439-447.
- [42] Gu Y, Zhao X, Gong C, Yi X (2020, November) Deepfake video detection using audio-visual consistency. In *International Workshop on Digital Watermarking*, Springer, Cham, pp 168-180.
- [43] Zhou Y, Lim SN (2021) Joint audio-visual deepfake detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp 14800-14809.
- [44] Gino J (2021) Audio-video deepfake detection through emotion recognition.
- [45] Mozley K (2021) Don't believe everything you hear: combining audio and visual cues for deepfake detection.
- [46] Blankertz B (2001) The constant Q transform. URL: http://doc.ml.tu-berlin.de/bbci/material/publications/Bla_constQ.Pdf.
- [47] Schörkhuber C, Klapuri A, Sontacchi A (2012, September) Pitch shifting of audio signals using the constant-q transform. In *Proc. of the DAFx Conference*.
- [48] Todisco M, Delgado H, Evans NW (2016, June) A new feature for automatic speaker verification anti-spoofing: constant Q cepstral coefficients. In *Odyssey*, Vol 2016, pp 283-290.
- [49] Wang M, Wang R, Zhang XL, Rahardja S (2019, November) Hybrid constant-Q transform based CNN ensemble for acoustic scene classification. In *2019 Asia-Pacific signal and information processing association annual summit and conference (APSIPA ASC)*, pp 1511-1516.
- [50] Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 1251-1258.
- [51] Ghosh S, Pal A, Jaiswal S, Santosh KC, Das N, Nasipuri M (2019) SegFast-V2: Semantic image segmentation with less parameters in deep learning for autonomous driving. *International Journal of Machine Learning and Cybernetics* 10(11): 3145-3154.
- [52] Szegedy C, Ioffe S, Vanhoucke V, Alemi AA (2017, February) Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*.
- [53] Arfianti UI, Novitasari DCR, Widodo N, Hafiyusholeh M, Utami WD (2021) Sunspot number prediction using gated recurrent unit (GRU) algorithm. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)* 15(2): 141-152.
- [54] Le XH, Ho HV, Lee G (2019, September) Application of gated recurrent unit (GRU) network for forecasting river water levels affected by tides. In *International Conference on Asian and Pacific Coasts*, Springer, Singapore, pp 673-680.
- [55] Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E (2016, June) Hierarchical attention networks for document classification. In *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies*, pp 1480-1489.
- [56] Hossin M, Sulaiman MN (2015) A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process* 5(2): 1.
- [57] Achlioptas P (2019) Stochastic gradient descent in theory and practice.
- [58] Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- [59] Ho Y, Wookey S (2019) The real-world-weight cross-entropy loss function: modeling the costs of mislabeling. *IEEE Access* 8: 4806-4813.
- [60] Khalid H, Kim M, Tariq S, Woo SS (2021b, October) Evaluation of an audio-video multimodal deepfake dataset using unimodal and multimodal detectors. In *Proceedings of the 1st workshop on synthetic multimedia-audiovisual deepfake generation and detection*, pp 7-15.
- [61] Aly M, Alotaibi NS (2022) A New Model to Detect COVID-19 Coughing and Breathing Sound Symptoms Classification from CQT and Mel Spectrogram Image Representation using Deep Learning. *International Journal of Advanced Computer Science and Applications*, pp 601-611.