

# XAI for supporting Gait Analysis of Patient with Schizophrenia

Francesca Rita<sup>1,\*</sup>, De Santis Laura<sup>1</sup>, Iannotti Federica<sup>2</sup> and Iasevoli Felice<sup>2</sup>

<sup>1</sup>Computer Science Department, University of Salerno, Fisciano, Italy

<sup>2</sup>Department of Neuroscience, Reproductive Science and Odontostomatology, University Federico II, Naples, Italy

## Abstract

The early identification and continuous observation of individuals diagnosed with schizophrenia can significantly enhance their quality of life. This document serves as an initial overview of a specific task within an ongoing project titled SPECTRA - Supporting Schizophrenia Patients' Care with Artificial Intelligence. The project's objective is to aggregate patient data from diverse sources, encompassing speech, emotional responses, and locomotion patterns, with a particular focus on gait analysis in this study. We utilize advanced deep learning algorithms to identify distinctive movement patterns and apply explainable AI techniques to elucidate the decision-making processes of the models. Our proposed methodology unfolds in two main stages: the gathering and pre-processing of data, followed by the classification of gait patterns with corresponding explanations. To classify gait, we employ spatio-temporal transformer models, and to elucidate these classifications, we generate visual explanations using SHAP (SHapley Additive exPlanations) images.

## Keywords

Schizophrenia, Gait Analysis, Deep Learning, explainable Artificial Intelligence, Spatio-temporal transformers.

## 1. Introduction

Schizophrenia (SZ) is a complex mental disorder that affects approximately 1 in 300 individuals (0.32%) globally, a statistic reported by the World Health Organization [1] in 2024. This prevalence rate escalates to 1 in 222 (0.45%) within the adult demographic.

Patients with SZ often experience a substantial decline in their Quality of Life (QoL), as the disorder is associated with significant social and occupational challenges [2]. Early diagnosis can lead to highly effective outcomes through pharmacological interventions, enhancing the QoL for many patients. Diagnosing SZ necessitates clinicians to engage in extensive interviews and observational sessions with patients, which is notably time-consuming. The repetitive nature of these assessments can induce a learning effect on patient responses. Furthermore, continuous patient monitoring is essential to mitigate the risk of sudden relapses, although regular medical evaluations are impractical due to their associated time and costs. The scarcity of mental healthcare providers is a pervasive issue worldwide, more so in developing nations,

---

*INI-DH 2024: Workshop on Innovative Interfaces in Digital Healthcare, in conjunction with International Conference on Advanced Visual Interfaces 2024 (AVI 2024), June 3-7, 2024, Arenzano, Genoa, Italy (2024)*

\*Corresponding author.

✉ francesca@unisa.it (F. Rita); l.desantis9@studenti.unisa.it (D. S. Laura); fede.iannotta.fi@gmail.com (I. Federica); felice.iasevoli@unina.it (I. Felice)

ORCID 0000-0002-6929-0056 (F. Rita); 0000-0002-7051-5013 (I. Felice)

 © 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

underscoring the urgent need for innovative tools to assist clinicians in the screening and monitoring of SZ patients. It has been documented that up to 80% of individuals with SZ exhibit Genuine Motor Abnormalities (GMA), which are also prevalent among ultra-high risk (UHR) groups and unaffected first-degree relatives possessing a genetic predisposition for SZ [3]. Consequently, gait analysis emerges as a valuable source of data for aiding the diagnosis and continuous monitoring of SZ. Although numerous gait analysis methodologies incorporating body sensors have been proposed, they can be obtrusive. Recent studies [4] [5] have successfully implemented non-intrusive techniques using one or multiple digital cameras.

This paper introduces ongoing research within the SPECTRA project, which focuses on aiding the diagnosis of Treatment-Resistant Schizophrenia (TRS) patients, who do not respond to anti-psychotic medications. The project aims to integrate various data sources to provide a comprehensive patient profile. Specifically, this paper discusses our proposed methodology for diagnosing and monitoring SZ through gait analysis. We further elucidate the deep learning-based classification process with visual explanations to foster clinician trust in our techniques.

The structure of the paper is as follows: Section 2 provides an overview of the relevant literature and background information. Section 3 delineates the objectives of the SPECTRA Project. Section 4 details the proposed approach, and Section 5 concludes the paper with a discussion on the implications, final observations, and prospective future research directions..

## 2. Background

In this section, we delve into the fundamental aspects of gait analysis and examine its relevance in the detection and continuous monitoring of Schizophrenia (SZ).

Gait analysis emerges as a critical instrument for decoding human locomotion patterns and their implications on a spectrum of health disorders. The interplay between gait characteristics and mental health conditions, notably including depression [6] [7], Alzheimer’s disease [8], and SZ [9], has garnered increasing attention in recent research. This method offers a non-intrusive and cost-effective avenue for mental health assessment, enabling the identification of nuanced deviations in walking patterns potentially indicative of mental health symptoms. It may be also used for detecting human cooperation behavior through video surveillance, which further illustrates the diverse applications of computer vision in healthcare and psychological studies [10]. Traditionally, gait analysis systems rely on data gathered by specialized equipment, such as body-worn sensors [11] or 3D sensing technologies [12]. These approaches, however, often require controlled environments, which can be costly and intrusive. An alternative and more accessible methodology was introduced by Miao et al. [4], utilizing a digital camera to capture footage of participants’ gait. To facilitate the computational analysis of human locomotion in these recordings, OpenPose [13], an open-source toolkit developed by Carnegie Mellon University, was employed to detect and monitor bodily movements. In a significant study by Martin et al. [3], gait patterns were analyzed through comprehensive movement assessments of 20 patients with SZ and 20 control participants. Utilizing motion capture technology, the research team evaluated mental health conditions using established scales such as the Positive And Negative Syndrome Scale (PANSS), Brief Psychiatric Rating Scale (BPRS), and Neurological Soft Signs (NSS), successfully identifying 16 quantifiable movement markers associated with SZ.

To augment understanding and interpretability of model decisions, explainable AI (XAI) techniques have been applied. These methods, particularly post-hoc approaches like the Interpretable Model-agnostic Explanations (LIME)[14] and SHapley Additive exPlanations (SHAP)[15], enhance model transparency by elucidating the impact of specific input features on the predictions. A notable application of these methodologies is illustrated in the work by Mishra et al. [16], where a multi-layer perceptron was adopted to classify people wearing two types of Knee Ankle Foot Orthosis. LIME emphasized the features that are more relevant for the model.

### 3. The SPECTRA project

Within the spectrum of schizophrenia (SZ), there are patients whose significant symptoms persist despite receiving adequate antipsychotic treatment. These patients are classified as having Treatment Resistant Schizophrenia (TRS) [17]. TRS is a severe condition that affects nearly 30% of individuals diagnosed with schizophrenia. Unfortunately, TRS is often diagnosed late in the course of the disorder, which hinders the timely switch to more effective treatments (such as clozapine) and non-pharmacological therapeutic approaches. This delay results in considerable individual suffering and substantial economic costs for communities. Early and accurate diagnosis of TRS is crucial. It allows clinicians to recommend more suitable pharmacological and non-pharmacological therapies, potentially improving the Quality of Life (QoL) for TRS patients and conserving valuable economic resources. Researchers have explored the use of sensors for mental disease detection and patient monitoring [18] [19] [20].

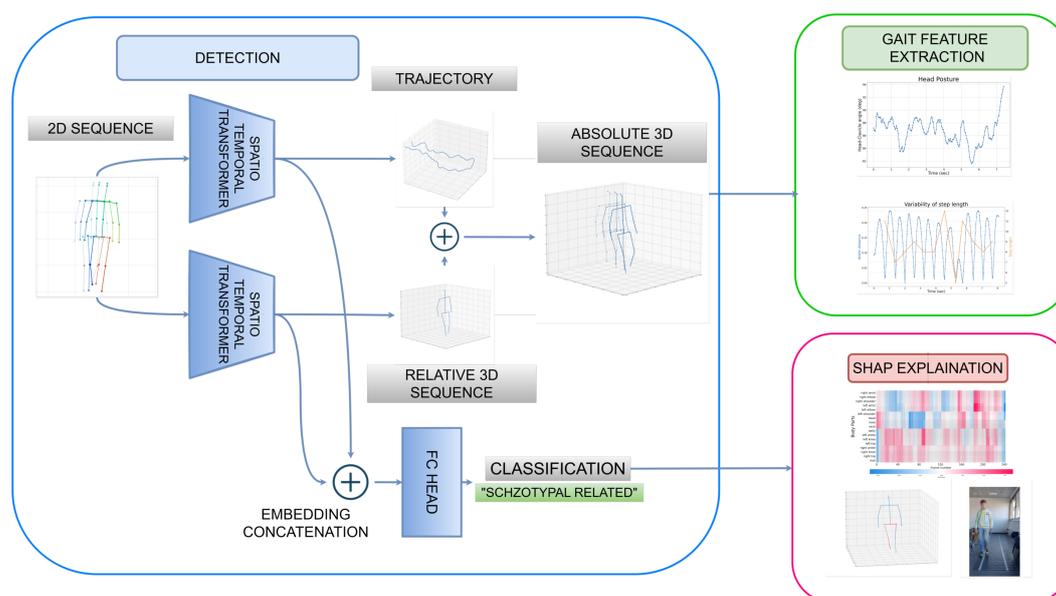
The SPECTRA Project aims to create a Decision Support System (DSS) using Artificial Intelligence (AI) techniques and cutting-edge IoT technologies. By combining standard clinical screening procedures with ICT-based assessment techniques driven by Machine Learning algorithms, SPECTRA seeks to enable early TRS diagnosis. The project involves a field study with real patients from the Unit for Treatment-Resistant Psychosis at the University “Federico II” of Naples. Eligible patients are categorized as either TRS or non-TRS, and the SPECTRA team collects historical clinical data (including Magnetic Resonance Images, questionnaire scores, demographic information, and geographical data) alongside real-time data from IoT sensors (such as ECG, temperature, EEG, and audio/video signals) during patient screenings. Conventional data and information obtained through ICT technologies are utilized to train Machine Learning and Deep Learning models for the early detection of Treatment Resistant Schizophrenia (TRS) patients. However, a significant challenge in adopting AI solutions for healthcare support lies in clinicians’ lack of trust in the black-box nature of these systems. To address this issue, another project aim is to develop interaction models that assist clinicians during TRS diagnosis by leveraging eXplainable Artificial Intelligence (XAI) methods [21]. These XAI techniques should highlight how AI black-box models generate predictions, potentially enhancing clinicians’ confidence in these novel technologies. Additionally, the project will create a dataset distinguishing TRS from non-TRS cases, which will be valuable for scientific research and experimentation in Machine Learning-based TRS diagnosis for individuals with schizophrenia

## 4. The proposed approach

In this study, our objective is to explore the gait characteristics of individuals diagnosed with schizophrenia by conducting a visual analysis of their walking patterns. The research is structured into two main stages: (i) data acquisition and preprocessing, and (ii) classification along with explanation. The sequential flow of this process is illustrated in Figure 1.

The initial phase involves recording two distinct groups: individuals diagnosed with schizophrenia (SZ) and a matched control group, utilizing comprehensive setup of multiple calibrated mobile cameras. Participants in both groups provide informed consent, with the assurance that their data will be anonymized and included in a publicly accessible dataset. We apply a machine learning algorithm designed for body detection to derive skeletal joint data from the captured video footage, resulting in 2D coordinates of the joints for each video frame. The dataset creation employs a setup of three cameras to ensure a comprehensive data collection, encompassing a diverse range of data samples. However, for practical application in real-world scenarios, we propose the use of a single-camera setup for gait analysis to facilitate ease of implementation in various settings.

Inspired by the spatial-temporal transformers model approaches in literature for motion analysis and recognition as VideoPose[22] and MotionBert[23], we leveraged a dual transformer network to analyze gait patterns, generate 3D skeletal representations, and classify participants. To explain the classification result, we employed the SHAP algorithm to investigate the decision-making process of our deep classification model.



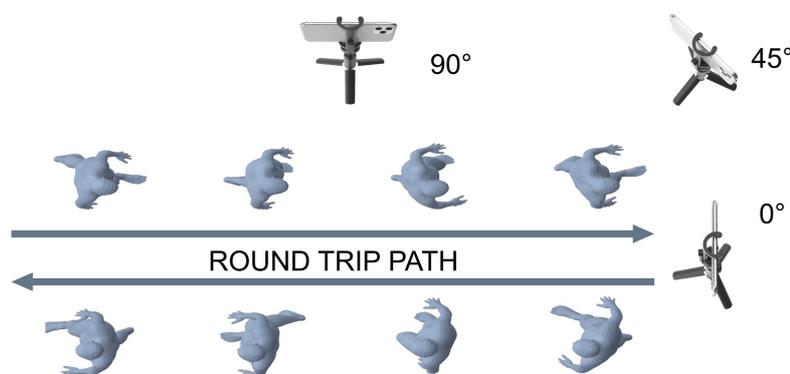
**Figure 1:** High level schema of proposed network

The network processes the 2D skeleton sequence as input, yielding both a classification of the subjects and 3D reconstructions of their corresponding skeletal movements. The 3D re-

construction facilitate the extraction of critical movement features, such as step count and walking speed, which are relevant for the clinician. The analysis provides these movement features in conjunction with SHAP visualizations of the key frames and joints with respect to the classification, offering domain experts enhanced insights to clarify the classification outcomes.

#### 4.1. Data collection and preprocessing

Data are collected using a setup comprising three calibrated mobile cameras, as shown in Figure 2. We select a quiet and bright environment. Before starting the recording participants freely walk in the environment. Their task consists in walking forward and backward for two minutes on a carpet long 5 meters and large one meter. The walk is recorded by three cameras situated as follows: a frontal camera, a central camera and the other at 45°.



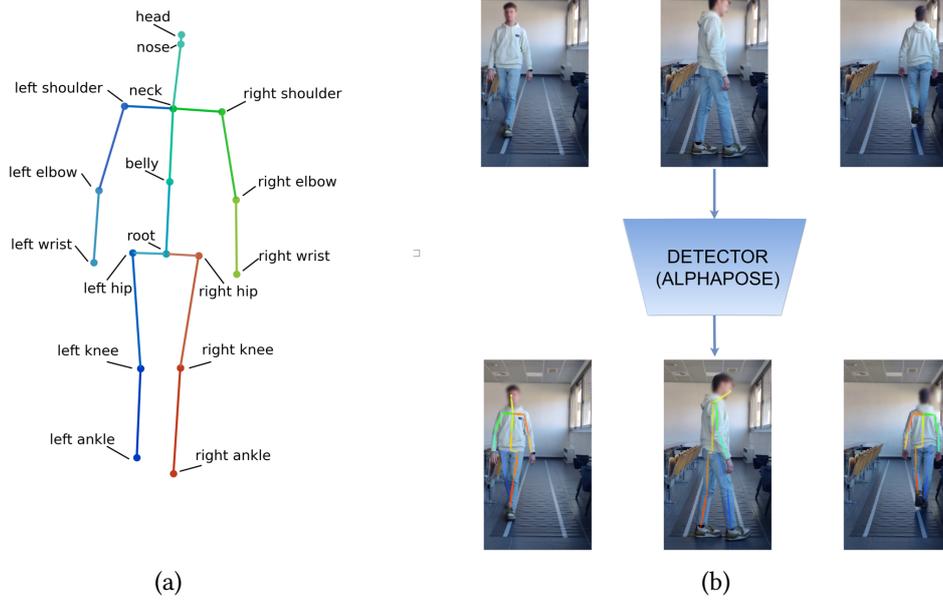
**Figure 2:** Data collection set-up.

The recorded video is then processed as follows:

- body detection, performed by using a machine learning algorithm, specifically Alphapose [24]. The format of the human body joints used coincides with that of the Human3.6 dataset [25] and is shown in Figure 3a.
- skeletal joint coordinates extraction, performed for each frame in the videos, providing 2D relative joints sequence information (Figure 3b).

#### 4.2. Gait Analysis

The network provides a classification of the patients. To enhance the performance of our algorithm, we plan to conduct fine-tuning procedures on the Transformer network using our collected dataset, this may be done by adding a linear layer or a multilayer perceptron (MLP). The transformer also reconstructs the 3D representation of the patient gaits from the recorded videos. This reconstruction resulted in a sequence of skeletal representations capturing the motion dynamics of each individual. This data will be used by the clinician for further analysis. It should be considered that the networks have been pre-trained on very large dataset



**Figure 3:** a) Posture Ankle estimation points; b) Skeleton sequence detection example.

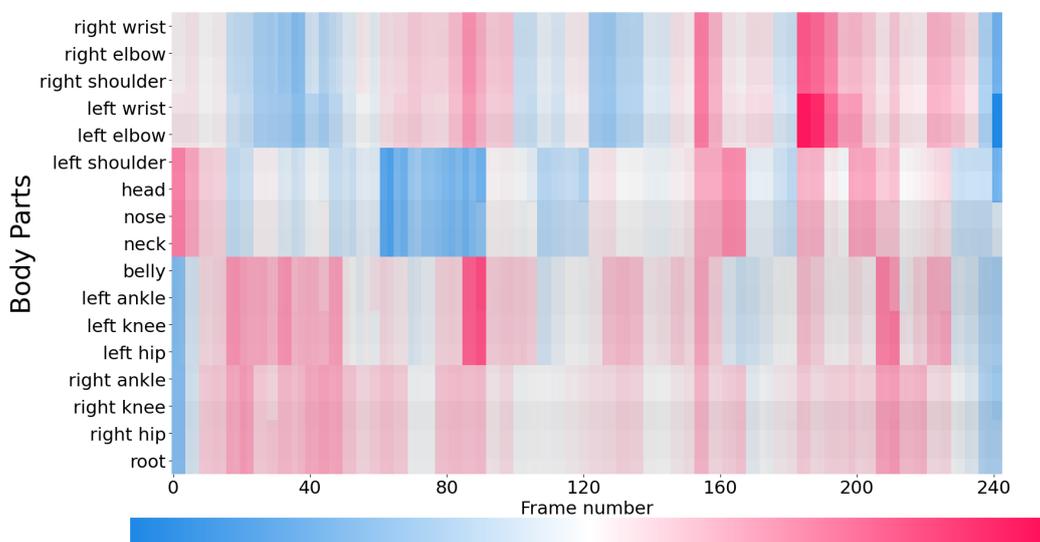
and exhibiting a very robust embedded representation of human body movement dynamic. Due to the restricted availability of video recordings and annotations—a common challenge in scenarios characterized by a scarcity of public datasets—our approach leverages one-shot learning techniques and supervised contrastive learning. These methods have previously demonstrated promise in tasks constrained by such difficulties, as evidenced by Sabater et al. [26] and Khosla et al. [27]. These studies validate the potential of employing these advanced learning strategies in environments with limited data.

### 4.3. The XAI interface

The SHapley Additive exPlanations (SHAP) algorithm plays a pivotal role in demystifying the decision-making process of our deep learning classification model. This approach is visually represented by the matrix in Figure 4, where the SHAP explanation on a sample video is shown. In particular, the columns represent frame numbers and the rows represent the joints under consideration, such as the right wrist. The matrix cells are colored according to the SHAP values which represent the influence of various body parts over time in a video classification model. In this analysis the colored cells have the following meaning:

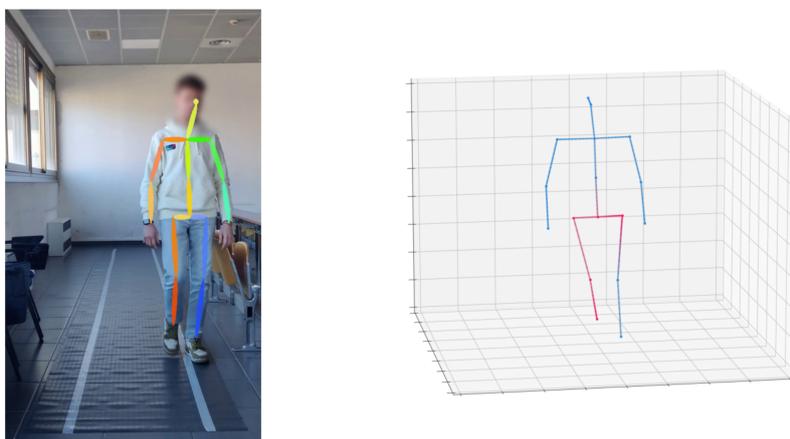
- **Neutral areas**, depicted in white on the heatmap, indicate points on the body that are not significantly relevant to the model’s classification decision throughout the video frames.
- **Red regions**, representing body movements that have positively contributed to the model’s decision-making process, strongly supporting the classification according to the model. This suggests that they are characteristic of the predicted class.

- **Blue regions**, indicating body movements that potentially misled the model’s classification, negatively impacting the decision, implying that the presence of these specific motions might be atypical for the predicted class or more common in other classes.



**Figure 4:** The SHAP video analysis at frame level

This SHAP based visual representation allows for an intuitive understanding of which body parts and their movements over time are considered decisive or inconsequential by the model, thereby offering insights into the model’s behavior and its reliance on specific features for making classifications. These emphasized frames serve as a focal point for clinicians, enabling them to delve deeper into the analysis and gain a clearer comprehension of the classifier’s results. Such an approach facilitates a more nuanced understanding of the diagnostic outcomes.



**Figure 5:** Relevant frames for the model classification.

Figure 5 illustrates this concept further by displaying the selected frames alongside the skeletal representations of the patients. This visual correlation allows clinicians to directly observe the specific gait patterns contributing to the classification, enhancing their ability to interpret and utilize the findings in a clinical setting.

#### 4.4. Further Patient's feature for the domain expert

As an additional value, the network produces the 3D reconstructions of the skeletal sequences by using the entire sequence of skeletal representations as 2D points. This representation is inputted to the gait feature extraction. These features may be adopted by the clinicians for further analysis. In particular, based on the results collected by Martin et al.[3], we selected the following features related to movement analysis, essential for understanding movement patterns, assessing gait, and identifying abnormalities or changes.

- **Head posture.** This refers to the angle between the head and the clavicle. It helps assess alignment and balance during movement.
- **Number of steps.** Simply put, it's the count of steps taken while walking. It's a fundamental measure of mobility.
- **Walking speed.** This parameter indicates how fast an individual walks. It's usually measured in meters per second (m/s) or kilometers per hour (km/h).
- **Arm sway.** Arm sway assesses the variability of the angle of the armpit or the variations in the elbow angle during movement. It provides insights into arm movement coordination.
- **Presence/absence of facial movements.** This parameter observes whether there are any noticeable facial expressions or movements during walking. It can be relevant for assessing overall coordination and comfort.
- **Step length.** Step length measures the distance covered by a single step. It's typically calculated from heel strike to the next heel strike of the same foot.
- **Variability of step length.** This parameter evaluates how consistent or variable step lengths are during walking. It can indicate gait stability and symmetry.

The analysis of these measures may provide further support to the clinician in the detection of anomalous patterns.

#### 4.5. Preliminary assessment

The SPECTRA project lasts 2 years, starting from November 2023, and patients' data are going to be collected during the first year. Thus, to set up the appropriate pipeline while waiting for the patients' data we start to use the data of the PsyMo dataset [28], consisting of walking sequences provided by recording 312 people accompanied by 6 psychological questionnaires they filled in.

## 5. Conclusion

In this paper, we described the methodology we plan to adopt in the ongoing SPECTRA project for performing gait analysis to detect people suffering from schizophrenia. We performed the

setup of the models and described the techniques we adopted.

It is important to highlight that we propose to use a single-camera setup for gait analysis. Results may be limited when compared to setups utilizing multiple cameras. However, this approach may be needed because it may be easier to take the patient's gait in a setting such as the patient's home or the rehabilitation center. This may be useful for supporting patient monitoring and the prediction of relapse. In the future, we plan to compare the detection performance of the two approaches (with multiple and single camera). In addition, the reliance on video footage for data acquisition may introduce challenges related to data quality, consistency, and privacy concerns, potentially affecting the robustness and generalizability of the findings. Concerning privacy issues, the research will follow the GDPR on patient data. In particular, the patient's data will be anonymized. Figure 3 shows an example of a blurred patient's face.

We also plan to experiment with the proposed deep learning pipeline on existing datasets and the data of patients suffering from SZ, on control samples, and of TRS and non-TRS patient

## Acknowledgments

This project has been financially supported by the European Union NEXTGenerationEU project and by the Italian Ministry of the University and Research MUR, a Research Projects of Significant National Interest (PRIN) 2022 PNRR, project n. D53D23017290001 entitled "Supporting schizophrenia PatiEnts Care wiTh aRtificial intelligence (SPECTRA)".

## References

- [1] World Health Organization., Schizophrenia., 2022. URL: <https://www.who.int/news-room/fact-sheets/detail/schizophrenia>.
- [2] J. Bobes, M. P. Garcia-Portilla, M. T. Bascaran, P. A. Saiz, M. Bouzoño, Quality of life in schizophrenic patients, *Dialogues in clinical neuroscience* 9 (2007) 215–226.
- [3] L. Martin, K. Stein, K. Kubera, N. F. Troje, T. Fuchs, Movement markers of schizophrenia: a detailed analysis of patients' gait patterns, *European Archives of Psychiatry and Clinical Neuroscience* 272 (2022) 1347–1364.
- [4] B. Miao, X. Liu, T. Zhu, Automatic mental health identification method based on natural gait pattern, *Psych J* 10 (2021) 453–464.
- [5] M. Zago, M. Luzzago, T. Marangoni, M. De Cecco, M. Tarabini, M. Galli, 3D Tracking of Human Motion Using Visual Skeletonization and Stereoscopic Vision, *Front Bioeng Biotechnol* 8 (2020) 181.
- [6] M. Belvederi Murri, F. Triolo, A. Coni, C. Tacconi, E. Nerozzi, A. Escelsior, M. Respino, F. Neviani, M. Bertolotti, K. Bertakis, L. Chiari, S. Zanetidou, M. Amore, Instrumental assessment of balance and gait in depression: A systematic review, *Psychiatry Res* 284 (2020) 112687.

- [7] T. Wang, C. Li, C. Wu, C. Zhao, J. Sun, H. Peng, X. Hu, B. Hu, A gait assessment framework for depression detection using kinect sensors, *IEEE Sensors Journal* 21 (2021) 3260–3270. doi:10.1109/JSEN.2020.3022374.
- [8] V. R. Varma, R. Ghosal, I. Hillel, D. Volfson, J. Weiss, J. Urbanek, J. M. Hausdorff, V. Zipunikov, A. Watts, Continuous gait monitoring discriminates community-dwelling mild Alzheimer’s disease from cognitively normal controls, *Alzheimers Dement (N Y)* 7 (2021) e12131.
- [9] M. Tonna, V. Lucarini, J. Lucchese, V. Presta, F. Paraboschi, F. Marsella, B. D. Daniel, M. Vitale, C. Marchesi, G. Gobbi, Posture, gait and self-disorders: An empirical study in individuals with schizophrenia, *Early Interv Psychiatry* 17 (2023) 447–461.
- [10] D. Freire-Obregón, M. Castrillón-Santana, P. Barra, C. Bisogni, M. Nappi, An attention recurrent model for human cooperation detection, *Computer Vision and Image Understanding* 197-198 (2020) 102991. doi:10.1016/j.cviu.2020.102991.
- [11] A. Ahmadi, F. Destelle, L. Unzueta, D. S. Monaghan, M. Teresa Linaza, K. Moran, N. E. O’Connor, 3d human gait reconstruction and monitoring using body-worn inertial sensors and kinematic modeling, *IEEE Sensors Journal* 16 (2016) 8823–8831. doi:10.1109/JSEN.2016.2593011.
- [12] V. Bijalwan, V. B. Semwal, T. K. Mandal, Fusion of multi-sensor-based biomechanical gait analysis using vision and wearable sensor, *IEEE Sensors Journal* 21 (2021) 14213–14220. doi:10.1109/JSEN.2021.3066473.
- [13] Z. Cao, T. Simon, S.-E. Wei, Y. Sheikh, Realtime multi-person 2d pose estimation using part affinity fields, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [14] M. T. Ribeiro, S. Singh, C. Guestrin, " why should i trust you?" explaining the predictions of any classifier, in: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.
- [15] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), *Advances in Neural Information Processing Systems*, volume 30, Curran Associates, Inc., 2017. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf).
- [16] A. Mishra, A. Shetkar, G. M. Bapat, R. Ojha, T. T. Verlekar, Xai-based gait analysis of patients walking with knee-ankle-foot orthosis using video cameras, *arXiv preprint arXiv:2402.16175* (2024).
- [17] F. Iasevoli, C. Avagliano, B. Altavilla, A. Barone, M. Ciccarelli, L. D’Ambrosio, D. N. Francesco, E. Razzino, M. Fornaro, A. de Bartolomeis, Evaluation of a few discrete clinical markers may predict categorization of actively symptomatic non-acute schizophrenia patients as treatment resistant or responders: A study by roc curve analysis and multivariate analyses, *Psychiatry research* 269 (2018) 481–493.
- [18] I. Amaro, R. Francese, G. Tortora, C. Tucci, L. D’Errico, M. Staffa, Supporting schizophrenia patients’ care with robotics and artificial intelligence, in: Q. Gao, J. Zhou, V. G. Duffy, M. Antona, C. Stephanidis (Eds.), *HCI International 2023 - Late Breaking Papers - 25th International Conference on Human-Computer Interaction, HCII 2023, Copenhagen, Denmark, July 23-28, 2023, Proceedings, Part II*, volume 14055 of *Lecture Notes in Computer*

- Science*, Springer, 2023, pp. 482–495.
- [19] A. A. Cantone, M. Esposito, F. P. Perillo, M. Romano, M. Sebillio, G. Vitiello, Enhancing elderly health monitoring: Achieving autonomous and secure living through the integration of artificial intelligence, autonomous robots, and sensors, *Electronics* 12 (2023). URL: <https://www.mdpi.com/2079-9292/12/18/3918>. doi:10.3390/electronics12183918.
  - [20] R. Francese, M. Risi, G. Tortora, F. D. Salle, Thea: empowering the therapeutic alliance of children with ASD by multimedia interaction, *Multim. Tools Appl.* 80 (2021) 34875–34907.
  - [21] A. M. Antoniadi, Y. Du, Y. Guendouz, L. Wei, C. Mazo, B. A. Becker, C. Mooney, Current challenges and future opportunities for xai in machine learning-based clinical decision support systems: a systematic review, *Applied Sciences* 11 (2021) 5088.
  - [22] D. Pavlo, C. Feichtenhofer, D. Grangier, M. Auli, 3d human pose estimation in video with temporal convolutions and semi-supervised training, in: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 7753–7762.
  - [23] W. Zhu, X. Ma, Z. Liu, L. Liu, W. Wu, Y. Wang, Motionbert: A unified perspective on learning human motion representations, in: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 15085–15099.
  - [24] H.-S. Fang, J. Li, H. Tang, C. Xu, H. Zhu, Y. Xiu, Y.-L. Li, C. Lu, Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time, *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022).
  - [25] C. Ionescu, D. Papava, V. Olaru, C. Sminchisescu, Human3.6m: Large scale datasets and predictive methods for 3d human sensing in natural environments, *IEEE transactions on pattern analysis and machine intelligence* 36 (2013) 1325–1339.
  - [26] A. Sabater, L. Santos, J. Santos-Victor, A. Bernardino, L. Montesano, A. C. Murillo, One-shot action recognition in challenging therapy scenarios, in: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 2777–2785.
  - [27] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, D. Krishnan, Supervised contrastive learning, *Advances in neural information processing systems* 33 (2020) 18661–18673.
  - [28] A. Cosma, E. Radoi, Psymo: A dataset for estimating self-reported psychological traits from gait, in: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024, pp. 4603–4613.