Explainable Artificial Intelligence in Movement Sciences

Zheng, Xiaoping

DOI:
10.33612/diss.958165448

Publication date:
2024

Chapter 7

General Discussion

Main Findings

The structure of this thesis and the findings of each chapter are briefly summarized in Fig. 1. The analysis begins with focusing on the gait patterns of healthy older adults. **Chapter 2** showed that deep learning (DL) outperformed conventional machine learning (CML) in classifying adults and older adults. Building on these results, **Chapter 3** utilized Explainable Artificial Intelligence (XAI) to investigate what DL models had learned from accelerometer signal data. This chapter found that DL models primarily used data surrounding heel contact for classification, indicating potential differences in acceleration and deceleration patterns between adults and older adults during walking. These gait analysis insights and the XAI methodologies could further aid in analyzing the movement of patients with chronic low back pain (CLBP). In **Chapter 4**, it was aimed to use CML and XAI to examine how human assumed central sensitization (HACS) was associated with the gait of patients with CLBP. The findings showed different gait patterns in patients with low or high levels of HACS (CLBP- and CLBP+), possibly associated with different motor control strategies characterized as “loose” or “tight”. Going beyond gait analysis, **Chapter 5** used unsupervised CML (with ante-hoc explainability) to explore the relationships between physical activity intensity (PAI), CLBP, and HACS. The findings indicate distinct PAI patterns, suggesting that patients in the CLBP+ group may exhibit an endurance pain response pattern. Given that severe levels of HACS in Chapters 4 and 5 were determined by the central sensitization inventory (CSI) questionnaire with a predetermined cut-off value of 40, **Chapter 6** used unsupervised CML (with ante-hoc explainability) to define a more accurate cut-off value for the population with CLBP. This analysis found two distinct subgroups within CLBP, possibly correlated with low and high levels of HACS, leading to the establishment of a new cutoff value of 35.

Gait Performance in Healthy Aging

The rapidly increasing aging population raises concerns about maintaining the quality of life in older adults, especially their mobility [1]. Aging is a continuous process that often involves the loss of muscle mass, reduced bone density, and declining nerve function [2]. These age-related alterations can lead to changes in gait patterns, such as lower walking speed, shorter stride/step length, increased gait variability, and increased gait instability compared to young controls [3, 4]. Given the limited availability of medical resources, automatic gait analysis plays a crucial role in monitoring the mobility of older adults [5] and contributed to a better understanding of aging.

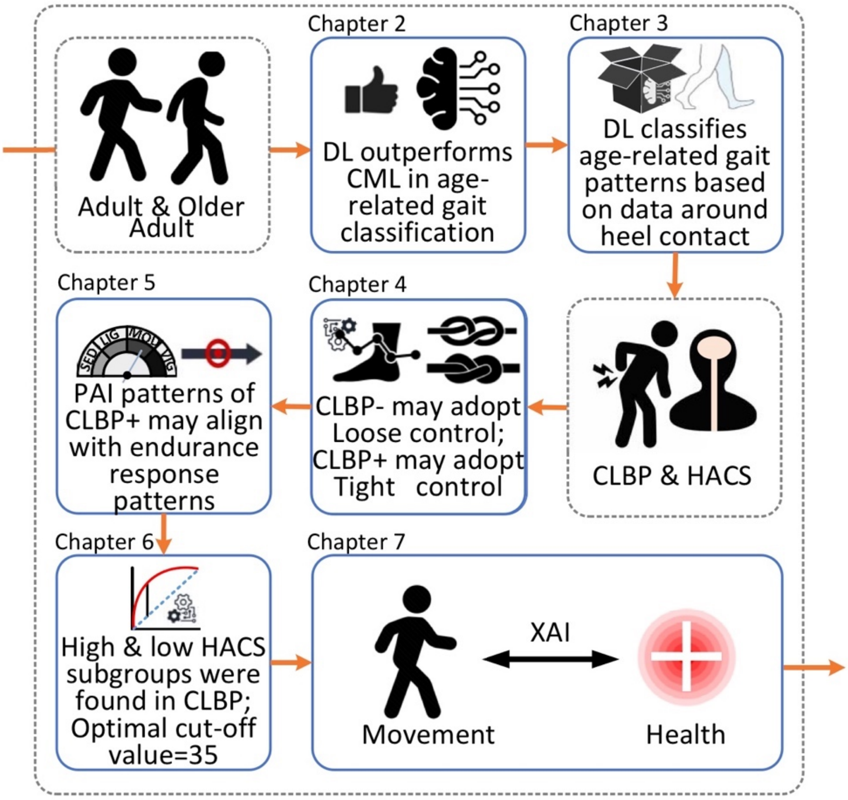

Figure 1. The summaries of findings this thesis. DL: deep learning; CML: conventional machine learning; CLBP: chronic low back pain; HACS: human assumed central sensitization; CLBP+: patients with CLBP and high levels of HACS; CLBP-: patients with CLBP and low levels of HACS; PAI: physical activity intensity; XAI: Explainable artificial intelligence.

CML has been widely used to classify age-related gait patterns based on gait outcomes [6, 7]. However, these gait outcomes are handcrafted, requiring extensive labor for design and selection [8]. Furthermore, the offline computation of gait outcomes may limit their use in real-time applications [9], such as monitoring gait of older adults in daily living environments. To overcome these challenges, DL has been introduced to gait analysis, utilizing accelerometer signal data as input, rather than gait outcomes (as discussed in **Chapter 2**). The results indicated that DL outperformed CML, achieving a high classification performance with an area under the curve (AUC) exceeding 0.94, compared to the highest AUC of 0.83 for the best CML models. These results not only highlighted the potential of DL in classifying age-related gait patterns but also suggested that DL had learned specific features that can efficiently capture the age-related changes in gait. However, the black-box nature of DL obscures the understanding of what the models have learned from the data.

To improve the transparency of DL, XAI was utilized in **Chapter 3**. The results showed that the accelerometer signal data spanning from the terminal swing to the loading response phase contributed most to the classification process. This implies that there may be notable differences in acceleration and deceleration patterns between adults and older adults during walking. The study also found that these distinct patterns within a single stride were sufficient for a convolutional neural network (CNN) model to accurately differentiate between adult and older adult groups, achieving an AUC of 0.89. Additionally, employing a recurrent neural network (RNN) model allowed for an examination of gait evolution by analyzing acceleration and deceleration patterns across strides, revealing subtle differences and relationships that might indicate postural control ability [10]. RNN achieved a high level of classification accuracy, with an AUC of 0.94. These distinct patterns could be associated with age-related changes, such as declining muscular capabilities, but our study lacked adequate data to confirm these links. To gain a more comprehensive understanding of these patterns, incorporating kinematics and kinetics [11], as well as electromyography (EMG) data [12], would be beneficial. For example, ground reaction forces (GRF) can provide insights into the forces involved in body weight absorption and push-off during walking, range of motion could track alterations at specific joints, and EMG can reveal muscle activation patterns.

Declining balance is associated with aging [13] and it is often observed that older adults exhibit less stable gait patterns [14]. **Chapter 3,** however, showed that DL mostly employed accelerometer signals in the anteroposterior (AP) and vertical (V) directions to distinguish older adults from adults, rather than relying on mediolateral (ML) direction data. To further explore gait stability differences between adults and older adults, additional steps could be taken. First, using long-term data (such as 24-hour recordings) from daily living environments may better reflect actual gait performance during daily living [15]. The walking task in our study, a simple 3-minute walk, lacked complexity and did not include perturbations encountered in daily life, such as navigating around obstacles or moving through doorways, which impose higher demands and often require multitasking. Second, this study used a single accelerometer to record walking data, which may not fully capture all aspects of balance-related postural control. Therefore, using more sensors, like additional accelerometers or gyroscopes, could yield a better understanding of the balance during gait of older adults.

Chapters 2 and 3 highlighted the superior performance of DL for classifying age-related gait patterns, and the use of XAI to enhance understanding of gait changes in aging. The findings indicated that distinctions in walking patterns between adults and older adults can be characterized by acceleration and deceleration patterns within a single stride or across multiple strides. This knowledge enhances the comprehension of how gait changes as individuals age.

Movement Characteristics of Patients with Chronic Low Back Pain

The methodologies of AI-driven gait analysis and XAI in Chapters 2 and 3, may shed light not only on exploring the relationship between gait and aging but also on movement and other health-related conditions, such as back pain.

Low back pain (LBP) is highly prevalent, with up to 84% of individuals experiencing LBP at least once in their lifetime [16]. In approximately 90% of patients, the cause of LBP is nonspecific [17], and around 20% of patients continue to report persistent back pain one year after the onset of acute LBP [18]. When LBP persists beyond three months, it is classified as chronic low back pain (CLBP). CLBP poses significant socioeconomic burdens and causes great individual suffering. Physical exercise is often recommended to manage CLBP but the effect size is modest [19].

CLBP is a heterogeneous condition [20]. Gait analysis reports inconsistent evidence when comparing patients with CLBP to healthy controls, including walking speed, step width, and stride variability [21-25]. Similarly, studies on PAI in patients with CLBP versus healthy controls have yielded mixed findings, with some indicating reduced daily PAI in patients with CLBP [26, 27] and others showing no significant difference [28, 29]. This inconsistent evidence highlights the heterogeneity and complexity of CLBP.

The presence of central sensitization (CS) in CLBP may be one of the key factors contributing to this heterogeneity, as CS is associated with long-lasting chronic pain [30]. CS is defined as an increased responsiveness of nociceptive neurons in the central nervous system to their normal or subthreshold afferent input [31]. However, due to the current inability to directly measure CS mechanisms in individual humans, the term human assumed central sensitization (HACS) is used [32]. Since movement may be changed due to pain, it was hypothesized that different levels of HACS might be associated with gait patterns (Chapter 4) and PAI patterns (Chapter 5) in patients with CLBP.

Chapter 4 used CML to classify gait patterns in patients with CLBP and with low or high levels of HACS (CLBP- and CLBP+, respectively). A satisfactory performance (an accuracy rate of 84.4%) was achieved, indicating distinct gait patterns between CLBP- and CLBP+ groups. XAI was employed to explain the classification process, revealing that CLBP- exhibited higher gait smoothness and stability, whereas CLBP+ showed a more regular, less variable, and more predictable gait pattern. These findings suggested different motor control strategies: “loose control” in CLBP- characterized by reduced trunk muscle excitability leading to reduced control over trunk movements [33], and “tight control” in CLBP+ marked by increased activation and co-contraction of trunk muscles for enhanced movement control [33]. These results could inform the categorization of CLBP patients into different treatment groups, but further research is needed for more direct validation of these motor control strategies. For example, the increased activation and co-contraction of trunk muscles might be observable through EMG [34]. Moreover, gait perturbations can be used to examine the presence of tight control. Tight control might result in increased trunk stiffness to counterbalance anticipated perturbations and hence will show larger trunk displacement due to the unanticipated perturbations [35].

In **Chapter 5**, unsupervised CML with ante-hoc explainability was employed to explore and clarify PAI patterns in CLBP- and CLBP+ groups. The findings showed that patients in the CLBP- group tended to break tasks into smaller bouts of activity, taking frequent and short rests in between, while patients in the CLBP+ group engaged in prolonged periods of activity with extended rest intervals. These patterns might be interpreted by the avoidance-endurance model (AEM) [36]. AEM postulates that a subgroup of patients shows a pattern of fear-avoidance responses caused by high fear of pain, leading to avoidance behaviour; another subgroup of patients shows a pattern of endurance responses with overuse and overload of physical structures, despite having pain. The endurance responses seemed to correspond with CLBP+ group, while the behaviour of CLBP- group might not necessarily align with fear-avoidance responses, as there was no evidence to support the presence of fear beliefs. These findings also need to be examined by future studies, since this chapter did not directly assess the fear or endurance belief. To assess avoidance and endurance beliefs more directly, the avoidance-endurance questionnaire [37] could be utilized. Alternatively, fear beliefs could be evaluated through responses to perturbations during gait, where individuals fearful of falling might demonstrate anticipatory postural adjustments prior to the perturbation [38, 39]

Chapters 4 and 5 provide insights into the movement characteristics of CLBP- and CLBP+ groups within their daily living environments. However, the relatively small sample size in these studies (n=42) is a limitation. Apart from this, further longitudinal research is required to explore potential causal links between movement changes and HACS. It is possible that HACS is a consequence of the observed movement patterns. Tight control presumably leads to increased muscle activation and co-contraction, resulting in higher spinal loading. This continuous muscle co-contraction, even at rest [40], can produce long-lasting peripheral noxious stimuli, potentially contributing to the development and/or persistence of HACS [41]. Additionally, patients with an endurance response pattern may subject their muscles to ongoing stress and repetitive strain during prolonged activities, leading to laxity and inflammation [42]. This persistent nociceptive input could also play a role in HACS development [36]. Conversely, HACS might also be a cause of the observed movement differences. Patients in CLBP+ group show higher levels of HACS. Considering HACS's mechanism, the relationship between movement and pain may become irrelevant, as pain can occur without tissue loading [43]. Thus, these patients might manage their trunk by increasing spinal loading, allowing them to continue activities for longer periods. Therefore, HACS could be either a result or a cause of changes in movement, or perhaps both.

In Chapters 4 and 5, HACS levels of CBLP- and CLBP+ were determined using a cut-off value of 40 from the CSI questionnaire [44] which serves as an indirect method for assessing HACS. It has been reported that the cut-off value of 40 for the CSI may vary depending on different types of musculoskeletal pain [45, 46], as well as across various cultural and national contexts [47]. Currently, the gold standard for diagnosing HACS is unavailable and hence, **Chapter 6** of this thesis discussed the exploration of HACS-related subgroups using unsupervised CML with ante-hoc explainability based on clinical outcomes (questionnaire data) to establish a more accurate cut-off value for the CSI. The findings found two distinct HACS-related groups and suggested adopting 35 as the new cut-off value for the Dutch-speaking population with CLBP.

The methodologies of AI, particularly XAI, as employed in chapters 4 to 6, provide a data-driven approach to enhance our understanding of complex issues related to HACS, CLBP, and movement characteristics. Ultimately, these methodologies enable researchers and clinicians to develop more personalized and effective strategies for assessing and managing conditions associated with HACS and CLBP.

Selection of AI Models for HMS

This thesis has demonstrated that movement analysis can benefit from the application of AI. The integration of AI into movement analysis begins with the selection of appropriate AI models.

Fig. 3 graphically illustrates the hierarchy of commonly used terms related to AI. AI is a broad concept, with subsets of machine learning, XAI, and others. Within machine learning, a spectrum of methodologies includes CML, DL, and others. Depending on the learning paradigm, AI models can be categorized into supervised learning models [48], unsupervised learning models [49], and others such as semi-supervised learning models. For XAI, it contains two main categories: ante-hoc explainability and post-hoc explainability [50]. Some supervised and unsupervised CML models are also categorized under ante-hoc explainability, e.g., K-nearest neighbors and K-means [51].

a. Supervised Learning or Unsupervised Learning

Supervised learning is tasked with recognizing patterns from labeled datasets. For instance, **Chapter 2** employed five supervised DL models to learn the gait patterns of adults and older adults. These models were trained using accelerometer data as input features, along with their respective labels ("adults" or "older adults"). Once trained, these models are capable of predicting labels for new accelerometer data inputs. The underlying assumption is that if there are distinct patterns that correlate with the labels, then the models should classify them with a high degree of accuracy. For example, if older adults and adults have different gait patterns, DL models are expected to classify them accurately, and vice versa.

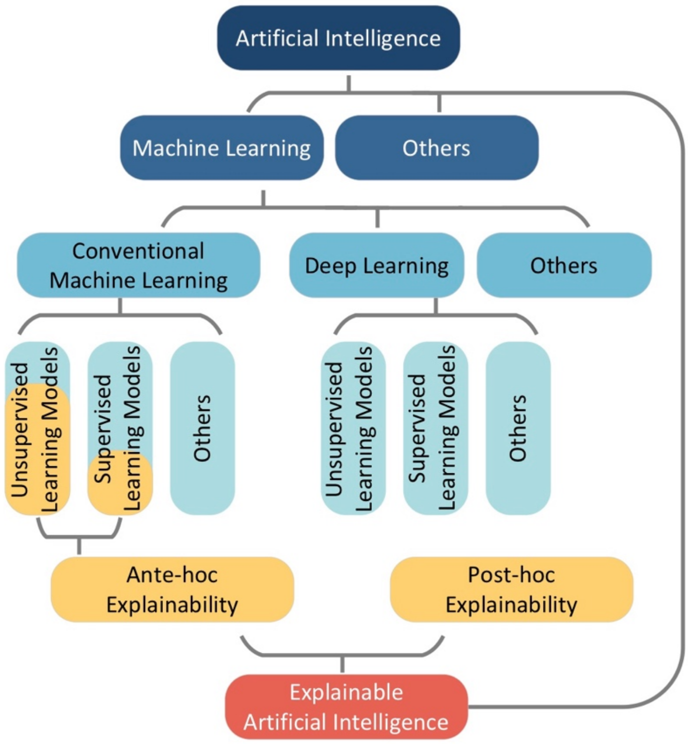

Figure 3. Hierarchy of artificial intelligence related terms.

Unsupervised learning, on the other hand, learns patterns within unlabeled data and automatically clusters them. In **Chapter 6,** four unsupervised CML models were employed to explore potential HACS-related subgroups. Clinical outcomes data from questionnaires were used as input, and the models determined the clusters automatically. The rationale for unsupervised learning is that within a dataset, if a subset of data samples is in close proximity to each other and distinct from others, they form a cluster.

In summary, the selection of a supervised or unsupervised model depends on the objectives of the analysis. If the goal is to recognize/predict specific gait patterns (**Chapters 2**, and **3**) or examine distinct gait patterns between groups (**Chapter 4**), supervised learning models are appropriate. Conversely, to uncover latent patterns (**Chapter 5**) or potential subgroups within the targeted population (**Chapter 6**), unsupervised learning models are preferable.

b. CML or DL

Once the decision is made between supervised or unsupervised learning, the next step is to select the specific models within CML or DL.

CML and DL each present different advantages and disadvantages. CML can be performing well even for small datasets. It is a critical consideration in medical fields where acquiring extensive gait data from patients can be challenging [52]. DL, on the other hand, has been shown to surpass CML in classifying gait patterns [53, 54], although this enhanced performance comes at the cost of increased model complexity, which can impede the interpretability of the models [51]. Consequently, **Chapter 2** utilized supervised DL models for accurate classification of age-related gait patterns, whereas **Chapter 4** employed supervised CML models to examine the relationships between CLBP, HACS, and gait outcomes, prioritizing explainability. Beyond objective constraints, such as dataset size, computing power, and etc., choosing between DL and CML often reflects a trade-off between explainability and performance [55]. Hence, for a better explainability, only unsupervised CML with ante-hoc explainability were used in this thesis (**Chapters 5**, and **6**).

c. Selection of Specific Models

This thesis evaluates the performance of various supervised CML models in **Chapters 2** and **4,** including support vector machines (SVM), random forests (RF), artificial neural networks, k-nearest neighbors, and naive bayes. These CML models have been extensively compared in other research, and RF and SVM are frequently utilized classifiers in gait analysis [6, 56]. Despite well-tuned RF and SVM potentially achieving similar levels of performance, as demonstrated in **Chapter 2**, RF is recommended over SVM for gait analysis, as suggested in **Chapter 4**. The performance of SVM is highly influenced by the choice of kernel function, which should align with the data's underlying nature (e.g., linear or non-linear) [57]. As the nature of the data is often not know, determining the optimal kernel function for SVM requires specialized knowledge and additional analysis. Conversely, RF is more user-friendly.

Regarding the supervised DL models, five supervised DL models within three categories were discussed in **Chapter 2**, including CNN, RNN (gate recurrent unit, long short-term memory, and bi-directional long short-term memory), and hybrid neural networks (HNN; convolutional long short-term memory). These supervised DL models are able to use time-series signal data as input. The time-dependent information within time-series data is critical for certain gait analyses and should be considered when selecting models. For instance, CNN excels at extracting local spatial-temporal features [58] and may capture independent gait outcomes (e.g., step length). Consequently, CNN could be effective at recognizing gait patterns observable within one or two strides, like asymmetric gait patterns [59]. RNN is specifically designed to learn both short- and long-term features [58] and may capture temporal-dependent gait outcomes (e.g., gait regularity). Therefore, RNN might benefit from analyzing extended time-series data to detect gradual changes in gait patterns, such as fatigue-related gait patterns [60]. HNN contains the structures of CNN and RNN. It is expected to benefit from both structures. However, **Chapter 2** did not support this idea. The unique characteristics of these models may lead to different intention of applications and may have different data processing requirements. For instance, CNN might achieve satisfactory performance with only one or two strides of data, while RNN and HNN might require longer data, potentially exceeding 10 seconds (8 strides), as the minimum for optimal performance (as noted in **Chapter 2**).

Limitations of XAI in Human Movement Sciences

This thesis showcases the practical implementation of XAI to enhance the transparency and explainability of human movement analysis, thereby making it more trustworthy and facilitating the extraction of scientific knowledge from the data. However, several challenges remain.

A fundamental challenge in XAI is the absence of a clear metric for evaluating the “interpretability” or “explainability” [61]. Due to the lack of a ground truth, the classic metrics (such as accuracy, sensitivity, and etc.) are not available. Hence, it is difficult to compare the performance among XAI approaches. In **Chapter 4**, the XAI approach, SHapley Additive exPlanations (SHAP) [62], was employed to explain the CML model (RF), instead of using the conventional Gini impurity [63]. This choice was made based on the authors’ experience, as SHAP has a stronger theoretical foundation [64]. Although we suggested that SHAP will offer superior explanations, it may be challenging to provide definitive proof of this. Additionally, due to the lack of metrics, it becomes a complex task to assess the trustworthiness of the explanations. The explanations provided by XAI may not align with domain expertise. These discrepancies may provide fresh insights when the explanations are trustworthy. However, these discrepancies could also be raised by biases or unknown errors. Due to the lack of metrics, caution is advised when using explanations provided by XAI methods. Recent efforts have been made in evaluating the effectiveness of XAI [65, 66]. Measurement techniques, such as the goodness checklist and the explanation satisfaction scale, seem to be a good step in the direction of evaluating XAI.

The explanatory capabilities of XAI are still limited. In **Chapter 3**, XAI was used to explain the classification process for age-related gait patterns based on accelerometer time-series. XAI provided explanations by highlighting the importance of specific segments of accelerometer data. However, the authors need to explain these XAI-generated explanations by manually linking these segments with the specific gait events and guessing why these gait events are important when classifying age-related gait patterns. The current explanations may not provide a complete picture of what happened during these gait events. XAI should not leave the majority explanation generation to users, as their diverse backgrounds and knowledge may lead to different interpretations and explanations [67]. Additionally, XAI generates explanations based on correlations and associations within the data, which may not be sufficient to unveil cause-effect relationships [68]. In **Chapter 4**, XAI identified the top 10 critical gait outcomes that have the most influence on the classification process. However, it remains challenging to determine whether changes in these gait outcomes are the causes or the consequences of HACS.

XAI is still in its early stages, especially in the context of human movement analysis. To effectively apply XAI in this field, collaborations with interdisciplinary fields like human-computer interaction and data sciences are essential. However, these collaborations can be challenging to establish, potentially hindering the application of XAI in human movement analysis. Therefore, the development of automated XAI solutions becomes crucial [69]. These services should be designed to offer substantial assistance to a broad spectrum of end-users, including non-technical experts. In line with this vision, steps have been taken to contribute. In **Chapters 2**, **3**, and **6** of this thesis, the authors have made the code openly accessible to fellow researchers, thereby facilitating the reuse and expansion of this work, and fostering collaborative efforts in the application of XAI to human movement analysis.

Future of XAI in Human Movement Sciences

In the domain of human movement analysis, numerous AI-driven systems, such as AI-driven gait analysis, have exhibited impressive levels of accuracy and performance. Looking forward, XAI is poised to clarify the opaque "black-box" nature of AI, laying the groundwork for the wider implementation of AI-driven solutions, not just in gait analysis but throughout various healthcare domains.

The widespread use of wearable smart devices (such as smart watches) has generated massive amounts of individual movement data and provided the computational capacity necessary for deploying AI-driven gait analysis. Trustworthy AI-driven gait analysis will facilitate personalized health monitoring, early disease prediction, and the assessment of treatment effectiveness, among other benefits. Leveraging both historical clinical data and the burgeoning influx of newly collected data, AI-generated insights have the potential to rapidly advance our understanding of health and movement. This, in turn, assists clinicians and computer scientists in further refining AI-driven gait analysis. These innovations are expected to not only improve gait analysis but also transform a broad spectrum of healthcare practices, making the processes of prognosis, diagnosis, treatment, patient follow-up, and clinical decision-making more streamlined, precise, and efficient.

AI in human movement analysis is becoming integral to clinical settings and everyday life. Rather than supplanting clinical experts, AI will support them in treating diseases and monitoring the health in daily live environments more effectively. It will not disrupt patient-clinician connections but instead help patients better understand their individual health. AI is not an isolated knowledge generator; it is a catalyst for deepening the understanding of movement scientists in health issues related to movement. Ultimately, AI contributes to a deeper comprehension of movement and paves the way for more effective treatments.

Conclusion

This thesis further highlighted the importance of AI in movement analysis and demonstrated the potential of XAI in enhancing our understanding of gait patterns in healthy older adults and movement characteristics in patients with back pain. It provided guidance on selecting appropriate AI models for movement (especially gait) analysis through a comparison of various AI models. Based on the insights from XAI, it revealed that differences in gait patterns between adults and older adults can be characterized by acceleration and deceleration patterns. In the context of CLBP, the findings indicated different motor control strategies and pain response patterns among patients. These findings highlighted the possibility of personalized treatment approaches in managing CLBP and HACS. The methodologies used in this thesis, including CML, DL, and XAI, bring us closer to understanding the complex interplay between movement, and aging or back pain. Looking ahead, further steps are necessary to developing more transparent and reliable AI tools in healthcare.

Reference

[1] U. Bechtold, N. Stauder, and M. Fieder, "Let's walk it: Mobility and the perceived quality of life in older adults," *International journal of environmental research and public health,* vol. 18, no. 21, pp. 11515, 2021.

[2] M. Intriago, G. Maldonado, R. Guerrero, O. Messina, and C. Rios, "Bone mass loss and Sarcopenia in Ecuadorian patients," *Journal of Aging Research,* vol. 2020, 2020.

[3] A. Aboutorabi, M. Arazpour, M. Bahramizadeh, S. W. Hutchins, and R. Fadayevatan, "The effect of aging on gait parameters in able-bodied older subjects: a literature review," *Aging clinical and experimental research,* vol. 28, no. 3, pp. 393-405, 2016.

[4] P. Terrier, and F. Reynard, "Effect of age on the variability and stability of gait: a cross-sectional treadmill study in healthy individuals between 20 and 69 years of age," *Gait & posture,* vol. 41, no. 1, pp. 170-174, 2015.

[5] A. Hanley, C. Silke, and J. Murphy, "Community-based health efforts for the prevention of falls in the elderly," *Clinical interventions in aging*, pp. 19-25, 2011.

[6] Y. Zhou, R. Romijnders, C. Hansen, J. van Campen, W. Maetzler, T. Hortobagyi, and C. J. C. Lamoth, "The detection of age groups by dynamic gait outcomes using machine learning approaches," *Scientific Reports,* vol. 10, no. 1, Mar, 2020.

[7] C. Prakash, R. Kumar, and N. Mittal, "Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges," *Artificial Intelligence Review,* vol. 49, pp. 1-40, 2018.

[8] F. Gu, M.-H. Chung, M. Chignell, S. Valaee, B. Zhou, and X. Liu, “A survey on deep learning for human activity recognition,” *ACM Computing Surveys (CSUR),* vol. 54, no. 8, pp. 1-34, 2021.

[9] Y. Hutabarat, D. Owaki, and M. Hayashibe, “Recent advances in quantitative gait analysis using wearable sensors: a review,” *IEEE Sensors Journal*, 2021.

[10] D. Borah, S. Wadhwa, U. Singh, S. L. Yadav, M. Bhattacharjee, and V. Sindhu, “Age related changes in postural stability,” *Indian J Physiol Pharmacol,* vol. 51, no. 4, pp. 395-404, 2007.

[11] K. A. Boyer, R. T. Johnson, J. J. Banks, C. Jewell, and J. F. Hafer, “Systematic review and meta-analysis of gait mechanics in young and older adults,” *Experimental gerontology,* vol. 95, pp. 63-70, 2017.

[12] A. Schmitz, A. Silder, B. Heiderscheit, J. Mahoney, and D. G. Thelen, “Differences in lower-extremity muscular activation during walking between healthy older and young adults,” *Journal of electromyography and kinesiology,* vol. 19, no. 6, pp. 1085-1091, 2009.

[13] R. W. Bohannon, P. A. Larkin, A. C. Cook, J. Gear, and J. Singer, “Decrease in timed balance test scores with aging,” *Physical therapy,* vol. 64, no. 7, pp. 1067-1070, 1984.

[14] S. M. Bruijn, O. Meijer, P. Beek, and J. H. van Dieen, “Assessing the stability of human locomotion: a review of current measures,” *Journal of the Royal Society Interface,* vol. 10, no. 83, pp. 20120999, 2013.

[15] I. Hillel, E. Gazit, A. Nieuwboer, L. Avanzino, L. Rochester, A. Cereatti, U. D. Croce, M. O. Rikkert, B. R. Bloem, and E. Pelosin, “Is every-day walking in older adults more analogous to dual-task walking or to usual walking? Elucidating the gaps between gait performance in the lab and during 24/7 monitoring,” *European review of aging and physical activity,* vol. 16, no. 1, pp. 1-12, 2019.

[16] B. F. Walker, “The prevalence of low back pain: a systematic review of the literature from 1966 to 1998,” *Clinical Spine Surgery,* vol. 13, no. 3, pp. 205-217, 2000.

[17] J. Hartvigsen, M. J. Hancock, A. Kongsted, Q. Louw, M. L. Ferreira, S. Genevay, D. Hoy, J. Karppinen, G. Pransky, J. Sieper, R. J. Smeets, M. Underwood, and W. Lancet Low Back Pain Series, “What low back pain is and why we need to pay attention,” *Lancet,* vol. 391, no. 10137, pp. 2356-2367, Jun 9, 2018.

[18] M. Von Korff, and K. Saunders, “The course of back pain in primary care,” *Spine,* vol. 21, no. 24, pp. 2833-2837, 1996.

[19] N. E. Foster, J. R. Anema, D. Cherkin, R. Chou, S. P. Cohen, D. P. Gross, P. H. Ferreira, J. M. Fritz, B. W. Koes, W. Peul, J. A. Turner, C. G. Maher, and W. Lancet Low Back Pain Series, “Prevention and treatment of low back pain: evidence, challenges, and promising directions,” *Lancet,* vol. 391, no. 10137, pp. 2368-2383, Jun 9, 2018.

[20] D. R. Fourney, G. Andersson, P. M. Arnold, J. Dettori, A. Cahana, M. G. Fehlings, D. Norvell, D. Samartzis, and J. R. Chapman, “Chronic low back pain: a heterogeneous

condition with challenges for an evidence-based approach," *Spine,* vol. 36, pp. S1-S9, 2011.

[21] C. J. Lamoth, O. G. Meijer, A. Daffertshofer, P. I. Wuisman, and P. J. Beek, "Effects of chronic low back pain on trunk coordination and back muscle activity during walking: changes in motor control," *European Spine Journal,* vol. 15, pp. 23-40, 2006.

[22] C. J. Lamoth, J. F. Stins, M. Pont, F. Kerckhoff, and P. J. Beek, "Effects of attention on the control of locomotion in individuals with chronic low back pain," *Journal of neuroengineering and rehabilitation,* vol. 5, no. 1, pp. 1-8, 2008.

[23] G. Christe, F. Kade, B. M. Jolles, and J. Favre, "Chronic low back pain patients walk with locally altered spinal kinematics," *Journal of biomechanics,* vol. 60, pp. 211-218, 2017.

[24] S. Ebrahimi, F. Kamali, M. Razeghi, and S. A. Haghpanah, "Comparison of the trunk-pelvis and lower extremities sagittal plane inter-segmental coordination and variability during walking in persons with and without chronic low back pain," *Human movement science,* vol. 52, pp. 55-66, 2017.

[25] S. P. Gombatto, T. Brock, A. DeLork, G. Jones, E. Madden, and C. Rinere, "Lumbar spine kinematics during walking in people with and people without low back pain," *Gait & posture,* vol. 42, no. 4, pp. 539-544, 2015.

[26] M. Soysal, B. Kara, and M. N. Arda, "Assessment of physical activity in patients with chronic low back or neck pain," *Turk Neurosurg,* vol. 23, no. 1, pp. 75-80, 2013.

[27] C. G. Ryan, P. M. Grant, P. M. Dall, H. Gray, M. Newton, and M. H. Granat, "Individuals with chronic low back pain have a lower level, and an altered pattern, of physical activity compared with matched controls: an observational study," *Australian Journal of Physiotherapy,* vol. 55, no. 1, pp. 53-58, 2009.

[28] J. A. Verbunt, K. R. Westerterp, G. J. van der Heijden, H. A. Seelen, J. W. Vlaeyen, and J. A. Knottnerus, "Physical activity in daily life in patients with chronic low back pain," *Arch Phys Med Rehabil,* vol. 82, no. 6, pp. 726-30, Jun, 2001.

[29] M. G. van Weering, M. M. Vollenbroek-Hutten, T. M. Tonis, and H. J. Hermens, "Daily physical activities in chronic lower back pain patients assessed with accelerometry," *Eur J Pain,* vol. 13, no. 6, pp. 649-54, Jul, 2009.

[30] J. A. Echeita, H. R. S. Preuper, R. Dekker, I. Stuive, H. Timmerman, A. P. Wolff, and M. F. Reneman, "Central Sensitisation and functioning in patients with chronic low back pain: protocol for a cross-sectional and cohort study," *Bmj Open,* vol. 10, no. 3, Mar, 2020.

[31] J. D. Loeser, and R.-D. Treede, "The Kyoto protocol of IASP basic pain Terminology☆," *Pain,* vol. 137, no. 3, pp. 473-477, 2008.

[32] I. Schuttert, H. Timmerman, K. K. Petersen, M. E. McPhee, L. Arendt-Nielsen, M. F. Reneman, and A. P. Wolff, "The Definition, Assessment, and Prevalence of (Human Assumed) Central Sensitisation in Patients with Chronic Low Back Pain: A Systematic Review," *Journal of Clinical Medicine,* vol. 10, no. 24, pp. 5931, 2021.

[33] J. H. van Dieen, N. P. Reeves, G. Kawchuk, L. R. van Dillen, and P. W. Hodges, "Motor Control Changes in Low Back Pain: Divergence in Presentations and Mechanisms," *Journal of Orthopaedic & Sports Physical Therapy,* vol. 49, no. 6, pp. 370-379, Jun, 2019.

[34] C. T. Candotti, J. F. Loss, A. M. S. Pressi, F. A. de Souza Castro, M. La Torre, M. de Oliveira Melo, L. D. Araújo, and M. Pasini, "Electromyography for assessment of pain in low back muscles," *Physical therapy,* vol. 88, no. 9, pp. 1061-1067, 2008.

[35] N. W. Mok, S. G. Brauer, and P. W. Hodges, "Failure to use movement in postural strategies leads to increased spinal displacement in low back pain," *Spine,* vol. 32, no. 19, pp. E537-E543, 2007.

[36] M. I. Hasenbring, and J. A. Verbunt, "Fear-avoidance and Endurance-related Responses to Pain: New Models of Behavior and Their Consequences for Clinical Practice," *Clinical Journal of Pain,* vol. 26, no. 9, pp. 747-753, Nov-Dec, 2010.

[37] M. I. Hasenbring, D. Hallner, and A. C. Rusu, "Fear-avoidance-and endurance-related responses to pain: development and validation of the Avoidance-Endurance Questionnaire (AEQ)," *European Journal of Pain,* vol. 13, no. 6, pp. 620-628, 2009.

[38] S. B. Swart, R. den Otter, and C. J. Lamoth, "Anticipatory control of human gait following simulated slip exposure," *Scientific Reports,* vol. 10, no. 1, pp. 9599, 2020.

[39] T. J. Ellmers, A. Maslivec, and W. R. Young, "Fear of falling alters anticipatory postural control during cued gait initiation," *Neuroscience,* vol. 438, pp. 41-49, 2020.

[40] A. Schinkel-Ivy, B. C. Nairn, and J. D. Drake, "Investigation of trunk muscle co-contraction and its association with low back pain development during prolonged sitting," *Journal of Electromyography and Kinesiology,* vol. 23, no. 4, pp. 778-786, 2013.

[41] J. Nijs, S. Z. George, D. J. Clauw, C. Fernández-de-las-Peñas, E. Kosek, K. Ickmans, J. Fernández-Carnero, A. Polli, E. Kapreli, and E. Huysmans, "Central sensitisation in chronic pain conditions: latest discoveries and their potential for precision medicine," *The Lancet Rheumatology,* vol. 3, no. 5, pp. e383-e392, 2021.

[42] M. I. Hasenbring, N. E. Andrews, and G. Ebenbichler, "Overactivity in Chronic Pain, the Role of Pain-related Endurance and Neuromuscular Activity An Interdisciplinary, Narrative Review," *Clinical Journal of Pain,* vol. 36, no. 3, pp. 162-171, Mar, 2020.

[43] J. Nijs, A. Lahousse, E. Kapreli, P. Bilika, İ. Saraçoğlu, A. Malfliet, I. Coppieters, L. De Baets, L. Leysen, and E. Roose, "Nociplastic pain criteria or recognition of central sensitization? Pain phenotyping in the past, present and future," *Journal of clinical medicine,* vol. 10, no. 15, pp. 3203, 2021.

[44] E. E. Bennett, K. M. Walsh, N. R. Thompson, and A. A. Krishnaney, "Central sensitization inventory as a predictor of worse quality of life measures and increased length of stay following spinal fusion," *World neurosurgery,* vol. 104, pp. 594-600, 2017.

[45] R. Neblett, H. Cohen, Y. Choi, M. M. Hartzell, M. Williams, T. G. Mayer, and R. J. Gatchel, "The Central Sensitization Inventory (CSI): establishing clinically significant values for

identifying central sensitivity syndromes in an outpatient chronic pain sample,” *J Pain,* vol. 14, no. 5, pp. 438-45, May, 2013.

[46] A. Mibu, T. Nishigami, K. Tanaka, M. Manfuku, and S. Yono, “Difference in the impact of central sensitization on pain-related symptoms between patients with chronic low back pain and knee osteoarthritis,” *Journal of Pain Research,* vol. 12, pp. 1757, 2019.

[47] R. Neblett, “The central sensitization inventory: A user's manual,” *Journal of Applied Biobehavioral Research,* vol. 23, no. 2, Jun, 2018.

[48] P. Cunningham, M. Cord, and S. J. Delany, "Supervised learning," *Machine learning techniques for multimedia: case studies on organization and retrieval*, pp. 21-49: Springer, 2008.

[49] H. B. Barlow, “Unsupervised learning,” *Neural computation,* vol. 1, no. 3, pp. 295-311, 1989.

[50] T. Speith, "A review of taxonomies of explainable artificial intelligence (XAI) methods." pp. 2239-2250.

[51] P. Lisboa, S. Saralajew, A. Vellido, R. Fernández-Domenech, and T. Villmann, “The coming of age of interpretable and explainable machine learning models,” *Neurocomputing,* vol. 535, pp. 25-39, 2023.

[52] P. Khera, and N. Kumar, “Role of machine learning in gait analysis: a review,” *Journal of Medical Engineering & Technology,* vol. 44, no. 8, pp. 441-467, 2020.

[53] C. Tunca, G. Salur, and C. Ersoy, “Deep learning for fall risk assessment with inertial sensors: Utilizing domain knowledge in spatio-temporal gait parameters,” *IEEE journal of biomedical and health informatics,* vol. 24, no. 7, pp. 1994-2005, 2019.

[54] R. Kaur, J. Levy, R. W. Motl, R. Sowers, and M. E. Hernandez, “Deep Learning for Multiple Sclerosis Differentiation Using Multi-Stride Dynamics in Gait,” *IEEE Transactions on Biomedical Engineering*, 2023.

[55] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining explanations: An overview of interpretability of machine learning." pp. 80-89.

[56] E. Balaji, D. Brindha, and R. Balakrishnan, “Supervised machine learning based gait classification system for early detection and stage classification of Parkinson’s disease,” *Applied Soft Computing,* vol. 94, pp. 106494, 2020.

[57] R. Zhang, and W. Wang, “Facilitating the applications of support vector machine by using a new kernel,” *Expert systems with applications,* vol. 38, no. 11, pp. 14225-14230, 2011.

[58] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, “Deep learning for time series classification: a review,” *Data mining and knowledge discovery,* vol. 33, no. 4, pp. 917-963, 2019.

[59] S. Viteckova, P. Kutilek, Z. Svoboda, R. Krupicka, J. Kauler, and Z. Szabo, “Gait symmetry measures: A review of current and prospective methods,” *Biomedical Signal Processing and Control,* vol. 42, pp. 89-100, 2018.

[60] F. A. Barbieri, P. C. R. Dos Santos, E. Lirani-Silva, R. Vitório, L. T. B. Gobbi, and J. H. Van Diëen, “Systematic review of the effects of fatigue on spatiotemporal gait parameters,” *Journal of back and musculoskeletal rehabilitation,* vol. 26, no. 2, pp. 125-131, 2013.

[61] L. Arras, A. Osman, and W. Samek, “CLEVR-XAI: A benchmark dataset for the ground truth evaluation of neural network explanations,” *Information Fusion,* vol. 81, pp. 14-40, 2022.

[62] S. M. Lundberg, and S.-I. Lee, “A unified approach to interpreting model predictions,” *Advances in neural information processing systems,* vol. 30, 2017.

[63] C. Kelly, and K. Okada, "Variable interaction measures with random forest classifiers." pp. 154-157.

[64] G. Owen, *Game theory*: Emerald Group Publishing, 2013.

[65] S. Mohseni, N. Zarei, and E. D. Ragan, “A multidisciplinary survey and framework for design and evaluation of explainable AI systems,” *ACM Transactions on Interactive Intelligent Systems (TiiS),* vol. 11, no. 3-4, pp. 1-45, 2021.

[66] R. R. Hoffman, S. T. Mueller, G. Klein, and J. Litman, “Metrics for explainable AI: Challenges and prospects,” *arXiv preprint arXiv:1812.04608*, 2018.

[67] D. Doran, S. Schulz, and T. R. Besold, “What does explainable AI really mean? A new conceptualization of perspectives,” *arXiv preprint arXiv:1710.00794*, 2017.

[68] C. Molnar, G. Casalicchio, and B. Bischl, "Interpretable machine learning–a brief history, state-of-the-art and challenges." pp. 417-431.

[69] C. Molnar, *Interpretable machine learning*: Lulu. com, 2020.