Anti-spoofing Detection Based on Hierarchical Spatio-Temporal Representation

Souad Khellat-Kihel[1], Ahmed Tibermacine[2]

[1]University of Science and Technology Mohamed-Boudiaf, Oran, Algeria
[2]LESIA Laboratory, Department of Computer Science, Biskra University, BP 145 RP, 07000, Biskra, Algeria.

Abstract

A complete biometric system ensures that only authorized individuals can access mobile devices. Face biometrics is a natural, user-friendly, and non-intrusive authentication method, making it a powerful biometric trait. However, recent studies have revealed its vulnerability to spoofing attacks, including printed photos, video replays, and 3D face masks. This paper proposes using the HMAX model to detect spoofing based on facial texture analysis. HMAX, inspired by the biological visual processing chain from the retinal stage to the inferotemporal cortex, encodes facial features sensitive to expressions and gaze direction. We further enhance this method by integrating HMAX with Long Short-Term Memory (LSTM) networks to build a spatio-temporal representation of facial dynamics for improved spoof detection. Extensive experiments on standard datasets demonstrate the feasibility and effectiveness of the proposed approach compared to state-of-the-art algorithms. Several experiments performed on real face images from standard datasets, also compared with state of the art algorithms, demonstrate the feasibility of the proposed approach in real applications.

Keywords

Security, anti spoofing, LSTM, Spatio-temporal representation,

1. Introduction

Biometric systems have become an essential component of security and authentication processes, with face recognition being one of the most widely adopted modalities. However, the increasing reliance on face recognition technology has also made it a target for spoofing attacks, where malicious actors attempt to deceive the system using fake representations of a legitimate user. These attacks, known as Presentation Attacks (PAs), pose a significant threat to security-sensitive applications, such as access control, banking transactions, and airport security. Presentation Attack Detection (PAD), commonly referred to as anti-spoofing detection, has emerged as a critical area of research to counteract these threats[1]. The goal of PAD is to distinguish between genuine and fake faces using various techniques, including motion analysis, texture-based feature extraction, and deep learning approaches. Spoofing attacks can take multiple forms, including printed photos, digital screen replays, 3D masks, and even highly sophisticated deepfake-generated faces[2]. A well-documented real-world case occurred in 2011 when a passenger successfully boarded a flight from Hong Kong to Canada by disguising himself as an elderly man using a high-quality mask, exposing vulnerabilities in biometric security measures, as illustrated in Figure 1.

Figure 1: An example of spoofing in a real-world case.

To address these security concerns, researchers have explored various methodologies for PAD, categorized broadly into traditional and deep learning-based approaches. One of the earliest approaches to PAD involves analyzing the texture and quality of facial images. Artur Costa-Pazo et al.[1] introduced two algorithms that leverage image-quality measures and texture analysis with Gabor-Jets filters for spoofing detection. Their study found that using an SVM-RBF classifier[3, 4, 5] resulted in an Equal Error Rate (EER) of 2.68%, demonstrating the effectiveness of such feature-based methods. Similarly, Boulkenafet et al.[6] proposed a color texture analysis technique utilizing the Color Local Binary Pattern (LBP) descriptor, which captures fine luminance variations between genuine and spoofed images. Their method achieved an EER of 0.4% on the Replay-Attack database and 6.2% on the CASIA database, highlighting

SYSTEM 2025: 11th Sapienza Yearly Symposium of Technology, Engineering and Mathematics. Rome, June 4-6, 2025
souad.khellat@univ-usto.dz (S. Khellat-Kihel);
ahmed.tibermacine@univ-biskra.dz (A. Tibermacine)
0000-0002-9586-6522 (S. Khellat-Kihel); 0009-0004-4729-7128 (A. Tibermacine)

the potential of handcrafted feature extraction techniques for PAD.

Another prominent approach to PAD involves motion-based methods that focus on liveness detection, as genuine faces exhibit dynamic patterns that are difficult to replicate in spoofing attacks. Chengyan Lin et al.[7] introduced a method based on the rank of sample matrices, observing that spoofed images have low-rank structures due to minimal frame variations, whereas live samples display higher rank values resulting from natural facial movements such as blinking and lip motion[8]. Another notable contribution by Samarth et al.[9] involved the use of Eulerian motion magnification, which enhances subtle facial expressions before extracting features using multiscale LBP. Their approach achieved a Half Total Error Rate (HTER) of 0% and 1.25% on benchmark datasets, demonstrating the robustness of motion-based PAD techniques[10, 11, 12].

With the rise of deep learning, convolutional neural networks (CNNs) have significantly improved PAD performance. Deep learning has also been widely applied in various domains such as computer vision[13, 14], robotic control[15], brain-computer interface (BCI)[16, 17, 18], EEG analysis[19, 20, 21], and sentiment analysis[22, 23, 24, 25, 26, 27], demonstrating its ability to extract meaningful representations from complex data structures. Leveraging these advancements, Jianwei Yang et al. [28] proposed a deep CNN architecture that learns discriminative features for classifying real and fake faces. Their model achieved an HTER of less than 5% on both the CASIA and Replay-Attack databases. More recent works, such as that of Liu et al. [29], introduced a hybrid CNN-RNN model to capture both spatial and temporal dependencies in facial videos, yielding state-of-the-art results across multiple datasets[30, 31, 32]. Similarly, Shao et al. [33]] explored a multi-modal framework that combines RGB, depth, and infrared (IR) imaging to enhance the robustness of PAD systems against varying attack scenarios.attack scenarios[34, 35].

Inspired by the human visual system, researchers have also investigated biological models for feature extraction in PAD. One such model is the Hierarchical Model and X (HMAX), which simulates the ventral stream of the visual cortex [36] and is particularly adept at capturing texture information[37, 38]. The HMAX model is based on a hierarchical structure that processes visual inputs through simple and complex cells, mimicking the way the brain interprets object textures. In this work, we propose to apply the HMAX model to cropped face images for texture-based spoofing detection, leveraging its ability to extract highly discriminative features.

Recent advances in PAD have been supported by the availability of large-scale spoofing datasets, enabling researchers to develop robust models that generalize across diverse attack types. Notable datasets include CASIA-SURF [39], Spoofing in the Wild (SiW)[40], and OULU-NPU[41], which contain extensive variations of spoofing attempts such as print attacks, replay attacks, and 3D mask attacks. However, a persistent challenge in PAD research is domain adaptation, as models trained on one dataset may not generalize well to unseen attack types. To tackle this issue, Yu et al.[42] introduced a domain adaptation framework designed to improve cross-dataset generalization, addressing the problem of dataset bias in PAD systems.

In this work, we aim to contribute to the field of anti-spoofing detection by utilizing a biologically inspired approach based on the HMAX model. Unlike conventional texture-based or CNN-based methods, our approach exploits the hierarchical structure of the human visual cortex to extract highly relevant features for distinguishing between real and spoofed faces. By integrating insights from biological vision systems and leveraging large-scale spoofing datasets, we strive to enhance the generalization capability of PAD systems while maintaining high accuracy against various attack types.

The remainder of this paper is structured as follows. Section 2 presents our proposed methodology, including the implementation details of the HMAX model and its application to spoofing detection. Section 3 describes the dataset and protocols used. Section 4 discusses the results. Finally, Section 5 concludes the paper with key findings and future research directions.

2. Proposed Approach

In this section we highlight the different proposed stages corresponding to the anti-spoofing detection algorithm. At first glance, the HMAX network is a biologically-inspired network which has been conceived to mimic the basic neural architecture of the ventral stream of the visual cortex. Also, the texture extracted with HMAX has a high discrimination performances. The general proposed architecture is depicted in Figure. 2.

2.1. Feature extraction based on hierarchical network

The HMAX model is an hierarchical model for object representation and recognition inspired by the neural architecture of the early stages of the visual cortex in the primates. The general architecture of the HMAX model is represented in Figure. 3. Proceeding to the higher levels of the model, the number and typicality of the extracted features change. Each layer is projected to the next layer by applying template matching or max pooling filters. Proceeding to the higher levels of the model, the number of (X,Y) pixel positions in a layer is reduced. The input to the model is the gray level image. S1 and C1

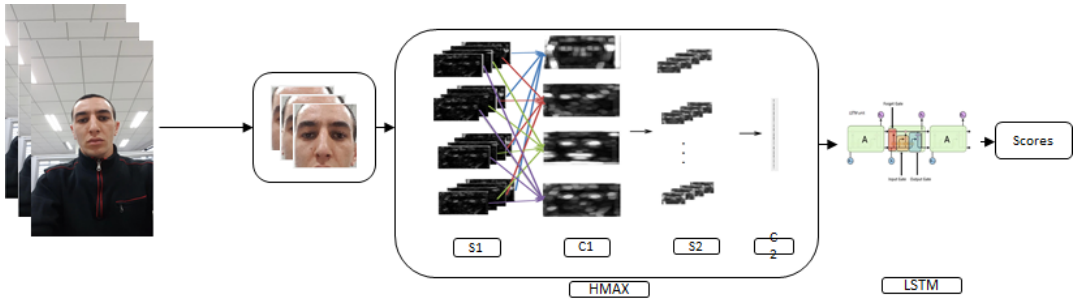

Figure 2: General proposed system for Anti-spoofing detection based on texture information.

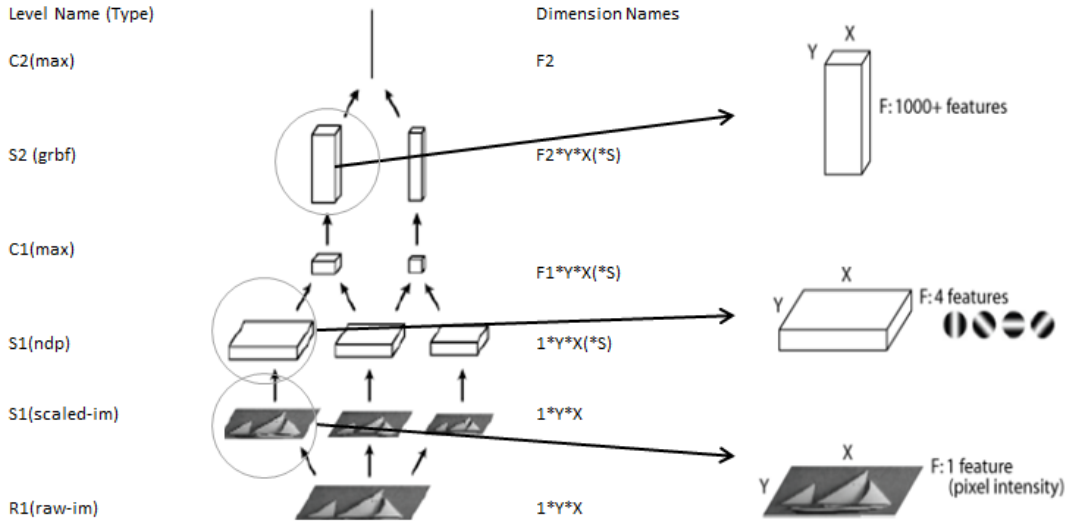

Figure 3: General architecture of the HMAX model.

represent the responses to a bank of Gabor filters tuned to different orientations. S2 and C2 are the responses to more complex filtering stages.

The first layer S1 in the HMAX network consists of a bank of Gabor filters applied to the full resolution image. The response to a particular filter G, of layer S, at the pixel position (X,Y) is given by:

$$R(X,Y) = \left| \frac{\sum X_i G_i}{\sqrt{\sum {X_i}^2}} \right| \quad (1)$$

The size of the Gabor filter is 11x11 and it is formulated as:

$$G(x,y) = exp(\frac{-(x^2\gamma^2 Y^2)}{2\sigma^2})cos(\frac{2\pi}{\lambda}X) \quad (2)$$

Where X=xcosθ-ysinθ and Y=xsinθ+ycosθ . x and y vary between -5 and 5, and θ varies between 0 and π. The parameters ρ (aspect ratio), σ (effective width), and λ (wavelengh) are set to 0.3, 4.5 and 5.6, respectively. For the local invariance (C1) layer, a local maximum is computed for each orientation. They also perform a subsampling by a factor of 5 in both the X and Y directions [10]. In the intermediate feature layer (S2 level), the response for each C1 grid position is computed. Each feature is tuned to a preferred pattern as stimulus. Starting from an image of size 256x256 pixels, the final S2 layer is a vector of dimension 44 x 44 x 4000.The response is obtained using:

$$R(X,P) = exp(\frac{||X-P||^2}{\sigma^2}) \quad (3)$$

The last layer of the architecture is the Global Invariance layer (C2). The maximum response to each intermediate feature over all (X,Y) positions and all scales is

calculated. The result is a characteristics vector that will be used for classification. For the implementation of the HMAX model we use the tool proposed in [43].

2.2. Classification

Firstly, we extract features using HMAX, then we add the Long-Short Term Memory (LSTM) with the extracted features as inputs to capture the temporal dynamic information for differentiation of genuine and fake faces. Each LSTM unit has a memory cell (C_t) and three gates [44]: the input gate (i_t), output gate (o_t) and forget gate (f_t). The memory cell (C_t) could store and output information allowing it to better discover long-range temporal relationships. LSTM mechanism is presented in Figure 4.

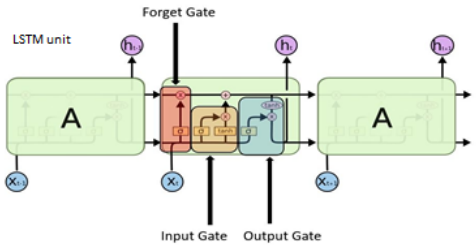

Figure 4: A diagram of an LSTM unit.

The gates serve to modulate the interactions between the memory cell itself and its enviornment. The input gate can allow incoming signal to alter the state of the memory cell or block it. On the other hand, the output gate can allow the state of the memory cell to have an effect on other neurons or prevent it.Finally, the forgot gate can modulate the memory cell's self-recurrent connection, allowing the cell to remember or forget its previous state, as needed. we used X^t and h^t as input and output vector respectively for timestep t, T are input weights matrices, R are recurrent weight matrices and b are bias vectors. Logic sigmoid $\sigma(x) = (\frac{1}{1+e^{-x}})$ and hyperbolic tangent $\phi(x) = (\frac{e^x - e^{-x}}{e^x + e^{-x}})$ are element-wise non-linear activation functions, mapping real values to (0,1) and (-1,1) separately. The dot product and sum of two vectors are denoted with $\bigodot$ and $\bigoplus$ respectively. Given inputs x^t, h^{t-1} and c^{t-1}, the LSTM unit updates for timestep t are:

$g^t = \phi(T_g x^t + R_g h^{t-1} + b_g)$ cell input
$i^t = \sigma(T_i x^t + R_i h^{t-1} + b_i)$ input gate
$f^t = \sigma(T_f x^t + R_f h^{t-1} + b_f)$ forgetgate
$c^t = g^t \bigodot i^t + c^{t-1} \bigodot f^t$ cell state
$o^t = \sigma(T_\circ x^t + R_\circ h^{t-1} + b_\circ)$ output gate
$h^t = \phi(c^t) \bigodot o^t$ cell output

3. Database and protocols

Different databases have been proposed for presentation attack detection (PAD) or anti-spoofing face detection. However, most existing ones are not dedicated in realistic conditions. The publicly available OULU-NPU face presentation attack database [45] consists of 5940 videos corresponding to 55 subjects recorded in three different environments (sessions) using high-resolution frontal cameras of six different smartphones. The high-quality of print and video replay attacks was created by two different printers and two different devices. Figure 5 shows examples corresponding to real accesses and attacks captured with a Samsung Galaxy S6 edge phone.

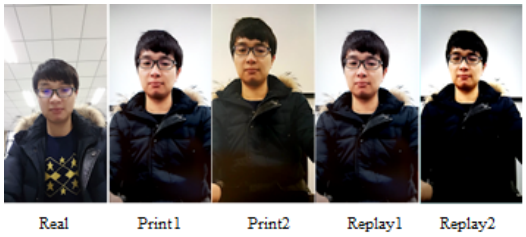

Figure 5: Simple images of real and attack videos captured with a Samsung Galaxy S6 edge phone.

Four protocols have been cited for this database [45] corresponding to the illumination variation effect, different displays and printers, the effect of camera device variation and the last protocol comprises all the previous ones. A challenge has been proposed to evaluate different PAD algorithms under some real-world variations, on the OULU-NPU dataset using its standard evaluation protocols and metrics. We tested the proposed anti-spoofing system on the different protocols proposed for the OULU-NPU dataset [45]. The protocols are defined as follow:

Protocol I: The first protocol is designed to evaluate the PAD methods under different environmental conditions, namely illumination and background scene. The database is recorded in three sessions.

Protocol II: The main goal in this protocol is to test attacks obtained from different resources (printers or displays). The effect of attack variation is evaluated by using unseen print and video-replay attack during the test phase.

Protocol III: The third protocol is dedicated to one of the critical issues in face PAD which is sensor interoperability. In each iteration, real and attack videos

	Session 1	Session 2	Session 3	Total
Training Set				
Client	889	854	0	1743
Imposter	855	893	0	1748
Testing Set				
Client	0	0	3362	3362
Imposter	0	0	5761	5761

Table 1
The number of images in training set and testing set.

obtained from five smartphones are used to run the algorithms. However, the models are constructed using the videos recorded with the remaining one.
Protocol IV: The last protocol englobes all the cases which are seen in the previous three protocols. Generalization of the face anti-spoofing algorithms is evaluated with previously unseen environmental conditions, attacks and input sensors.

The NUAA database [39] is publicaly available photograph imposter database. This database has been collected in three different sessions with an interval of 15 days between the two sessions was conducted and the place and illumination conditions of each session are different. The database is composed of 15 subjects. Some example of images are depicted in Figure 6. The distribution of the train set and the test is represented in Table I.

Figure 6: Sample images of genuine and fake faces. The top row consists of the real cases and the down row represents fake faces.

For performance evaluation we used a recent standardized ISO/IEC 30107-3 metrics proposed in [46] to compare the proposed framework with the various presented results.These metrics consists of Attack Presentation Classification Error Rate (APCER) and Bona Fide Presentation Classification Error Rate (BPCER), Equation (4) and (5) represent the APCER and BPCER metrics respectively.

$$APCER_{PAI} = \frac{1}{N_{PAI}} \sum_{i=1}^{N_{PAI}} (1 - Res_i) \qquad (4)$$

$$BPCER = \frac{\sum_{i=1}^{N_{BF}} Res_i}{N_{BF}} \qquad (5)$$

Where N_{PAI},is the number of the attack presentations for the given PAI, N_{BF} is the total number of the bona fide presentations. Res_i takes the value of 1 if the ith presentation is classified as an attack presentation and 0 is classified as bona fide presentation. These two metrics correspond to False Acceptance Rate (FAR) and False Rejection Rate (FRR) commonly used in biometrics evaluation systems. Also, Average Classification Error Rate (ACER) is proposed in the challenge conducted in [41], which is the average of the APCER and the BPCER.

4. Experimental results

In this section we summarized the different works conducted in the challenge to assume the effectiveness of proposed system. In [41], a detailed algorithm description applied to the OULU-NPU database was carried out. The described algorithms have been proposed within a challenge. Some groups use a Local Phase Quantization (MBLPQ and PML), others rely on a convolutional and deep CNN (Convolutional Neural Networks) models (VSS, NWPU, SZUCVI, Record, CPqD and mixedFASNet). In the Baseline [2], $Massy_HNU$ and HKBU groups proposed some systems based on LBP (Local Binary Pattern) algorithm. Furthermore, Binarized statistical image features was applied by the MFT-FAS group, while GRADIANT and Idiap groups were based on fusion between motion and texture information[1]. In Table 2, Table 3, Table 4 and Table 5, we propose a comparaison between the best obtained results from each category and the results obtained from our proposed approach. The category corresponds to feature extraction method such as the LBP, CNN and LPQ. For classification an LSTM network was used.

As mentioned in Table 2, Table 3, Table 4 and Table 5 the proposed architectures surpass the previous proposed methods mainly because the LSTM is based on video sequences or the features obtained during time hence the emotion in natural cases are analysed. From the obtained

[1]These frameworks abbreviations and metrics are used in [41]

	EER	Display APCER	Print APCER	Overall APCER	Overall BPCER	Overall ACER
LBP	4.4	5	1.3	5	20.8	12.9
CNN	1.3	0.00	0.00	0.00	17.5	8.8
LPQ	0.6	7.5	11.3	11.3	9.2	10.2
HMAX LSTM	0.46	0.57	0.6	0.80	0.09	0.10

Table 2
Comparaison between the results obtained by applying the first protocol.

	EER	Display APCER	Print APCER	Overall APCER	Overall BPCER	Overall ACER
LBP	4.1	15.6	22.5	22.5	6.7	14.6
CNN	1.3	6.4	9.7	9.7	2.5	6.1
LPQ	0.9	11.4	9.4	11.4	3.9	7.6
HMAX LSTM	0.40	4.50	2.30	1.05	1.38	0.69

Table 3
Comparaison between the results obtained by applying the second protocol.

results it is obvious that the LSTM can achieve a low EER. Also, the approach is based on a biological aspect not only by simulating the visual perception by HMAX but also by treating the frames during time using LSTM. The LSTM showed good performances comparing to approaches based on LBP, CNN and LPQ. For the NUAA database, the Fourier spectra analysis method introduced in [47] gives a classification rate of 76.7% when the DoG features proposed in [39] obtain about 10% higher than the first approach. In our case the proposed hierarchical spatio-temporal representation achieve a very high performances around 90% as a rate of classification between the genuine and fake faces.

5. Conclusion

The study of weaknesses of biometric systems against spoofing attacks has been very active field of research in recent years. This focus has led to investigate in Anti-spoofing applications based on faces. However, even if the research is progressing, the development of such applications with a high accuracy may be a challenging task. This mainly due to the distinctiveness difficulty between the fake and genuine images even by human eye, also the high quality of the 3D masks. In this paper, a framework based on biologically spatio-temporal representation has been developed to study the anti-spoofing based on faces. This approach is based on the combination between the HMAX and the LSTM. The experimental evaluation carried out shows a great efficiency comparing to the proposed methods in the litterature.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT, Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

	EER	Display APCER	Print APCER	Overall APCER	Overall BPCER	Overall ACER
LBP	3.9	9.3	11.8	14.2	8.6	11.4
CNN	1.4	1.7	5.3	5.3	7.8	6.5
LPQ	1.1	8.2	15.3	15.7	15.8	15.8
HMAX LSTM	0.90	2.05	4.60	6.00	3.33	1.66

Table 4
Comparaison between the results obtained by applying the third protocol.

	EER	Display APCER	Print APCER	Overall APCER	Overall BPCER	Overall ACER
LBP	4.7	19.20	22.5	29.2	23.3	26.3
CNN	2.8	10	4.2	10	35.8	22.9
LPQ	0.8	59.2	38.3	61.7	13.3	37.5
HMAX LSTM	2.65	3.00	14.80	15.70	7.40	5.00

Table 5
Comparaison between the results obtained by applying the fourth protocol.

References

[1] A. Costa-Pazo, S. Bhattacharjee, E. Vazquez-Fernandez, S. Marcel, The replay-mobile face presentation-attack database, in: International Conference of the Biometrics Special Interest Group (BIOSIG), 2016.

[2] Z. Boulkenafet, J. Komulainen, A. Hadid, Face anti-spoofing based on color texture analysis, in: IEEE International Conference on Image Processing (ICIP), 2015.

[3] S. Russo, S. Ahmed, I. E. Tibermacine, C. Napoli, Enhancing eeg signal reconstruction in cross-domain adaptation using cyclegan, in: 2024 International Conference on Telecommunications and Intelligent Systems (ICTIS), IEEE, 2024, pp. 1–8.

[4] E. Iacobelli, V. Ponzi, S. Russo, C. Napoli, Eye-tracking system with low-end hardware: development and evaluation, Information 14 (2023) 644.

[5] S. Bouchelaghem, I. E. Tibermacine, M. Balsi, M. Moroni, C. Napoli, Cross-domain machine learning approaches using hyperspectral imaging for plastics litter detection, in: 2024 IEEE Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS), IEEE, 2024, pp. 36–40.

[6] K. Patel, A. K. Jain, Secure smartphone unlock: Robust face spoof detection on mobile, 2015.

[7] L. Chengyan, Y. Lu, J. Wu, Y. Xu, Low rank analysis of eye image sequence - a novel basis for face liveness detection, in: Biometric Recognition - 10th Chinese Conference, China, 2015.

[8] S. eddine Boukredine, E. Mehallel, A. Boualleg, O. Baitiche, A. Rabehi, M. Guermoui, A. Douara, I. E. Tibermacine, Enhanced performance of microstrip antenna arrays through concave modifications and cut-corner techniques, ITEGAM-JETIA 11 (2025) 65–71.

[9] S. Bharadwaj, T. I. Dhamecha, M. Vatsa, R. Singh, Computationally efficient face spoofing detection with motion magnification, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW '13), Washington, DC, USA, 2013, pp. 105–110.

[10] A. Tibermacine, N. Djedi, Neat neural networks to control and simulate virtual creature's locomotion, in: 2014 International Conference on Multimedia Computing and Systems (ICMCS), IEEE, 2014, pp. 9–14.

[11] B. Nail, B. Djaidir, I. E. Tibermacine, C. Napoli, N. Haidour, R. Abdelaziz, Gas turbine vibration monitoring based on real data and neuro-fuzzy system, Diagnostyka 25 (2024).

[12] A. Tibermacine, S. M. Amine, An end-to-end trainable capsule network for image-based character recognition and its application to video subtitle recognition., ICTACT Journal on Image & Video Processing 11 (2021).

[13] A. Tibermacine, W. Guettala, I. E. Tibermacine, Efficient one-stage deep learning for text detection in scene images., Electrotehnica, Electronica, Automatica 72 (2024).

[14] A. Tibermacine, M. A. Selmi, An end-to-end trainable capsule network for image-based character recognition and its application to video subtitle recognition, ICTACT Journal on Image & Video Processing 11 (2021).

[15] W. Guettala, et al., Real-time human detection by unmanned aerial vehicles, in: 2022 International Symposium on iNnovative Informatics of Biskra (ISNIB), IEEE, 2022, pp. 1–6.

[16] R. Brociek, G. D. Magistris, F. Cardia, F. Coppa, S. Russo, Contagion prevention of covid-19 by means of touch detection for retail stores, in: CEUR Workshop Proceedings, volume 3092, 2021, p. 89 – 94.

[17] N. Boutarfaia, S. Russo, A. Tibermacine, I. E. Tibermacine, Deep learning for eeg-based motor imagery classification: Towards enhanced human-machine interaction and assistive robotics, CEUR Workshop Proceedings 3695 (2023) 68 – 74.

[18] I. Naidji, A. Tibermacine, W. Guettala, I. E. Tibermacine, et al., Semi-mind controlled robots based on reinforcement learning for indoor application., in: ICYRIME, 2023, pp. 51–59.

[19] A. Tibermacine, D. Akrour, R. Khamar, I. E. Tibermacine, A. Rabehi, Comparative analysis of svm and cnn classifiers for eeg signal classification in response to different auditory stimuli, in: 2024 International Conference on Telecommunications and Intelligent Systems (ICTIS), IEEE, 2024, pp. 1–8.

[20] N. Brandizzi, V. Bianco, G. Castro, S. Russo, A. Wajda, Automatic rgb inference based on facial emotion recognition, in: CEUR Workshop Proceedings, volume 3092, 2021, p. 66 – 74.

[21] A. Tibermacine, I. E. Tibermacine, M. Zouai, A. Rabehi, Eeg classification using contrastive learning and riemannian tangent space representations, in: 2024 International Conference on Telecommunications and Intelligent Systems (ICTIS), IEEE, 2024, pp. 1–7.

[22] N. Brandizzi, S. Russo, G. Galati, C. Napoli, Addressing vehicle sharing through behavioral analysis: A solution to user clustering using recency-frequency-monetary and vehicle relocation based on neighborhood splits, Information (Switzerland) 13 (2022). doi:10.3390/info13110511.

[23] N. Brandizzi, S. Russo, R. Brociek, A. Wajda, First studies to apply the theory of mind theory to green and smart mobility by using gaussian area clustering, in: CEUR Workshop Proceedings, volume 3118,

2021, p. 71 – 76.
[24] S. Russo, I. E. Tibermacine, A. Tibermacine, D. Chebana, A. Nahili, J. Starczewscki, C. Napoli, Analyzing eeg patterns in young adults exposed to different acrophobia levels: a vr study, Frontiers in Human Neuroscience 18 (2024) 1348154.
[25] A. Alfarano, G. De Magistris, L. Mongelli, S. Russo, J. Starczewski, C. Napoli, A novel convmixer transformer based architecture for violent behavior detection, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 14126 LNAI, 2023, p. 3 – 16. doi:10.1007/978-3-031-42508-0_1.
[26] I. E. Tibermacine, A. Tibermacine, W. Guettala, C. Napoli, S. Russo, Enhancing sentiment analysis on seed-iv dataset with vision transformers: A comparative study, in: Proceedings of the 2023 11th international conference on information technology: IoT and smart city, 2023, pp. 238–246.
[27] V. Marcotrigiano, G. D. Stingi, S. Fregnan, P. Magarelli, P. Pasquale, S. Russo, G. B. Orsi, M. T. Montagna, C. Napoli, C. Napoli, An integrated control plan in primary schools: Results of a field investigation on nutritional and hygienic features in the apulia region (southern italy), Nutrients 13 (2021). doi:10.3390/nu13093006.
[28] J. Yang, Z. Lei, S. Z. Li, Learn convolutional neural network for face anti-spoofing, CoRR (2014). ArXiv preprint.
[29] Y. Liu, et al., Hybrid cnn-rnn for face anti-spoofing, IEEE Transactions on Biometrics, Behavior, and Identity Science (2021).
[30] B. Nail, M. A. Atoussi, S. Saadi, I. E. Tibermacine, C. Napoli, Real-time synchronisation of multiple fractional-order chaotic systems: an application study in secure communication, Fractal and Fractional 8 (2024) 104.
[31] G. Capizzi, C. Napoli, S. Russo, M. Woźniak, Lessening stress and anxiety-related behaviors by means of ai-driven drones for aromatherapy, in: CEUR Workshop Proceedings, volume 2594, 2020, p. 7 – 12.
[32] A. Tibermacine, N. Djedi, Neat neural networks to control and simulate virtual creature's locomotion, in: 2014 International Conference on Multimedia Computing and Systems (ICMCS), IEEE, 2014, pp. 9–14.
[33] S. Shao, et al., Multi-modal face anti-spoofing using rgb, depth, and infrared data, IEEE Transactions on Pattern Analysis and Machine Intelligence (2022).
[34] A. Tibermacine, N. Djedi, Gene regulatory network to control and simulate virtual creature's locomotion, International Journal of Artificial Intelligence & Applications (2015).
[35] W. Guettala, A. Sayah, L. Kahloul, A. Tibermacine, Real time human detection by unmanned aerial vehicles, in: 2022 International Symposium on iNnovative Informatics of Biskra (ISNIB), IEEE, 2022, pp. 1–6.
[36] M. Riesenhuber, T. Poggio, Hierarchical models of object recognition in cortex, Nature Neuroscience 2 (1999) 169, 321–354.
[37] C. Napoli, V. Ponzi, A. Puglisi, S. Russo, I. Tibermacine, et al., Exploiting robots as healthcare resources for epidemics management and support caregivers, in: CEUR Workshop Proceedings, volume 3686, CEUR-WS, 2024, pp. 1–10.
[38] B. Ladjal, I. E. Tibermacine, M. Bechouat, M. Sedraoui, C. Napoli, A. Rabehi, D. Lalmi, Hybrid models for direct normal irradiance forecasting: A case study of ghardaia zone (algeria), Natural Hazards 120 (2024) 14703–14725.
[39] X. Tan, Y. Li, J. Liu, L. Jiang, Face liveness detection from a single image with sparse low rank bilinear discriminative model, in: 11th European Conference on Computer Vision (ECCV'10), Crete, Greece, 2010.
[40] SiW Dataset, Spoofing in the wild, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
[41] Oulu-npu database, https://sites.google.com/site/oulunpudatabase/welcome, ????
[42] X. Yu, et al., Domain adaptation for face antispoofing, IEEE Transactions on Neural Networks and Learning Systems (2023).
[43] Hmax toolbox, http://maxlab.neuro.georgetown.edu/hmax.html, ????
[44] X. Tu, et al., Enhance the motion cues for face antispoofing using cnn-lstm architecture, CoRR (2019). ArXiv:1901.05635.
[45] Z. Boulkenafet, J. Komulainen, L. Li, X. Feng, A. Hadid, Oulu-npu: A mobile face presentation attack database with real-world variations, in: IEEE International Conference on Automatic Face and Gesture Recognition, 2017.
[46] Z. Boulkenafet, et al., A competition on generalized software-based face presentation attack detection in mobile scenarios, in: IEEE International Joint Conference on Biometrics (IJCB), 2017, pp. 688–696.
[47] J. Li, Y. Wang, T. Tan, A. K. Jain, Live face detection based on the analysis of fourier spectra, in: SPIE Conference, 2004, pp. 296–303.