

ID N. 24



Istituto di Calcolo
e Reti ad Alte Prestazioni

UNIVERSITÀ CAMPUS BIO-MEDICO DI ROMA

DEPARTMENT OF ENGINEERING

**INSTITUTE FOR HIGH-PERFORMANCE
COMPUTING AND NETWORKING**

**DEPARTMENT OF ENGINEERING, ICT AND
TECHNOLOGIES FOR ENERGY AND TRANSPORT**

**Italian National Ph.D. in Artificial Intelligence
Health and Life Sciences
XXXVII Cycle**

**Revolutionizing Elderly Home Care with AI Systems:
Towards Smart Monitoring and Personalized
Assistance in Rehabilitation Therapy**

Supervisors

Massimo Esposito

Umberto Maniscalco

Candidate

Ciro Mennella

January, 2025

To my family.

Acknowledgements

I would like to express my sincere gratitude to my supervisors Massimo and Umberto, for their invaluable technical and professional guidance throughout my PhD studies. The profound esteem and friendship we share with each other have made this journey even more rewarding and enriching. I am deeply grateful for the opportunity to work with them and hope that this collaboration will lead to many more successful research endeavors.

I would also like to thank my wife Martina, for her steady support throughout this gratifying and challenging journey. Her extraordinary dedication to family responsibilities, especially during the demanding periods of my research, has been invaluable. I am truly grateful for her understanding and love.

Last but not least, I want to thank my daughter Chiara, for filling my life with the purest shape of love. Every goal I achieve, both in my professional and personal life, will always be thanks to her inspiration. I want to promise her that I will make an effort to spend more enjoyable free-time with her.

I am immensely grateful to all of you.

Abstract

As the age of the global population increases, the demand for innovative and practical solutions to support elderly care has attracted significant interest. Recent advances in digital technologies and computational science offer promising opportunities to extend the continuum of care at patients' homes.

This dissertation investigates the potential role of artificial intelligence (AI) methods and technology can have in rehabilitation medicine. The hypothesis is that AI can provide solutions for smart monitoring and assistance in home-based rehabilitation programs, empowering patients to actively participate in individualized treatment, improve self-management, and autonomously adhere to personalized long-term therapies.

State-of-the-art deep learning algorithms for computer vision and pattern recognition were analyzed and proposed in the monitoring and evaluation of human motion behavior during physical therapy exercises. The research also explored bias mitigation and resilience techniques in algorithm implementation to support the integration of effective and equitable AI technologies into future clinical applications. At the same time, the ethical challenges posed by these technological advancements were examined, focusing on the core principles and strategies to ensure the fairness of AI technologies that support clinical decision-making. A critical evaluation was also conducted on the regulatory framework within which these technologies could be effectively implemented and governed.

This thesis pioneers the development and application of AI-driven technologies in rehabilitation medicine. Its aim is to create more accessible, responsive, and impactful healthcare solutions to address the pressing global need for high-quality, cost-effective, and efficient treatments to manage the growing healthcare demands of non-communicable chronic diseases driven by population aging.

Contents

1	Background	10
1.1	Context and motivation	10
1.2	Aims and objectives	11
1.3	Thesis organization	12
2	Introduction	14
2.1	The impact of AI on healthcare	14
2.1.1	Foundations	14
2.1.2	Clinical applications	17
2.1.3	Ethical principles and guidelines	24
2.1.4	Legislative measures	29
2.2	The role of AI in rehabilitation	33
2.2.1	New trends in rehabilitation	33
2.2.2	Analysis of the state-of-the-art	35
2.2.3	AI-based systems and technologies	36
2.2.4	Algorithms	40
2.2.5	Clinical applications	43
2.2.6	Drawbacks	44
3	Data for human motion analysis	46
3.1	Data modeling and collection	46
3.1.1	Approaches for motion evaluation	46
3.1.2	A dataset for shoulder rehabilitation	48
3.1.3	A dataset for sarcopenia prevention	53
3.2	Synthetic data generation: a new research challenge	57
3.2.1	Existing methods	57
3.2.2	The proposed approach	61
3.2.3	Qualitative results	72

3.2.4	Quantitative results	75
3.2.5	Data validation	79
4	AI-based monitoring and assistive system: a proof-of-concept	83
4.1	General overview	83
4.1.1	Motion monitoring	85
4.1.2	Motion analysis	87
4.2	Experimental study	88
4.2.1	Data pre-processing	88
4.2.2	Deep learning algorithm	89
4.2.3	Model training and evaluation	90
4.2.4	Results	95
4.3	Comparative analysis with related works	95
5	Advancing activity recognition in healthcare: addressing fairness and robustness challenges	101
5.1	Bias and generalizability issues	101
5.2	Evaluating bias in human motion data	103
5.2.1	Dataset	103
5.2.2	Algorithms	108
5.2.3	Experimental design	113
5.2.4	Bias analysis	116
5.3	Enhancing cross-subject algorithm robustness	121
5.3.1	Dataset	121
5.3.2	Multiscale algorithms	122
5.3.3	Training and evaluation methods	126
5.3.4	Results	128
6	Discussion	131
6.1	Contribution summary	131
6.2	Limitations and future studies	132
6.3	Open challenges	135
6.3.1	The doctor-patient relationship	135
6.3.2	AI driving medical decision-making	139
6.3.3	Health data records	142
7	Conclusions	146

List of Figures

2.1	The AI impact on future healthcare system.	18
2.2	Category and characteristics of AI decision-support systems.	20
2.3	Literature evidences of diseases diagnosis using AI.	22
2.4	Ethical principles and guidelines.	27
2.5	Evidencing key European acts to govern AI in healthcare.	30
2.6	PRISMA flowchart of the results from the literature search.	37
2.7	Distribution of the number of articles published over years.	38
2.8	Results synthesis.	38
2.9	Categorization of machine learning algorithms.	40
3.1	A chronological sequence of left-side shoulder flexion, labeled within the context of ROM classes.	50
3.2	Distribution of ROM classes for gesture type and subject category.	52
3.3	Exercise correct repetition.	56
3.4	Compensatory movement.	56
3.5	Workflow of the proposed methodology.	61
3.6	Illustration of the pose-guided image generation method using the pose-guided conditioned diffusion model.	62
3.7	ControlNet architecture.	64
3.8	Samples of generated images in ‘camera view 1’.	73
3.9	Samples of generated images in ‘camera view 2’.	74
3.10	Analysis of joint-class Detection Score (DS-joint).	75
3.11	Correlation plot and distance correlation metric between original and synthetic dataset with background.	79
3.12	Correlation plot and distance correlation metric between original and synthetic dataset with masked background.	81
3.13	t-SNE plots to visualize movement classes boundaries preservation in the synthetic dataset generated from the original dataset.	82

4.1	Proposal of a system that counts valid and invalid repetitions and recognize compensatory motion patterns of the rehabilitation exercises.	84
4.2	COCO Keypoints used in MoveNet.	86
4.3	Outliers detection plot using Hotelling T^2 statistics.	92
4.4	Incorrect classes predicted by the ROM classifier for each exercise.	97
4.5	Incorrect classes predicted by the Compensation classifier for each exercise.	97
5.1	Distribution of exercise duration, measured in frames, within correct and incorrect motion classes.	106
5.2	Distribution of correct and incorrect repetitions for each exercise.	106
5.3	Duration of correct sequences, measured in frames, categorized by subject type across diverse rehabilitation exercises.	107
5.4	Duration of incorrect sequences, measured in frames, categorized by subject type across diverse rehabilitation exercises.	107
5.5	Architecture of tCNN Model.	109
5.6	Architecture of CNN-LSTM Model.	110
5.7	Architecture of ConvLSTM Model.	111
5.8	Architecture of Transformer model.	112
5.9	Experimental pipeline.	115
5.10	Barplot of demographic bias analysis.	117
5.11	Barplot of disability related bias.	119
5.12	Visualization of motion length in terms of the number of frames across exercises.	123
5.13	An example of motion variability in the dataset.	123
5.14	Architecture of multiscale tCNN model.	125
5.15	Architecture of multiscale CNN-LSTM model.	125
5.16	Architecture of multiscale ConvLSTM model.	126
5.17	Architecture of multiscale Transformer model.	127
5.18	Comparison of traditional and multiscale model architectures for accuracy among healthy and patient subjects using LOSO cross-validation.	130

List of Tables

2.1	Description of AI system autonomy levels.	15
2.2	Description of existing algorithm categories.	15
2.3	Electronic search strategy.	35
2.4	Key differences between this review and existing literature.	36
3.1	Description of exercises included in the new dataset for shoulder therapy. . .	49
3.2	Description of the new dataset in terms of samples for each ROM class category along the various gesture types.	51
3.3	Exercises description and characteristics.	54
3.4	Dataset description in terms of classes of movement and number of samples.	65
3.5	Parameters configuration of latent diffusion model and ControlNets to generated synthetic images.	67
3.6	Inception score (IS) and Detection Score for person class (DS-person) in original and generated dataset.	75
3.7	Structural similarity index (SSIM) and Fréchet Inception distance (FID) calculated between original and generated dataset over each classes of movement in each subset.	76
3.8	Structural similarity index (SSIM) and Fréchet Inception distance (FID) calculated between original and generated dataset over each classes of movement in each subset.	77
4.1	MLP model configuration in Keras library.	89
4.2	Data augmentation parameters used in Keras library.	91
4.3	Datasets characteristics in terms of number of samples for ROM classifier training.	93
4.4	Datasets characteristics in terms of number of samples for Compensation classifier training.	94

4.5	Classification results in terms of accuracy, precision, recall, and F1-score obtained for the ROM classifiers and the compensation classifiers.	95
4.6	Classification results in terms of accuracy, precision, recall, and F1-score obtained for each class of the ROM classifiers.	96
4.7	Classification results in terms of accuracy, precision, recall, and F1-score obtained for each category of the compensation classifiers.	96
5.1	Demographic frequencies in the selected IRDS subset.	105
5.2	Distribution of wheelchair disability and physical impairment side among patients in the selected IRDS subset.	105
5.3	Description of the heterogeneity introduced in training data subsets considering demographic attributes of healthy and patient subjects.	107
5.4	Description of the heterogeneity introduced in training data subsets considering disability (wheelchair usage or not) and physical impairment characteristics (impairment on the right side, impairment on the left side, bilateral impairment, and impairment involving only lower limbs).	108
5.5	Classes of correct and incorrect exercise actions used to train the activity recognition model.	114
5.6	Experimental parameters used for models training.	114
5.7	Performance metrics, including mean accuracy, precision, and recall, of models trained while considering the influence of demographic attributes.	116
5.8	Performance metrics, including mean accuracy, precision, and recall, of models trained while considering the influence of disability and physical impairment bias.	118
5.9	Description of rehabilitation activities.	121
5.10	Demographic characteristics of healthy subjects.	122
5.11	Demographic characteristics of patient subjects.	122
5.12	Mean and standard deviation of LOSO cross-validation results using traditional architecture.	128
5.13	Mean and standard deviation of LOSO cross-validation results using multiscale architecture approach.	128
5.14	Results of activity recognition algorithms trained using traditional architecture for all rehabilitation exercises.	129
5.15	Results of activity recognition algorithms trained using multiscale architecture for all rehabilitation exercises.	129

Chapter 1

Background

1.1 Context and motivation

Demographic ageing and the prevalence of noncommunicable diseases are a global trend. The European population growth is slowing down, while population ageing accelerates [53]. In the European Union, the number of people aged over 65 years old will almost double over the next 50 years, from 85 million in 2008 to 151 million in 2060 [240]. Therefore, frailty is highly prevalent among elderly people, and it is also recognized as a major public health concern by the World Health Organization. The progressive decline in physical and cognitive skills prevents elderly people from living independently and lead to the increasing demand for healthcare assistance [257]. Disability trends are putting significant pressure on age-related public expenditure in the EU, which is estimated that, by 2060, will reach 12.9% of gross domestic product (GDP) for pensions, 8.3% of GDP for health care and up to 3.4% of GDP for long-term care [78]. So there is an emerging consensus that preventing disability in older adults could reduce related healthcare services costs [317, 119]. In response to the aging of populations and the resulting shift to high prevalence rates of disability, there is a need to shift the healthcare delivery to the elderly to a more accessible and convenient services strategy in terms of efficacy, efficiency, and costs.

Home health care is an accepted strategy worldwide for a cost-effective approach to manage the growing services need for the healthcare of elderly people. However, it was also evidenced that elderly people sometimes unsuccessfully follow their home-based care plans due to issues related to the confusion about treatments, forgetfulness, loss of motivation, social isolation [166]. New ways to provide and obtain successfully long-term programs and better outcomes in the healthcare of older people should be investigated.

The recent developments in sensing, mobile, embedded devices have attracted consid-

erable attention toward innovative health applications, representing an hot research field today. Due to its support for the integration and cost-effective delivery of services, technology is recognized as a key asset in the achievement of universal health coverage for all ages. Recently, the WHO guidelines on Integrated Care for Older People (ICOPE) recognized the importance of technology in providing long-term healthcare services integrated in patient's everyday life [242]. The introduction of smart healthcare applications for elderly people offers the possibility to implement individual integrated programs especially in remote and home settings.

Technological advances, particularly in engineering and computer science, are starting to be applied to support rehabilitation research and practice [239]. In the last decade, research on the use of new technologies to improve key rehabilitation processes and outcomes has steadily grown.

Mobile systems and unobtrusive devices, coupled with specific software applications, are increasingly available to collect data and provide users with information to assess health and monitor progress toward individualized rehabilitation goals. Additionally, socially and physically assistive systems are emerging to aid individuals in recovering from illness and injury to restore function loosing caused by sensory, motor, or cognitive impairments [164].

In response to the growing demand for healthcare assistance for chronic diseases related to demographic aging and the trend of applying technology to healthcare services, this thesis investigates the latest advancements in artificial intelligence (AI) research in order to propose alternative and innovative methods to enhance long-term rehabilitation programs.

1.2 Aims and objectives

AI-powered smart healthcare applications are expected to offer the possibility of implementing efficient, individualized, and integrated programs in remote and home settings. Following this, the thesis aims to highlight the role of emerging AI solutions in supporting exercise monitoring and assistance in rehabilitation medicine.

The primary objective is to investigate AI methods and techniques that facilitate home-based rehabilitation programs delivered remotely, without the need for direct clinical supervision. Such decentralized healthcare services are particularly beneficial for the long-term management of chronic diseases in elderly care, providing essential support for patients in their home environments.

The research will investigate AI-driven methodologies for unobtrusively monitoring prescribed exercise therapy, with a particular emphasis on recent advancements in deep learning, computer vision, and pattern recognition techniques.

The overall contributions of the thesis are as follows:

1. Identify key AI methods and strategies that can support monitoring and assistance in rehabilitation medicine;
2. Analyze various methods and computational techniques for human motion analysis, while addressing fairness and equity issues in the implementation of human-centered systems in healthcare;
3. Discuss the ethical implications of fully automated systems in home-based rehabilitation services and outline the regulatory landscape regarding AI and medical devices, providing a framework to govern these technologies.

The contributions of this thesis aim to advance the development of AI-based therapeutic applications that can support and optimize home-based rehabilitation in future healthcare services. AI-powered monitoring and decision-support systems are expected to streamline clinical workflows, enhance patient safety, aid in diagnoses, and enable personalized treatment [193]. The scope is to guide the future development of healthcare applications powered by AI that facilitate the implementation of efficient, individualized rehabilitation programs in remote and home settings.

1.3 Thesis organization

The results of the research conducted throughout this project have been published in international peer-reviewed journals, contributing significantly to the body of work presented in this thesis.

The thesis is structured as follows.

- Chapter 2 introduces the broader landscape of AI innovations in healthcare, with a particular focus on the emerging applications revolutionizing the field of rehabilitation medicine. The content of the following published articles [213, 212] significantly contributed to the development of this chapter.
- Chapter 3 presents methodologies for modeling data that can be used to train and develop automated systems for motion analysis. This chapter presents the results of the following published articles [214, 210, 211].
- Chapter 4 describes the development and experimentation of a proof-of-concept system designed to monitor and evaluate exercises in an unobtrusive and automated manner.

The pipeline of the system and the results of the experiments are presented. Relevant content from the following published article [210] contributes to this section.

- Chapter 5 addresses the significant challenge of ensuring fairness in AI systems, particularly in the context of monitoring and assistive healthcare applications. This chapter provides methods and solutions to promote unbiased and equitable systems, offering insights from data and algorithmic perspective. The following published article [209] contributes to this section, while additional results presented in this chapter are currently under review for future publication.
- Chapter 6 summarizes the key contributions of the thesis and offers a comprehensive discussion of the overall work, exploring both the technical and ethical implications of the research. It also outlines potential directions for future research endeavors.
- Chapter 7 concludes the thesis by emphasizing the key contributions of this research and their implications for future applications.

Chapter 2

Introduction

2.1 The impact of AI on healthcare

2.1.1 Foundations

The concept of AI was first discussed in 1956 [205], referring to technology used to mimic human behavior. Since then, the field has made remarkable strides in development. As a subfield of AI, Machine Learning (ML) was conceptualized by Arthur Samuel in 1959 [276]. He emphasized the importance of systems that automatically learn from experience instead of being programmed. In the 1980s, ML demonstrated great potential in computer foresight and predictive analytics, including clinical practice and machine translation [30]. Deep Learning (DL), a subfield of ML, has ushered in new breakthroughs in information technology. DL may study underlying features in data from multiple processing layers using neural networks, similar to the human brain [169]. Since the 2010s, DL has garnered immense attention in many fields, especially in image recognition and speech recognition [280].

The term AI generally refers to the performance of tasks that are commonly associated with intelligent beings by software and/or devices. A specific definition of AI in a recommendation by the OECD Council on Artificial Intelligence states: ‘An artificial intelligence system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions that affect real or virtual environments’.

The basis of AI technologies are algorithms and they may operate with different levels of autonomy (see Table 2.1) [161].

Two main approaches are actually proposed to develop AI algorithms: the rule-based approach and the ML-based approach whose definitions are described in Table 2.2 [90]. These algorithms are translated into computer code that contains instructions for the rapid analysis and transformation of data into conclusions, information, or other results.

Table 2.1: Description of AI system autonomy levels.

Level	Definition	Example	Final decision
0	No presence of AI	Standard care.	Human
1	AI suggests decision to human	Clinicians consider AI recommendations but ultimately make the final decision on treatment and therapy.	Human
2	AI makes decisions, with permanent human supervision	AI makes clinical decisions on treatment and therapy, with human doctors providing ongoing supervision.	Human
3	AI makes decisions, with no continuous human supervision but human backup available	AI autonomously makes clinical decisions regarding treatment and therapy, but it can alert human users in case of uncertainty, minimizing the need for constant supervision.	AI
4	AI makes decisions, with no human backup available	AI autonomously govern the clinical decision with no human backup.	AI

Table 2.2: Description of existing algorithm categories.

Category	Definition	Example
Rule-based algorithm [162]	The system follows a set of rules predefined by experts.	The expert defines the knowledge representation of a phenomenon and integrates this model into the system.
ML-based algorithm [185]	The AI system facilitates the incorporation of intricate knowledge representations using statistics and probability theory.	The focus is not on defining a prior knowledge model, but rather on the collection of data and their integration into a training set. This approach allows for the continuous development of knowledge in a specific application domain through the ongoing utilization of data.

The huge amounts of data and the ability to analyze data rapidly feed AI to perform increasingly complex tasks [164]. For this reason, AI in medicine raises the idea that AI replaces doctors and human decision-making. However, applications of AI are still relatively new and AI is not yet routinely used in clinical decision-making. Few of these systems have been evaluated in clinical studies [212].

The use of AI in clinical care is expected to bring about the major changes schematically

reported in Figure 2.1. There are four main trends recognized by the WHO: i) the evolution of the role of the patient in the clinical care process; ii) the shift from hospital to community-based care; iii) the use of AI to provide clinical care outside the formal healthcare system; iv) the use of AI for resource allocation and prioritization [270]. Each of these trends has ethical implications, which will be discussed later.

Patients already take significant responsibility for their care, including taking medication, improving their diet and nutrition, engaging in physical activity, treating wounds, or administering injections. However, AI could further change the way patients independently manage their medical conditions, particularly in chronic disease conditions. AI could help in self-care, for instance through conversational agents such as ‘chatbots’, health monitoring tools and risk prediction technologies for prevention programs [243]. While the move to patient-based care may be seen as empowering and beneficial for some patients, others may find the added responsibility stressful and may limit an individual’s access to formal healthcare services.

Telemedicine is part of a broad revolution marking the transition from hospital to home care, and the use of AI technologies is helping to accelerate this journey. The shift to home care has been partly facilitated by the increase in the use of search engines (which rely on algorithms) for medical information, as well as the growth in the number of text or voice chatbots for healthcare [283]. In addition, AI technologies can play a more active role in managing patients’ health outside clinical settings, such as in ‘just-in-time adaptive interventions’. These rely on sensors to provide patients with specific interventions according to previously collected data [224]. The growth and use of wearable sensors and devices may improve the effectiveness of ‘just-in-time adaptive interventions’, but also raise concerns in light of the amount of data these technologies are collecting, how they are being used, and the burden these technologies may shift to patients.

The increasing use of digital self-management applications and technologies also raises broader questions about whether these technologies should be regulated as clinical applications, thus requiring more regulatory control, or as ‘wellness applications’, requiring less regulatory control. Many digital self-management technologies are likely to fall in a ‘grey area’ between these two categories and may present a risk if they are used by patients for their own disease management or clinical care, but remain largely unregulated or could be used without prior medical advice. These concerns are exacerbated by the distribution of such applications by entities that are not part of the formal healthcare system. Indeed, AI applications in healthcare are no longer used exclusively in healthcare (or home care) systems, as AI technologies for health can easily be acquired and used by entities in the non-health system. Emerging is therefore the issue concerning the use of AI to extend ‘clinical’

care beyond the formal healthcare system.

Finally, with the trend towards self-management, the use of mobile and wearable devices driven by software capable of acquiring and processing data through sophisticated algorithms has increased [268]. Self-management systems empower individuals to take an active role in managing their health. They provide the necessary resources and tools for patients to track their health, adhere to treatment plans, and make informed decisions about their well-being. This approach fosters a proactive approach to healthcare, empowering patients to become partners in their health management. Wearable technologies include those placed in the body (artificial limbs, smart implants), on the body (insulin pump patches, electroencephalogram devices), or near the body (activity trackers, smart watches, and smart glasses). Wearable devices will create more opportunities to monitor a person's health and capture more data to predict health risks, often more efficiently and in a more timely manner. Although such monitoring of 'healthy' individuals may generate data to predict or detect health risks or improve a person's treatment when necessary, it raises concerns as it allows for near-constant surveillance and the collection of excessive data that would otherwise have to remain unknown or uncollected. Such data collection also contributes to the growing practice of 'bio-surveillance', a form of surveillance of health and other biometric data, such as facial features, fingerprints, temperature and pulse [93]. The growth of biosurveillance raises significant ethical and legal concerns, including the use of such data for medical and non-medical purposes for which explicit consent may not have been obtained or the re-use of such data for non-health purposes by a government or company, such as within the criminal justice or immigration systems. Therefore, such data should be subject to the same levels of protection and security as data collected on an individual in a formal clinical care setting.

2.1.2 Clinical applications

Digital technologies are at the forefront of transforming healthcare practices. Recent innovations hold the promise of improving preventive measures, facilitating early detection of severe illnesses, and enabling remote management of chronic conditions beyond the confines of traditional healthcare settings. These developments open up new possibilities for delivering healthcare services at any time and in any place, aligning with the era of disruptive and minimally invasive medicine.

Within the realm of digital health technologies, a diverse array of innovative healthcare tools has emerged, such as health information technologies, telemedicine applications, robotic platforms, mobile and wearable devices, and Internet of Things (IoT) networks. While these technologies may differ significantly from a technical perspective, they all share a common

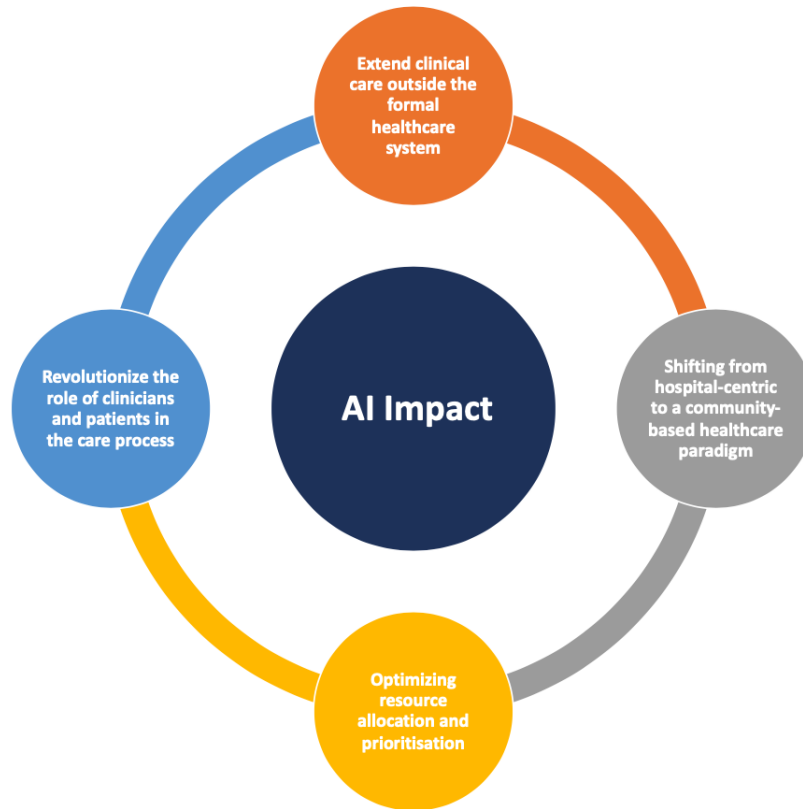


Figure 2.1: The AI impact on future healthcare system.

objective: to provide decision support in the context of healthcare practice. They accomplish this by gathering data during their use, consequently enriching the informativeness and effectiveness of medical practice through the analysis of the recorded information.

A decision support system in healthcare can be considered a computerized tool or software that assists healthcare professionals, including doctors, nurses, and administrators, in making informed and evidence-based decisions related to patient care, treatment options, and healthcare management. It integrates patient data, medical knowledge, and analytical tools to provide real-time information and recommendations, aiding healthcare providers in diagnosing conditions, creating treatment plans, and optimizing healthcare operations. Decision support systems have the potential to enhance the quality of care, reduce errors, and improve efficiency by offering valuable insights and suggestions based on the latest medical research and patient data.

In detail, AI is playing a pivotal role in revolutionizing decision support systems in healthcare, fundamentally transforming the way data is collected, analyzed, and utilized. The integration of AI introduces several significant advancements. AI algorithms can process vast amounts of patient data rapidly and with a high degree of accuracy. This allows for more

nanced and precise diagnostic and treatment recommendations. AI can identify patterns and anomalies in patient data that might be challenging for human professionals to detect. AI enables the tailoring of treatment plans to individual patients. By analyzing a patient's unique health data, genetic information, and treatment history, AI can suggest personalized therapies that are more effective and have fewer side effects. This personalized approach can lead to better outcomes and a higher quality of care. AI-driven decision support systems can use predictive data analysis to anticipate potential health issues. By continuously monitoring and analyzing patient data, AI can alert healthcare providers to early signs of disease or complications, enabling proactive intervention and preventive measures. AI can monitor patients in real-time, whether they are in a healthcare facility or at home. Wearable devices and sensors connected to AI systems can provide continuous updates on a patient's health status. This real-time monitoring is particularly valuable for chronic disease management and remote patient care. AI technology allows also to extraction of valuable information from unstructured data sources such as medical notes and reports. This capability simplifies the retrieval of relevant patient information, aiding in diagnosis and treatment planning. AI-powered care assistive systems enhance the patient experience during remote consultations. They can assist in collecting and interpreting patient data during telehealth visits, ensuring that healthcare providers have access to the information they need for informed decision-making.

To delve deeper into the discussion on AI technologies, these systems can be methodically categorized into two primary groups: diagnosis support systems, and care assistive systems.

The subsequent paragraphs offer an in-depth examination of each system category, elucidating their distinct healthcare functions as shown in Figure 2.2. A comprehensive overview of the AI technologies adopted for each category and current trends as evidenced in existing literature are reported.

Diagnosis support systems

Diagnosis support systems are designed to assist healthcare professionals in accurately diagnosing medical conditions, often by analyzing patient data, medical records, and clinical information. They aid in the formulation of precise diagnoses and the selection of appropriate treatment plans.

The use of AI in disease diagnosis and treatment has been a focus of research since the 1970s when MYCIN, developed at Stanford, was used for diagnosing blood-borne bacterial infections [42].

In various medical fields, researchers have harnessed a range of AI-based techniques to detect diseases that require early diagnosis. Figure 2.3 offers a comprehensive summary of

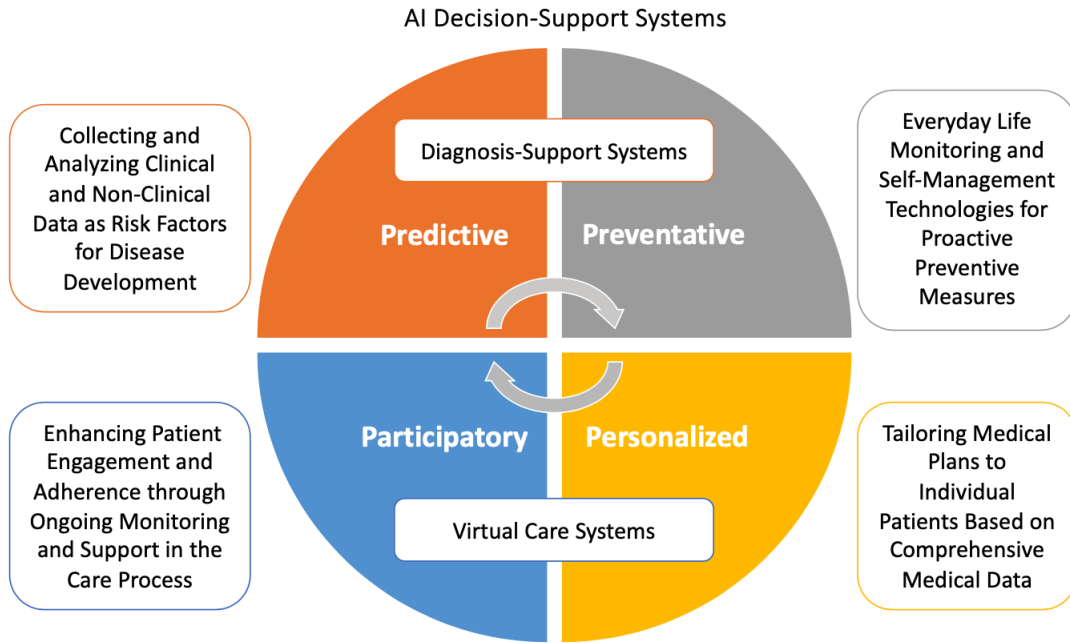


Figure 2.2: Category and characteristics of AI decision-support systems.

the distribution of medical areas of interest, considering the results of a recent systematic review [158].

AI is proving to be a valuable tool for image analysis and is increasingly employed by professionals in radiology for early disease diagnosis and the reduction of diagnostic errors in preventive medicine. AI also aids in analyzing images and signals from various diagnostic tools to support decision-making. For instance, the Ultromics platform, implemented in Oxford, utilizes AI to analyze echocardiography scans, detecting heartbeat patterns and ischemic heart disease [98]. AI has shown promising results in the early detection of diseases such as breast and skin cancer, eye diseases, and pneumonia using various imaging techniques [322, 80, 260]. More recently, a system integrated with the advanced DL model enhanced precision in identifying ductal carcinoma in breast cancer imaging, providing valuable insights for practitioners [256]. Furthermore, AI is becoming an integral part of clinical practice, aiding in diagnostic and therapeutic imaging analysis within the context of the prostate cancer pathway [304]. This integration facilitates enhanced risk stratification and enables more precisely targeted subsequent management.

Further, surprising speech pattern analysis with AI has been shown to predict psychotic occurrences and identify features of neurological diseases like Parkinson’s disease [28, 47]. In a recent study, AI-based models predicted the onset of diabetes [57]. Additionally, AI has been instrumental in assisting the public in the battle against the virus, as it has played a

crucial role in the diagnosis of COVID-19 using various imaging techniques, such as computed tomography (CT), X-rays, magnetic resonance imaging (MRI), and ultrasound (US) [108, 145, 154].

AI is significantly impacting clinical decision-making and disease diagnosis. It can process, analyze, and report large volumes of data from different sources, aiding in disease diagnosis and clinical decision-making. AI has the potential to assist physicians in making more informed clinical decisions and, in some cases, may even replace human decisions in therapeutic domains [281]. Moreover, investigations employing computer-aided diagnostics have shown remarkable sensitivity, accuracy, and specificity in detecting subtle radiographic abnormalities, contributing to advancements in public health. However, it is worth noting that the assessment of AI outcomes in imaging studies often focuses on lesion detection, potentially overlooking the biological severity and type of a lesion. This approach may lead to a skewed interpretation of AI output. Additionally, the use of non-patient-related radiological and pathological endpoints may increase sensitivity but at the cost of higher false positives and potentially overestimate diagnosis by detecting minor abnormalities that could mimic subclinical disease [237]. Furthermore, AI has found application in motion analysis, with machine learning-driven video analysis showcasing the capacity of computers to automate the identification of gait abnormalities and associated pathologies in individuals afflicted with orthopedic and neurological disorders [6, 164].

Despite significant advancements in recent years, the field of precise clinical diagnostics still faces several challenges that demand continuous improvement to effectively combat emerging illnesses and diseases. Even healthcare professionals acknowledge the obstacles that need to be addressed before illnesses can be accurately detected in collaboration with AI. Currently, doctors are not entirely reliant on AI-based approaches because they are uncertain about their ability to predict diseases and associated symptoms. Therefore, substantial efforts are needed to train AI-based systems, enhancing their accuracy in disease diagnosis. Consequently, future AI-based research should take into account the aforementioned limitations to establish a mutually beneficial relationship between AI and clinicians. Utilizing a unified model for disease diagnosis across different institutions can significantly improve accuracy, thereby aiding in the early diagnosis of diseases.

Care assistive systems

Care assistive systems encompass a diverse range of versatile systems that offer comprehensive monitoring and support capabilities across various healthcare settings. They enable real-time monitoring and provide assistance, fostering patient engagement and compliance in care. Whether the patient is physically present with the healthcare provider or in a remote

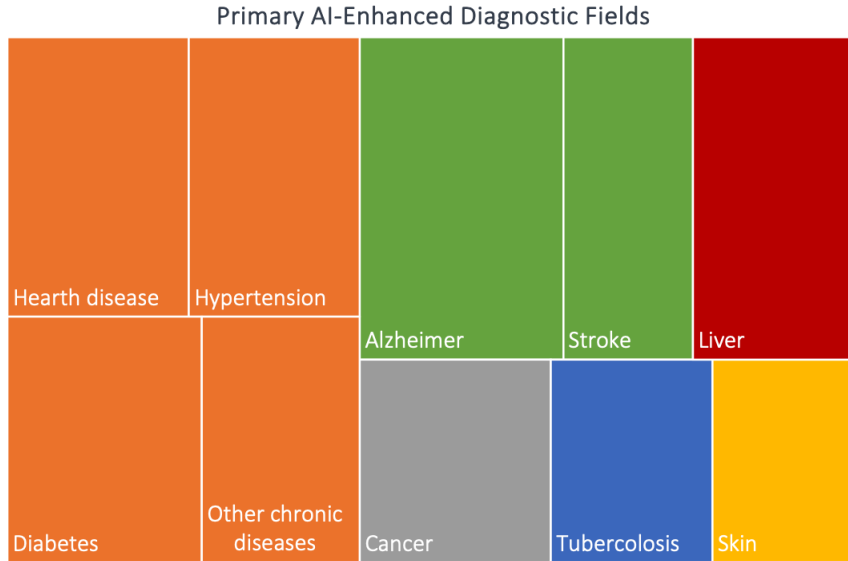


Figure 2.3: Literature evidences of diseases diagnosis using AI.

location, these systems ensure accessibility and convenience, thereby improving the overall patient experience.

Due to remarkable technological advancements, AI has ushered in innovative applications within the realm of care assistive [212]. Significant progress was made in wearable devices capable of measuring physiological changes and facilitating real-time patient monitoring [227]. Remote patient monitoring, a subset of telehealth, enables healthcare providers to remotely monitor and assess patient conditions, reducing the reliance on traditional in-person visits. This approach leverages sensors and communication technologies, simplifying the process of remotely collecting and evaluating health data, and empowering patients to take control of their health [8, 21].

Traditionally, patient monitoring systems relied heavily on clinicians' time management and invasive methods requiring skin contact. However, patient remote monitoring in healthcare now incorporates innovative IoT techniques, including contact-based sensors, wearable devices, and telehealth applications. These technologies enable the examination of vital signs and physiological variables, such as motion recognition, which supports medical decision-making and therapeutic strategies for various conditions, including psychological illnesses and movement disorders [286]. Healthcare providers have also harnessed remote patient monitoring platforms to ensure the continuity of patient care during the COVID-19 pandemic.

Conventional AI techniques are commonly employed in virtual applications to detect early signs of patient deterioration, understand patient behavior patterns through reinforcement

learning, and tailor the monitoring of patient health parameters via federated learning. AI plays a pivotal role in managing chronic diseases, including diabetes mellitus, hypertension, sleep apnea, and chronic bronchial asthma, through non-invasive, wearable sensors [146]. Smart homes equipped with sensors that monitor physiological variables such as respiratory rate, pulse rate, breathing waveform, blood pressure, and ECG can aid residents in their daily activities and alert caregivers when assistance is required. Additionally, smart mobile and wearable devices allow users to collect data and monitor progress toward personalized therapy goals [12, 189]. Inertial sensors in wearable technology can assess individuals' adherence to exercise regimens, particularly in rehabilitation programs [104, 294]. Recent reviews highlight the potential of AI integration into wearable technologies while addressing the issue of user retention, and underscore the need for patient education to improve AI acceptance [308, 295, 94]. AI-driven data processing from sensors has the potential to monitor patterns in physiological measurements, positional data, and kinematic data, offering insights into improving athletic performance. AI can enhance injury prediction models, increase the diagnostic accuracy of risk stratification systems, enable continuous patient health monitoring, and enhance the patient experience. Despite these benefits, the adoption of AI in wearable devices faces challenges such as missing data, socioeconomic bias, data security, outliers, signal noise, and the acquisition of high-quality data using wearable technology [52, 100]. Patient acceptance also presents a critical hurdle to the widespread adoption of these technologies. However, AI's transformative potential in remote patient monitoring is accompanied by challenges related to privacy, signal processing, data volume, uncertainty, imbalanced datasets, feature extraction, and explainability [286].

The integration of AI language models can further revolutionize patient care, with the potential for digital applications to assist patients in managing their treatment regimens. These virtual assistants, similar to personal healthcare advisors, can provide reminders for medication adherence and offer health status updates. The rise of virtual patient assistants using the AI natural language model exemplifies the application of AI in healthcare. These virtual assistants can guide patients in managing chronic conditions like diabetes, prescribe over-the-counter medications, and offer guidance on remote therapy sessions. Physically and socially supportive robots have proven invaluable in aiding individuals in their recovery from injuries or illnesses, bridging cognitive, motor, or sensory deficits. These technologies contribute significantly to enhancing functional abilities, independence, and overall well-being [189]. ML methods have also found application in the evaluation of patient data, clinical decision support, and diagnostic imaging. In therapy, artificial cognitive applications can assess rehabilitation sessions based on machine-generated signals [12]. The incorporation of AI-driven tools, such as the AI natural language model, into rehabilitation sessions can complement

traditional therapy, offering personalized guidance and monitoring for patients during their recovery journey [210]. Furthermore, the AI language model can assist individuals in practicing speech and language skills, making it accessible both at home and outside. Studies have demonstrated ChatGPT's ability to redraft text empathetically, enhancing peer-to-peer mental health support and various community-based self-managed therapy tasks, including cognitive behavioral therapy [287]. Furthermore, the utilization of apps and online portals for patient-physician communication has significantly improved patient engagement rates. Healthcare apps securely store and distribute patient data in the cloud, providing patients with easy access to their health information and facilitating better health outcomes. AI-based medical consultation apps allow patients to obtain information (non-emergency) and even offer medication reminders. AI language model's integration into healthcare apps streamlines time-consuming tasks like summarization, note-taking, and report generation, making healthcare more efficient. These apps also assist patients in symptom checking, appointment scheduling, medication management, patient education, and the self-management of chronic diseases [131]. Various digital platforms, including mobile apps, voice assistants, and websites, can facilitate access to care assistive. However, the utilization of virtual assistants in healthcare is not without its challenges, encompassing ethical concerns, data interpretation, privacy, security, consent, and liability issues [131].

AI is sparking a transformative wave in patient care, touching upon many healthcare domains, ranging from remote monitoring to injury prevention and virtual assistance. Through ongoing innovation and comprehensive education, AI stands poised to elevate healthcare services' quality and global reach. However, it simultaneously raises many issues regarding ethical and regulatory aspects.

2.1.3 Ethical principles and guidelines

The ethical foundation in medicine rests upon a set of fundamental principles that guide healthcare professionals in delivering compassionate and patient-centered care. Central to these principles is respect for patient autonomy, acknowledging individuals' rights to make informed decisions about their healthcare. Concurrently, the principle of beneficence underscores the healthcare provider's duty to act in the best interests of patients, promoting well-being and striving to maximize positive outcomes while minimizing harm, in adherence to the principle of non-maleficence [27].

Justice in medical ethics emphasizes the equitable distribution of healthcare resources, treatments, and opportunities, addressing disparities and ensuring universal access. Veracity and confidentiality highlight the importance of honest and transparent communication while

safeguarding patient privacy.

Fidelity, or professional faithfulness, underscores the commitment of healthcare professionals to fulfill their duties and obligations, maintaining trust within the physician-patient relationship and the broader healthcare system. These principles collectively form the ethical compass that guides decision-making in medicine, ensuring a balance between individual rights, societal equity, and the integrity of the medical profession. As the healthcare landscape evolves, these principles remain essential, fostering ethical practices that prioritize patient welfare and uphold the core values of the medical profession. Continuous reflection on these ethical considerations ensures that healthcare professionals navigate the complexities of medical practice with integrity, compassion, and an unwavering commitment to ethical standards.

When developing digital technologies for healthcare, it is essential to take into account the requirements for monitoring patient safety, privacy, traceability, accountability, and security. Furthermore, plans should be established to address any breaches that may occur.

The WHO initiated in 2019 the development of a framework to facilitate the integration of digital innovations and technology into healthcare. The WHO's guidelines for digital interventions in healthcare emphasize the importance of evaluating these technologies based on factors such as benefits, potential drawbacks, acceptability, feasibility, resource utilization, and considerations of equity. These recommendations underscore that these digital tools should be perceived as essential aids in the quest for universal health coverage and long-term sustainability[109].

The ethical principles for applying AI in the field of healthcare are designed to provide guidance to developers, users, and regulators to enhance the design and utilization of these technologies while ensuring proper oversight.

At the heart of all ethical principles lie human dignity and the intrinsic worth of every individual. These foundational values underpin the ethical guidelines that outline duties and responsibilities within the sphere of developing, implementing, and continually evaluating AI technologies for healthcare. The European Regulation enacted on April 21, 2021, categorizes AI products with precision based on their potential risk to fundamental rights such as health and safety, dignity, freedom, equality, democracy, the right to be free from discrimination, and data protection.

Given this classification, ethical principles play a pivotal role for all stakeholders engaged in the responsible advancement, deployment, and assessment of AI technologies for healthcare. This inclusive group encompasses physicians, system developers, healthcare system administrators, health authority policymakers, as well as local and national governments. Ethical principles should serve as catalysts, encouraging and aiding governments and public

sector agencies in adapting to the rapid evolution of AI technologies through legislation and regulation. Moreover, these principles should empower medical professionals to judiciously employ AI technologies in their practice.

Within a general framework of the use of AI techniques in the service of society, six fundamental principles in favor of the ethical development of such technologies have been identified in the literature. Some of them are fundamental principles commonly used in bioethics: beneficence and non-maleficence (i.e. to do no harm and minimize the benefit/risk trade-off), autonomy (respecting the individual's interest in making decisions), justice (ensuring fairness and that no person or group is subjected to discrimination, neglect, manipulation, domination or abuse). Other principles, on the other hand, principles draw from moral and legal standards. These principles emphasize maintaining the epistemological aspect of intelligibility, encompassing both the need for clear explanations of technology's operations and the responsibility to trace cause-and-effect relationships resulting from technology's actions. Furthermore, they underscore the importance of safeguarding and upholding individual privacy to empower individuals to retain control over sensitive information concerning themselves, thereby preserving their capacity for self-determination and, in turn, respecting their autonomy [5].

In a recent WHO initiative conducted in June 2021, a comprehensive set of indications, recommendations, and guidelines about the development, application, and utilization of AI technologies in medicine was established [270]. The WHO work offered a detailed exploration of fundamental ethical principles designed to guide the development and implementation of AI technologies. Strongly endorsing the adoption of this updated document within the AI domain, the subsequent paragraphs provide a comprehensive review of the key ethical guidelines delineated by WHO for a more thorough examination of the topic. Figure 2.4 provides a schematic summary of recommended ethical principles and guidelines for AI in healthcare.

Protection of autonomy. The integration of AI may lead to scenarios where decision-making is either transferred to or shared with machines. Upholding the principle of autonomy necessitates that any expansion of machine autonomy should not compromise human autonomy. In the context of healthcare, this implies that individuals should retain complete control over healthcare systems and medical choices. AI systems should be meticulously and consistently designed to align with established principles and human rights, specifically focusing on their role in aiding individuals, whether they are healthcare professionals or patients, in making well-informed decisions. Respecting autonomy also encompasses the accompanying responsibilities of safeguarding privacy, maintaining confidentiality, and ensuring informed

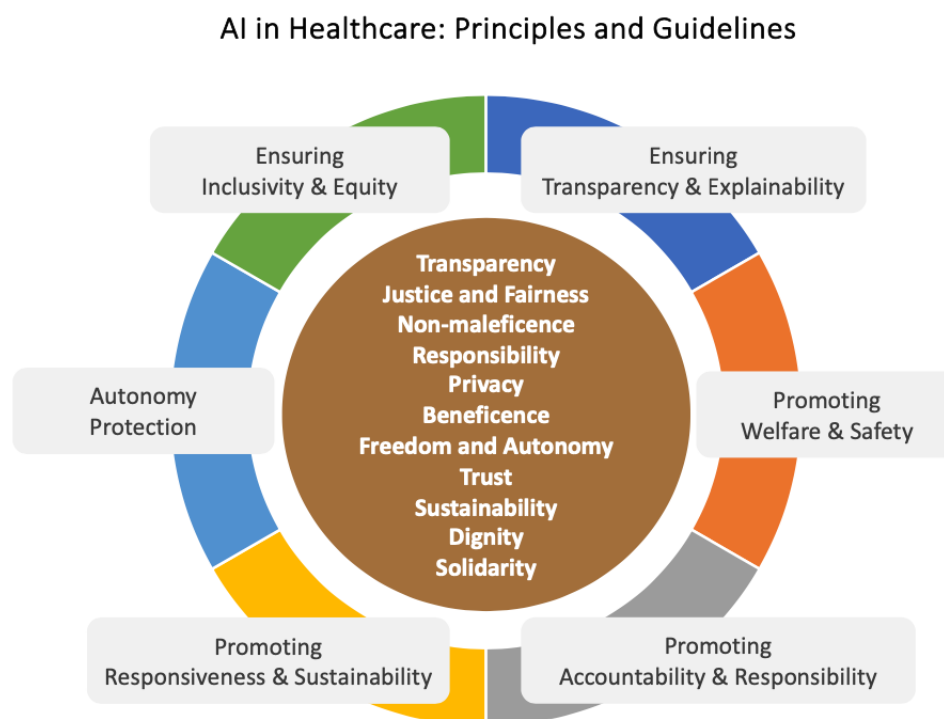


Figure 2.4: Ethical principles and guidelines.

and valid consent through the implementation of appropriate legal frameworks for data protection.

Promoting welfare, safety and public interest. AI technologies must adhere to strict regulatory standards concerning safety, accuracy, and effectiveness before they are made available to the public. These requirements are essential to ensure not only initial quality control but also to promote continuous improvement. As a result, individuals and entities involved in funding, development, and utilization of AI technologies carry an ongoing duty to evaluate and oversee the performance of AI algorithms to confirm their intended functionality.

Ensure transparency, explainability and intelligibility. AI should be comprehensible to developers, users, and regulators. Achieving transparency necessitates the provision of adequate information, which should be documented or made available before an AI technology is designed and implemented. This commitment to transparency not only enhances the overall system quality but also serves as a protective measure for patient safety and public health. For instance, system evaluators rely on transparency to identify and rectify errors, and government regulators depend on it to carry out effective oversight.

AI technologies should be as explainable as possible, and tailored to the comprehension

levels of their intended audience. Striking a balance between full algorithmic explainability (even at the cost of some accuracy) and enhanced accuracy (possibly at the expense of explainability) is a critical consideration.

Promoting accountability and responsibility. Accountability can be established through the application of ‘human assurance’, which entails the evaluation of AI technologies by both patients and physicians during their development and implementation. In the context of human assurance, regulatory principles are employed both upstream and downstream of the algorithm, creating human supervision points. The primary aim is to ensure that the algorithm maintains its medical effectiveness, remains open to evaluation, and upholds ethical accountability. Consequently, the utilization of AI technologies in the field of medicine necessitates accountability within intricate systems where responsibility is distributed among various stakeholders.

In cases where AI technologies make medical decisions that result in harm to individuals, accountability processes should unequivocally identify the respective roles of producers and clinical users in the harm incurred. To avert the diffusion of responsibility, where ‘everyone’s problem becomes nobody’s responsibility’, a robust accountability model, often referred to as ‘collective responsibility’, holds all parties involved in the development and deployment of AI technologies accountable. This approach encourages all stakeholders to act with integrity and minimize harm. It is worth noting that this remains a continually evolving challenge and is yet to be fully addressed in the laws of many countries.

To ensure appropriate redress for individuals and groups adversely affected by decisions made by algorithm-based systems, mechanisms for compensation must be put in place. This should encompass access to prompt and effective remedies and redress facilitated by both government bodies and companies employing AI technologies in the healthcare sector.

Ensuring inclusivity and equity. Inclusivity dictates that AI employed in healthcare should be intentionally designed to promote the broadest possible and equitable utilization and access, irrespective of factors such as age, gender, income, ability, or other distinguishing characteristics. AI technologies should not solely cater to the requirements and usage patterns prevalent in high-income settings; they must also be adaptable to accommodate various devices, telecommunications infrastructures, and data transfer capabilities, particularly in less economically privileged environments. Both industry and governments are responsible for bridging the ‘digital divide’ within and between countries to ensure that new AI technologies are accessible on an equitable basis.

AI developers must ensure that AI data, especially training data, are free from sampling

bias and characterized by accuracy, comprehensiveness, and diversity. Special provisions should be in place to safeguard the rights and well-being of vulnerable populations, coupled with mechanisms for redress in cases where biases and discrimination emerge or are alleged.

Promoting responsiveness and sustainability. AI responsiveness necessitates that designers, developers, and users engage in a continuous, systematic, and transparent assessment of AI technology to ensure that it functions effectively, and appropriately, in accordance with the expectations and requirements of the specific context in which it is deployed. When an AI technology proves to be ineffective or causes dissatisfaction, the obligation to be responsive involves instituting a structured process to resolve the issue, which may include discontinuing the technology’s use. Therefore, AI technologies should only be introduced if they can be seamlessly integrated into the healthcare system and receive adequate support. Regrettably, in under-resourced healthcare systems, new technologies are frequently underutilized, left unrepaired, or not upgraded, which squanders precious resources that could have been invested in other beneficial interventions.

Sustainability also hinges on the proactive response of governments and companies to anticipate workplace disruptions. This includes providing training for healthcare professionals to adapt to the integration of AI and addressing potential job losses due to the adoption of automated systems for routine health functions and administrative tasks.

2.1.4 Legislative measures

It is imperative that AI models maintain simplicity in their properties and functions to ensure ease of operation by healthcare providers [38]. Nevertheless, several hurdles hinder the widespread adoption of AI in healthcare. These challenges encompass capacity limitations in developing and maintaining infrastructure to support AI processes, elevated costs associated with data storage and backup for research purposes, and the substantial expenses required to enhance data reliability [298]. AI algorithms, while powerful, are not without their limitations, including limited applicability beyond the training domain, and susceptibility to bias [198]. To address these challenges, healthcare stakeholders should develop and execute a carefully planned strategy for AI implementation in healthcare to handle the cost, technological infrastructure, and AI system integration into clinical workflows. Furthermore, clinicians frequently experience mistrust and lack of understanding concerning AI-based clinical decision support systems, primarily due to undisclosed risks and the reliability of such systems [58]. This skepticism acts as a substantial obstacle to broad adoption. To address this, there is a growing focus on implementing explainable AI solutions to boost user trust and navigate these issues [103].

All these pieces of evidence emphasize that the increasing integration of AI technologies into healthcare necessitates effective governance to address regulatory, ethical, and trust-related concerns [262, 243]. A recent study also underscores the critical role of governing AI technologies at the healthcare system level in ensuring patient safety, healthcare system accountability, bolstering clinician confidence, enhancing acceptance, and delivering substantial healthcare benefits [201].

Maintaining control over regulated domains, especially healthcare, underscores the urgency of implementing national and international regulations. These regulations are essential for the responsible integration of AI-driven applications in healthcare while upholding the tenets of medical ethics [125]. This section concentrates on the proposal and refinement of five essential European acts, which together can form a unified regulatory framework for governing AI in healthcare (see Figure 2.5).

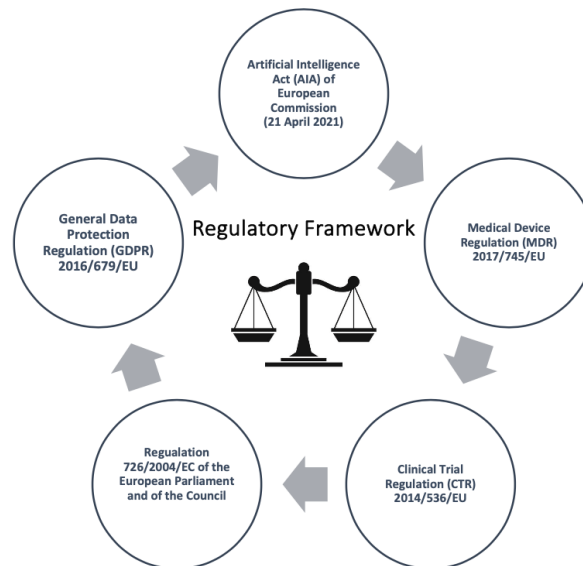


Figure 2.5: Evidencing key European acts to govern AI in healthcare.

In this endeavor, the European Union (EU) has already initiated actions by enacting the General Data Protection Regulation (GDPR) in 2018. GDPR is specifically crafted to protect personal data handled by data processors or controllers operating within the EU, setting a precedent for substantial regulatory changes in the United States and Canada [91].

Moreover, the European Commission has introduced the recent Artificial Intelligence Act (AIA), which is designed to address a range of risks linked to the extensive implementation of AI [3, 2]. This set of regulations advocates for the responsible deployment of AI and endeavors to prevent or alleviate potential harms arising from specific technology applica-

tions. According to this proposed act, high-risk AI systems are obligated to undergo pre-deployment compliance assessments and post-market monitoring to ensure their adherence to all the requirements outlined by the AIA [279].

Furthermore, the European Union has recently implemented the Medical Device Regulation (MDR) 2017/745/EU [4], which replaces both the Medical Device Directive (MDD) 93/42/EC [1], and the Active Implantable Medical Device Directive (AIMDD) 90/385/EC [112] which aims to enhance the regulation of the medical device market .

In general, the implementation of the MDR does not remove any requirements from the replaced regulatory acts. Instead, it introduces new ones, emphasizing a life-cycle approach to device safety. To summarize, the primary changes brought about by the MDR's entry into force and application for medical devices are as follows:

- i) the MDR enhances controls to ensure the safety and effectiveness of devices.

- ii) the mechanism allowing for the acceleration of placing devices on the market or putting them into service through equivalence to existing devices is no longer applicable for all medical devices.

- iii) post-marketing clinical follow-up is extended to all medical devices, leading to an increased importance and number of clinical evaluations and investigations.

According to Article 51(1) of the MDR, medical devices are classified into four main classes: I, IIa, IIb, and III. This classification depends on their intended purpose and inherent risks, based on the criteria specified in Annex VIII. In general, class I includes most non-invasive and non-active devices, representing the lowest risk. Class II(a) devices are of medium risk, while class II(b) devices are medium to high risk. Class III is reserved for high-risk devices. Regarding software, it is classified as class II(a) when it is intended to provide information for diagnostic or therapeutic decisions, unless these decisions could result in death or irreversible deterioration of a patient's health (in which case it is classified as class III), or if they lead to serious health deterioration or surgery (in which case it is classified as class II(b)). Furthermore, if the software is designed for monitoring physiological processes, it is classified as Class II(a), unless it is meant for monitoring vital physiological parameters, where changes in these parameters could pose an immediate danger to the patient (in which case it is classified as Class IIb).

All other software falls under class I (as per rule 11 of Annex VIII). However, it is essential to note that, concerning AI-based medical devices, class I risk classification is typically not applicable. From this point of view, however, it should be considered that, as far as AI-based

MD is concerned, it is not plausible that risk class I could be applicable [199]. Nonetheless, AI technologies are medical devices in the light of their definition under the MDR. However, this conclusion requires further clarification. Specifically, the choice of applicable law hinges on whether a medical product incorporated into an AI device is considered ‘ancillary’, leading to the application of the MDR, or ‘non-accessory’, which triggers the application of laws related to medicines for human use [215]. In this context, one could argue that, according to Article 1 of the MDR, if the medicinal product incorporated into the device has an ancillary role concerning the device, it falls under the evaluation and authorization process defined by the MDR. However, if it serves a principal (i.e., non-accessory) function, the comprehensive product will be regulated by Directive 2001/83/EC of the European Parliament and of the Council, which pertains to medicinal products for human use, or by Regulation 726/2004/EC, which governs Community procedures for authorizing and overseeing medicinal products for human and veterinary use, including the establishment of the European Medicine Agency (EMA) when applicable.

Furthermore, when discussing accountability and transparency in AI, the Clinical Trial Regulation (CTR) offers relevant suggestions for regulation.

Of particular significance are the regulations surrounding two key aspects of the product lifecycle, particularly during the pre-market phase: the responsibilities of sponsors and investigators. A sponsor of a clinical study or experiment can be an individual, a company, an institution, or an organization that assumes the responsibility for initiating, managing, and financing the said investigation or experiment. On the other hand, an investigator in a clinical study or experiment is an individual responsible for conducting it at a clinical investigation site. Under the MDR, the sponsor, whether acting alone or in conjunction with the investigator, is obliged to adhere to various obligations concerning the execution of clinical investigations. These obligations, outlined in Article 11 of the MDR and Article 25 of the AIA, encompass the following tasks:

- i)** maintaining a copy of the EU Declaration of Conformity and technical documentation available for review by the national competent authority in the context of market surveillance;
- ii)** furnishing the national competent authority, upon request, with all the necessary information and documentation to demonstrate the conformity of a high-risk AI system;
- iii)** collaborating with the national competent authority, upon request, on any actions about the AI system.

Finally, the accountability process for regulating AI products may be extended to users,

which encompasses health professionals, laypersons, legal entities, public authorities, agencies, and other organizations that utilize medical devices under their jurisdiction.

Article 29 AIA lays down a series of responsibilities for users employing high-risk AI systems. These responsibilities include adherence to the provided instructions for use, ensuring the data inputted aligns with the system's intended purpose, diligent monitoring of the system's operation as per the instructions, immediate notification of the developer or distributor, and discontinuation of the AI medical device use if there are concerns about its application not complying with the instructions for use. Users are also expected to report identified serious risks to the sponsor and investigator, utilize relevant information for fulfilling their obligations, and comply with data protection regulations specified in Article 35 of Regulation (EU) 2016/679 or Article 27 of Directive (EU) 2016/680, where applicable.

2.2 The role of AI in rehabilitation

2.2.1 New trends in rehabilitation

The current model of care is resource-demanding and already faces major challenges in coping with an increasing number of patients due to demographic change, the paucity of skilled healthcare professionals and economic pressure to minimize healthcare costs [63].

Traditional rehabilitation services strongly rely on physical therapy sessions, which are based on one-to-one interactions with healthcare practitioners during a hospital stay (above all during the acute and sub-acute phase of the disease) or as part of periodic visits to specialized clinics (mostly during the chronic phase). Despite the clinical-centered model allowing healthcare professionals to closely monitor and support patients and their families physically and emotionally, it presents some limitations. A clinical-centered model that heavily relies on physical access to medical facilities and prolonged interactions with trained specialists makes events such as the coronavirus pandemic (COVID-19) and the increased demand for assistance to chronic patients challenging to manage [138, 45]. Distancing measures necessary during a pandemic phase and the increasing management issues avoiding the overload of the healthcare system could lead to a reduction in rehabilitation visits. Moreover, the lack of human resources could affect the dose and quality of therapy patients receive, limiting rehabilitation to relatively short periods that might be insufficient to achieve functional recovery [311, 231].

Despite the growing evidence that intensive high-dose rehabilitation positively impacts functional recovery even long after the disease, therapy dose is typically rather low at all stages of the continuum of care. In fact, after clinic discharge, there is limited support

for ongoing rehabilitation management at home, which is necessary for the maintenance of functional outcomes in physical activities [66, 325]. To this end, there is a need for a sustainable approach that facilitates patients' long-term management more efficiently [204, 121, 203]. Many efforts are needed to shift care services toward a decentralized model opposite the conventional therapy approach.

New developments are emerging as possible answers to rehabilitation service limitations related to time, distance, difficult terrains, costs, and limited access to clinical facilities [328]. Some proposed solutions span from a live communication stream in which rehabilitation professionals guide and encourage patients while monitoring their progress to more advanced solutions involving various sensor and information and communication technologies (ICTs), sometimes coupled with AI algorithms to support the monitoring and assistance in remote contexts [15, 310].

AI applications are considered to play a key role in further establishing and supporting a decentralized rehabilitation model in which intelligent connected tools will be employed to assist clinical decision-making, and health outcomes monitoring [140, 226]. Many AI-based methods and solutions have been proposed in recent years to support the future challenge of enabling assisted physical therapy and assessments in a minimally supervised and decentralized manner, ideally at the patient's home. Some existing works in literature have provided an overview of the role of machine learning algorithms combined with specific technologies used for rehabilitation issues, such as wearable sensors [67, 34] and vision-based motion capture technologies [197]. More specifically, a recent work emphasized machine learning methods for movement evaluation in rehabilitation programs using motion capture systems [177]. Other works highlighted the progress in machine learning for automated evaluation of patient performance and recovery, introducing data analytics to improve the effectiveness and efficiency of care for physical rehabilitation [137].

However, no prior studies have offered a comprehensive review of machine learning methods and applications for remote monitoring and assistance in rehabilitation that could support the advancement of this dissertation. To fill this gap, Section 2.2 of this thesis presents a detailed and extensive analysis of existing literature on the topic. The analysis of current trend in literature provides an in-depth exploration of the evolving landscape of AI-driven solutions, serving as a valuable foundation to guide the research endeavors undertaken during this doctoral study.

2.2.2 Analysis of the state-of-the-art

This section offers a comprehensive overview of recent developments in AI-based methodologies and emerging technologies relevant to rehabilitation.

In line with the thesis objective, a systematic analysis of current research efforts provides essential guidance for the implementation and evaluation of AI-driven solutions for remote and decentralized rehabilitation services.

A literature search was conducted in PubMed, IEEE Xplore, and Scopus databases by covering the period from 2010 to May 2022. A structured search strategy was performed in each electronic database concerning AI-based machine learning methods and applications supporting remote monitoring and assistance in the rehabilitation context. The same search string was used for PubMed, IEEE Xplore, and Scopus, with the only difference due to the syntax required by the database. Performed queries are shown in Table 2.3.

Table 2.3: Electronic search strategy.

Database	Query
PubMed	(“artificial intelligence” OR “deep learning” OR “machine learning”) AND ((tele OR remote OR home) AND (evaluation OR monitoring OR supervision OR assessment OR interaction OR feedback OR communication)) AND (“physical therapy” OR exercise OR “physical rehabilitation”)
IEEE Xplore	(artificial intelligence OR deep learning OR machine learning) AND (tele OR remote OR home) AND (evaluation OR monitoring OR supervision OR assessment OR interaction OR feedback OR communication) AND (physical therapy OR exercise OR physical rehabilitation)
Scopus	TITLE-ABS-KEY ((“artificial intelligence” OR “deep learning” OR “machine learning”) AND ((tele OR remote OR home) AND (evaluation OR monitoring OR supervision OR assessment OR interaction OR feedback OR communication)) AND (“physical therapy” OR exercise OR “physical rehabilitation”)) AND (LIMIT-TO (LANGUAGE , “English”))

Most of the current literature addressing AI technologies and methods in rehabilitation is focused on specific applications or a specific area of research. In this thesis is presented a broader overview of different research areas and applications.

The main contributions of the systematic literature review are as follows:

- AI-based systems and technologies for monitoring and assistance in rehabilitation are identified and discussed along with their practical implications;

- machine learning algorithms that find their application in rehabilitation are classified and discussed according to the state-of-art literature;
- machine learning methods and approaches are outlined and discussed according to the clinical application scopes from a practical point of view;

Table 2.4 highlights the major differences with other review articles about AI and rehabilitation topics.

Table 2.4: Key differences between this review and existing literature.

Reference	Main differences
[34]	The objective of the study was to survey exercise evaluation in post-stroke rehabilitation by means of wearable devices and machine learning algorithms.
[197]	The study reviewed the evolution of RGB-D sensors for musculoskeletal health monitoring, also evidencing computational algorithms based on machine learning.
[177]	The study reviewed computational approaches for evaluating patient performance in rehabilitation programs using motion capture systems.
[137]	The study provided a non-systematic brief overview of big data analysis to optimize outpatient rehabilitation.

The entire review process used to analyze the state of the art is schematically presented step by step in Figure 2.6.

The overall findings from the literature are summarized in Figure 2.7. This figure illustrates the distribution of published works over the years and provides a qualitative assessment of these publications based on a quality quartile metric, which evaluates the impact of the journals. A preliminary summary of emerging technologies and applied methodologies supporting AI-driven rehabilitation services for remote therapy delivery is provided in Figure 2.8.

In the next sections, a more detailed presentation and analysis of state-of-the-art technologies and applications are presented and analyzed. Specifically, Section 2.2.3 presents an overview of relevant technologies, Section 2.2.4 reviews the algorithms used, and Section 2.2.5 explores the clinical scopes of AI technologies.

2.2.3 AI-based systems and technologies

The remote delivery of rehabilitation services should enable therapists to optimize the timing, intensity, and duration of therapy which is often not possible due to the constraints of face-to-face treatment protocols of the traditional healthcare model. Therefore, ICTs used in

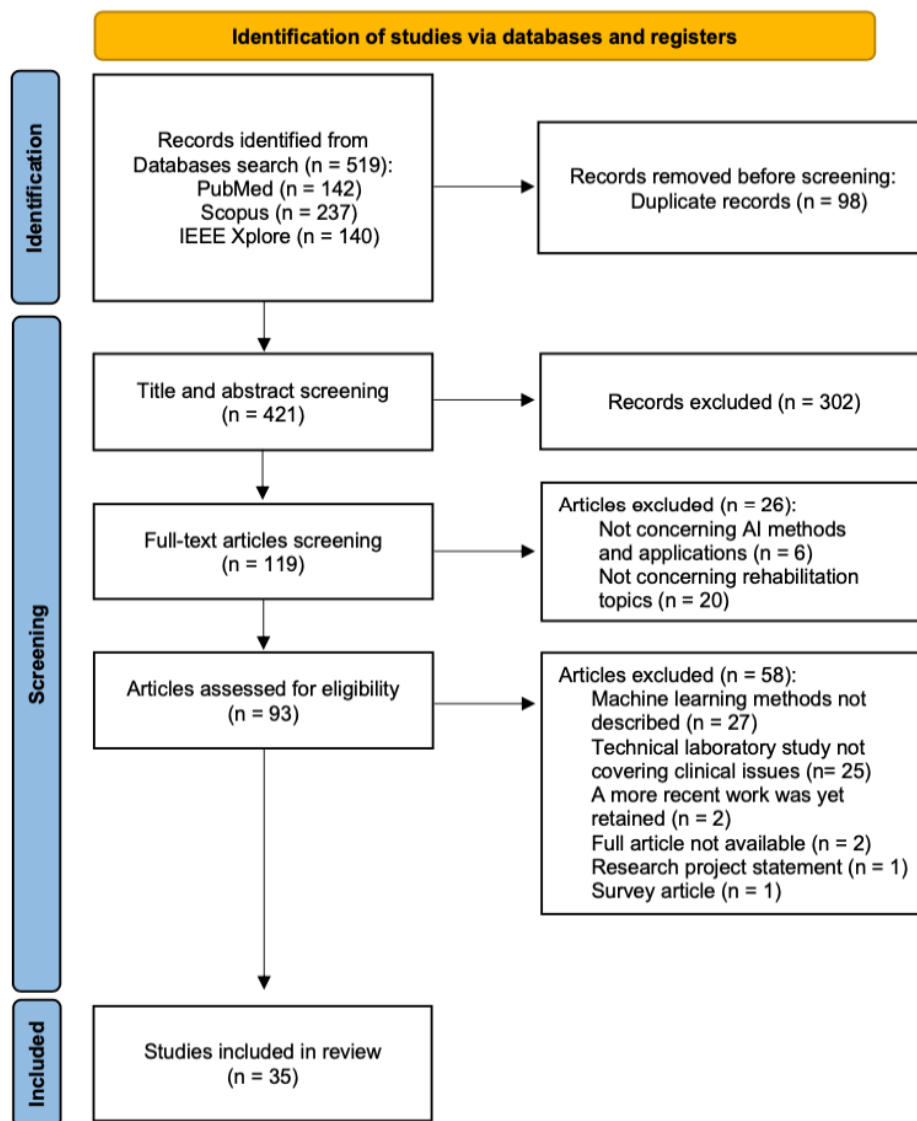


Figure 2.6: PRISMA flowchart of the results from the literature search.

decentralized rehabilitation services allow facing traditional logistic problems making treatments access equitable to geographically remote individuals.

Applications to date used to support decentralized services can be categorized as: i) synchronous systems using real-time electronic communication *via* simple webinar platforms through which healthcare professionals provide audio and video support for the patient, and ii) asynchronous systems using store-and-forward communication involving data and collection through distributed devices [303, 159]. Healthcare professionals typically use synchronous systems to interact with patients on a live communication stream, guiding and encouraging patients by monitoring their progress. Asynchronous approaches, in general, in-

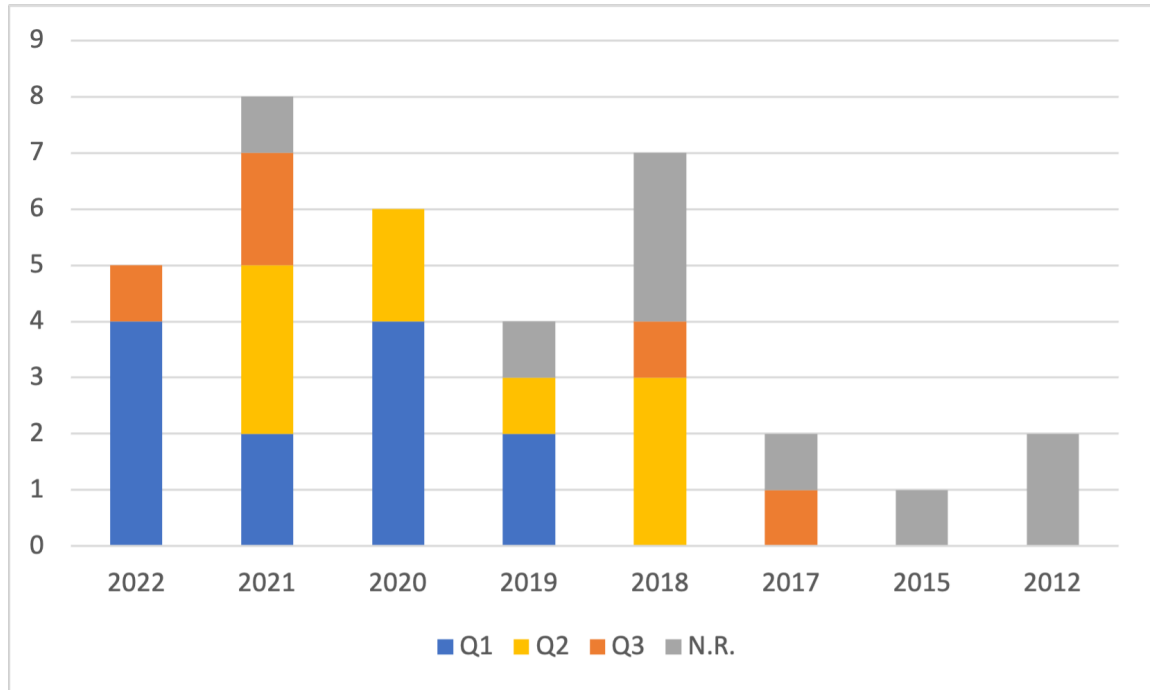


Figure 2.7: Distribution of the number of articles published over years.

Note. The quality quartile ranking according to Scimago Journal & Country Rank (SJR) database is reported.

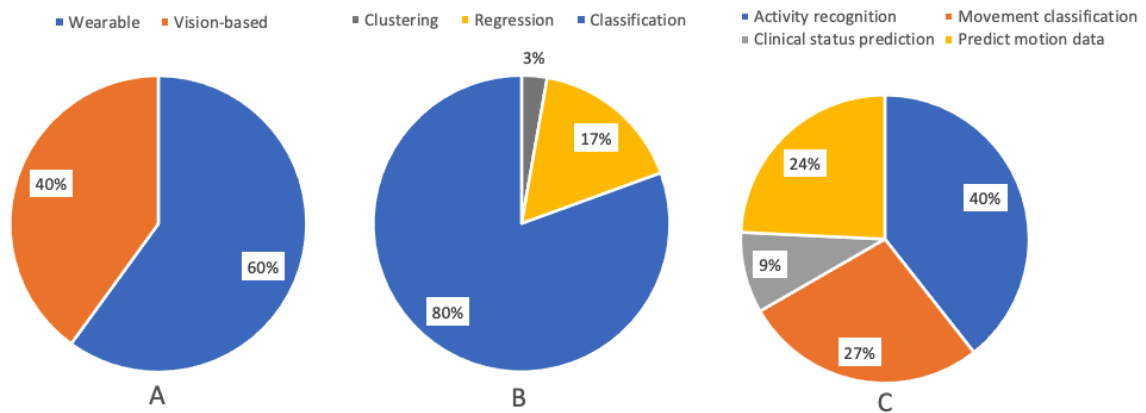


Figure 2.8: Results synthesis.

Note. A) capture motion technologies, B) machine learning tasks, C) machine learning applications.

volve the acquisition, transmission, and elaboration of data to generate an appropriate action in patient care with associated decision support [310]. Such approaches have attracted recent researchers' interest, as evidenced by the results of reviewed articles. Some attempts have been made to provide real-time audio and video suggestions with asynchronous technologies.

Retrieved systems have been used to guide and assist motion while collecting objective data on movement quality and encouraging remote rehabilitation sessions *via* digital therapeutic approaches, providing access to quality therapy from a distance without needing constant supervision by a therapist on site. Different digital strategies have been individualized: exergame solutions that engage the patient to perform some functional movements providing relevant video and audio feedback during a serious game session [96, 319, 192, 253, 320, 222]; digital coach solutions that, *via* speech [32, 326, 69], sometimes involving social humanoid robot to encourage activity participation [200], or app-based text notifications [82], offer automated conversational interaction to replace some human care tasks (reminders and motivational messages for medication, nutrition, and exercise, routine condition checks and health maintenance based on personal monitoring data).

Over the past decade, significant efforts focused on developing unobtrusive, effective, and objective motion-modeling systems have been made, taking advantage of the progress made in sensor technology which has become more compact and more efficient [195]. In almost all the included works, motion capture technologies based on wearable sensors and optical sensor-based systems have been mainly individualized as useful systems to support the remote monitoring of physical rehabilitation tasks. Such technologies allow for providing crucial information about the patient's movement during physical exercise sessions by analyzing significant kinetic (force, torque, moments, *etc.*) and kinematic (angle, orientation, velocity, *etc.*) features collected in an unobtrusively way [208].

Wearable systems are based on compact, lightweight sensors directly mounted on the interesting body part to analyze, sometimes resulting in embedded smart textile systems such as garments and gloves [17, 278, 288] or wrist-worn devices such as smart bands and smartwatches [39, 48, 222, 40, 165, 19], or as the hardware part of immersive virtual reality systems (haptic interfaces and headset) [253]. The Inertial Measurement Unit (IMU) was the most common motion capture technology thanks to its low cost and wearability. IMUs combine linear acceleration from the accelerometer and the angular turning rates from gyroscopes resulting in accurate motion data. Moreover, IMUs coupled with different sensors such as optical fiber sensors [17] and flex and force resistive sensors [278] are useful to acquire more information about the amount of deflection or bending and the force applied during a rehabilitation task. Acceleration signals were also used and coupled to electrocardiogram (ECG) data in order to better recognize the patient's physical activity form [217].

Recently, the advancements in optical sensor-based motion capture technologies attracted researchers to perform unobtrusive motion analysis. The vision-based sensor technology uses the contactless approach for motion capture; at the same time, it provides reliable movement tracking without influencing its naturalness, extraction of kinetic and kinematic parameters,

and accuracy in non-controlled environments [220]. The Microsoft Kinect camera was the vision-based system majorly used in motion tracking. Meanwhile, the Leap Motion system was more specifically applied in the field of hand rehabilitation [149].

Besides all these types of motion capture technologies, the recent developments in the computer vision field extended the use of low-cost RGB and depth sensors to perform motion analysis by means of human pose estimation algorithms based on deep learning frameworks [32, 130, 320]. Human pose estimation is a field of computer vision that aims to predict the poses of human bodies by extracting joints from images and videos for motion analysis [70]. Contrarily to wearable sensors, AI-based human motion modeling enables commercial systems equipped with a camera and low-cost hardware, such as tablets and smartphones, to perform inexpensive and unobtrusive home-based monitoring in patients' daily life [32, 315, 133].

2.2.4 Algorithms

Machine learning is an artificial intelligence application that allows systems to automatically learn and improve from experience without being explicitly programmed to do so using previously selected features. Depending on scopes and tasks, machine learning algorithms can be divided into unsupervised and supervised learning. Unsupervised learning is well known for feature extraction, while supervised learning is suitable for predictive modeling by building relationships between features and the target of interest [135]. Retrieved machine learning algorithms are schematically presented in Figure 2.9.

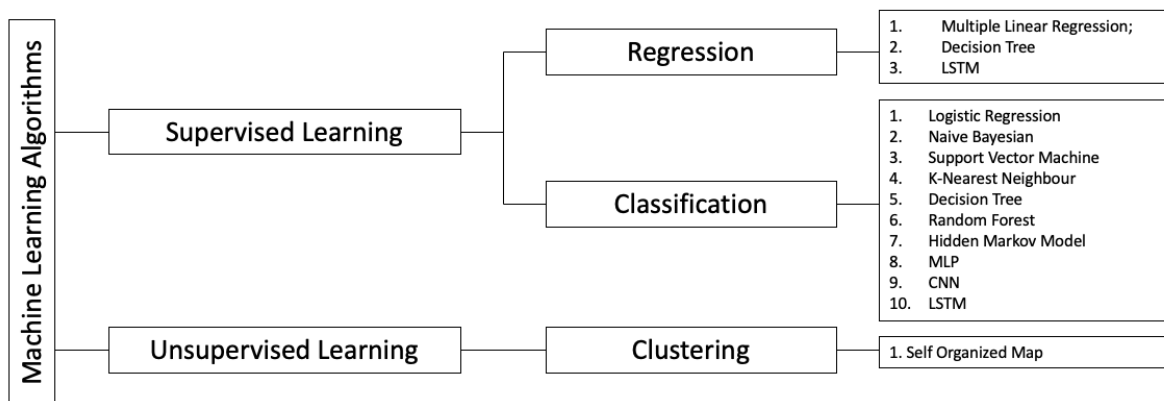


Figure 2.9: Categorization of machine learning algorithms.

Note. LSTM: long short-term memory, MLP: multilayer perceptron, CNN: convolution neural network.

In only one study that emerged, an unsupervised machine learning technique known as

Self Organized Map (SOM) was used to individualize some clusters according to patients' demographic and clinical characteristics [96]. SOMs are artificial neural networks that generate a feature map to produce a low-dimensional and discretized representation of the input training samples for clustering, visualization, and classification scopes [216].

Almost all the articles involved supervised machine-learning techniques for classification tasks. Regression analysis modeling relationships between dependent and independent variables involved linear regression algorithms in analyzing and learning from the existing training data. Linear regression algorithms have been applied in fewer works aiming, for example, to reduce the system complexity of an exergame platform by estimating ground reaction forces from kinematic features [319], or to solve self-occlusion errors in motion analysis mapping the noisy Kinect-based measurements to a more accurate benchmark system measurements [51].

Random Forest (RF) and, more specifically, Decision Trees (DT) algorithms were the most used classifiers. DT algorithm was also used for a regression problem concerning the feasibility of using motion tracking data to predict joint angles and torques during an immersive virtual exergame [253]. DT is one of the oldest machine learning algorithms; it bases its decision logic on a tree-like architecture [151]. Its ease of interpreting and its rapid learning speed made it popular to use in the healthcare domain, especially in multi-class activity recognition problems in rehabilitation. The reason is that when examining the tree for a classification sample, the results of each node will provide relevant information to be inferred about its class. RFs are sets of randomized DTs combined using bagging techniques to reduce overfitting problems when the dataset is relatively large [247]. An RF model is particularly useful for categorizing classes of patients according to their clinical assessment characteristics [327]. Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) were the other traditional supervised machine learning algorithms most used in selected studies, followed by logistic regression classifiers, Naive Bayes (NB) algorithm, and the Hidden Markov Model (HMM) algorithm. HMM is a Markov chain process that finds application in posture recognition and characterization of skeletal tracking spatial-temporal data as a parametric stochastic model [180]. HMM was applied as a probabilistic approach to model a given action into hidden states representing an arbitrary decomposition of the whole movement into successive phases. For example, an HMM algorithm was applied to classify the performed exercise into six possible correctness classes according to its execution [271]. Also, logistic regression was used for movement evaluation [278, 222] and activity monitoring [172]. Logistic regression is considered the baseline supervised machine learning algorithm to classify an observation, and its outputs can be easily interpreted as probabilities of the occurrence of a class [230]. In the case where authors handled multiple classes [217], a

generalization of logistic regression was used: the softmax regression algorithm [72]. A more simplest approach used to make class-probability inference was the NB algorithm. The NB classifier is fast and easy to implement, resulting in particularly used for the implementation of a recommendations system. However, one of its biggest disadvantage is that the algorithm is based on Bayes' theorem that assumes independence between the features which are dependent in most of the real-life cases [70].

The KNN is another algorithm often applied in classification problems. KNN is a non-parametric classification algorithm known for its simplicity and effectiveness, used to classify data based on closest or neighboring training examples in each region [300]. The KNN classifier is based on distance metrics and was widely used in real-time applications for activity monitoring [69, 254, 26, 172, 40] as it is free from the underlying assumptions about the distribution of the dataset. Finally, Support Vector Machines (SVM) was the third most used traditional machine learning classifier applied for classification tasks in activity recognition [69, 31, 192, 254, 26, 40], movement classification [326], and also for clinical assessment scopes [82]. SVM algorithms present good generalization ability for sequential data structures and datasets that are not too large and for linear and non-linear problems.

Artificial neural networks were also a common choice among researchers interested in implementing solutions for monitoring and assisting decentralized physical rehabilitation. Artificial neural networks are a subset of machine learning algorithms inspired by neuroscientific studies concerning the functioning of the neurons' brain, whose branch is known as deep learning [116]. The recent advancements in deep learning, such as hybrid and lightweight deep neural networks, have been exploited to develop a framework for exercise performance evaluation and smart assistance for home-based rehabilitation sessions. Convolutional neural network (CNN) was the commonest architecture applied in selected papers. CNNs are designed to learn spatial hierarchies of features automatically and adaptively through back-propagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers [331]. The CNNs have been trained on large-scale datasets and on graph-structured data for motion analysis issue [315, 152]. CNNs achieved outstanding accuracies in the computer vision field for human detection and pose estimation that is useful to compute position and orientation data of interesting joints [32, 327, 320, 200], but also for activity monitoring [39, 48, 254, 40] and movement evaluation [37]. Particularly CNN architectures such as ResNet [115] and AlexNet [155] achieved the highest accuracy respectively in pose estimation method for tracking human motion [320] and activity monitoring by processing some kinematic data [254]. Also, recurrent neural network (RNN) models achieved high accuracy in activity monitoring, which, thanks to the redundant connections between the network's neurons, can capture temporal dependencies in input data [319, 312].

More in detail, long short-term memory (LSTM), which is an RNN variant model [173], is particularly useful for analyzing temporal sequences between motion data frames [200]. The latest common deep learning approach among articles was the multilayer perceptron (MLP). MLP showed to achieve high accuracy above all in detecting specific movements [31]. MLP is a feed-forward fully connected network consisting of an input layer that receives data, an output layer that makes the decision or prediction about the input signal, and one or more hidden layers between these two that are considered the network's computational engine [114]. MLP, not requiring preliminary feature engineering, is considered easier to develop and apply in smart assistant applications than other deep learning methods. Also, a more simple learning algorithm, called extreme learning machine (ELM) that is composed of a single-hidden layer feed-forward neural networks [143], was able to achieve high accuracy in multi-class classification task in work [41].

2.2.5 Clinical applications

Besides technical requirements, it is fundamental to integrate clinical knowledge into technologies to be used at home in a minimally supervised way. Hence, the role of AI in clinical practice should be to provide a combination of medical, psychological, and technical knowledge in the form of embedded algorithms analyzing and processing online the data generated by digital technologies. AI is expected to play a crucial role in clinical decision-making, the online adaptation of therapy exercises, and the monitoring of progress through the extraction of validated assessment scores [164]. Based on the reviewed papers, three main approaches have been introduced to manage the remote monitoring and evaluation of physical rehabilitation therapies by means of machine learning applications. These approaches could be categorized as activity recognition, movement classification, and clinical functional assessment.

Machine learning algorithms address activity recognition tasks aimed to identify specific rehabilitation actions to remotely track patients' adherence to the prescribed therapy and sometimes to measure the treatment outcomes based on patients' activity in real-world life [238]. Most of the reviewed papers covered the problem of recognizing activities for specific body parts like strength training exercises for upper and lower limb [69, 254, 200, 217], exercises to improve range of motion [192], flexibility and balance exercises [288, 200], specific rehabilitation exercise for upper limb and lower impaired limb in stroke [31, 48, 149], and shoulder impairments [39, 26, 40]. Other articles monitored activities of daily living (ADL), detecting general activities such as standing, sitting, squatting, walking, and running useful for health promotion programs in the elderly [17, 200], and other routine activities involving

upper limb goal-directed tasks in stroke rehabilitation [172].

Movement classification approaches aimed to assess the quality of rehabilitation exercises performed remotely. Algorithms addressing movement classification tasks have been used to evaluate exercise performance in terms of the well and poorly-executed tasks. Some approaches relied on classifying the movement into discrete score classes based on clinician rating, such as low, moderate, and high performance of hand movements in stroke rehabilitation [41, 278] and well or badly performed upper limb movements while playing exergame [222]. Always based on clinical expertise knowledge, other approaches classify movements in correct or erroneous execution classes defined by a set of pre-established rules [326, 133, 271, 54]. For example, a set of predefined rules was used to label correct and compensatory movements in different possible classes to provide real-time feedback during remote hip rehabilitation exercises [271]. Noteworthy, another approach based on the movement classification task was applied to develop an accurate segmentation of exercise repetitions to evaluate the patient's performance in a detailed way [37, 25].

The last interesting approach in reviewed papers aimed to manage rehabilitation interventions through preventive and personalized solutions by predicting the patient's clinical functional status. A personalized exergame based on classifying the user profile according to related impairment characteristics was introduced to adapt therapeutic settings during a motor neurorehabilitation session [96]. An accurate and quantitative balance assessment based on a validated clinical instrument mini-BEST test was proposed to recognize the level of patient impairment and predict the fall risk [327]. Finally, a multi-class classifier was used to separate new users into three groups based on their initial activity level (vigorous, mild, and sedentary users) to develop a conversational agent-assisted health coaching system to support behavioral change *via* personalized recommendations [82].

2.2.6 Drawbacks

In the analysis of literature emerged some drawbacks of the current research in AI applications. Major issues regard the size and characteristics of the samples and the low generalizability of results in a relevant operational environment.

In general, the data are mainly collected from a limited sample composed mainly of healthy subjects that did not reflect the pathological behavior of the end-user population target for which the technologies have been developed.

Few systems have been tested and validated on relevant clinical population targets: stroke [41, 31, 48, 172], Parkinson's disease [326, 327], cerebral palsy [192], patients with shoulder impairments [39, 26], osteoarthritis [32, 25], chronic obstructive pulmonary disease [133].

Other studies recruited subjects with various motor impairments related to orthopedic and neurological conditions [96, 32, 222, 54], thus addressing an enormous heterogeneity in collected data.

Other problems concerning the generalizability of results have been addressed as related to the reproducibility of the data collected in limited or different conditions (such as settings, users, and operators).

Considering the technology readiness level (TRL) of all the reviewed solutions, it generally spans from level 3 (experimental proof of concept) to level 4 (technology validated in the laboratory) [153]. A very minority of studies validated and demonstrated the application of proposed AI-based solutions in relevant operation environments such as remote conditions and home-based settings [192, 82, 39, 48].

Chapter 3

Data for human motion analysis

3.1 Data modeling and collection

3.1.1 Approaches for motion evaluation

A key prerequisite for evaluating the progress of patients in home exercise programs is the provision of an efficient and comprehensive performance evaluation approach. Mathematical modeling of human motions is a challenging research topic that has been investigated for a wide range of applications. Nevertheless, from a general point of view, modeling of human motions remains a challenging problem, due to several aspects related to their intrinsic properties.

First, the selection of the approach to model and analyze rehabilitation performance depends on the nature of the exercise, whether it involves simple or complex movements. A complex movement is characterized by patterns that can differ between individuals due to the specific functional activity that is being performed [60, 134]. These movements involve multiple joints and muscles that work together to accomplish a specific task or functional activity. Complex movements are more context-dependent and may vary depending on individual abilities, body mechanics, and motor skills. Activities such as drinking, walking or picking up objects are examples of complex movements because they require coordination and integration of various body movements to achieve the desired outcome. On the other hand, a simple movement can be defined as a physiological bending and extension along the range of motion of a particular joint or body part [60]. This kind of movement typically follows a standard pattern and is consistent across individuals. Examples of simple movements include basic joint flexion and extension, such as bending and straightening the elbow or knee.

Existing approaches for evaluating rehabilitation exercises can generally be categorized into three distinct areas: comparison, categorization, and scoring [70].

Comparison techniques involve assessing a patient’s exercise performance by contrasting their motion sequences with predefined template motion sequences. These templates are typically derived from accurate executions of exercises by healthy individuals, clinicians, or patients under clinical supervision. Categorization methods, on the other hand, are designed to classify patient activities into discrete groups, such as correct and incorrect movement patterns. Scoring approaches aim to provide automated evaluations of a patient’s motion using personalized clinical or non-clinical scoring systems, utilizing both continuous and categorical scales.

Unlike comparative and scoring analyses, categorization methods play a crucial role in delivering real-time patient feedback and guidance during rehabilitation sessions. The selection of the most suitable computational technique for monitoring and assessing human movements significantly impacts the collection and organization of datasets.

Several datasets have been published in the field of rehabilitation research. Datasets like SPHERE [245], MSR-3D [175], and UTD-MHAD [50] have been compiled and used for monitoring applications, primarily offering graphical or statistical comparisons of patient actions without providing definitive patient evaluations. For example, the SPHERE dataset includes examples of both normal and physically impaired movements for activities like walking, ascending stairs, and sit-to-stand transitions. The KIMORE [44] and UI-PRMD [313] datasets have been designed to develop an evaluation system that assigns a score to exercise performance.

Conversely, other datasets like the Toronto Rehab Stroke Pose (TRSP) [74], Kinect 3D Active (K3Da) [174], and IntelliRehabDS (IRDS) [218] were created to classify movements into categorical classes. For example, the TRSP dataset contains pose estimations of stroke patients and healthy subjects performing upper limb rehabilitation movements, annotated for regular and compensatory patterns of motion during an assisted-robot rehabilitation session. The IRDS dataset contains gestures of simple rehabilitation movements performed by both healthy and pathological subjects under various conditions (sitting or standing), labeled for correct and incorrect executions.

The currently available published datasets exhibit limitations in motion data modeling, which are crucial to provide real-time patient feedback and guidance during rehabilitation sessions. In this thesis a methodology for modeling motion data is proposed to support AI-driven assistive systems for home-based rehabilitation programs. The approach enables real-time feedback to patients, allowing them to correct their performance even in the absence of direct clinical supervision.

The modeling process focuses on establishing a set of categorical classes for movements associated with different phases of exercise. The proposed modeling approach aims to im-

plement a set of categorical movement classes linked to specific exercise phases, enabling the detailed evaluation of exercise repetitions based on the quality of motion patterns [143].

To support the objectives of this thesis, a subset of a publicly available dataset was adapted for rehabilitation exercises (see Section 3.1.2), and a new dataset focused on preventing physical decline was collected (see Section 3.1.3).

3.1.2 A dataset for shoulder rehabilitation

The IRDS differs from other datasets is its inclusion of natural motions performed by both patients and healthy subjects, encompassing correct and incorrect patterns. Moreover, the dataset features patients with diverse pathological conditions, such as stroke and spinal cord injury, offering a valuable resource for enhancing the monitoring and assistance of upper limb rehabilitation across various populations, including those with neurological and musculoskeletal disorders.

While the current published form of the IRDS dataset serves as a valuable resource for categorizing exercises and assessing their correctness, it falls short of addressing the needs of digital assistive rehabilitation systems. These systems require a finer evaluation, recognizing repetitions of each exercise and assessing them for the completeness of range of motion (ROM) while analyzing incorrect patterns of motion, as suggested by recent works [37, 176]. The overall aim is to provide a dataset that facilitates the evaluation of whether an exercise is performed correctly according to its prescription or not.

A subset of the IRDS dataset, containing shoulder movement repetitions performed by 29 subjects (15 patients and 14 healthy individuals), was selected for modeling. Recorded using a Microsoft Kinect One sensor at 30 frames per second, the dataset includes 3D coordinates of 25 body joints and corresponding depth maps for each frame. Each movement is carefully annotated with details such as exercise type (shoulder flexion, abduction, or forward elevation), limb side involved (right, left, or both), subject position (sitting or standing), and motion pattern correctness (correct or incorrect). Shoulder flexion and abduction exercises involve performing gestures separately with the left and right body sides, while shoulder forward elevation requires simultaneous movement of both sides. Table 3.1 summarizes the rehabilitation exercises included in the dataset.

The data modeling process aims to enable a comprehensive evaluation of the range of motion (ROM) and compensatory movement patterns during a shoulder rehabilitation exercises session.

To increase the granularity of the analysis, a set of categorical ROM classes was associated with specific exercise phases in an individual's performance. This approach enables a

Table 3.1: Description of exercises included in the new dataset for shoulder therapy.

Gesture Index	Gesture Name	Description
2	Shoulder Flexion Left	Flexion and extension movement of left shoulder while keeping the arm straight in front of the body
3	Shoulder Flexion Right	Flexion and extension movement of right shoulder while keeping the arm straight in front of the body
4	Shoulder Abduction Left	The left arm is raised away from the side of the body while keeping the arm straight
5	Shoulder Abduction Right	The right arm is raised away from the side of the body while keeping the arm straight
6	Shoulder Forward Elevation	With hands clap together, the arms are kept straight and raised above the head, keeping the elbows straight

detailed assessment of each movement, considering both the accuracy and completeness of its execution. In the IRDS dataset, each frame, capturing both correct and incorrect movements, underwent labeling for ROM execution classes. Depth images were assessed to assign classes such as complete concentric phase (up), complete eccentric phase (down), incomplete eccentric phase (incomplete down), and incomplete concentric phase (incomplete up). These labels serve three main objectives in rehabilitation evaluation: segmenting exercise routines into repetitions, assessing each repetition’s completion status and providing feedback for performance enhancement, indicating which phase of the repetition is considered invalid—concentric, eccentric, or totally incomplete.

Furthermore, incorrect movements were labeled to distinguish various classes of ROM, allowing for a detailed evaluation of repetitions not only in terms of motion completeness but also in terms of the percentage of correct and incorrect patterns. This approach provides insights into the quality of motion and facilitates tracking advancements during the rehabilitation process.

The methodology is based on the hypothesis that separating the evaluation of ROM and the recognition of compensatory movements into distinct stages can enhance the accuracy of AI-assisted rehabilitation systems. This separation reduces the complexity of each task by minimizing the number of classes, thereby potentially improving overall system performance.

Formally, are considered a set of input–output pairs $D = \{(x_i, y_i)\}_{i=1}^N$ where $x_i \in R^n$ are the N samples of the input feature space obtained from the 3D video motion analysis and $y_i \in C = \{C_j | 1 \leq j \leq j\}$ are the class of labels to which these features correspond.

The evaluation look for a decision function $f : R^n \rightarrow C$ which given a sample, that in this work contains features of body joints' 3D pose data, determines the output class that includes the kind of execution performed: $x \rightarrow y = f(x, w)$. Throughout the following, let $X(z) = \{X_t, m_t, c_t\}$ denote a labeled sequence of N frames of the considered set of the three shoulder exercises $z \in \{S1, S2, S3\}$, where X_t denotes the acquired RGB-D image at frame t , and m_t represents the ROM phase of the exercise execution. Moments of a given exercise comprise both the key ROM phases of the exercise, resulting as $m_t \in C = \{C_j | 1 \leq j \leq 4\}$. Moreover, $c_t \in C = \{C_j | 0 \leq j \leq 2\}$ denoting whether the t frame belongs either to a physiological or compensatory motion pattern class.

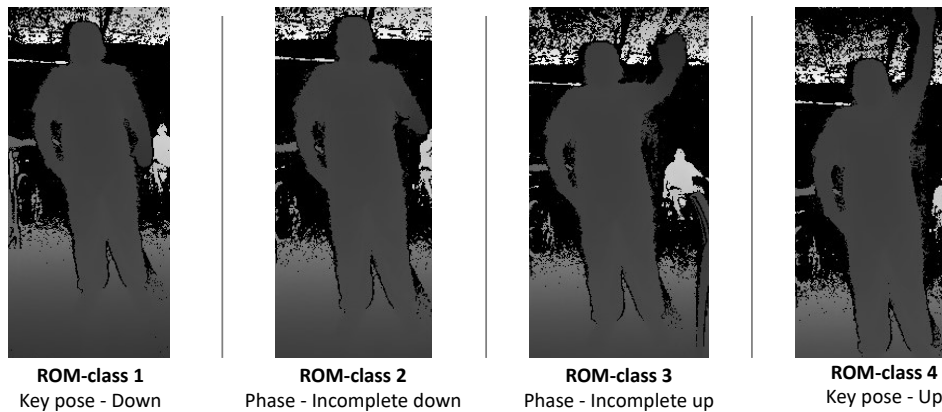


Figure 3.1: A chronological sequence of left-side shoulder flexion, labeled within the context of ROM classes.

Note. In the example are showed depth images of patient with ID=215.

At the final stage of the modeling process, the new dataset consists of 1488 files for shoulder rehabilitation movements. Each file is associated with i) subject ID, ii) gesture type, iii) number of repetitions, and iv) label of correct movement and incorrect movements, following the structure and nomenclature of the original dataset (*'SubjectID.GestureType.RepetitionNo.CorrectMovement.Position.csv'*).

The data contents of each file, presented in *csv* format, can be described as follows: (i) each file contains n frames, (ii) each frame contains spatial information of 25 joints along the three axes (x, y, z), (iii) each frame contain temporal information of ROM phase execution.

The ROM phase executions are available along with depth image files. Figure 3.1 displays

a chronological sequence of motion for ROM classes, exemplifying their temporal order in execution.

The comprehensive dataset composition is outlined in Table 3.2, providing a detailed breakdown of ROM classes for both the healthy and patient categories.

Table 3.2: Description of the new dataset in terms of samples for each ROM class category along the various gesture types.

Gesture Index	ROM_class	Healthy	Patients
2	down	1098	684
	up	4022	2793
	incomplete_down	2959	17790
	incomplete_up	1349	9075
3	down	951	398
	up	3785	3139
	incomplete_down	2873	14238
	incomplete_up	1321	6819
4	down	929	1267
	up	4286	552
	incomplete_down	3757	6993
	incomplete_up	1498	1945
5	down	893	380
	up	3967	2879
	incomplete_down	3342	4314
	incomplete_up	1475	2282
6	down	3162	3545
	up	4687	2792
	incomplete_down	995	2730
	incomplete_up	1903	3141

Finally, the distribution of data for each ROM class was systematically examined within each patient and healthy subject category for the various exercise types. Figure 3.2 depicts a bar plot that distinguishes between patient and healthy category IDs for each gesture type.



Figure 3.2: Distribution of ROM classes for gesture type and subject category.

3.1.3 A dataset for sarcopenia prevention

The dataset presented in this section was designed to evaluate motion performance within a rehabilitation protocol involving resistance training exercises.

Resistance training is defined as a strength training exercise with the use of progressive overload in which the muscles create the force against external load. Resistance training interventions might be particularly beneficial to reduce the causes (e.g., loss of muscle mass) and consequences (e.g., loss of muscular strength or functionality) that both syndromes usually produce, even at early stages [246, 299]. The exercise protocol is proposed as an effective strategy to treat sarcopenia and physical frailty [202, 334, 206]. Furthermore, the protocol was developed to be scalable for home-based rehabilitation programs and to fit the needs of a large range of individuals with reduced motor ability such as elderly people [124].

The exercises were selected according to the guidelines of the American College of Sports Medicine (ACSM) for resistance training for older subjects [228]. The protocol consists of 6 rehabilitation exercises: 3 for upper limbs (biceps curl (E1), triceps curl (E2), front raise (E3)), and 3 for lower limbs (squat (E4), leg extension (E5), leg curl (E6)). These exercises targeted the major muscle groups, such as the legs, back, abdomen, chest, shoulders, and arms.

Resistance training exercises are typically characterized by the execution of a muscle contraction (concentric phase) followed by muscle relaxation (eccentric phase) repeated a defined number of times. So, the repetition of a certain exercise corresponds to the period between two consecutive key phases (concentric and eccentric phases) that a subject should accomplish. The concentric and eccentric phases are correctly performed when the subject achieves the full ROM of the joint interested in the exercise. For this reason, it is important to distinguish series performed along the entire ROM as correct executions from the wrong performances that are represented by incomplete executions in concentric/eccentric phases. Furthermore, the patient intending to achieve the target movement during the concentric phase of resistance training exercises could accomplish compensatory movements that have a deleterious outcome on the rehabilitation program. A meticulous evaluation should promptly recognize the performance of compensatory movements to preserve the beneficial effects of rehabilitation exercise. A set of pre-defined categories of the compensatory patterns of motion were defined for each exercise in agreement with existing guidelines [60].

Under the scope of the proposed approach, a dataset containing the video of 6 rehabilitation exercises performed in the correct and wrong manner was acquired. The dataset was obtained by recording a rehabilitation expert performing the pre-defined variants of the exercise motion. Five of the 6 exercises requested the person in a stand position (S1). Exercises

E2 and E6 were recorded with the therapist in the seated position (S2). Each video was recorded from the best camera perspective which allows us to perfectly monitor the exercise. Three scenarios were considered for selecting the most relevant point of view to describe subjects' movements in front of the camera: a lateral view with the right side perpendicular to the camera (V1), a frontal view facing the camera (V2), and a lateral view with the right side oblique to the camera (V3). For V1 and V3, only the right side is completely visible in the image. All samples were recorded under a resolution of 1280x720 at 30 frames per second. All frames of each video were labeled in agreement with a physical rehabilitation expert concerning temporal and categorical motion domains. The labeling process was carefully conducted twice by a rehabilitation expert to ensure accuracy and avoid any errors in labeling. The characteristics of each exercise, the information regarding scenario settings, as well as the classes corresponding to the categories for the evaluation of the completeness of ROM and compensatory movements, are all described in Table 3.3.

Table 3.3: Exercises description and characteristics.

Exercise name (N)	Camera view and Scenario	List of ROM-classes	List of compensatory movements
Biceps curl (E1)	Camera view: V1 (frontal), V3 (lateral oblique); Scenario: S1 (standing)	Down (ROM-class 1): elbow extension ($\leq 30^\circ$) with the forearms parallel to the trunk; Incomplete down (ROM-class 2): incomplete elbow extension; Incomplete up (ROM-class 3): incomplete elbow flexion; Up (ROM-class 4): elbow flexion ($\geq 130^\circ$) with arms parallel to the trunk.	C1: excessively shoulder flexion with arms not parallel to the trunk during ROM-classes 3-4; C2: excessively trunk extension during ROM-classes 3-4.
Triceps curl (E2)	Camera view: V2 (lateral perpendicular), V3 (lateral oblique); Scenario: S2 (seated)	Down (ROM-class 1): elbow flexion ($\geq 130^\circ$) with the arms in line to the trunk; Incomplete down (ROM-class 2): incomplete elbow flexion; Incomplete up (ROM-class 3): incomplete elbow extension; Up (ROM-class 4): elbow extension ($\leq 30^\circ$) with the forearms parallel to the trunk.	C1: excessively shoulder abduction with the arms and trunk not aligned below them during ROM-classes 3-4.

Continuing of Table 3.3

Exercise name (N)	Camera view and Scenario	List of ROM-classes	List of compensatory movements
Frontal raise (E3)	Camera view : V1 (frontal), V3 (lateral oblique); Scenario: S1 (standing)	Down (ROM-class 1): shoulder extension with the forearms parallel to the trunk; Incomplete down (ROM-class 2): incomplete shoulder extension; Incomplete up (ROM-class 3): incomplete shoulder flexion; Up (ROM-class 4): shoulder flexion at 90° with elbow aligned to shoulder and wrist joints.	C1: excessively trunk extension during ROM-classes 3-4; C2: elbow flexion not aligned to shoulder and wrist joints during ROM-classes 3-4.
Squat (E4)	Camera view: V1 (frontal), V3 (lateral oblique); Scenario: S1 (standing)	Up (ROM-class 1): standing upright position; Incomplete up (ROM-class 2): incomplete upright position; Incomplete down (ROM-class 3): incomplete squat position; Down (ROM-class 4): squat position with bent knees and hips aligned below them.	C1: excessively trunk flexion during ROM-classes 3-4; C2: knees and ankles joints not aligned below them during ROM-classes 3-4.
Leg extension (E5)	Camera view: V1 (frontal), V3 (lateral oblique); Scenario: S2 (seated)	Down (ROM-class 1): seated position with bent knee (90°); Incomplete down (ROM-class 2): incomplete flexion of knee; Incomplete up (ROM-class 3): incomplete extension of knee; Up (ROM-class 4): seated position with extended knee aligned to hip and ankle joints.	C1: hip flexion with knee not aligned to hip joint during ROM-classes 3-4.
Leg curl (E6)	Camera view: V1 (frontal), V3 (lateral oblique); Scenario: S1 (standing)	Down (ROM-class 1): standing upright position; Incomplete down (ROM-class 2): incomplete flexion of knee; Incomplete up (ROM-class 3): incomplete extension of knee; Up (ROM-class 4): seated position with bent knee aligned to hip joint.	C1: hip extension with knee not aligned to hip joint during ROM-classes 3-4.

During the repetition of a certain exercise, key poses represent the most important human poses that everyone should accomplish, and phases correspond to the period between two consecutive key poses. Therefore, frames with a common phase or key pose will have the

same temporal label accordingly. On the other side, repetitions can also be analyzed in terms of validity at a frame-level, since a set of categorical motion categories of ROM and compensatory movements can be associated with each frame depending on the individual's performance, as shown in Table 3.3.

The temporal and categorical motion domains to evaluate the valid repetitions taking into account the ROM are visually depicted in Figure 3.3.

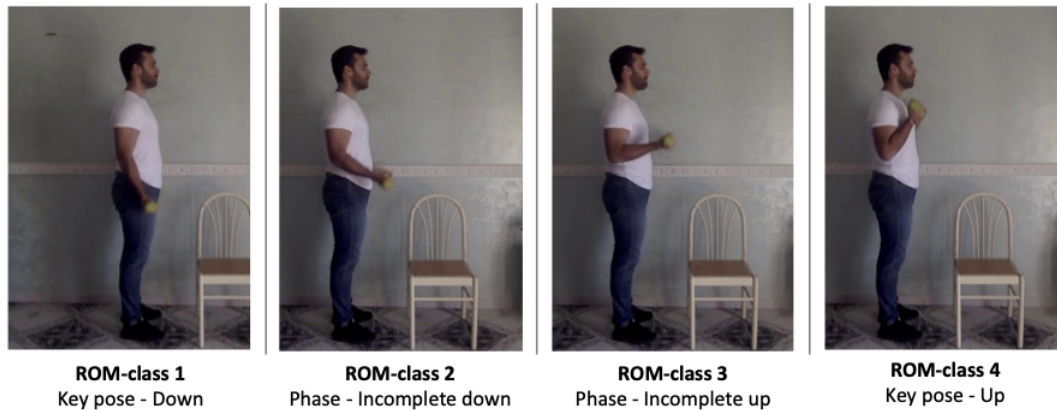


Figure 3.3: Exercise correct repetition.

Note. The figure is a visual description example of a biceps curl (E1) repetition, highlighting: i) the difference between key poses and phases, ii) the difference between each class of ROM execution.

For instance, during the execution of the frontal raise exercise (E3), individuals might excessively extend the trunk (C1) or bent elbow flexion (C2) in ROM-classes 3-4 instead of following the physiological pattern of movement (see Figure 3.4).



Figure 3.4: Compensatory movement.

Note. An example of compensatory movement during the execution of a frontal raise (E3).

This dataset was implemented to develop a preliminary proof-of-concept system capable of monitoring and evaluating exercise execution, as described in Chapter 4. For detailed

statistical characteristics of the dataset, including the number of samples for each movement class and exercise type, refer to Section 4.2.4.

3.2 Synthetic data generation: a new research challenge

3.2.1 Existing methods

The recent application of deep learning techniques in various domains has been significantly bolstered by the advent of data augmentation methods [292, 156]. Data augmentation serves as a powerful tool, enabling the generation of diverse training samples from existing data, thereby enhancing the performance of machine learning models. While traditional augmentation techniques, such as rotations and flips, have proven effective in many applications, they often fail to capture the full range of semantic variations present in the rehabilitation data such as the diverse range of exercise motion patterns, patient demographic, and physical characteristics, or the context of the therapy setting.

The augmentation of datasets containing labeled images of rehabilitation exercises offers several benefits. Firstly, it addresses the issue of insufficient training data, which is a common challenge in real-world rehabilitation settings where access to large-scale, diverse datasets is limited. Synthetic data serves as a valuable resource to improve the diversity of data in a supervised learning setting without the time-consuming labeling process. Synthetic data effectively expands the training dataset and promotes a more comprehensive representation of the underlying data distribution by synthesizing new samples that accurately capture variations in human movements. Synthetic data can be seen as a logical extension of traditional data augmentation techniques, which are standard in image processing. These augmentation methods, such as rotations, flips, and zooms, introduce minor variations to existing data to enhance model performance and generalization. However, their ability to generate diverse samples along key semantic axes may be limited. Synthetic data, on the other hand, enables the creation of entirely new instances that encompass a broader range of pose variations, postures, and gestures, enhancing the overall diversity of the dataset.

Researchers have recognized the significance of synthetic data as a tool to augment existing datasets. Studies such as those by Shorten and Khoshgoftaar [290], Perez and Wang [250], Inoue [128], and Müller and Holzinger [223] have explored various approaches to leveraging synthetic data for improving classification and recognition tasks. These works highlight the benefits of synthetic data in enhancing the generalization and performance of supervised learning models.

Moreover, Seib et al. [282] provide a comprehensive overview of different approaches for enriching real data with synthetic data. Their study discussed techniques ranging from generative adversarial networks (GANs) to physics-based simulations, demonstrating the broad applicability of synthetic data across various domains. By combining real and synthetic data, researchers can create more comprehensive and diverse datasets that can further bolster the performance and robustness of machine learning models.

In the context of deep learning, the success of models largely hinges on the size of the training dataset available. In the medical field, however, obtaining extensive and diverse medical data poses significant challenges due to privacy concerns and limited accessibility. To tackle these issues, researchers have turned to GANs within the medical domain. GANs have shown their versatility in various applications, including addressing legal and ethical data-sharing dilemmas, translating data between different modalities, and enhancing the performance of deep learning models. Nonetheless, generating meaningful medical data is exceptionally challenging, as medical diagnoses often rely on detecting subtle changes in complex organ structures. Additionally, GANs suffer from inherent architectural problems, such as failing to capture true diversity and exhibiting issues like mode collapse and unstable training behaviors. Therefore, generating high-quality synthetic medical data has become a critical focus.

Recent breakthroughs in denoising diffusion probabilistic models (DDPMs) and latent DDPMs have achieved state-of-the-art results, surpassing GANs in the realm of natural images. However, a comprehensive, large-scale comparison between latent DDPMs and GANs for medical images across various medical domains is still lacking.

Initial studies have started to compare the performance of Diffusion Probabilistic Models (DDPMs) and Generative Adversarial Networks (GANs) in specific medical image synthesis scenarios. For instance, Pinaya et al. utilized a latent DDPM to generate brain MRI images, surpassing the performance of Least Squares GAN and Variational Autoencoder [252]. Similarly, Dorjsembe et al. employed a DDPM to create brain MR images, with the DDPM outperforming GAN in quantitative evaluations but falling short when compared to a Cycle Consistent Embedding GAN [76]. Akbar et al. observed that DDPMs exhibited a stronger capacity to memorize training images when generating brain MR images in contrast to StyleGAN [7].

In the context of training a classification network using either real or synthetic radiographs, Packhäuser et al. achieved competitive results with images generated by a latent DDPM but encountered limitations with those produced by a Progressively Growing GAN (ProGAN) [244]. Furthermore, Moghadam et al. demonstrated the superior performance of DDPMs over ProGAN when generating histopathological images [221]. Collectively, these

studies underscored the potential of DDPMs as a promising alternative to GANs within the medical domain, particularly when leveraging latent DDPMs. These specialized models, distinct from conventional DDPMs, enable the generation of larger images and volumetric data through sampling within a compressed latent space, making them especially valuable for medical data synthesis.

Overall, synthetic data represents a powerful means of enriching existing datasets. Its utilization extends the capabilities of traditional data augmentation methods and addresses the challenges of limited labeled data. By integrating synthetic data, researchers can significantly improve the diversity, generalization, and performance of machine learning models across various applications, including rehabilitation frameworks. The fusion of real and synthetic data lays the foundation for the development of more resilient and dependable models, ultimately propelling the field of artificial intelligence (AI) methods in this domain.

In rehabilitation medicine, the integration of synthetic data can alleviate the challenges associated with limited labeled data availability. The reliance on small, potentially biased datasets can be mitigated by synthesizing additional training samples that capture a wide array of human movements with different characteristics. This not only enhances the diversity but also may improve the generalization capabilities of the existing datasets.

Image and video synthesis is a thriving area of research in computer vision and machine learning. It has gained significant attention due to the development of deep generative models and its academic and practical value. Among the various object categories, the synthesis of high-fidelity human images has been a common focus, with researchers exploring different deep generative models, task settings, and applications.

In recent years, the field of image synthesis has made remarkable progress, thanks to the emergence of deep generative models like GANs, and Variational Autoencoders (VAEs) [105, 110, 147, 232]. Notable advancements, such as the DDPM [123], text-to-image generation models as latent diffusion [267] and DALL-E [261], have further enhanced stability, scalability, and performance of image synthesis tasks. These models have spurred research in visual generation tasks, with human image generation being a particularly popular research topic. Human image generation involves simulating complex variations in human body appearance, such as articulated and non-rigid deformations, clothing attributes, background changes, illumination variations, and viewpoint differences across different cameras. This research area has gained significant attention in both computer vision and computer graphics communities.

Deep person generation has also attracted extensive research due to its broad applications in virtual agents, video conferencing, online shopping, and art/movie production. With the advancement of deep learning, it has become easier to generate or manipulate visual appearances, including the face, pose, and clothing, of persons in images.

In particular, the pose-guided person image generation method, originally introduced in the domain of computer vision, has emerged as a promising approach to generate realistic-looking person images with varying poses [134, 284]. Pose-guided person generation aims to generate full-body photo-realistic person images or videos guided by target poses. This method was developed for motion-aware generation tasks like dancing and sports synthesis. It has applications in movie production, online fashion shopping, and more. Existing methods for pose-guided person image generation can be categorized into three groups: top-down, bottom-up, and hybrid methods [284]. Top-down methods directly learn a mapping from input to output images using GAN-like networks [190, 142, 182]. Bottom-up approaches decompose the pipeline into intermediate components and gradually build up the final result [22, 49, 184]. Hybrid methods combine the advantages of both approaches [229, 330, 332]. All these methods face challenges in handling the diversity and complexity of human poses, especially when there are large gaps between source and target poses. Specifically, top-down methods may sacrifice texture details, while bottom-up methods are sensitive to occluded areas.

Recently, a new neural network architecture called ControlNet was adapted to learn the mapping between source person images and target poses for pose-guided person image generation task [336]. Specifically, the integration of ControlNet and latent diffusion models in controlling poses for image generation represents an innovative advancement in the field. ControlNet, originally introduced as a self-supervised learning approach for visual representation learning, has been adapted to specifically address pose-guided person image generation. It operates by learning the mapping between source person images and target poses through training on paired image-pose datasets. With a combination of convolutional and fully connected layers, ControlNet effectively captures both local features and global dependencies, ensuring accurate pose alignment during image generation. Latent diffusion models, on the other hand, have emerged as a solution to challenges faced by generative models, such as mode collapse and training instability. These models employ a diffusion process in which noise vectors are progressively transformed into desired output samples. By exploring the complete data distribution through iterative transformations, latent diffusion models generate diverse and high-quality samples [296, 306].

The combined application of ControlNet and latent diffusion models holds great potential for various domains, including rehabilitation. By precisely controlling poses in synthetic image generation, tailored training scenarios for rehabilitation purposes can be created. Healthcare researchers can simulate a range of poses, movements, and exercises. The generated synthetic images that exhibit accurate pose alignment, visual fidelity, and diversity, may offer a realistic and customizable dataset for effective rehabilitation training of machine

learning-based systems. Moreover, this approach addresses the limitations of relying solely on real-world data, providing a scalable and adaptable solution.

3.2.2 The proposed approach

In this thesis is introduced a novel approach to synthetic data generation in the context of rehabilitation medicine. The proposed method aims to address the challenge of limited data availability in the area. Specifically, for the first time diffusion models to generate synthetic data that mimics realistic-looking human movements in a rehabilitation context is provided in this thesis.

This is the first work that presents a synthetic generated dataset that mimics realistic-looking human movements during rehabilitation exercises. A pose-guided person image method that generates variations of real images using diffusion models and ControlNet architecture is described.

This method introduces significant innovations in the field. Firstly, it pioneers the concept of data augmentation of images containing biomechanical data using pose-guided image generation methods. Secondly, by harnessing the power of cutting-edge techniques like ControlNet and latent diffusion models, the generation of synthetic images with precise and accurate pose control becomes feasible. The integration of these models should enable fine-grained manipulation of poses within the generated images, offering unprecedented control over the desired body movements and joint positions. Thirdly, a qualitative and quantitative analysis were provided to validate the generated dataset.

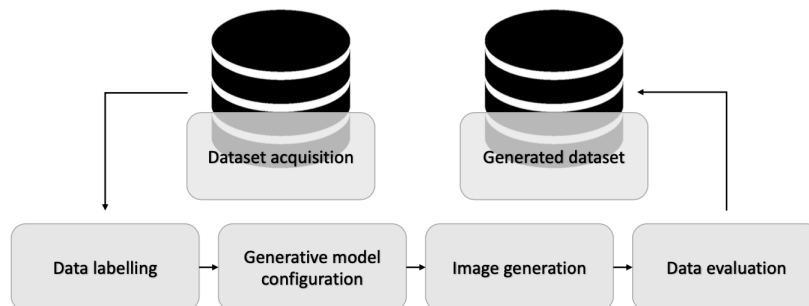


Figure 3.5: Workflow of the proposed methodology.

The objective was to modify the appearance of foreground persons and backgrounds via text-to-image prompts of the diffusion model to introduce more variability in the physical and demographic characteristics of subjects and also vary the setting scenario of the rehabilitation intervention. Furthermore, by conditioning the image generation process on pose information

via the ControlNet architecture this method enabled the synthesis of images that accurately reflect the spatial consistency of human joint relationships of a specific class of movement of a predefined rehabilitation exercise. Specifically, the proposed data generation method was investigated on the dataset preliminary described in Section 3.1.3. The entire workflow of the proposed experimental method is synthetically described in Figure 3.5.

This Section is structured as follows: in paragraph 3.2.2 is described the overall framework to generate synthetic data of rehabilitation exercises with a particular focus on ControlNet architecture used to condition text-to-image latent diffusion model; in paragraph 3.2.2 is described the experimental setup of data generation; paragraph 3.2.2 describes the evaluation metrics used to analyze the robustness of the method to generate images of human exercise movements.

Data generation framework

The data generation framework is based on a latent diffusion model used in combination with the ControlNet architecture. The diagram of Figure 3.6 illustrates the main characteristics of the entire framework using the latent diffusion model in combination with ControlNet.

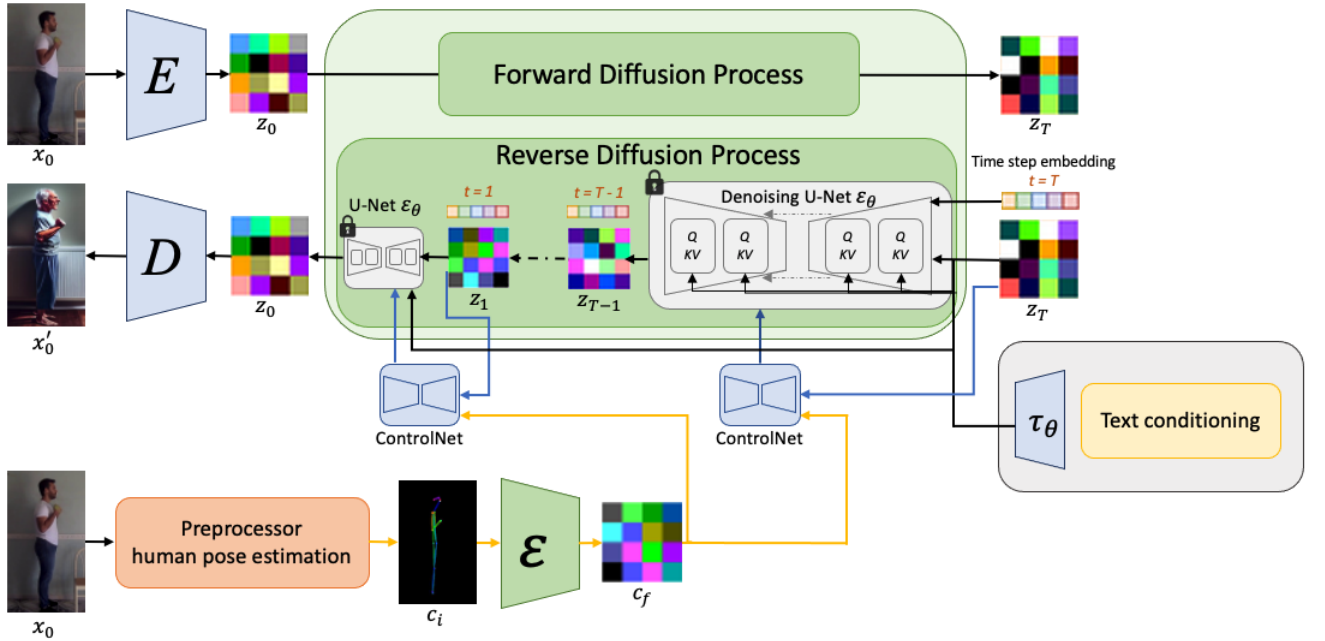


Figure 3.6: Illustration of the pose-guided image generation method using the pose-guided conditioned diffusion model.

A network ϵ is used to encode the input conditions c_i into feature maps c_f because UNet of latent diffusion accepts a latent feature (64×64) instead of the original image. Image

diffusion models learn to progressively denoise images to generate samples. The denoising can happen in pixel space or a “latent” space encoded from training data. Latent diffusion uses latent images as the training domain. Given an image x_o encoded in z_o latent feature, diffusion algorithms progressively add noise to the encoded image and produce a noisy latent feature z_t , with t being how many times the noise is added. When t is large enough, the image approximates pure noise. Given a set of conditions including time step t , text prompts c_t , as well as a task-specific condition c_f , image diffusion algorithms learn a network ϵ_θ to predict the noise added to the noisy latent feature z_t with

$$L = E_{z_0, t, c_t, c_f, \epsilon \sim N(0,1)} [\|\epsilon - \epsilon_\theta(z_t, t, c_t, c_f)\|^2] \quad (3.1)$$

where L is the overall learning objective (loss function) of the entire diffusion model. This learning objective can be directly used in fine-tuning diffusion models on task-specific applications.

In detail, the latent diffusion model is a powerful text-to-image diffusion model that has been trained on a massive dataset of billions of images. This model follows a U-net architecture, featuring an encoder, a middle block, and a skip-connected decoder. Both the encoder and decoder consist of 12 blocks each, and the complete model incorporates 25 blocks, including the middle block. Among these blocks, 8 are responsible for down-sampling and up-sampling convolution layers, while the remaining 17 are main blocks, each consisting of 4 ResNet layers, and 2 Vision Transformers (ViTs) that employ cross-attention and self-attention mechanisms.

Text inputs are encoded using OpenAI CLIP [259], and diffusion time steps are encoded using positional encoding. Initially, the text inputs are converted into embedding vectors through a language model denoted as τ_θ (CLIP), after which they are mapped into the U-Net using a (multi-head) Attention(Q, K, V) layer. The parameters in each level of the UNet are locked in θ and then duplicated into a trainable copy θ_c . Specifically, ControlNet generates this trainable copy, which encompasses the 12 encoding blocks and 1 middle block of the latent diffusion model. The copied θ_c is trained with an external condition vector c_f , referred to as the control map, which, in our approach, contains the positions of human key points as the input condition c_i .

To connect the UNet blocks, a unique type of convolution layer called ‘zero convolution’ is employed [336]. The zero convolution operation is denoted as $Z(\cdot; \cdot)$, and two instances of parameters θ_{z1}, θ_{z2} are used to compose the ControlNet structure with the following loss function:

$$y_{cf} = F(x; \theta) + Z(F(x + Z(c_f; \theta_{z1}); \theta_{cf}); \theta_{z2}) \quad (3.2)$$

In the above equation, y_{cf} represents the output of the UNet block, as visualized in Figure 3.6.

During the initial training step, as the weight and bias of the zero convolution layer are initialized as zeros, the feed-forward process is identical to the one without the ControlNet encoder. However, after backpropagation updates the trainable copy and the zero convolution layers in the ControlNet, the zero convolution weights gradually transition to optimized values.

The overall functioning of ControlNet architecture in conditioning the latent diffusion model is schematically showed in Figure 3.7.

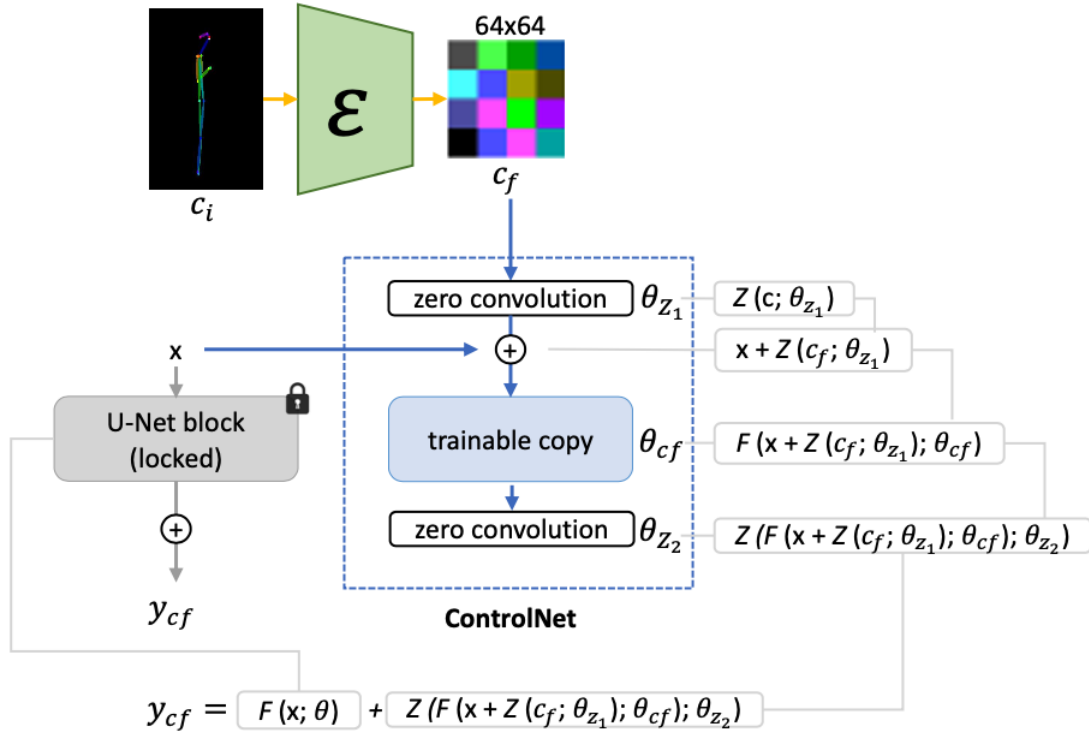


Figure 3.7: ControlNet architecture.

Experimental setup and data generation

The proposed synthetic dataset was generated from the preliminary collected dataset including the resistance training exercise protocol developed. The characteristics of dataset regarding the classes and the number of samples corresponding to each category of exercise movements, are all described in Table 3.4.

Table 3.4: Dataset description in terms of classes of movement and number of samples.

Exercise	Class of movement	Number of samples	
		Camera view 1	Camera view 2
Biceps curl (E1)	Down	292	185
	Incomplete down	250	113
	Incomplete up	409	253
	Up	176	218
	Compensation 1 (C1)	672	456
	Compensation 2 (C2)	352	426
Triceps curl (E2)	Down	237	235
	Incomplete down	221	195
	Incomplete up	369	406
	Up	245	320
	Compensation 1 (C1)	676	421
Frontal raise (E3)	Down	269	312
	Incomplete down	397	350
	Incomplete up	299	240
	Up	313	199
	Compensation 1 (C1)	651	578
	Compensation 2 (C2)	631	286
Squat (E4)	Down	168	132
	Incomplete down	414	489
	Incomplete up	469	440
	Up	280	181
	Compensation 1 (C1)	765	393
	Compensation 2 (C2)	261	377
Leg extension (E5)	Down	477	416
	Incomplete down	165	348
	Incomplete up	222	166
	Up	198	116
	Compensation 1 (C1)	564	436
Leg curl (E6)	Down	627	498
	Incomplete down	213	163
	Incomplete up	177	138
	Up	118	156
	Compensation 1 (C1)	711	430

As a basis for image generation, a pre-trained latent diffusion model, recently released, was used [318]. This is a model that allows images to be created and modified based on image and text prompts. It is a latent diffusion model trained on a subset (LAION-Aesthetics) of the LION5B text-to-image dataset [267], and uses the text encoder CLIP ViT-L/14 to encode the text prompts [259, 274]. To improve the preservation of spatial consistency in human joint relationships within a specific class of movement in the generated images, the ControlNet architecture was integrated to introduce an additional control based on human poses to the diffusion process [336]. With this method, keypoints are extracted from the input image using a 2D human pose estimation network, and saved as a control map containing the positions of human joints. It is then fed to latent diffusion as an extra conditioning together with the text prompt and images are generated based on these two conditions. Text prompts were used to condition the diffusion model, allowing it to modify the foreground aspects to represent an older man or woman and adjust the background setting to depict a traditional home environment. Preliminary tests were executed to find the optimal configuration of setting parameters to generate synthetic images.

The Classifier Free Guidance scale (CFG) is a parameter to control how much the diffusion model should respect the text prompt and it ranges between 0 and 30. CFG value was selected to strictly follow the text-prompt condition. The sampling steps and sampling method of the denoising process determine the quality of the generated image. Quality improves as the sampling step increases. The selected number of steps with the Euler sampler reached a high-quality, sharp image. The image size corresponds to the size of the output image. The random seed determines the initialized noise pattern and hence the final generated image is a new one at each image processing output. Control weight parameter, starting and ending control step parameters control how much the control map is followed relative to the text prompt. The control weight and starting and ending control step parameters range between 0 and 1 it was set so that the control map strictly follows the text-prompt. The control mode controls the importance of the text prompt and the control map during the denoising process so it was set to be balanced. Resize mode controls what to do when the size of the input image or control map is different from the size of the images to be generated. The resize mode was set to scale the width and height of the control map to fit the generated image size.

Finally, all synthetic data were generated using the settings described in Table 3.5. The overall experiment was run on NVIDIA T4 GPU and it took about 12 seconds to create a single image. This resulted in a total time of more than 74 hours for the creation of the whole synthetic dataset.

Table 3.5: Parameters configuration of latent diffusion model and ControlNets to generated synthetic images.

Parameter	Configuration
Text prompt	‘1 old man or woman at home’
CFG scale	25
Sampling steps	20
Sampling method	Euler
Image size	512x512
Batch size	1
Seed	-1 (random)
Control weight	1
Starting control step	0
Ending control step	1
Control mode	Balanced
Resize mode	Just resize

Evaluation methods and techniques

Evaluating the results of generative models is always a challenging task. Generally, the evaluation metrics are divided into qualitative and quantitative ones. The qualitative evaluation metrics, based on human visual evaluations, can reflect the users’ opinion on the generation results directly. On the other hand, the quantitative evaluation metrics are more objective than the qualitative ones however presenting limited robustness [323]. Sometimes, the evaluation results of one metric might be not consistent with the results of other metrics [33, 302].

Synthetic data generation involves creating artificial data that closely resembles real data in terms of statistical and mathematical properties. However, the quality of the synthetic data may not always match that of the original data, and it may not accurately capture the various relationships between different data features. This lack of fidelity can pose challenges for stakeholders who rely on synthetic data for obtaining reliable and high-quality results. Therefore, it is crucial to have metrics that can assess the quality of synthetic data, helping both data generators and users evaluate its usefulness.

One important aspect to consider is how well the synthetic data captures the interdependency between features and the overall context, in addition to matching the statistical properties of the real dataset. These metrics can provide insights into whether the synthetic

data accurately represents the complex relationships present in the original data, ensuring that it can be effectively used for various applications and analyses.

Further analysis was performed to evaluate the consistency and validity of motion data between the original and synthetic datasets.

Paragraph 3.2.2 describes the qualitative evaluation methods. Paragraph 3.2.2 describes metrics used to quantitatively evaluate the generated synthetic images. Paragraph 3.2.2 describes methods used for motion data validation.

Qualitative evaluation

An expert-based evaluation method was addressed to perform a qualitative analysis of the generated images. As a rehabilitation specialist with extensive expertise in biomechanical and motion analysis, I personally undertook the assessment of the generated images.

The qualitative validation process involved the examination of a random sample of 50 images for each category of movement within the six rehabilitation exercises.

The objectives of this evaluation encompassed several key aspects: i) assessing the realism in depicting older individuals engaging in specific exercises within predefined poses; ii) evaluating the introduced variations in the physical and demographic attributes of the depicted individuals; iii) scrutinizing the portrayal of the rehabilitation environment within the images.

This expert-driven evaluation aimed to ensure the quality, fidelity, and realism of the generated images while also examining their biomechanical characteristics and the diversity introduced into the dataset.

Image evaluation metrics

Considering the metrics reported in the literature, following evaluation metrics were selected and applied for the analysis: the Inception score (IS) [275], the Fréchet Inception distance (FID) [120], the Structural similarity index (SSMI) [324], the detection score for person-class and joint-class [284]. The performance of these metrics was evaluated considering also the background influence by masking it out with a state-of-the-art human segmentation model [258].

The IS is a popular quantitative evaluation metric in the research of human image generation, it evaluates both the quality and diversity of the image samples. IS calculates the distribution of generated images via a pre-trained Inception V3 network [23]. The calculation of the IS is based on a statistics formula called the Kullback-Leibler (KL) divergence [64] and it measures the similarity of two probability distributions: the average class probability

(quality) and the marginal entropy of the class distribution (diversity) (see Equation 3.3.

$$IS = \exp\left(\frac{1}{N} \sum_{t=1}^N D_{KL}(p(y|x^t) || \hat{p}(y))\right) \quad (3.3)$$

The IS equation 3.3 contains two main items, $(p(y|x^t))$ and $\hat{p}(y)$ that represent respectively the posterior distribution and the empirical distribution of classification labels, where x indicates the generated image samples and y indicates their labels. Therefore, based on the calculation of KL-divergence, the highest IS values express good quality and diversity in the dataset. The IS provides an overall assessment of the dataset so it was calculated for the whole original and synthetic generated datasets for comparative analysis.

The FID is a metric to measure the quality of images synthesized by generative models. The FID calculates the Fréchet distance between the distribution of generated images and real images to capture their similarity. It assumes that the real images and generated images follow the Gaussian distributions and calculates the Fréchet distance between the two distributions [92]. Similar to IS, FID uses a pre-trained inception network to extract the features of real images and generated images. The multidimensional Gaussian distributions of generated images and real images are defined as $N(\mu_x, \Sigma_x)$ and $N(\mu_y, \Sigma_y)$, respectively. Then, FID is calculated as follows:

$$FID(x, y) = \|\mu_x - \mu_y\|_2^2 + \text{tr}(\Sigma_x + \Sigma_y - 2\sqrt{\Sigma_x \Sigma_y}) \quad (3.4)$$

The FID was applied to compare each ground-truth image to the corresponding generated samples. The mean and standard deviation were reported for the analysis.

The SSIM measures the quality of the generated image compared with the original image, which is widely used in image synthesis. Specifically, SSIM calculates the similarity between synthesized image and ground-truth real image in three dimensions. It considers the luminance, contrast, and structure of the images during evaluation. The SSIM index can be formalized as follows:

$$SSIM(x, y) = I(x, y)^\alpha C(x, y)^\beta S(x, y)^\gamma$$

$$\begin{aligned} I(x, y) &= \frac{2\mu_x 2\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \\ C(x, y) &= \frac{2\sigma_x 2\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \\ S(x, y) &= \frac{2\sigma_{xy} + C_3}{\sigma_x^2 + \sigma_y^2 + C_3} \end{aligned} \quad (3.5)$$

In Equation 3.5, x and y indicate the generated image and referenced image, respectively. The α , β , γ parameters are usually set as 1. μ and σ indicate the mean and standard deviation of corresponding images and σ_{xy} denotes the correlation coefficient between x and y . C_1, C_2, C_3 are the constants. $I(x, y), C(x, y), S(x, y)$ evaluate the similarity of luminance, contrast, and structure between the images, respectively. The SSIM index is a bounded metric ranging between 0 and 1 which higher value indicates that two images are identical. The evaluation of SSIM was applied to compare each ground-truth image to the corresponding generated samples. The mean and standard deviation were used for the analysis.

The DS for person-class (DS-person) measures the confidence of a pre-trained model used for person detector in image samples, by computing the average person-class detection scores on generated images. The model used for person detection was the YOLO model developed for real-time detection tasks of several object classes [263].

The DS for joint-class (DS-joint) measures the average confidence of a pre-trained model used to detect human joints by a pose estimation model. The joint estimation model used was the MoveNet Thunder model which is a recent state-of-the-art bottom-up estimation model that uses heatmaps to localize human keypoints with high performance in accuracy [321]. The confidence prediction of COCO's standard 17 keypoints locations was also analyzed.

The DS for person-class and joint class was computed for each original and synthetic image sample for comparative analysis. The mean and standard deviation values are reported.

Motion data validation

Synthetic data generation is a valuable technique for creating data that mimics real data. However, assessing its quality, including its ability to capture relationships between features, is essential for ensuring its reliability and utility in various domains.

Considering that the proposed dataset was developed to enable systems to automatically evaluate rehabilitation exercise performance, further validation analysis on the consistency

of motion data between original and synthetic datasets is required.

To this scope, the human pose estimation model MoveNet Thunder was used to extract motion data from the original and synthetic images. Specifically, the model extracts the x and y coordinates of the 17 body keypoints in the image matrix. Each keypoint provided is denoted by $p_k^t = [x_k^t y_k^t]$ that represents 2D coordinates in the image of a body keypoint k from a set of joints $K = 17$. These coordinates were used to perform validation analysis in terms of motion data consistency.

The validation of motion data employed two key methods: i) intra-class consistency was assessed by investigating the pairwise mutual information of motion features; ii) inter-class feature distribution was evaluated by examining the preservation of class boundaries within the motion data.

Mutual information is a measure that reveals the inter-dependencies between various data features. It quantifies the amount of information shared between two random variables, allowing us to investigate their mutual dependencies. When generating synthetic data, it is desirable for the mutual dependencies between features to be maintained, resembling those observed in the real data. To assess this, a heatmap illustrating the relationships between features in the synthetic data can be visualized. This visualization provides a clear representation of how the synthetic data captures the mutual dependencies observed in the original dataset. Moreover, the distance correlation metric (dCor) was used to perform a statistical test to analyze the dependency between features correlation matrices of the original and synthetic dataset [297]. The distance correlation measures both linear and nonlinear association between two vectors. Distance correlation ranges from 0 to 1, where 0 implies independence between X and Y and 1 implies that the linear sub-spaces of X and Y are equal.

The t-distributed stochastic neighbor embedding (t-SNE) method was employed to visualize the boundaries of each movement class within both the original and synthetic datasets. This technique involved mapping the high-dimensional data into a two-dimensional map, grouping the data points by their respective movement classes [122]. The t-SNE algorithm implementation from the Scikit-learn library was utilized for this purpose. To ensure reproducibility, a fixed seed value was used due to the stochastic nature of the t-SNE algorithm. After carefully experimenting with different values, the perplexity parameter was ultimately set to 100 to enhance the visualization of the specific pattern of interest.

3.2.3 Qualitative results

A total of 22,352 images were generated from processing a pre-labeled class of movements for each of the 6 rehabilitation exercises via the latent diffusion model using text and human pose conditioning.

Synthetic images were generated to simulate the demographic and physical characteristics of elderly people doing exercise in various home settings. The foreground person characteristics and the background setting were manipulated via the text prompt conditioning of the latent diffusion model using respectively the keywords ‘old’, ‘man/woman’, and ‘home setting’.

The text prompt condition introduced more variability in the physical and demographic characteristics of generated persons as it can be noted from images of the generated dataset that are randomly plotted in Figures 3.8, 3.9.

The keyword ‘old’ served as a conditioning factor for the age of the generated individuals, resulting in all images depicting elderly people appearing to be over the age of 65 years. Furthermore, the text controlling gender output played a pivotal role in balancing the distribution of male and female categories within the generated dataset. The text conditions, used to control the age and sex output, contributed to modifying the physical aspect of synthesized individuals. In fact, from a biomechanics point of view noteworthy was the variability in physical characteristics introduced by the proposed method. The synthetic dataset contains individuals with varying body mass compositions, as well as distinct morphological characteristics such as more pronounced lumbar lordosis and thoracic kyphosis curves, which are typically more prevalent in the elderly.

The processing of posture information and joint relationships for each class of movement proceeded seamlessly, without encountering any significant issues or variations.

Furthermore, the proposed approach introduced variations in additional elements of home scenarios, such as sofas, columns, windows, chairs, and decorative elements. The generated images also exhibit diversity in individuals’ clothing types and colors, including t-shirts, shorts, and pants. For example, as evident in the plot (d) of Figure 3.9, the clothing worn by the individuals appears to be more suitable for a gym or sportswear setting.

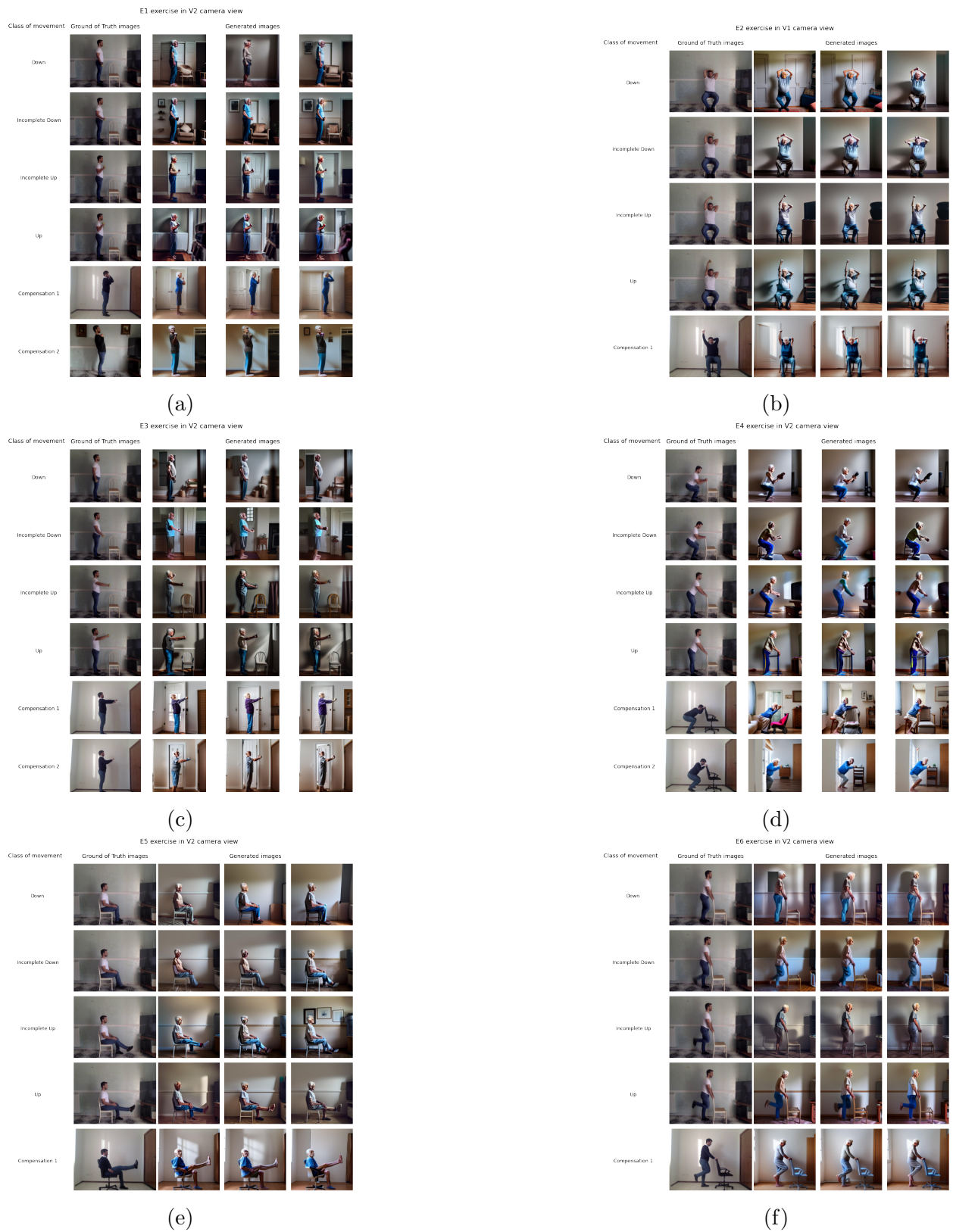


Figure 3.8: Samples of generated images in 'camera view 1'.

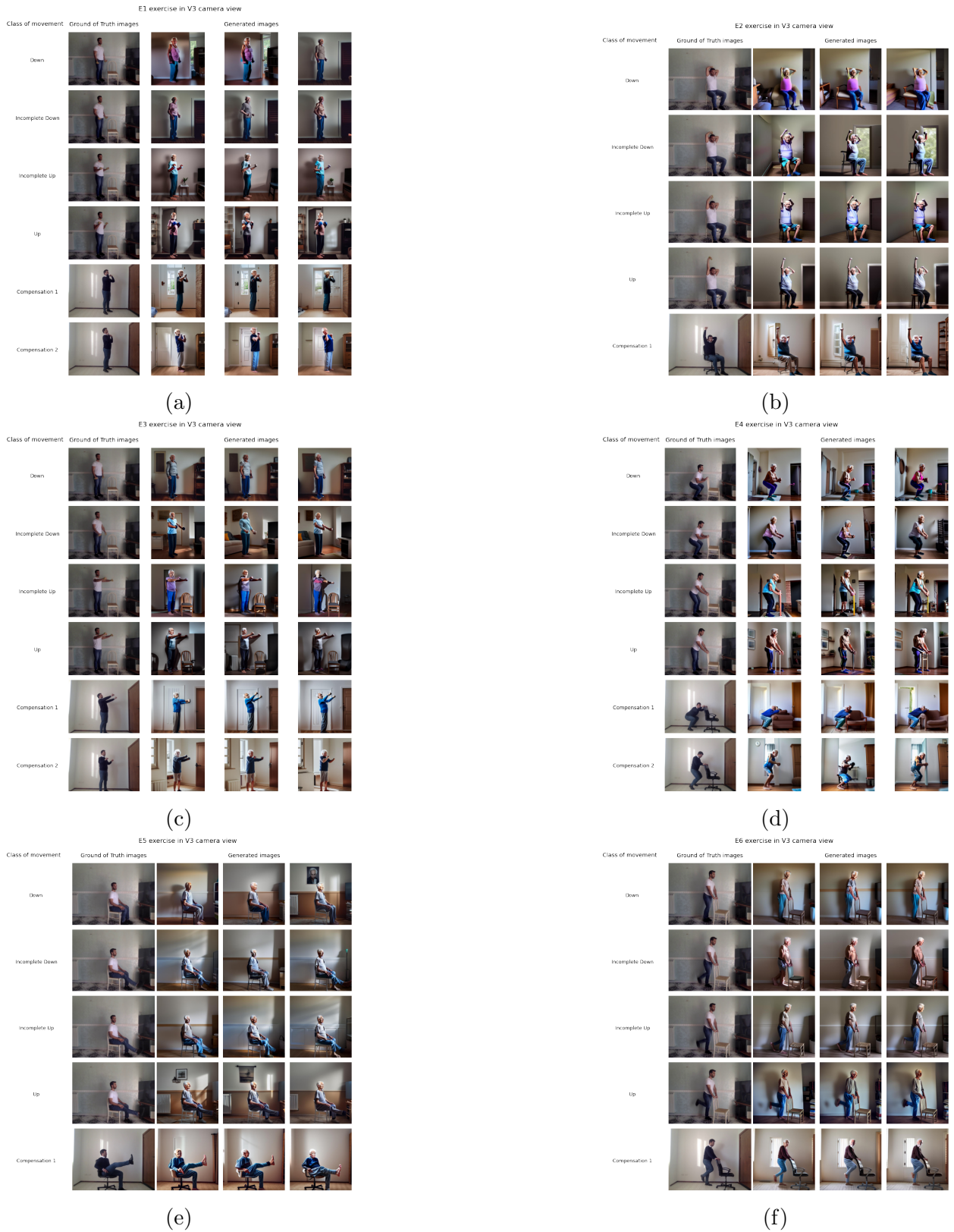


Figure 3.9: Samples of generated images in ‘camera view 2’.

3.2.4 Quantitative results

Generally, quantitative metrics showed comparable results across ground-of-truth data and generated data.

Table 3.6 shows results of IS and DS-person metrics for a preliminary evaluation of the quality of the proposed method to generate images of elderly subjects doing exercise in the home environment.

Table 3.6: Inception score (IS) and Detection Score for person class (DS-person) in original and generated dataset.

Dataset	IS	mask-IS	DS-person	mask-DS-person
Original	171.87	181.11	0.844	0.858
Synthetic	162.94	272.14	0.789	0.861

Results of the DS-joint score comparing the original and synthetic dataset in the background and masked background conditions are summarized in Figure 3.10.

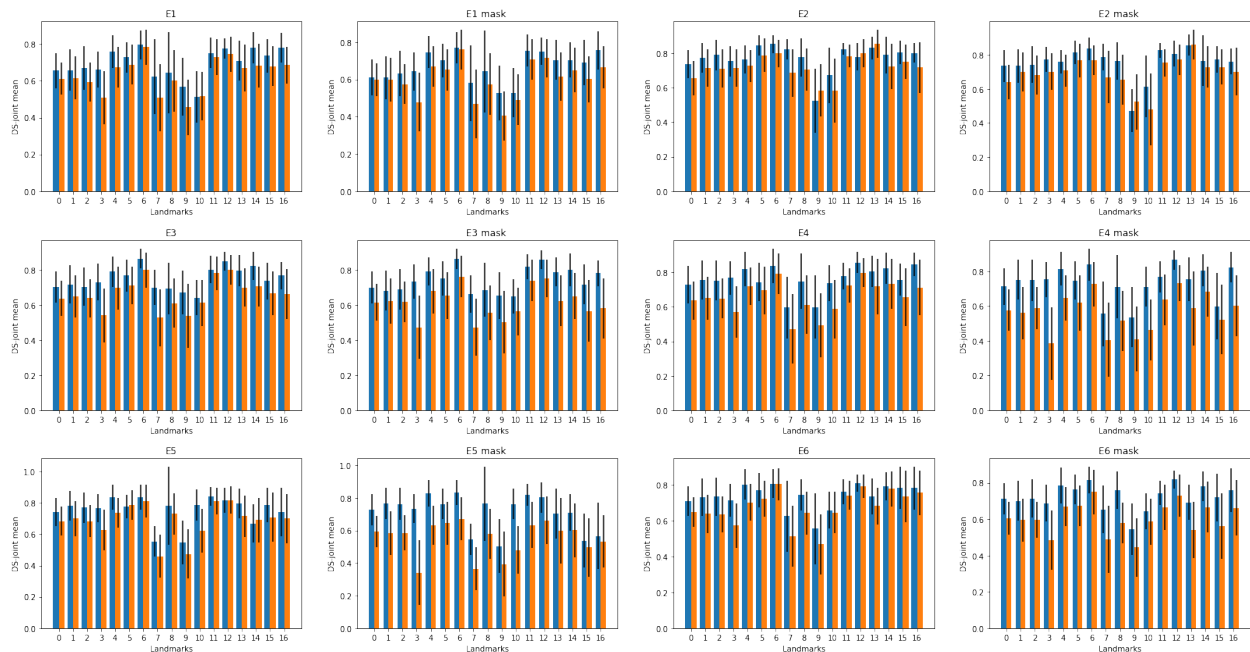


Figure 3.10: Analysis of joint-class Detection Score (DS-joint).

Note. DS-joint comparison between original (blue bar) and synthetic (orange bar) dataset considering background effects (on the left column results with background, on the right column with masked background). Joint classes: nose = 0, left eye = 1, right eye = 2, left ear = 3, right ear = 4, left shoulder = 5, right shoulder = 6, left elbow = 7, right elbow = 8, left wrist = 9, right wrist = 10, left hip = 11, right hip = 12, left knee = 13, right knee = 14, left ankle = 15, right ankle = 16.

Table 3.7 and Table 3.8 provide a detailed presentation of the FID and SSIM metrics, which serve as a comprehensive comparison between the generated images and the ground truth images.

Table 3.7: Structural similarity index (SSIM) and Fréchet Inception distance (FID) calculated between original and generated dataset over each classes of movement in each subset.

Exercise	Class of movement	SSSIM	Camera view 1		
			mask-SSIM	FID	mask-FID
E1	Down	0.545	0.881	299.06	304.35
	Incomplete down	0.517	0.868	276.18	341.26
	Incomplete up	0.530	0.878	318.15	285.45
	Up	0.480	0.876	347.84	319.59
	C1	0.673	0.918	211.85	164.27
	C2	0.578	0.865	234.06	186.02
E2	Down	0.573	0.868	285.27	291.39
	Incomplete down	0.636	0.848	276.44	279.79
	Incomplete up	0.655	0.858	237.79	266.96
	Up	0.598	0.842	302.18	200.64
	C1	0.675	0.860	188.82	179.63
E3	Down	0.567	0.860	289.78	293.04
	Incomplete down	0.539	0.861	254.50	269.25
	Incomplete up	0.598	0.825	290.51	245.26
	Up	0.563	0.838	288.64	267.18
	C1	0.637	0.894	199.60	152.56
	C2	0.662	0.892	214.68	191.89
E4	Down	0.616	0.872	295.02	343.44
	Incomplete down	0.605	0.864	246.47	249.49
	Incomplete up	0.584	0.850	237.93	178.42
	Up	0.583	0.840	262.56	199.50
	C1	0.529	0.783	228.84	285.33
	C2	0.669	0.873	293.39	247.27
E5	Down	0.655	0.840	288.39	277.77
	Incomplete down	0.653	0.848	223.68	272.38
	Incomplete up	0.622	0.823	224.73	286.01
	Up	0.632	0.832	196.09	289.48
	C1	0.671	0.823	169.46	195.25
E6	Down	0.590	0.852	134.88	178.52
	Incomplete down	0.617	0.845	108.09	216.15
	Incomplete up	0.626	0.866	137.30	139.35
	Up	0.622	0.854	117.00	209.41
	C1	0.681	0.826	181.26	145.43

Table 3.8: Structural similarity index (SSIM) and Fréchet Inception distance (FID) calculated between original and generated dataset over each classes of movement in each subset.

Exercise	Class of movement	Camera view 2			
		SSSIM	mask-SSIM	FID	mask-FID
E1	Down	0.569	0.869	179.86	203.63
	Incomplete down	0.630	0.883	159.45	209.16
	Incomplete up	0.597	0.875	223.98	173.52
	Up	0.532	0.889	236.14	181.66
	C1	0.616	0.901	228.14	172.92
	C2	0.628	0.880	258.54	181.92
E2	Down	0.581	0.832	248.68	213.93
	Incomplete down	0.609	0.812	224.81	242.18
	Incomplete up	0.634	0.825	207.32	223.47
	Up	0.667	0.849	280.37	153.19
	C1	0.608	0.833	196.27	228.52
E3	Down	0.597	0.875	145.20	188.74
	Incomplete down	0.539	0.869	129.79	152.17
	Incomplete up	0.584	0.792	159.71	171.14
	Up	0.542	0.814	179.57	164.75
	C1	0.658	0.896	178.51	140.31
	C2	0.628	0.892	184.94	165.86
E4	Down	0.553	0.879	194.66	192.39
	Incomplete down	0.555	0.872	141.55	172.23
	Incomplete up	0.512	0.836	142.72	207.51
	Up	0.512	0.838	249.03	220.17
	C1	0.568	0.771	256.70	318.58
	C2	0.654	0.844	259.71	257.03
E5	Down	0.613	0.833	252.29	196.09
	Incomplete down	0.663	0.835	228.59	189.81
	Incomplete up	0.634	0.827	233.59	189.80
	Up	0.602	0.828	238.85	192.68
	C1	0.666	0.821	181.17	205.67
E6	Down	0.632	0.839	99.75	195.62
	Incomplete down	0.641	0.817	171.73	176.82
	Incomplete up	0.623	0.808	188.43	203.07
	Up	0.593	0.808	172.28	192.36
	C1	0.635	0.841	180.93	224.76

Both the original and synthetic datasets achieved high IS, indicating a commendable overall quality of the generated images (see Table 3.6). Additionally, the DS-person and DS-joint scores exhibit similar performance between the original and synthetic datasets, suggesting that the proposed method successfully captures the key characteristics of indi-

viduals' movements and joint relationships. However, upon isolating the foreground person and applying a masked background, the IS score is notably higher in the generated dataset. This indicates a greater diversity in the generated individuals, showcasing the versatility and variability introduced by the synthesis process. Importantly, the results of DS-joint metrics indicate no significantly negative effects from the background, reaffirming the method's ability to preserve the spatial consistency of joint relationships (see Figure 3.10). Conversely, the DS-person score is higher, suggesting the potential influence of background disturbances on person detection in the generated images (see Table 3.6).

The FID score and SSIM scores were utilized to compare the generated images with their ground truth counterparts, enabling an assessment of the quality and similarity between the two datasets. These metrics offer valuable insights for each class of movement within various subsets of both the original and the generated dataset. The FID score offers insights into the overall distributional similarity, while the SSIM score focuses on perceptual similarity in terms of structural details. Together, these metrics provide a comprehensive evaluation of the fidelity, quality, and visual resemblance of the generated images to the ground truth images. A detailed report of these metrics can be found in Table 3.7, enabling a quantitative assessment of the images within each class of movement across the different rehabilitation exercises. Additionally, the effect of the camera view perspective on image quality and diversity is considered. Notably, the synthesis of images from 'camera view 1' and 'camera view 2' perspectives did not exhibit significant differences in the quantitative results. This finding demonstrates the robustness of the proposed method in generating foreground persons across different camera perspectives, including frontal, lateral perpendicular, and lateral oblique positions. Furthermore, the method proves effective in scenarios where the individual is in an upright or seated position. Considering that the proposed method aimed to modify both the background and foreground person, the evaluation also included an assessment of the background's effects on image quality and diversity. Overall, when observing the results of the SSIM scores for images with the presence of background versus those with masked backgrounds, it can be inferred that the background contributes to the greater diversity in the synthetic images. In contrast, when evaluating the SSIM score with masked backgrounds, the similarity index of the synthesized individual significantly exceeds the mean of 0.80. This indicates a remarkably close similarity between the original and synthesized foreground persons.

3.2.5 Data validation

Mutual dependencies (inter-class consistency) between motion data of the original and synthetic dataset considering different camera perspectives are observed in heatmap plots of Figure 3.11 and Figure 3.12 with the respective correlation distance measures.

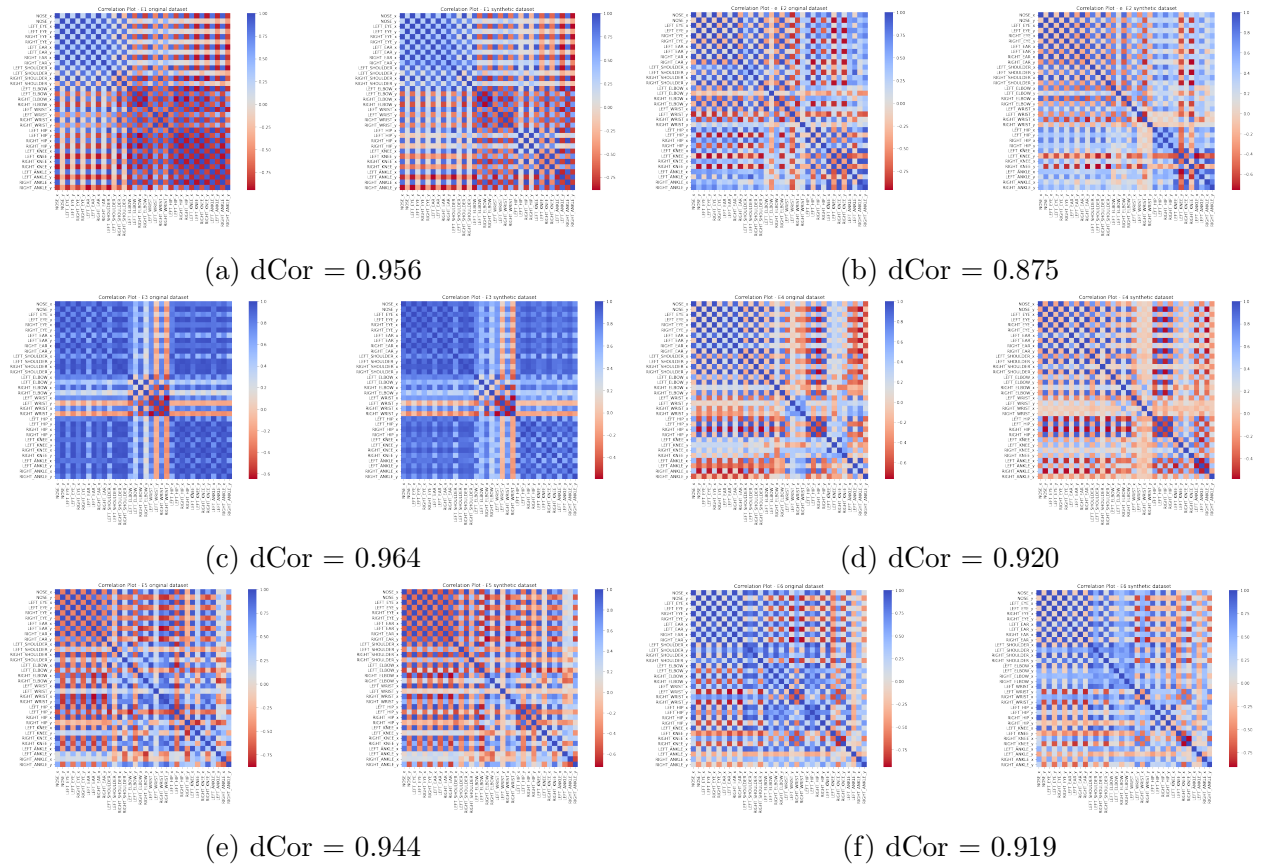


Figure 3.11: Correlation plot and distance correlation metric between original and synthetic dataset with background.

Note. Results considering background and masked background are shown respectively on left and right side of the figure. Notes: (a)=E1 with background vs E1 masked-background; (b)=E2 with background vs E2 masked-background; (c)=E3 with background vs E3 masked-background; (d)=E4 with background vs E4 masked-background; (e)=E5 with background vs E5 masked-background; (f)=E6 with background vs E6 masked-background.

The results demonstrate good comparability between the generated images and the ground truth dataset, showcasing the accurate reflection of spatial consistency in human joint relationships for each exercise movement. Importantly, this evaluation also considers the influence of the background, and the absence of significant negative background effects on motion data is observed. This is evident in Figures 3.11 and 3.12, which compare the

original and generated datasets with and without the background.

The preservation of mutual information is further validated by a distance correlation metric, with a value above the mean of 0.90 between the original and generated motion data. These findings indicate that the synthetic data successfully maintains the mutual information of motion data, effectively capturing the essence of the original movements.

Moreover, the analysis also considered inter-class consistency, which includes an assessment of class boundary preservation of motion data. Movement class boundaries of the original and the generated synthetic datasets are shown in t-SNE plots of Figure 3.13. The plots showed the preservation of boundaries between various movement classes in the synthetic dataset when compared to the real dataset. Furthermore, the plots revealed a heightened variability of information within each class within the generated dataset.

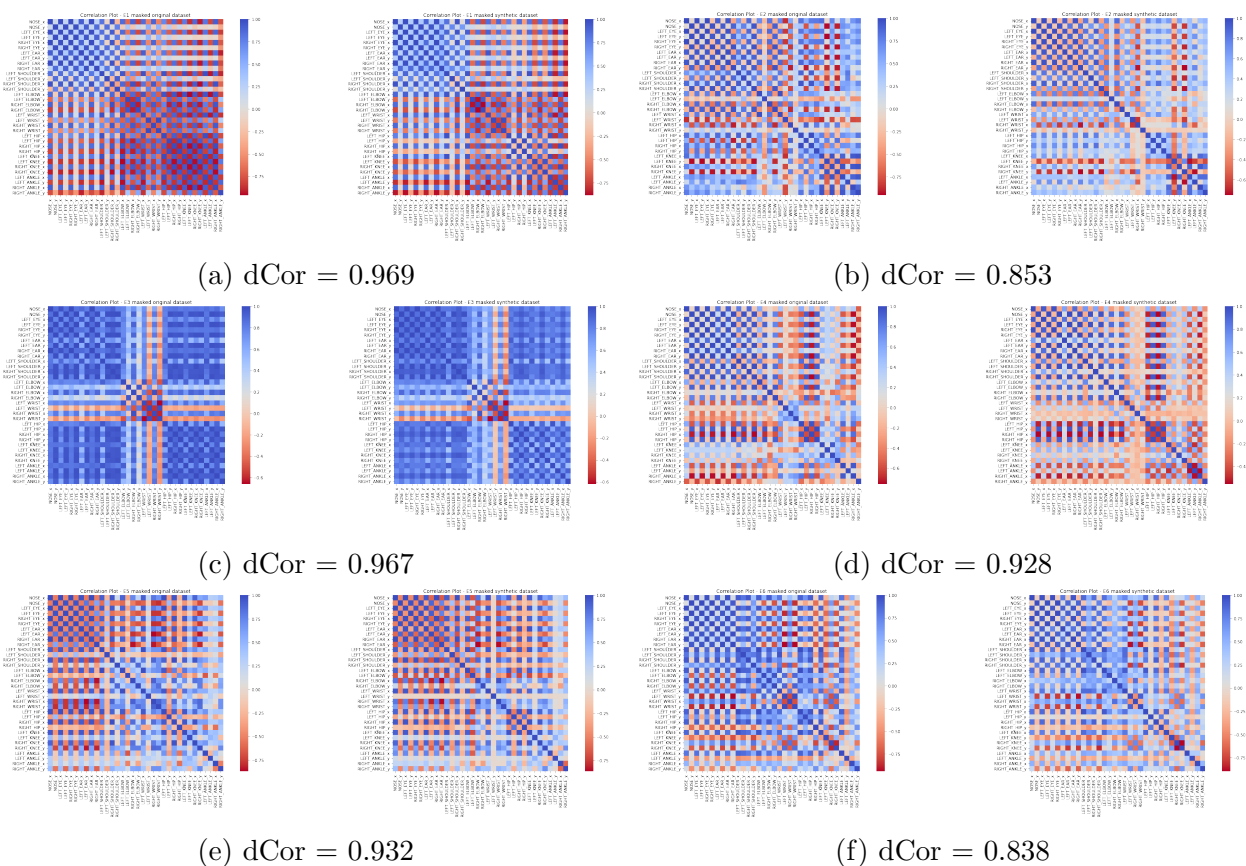


Figure 3.12: Correlation plot and distance correlation metric between original and synthetic dataset with masked background.

Note. Results considering background and masked background are shown respectively on left and right side of the figure. Notes: (a)=E1 with background vs E1 masked-background; (b)=E2 with background vs E2 masked-background; (c)=E3 with background vs E3 masked-background; (d)=E4 with background vs E4 masked-background; (e)=E5 with background vs E5 masked-background; (f)=E6 with background vs E6 masked-background.

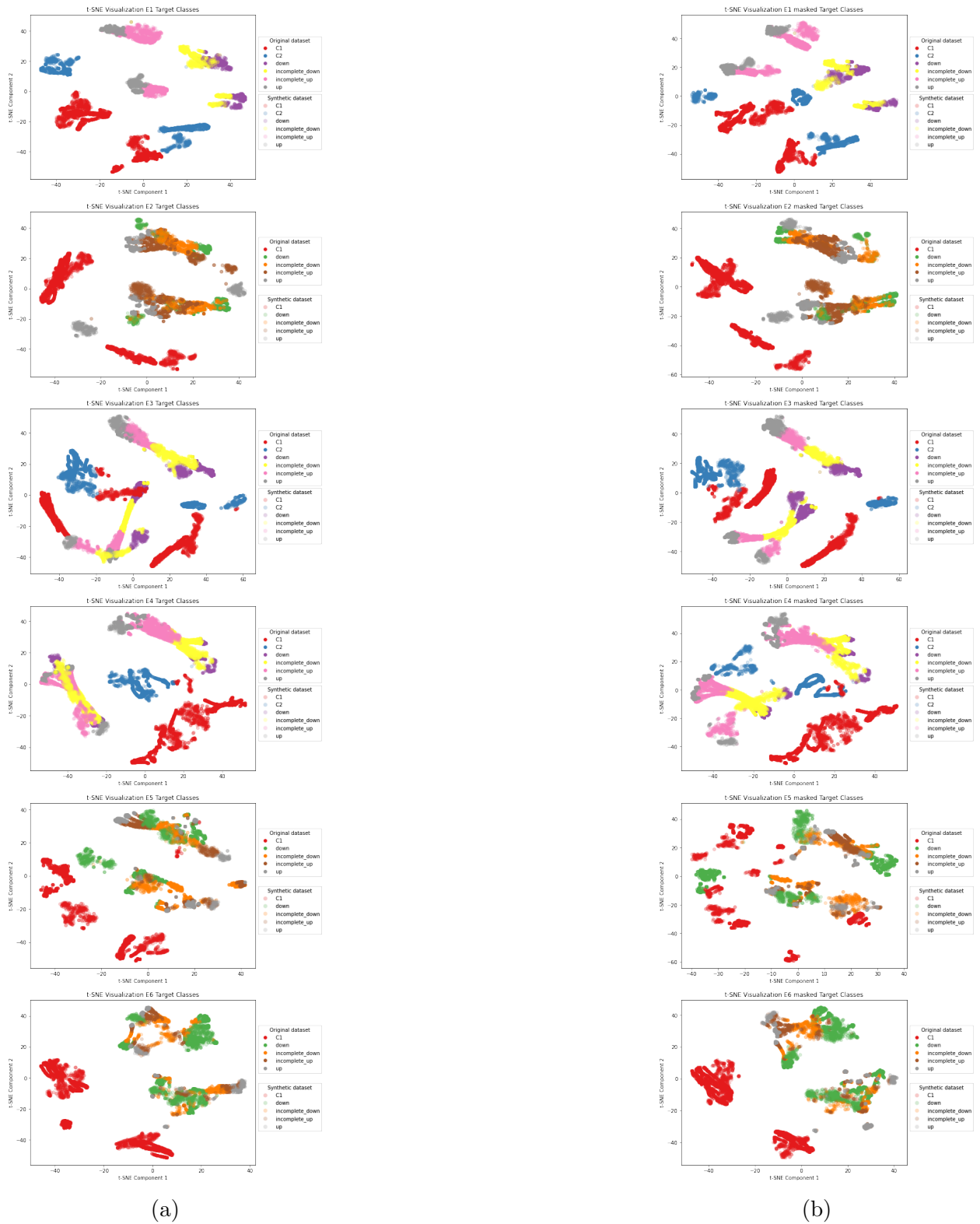


Figure 3.13: t-SNE plots to visualize movement classes boundaries preservation in the synthetic dataset generated from the original dataset.

Note. Results considering background and masked background are shown respectively on left (a) and right (b) side of the figure.

Chapter 4

AI-based monitoring and assistive system: a proof-of-concept

4.1 General overview

The emergence of digital biofeedback systems marks a new era in rehabilitation technologies, aiming to provide remote guidance and support to patients during their rehabilitation journey [101, 126]. Such guidance has been proven to improve the likelihood of patients performing exercises correctly. Nevertheless, achieving real-time motion assessment and live feedback regarding the quality of exercises remains an ongoing challenge. To address the complexities of both home-based and in-clinic rehabilitation programs, the development of reliable systems capable of capturing human movements, automatically analyzing recorded data, and evaluating the quality of movement performance is crucial.

Advancements in low-cost sensors integrated with motion tracking functionality offer a promising avenue for the development of such systems. Additionally, the creation of efficient computational algorithms for modeling and analyzing human motion becomes central to resolving the challenge of evaluating the quality of rehabilitation movements.

Assistive systems play a vital role in supporting patients with in-home exercises [291, 102]. To be effective, assistive systems need to be adequate, affordable, and easily accessible, while also ensuring user engagement [266]. Furthermore, it is essential for these systems to assess patients' performance and provide therapists with the necessary data to track progress and make informed clinical decisions [234, 338].

The present Section introduces an innovative, cost-effective application designed to automatically evaluate rehabilitation exercises within prescribed routines. By unobtrusively monitoring a person's exercises during therapeutic sessions, the system assesses their perfor-

mances to prescribed motions, distinguishing between correct and incorrect movements. The primary contribution lies in the development of a proof-of-concept system capable of analyzing human motion using AI, providing valuable support to patients without the constant need for supervision during rehabilitation.

The proposed technological solution composed of a single laptop with a built-in webcam aims to achieve three main objectives: first, gathering comprehensive data about human motion through exercise monitoring [26, 196]; second, evaluating the quality of exercise performance in terms of motion offering corrective feedback to patients [249, 95, 191, 160].

The proposed system relies on the architecture illustrated in Figure 4.1, which comprises five main modules: (i) a 2D human pose network, ii) data pre-processing; (iii) a ROM-classifier module, (iv) a compensation-classifier module, and v) a repetition counting and validation module to analyze ROM and compensatory patterns.

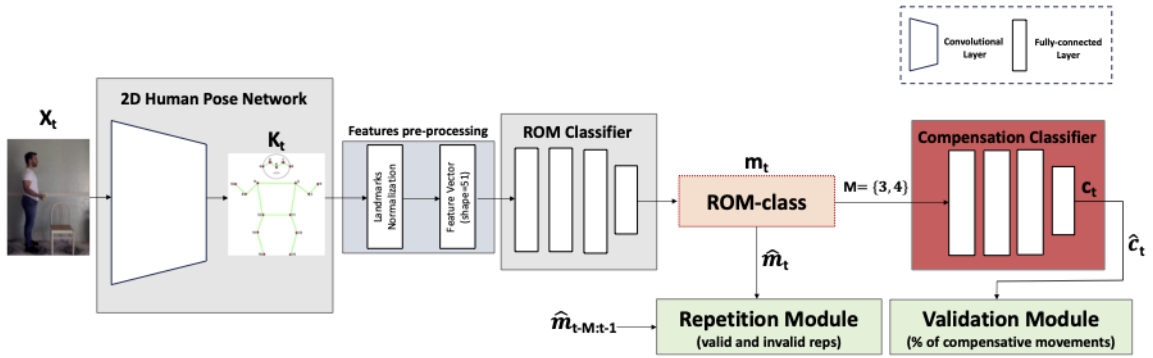


Figure 4.1: Proposal of a system that counts valid and invalid repetitions and recognize compensatory motion patterns of the rehabilitation exercises.

Note. The architecture is composed of five different modules: (i) a 2D human pose network, ii) data pre-processing module, (iii) a ROM-classifier module, (iv) a compensation-classifier module, v) a repetition counting and validation module. $m_t \in C = \{C_j | 1 \leq j \leq 4\}$ represents ROM classes of a given exercise comprising key poses or phases of the exercise. $c_t \in C = \{C_j | 0 \leq j \leq 2\}$ denotes whether the t frame belongs either to a physiological or compensatory motion pattern class.

Throughout the following, let $X(z) = \{X_t, m_t, c_t\}$ denote a labeled sequence of N frames of the considered set of exercises $z \in \{E1, E2, E3, E4, E5, E6\}$, where X_t denotes the acquired RGB image at frame t , m_t represents the current ROM class of the exercise execution, and c_t indicates the class of motion corresponding to physiological or compensatory movement patterns. First, the acquired image X_t is fed through the 2D human pose network to obtain a set of K keypoints coordinates. Second, a ROM-classifier operates on top of skeleton-based features to predict the ROM motion category. Furthermore, ROM-class 3 and ROM-class 4 of each exercise are evaluated to belong to the good or compensatory

patterns category via the compensation classifier. Finally, the repetition counting module receives the predicted ROM-class which outputs the current number of valid and invalid repetitions via an internal state machine and the validation module tracks the percentage of erroneous patterns present in each exercise repetition. The internal state machine outputs and increments the current number of valid and invalid repetitions if the following conditions are satisfied: (i-a) there is a transition from the last to the first ROM-class of the exercise (valid repetition), (i-b) there is a transition from the intermediate ROM-classes of the exercise (invalid repetition); (ii) all the ROM-classes of the repetition appear in the right order within the current repetition. This module updates the current number of repetitions at each frame t by having the ROM-class prediction of the current frame m^t , as well as the M past ROM-class predictions $m_{t-M:t-1}$. Finally, the system counts those repetitions correctly if all the ROM-classes of the exercise appear in the right order in a given time window (3 past frames, current frame, and 3 future frames) [85].

In summary, the system quantitatively assess motor patterns of rehabilitation exercises through 2D video analysis in real time. The main goal is to monitor unobtrusively the exercise carried out by a person in a therapeutic session and evaluate whether it is being performed according to its prescription (correct performance) or not (wrong performance). The evaluation corresponds to the assessment of the range of motion (ROM) and the compensatory movement patterns that a subject could accomplish during the rehabilitation session.

The system was developed and validated to monitor resistance training exercise described in Section 3.1.3 to prevent physical decline in elderly patients.

The following sections provide a detailed description of the motion monitoring approach implemented in the system (see Section 4.1.1) and the motion evaluation method developed to deliver biofeedback, enabling patients to correct their performance (see Section 4.1.2).

4.1.1 Motion monitoring

Pose estimation refers to computer vision techniques that detect human joints in images and videos. It is important to know pose estimation merely estimates where key body joints are and do not recognize a person's identity in an image or video.

To this scope, a convolutional neural network model that runs on RGB images and predicts the human joint locations of a single person was applied for pose estimation.

After a meticulous evaluation of various state-of-the-art pose estimation models, the pre-trained Google-based inference model MoveNet was adopted for motion monitoring, which provides joint landmarks kinematics. This choice was primarily driven by MoveNet's outstanding accuracy and exceptional performance on edge devices [321]. Further reinforcing

our decision, a recent study compared several skeleton-based human pose estimation models, including OpenPose, PoseNet, MoveNet, and MediaPipe Pose. The study’s findings revealed that MoveNet exhibited the highest performance in detecting various human poses in static images and videos [59].

MoveNet is a bottom-up estimation model that uses heatmaps to localize human key points. The model is designed to be run in the browser using Tensorflow.js or on devices using TF Lite in real time. The variant MoveNet Thunder resulting in a higher capacity model that performs better prediction quality while still achieving high real-time speed was selected [59].

The MoveNet Thunder version architecture takes in input an RGB frame of video or an image with a resolution of 256x256. MoveNet output consists of lists containing the coordinate predictions of COCO’s standard 17 keypoint locations, and their corresponding prediction confidence (see Figure 4.2).

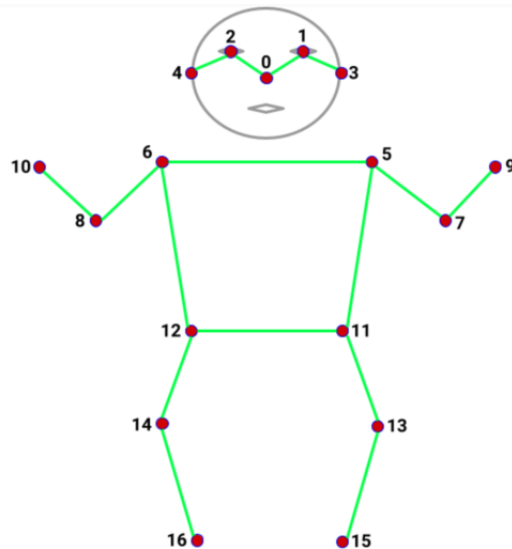


Figure 4.2: COCO Keypoints used in MoveNet.

Note. Keypoints list: nose = 0, left eye = 1, right eye = 2, left ear = 3, right ear = 4, left shoulder = 5, right shoulder = 6, left elbow = 7, right elbow = 8, left wrist = 9, right wrist = 10, left hip = 11, right hip = 12, left knee = 13, right knee = 14, left ankle = 15, right ankle = 16.

MoveNet was trained on two datasets: COCO and an internal Google dataset called Active. While COCO is the standard benchmark dataset for general detection, the Active dataset containing yoga, fitness, and dance videos from YouTube makes MoveNet suitable for physical activity and exercise monitoring. The architecture consists of a feature extractor and a set of prediction heads that inference the landmarks of the 17 body joints.

The feature extractor in MoveNet is MobileNetV2 [277] with an attached feature pyramid network, which allows for a high resolution (output stride 4), semantically rich feature map output. There are four prediction heads attached to the feature extractor, responsible for densely predicting: i) person center heatmap: predicts the geometric center of person instances; ii) keypoint regression field: predicts a full set of keypoints for a person, used for grouping keypoints into instances; iii) person keypoint heatmap: predicts the location of all keypoints, independent of person instances; iv) 2D per-keypoint offset field: predicts local offsets from each output feature map pixel to the precise sub-pixel location of each keypoint.

4.1.2 Motion analysis

The system aims to evaluate the performance quality in two different stages of classification. The proposal divides the complete process of determining the completeness of ROM and the recognition of compensatory movements into two different stages of classification. It is based on the hypothesis that separating the evaluations into two different stages, both classifications would improve their accuracy rates due to the reduction of the number of classes.

Formally, were considered a set of input–output pairs $D = \{(x_i, y_i)\}_{i=1}^N$ where $x_i \in R^n$ are the N samples of the input feature space obtained from the 2D video motion analysis and $y_i \in C = \{C_j | 1 \leq j \leq J\}$ are the class to which these features correspond. The number of classes $J \in N$ depends on each proposal. The classification algorithms look for a decision function $f : R^n \rightarrow C$ which given a sample, that in this work contains features of body joints' 2D pose data, determines the output class that includes the kind of execution performed: $x \rightarrow y = f(x, w)$. During the training process, the algorithm finds the parameters w that best fit the given training data set. These methods aim to find a function f capable of generalizing its good accuracy to the given new data, which corresponds to a person's motion features derived from a 2D pose estimation module.

In summary, to assess a patient's exercise performance from 2D videos, the methodology is composed of the following steps: body keypoint extraction, data normalization, and classification. A neural network-based approach was investigated to classify the correctness of the prescribed exercise. The classifier must be robust enough to assign a label to each frame denoting the wrong performance and indicate good movement quality (normal movement patterns without erroneous motion categories such as limited ROM executions or compensatory movements).

Two different models that take the preprocessed landmark coordinates as input and predict the pose class that the person performs were addressed respectively to assess the ROM

execution and compensatory movements for each of exercises. To identify multiple motion patterns from video frames, a multilabel classification problem is addressed. First, a multilabel classifier assesses the ROM execution to distinguish correctly performed series from those performed erroneously. Simultaneously, a second multilabel classifier detects the presence of compensatory movements by identifying the patterns associated with compensation.

4.2 Experimental study

4.2.1 Data pre-processing

The selected pose estimation module MoveNet Thunder was used to extract the body joints' pose data from each frame of the dataset.

The 2D position of 17 body keypoints (body skeleton) in the image coordinate system $\{I\}$ (Figure 4.2). Each keypoint provided is denoted by $o_k^t = [p_k^t z_k^t]'$ = $[x_k^t y_k^t z_k^t]'$. Here, $p_k^t = [x_k^t y_k^t]'$ denotes the transposed vector of 2D coordinates in the image of a body keypoint k from a set of joints K , t is the frame number, and z_k^t is a confidence score of keypoint detection. A moving average filter with a window of five frames was applied to reduce the noise and measurement errors of the pose estimation module in keypoints inferences [62, 171].

In a real-world setting, subjects have body parts of different sizes and occupy different locations regarding the camera. Accordingly, landmarks were normalized and scaled to a constant pose size based on the torso's length size and the maximum distance of a landmark from the pose center. Each keypoint was transformed from the image coordinate system $\{I\}$ to the body coordinate system $\{B\}$. The image coordinate system $\{I\}$ represents the Cartesian pixel coordinate system with its origin (0,0) located in the upper left corner. Conversely, the body coordinate system $\{B\}$ is also a Cartesian pixel coordinate system, but with its origin at the midpoint between the left hip ($k = 11$) and right hip ($k = 12$) ($c^t(p_{11}^t, p_{12}^t)$ see Equation 4.1). The introduction of the body coordinate system $\{B\}$ aims to address variations in subject positions when facing the camera. To further account for differences in body part dimensions, each keypoint coordinate in the $\{B\}$ system was normalized based on the subject's spine length ($d^t(p_0^t, p_c^t)$ see Equation 4.2) measured at each time step t . This method allows the algorithm to generalize with different users' body length measurements, distance from the camera, as well as other relative factors.

$$c^t = \left(\frac{x_{11}^t + x_{12}^t}{2}, \frac{y_{11}^t + y_{12}^t}{2} \right) \quad (4.1)$$

$$d^t = \sqrt{(x_0^t - x_c^t)^2 + (y_0^t - y_c^t)^2} \quad (4.2)$$

4.2.2 Deep learning algorithm

A feedforward fully connected multilayer perceptron (MLP) classification algorithm was implemented in the Keras library to address the movement classification tasks [114]. Some studies recently evidenced that MLP achieves high accuracy in classification movement tasks for evaluation purposes in rehabilitation [31, 278, 192]. Each model was built using Python libraries such as TensorFlow, Keras, NumPy, and Scikit Learn [248, 55].

The comprehensive configuration details of the neural network models used for the ROM classifier and the compensation classifier are summarized in Table 4.1. To determine the optimal configuration, various model architectures were systematically explored, ranging from one to three layers, with hidden units varying from 16 to 512. An adaptive learning rate approach was also employed, using multiple initial learning rate values, including 0.001, 0.005, 0.01, 0.05, and 0.1. This extensive exploration of hyperparameters was conducted using the Keras library, ensuring an optimized configuration for our neural network models.

Table 4.1: MLP model configuration in Keras library.

Model architecture	Parameters / Activation function
Input layer	34
Hidden layer	128 / ReLU
Drop out layer	0.5
Hidden layer	64 / ReLU
Drop out layer	0.5
Output layer	$C = \{C_j 1 \leq j \leq J\}$ / Softmax
Train configuration	Parameters
Epochs	200
Batch size	16
Early stopping	monitor='val accuracy' patience=5
Optimizer	Adam
Loss	Categorical cross entropy
Metrics	Accuracy

The used MLP architecture consists of an input layer of 128 neurons, a hidden layer of 64 neurons, and an output layer (size depending on the number of classes to predict).

Preprocessed landmarks have been flattened in a feature vector representation used to feed the input layer. A batch size of 16 was used to train the model over epochs. Adam algorithm was selected for the model optimization and the early stopping technique was used to stop model training once its performance stops improving on the hold-out validation dataset (15% of the total of the training dataset).

Dropout regularization for reducing overfitting and improving the generalization of the model was used in the input and hidden layers by setting the hyperparameter at 0.5.

Rectified Linear Unit (ReLU) activation functions were used in the input and hidden layer. Softmax activation function was used in the output layer. About the classification task addressed by the model the categorical cross-entropy loss function was selected, and the accuracy metric 4.3 was used to assess the general model's performance.

Also, a set of metrics appropriated to a multi-class problem to evaluate our classification models' performance was adopted: Precision, Recall, and F1 score [118]. Metric Precision 4.4 is the percentage of predicted labels truly significant for the sample. Recall 4.5 expresses the classifier's ability to detect all positive samples. Score F1 4.6 is a weighted harmonic mean Precision and Recall, which measures classification accuracy.

$$acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.3)$$

$$prec = \frac{TP}{TP + FP} \quad (4.4)$$

$$rec = \frac{TP}{TP + FN} \quad (4.5)$$

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (4.6)$$

True positive (TP), true negative (TN), false positive (FP), and false negative (FN) were used in the above equations to denote the performance of the classification model.

The evaluation of performance metrics was conducted on test sets obtained by employing a stratified method to divide the original dataset into train and test subsets. This approach effectively maintains the class distribution from the original dataset in both subsets, making it particularly valuable when dealing with imbalanced datasets, where certain classes may have significantly fewer samples than others.

4.2.3 Model training and evaluation

In deep learning, artificial neural network models need a considerable number of samples for training to obtain satisfactory results, or else there may be overfitting, and the model's generalization ability may drastically decrease. However, for certain actions, it is difficult to obtain a quantity of available and quality samples, so data augmentation methods are used to enlarge the dataset. A data augmentation strategy is proposed to generate new frame samples and prevent the MLP model from overfitting and generalizing in various

configuration settings. The strategy consists of mimicking a person acting in different places and sides in the field of view and at different distances from the camera view. The Keras ‘Image Data Generator’ library was used to apply transformations to obtain augmented data [55]. Horizontal and vertical shifts were applied to simulate an action performed in different places of view. A horizontal flip was applied to simulate the exercise also performed with the left body side and not only performed by the right side as obtained in the video dataset. A range for zooming in and out each frame of the dataset was applied to simulate the various distances from the camera that the subject could accomplish. The parameters of the arguments of each transformation are synthesized in Table 4.2.

Table 4.2: Data augmentation parameters used in Keras library.

Argument	Parameter value
width shift range	0.2
height shift range	0.2
zoom range	[0.9,1.8]
horizontal flip	True
fill mode	‘nearest’

Noteworthy is to highlight that each exercise dataset was split into train and test sets and therefore data augmentation was applied on each split set, including the test dataset. The dataset was splitted in train and test sets using a stratified method, which effectively preserves the class distribution from the original dataset in both subsets. This approach enhances the model’s ability to generalize well to new, unseen data and ultimately contributes to the robustness of our findings. In this way, the model’s performance was evaluated taking into account both the original test dataset (not augmented), containing the frames acquired in an optimal setup configuration, and a generated test dataset containing the augmented frame used to simulate a real application that no expects setup configurations. This method allows us to assess whether setup configurations should be necessary to consider in the development process of the system that integrates the proposed system to monitor and assess rehabilitation exercises at home.

Some of the applied transformations introduced some measurement errors during the inference of the 2D human pose network so an outlier detection method was applied to the train dataset before compiling the model. Considering the multivariate characteristics of the dataset a principal component analysis (PCA) was used to reduce the dimensionality of the data and the Hotelling T^2 test was applied to detect outliers across the multidimensional space of PCA [87, 86, 307]. The Hotelling T^2 computes the chi-square tests and P-values

(level of significance fixed at $\alpha < 0.05$) across the top five PCA components which allows to determine the detection of outliers with its ranking (strongest to weak). A visual example of the outlier detection algorithm is shown in Figure 4.3.

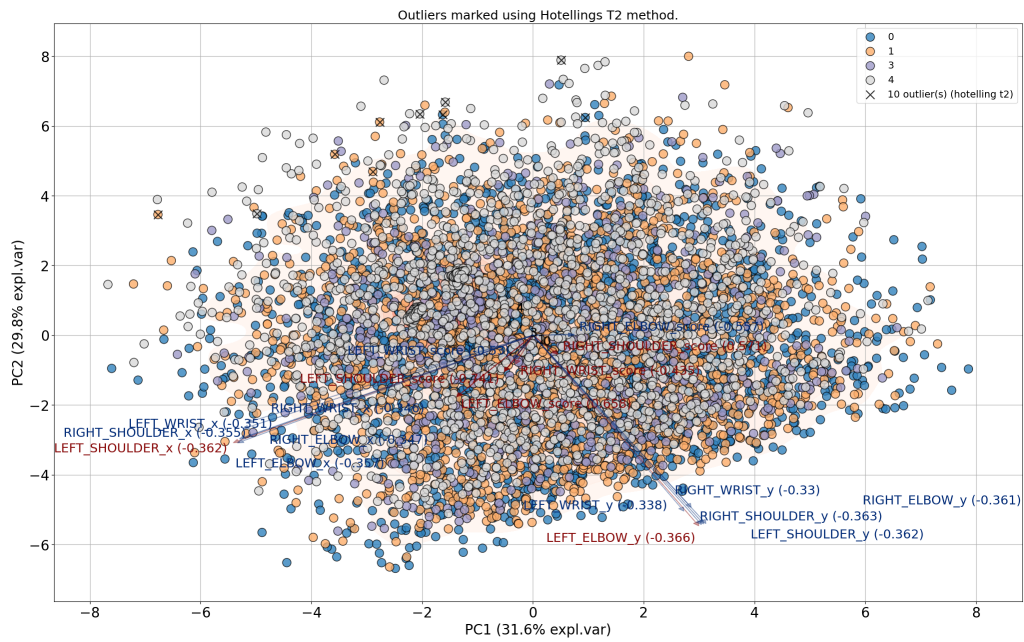


Figure 4.3: Outliers detection plot using Hotelling T^2 statistics.

Note. The figure reports an example of outlier detection in ROM classes of the E1 dataset.

The overall characteristics of data sets used to train the ROM classifier and the compensation movement classifier are presented respectively in Table 4.3 and Table 4.4.

Table 4.3: Datasets characteristics in terms of number of samples for ROM classifier training..

Exercise	ROM class	Data original		Data augmented	
		Train	Test	Train	Test
E1	Down (M=1)	439	109	2634	654 9
	Incomplete down (M=2)	369	92	2214	552
	Incomplete up (M=3)	530	132	3180	792
	Up (M=4)	316	78	1896	468
E2	Down (M=1)	378	94	2268	564
	Incomplete down (M=2)	333	83	1998	498
	Incomplete up (M=3)	620	155	3720	930
	Up (M=4)	452	113	2712	678
E3	Down (M=1)	465	116	2790	696
	Incomplete down (M=2)	598	149	3588	894
	Incomplete up (M=3)	432	107	2592	642
	Up (M=4)	410	102	2460	612
E4	Up (M=1)	369	92	2214	552
	Incomplete up (M=2)	728	181	4368	1086
	Incomplete down (M=3)	723	180	4338	1080
	Down (M=4)	240	60	1440	360
E5	Down (M=1)	715	178	4290	1068
	Incomplete down (M=2)	411	102	2466	612
	Incomplete up (M=3)	311	77	1866	462
	Up (M=4)	252	62	1512	372
E6	Down (M=1)	900	225	5400	1350
	Incomplete down (M=2)	301	75	1806	450
	Incomplete up (M=3)	252	63	1512	378
	Up (M=4)	220	54	1320	324

Table 4.4: Datasets characteristics in terms of number of samples for Compensation classifier training.

Exercise	Compensation class	Data original		Data augmented	
		Train	Test	Train	Test
E1	C1 (shoulder flexion)	903	225	5418	1350
	C2 (trunk extension)	659	164	3954	984
	Good patterns	845	211	5070	1266
E2	C1 (shoulder abduction)	878	219	5268	1314
	Good patterns	1075	268	6450	1608
E3	C1 (trunk extension)	984	245	5904	1470
	C2 (elbow flexion)	734	183	4404	1098
	Good patterns	841	210	5046	1260
E4	C1 (trunk flexion)	927	231	5562	1386
	C2 (knees displacement)	511	127	3066	762
	Good patterns	963	240	5778	1440
E5	C1 (hip flexion)	800	200	4800	1200
	Good patterns	571	142	3426	852
E6	C1 (hip extension)	913	228	5478	1368
	Good patterns	460	115	2760	690

4.2.4 Results

The proposed MLP classification algorithm used to evaluate motion execution showed good performance in classifying both ROM classes and also compensatory movements.

Classification results obtained for the ROM classifier and the compensation classifier are reported in Table 4.5 in terms of accuracy, precision, recall, and F1-score.

Table 4.5: Classification results in terms of accuracy, precision, recall, and F1-score obtained for the ROM classifiers and the compensation classifiers.

Model performance	Test set	ROM classifier				Compensation-classifier			
		<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
E1	Original	0.94	0.95	0.93	0.94	0.97	0.98	0.97	0.97
	Augmented	0.90	0.93	0.88	0.89	0.95	0.96	0.94	0.94
E2	Original	0.85	0.85	0.85	0.85	1.00	1.00	1.00	1.00
	Augmented	0.76	0.79	0.75	0.75	0.97	0.98	0.97	0.97
E3	Original	0.95	0.96	0.95	0.95	0.95	0.95	0.95	0.95
	Augmented	0.93	0.94	0.93	0.93	0.90	0.91	0.90	0.90
E4	Original	0.88	0.87	0.91	0.88	0.99	0.99	0.99	0.99
	Augmented	0.86	0.84	0.87	0.85	0.98	0.98	0.96	0.97
E5	Original	0.86	0.81	0.81	0.81	1.00	1.00	1.00	1.00
	Augmented	0.81	0.76	0.78	0.76	0.99	0.99	0.99	0.99
E6	Original	0.89	0.86	0.86	0.86	1.00	1.00	1.00	1.00
	Augmented	0.87	0.83	0.81	0.81	0.98	0.98	0.97	0.98

Classification performances of each model over ROM-classes and compensation movements categories are respectively reported in detail in Table 4.6 and Table 4.7.

To better understand each model’s performance on each exercise, incorrect samples predicted by the ROM classifiers and compensation classifiers are plotted respectively in Figure 4.4 and Figure 4.5.

4.3 Comparative analysis with related works

Providing qualitative and quantitative evaluation of rehabilitation exercises to patients is fundamental for supporting the effective implementation of at-home rehabilitation programs.

The advances in machine learning and computer vision techniques have inspired an increased interest in the automated evaluation of rehabilitation exercises. Despite the progress, there are still open questions and numerous challenges to overcome before the broad deployment of these systems in home-based and in-clinic settings.

An innovative deep learning-based system for monitoring and assessing resistance training exercises has been proposed in this research thesis.

Table 4.6: Classification results in terms of accuracy, precision, recall, and F1-score obtained for each class of the ROM classifiers.

Model performance	ROM class	Test original			Test augmented		
		<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
E1	Down (M=1)	0.91	0.97	0.94	0.89	0.98	0.93
	Incomplete down (M=2)	0.96	0.89	0.93	0.98	0.86	0.92
	Incomplete up (M=3)	0.92	1.00	0.96	0.84	1.00	0.91
	Up (M=4)	1.00	0.86	0.92	0.99	0.68	0.81
E2	Down (M=1)	0.84	0.96	0.90	0.82	0.95	0.88
	Incomplete down (M=2)	0.76	0.73	0.75	0.72	0.63	0.67
	Incomplete up (M=3)	0.84	0.87	0.86	0.68	0.87	0.77
	Up (M=4)	0.95	0.82	0.88	0.93	0.54	0.68
E3	Down (M=1)	0.94	1.00	0.97	0.92	0.96	0.94
	Incomplete down (M=2)	0.93	0.95	0.94	0.91	0.93	0.92
	Incomplete up (M=3)	0.96	0.90	0.93	0.92	0.91	0.92
	Up (M=4)	1.00	0.96	0.98	0.99	0.93	0.96
E4	Up (M=1)	0.84	0.90	0.91	0.79	0.99	0.88
	Incomplete up (M=2)	0.99	0.78	0.87	0.96	0.79	0.87
	Incomplete down (M=3)	0.86	0.91	0.89	0.87	0.88	0.87
	Down (M=4)	0.79	0.95	0.86	0.74	0.80	0.77
E5	Down (M=1)	0.99	0.97	0.98	0.99	0.92	0.95
	Incomplete down (M=2)	0.87	0.86	0.87	0.76	0.81	0.78
	Incomplete up (M=3)	0.64	0.71	0.67	0.60	0.50	0.55
	Up (M=4)	0.75	0.71	0.73	0.67	0.87	0.76
E6	Down (M=1)	0.95	0.94	0.95	0.93	0.97	0.95
	Incomplete down (M=2)	0.74	0.81	0.78	0.79	0.72	0.76
	Incomplete up (M=3)	0.86	0.76	0.86	0.83	0.63	0.71
	Up (M=4)	0.88	0.93	0.90	0.76	0.91	0.83

Table 4.7: Classification results in terms of accuracy, precision, recall, and F1-score obtained for each category of the compensation classifiers.

Model performance	Compensation class	Test original			Test augmented		
		<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
E1	C1	0.94	1.00	0.97	0.88	1.00	0.94
	C2	1.00	0.91	0.95	1.00	0.82	0.90
	Good patterns	1.00	1.00	1.00	1.00	1.00	1.00
E2	C1	1.00	1.00	1.00	1.00	0.94	0.97
	Good patterns	1.00	1.00	1.00	0.95	1.00	0.98
E3	C1	0.98	0.96	0.97	0.98	0.87	0.92
	C2	1.00	0.88	0.94	0.95	0.87	0.91
	Good patterns	0.89	1.00	0.94	0.80	0.96	0.87
E4	C1	0.98	1.00	0.99	0.94	1.00	0.97
	C2	1.00	0.97	0.98	1.00	0.89	0.94
	Good patterns	1.00	1.00	1.00	1.00	1.00	1.00
E5	C1)	1.00	0.99	1.00	1.00	0.99	1.00
	Good patterns	0.99	1.00	1.00	0.99	1.00	0.99
E6	C1	1.00	1.00	1.00	0.97	1.00	0.99
	Good patterns	1.00	1.00	1.00	1.00	0.94	0.97

The results of each stage of this method were individually evaluated: first, the ones corresponding to the exercises ROM-class recognition and then, those results obtained during

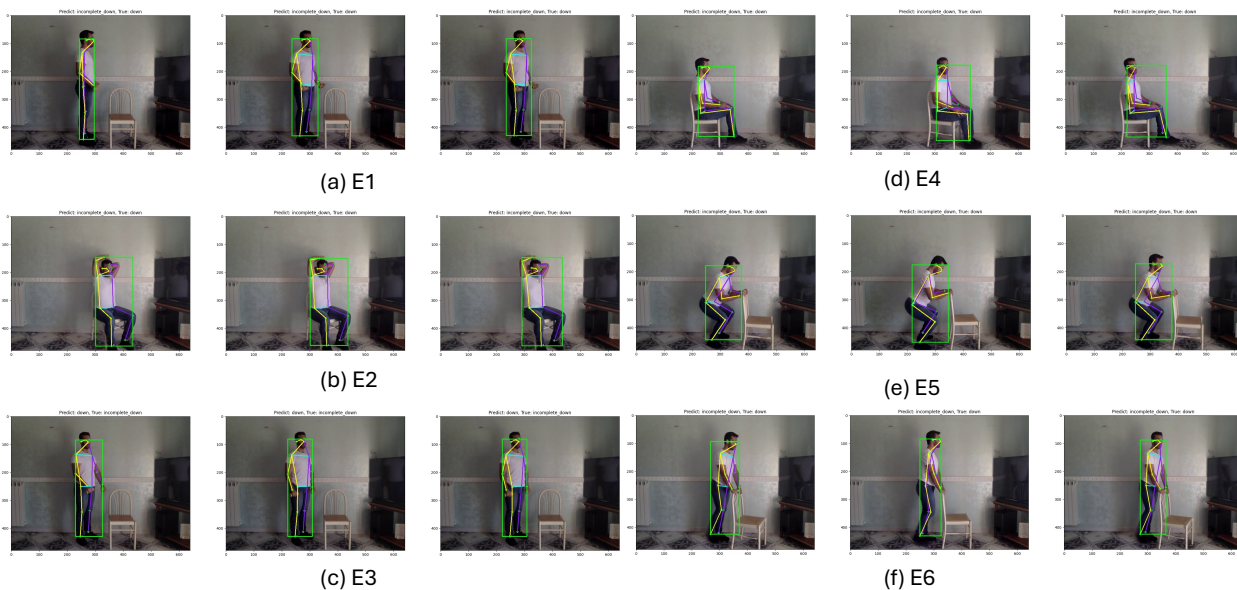


Figure 4.4: Incorrect classes predicted by the ROM classifier for each exercise.

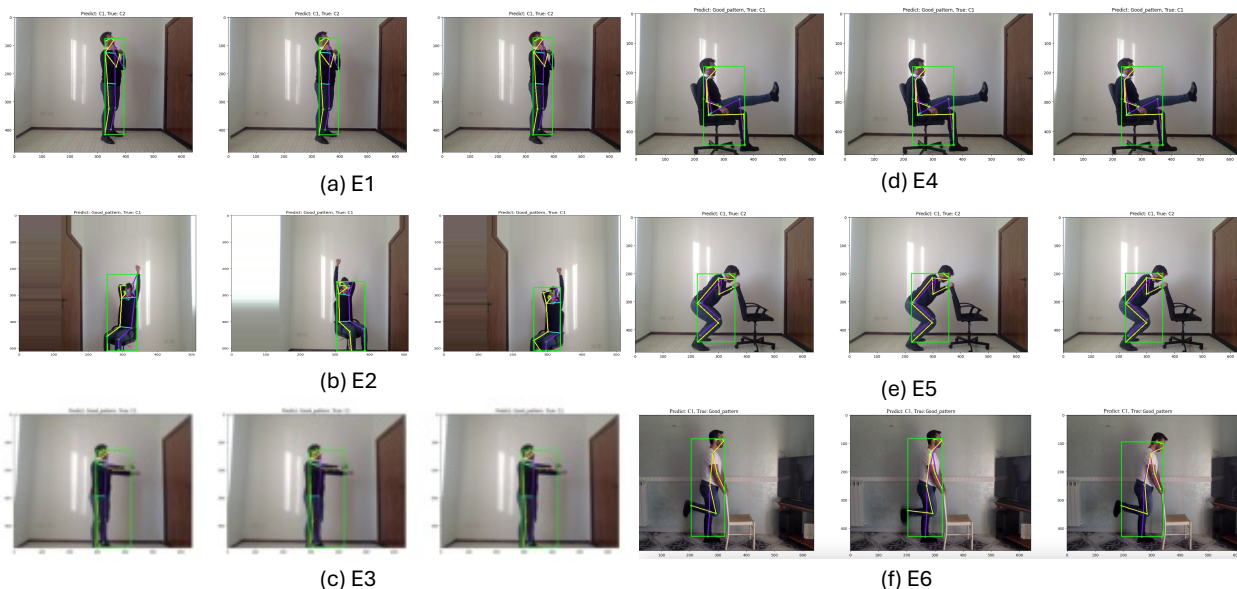


Figure 4.5: Incorrect classes predicted by the Compensation classifier for each exercise.

the compensation evaluation. The proposed method showed high precision to recognize both ROM classes and compensatory movement patterns, following recent findings evidencing the high performance of the MLP model in classification tasks applied to the movement analysis [31, 278, 192]. Concerning its metrics, most models achieved accuracy, F1-score, precision, and recall above 90% when the best setup configuration is provided for the exercise monitoring (compare the results between test original and test augmented in Table 4.5).

The results presented in Table 4.5 demonstrate that our approach achieves comparable

performance to state-of-the-art discrete movement classification methods used for exercise evaluation. For instance, in existing literature, certain approaches have shown high accuracy in classifying movements into correct or erroneous execution classes [326, 271, 139].

For example, in the work by [326], the authors achieved an accuracy above 90% in classifying error movements using a support vector machine (SVM) algorithm. Similarly, in [271], authors used a Hidden-Markov Model approach to classify correct and incorrect movements with an accuracy of 92%. In [139], a k-nearest neighbors classifier was applied for exercise classification after filtering the data noise and applying dimensionality reduction, resulting in a classification accuracy of 96%.

Some other works focus on evaluating the phase of each exercise to monitor subjects' performance in a detailed and highly accurate manner, as proposed in our study [37, 301]. In [37], the authors achieved an overall accuracy of 87% in classifying segmented motions occurring during a shoulder rehabilitation exercise. Moreover, [301] utilized multi-label classifiers to detect subtle errors in exercise performances, achieving a sensitivity of 75%, specificity of 90%, and accuracy of 80%.

Our results demonstrate that our approach is competitive with these state-of-the-art methods in accurately classifying exercise movements, which is crucial for effective rehabilitation monitoring.

The analysis reveals consistent differences in model performance across exercises. Table 4.5 shows that exercises involving more distinct motion classes achieve superior metrics in ROM-class classification. For instance, the ROM classifiers for E1 and E3 exercises exhibit accuracy levels above 95%, primarily attributed to the substantial variations in motion features observed within each movement category. This finding is also supported by the models' performance in recognizing compensatory movements, where little substantial difference exists between exercises, as shown in Table 4.7. The high accuracy in distinguishing between correct and compensatory performances for each exercise is attributed to the markedly distinct features they exhibit. Conversely, exercises with minor differences in motion features between movement classes show poorer classification performance. Consequently, consecutive ROM classes in these exercises are closer together compared to exercises with well-distinguished motion classes [69]. A clear illustration of this can be seen in the varying metrics of the ROM classifier for exercises E5 and E6, as presented in Table 4.6.

Moreover, the ROM classifier's classification metrics might be negatively impacted by partially occluded and incorrectly detected body joints, as observed in exercises E2, E5, and E6 (Figure 4.4). To address this concern, a mean average filter with a five-frame window was applied, effectively smoothing the data and producing more reliable results.

Additionally, the system demonstrates remarkable generalizability across various condi-

tions, successfully simulating patients placed at different angles and distances in front of the camera. Nevertheless, to ensure precise evaluations of exercises E2, E5, and E6, which displayed decreased performance in the augmented dataset compared to the optimal condition, fine-tuning the patient's configuration in front of the camera becomes critical. This setup will significantly improve result reliability when assessing these specific exercises.

The proposed system presents two major advantages for future applications: one from the technical side and the other from the side of the exercise assessment.

The used 2D pose estimation network is an accurate model that runs well on hardware accelerators supported by TensorFlow Lite, including CPU and GPU. MoveNet is a model deployable on edge device running in real-time (30+ FPS) on most modern desktops, laptops, and phones, which proves crucial for live fitness, sports, and health applications. Also, the MLP model introduced for movement classification is easier to develop, not requiring preliminarily feature engineering like the other machine learning methods, and to adopt in smart assistant applications due to its relatively low computational costs. All mentioned characteristics make the proposed deep learning pose estimation and classification approach very useful to develop low-cost and smart applications for healthcare purposes such as the monitoring and assessment of rehabilitation exercises in home-based settings.

The other relevant aspect of the proposed methods is related to their applicability in clinical practice. The proposed method can promptly provide feedback regarding an individual's performance during exercise, simulating the routine activities of a rehabilitation expert. While a patient is performing a rehabilitation exercise, the data captured by the camera will be processed by the pose estimation algorithm to extract landmark features that are fed into a movement classification engine frame-by-frame for a detailed motion analysis based on predefined evaluation rules. The model trained to recognize correct concentric and eccentric key poses may be used to extract individual repetitions from a continuous motion sequence of an exercise to evaluate them in terms of completely or incompletely executed repetitions. Furthermore, the classifiers that recognize compensatory poses may provide real-time feedback with much more specific information regarding exactly how the motion deviates from the correct execution and quantify the quality of the overall exercise performance by analyzing the pose validity at each frame level. At last, the proposed approach avoids privacy issues related to clinical practice [56]. The pose estimation module elaborates only landmarks features from monocular RGB images for motion analysis [43]. This characteristic makes the approach highly robust for preserving subjects' privacy in their environment during decentralized rehabilitation treatments.

Before testing the presented approach in a relevant clinical application, future studies are required to investigate the performance of the proposed technique deployed on a proto-

type system that provides real-time evaluation of resistance training rehabilitation exercises. Further, setup configurations of the patient in front of the camera should be considered necessary for correctly monitoring and assessing in-home rehabilitation exercises in future validation studies. Moreover, also the efficiency and user-friendliness of this type of system also should be discussed as crucial research topics in the next studies.

Chapter 5

Advancing activity recognition in health-care: addressing fairness and robustness challenges

5.1 Bias and generalizability issues

Fairness in the research community is a recently established area of machine learning that studies how to ensure that biases in the data and model inaccuracies do not lead to models that treat individuals unfavorably based on characteristics such as e.g. race, gender, disabilities, and sexual or political orientation [236, 81]. The integration of explanations, enabled by explainable AI, offers a pathway to strengthen trust in AI-driven decision-making processes [13].

The pivotal relationship between transparency, explanation, and fairness in AI decisions is paramount for future development. Prioritizing both user trust and perceived fairness is vital for the responsible use of AI systems, an aspect that has been frequently overlooked in the recent past. Fairness significantly affects user trust, particularly when fairness levels are lower, resulting in diminished trust.

Researchers and practitioners are showing an increasing interest in algorithmic fairness, proposing solutions for many different tasks [207, 46, 251]. These works typically focus on describing and classifying important measures of algorithmic fairness and methods to enhance it. Advances in applications like deep face recognition, loan and credit assessment, and product recommendation systems have emphasized the crucial importance of assessing datasets for potential biases [314]. Studies have highlighted that biases can manifest in various forms, including social, measurement, representation, label, algorithmic, evaluation,

deployment, and feedback bias in the underlying datasets [314].

Recent research endeavors have prioritized the mitigation of representation bias to foster the development of fair and equitable machine learning models [141]. This undertaking entailed curating a balanced dataset encompassing diverse age groups, genders, and ethnic backgrounds, drawing from various freely available sources and public image datasets.

In the domain of activity recognition, prior research has predominantly delved into dynamic inductive biases, encompassing factors such as sensor types, positions, segment sizes, and preprocessing techniques [113]. However, the bias introduced by the selection of subjects for activity recognition has received less attention. Recent studies in gait activity-based person re-identification have emphasized the distinctive nature of each individual’s motion behavior [11]. Various physical characteristics also play a significant role in influencing human motion, as discussed in other literature [264, 337, 225].

Koskimäki et al. investigated bias avoidance in classification algorithms using a dataset of 10 subjects [150]. Similarly, Alam et al. proposed mitigating bias in motion features for activity recognition systems targeting older adults [10]. Lai et al. explored bias in human action recognition using 17 subjects from the UP-Fall Detection Dataset [163]. Braganca et al. examined potential bias issues using the SHOAIB and WISDM datasets [35]. These studies collectively highlight the importance of addressing bias and individual differences in developing fair and effective activity recognition systems.

Understanding and addressing differences in motion features between healthy and impaired subjects is crucial for developing accurate and fair activity recognition systems in healthcare. Researchers aim to create robust models that provide reliable and equitable healthcare solutions by focusing on these factors [186, 84]. Recent studies have explored the notion of subject similarity to weight training data, prioritizing data analogous to the user being tested [84]. These studies assess personalization models based on user similarity in physical attributes and signal patterns. Using metrics such as euclidean distance between feature vectors of subjects based demographics researchers have visualized multidimensional scaling over physical characteristics [83].

Overall, it is crucial to acknowledge that motion behavior is significantly influenced by the health conditions of subjects. Previous research has not sufficiently addressed the need for solutions that can mitigate bias stemming from differences in disability and impairment characteristics among subjects in healthcare-related activity recognition systems.

This thesis presents two distinct strategies for promoting fairness in activity recognition systems for healthcare applications. The first strategy focuses on evaluating subject-specific characteristics that influence model bias, aiming to reduce bias during data preparation and model training. The second strategy proposes innovative model architectures designed to

mitigate performance disparities and enhance cross-subject resilience of activity recognition systems.

In the Section 5.2 is evaluated the influence of subjects' physical and demographic characteristics on training different deep learning architectures used in activity recognition tasks. An existing published dataset containing information on subjects' health status (patients or healthy individuals), demographics (age and gender), and health-related attributes (such as wheelchair usage and affected body side due to physical impairments) is utilized to examine the impact of these factors on activity recognition. The investigation focuses on assessing how disability characteristics and demographic attributes influence activity classification performance, considering dataset bias in classifier models within the rehabilitation domain.

In Section 5.3, a novel approach is explored to enhance cross-subject resilience in activity recognition systems, specifically addressing the variability in motion behavior caused by diverse health conditions. To achieve this, a multiscale feature learning technique is employed, following recent evidence that highlights the importance of capturing both short- and long-term motion patterns [79]. This approach aims to allow the system to better accommodate the wide range of individual differences in motion behavior, ensuring more robust and accurate activity recognition across various health conditions.

5.2 Evaluating bias in human motion data

5.2.1 Dataset

In the healthcare domain, datasets used for activity recognition often lack sufficient subject numbers and comprehensive annotations on demographic and health-related characteristics. For instance, Vakanski et al. presented the UI-PRMD dataset, which includes 10 healthy subjects with annotations limited to demographic data [313]. Similarly, Antunes et al. introduced the AHA-3D dataset, consisting of 21 healthy subjects annotated solely for age attributes, without information on other characteristics such as physical impairments [14]. Notably, physical and health-related attributes can significantly influence motion, as evidenced by recent studies [11, 264].

While advancements have been made, such as the dataset by Capecchi et al., which includes 44 healthy and 34 patient subjects with demographic annotations, further annotations on health-related characteristics are often lacking [44]. The Toronto Rehab Stroke Posture Detection Dataset (TRSPD) stands out as one of the few datasets providing annotations on clinical and disability features of patient subjects, including details on affected body parts and mobility status [74]. Despite its strengths, the TRSPD has a smaller patient cohort when

compared with the IRDS Dataset 3.1.2, which provides comprehensive annotations covering demographics, disability status, and affected body parts. Thus, the IRDS dataset emerged as the most suitable free and open resource available, offering comprehensive annotations covering demographic characteristics, disability status, and affected body parts.

The IRDS dataset encompasses patients with diverse pathological conditions, including stroke and spinal cord injury, making it a valuable resource for improving the monitoring and assistance of upper limb rehabilitation across various populations, including those with neurological and musculoskeletal disorders.

The selected subset consists of repetitions of shoulder movements performed by 28 subjects, comprising 14 patients and 14 healthy controls. Recordings pertaining to patient ID 205 were omitted from the analysis as they comprised only a single correct repetition for one of the exercises. The dataset provides 3D coordinates of 25 body joints and corresponding depth maps for each frame, recorded using a Microsoft Kinect One sensor at 30 frames per second. Each movement is annotated with exercise type (such as shoulder flexion right (SFR), shoulder flexion left (SFL), shoulder abduction right (SAR), shoulder abduction left (SAL), shoulder forward elevation (SFE)), the correctness of pattern motion (correct or incorrect), the subject's demographics (gender and age), and subject's health-related characteristics (including wheelchair usage, and the body side affected by physical impairment).

Table 5.1 provides information about age and gender attributes. Additionally, Table 5.2 outlines details about physical disabilities and impairments among patient subjects.

There is a notable trend wherein incorrect gestures tend to be longer than the correct ones, likely due to patients either struggling to perform the gesture or taking longer to prepare for it. As illustrated in Figure 5.1, most incorrect gestures exhibit significantly longer durations compared to correct executions, with exercises SFL and SFR particularly notable in this regard.

Given the various physical limitations of patients, not all of them completed the same number of exercise repetitions (sequences); a disparity was also observed among the controls due to differing availability times. Each subject attempted to perform gestures a certain number of times, with these repetitions labeled as correct or incorrect. Figures 5.2 and 5.3 respectively depict the distribution of the number of correct and incorrect repetitions for each exercise, while Figure 5.4 illustrates the contrast in distribution between correct and incorrect repetitions among healthy and patient subjects across each exercise.

Training data were curated following a measure of heterogeneity. Heterogeneity measure, a statistical concept, denotes the non-uniform distribution of qualities within a dataset. Consequently, the curated training set utilized for neural network training will encompass various degrees of heterogeneity based on the health, physical, and demographic attributes

Table 5.1: Demographic frequencies in the selected IRDS subset.

Gender	Healthy		Patients	
	N	ID	N	ID
Male	7	101-103-106- 107-301-302- 305	11	201-202-203- 204-207-209- 210-212-213- 214-216
Female	7	102-104-105- 303-304-306- 307	3	206-211-215
Age	Healthy		Patients	
	N	ID	N	ID
<30	11	103-105-106- 107-301-302- 303-304-305- 306-307	3	203-209-214
30-39	3	101-102-104	3	206-215-216
40-59	0	n/a	6	201-202-207- 211-212-213
>60	0	n/a	2	204-210

Table 5.2: Distribution of wheelchair disability and physical impairment side among patients in the selected IRDS subset.

Disability	Patients	
	N	ID
Wheelchair	7	201-203-209-210-212-215-216
No wheelchair	7	202-204-206-207-211-213-214
Impairment side	Patients	
	N	ID
Right	5	206-207-209-213-215
Left	2	202-211
Bilateral	2	210-212
Lower limbs	2	201-203
n/r	3	204-214-216

of the subjects. The hypothesis is that classifier performance will vary when trained with data from differing levels of heterogeneity. Moreover, as the heterogeneity of physical characteristics among subjects increases within a given training set, the activity classifier model’s performance will improve when tested on an unfamiliar set of subjects with diverse physical attributes.

To test the hypothesis, various training subsets were curated based on measures of het-

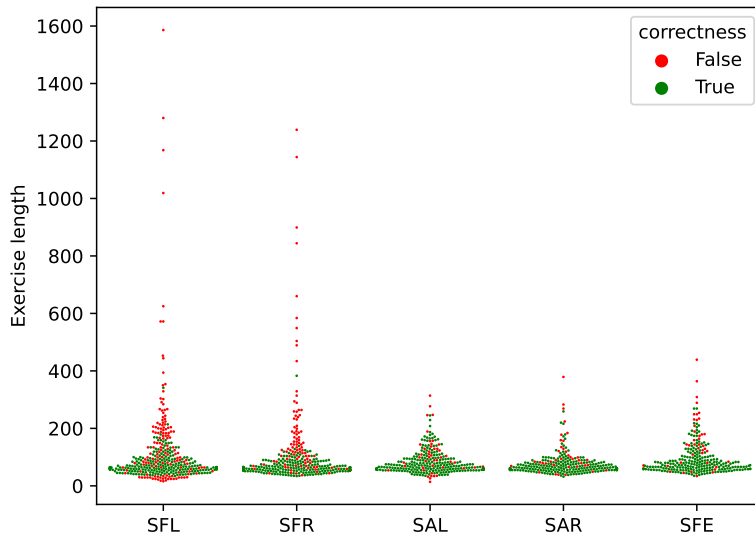


Figure 5.1: Distribution of exercise duration, measured in frames, within correct and incorrect motion classes.

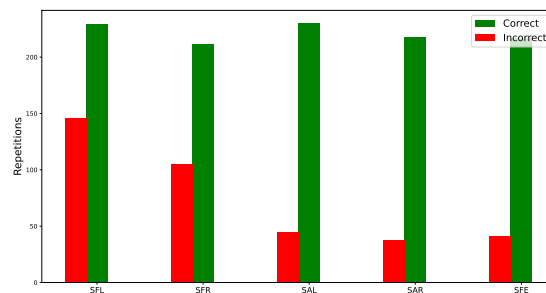


Figure 5.2: Distribution of correct and incorrect repetitions for each exercise.

erogeneity.

An initial assessment was conducted to evaluate how demographic attributes, such as gender and age, influence model biases during training across subjects. For this purpose, 12 subjects (6 healthy and 6 patients) were included in the training set, while the remaining 14 subjects were designated for testing. The diversity within the training sample is characterized by unique identifiers based on distinct demographic characteristics, as shown in Table 5.3.

Following this, an investigation was conducted into how disabilities related to wheelchair use and physical impairments among patients influence model biases during training. Consequently, 8 patient subjects were included in the training set, while the remaining 6 subjects were allocated for testing. The diversity within the training sample is detailed in Table 5.4.

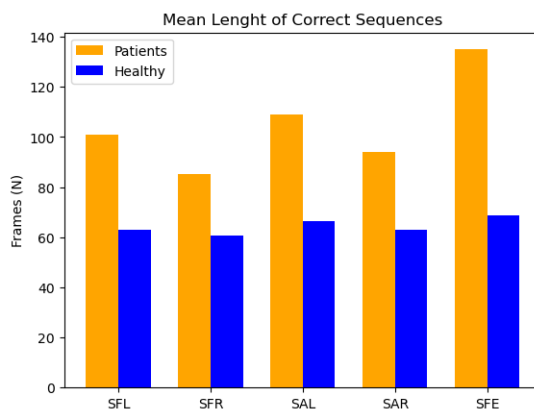


Figure 5.3: Duration of correct sequences, measured in frames, categorized by subject type across diverse rehabilitation exercises.

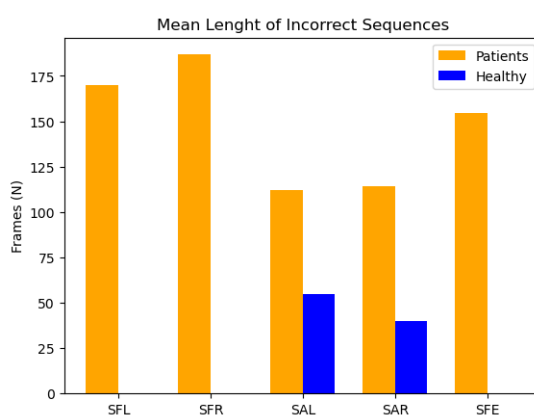


Figure 5.4: Duration of incorrect sequences, measured in frames, categorized by subject type across diverse rehabilitation exercises.

Table 5.3: Description of the heterogeneity introduced in training data subsets considering demographic attributes of healthy and patient subjects.

Group	Heterogeneity
1	Healthy and patient subjects exhibit discrepancies across all demographic attributes.
2a	Healthy and patient subjects have identical demographic characteristics (balanced for age and gender).
2b	Healthy and patient subjects have the same characteristics for one demographic attribute (age balanced)
2c	Healthy and patient subjects have the same characteristics for one demographic attribute (gender balanced).

Table 5.4: Description of the heterogeneity introduced in training data subsets considering disability (wheelchair usage or not) and physical impairment characteristics (impairment on the right side, impairment on the left side, bilateral impairment, and impairment involving only lower limbs).

Group	Heterogeneity
3a	Patient subjects have identical disability characteristics (balanced for disability).
3b	Patient subjects have variations across disability characteristics, with discrepancies observed in the wheelchair usage (unbalanced for disability).
4a	Patient subjects have identical body impairment characteristics (balanced for impairment side).
4b	Patient subjects have variations across body impairment characteristics, with discrepancies observed in the side of impairment (unbalanced for impairment side).

For instance, groups designated as ‘2a’, ‘3a’, and ‘4a’ indicate a lack of heterogeneity within the selected training set, where all subjects share identical characteristic values. Conversely, ‘1’, ‘3b’, and ‘4b’ signify a heterogeneous training set, where each subject possesses distinct characteristics. The groups identified as ‘3b’ and ‘4b’ can be viewed as comprising various subgroups depending on the distribution of characteristics within each subset. For instance, in ‘3b’, it can be observed whether patients who use wheelchairs due to disabilities have higher or lower frequencies compared to patients who do not use wheelchairs. Likewise, the same analysis can be considered for the distribution of impairment sides within ‘4b’.

To mitigate selection bias, subjects for the training sample were randomly selected, considering the experiment’s heterogeneity. Consequently, experiments were conducted on the maximum number of feasible training sample sets within the dataset. This process yielded five different experiments for each heterogeneity group, employing a 5-fold cross-validation for each model training.

5.2.2 Algorithms

Deep learning techniques to analyze multi-channel time series data were used for human activity recognition [157]. Unlike traditional classifier models that rely on manually engineered features, neural network architectures autonomously extract relevant features from input data prior to the classification stage during supervised learning.

A set of neural network architectures tailored for activity recognition was employed, including temporal convolutional neural networks (tCNN), hybrid Long Short-Term Mem-

ory (LSTM) networks such as CNN-LSTM and Convolutional LSTM (ConvLSTM), and Transformer models. Each of these architectures offers distinct advantages and has been extensively studied in the research literature for human activity recognition tasks.

Temporal CNN model

In sequence modeling tasks, a variation on CNN called a temporal convolutional neural network (tCNN) was used in the study [20]. tCNN has a memory significantly longer than recurrent architectures with the same capacity. Likewise, in different sequence modeling tasks, tCNN achieves better results than RNN, LSTM, since tCNN is parallel, adaptable in terms of receptive field size, stability of gradients, low memory requirements during training, and acceptability of inputs of varying lengths. The tCNN, designed to process temporal sequences effectively, operates convolutional operations across both spatial and temporal dimensions, showcasing promising results in capturing temporal dependencies and extracting discriminative features from time-series data [168, 183].

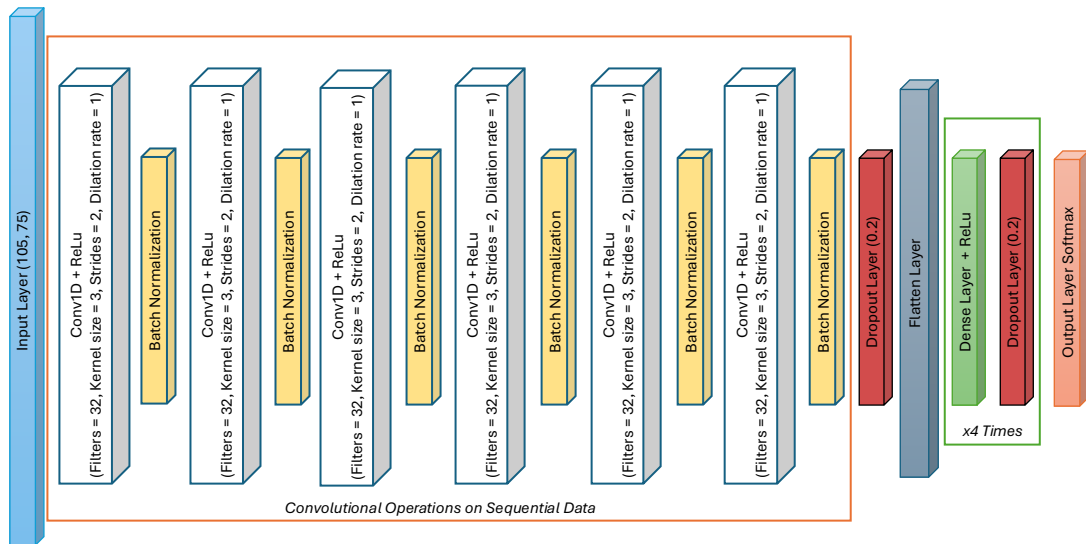


Figure 5.5: Architecture of tCNN Model.

The tCNN architecture utilized is schematically presented in Figure 5.5. The tCNN comprises a structured input layer that receives data representing 3D joint coordinates over time steps. Through six CNN layers with batch normalization, essential features are extracted, while dropout layers mitigate overfitting. A flattened layer prepares the data for 4 fully connected layers, where features are refined using ReLU activation and L2 regularization to bolster model robustness. Finally, the softmax activation in the output layer ensures precise activity classification predictions.

Hybrid LSTM models

LSTM is a type of recurrent neural network (RNN) architecture designed to overcome the vanishing gradient problem in traditional RNNs. It consists of memory cells with self-connections called ‘gates’ that regulate the flow of information, allowing it to remember information over long sequences. LSTM is well-suited for sequential data modeling tasks where long-range dependencies are important, such as natural language processing, speech recognition, and time series prediction.

CNN-LSTM

CNN-LSTM combines convolutional layers with LSTM layers. The CNN layers are typically used for feature extraction from input sequences, capturing spatial patterns if the input data has a spatial structure (e.g., images, sensor data). The LSTM layers are then used to capture temporal dependencies in the extracted features. CNN-LSTM architectures are commonly used for tasks involving spatio-temporal data, such as video classification, action recognition, and weather forecasting[289, 73].

The CNN-LSTM architecture employed is presented in Figure 5.6.

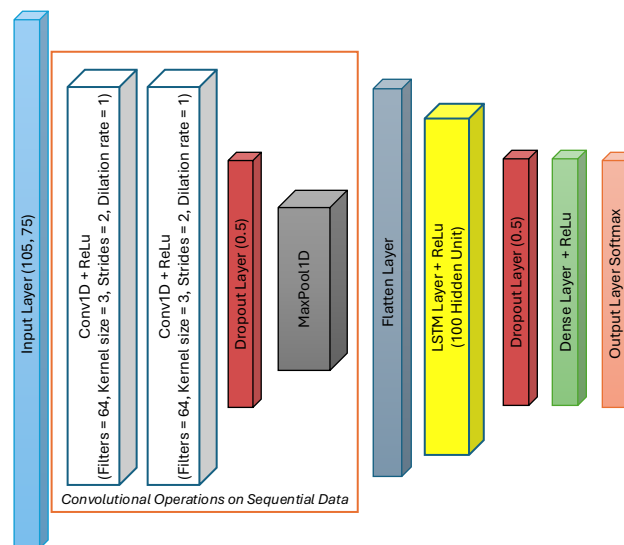


Figure 5.6: Architecture of CNN-LSTM Model.

The input layer applies a 1D convolution operation (Conv1D) with 64 filters and a kernel size of 3, followed by ReLU activation. Two subsequent convolutional layers maintain the same configuration. A dropout layer with a rate of 0.5 is inserted after the second convolutional layer to mitigate overfitting. Max pooling with a pool size of 2 is then used to downsample the features. The output of the convolutional layers is flattened to a vector

format for input into the subsequent LSTM layer. The LSTM layer, comprising 100 units, processes the sequential data, capturing temporal dependencies and patterns. A dropout layer with a dropout rate of 0.5 follows the LSTM layer to further prevent overfitting. Two dense fully connected layers follow the LSTM layer. The first dense layer consists of 100 units with ReLU activation, aiding in feature extraction and nonlinear mapping. The second dense layer utilizes softmax activation, generating final output probabilities for each class.

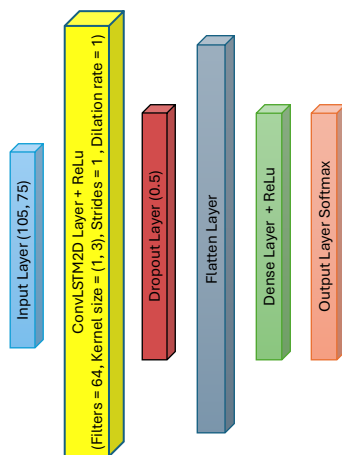


Figure 5.7: Architecture of ConvLSTM Model.

Convolutional LSTM

ConvLSTM is a variant of LSTM that incorporates convolutional operations within the LSTM units. In ConvLSTM, the input-to-state and state-to-state transitions are convolutions rather than fully connected operations. This allows the model to learn spatial-temporal patterns directly from sequential data, without the need for separate CNN layers. ConvLSTM is particularly useful for tasks where both spatial and temporal dependencies are important and can be learned jointly, such as video prediction, precipitation nowcasting, and spatiotemporal anomaly detection [289].

The ConvLSTM architecture utilized in comprises an initial 2D ConvLSTM layer with ReLU activation. This layer processes spatio-temporal data defined by the number of time steps in the sequence, the length of each sequence, and the number of features in each time step. To mitigate overfitting, a dropout layer with a rate of 0.5 is incorporated. The output of the ConvLSTM layer is then flattened into a 1D array. Following this, two dense (fully connected) layers are added. The first dense layer consists of 100 units with ReLU activation, aiding in feature extraction and nonlinear mapping. Finally, the second dense layer employs softmax activation to generate final output probabilities for each class. The

overall architecture of ConvLSTM model is presented in Figure 5.7.

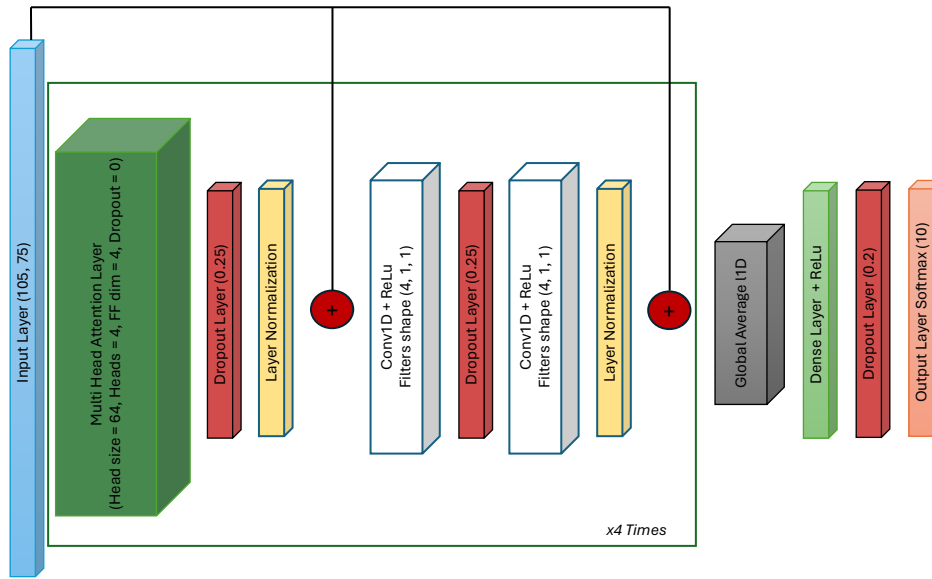


Figure 5.8: Architecture of Transformer model.

Transformer model

Originally developed for natural language processing tasks, Transformer models have been adapted for sequence modeling across different domains. By leveraging self-attention mechanisms, Transformers efficiently capture long-range dependencies in sequential data without relying on recurrent connections. Their effectiveness in processing sequential data has shown promising results in tasks such as activity recognition [285].

The Transformer architecture follows an encoder-decoder paradigm. The encoder extracts features from input sequences, while the decoder utilizes these features to generate output sequences. Each encoder block employs self-attention to enhance tokens with contextual information. Additionally, positional encoding is injected into each embedding, enabling the model to understand the input sequence's positional information without redundancy. The Transformer architecture processes input data without recurrence or convolution, making it versatile for handling various types of sequential data.

The utilized Transformer model consists of an encoder layer featuring 4 attention heads, along with a position-wise feed-forward network and two residual connections, each complemented by Layer Normalization. The feed-forward network incorporates convolutional layers with ReLU activation functions, enhancing the model's ability to capture non-linear relationships within the data. Subsequently, an MLP with two Dense layers, each with 128

dimensions and ReLU activation, is applied, incorporating a dropout rate of 0.4 to improve generalization. Finally, a Softmax Dense Layer performs classification across the activity classes, addressing the exercise classification task.

The overall Transformer architecture is visually summarized in Figure 5.8.

5.2.3 Experimental design

Experiments were carried out using the previously described training subsets, which were designed to incorporate diverse heterogeneity within each group under consideration. Multi-class classifications were performed using tCNN, CNN-LSTM, ConvLSTM, and Transformer models across each of the 5 folds, ensuring consistency with the established heterogeneity groups. This cross-validation approach aims to yield more accurate results, particularly given the limited size of the training subsets generated. The model was trained on the 10 classes representing both correct and incorrect exercise actions, as outlined in Table 5.5.

The multi-channel time series data comprises 3D raw coordinates (x, y, z) of 25 body joints recorded by the Kinect sensor, resulting in a set of 75 features utilized for training each deep learning model. The input tensor for the time-series deep learning model includes three main dimensions: the number of samples, the length of the sequence to be processed, and the number of features. Data pre-processing was executed to adjust input data in alignment with the architecture type specified in Table 5.6. Different batch sizes were employed during training on datasets comprising both healthy and patient subjects, as well as datasets containing only patient subjects. This discrepancy in batch size stems from the disparity in the number of samples, which is lower when considering only patients.

To address the issue of imbalance in exercise activity classes, a Naive Bayesian method was applied to model the weight of each class in the classification [136]. This technique helps mitigate the effects of class imbalance, ensuring more robust and accurate predictions.

Given the diverse repetition lengths observed between correct and incorrect motion patterns among both patients and healthy subjects, as shown in Figures 5.3 and 5.4, multiple pre-model evaluations were conducted using varying sequence lengths to identify the optimal configuration. The best results were achieved using the mean sequence length of patients, consisting of 105 frames. Therefore, a sequence length of 105 frames was selected for the experimental analysis. Moreover, a time series length normalization technique was implemented using the zero pre-padding method. Zero padding is a widely used approach for pre-processing time series data with varying lengths. This method was chosen based on recent research indicating its superior performance compared to other methods [129]. The rationale behind zero pre-padding is that the zeros introduced should be disregarded by the

Table 5.5: Classes of correct and incorrect exercise actions used to train the activity recognition model.

Class (N)	Exercise action
0	SFL performed following the correct pattern of motion.
1	SFL performed following an incorrect pattern of motion.
2	SFR performed following the correct pattern of motion.
3	SFR performed following an incorrect pattern of motion.
4	SAL performed following the correct pattern of motion.
5	SAL performed following an incorrect pattern of motion.
6	SAR performed following the correct pattern of motion.
7	SAR performed following an incorrect pattern of motion.
8	SFE performed following the correct pattern of motion.
9	SFE performed following an incorrect pattern of motion.

Table 5.6: Experimental parameters used for models training.

Parameter	tCNN	CNN-LSTM	ConvLSTM	Transformer
Input dimension	n x 105 x 75	n x 105 x 75	n x 105 x 75	n x 105 x 75
Batch size	32-8	32-16	32-16	32-16
Optimization algorithm	Adam	Adam	Adam	Adam
Learning rate	1×10^{-4}	1×10^{-4}	1×10^{-4}	1×10^{-4}
Loss function	Categorical cross-entropy	Categorical cross-entropy	Categorical cross-entropy	Categorical cross-entropy
Class weight balancing	Naive Bayesian	Naive Bayesian	Naive Bayesian	Naive Bayesian
Epochs	100	100	100	100
Patience (epochs)	True (10)	True (10)	True (10)	True (10)

model’s weights during training.

The Keras Python library was used to develop and train deep learning architectures for activity recognition tasks. To ensure consistency in experimentation, a fixed seed was applied

for neural network initialization.

The entire pipeline of the experimental study is simply summarized in Figure 5.9.

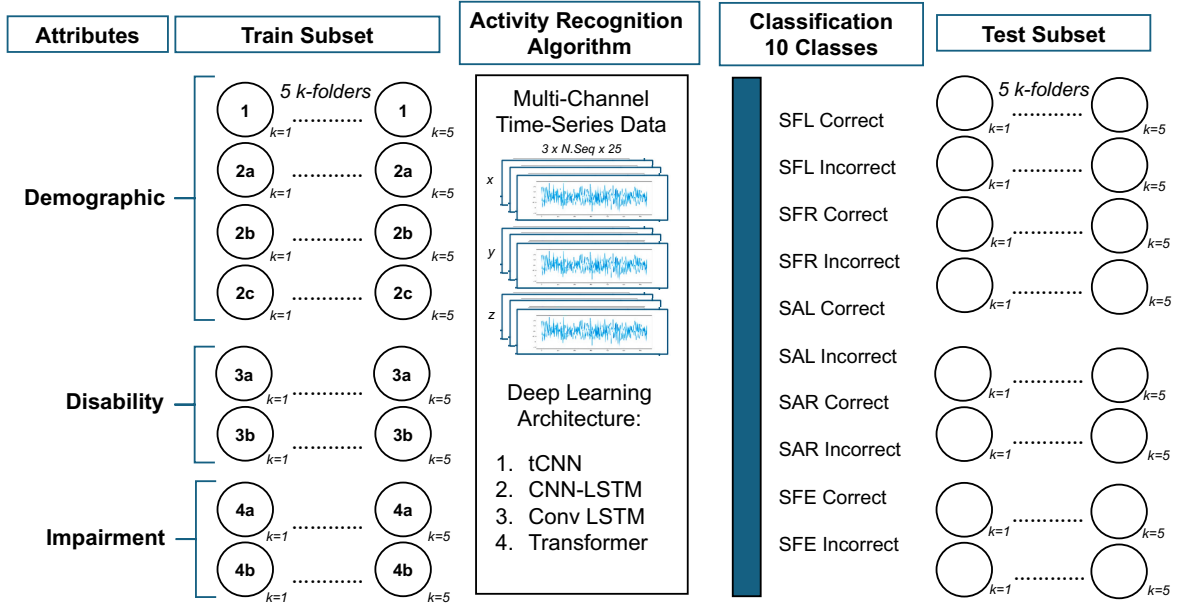


Figure 5.9: Experimental pipeline.

Note. i) Eight training subsets were extracted from the IRDS dataset, which was specifically curated to encompass demographic, disability, and impairment characteristics of subjects. This curation was performed to assess their influence on the activity recognition algorithm. Cross-validation using five folds was employed for the experimental analysis. ii) State-of-the-art architectures designed for processing multi-channel time series data were carefully selected. These included tCNN, CNN-LSTM, ConvLSTM, and Transformer models. These architectures were chosen to classify ten classes of exercise activities, distinguishing between correct and incorrect execution. iii) Cross-validation approach was utilized to evaluate the performance of each model. This evaluation was conducted using the remaining subjects in the test sets, ensuring robustness and generalization of the models' performance across diverse datasets.

K-fold cross-validation was applied to evaluate the models, generating five folds for each training set. Recordings from different subjects were used for both training and test sets to ensure a comprehensive assessment of the model's generalizability. A stratified method was employed to split the training and validation sets.

The categorical cross-entropy loss function was selected for the classification task, while the general model performance was evaluated using the categorical accuracy metric.

To further assess classification performance, a set of metrics designed for multi-class problems was employed, including Precision, Recall, and F1 score [118].

5.2.4 Bias analysis

Demographic bias

The performances of models trained on the various curated training subsets ‘1’, ‘2a’, ‘2b’, ‘2c’, following heterogeneity measures in demographic attributes, are summarized in Table 5.7.

Table 5.7: Performance metrics, including mean accuracy, precision, and recall, of models trained while considering the influence of demographic attributes.

Group	Train size (N. Sequences)	Metrics	tCNN	CNN-LSTM	ConvLSTM	Transformer
1	663	Accuracy	0.68	0.88	0.80	0.77
		Precision	0.69	0.88	0.80	0.79
		Recall	0.67	0.87	0.79	0.77
2a	589	Accuracy	0.69	0.70	0.65	0.73
		Precision	0.71	0.70	0.66	0.75
		Recall	0.67	0.70	0.65	0.72
2b	595	Accuracy	0.67	0.69	0.65	0.71
		Precision	0.69	0.70	0.67	0.73
		Recall	0.65	0.68	0.64	0.69
2c	606	Accuracy	0.67	0.77	0.67	0.64
		Precision	0.69	0.77	0.69	0.65
		Recall	0.65	0.76	0.66	0.62

Figure 5.10 illustrates the trends of inductive demographic attribute influence on model performance by displaying the mean and standard deviation of accuracy and F1-score metrics across each train subset group.

Looking at the data provided in Table 5.7, it can be observed several trends in the performance metrics across different groups. There appear to be discernible differences in performance metrics across various training set groups considering the demographic attributes of subjects (see Figure 5.10).

In group ‘1’, characterized by maximal variability in age and gender among healthy and patient subjects, all models exhibit relatively high-performance metrics. Moving to group ‘2a’, where patients and healthy subjects share identical age and gender attributes, all models demonstrate moderate performance metrics. Accuracy, precision, and recall hover around similar values across models. Similarly, in group ‘2b’, where subjects are balanced solely for age attributes, all models exhibit moderate performance metrics, although slightly lower than group ‘2a’. Group ‘2c’, where patients and healthy subjects are balanced for gender attributes, echoes similar trends.

Overall, the results suggests that while CNN-LSTM consistently demonstrates strong

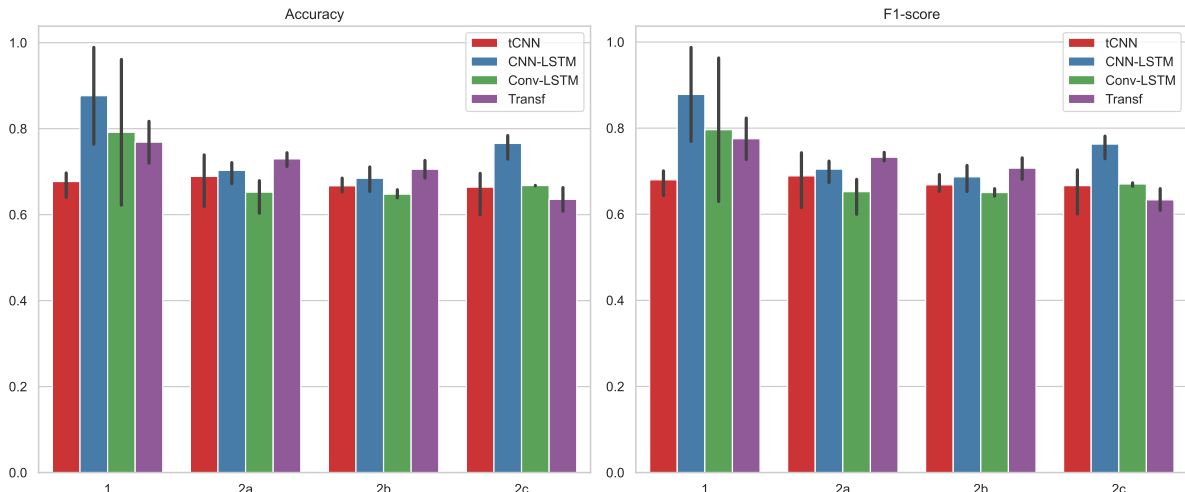


Figure 5.10: Barplot of demographic bias analysis.

Note. The figure shows mean and standard deviation statistics of accuracy and F1-score metrics across each train subset group, curated to account for demographic bias among patient and healthy subjects.

performance across all groups, the performance of other models may vary depending on the specific characteristics of the dataset, particularly regarding age and gender variability among healthy and patient subjects. Additionally, the results indicate that demographic bias may have a varying impact on model performance, with certain models exhibiting better resilience in handling such biases than others.

Age and gender can influence the way individuals perform various activities, including motion-based tasks [293, 75]. For example, older individuals may exhibit different movement patterns compared to younger ones due to factors such as muscle strength, flexibility, and balance changes associated with aging. Similarly, differences in anatomy, physiology, and societal roles between genders can lead to variations in motion characteristics.

When training an activity recognition model on a dataset that reflects a wide range of ages and genders, the model learns to recognize and accommodate these subtle differences in motion patterns. This exposure to diverse demographic characteristics helps the model generalize better to new, unseen subjects because it has learned to identify and adapt to the variability present in the population.

On the other hand, if the training dataset predominantly consists of subjects from a narrow age range or a single gender group, the model may become biased towards the specific characteristics of that group. As a result, when faced with unseen subjects from different age groups or genders, the model may struggle to accurately recognize their activities due to its limited exposure to the diverse range of motion patterns present in the population.

In essence, training models on datasets with heterogeneous distributions of age and gender attributes allows them to better capture the full spectrum of motion characteristics present in the real world, leading to improved generalization performance on unseen subjects. Conversely, training on datasets with homogeneous age and gender distributions may result in poorer generalization due to the model’s limited ability to adapt to diverse demographic characteristics.

To address the challenges posed by demographic bias, future research should focus on strategies tailored to mitigate the impact of age and gender variability.

Central to this effort is the creation of datasets that comprehensively capture a diverse spectrum of age and gender demographics. Ensuring representation from various population groups in these datasets is crucial as it enables the training of models on data that closely mirrors real-world demographics. This approach not only enhances model generalizability but also serves to mitigate bias inherent in training data.

Disability bias

The summary of findings evaluating the impact of disability and physical impairment attribute bias on model performance is presented in Table 5.8.

Table 5.8: Performance metrics, including mean accuracy, precision, and recall, of models trained while considering the influence of disability and physical impairment bias.

Group	Train size (N. Sequences)	Metrics	tCNN	CNN-LSTM	ConvLSTM	Transformer
3a	353	Accuracy	0.30	0.32	0.22	0.30
		Precision	0.31	0.33	0.22	0.31
		Recall	0.26	0.32	0.21	0.28
3b	303	Accuracy	0.37	0.51	0.36	0.44
		Precision	0.39	0.51	0.37	0.45
		Recall	0.35	0.50	0.35	0.41
4a	333	Accuracy	0.40	0.44	0.31	0.46
		Precision	0.43	0.45	0.32	0.48
		Recall	0.34	0.43	0.30	0.39
4b	321	Accuracy	0.43	0.43	0.44	0.48
		Precision	0.44	0.43	0.45	0.50
		Recall	0.37	0.42	0.43	0.45

Trends regarding the influence of inductive disability and physical impairment bias on model performance are illustrated in Figure 5.11, which also depicts the mean and standard deviation values.

The data presented in Table 5.8 reveals notable trends in performance metrics across different groups and models. There appear to be discernible differences in performance

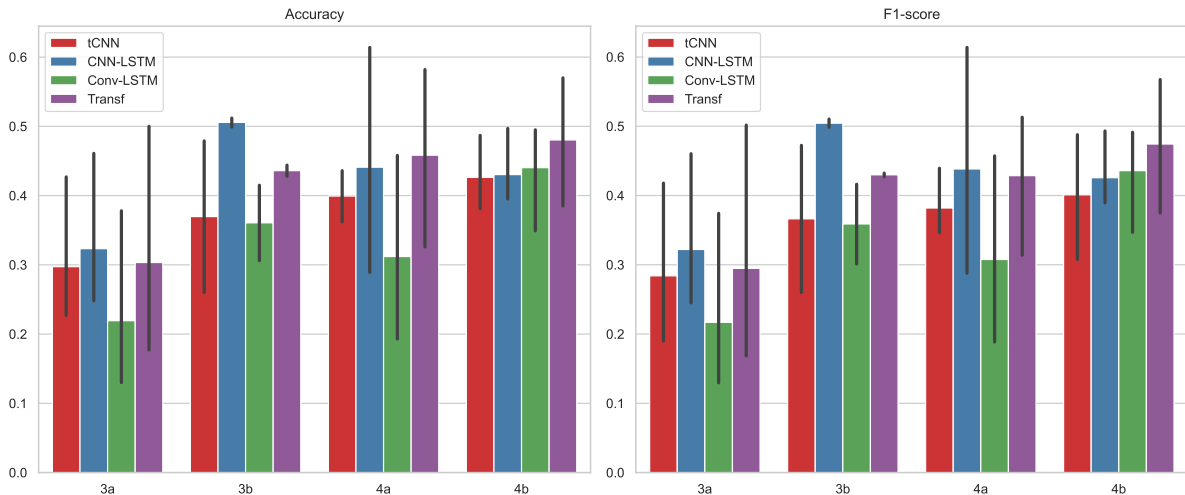


Figure 5.11: Barplot of disability related bias.

Note. The figure shows mean and standard deviation statistics of accuracy and F1-score metrics across each train subset group, curated to account for disability and physical impairment bias among patient subjects.

metrics across various training set groups considering disability and physical impairment characteristics of patient subjects, as shown in Figure 5.11.

Group ‘3a’ consistently displays lower performance metrics compared to other groups. Models trained in this group tend to exhibit the lowest accuracy, precision, and recall among all architectures. This group involves matching subjects with disabilities, particularly those in wheelchairs, with subjects without wheelchair usage. The presence of this specific uniformity in disability attribute process likely impacts the performance metrics within this group. In contrast, models trained on group ‘3b’ show moderate to high accuracy, precision, and recall. This group consists of subjects not balanced for disability, resulting in varying frequencies of disabilities within the dataset. This variability may introduce additional heterogeneity, positively affecting model performance. Similarly, models trained on group ‘4a’ demonstrate moderate to high accuracy, precision, and recall. This group involves matching subjects for physical impairment, ensuring an equal number of subjects for each type of impairment. By balancing the representation of different types of physical impairments, this matching process aims to mitigate bias and variability within the dataset. Group ‘4b’ presents the highest variability in frequency for each subgroup of physical impairment, resulting in a diverse and heterogeneous dataset. Models trained on this group may exhibit variable performance metrics, with accuracy, precision, and recall contingent upon the specific characteristics and frequencies of physical impairments represented in the dataset.

Disability and physical impairment can significantly affect an individual’s motion patterns

due to various factors such as limited mobility, altered biomechanics, and compensatory movements [316]. For instance, individuals with mobility impairments may exhibit different gait patterns or have difficulties performing certain movements compared to those without impairments. Similarly, individuals with conditions like arthritis or joint stiffness may have restricted range of motion, affecting their movement capabilities.

These attributes can influence model learning in activity recognition algorithms in several ways. Firstly, if the training dataset includes diverse samples encompassing individuals with disabilities and physical impairments, the model can learn to recognize and accommodate the unique motion patterns associated with these conditions. This exposure helps the model generalize better to unseen individuals with similar attributes, thus improving its overall performance.

Conversely, if the training dataset lacks sufficient representation of individuals with disabilities or physical impairments, the model may struggle to accurately recognize their movements. This can lead to biases and inaccuracies in activity recognition, particularly when the model encounters unseen individuals with such attributes.

In essence, prioritizing the curation of datasets that encompass a wide range of physical characteristics and health-related attributes, throughout both dataset creation and model training stages of activity recognition algorithms, represents a pivotal step towards advancing the creation of more equitable and unbiased systems for monitoring and evaluating patient motion.

It is crucial to recognize that physical and health-related attributes influence motion in subtle yet significant ways. Understanding and studying these attributes is inherently complex, yet imperative in ensuring accurate and inclusive motion monitoring systems.

In future research endeavors, it is essential to acknowledge that health-related characteristics impact motion and model learning dynamics profoundly. Therefore, datasets should encompass comprehensive annotations of patients' health-related attributes, extending beyond demographics. By incorporating detailed information about health conditions, disabilities, impairments, and other relevant factors, researchers can more accurately analyze the relationship between these attributes and motion patterns, enabling a deeper understanding of their impact on activity recognition systems.

5.3 Enhancing cross-subject algorithm robustness

5.3.1 Dataset

Analysys were performed on the IRDS dataset containing nine rehabilitation exercises (detailed in Table 5.9) performed by healthy and patient subjects.

Table 5.9: Description of rehabilitation activities.

Label	Exercise Name	Description
0	Elbow Flexion Left (EFL)	Flexion and extension movement of left elbow.
1	Elbow Flexion Right (EFR)	Flexion and extension movement of right elbow .
2	Shoulder Flexion Left (SFL)	Flexion and extension movement of left shoulder while keeping the arm straight in front of the body.
3	Shoulder Flexion Right (SFR)	Flexion and extension movement of right shoulder while keeping the arm straight in front of the body.
4	Shoulder Abduction Left (SAL)	The left arm is raised away from the side of the body while keeping the arm straight.
5	Shoulder Abduction Right (SAR)	The right arm is raised away from the side of the body while keeping the arm straight.
6	Shoulder Forward Elevation (SFE)	With hands clap together, the arms are kept straight and raised above the head, keeping the elbows straight.
7	Side Tap Left (STL)	The left leg is moved laterally and then returned to its original position while maintaining balance.
8	Side Tap Right (STR)	The left leg is moved laterally and then returned to its original position while maintaining balance.

Exercises performed by 22 of the 29 subjects were taking into account for the analysis, as only these participants comprised in the IRDS dataset completed all the exercises.

The demographic characteristics of the subjects are detailed in Table 5.10 for healthy subjects and in Table 5.11 for patients.

Due to the real-world collection of the data, there is considerable variability in the motion data. Figure 5.12 shows that motion duration vary significantly among the different exercises, influenced by both inter-subject and intra-subject variability. This variability arises from differences in health conditions and intrinsic motion patterns. Variability is observed both within the same movement repeated by the same subject multiple times, referred to as within-subject variability, and between different subjects performing the same movement, known as between-subject variability. An example of within-subject variability is illustrated in left side of Figure 5.13, which plots the x-axis data of the right wrist of a healthy subject performing exercise SAR correctly while standing. This figure demonstrates variations not only in the length (number of frames) but also in the positional values (coordinates) of the data. Meanwhile, in the right side of Figure 5.13 illustrates an example of between-subject

Table 5.10: Demographic characteristics of healthy subjects.

Person ID	Age-Group	Sex
101	30–39	male
102	30–39	female
103	20–29	male
104	30–39	female
105	20–29	female
106	20–29	male
107	20–29	male
301	20–29	male
302	20–29	male
303	20–29	female
304	20–29	female
305	20–29	male
306	20–29	female
307	20–29	female

Table 5.11: Demographic characteristics of patient subjects.

Patient ID	Age-Group	Sex
204	60+	male
206	30–39	female
211	50–59	female
212	40–49	male
213	50–59	male
214	20–29	male
215	30–39	female
216	30–39	male

variability, where five subjects performed the exercise SAR correctly multiple times, revealing significant differences in how each subject executes the exercise.

5.3.2 Multiscale algorithms

Most algorithms are used to extract temporal features using a single stride length in convolutional operations, which simplifies the analysis but does not account for variations in motion features between different subjects. As a result, they overlook individual differences in gait, speed, and movement style, which are crucial for accurately recognizing activities across diverse populations, especially in healthcare applications.

Multiscale feature learning represents a key challenge in developing trustworthy activity recognition algorithms, particularly in healthcare applications. Unlike the traditional

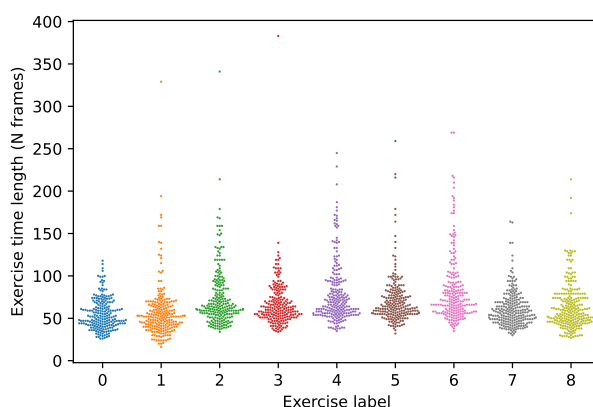


Figure 5.12: Visualization of motion length in terms of the number of frames across exercises.

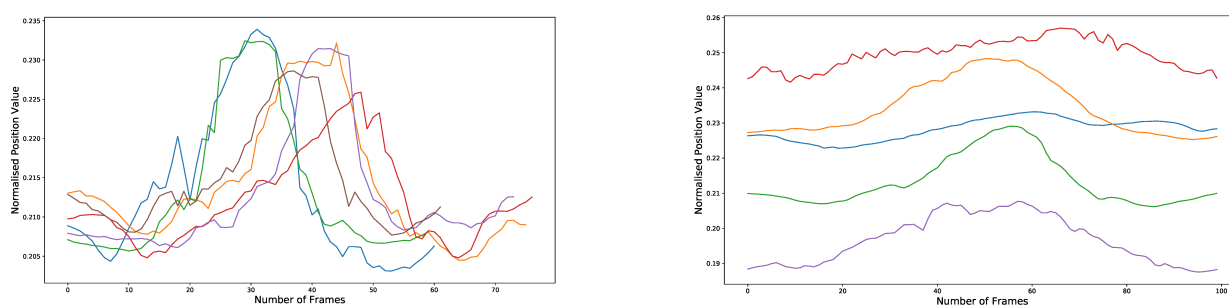


Figure 5.13: An example of motion variability in the dataset.

Note. The figure shows intra-subject variability on the left and inter-subject variability on the right.

algorithm, the multiscale algorithm has two parallel branches that extract features at different scales to identify meaningful patterns in the data. Generally, one branch performs convolution operations with a reduced stride to extract detailed features from short-term dependencies, while the other branch uses a larger stride to capture long-term dependencies. Customized architectures based on a multiscale feature learning approach are evaluated, specifically designed for processing multi-channel time-series data. The experiments underscore the importance of advanced methods to improve activity recognition in healthcare, where accurate exercise monitoring and performance evaluation are critical for delivering effective and personalized treatments.

To summarize, the top branch is designed to capture fine details of movement at a higher computational cost, while the lower branch sacrifices some information to gain a global view of the movement. The theoretical basis for using parallel branches with varying speeds involves two key components: the varying stride in parallel branches and the use of dilated

convolutions to expand the receptive fields [335].

Formally, a 1D convolution with varying stride is shown in Eq.5.1, where x is the input to the convolutional layer of size n , the kernel h has length k , and s represents the stride, or the number of positions moved after each convolution operation. A stride $s > 1$ results in the loss of information, which can be interpreted as sacrificing fine-grained details to capture the overall picture, akin to a moving average filter with a non-overlapping window. The formal definition of a dilated convolution is given in Eq.5.2 [335].

$$y(n) = \begin{cases} \sum_{i=0}^k x(n+1)h(i), & \text{if } n = 0 \\ \sum_{i=0}^k x(n+i+(s-i))h(i), & \text{if } n \neq 0 \end{cases} \quad (5.1)$$

$$(x *_l h)(y) = \sum_{i=0}^{f-1} x_{y-l \cdot i} \cdot h(i) \quad (5.2)$$

Multiscale temporal CNN model The multiscale version of the tCNN consists of two parallel branches of convolutional operations on sequential data. The top branch has six CNN layers with a stride of 1, focusing on short-term dependencies and capturing fine-grained features. In contrast, the lower branch uses a stride of 3 to capture long-term dependencies and extract higher-level features. Both branches culminate in a batch normalization layer and a global average pooling layer before being concatenated. The combined output is then fed forward to the flattening layer, which processes the data for the final output. The entire architecture of the multiscale tCNN model is illustrated in Figure 5.14.

Multiscale CNN-LSTM The multiscale version of the CNN-LSTM consists of two parallel branches of convolutional operations on sequential data. One branch has two hidden layers, applying a 1D convolution operation (Conv1D) with 64 filters and a kernel size of 3 with a stride of 1, focusing on short-term dependencies and capturing fine-grained features. In contrast, the second branch uses a stride of 3 to capture long-term dependencies and extract higher-level features. Each branch culminates in a flattened layer before being concatenated. The combined output is then fed into the LSTM layer, which processes the data for the final output. The entire architecture of the multiscale CNN-LSTM model is illustrated in Figure 5.15.

Multiscale convolutional LSTM The multiscale ConvLSTM model comprises two parallel branches of 2D ConvLSTM layers with ReLU activation. The first branch extracts features using a 1x1 convolution with a stride of 3, focusing on long-term dependencies, while the second branch uses a stride of 1 for capturing short-term dependencies. After the

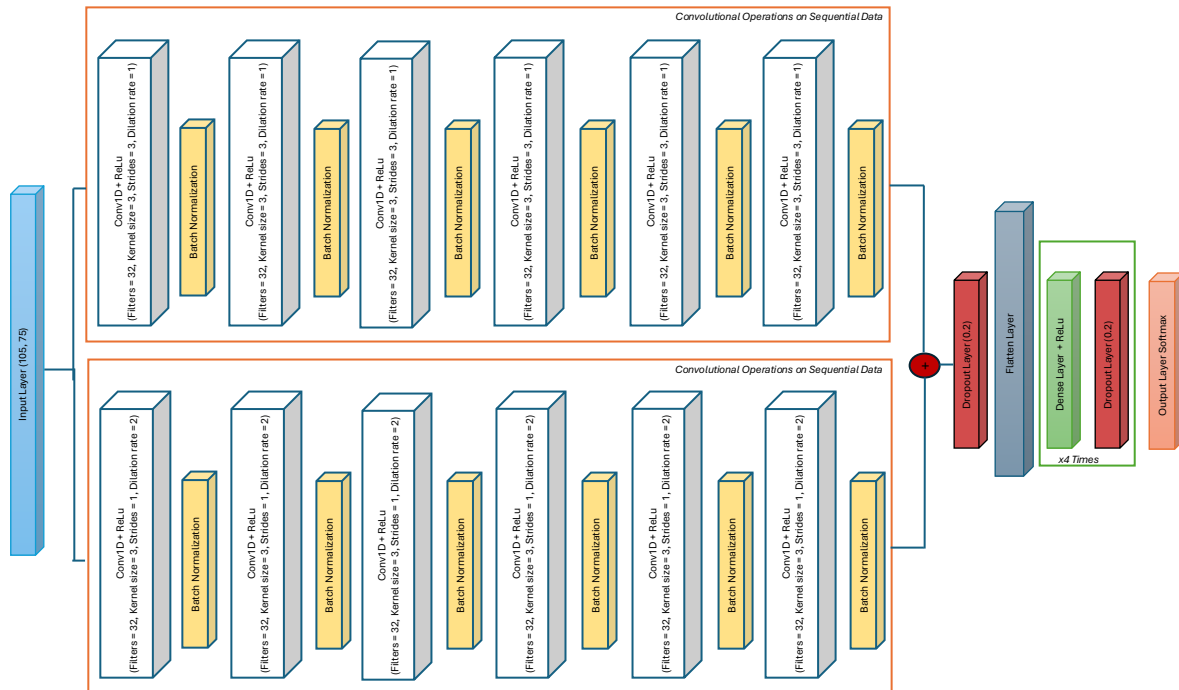


Figure 5.14: Architecture of multiscale tCNN model.

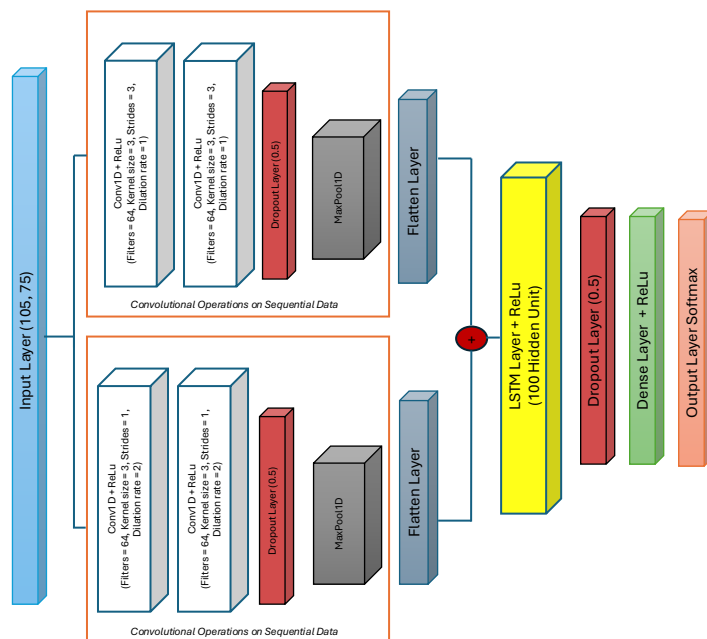


Figure 5.15: Architecture of multiscale CNN-LSTM model.

convolution operations, each branch converges into a flattened layer, and the outputs of these two branches are concatenated before producing the final output. The overall architecture of the ConvLSTM model is presented in Figure 5.16.

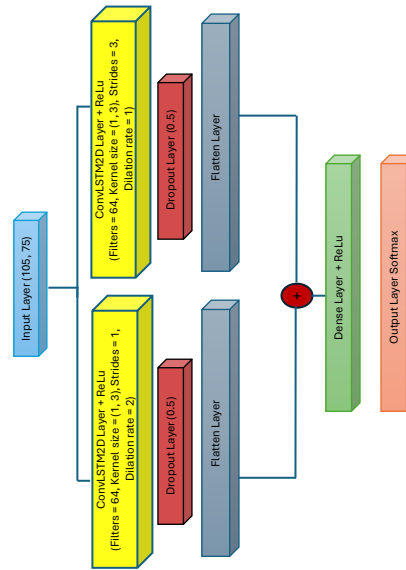


Figure 5.16: Architecture of multiscale ConvLSTM model.

Multiscale Transformer model The multiscale transformer consists of two parallel branches. The top branch performs a 1D convolution with a stride of 1, followed by a dilated convolution with a dilation rate of 2. In contrast, the second branch starts with a 1D convolution with a stride of 3. Aside from these differences, the two branches are identical: the top branch continues with a positional encoding layer, which adds the output of the dilated 1D convolution to a positional signal. The overall architecture of the multiscale transformer is presented in Figure 5.17.

5.3.3 Training and evaluation methods

Experiments were conducted using the previously described IRDS, which contains nine rehabilitation activities performed by 22 subjects.

Multi-class classifications were carried out using tCNN, CNN-LSTM, ConvLSTM, and Transformer models in both traditional and multiscale architectures processing multi-channel time series data.

The multi-channel time series data comprises 3D raw coordinates (x, y, z) of 25 body joints recorded by the Kinect sensor, resulting in a set of 75 features used to train each deep learning model. The input tensor for the time-series deep learning models includes three main dimensions: the number of samples, the length of the sequence to be processed, and the number of features. Data pre-processing was performed to align the input data with the architecture type. The training parameters were fixed as described in Table 5.6 showed in Section 5.2.3. Various batch sizes were employed during pre-analysis to determine the best

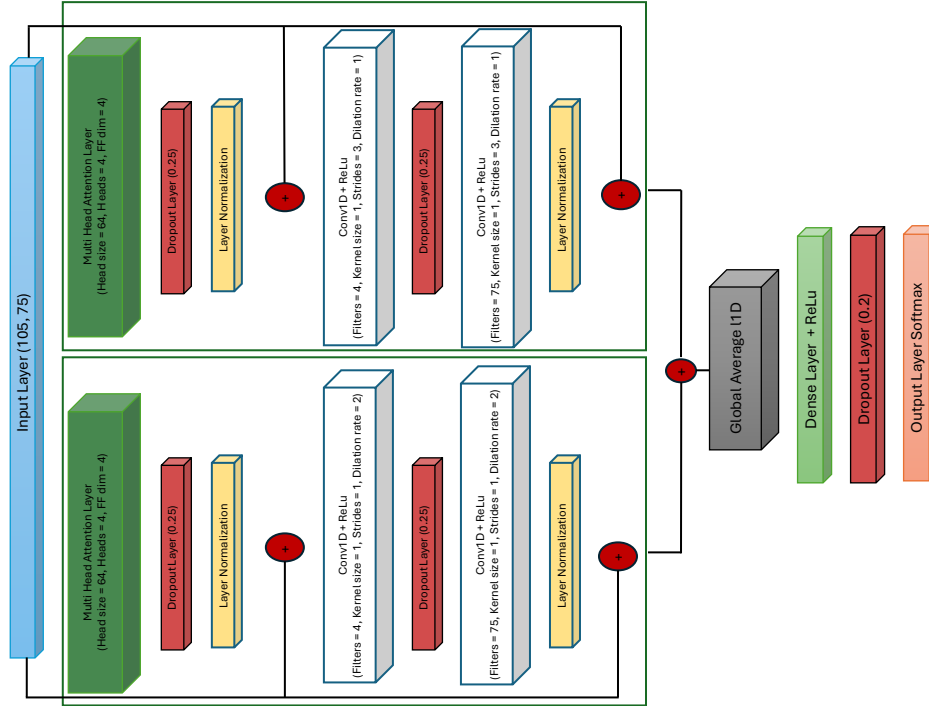


Figure 5.17: Architecture of multiscale Transformer model.

configuration for training the models.

Given that the exercise classes are unbalanced, a Naive Bayesian method was applied to model the weight of each class in the classification [136]. This technique helps mitigate the effects of class imbalance, ensuring more robust and accurate predictions.

Given the varying repetition lengths observed in motion patterns among patients and healthy subjects, multiple pre-model evaluations were investigated using different lengths to find the optimal configuration. Ultimately, the best results were achieved by using the mean sequence length of patients, which was 105 frames. Therefore, a sequence length of 105 frames was chosen for the experimental analysis.

Additionally, a time series length normalization technique was implemented using the zero pre-padding method. Zero padding is a commonly employed approach in preprocessing time series data with varying lengths. Recent research suggests that zero pre-padding offers superior performance compared to alternative methods [129]. The rationale behind zero pre-padding is that the introduced zeros should be disregarded by the model's weights during training.

To ensure robust evaluation and generalize our models' predictive ability, a cross-validation method was applied. Cross-validation involves partitioning the dataset into smaller subsets. The leave-one-subject out (LOSO) cross-validation was applied since each subject exhibits

unique motion patterns. During each validation loop, all subsets except one are used for training, and the remaining subset is used for validation [248, 97]. Validating the models on individual subject patterns provides insights into classification performance and generalizability.

Given that each algorithm underwent training with LOSO cross-validation, involving multiple iterations of training and testing, the average and standard deviation scores for accuracy, precision, recall, and F1-score are presented in results.

5.3.4 Results

General performance of traditional and multiscale algorithms are reported respectively in Table 5.12 and in Table 5.13.

Table 5.12: Mean and standard deviation of LOSO cross-validation results using traditional architecture.

Model	Accuracy	AUC	Precision	Recall
tCNN	0.92 (0.12)	0.98 (0.04)	0.92 (0.11)	0.91 (0.12)
CNN-LSTM	0.91 (0.12)	0.98 (0.04)	0.92 (0.12)	0.90 (0.13)
ConvLSTM	0.88 (0.16)	0.98 (0.04)	0.90 (0.14)	0.88 (0.16)
Transformer	0.81 (0.15)	0.97 (0.05)	0.83 (0.15)	0.80 (0.16)

Table 5.13: Mean and standard deviation of LOSO cross-validation results using multiscale architecture approach.

Model	Accuracy	AUC	Precision	Recall
Multiscale tCNN	0.91 (0.13)	0.98 (0.03)	0.93 (0.11)	0.91 (0.13)
Multiscale CNN-LSTM	0.90 (0.14)	0.98 (0.04)	0.91 (0.13)	0.89 (0.14)
Multiscale ConvLSTM	0.89 (0.14)	0.98 (0.04)	0.89 (0.13)	0.88 (0.15)
Multiscale Transformer	0.93 (0.11)	0.98 (0.04)	0.93 (0.11)	0.93 (0.12)

Detailed information about algorithm performance on exercise recognition is provided in Table 5.14 for traditional algorithms, and in Table 5.15 for the algorithms using multiscale feature learning.

A performance comparison between traditional and multiscale approaches in terms of cross-subject resilience is shown in Figure 5.18. The figure illustrates the accuracy values obtained by the evaluated models for both healthy and patients, highlighting the ability of multiscale feature learning to generalize better across subjects.

Multiscale architectures, especially the multiscale CNN-LSTM and the multiscale Transformer exhibit superior performance compared to traditional architectures. The multiscale

Table 5.14: Results of activity recognition algorithms trained using traditional architecture for all rehabilitation exercises.

Exercise	tCNN		CNN-LSTM		ConvLSTM		Transformer	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
0	0.85	0.85	0.88	0.89	0.84	0.84	0.82	0.80
1	0.89	0.90	0.93	0.92	0.87	0.88	0.81	0.80
2	0.91	0.91	0.85	0.85	0.86	0.84	0.79	0.73
3	0.91	0.87	0.92	0.89	0.89	0.82	0.79	0.74
4	0.86	0.88	0.89	0.87	0.87	0.86	0.73	0.73
5	0.88	0.90	0.90	0.91	0.91	0.89	0.76	0.72
6	0.98	0.98	0.99	0.98	0.98	0.98	0.98	0.98
7	0.87	0.87	0.92	0.94	0.88	0.89	0.84	0.82
8	0.98	0.93	0.94	0.91	0.86	0.87	0.83	0.80

Table 5.15: Results of activity recognition algorithms trained using multiscale architecture for all rehabilitation exercises.

Exercise	Multiscale tCNN		Multiscale CNN-LSTM		Multiscale ConvLSTM		Multiscale Transformer	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
0	0.88	0.89	0.88	0.88	0.85	0.83	0.89	0.90
1	0.87	0.85	0.89	0.90	0.83	0.83	0.94	0.93
2	0.93	0.92	0.85	0.84	0.87	0.86	0.88	0.88
3	0.88	0.88	0.90	0.88	0.90	0.84	0.93	0.90
4	0.90	0.91	0.90	0.87	0.86	0.87	0.90	0.91
5	0.91	0.90	0.93	0.92	0.89	0.87	0.88	0.90
6	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
7	0.95	0.95	0.91	0.92	0.88	0.90	0.94	0.95
8	0.94	0.91	0.93	0.88	0.94	0.91	0.96	0.97

models have significantly improved activity recognition accuracy, particularly for exercises that show considerable variability between subjects and variations in motion duration, such as those involving shoulder movements. These results highlight the effectiveness of multiscale architectures in enhancing activity recognition accuracy and robustness, which is especially beneficial for complex rehabilitation gestures like shoulder movements, which are characterized by their significant variability in movement patterns across different subjects and clinical conditions [219].

Proposed algorithms also outperformed state-of-the-art experiments based on the IRDS dataset. For example, systems proposed by Khanghah et al. (2023) achieved average accuracy values of 90% and 83% using 10-Fold and LOSO cross-validation, respectively [24], and average accuracies of 89% and 88% for 10-Fold and LOSO, respectively, when using time-series data derived from kinematics of 3D joint coordinates [144]. Similarly, Sadawi et al. (2019) used tCNN, LSTM, and k-nearest neighbors algorithm (KNN) on the IRDS dataset to develop a system able to recognize nine exercise rehabilitation exercises, showing an average accuracy of 85% for tCNN, 85% for LSTM, and 94% for KNN using raw path data of joint kinematics and LOSO cross-validation [272].

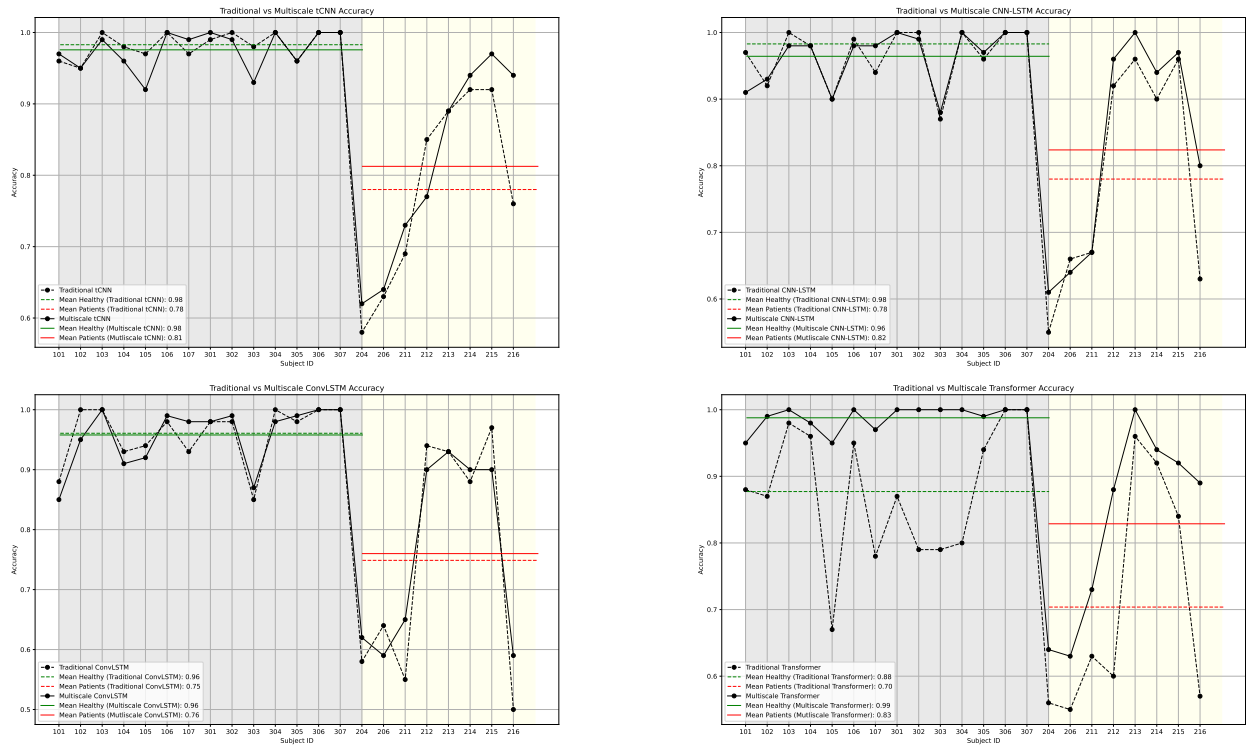


Figure 5.18: Comparison of traditional and multiscale model architectures for accuracy among healthy and patient subjects using LOSO cross-validation.

Furthermore, the multiscale approach demonstrated superior performance generalization across healthy individuals and patients with varying levels of disability, as illustrated in Figure 5.18. Specifically, multiscale algorithms exhibited enhanced performance consistency, highlighting their resilience in motion recognition across subjects with diverse clinical conditions. The ability of these architectures to learn patterns at different scales holds promise for developing fair and reliable activity recognition systems in healthcare, ensuring high accuracy across individuals with varying physical capabilities. However, these results are based on the IRDS dataset, which has inherent limitations, including significant variability in patient positions and recording settings. Although the proposed methods improved cross-subject performance in activity recognition, some subjects continued to exhibit suboptimal results, consistent with findings from previous studies [218, 24, 144, 272]. Future research should validate the multiscale learning approach using additional datasets to strengthen its robustness and generalizability.

Overall, the findings highlight the critical need for developing methods that ensure fairness in activity recognition algorithms for healthcare applications. Addressing disparities caused by impairments and health-related characteristics remains essential for creating inclusive and effective monitoring systems.

Chapter 6

Discussion

6.1 Contribution summary

The results of this research represent a pioneering advancement in the application of AI-driven technologies within the field of rehabilitation medicine.

A thorough literature review was conducted to explore emerging AI solutions and their associated ethical challenges within rehabilitation. This review established a foundational understanding of current technologies, identified research gaps, and highlighted potential ethical considerations regarding AI applications in healthcare. By synthesizing insights from various studies, the review evidenced the effectiveness and limitations of current AI methodologies in rehabilitation settings.

The research incorporated advanced data modeling techniques to enhance the evaluation algorithms for motion analysis. A pre-published dataset was utilized and modeled to support real-time motion evaluation systems, enabling immediate feedback on the execution of exercises. Additionally, a new dataset was collected focusing on rehabilitation exercises aimed at preventing physical decline in elderly individuals. This dataset underwent data synthesis techniques to augment environmental and physical characteristics, simulating the challenges faced by elderly individuals in home-based settings.

The study also addressed critical issues of fairness and robustness within clinical systems. It underscored the importance of expanding the annotations of data to include physical attributes and disabilities that influence the execution of motion. By addressing these factors, the research aimed to mitigate potential biases in AI-assisted assessments. This emphasis on fairness is crucial to develop AI systems that are equitable and effective across diverse patient populations. The research specifically explored pattern recognition techniques applied to multi-channel time-series data, utilizing advanced deep learning algorithms to enhance

motion tracking and evaluation.

Addressing fairness and robustness remains a key challenge in the advancement of AI-based support and assistive technologies in healthcare. Future work should expand these findings to include more diverse datasets, explore novel strategies to mitigate bias, and test the models in real-time systems. Solutions to mitigate demographic and disability-related biases were addressed by focusing on dataset manipulation, suggesting that future research should annotate such characteristics to train unbiased systems or incorporate more balanced attributes. In addition, techniques were explored to reduce performance variation among subjects with different motor abilities, ensuring that AI systems remain equitable and accessible to individuals with varying levels of disability. This approach is particularly important to ensure that AI-assisted rehabilitation systems do not create barriers for people with disabilities.

The ethical implications of AI-assisted rehabilitation systems were critically examined throughout the research process. This included concerns about patient privacy, data security, and potential biases in algorithmic decision making. While AI-powered decision support systems offer substantial benefits in clinical and care settings, they also pose significant challenges that require a thoughtful examination of ethical, legal, and regulatory aspects. Effective governance is essential to foster the acceptance and implementation of AI in healthcare. The research systematically addressed these ethical considerations, identifying both strengths and limitations. It emphasized the importance of establishing comprehensive ethical frameworks as AI technologies continue to evolve in healthcare. This is vital to ensure that AI applications enhance clinical outcomes and protect patient rights and privacy. Open challenges in this area are further explored in the below Section 6.3.

6.2 Limitations and future studies

Emerging AI-based computer vision approaches, which currently represent an intensive research topic, are promising for rehabilitation applications to provide home-based, inexpensive, and unobtrusive monitoring of patients [117].

Unobtrusive technologies coupled with algorithms that recognize performed activities will make the remote evaluation of patient adherence to prescribed exercise affordable, which was based on patients' self-reported outcomes in the absence of a clinician [34].

This thesis explored principally human motion modeling techniques to provide prompt corrective feedback about the correctness of motion execution in full ROM of motion and postural compensatory movements. Further evaluation should also concentrate on other aspects that influence the correctness of motion, such as the speed of motion execution.

Future research should prioritize developing techniques that enable quantitative evaluation of movement performance, complementing clinical outcome assessments provided by health-care professionals. This approach would facilitate the identification of digital biomarkers, offering deeper insights into therapy effectiveness and recovery from functional limitations. Also, assessment-driven therapy adaptation algorithms integrated into assistive solutions should ensure that the physical therapy remains at an optimal level of challenge to maintain motivation for long-term therapy programs necessary in chronic conditions [99]. Movement recognition end evaluation algorithms integrated into digital assistive systems could support real-time evaluations and provide personalized feedback. They can also drive emotional support to increase engagement during home-based rehabilitation therapy in clinical populations characterized by fragile psycho-physical conditions [309, 170]. For instance, also the integration of AI-based voice assistants could profoundly impact patient engagement during remote exercise by providing instructions on movement sequences and posture correctness and even offering suggestions to enhance exercise quality and identify areas requiring improvement.

Transparency and interpretability are also crucial to building trust in AI systems, particularly in medical applications for rehabilitation monitoring and evaluation. Although this thesis did not primarily implement explicit explainable AI (XAI) techniques, it prioritized an expert-driven approach by grounding motion data modeling and assessment in clinical knowledge to enhance the interpretability and transparency of proposed methodologies. As detailed in Section 3.1, a motion categorization approach was adopted to recognize and assess rehabilitation activities in real-time. This approach involved identifying movement phases and then classifying performance as correct or incorrect based on prior biomechanical knowledge of the specific motion being evaluated. Thus, the proposed AI methods were based on motion modeling and evaluation approaches using an interpretable expert-driven perspective by integrating biomechanics and rehabilitation science knowledge into human motion analysis. This methodology differs from the research works focusing mainly on automated scoring in evaluating movement, which can yield 'black-box' results with limited clinical relevance. In contrast, the methods proposed in this thesis prioritize grounding AI systems in support, rather than replace, medical knowledge and clinical practice. Based on expert knowledge, this system aims to ensure the reliability and trustworthiness of the proposed AI methodologies to recognize rehabilitation activities.

Nevertheless, in future works, it is essential to acknowledge that further efforts should explore XAI techniques to enhance the transparency and clinical applicability of AI-based activity recognition systems from an algorithmic perspective, an open research area, particularly for systems processing multidimensional signal data, as highlighted by recent literature [194]. Frameworks such as GNN-XAR and DeXAR demonstrate promising advances in inter-

pretability for sensor-based activity recognition [88, 16]. Likewise, adaptations of techniques such as Grad-CAM and the development of Video-TCAV contribute to the explainability of video-based action recognition [132, 9, 273]. In addition, model-agnostic tools like SHAP and LIME offer flexible means of interpreting various activity recognition models [187, 265], while attention mechanisms in deep learning models enhance interpretability by highlighting key movement features that influence predictions [71]. Despite growing evidence in the literature, significant challenges remain, particularly in designing XAI methods that are well-suited for evaluating algorithms processing human motion, given its inherent complexity. Ensuring that explanations are technically accurate, clinically meaningful, and actionable is crucial. Future research should focus on developing domain-specific explainability techniques and rigorously evaluating their effectiveness through relevant metrics to advance more transparent, trustworthy, and effective AI-driven rehabilitation systems that surpass the limitations of current methods.

As the future unfolds, ubiquitous systems offering remote assistance in rehabilitation procedures are expected to significantly complement traditional approaches for evaluating home rehabilitation programs. The real challenge of AI applications results in the incorporation of clinical skills and knowledge to assist and manage patients in a minimally supervised and more efficient way. Besides the potentiality of AI-based solutions to support the remote monitoring and evaluation of physical therapy programs, further developments of underlying algorithms and clinical validation of methods will be required for broader adoption. As also evidenced by recent literature a very minority of studies validated and demonstrated the application of proposed AI-based solutions in relevant operation environments such as remote conditions and home-based settings. For future developments of this research project and to transfer achieved innovations to clinical practice, it is necessary to collaborate with clinical organizations to validate these types of technologies in a relevant environment as well as to collect data in a real-world context without simulation. Future studies must focus on testing AI-based applications in relevant operational environments, such as remote conditions and home settings. Clinical trials comparing the effects on specific clinical population targets of such emerging solutions versus traditional rehabilitation approaches are required to lead AI-based applications to clinical acceptance. Social and ethical considerations will also need to be considered, as AI-based applications will change the interactions between healthcare professionals and patients in decentralized contexts.

6.3 Open challenges

A plethora of challenges emerges in the healthcare landscape with the growing integration of AI. These challenges span technological advancements, regulatory considerations, and ethical dimensions.

This section delves into the implications of AI technologies in clinical practice, focusing the discussion on two pivotal aspects: the doctor-patient relationship and AI-driven clinical decision-making. Additionally, challenges related to health record data are addressed, offering insights into current practices regarding the utilization process of medical data.

6.3.1 The doctor-patient relationship

The patient-doctor relationship is a critical aspect of healthcare, characterized by mutual trust, effective communication, and collaboration. Establishing trust between patients and healthcare providers is foundational for successful medical care. It involves open and transparent communication, empathy, and a shared understanding of the patient's concerns, values, and treatment preferences [106].

Issues related to trust within the patient-doctor relationship are multifaceted. Patients place trust in the expertise and knowledge of their healthcare providers, relying on their guidance for accurate diagnoses and effective treatments. At the same time, healthcare providers trust in the information shared by patients to make informed decisions about their care. Exploring these trust dynamics is essential for a comprehensive understanding of the patient-doctor relationship, especially when integrating AI into clinical decision support systems.

The incorporation of AI prompts a reassessment of the conventional doctor-patient relationship.

Some argue that the traditional model of a nurturing relationship has become outdated. It appears that the concept of patients seeking a doctor's expert advice and placing themselves under their care is evolving into a model where patients actively participate in generating health knowledge and acquiring expertise to manage their illnesses [89].

The healing relationship must be understood as an idealistic picture of the relationship between 'expert doctors' and 'vulnerable patients'. This concept encompasses the factors that drive patients to seek professional medical assistance or leverage knowledge and technology for self-care. Whether one opts for professional medical services or self-directed care, the fiduciary responsibilities arising from this vulnerability remain consistent, regardless of the sources of expertise. These sources can include medical professionals, repositories of

medical information and guidance, or other technologies and systems that support self-care, like telemedicine or easily accessible medical information on the internet.

In this context, the importance of redefining how fiduciary obligations of medicine are met takes on renewed significance, especially with the future deployment of AI in medicine. Questions of relevance have been raised, for instance, about the validity and effectiveness of medical knowledge available through Internet portals. Furthermore, as medical information becomes increasingly accessible through various means, the role of expertise as an indicator of trustworthiness does not undergo any change.

Built upon this foundation, the healing relationship model can be construed as an illustration of the ethical attributes and responsibilities intrinsic to medical practice. Traditionally, these principles were embodied by healthcare professionals, but they are now extending across diverse platforms and individuals, encompassing web portals, consumer device creators, wellness service providers, and others. While contemporary medicine has progressed beyond the conventional doctor-patient model delineated in the therapeutic relationship, the responsibilities associated with this connection have not evaporated. Instead, the diffusion and transfer of these responsibilities to new technological participants in medicine raise concerns regarding the oversight of AI integration in healthcare.

The selection of metrics becomes paramount when assessing the impact of AI and algorithmic technologies on the doctor-patient relationship. The rationale for incorporating AI into healthcare and expanding its role becomes evident if it is solely gauged in terms of cost-effectiveness or utility. Nonetheless, while algorithmic technologies might enhance efficiency and reduce the cost of treating more patients, they could potentially erode the non-mechanical aspects of care. One can distinguish between the effects of algorithmic systems (and their utility components) that contribute to the welfare of the patient or the practice of medicine, governed by established ethical standards, and those that benefit medical institutions and healthcare services.

The moral engagement intrinsic to the doctor-patient relationship, where treatment ideally stems from the practitioner's contextual and historically informed evaluation of the patient's condition, is challenging to replicate in interactions with AI systems. The role of the patient, the factors prompting individuals to seek medical assistance, and the vulnerability of patients remain unaltered with the introduction of AI as a mediator or enhancer of medical care. What does transform is how care is administered, the possibilities for its delivery, and who delivers it. The delegation of care skills and responsibilities to AI systems can be disruptive in multiple respects.

The deployment of AI machines and robots in medicine, particularly when they demonstrate enhanced efficiency, precision, speed, and cost-effectiveness, appears appealing when

considering their substitution for humans in tasks that are repetitive, tedious, perilous, demeaning, or exhausting. If used judiciously, AI has the potential to reduce the time healthcare professionals allocate to bureaucratic, routine tasks, or those that expose them to unnecessary risks, thereby offering a lower-risk, patient-focused environment.

Traditionally, clinical care and the doctor-patient relationship are ideally rooted in the physician's contextual and historically informed evaluation of the patient's condition. This type of care is challenging to replicate in technologically mediated healthcare. Patient data representations inherently confine the doctor's comprehension of the patient's case to quantifiable attributes. This can pose challenges as clinical assessments increasingly rely on data representations derived from sources like remote monitoring technologies or data collected without in-person interactions.

Patient data representations can be perceived as an 'objective' gauge of an individual's health and overall well-being, potentially diminishing the significance of contextual health elements or the perspective of the patient as a socially situated individual. These data representations can generate an 'illusion of certainty', where 'objective' monitoring data are regarded as an accurate portrayal of the patient's condition, often overlooking the patient's interpersonal context and other unspoken knowledge [188].

Healthcare providers encounter this challenge when integrating AI systems into patient care protocols. The volume and intricacy of data and suggestions derived from technology can complicate the identification of crucial missing contextual details regarding a patient's condition. Depending solely on data obtained from health apps or monitoring technologies (e.g., smartwatches) as the primary information source about a patient's health may lead to the oversight of facets of a patient's well-being that are not readily measurable. This encompasses vital aspects of mental health and overall well-being, such as the patient's emotional, mental, and social status. Consequently, a 'decontextualization' of the patient's state may transpire, wherein the patient relinquishes some influence over how their condition is conveyed and comprehended by healthcare professionals and caregivers. These scenarios imply that the interactions essential for cultivating the fundamental trust inherent in a traditional doctor-patient relationship may be impeded by technological intervention. Technologies that obstruct the conveyance of 'psychological signals and emotions' may impede the physician's comprehension of the patient's condition, jeopardizing 'the establishment of a physician-patient relationship grounded in trust and the pursuit of healing' [89].

Serving as an intermediary position between the doctor and the patient, AI systems modify the dynamics between healthcare providers and patients, as they delegate a portion of the patient's continuous care to a technological platform. This shift could lead to a growing divide between healthcare professionals and patients, potentially implying a missed chance

to cultivate an intuitive grasp of the patient's health and overall well-being [188].

Relying on AI systems for clinical care or expert diagnostic capabilities may hinder the cultivation of expertise, professional networks, and the establishment of 'best practice' standards in the field of medicine. This phenomenon is commonly known as 'de-skilling' and contradicts the principles of 'human-centered AI' advocated by the WHO. Human-centered AI aims to bolster and enrich human skills and competency development rather than diminishing or substituting them [61].

As care becomes increasingly technologically mediated and involves non-professional entities, the development, maintenance, and enforcement of internal standards aimed at fulfilling moral obligations to patients may face potential compromises. There is a looming prospect that algorithmic systems could supplant the traditional roles of healthcare professionals, with a primary focus on efficiency and cost-effectiveness.

To safeguard against the erosion of comprehensive, patient-centered care, new care providers and entities outside the conventional medical community must place significant emphasis on upholding these moral obligations to benefit and respect patients. In doing so, they can ensure that healthcare remains not only technically 'efficient' but also holistic and genuinely beneficial.

A key role of human clinical expertise is to safeguard the well-being and safety of patients. When this human expertise is compromised through skill degradation or displaced by automation bias, it becomes essential for tests and clinical effectiveness trials to step in and bridge the gap, thus ensuring patient safety [148]. This trade-off is mirrored in the context of transparency and precision. Some scholars contend that medical AI systems may not require comprehensible explanations if their clinical accuracy and effectiveness can be consistently verified [233].

Automation in data acquisition, interpretation, diagnosis processing, and therapy identification cannot operate in complete isolation from human involvement. It necessitates continuous validation and, therefore, does not diminish the significance of the doctor-patient relationship's uniqueness. Each individual's ailment is inherently distinctive, and personal interaction remains the fundamental component of every diagnosis and treatment. In this context, machines cannot replace humans in a relationship founded on the interplay of complementary realms of autonomy, competence, and responsibility.

AI should be regarded solely as a tool to aid physicians in their decision-making processes, which must always be under human control and supervision. Ultimately, the responsibility for making the final decision rests with the doctor, as the machine's role is exclusively supportive, involving the gathering and data analysis, serving in an advisory capacity. It is important to underscore that an 'automated cognitive assistance system' in diagnostic and

therapeutic procedures does not equate to an ‘autonomous decision-making system’ [148]. It assists by collating clinical and documentary data, comparing them with statistics on similar patients, and expediting the physician’s analytical process.

An important concern arises: what happens when AI outperforms the capabilities of a doctor? In specific scenarios, this is indeed a technical possibility that must be considered. It is within this particular realm that the feared ‘replacement’ of machines for humans might potentially occur in the future.

However, a more immediate consequence might involve the delegation of decision-making to technology. Entrusting complex tasks to AI systems can result in the erosion of human and professional qualities. To maintain the doctor-patient relationship as one grounded in trust, in addition to care, it is crucial to preserve the essential role of the ‘human doctor’. Only the human doctor possesses the unique capacities of empathy and genuine understanding that cannot be replicated by AI. While predetermined standards of behavior and codes of conduct, such as protocols and guidelines, provide support based on knowledge and experience in professional practice, the demands of diagnosis and treatment often necessitate deviation from these predetermined models.

It would be of great concern if the space seemingly left to the presumed neutrality of machines led to the ‘neutralization’ of the patient. The vast potential offered by AI should be viewed as a valuable opportunity through which technology can expand ethical horizons, enhancing the patient’s opportunity to be heard and fostering a deeper connection with the progression of their illness. In this context, AI serves as a valuable tool that saves the physician time on routine tasks, allowing more time to be dedicated to the doctor-patient relationship.

6.3.2 AI driving medical decision-making

A substantial portion of the momentum in AI stems from the belief that applying these technologies in diagnosis, care, or healthcare systems could enhance clinical and institutional decision-making.

Physicians and healthcare professionals are susceptible to various cognitive biases, leading to diagnostic errors. The National Academy of Sciences (NAS) has reported that approximately 5% of US adults seeking healthcare guidance are subject to misdiagnosis, with such errors contributing to 10% of all patient fatalities [305].

AI has the potential to mitigate inefficiencies and errors, leading to a more judicious allocation of resources, provided that the foundational data is both accurate and representative. Accountability plays a pivotal role in holding individuals and entities responsible for any

adverse consequences of their actions. It is an indispensable element for upholding trust and safeguarding human rights.

Nevertheless, certain attributes of AI technologies introduce complexities to notions of accountability. These attributes include their opacity, reliance on human input, interactions, discretionary functions, scalability, capacity to uncover concealed insights and software intricacies. One particular challenge in ascribing responsibility arises from the ‘control problem’ associated with AI. In this scenario, AI developers and designers may not be held accountable, as AI-driven systems can operate independently and evolve in ways that the developer may assert are unpredictable [235]. Assigning responsibility to the developer could provide an incentive to take all possible measures to minimize harm to the patient. Such expectations are already well established for manufacturers of other commonly used medical technologies, including drug and vaccine manufacturers, medical device companies, and medical equipment manufacturers.

Another challenge is the ‘traceability of harm’, which is a persistent issue in complex decision-making systems within healthcare and other domains, even in the absence of AI. Due to the involvement of numerous agents in AI development, ascribing responsibility is a complex task, entailing both legal and moral dimensions. The diffusion of responsibility can have adverse consequences, including the lack of compensation for individuals who have suffered harm, incomplete identification of the harm and its root causes, unaddressed harm, and potential erosion of societal trust in such technologies when it appears that neither developers nor users can be held accountable [333].

Physicians routinely employ various non-AI technologies in the diagnosis and treatment process, ranging from X-rays to computer software. When a medical professional commits an error in the utilization of such technology, they can be held accountable, especially if they have received training in its application [107]. However, in cases where an error arises from the algorithm or data used to train an AI technology, it may be more appropriate to assign responsibility to those involved in the development or testing of the AI system. This approach avoids placing the onus on the doctor to assess the AI technology’s effectiveness and usefulness, as they might not possess the expertise to evaluate complex AI systems [269].

Numerous factors argue against placing exclusive responsibility on physicians for decisions made by AI technologies, many of which apply to assigning responsibility for the use of healthcare technologies beyond AI. To begin with, physicians do not wield control over an AI-driven technology or the recommendations it provides [36]. Nonetheless, physicians should not be entirely absolved of liability for inaccuracies in the content, as this is necessary to prevent ‘automation bias’ and encourage critical evaluation of whether the technology aligns with their requirements and those of the patients [333]. Automation bias occurs when a

physician disregards errors that should have been identified through human-guided decision-making. While it is crucial for doctors to have trust in an algorithm, they should not set aside their own experience and judgment to blindly endorse a machine's recommendation. Automation bias occurs when a physician disregards errors that should have been identified through human-guided decision-making. While it is crucial for doctors to have trust in an algorithm, they should not set aside their own experience and judgment to blindly endorse a machine's recommendation [269].

Certain AI technologies may present not just a single decision but a range of options from which a physician must make a selection. When a physician makes an incorrect choice, determining the criteria for holding them accountable becomes a multifaceted challenge.

The complexity of attributing liability is magnified when AI technology is integrated into a healthcare system. In such cases, the developer, the institution, and the physician may all have contributed to a medical error, making it difficult to pinpoint full responsibility [305]. Consequently, accountability might not rest solely with the provider or developer of the technology; instead, it may lie with the government agency or institution responsible for selecting, validating, and implementing the technology.

Fortunately, as of today, the shift of decision-making in healthcare from humans to machines has not reached its culmination. While AI today is primarily proposed to augment human decision-making in the practice of public health and medicine, epistemic authority has begun to debate the issue of why, in some circumstances, AI systems (as with the use of computer simulations) can move humans from the center of knowledge production [18, 77]. The debates evidence a possible complete transfer of routine medical tasks to AI, prompting questions about the legality of such full delegation. Modern laws increasingly acknowledge an individual's right not to be solely subjected to automated decisions when these decisions hold substantial consequences. Additional concerns may emerge if human judgment is progressively supplanted by machine-driven assessment, giving rise to broader ethical considerations associated with the loss of human oversight, particularly if predictive healthcare becomes standard practice.

Nonetheless, it is improbable that AI in the field of medicine will attain complete autonomy; instead, it may reach a stage of conditional automation or necessitate ongoing human support [127].

Substituting human judgment with AI and relinquishing control over certain aspects of clinical care offers distinct advantages. Humans are capable of making decisions that may be less equitable and more biased compared to machines (the concern about bias in AI usage is further elaborated below).

The utilization of AI systems for specific, well-defined decisions can be entirely justified

if there is compelling clinical evidence indicating that the system outperforms a human in that particular task. However, the transition to the application of AI technologies for more intricate aspects of clinical care presents a set of challenges.

One such challenge is the potential emergence of a ‘disagreement between equals’ between two competent experts: an AI machine and a doctor. In such scenarios, there is no feasible method for harmonizing decisions or engaging in a rationale with the algorithm since accessing it or committing to a change of mind is not possible. Furthermore, there are no clear-cut rules for determining which entity is correct, and whether a patient should place trust in technology or a doctor. The choice may hinge on factors that are not rooted in the ‘competence’ of either the machine or the doctor. Some argue that the algorithm’s recommendation should be favored, as it incorporates the expertise of multiple individuals and a wealth of data points. However, opting for one over the other may result in an undesirable outcome. If the physician disregards the machine’s recommendation, the AI’s value addition can be limited. If the physician concurs with the machine’s decision, it might erode their authority and diminish their responsibility [305].

The delegation of decision-making to AI-driven technologies and the subsequent loss of human control could impact various aspects of clinical care and the healthcare system.

While offering individuals more opportunities to share data and access autonomous health advice might enhance their sense of empowerment and self-care, it could also potentially trigger anxiety and fatigue [107]. As these technologies accumulate more personal data and incorporate it into the decision-making process of physicians, patients might find themselves gradually marginalized in shared decision-making, potentially compromising their ability to exercise free will or autonomy in health-related decisions [305].

The utilization of AI in medicine, especially when its use is not disclosed, poses challenges to the fundamental principles of informed consent and broader public trust in healthcare. While some may view the reduction of physician control over patients as a way to promote patient autonomy, there is an equally significant risk of relinquishing decision-making to AI technology. This risk is heightened if the technology is presented to the patient as offering a superior understanding of their health status and prognosis compared to a physician [269].

6.3.3 Health data records

The integration of AI technology in the healthcare sector faces notable limitations, particularly in the realm of managing health record data. Health data encompasses a patient’s complete medical history, including details from physical examinations, investigations, and treatment records—all stored in electronic form.

The 2018 WHO-UNICEF Global Primary Health Care (PHC) Conference underscored the importance of harnessing AI techniques to augment existing systems and leverage data for electronic decision support and analytics [241]. The overarching objective is to attain Universal Health Coverage and advance healthcare equity for both individuals and populations. However, ethical challenges arise in the context of health records. The use of electronic health records introduces ethical issues related to beneficence, autonomy, fidelity, and justice. Autonomy is compromised when patients' health data are shared or linked without their knowledge. Fidelity is at risk when health organizations fail to take adequate precautions to safeguard identifiable health data. Justice is undermined when disparities exist in access to health information and services based on income, language, age, geography, literacy, and disability [167].

A comprehensive examination of health records and their implications necessitates the involvement of various stakeholders, including health personnel, leaders, policymakers, ethicists, and engineers.

Despite significant strides in computerization, challenges persist throughout the data life cycle—from collection and documentation to storage, management, sharing, and utilization [68, 179]. Efforts to address these challenges are essential to fully realize the potential benefits of AI in healthcare.

Legislation for AI regulation in the EU emphasizes trustworthy and ethical AI with explicit processes and human oversight for high-risk systems. Equitable AI is proposed as a solution, emphasizing fairness in capacity-building strategies. Integrating data quality assessment and management with information governance in the big data environment is essential for ethical health information ecosystems.

To effectively navigate the complexities of integrating AI technology into healthcare, a comprehensive approach that combines data usage, data quality management, governance, and ethics is essential. In this context, key recommendations have been identified to guide stakeholders in the enhancement of data utilization in healthcare: transparency, use limitation, access and correction, data quality, and security [329].

Healthcare professionals play significant roles as creators, collectors, managers, and users of observational health data, raising the question of whether new privacy consent models are needed at different points in the data life cycle. Determining the adequacy of consent, its informativeness, and relevance becomes a critical consideration, prompting the need for clear criteria and decision points throughout the life cycle.

In promoting good governance, it is essential to establish consent or employ other legal processes, such as opting out or utilization for public health purposes, to regulate access and sharing. Adhering to ethical principles, data quality categories, and information gover-

nance is crucial, ensuring their relevance at various stages in the data life cycle. Innovative technologies like blockchain are actively being utilized to expedite the consent process for clinical trials, potentially alleviating concerns related to personal data privacy associated with integrated service delivery [29].

Establishing and maintaining trust is paramount in implementing sustainable data creation and collection practices.

It is crucial to recognize that community-driven health data repositories may not provide the level of privacy citizens assume, particularly in the context of free online services with user agreements allowing the service owner to utilize collected data. The deployment of applications and systems facilitating unethical and unlawful exchange of digital information erodes trust. Building and maintaining trust in the health information exchange network requires reciprocity, transparency, and mutual trust among all actors, including data custodians and providers. The trustworthiness of data custodians hinges on their competence, commitment, and motives. Patients and service users generally have high levels of trust in the professionalism of clinical teams and health services. However, unnecessary tests and procedures resulting from unclear data or misinterpretations of unstructured data can compromise patient autonomy and hinder informed decision-making.

In the United States, legislation such as the Health Insurance Portability and Accountability Act (HIPAA) specifically excludes protections once data leaves a covered entity. Controversy surrounds the transmission of data to large companies for data mining, primarily due to the lack of explicit consent from patients, placing control in the hands of system designers and owners. In contrast, the EU's GDPR empowers individuals with control over their personal data.

Careful consideration is needed for all data processes, as they may harbor unrecognized risks.

Transferring data from information systems to data repositories demands a commitment to security, safety, and accuracy [255]. Privacy is a paramount concern, and while privacy-preserving linkage techniques exist for integrating observational data from information systems, their accuracy and security are not always guaranteed [65]. Beyond the threat of re-identification during data, there exists the potential for data loss and compromise of data integrity.

For instance, a recent study underscored inaccuracies in cohort identification when utilizing vocabulary mappings within a common data model during the data process [111]. These mappings, integral to data, may suffer from inaccuracies stemming from programming bugs and errors that escape detection during quality assurance stages.

It is imperative to thoroughly identify, assess, and establish contingency plans to mitigate

all risks associated with the data process. Vigilance in recognizing potential pitfalls and implementing safeguards is essential for ensuring the reliability and security of the data transfer process [178].

Transparent data processing is imperative, along with the ethical and secure sharing of research data, methodologies, and algorithms. Achieving reproducibility and generalizability while safeguarding patient privacy, financial investments, and intellectual property is essential.

XAI algorithms allow doctors to understand and validate results, aligning with the ‘learned intermediary’ principle where clinicians play a central role in decision-making. However, discerning between biased or inadequately explained AI guidance and the clinician’s interpretation poses challenges. Collaborative efforts among clinicians can potentially address these biases and enhance the design of XAI, ensuring a critical appraisal of AI guidance in primary care that augments rather than undermines the patient-physician relationship [181].

Conscious efforts to develop discrimination-aware algorithms aim to reduce bias in databases and associated applications, ensuring that data mining models do not lead to discriminatory decisions against vulnerable groups, even if the underlying dataset is inherently biased. Empirical studies evaluating AI practices are necessary to highlight ethical concerns and propose solutions. Trials comparing AI-based practices with conventional methods are vital for evidence-based policymaking in the face of poorly regulated AI practices. Reducing the opacity and complexity of AI methods, including deep machine learning and neural networks, is imperative.

The ethical governance of AI platforms necessitates modifications to traditional medical research ethics principles. This includes considerations for informed consent, protecting individual and group-level harms and benefits, ensuring patient empowerment, preserving the patient-doctor relationship, and upholding research subject rights in AI-supported projects. Data protection regulations for research and personal health services must be robust and effectively enforced.

In conclusion, adopting an integrated strategy encompassing data utilization, quality assurance, governance, and ethical considerations is pivotal. Key recommendations involve ensuring consent remains a focal point throughout the data life cycle, fostering sustainable data practices, meticulously examining data processes, seamlessly integrating data governance with quality management, navigating ethical challenges posed by AI, and applying a structured ethical framework aligned with the data life cycle. These proactive measures collectively strive to uphold trust in current systems and anticipated developments.

Chapter 7

Conclusions

This thesis represents a pioneering exploration of emerging AI-driven computational methods and applications designed to assist rehabilitation medicine.

This thesis investigates AI algorithms designed for motion monitoring and evaluation in rehabilitation exercises. These algorithms were explored to assess exercise quality, facilitating personalized adjustments to rehabilitation plans based on individual patient data. Through real-time analysis of movement patterns, the proposed technologies support continuous patient monitoring, offering timely feedback during exercise sessions. This approach not only ensures adherence to prescribed therapeutic regimens but also enhances the overall quality of care by promoting precise, data-driven decision-making in rehabilitation programs.

The integration of AI-powered monitoring systems into home-care settings has the potential to alleviate the burden on healthcare infrastructure, improve patient outcomes, and bridge the accessibility gap for individuals in remote or underserved regions. These innovations pave the way for a sustainable model of delivering personalized rehabilitation services, fundamentally redefining care delivery within modern healthcare ecosystems. This thesis aims to drive a paradigm shift from traditional clinic-centered care to a home-based model by demonstrating the effectiveness of AI assistive solutions in rehabilitation. This transformation holds the potential to provide more tailored, efficient, and accessible healthcare services, meeting the evolving needs of diverse patient populations while optimizing resource utilization. AI assistant tools will provide guidance, encouragement, and reminders, fostering greater self-empowerment and adherence to prescribed therapy regimens. This is particularly vital for elderly patients who may face challenges in adhering to rehabilitation plans due to diminished motivation to engage consistently in therapeutic exercises at home. AI-assisted systems can provide patients with personalized feedback on their performance, motivating them to adhere more consistently to their prescribed rehabilitation protocols. This approach has the potential to improve clinical outcomes and enhance patient satisfaction and overall

quality of life. The integration of AI into home-based rehabilitation services holds significant potential to reduce operational costs while enhancing service delivery efficiency. This aligns with European policy initiatives to improve care for aging populations and address the growing need for innovative healthcare solutions that can adapt to societal changes driven by demographic aging.

The broader impact of this research lies in redefining rehabilitation services by fostering more accessible, responsive, and efficient healthcare solutions. It addresses the urgent global need for high-quality, cost-effective, and scalable home-based treatments, which are increasingly crucial as healthcare systems adapt to the growing demands of an aging population.

Bibliography

- [1] Council directives 90/385/eec and 93/42/eec., .
- [2] European commission, commission staff working document, impact assessment, accompanying the proposal for a regulation laying down harmonized rules on artificial intelligence (artificial intelligence act), swd (2021) 84 final, part 1/2, p. 1., .
- [3] European commission. proposal for a regulation laying down harmonised rules on artificial intelligence (artificial intelligence act); com (2021) 206 final; european commission: Brussels, belgium, 2021., .
- [4] Regulation (eu) 2017/745 of the european parliament and of the council of 5 april 2017 on medical devices, amending directive 2001/83/ec, regulation (ec) no 178/2002 and regulation (ec) no 1223/2009 and repealing council directives 90/385/eec and 93/42/eec.
- [5] What is 'biosurveillance'? the covid-19 measures getting under our skin. <https://medium.com/digital-freedom-fund/what-is-biosurveillance-c8bffe70d16f>, 2020.
- [6] Rajeev Aggarwal, Suvarna Shyam Ganvir, et al. Artificial intelligence in physiotherapy. *Physiotherapy-The Journal of Indian Association of Physiotherapists*, 15(2):55, 2021.
- [7] Muhammad Usman Akbar, Wuhao Wang, and Anders Eklund. Beware of diffusion models for synthesizing medical images—a comparison with gans in terms of memorizing brain tumor images. *arXiv preprint arXiv:2305.07644*, 2023.
- [8] Ahmed Al Kuwaiti, Khalid Nazer, Abdullah Al-Reedy, Shaher Al-Shehri, Afnan Al-Muhanna, Arun Vijay Subbarayalu, Dhoha Al Muhanna, and Fahad A Al-Muhanna. A review of the role of artificial intelligence in healthcare. *Journal of Personalized Medicine*, 13(6):951, 2023.

- [9] Md Tasnim Alam, Yaa Takyiwaa Acquaaah, and Kaushik Roy. Image-based human action recognition with transfer learning using grad-cam for visualization. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pages 117–130. Springer, 2024.
- [10] Mohammad Arif Ul Alam. Ai-fairness towards activity recognition of older adults. In *MobiQuitous 2020-17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, pages 108–117, 2020.
- [11] Claudia Álvarez-Aparicio, Ángel Manuel Guerrero-Higueras, Miguel Ángel González-Santamarta, Adrián Campazas-Vega, Vicente Matellán, and Camino Fernández-Llamas. Biometric recognition through gait analysis. *Scientific Reports*, 12(1):14530, 2022.
- [12] Dustin Anderson. Artificial intelligence and applications in pm&r. *American Journal of Physical Medicine & Rehabilitation*, 98(11):e128–e129, 2019.
- [13] Alessa Angerschmid, Jianlong Zhou, Kevin Theuermann, Fang Chen, and Andreas Holzinger. Fairness and explanation in ai-informed decision making. *Machine Learning and Knowledge Extraction*, 4(2):556–579, 2022.
- [14] João Antunes, Alexandre Bernardino, Asim Smailagic, and Daniel P Siewiorek. Aha-3d: A labelled dataset for senior fitness exercise recognition and segmentation from 3d skeletal data. In *BMVC*, page 332, 2018.
- [15] Rob Argent, Ailish Daly, and Brian Caulfield. Patient involvement with home-based exercise programs: Can connected health interventions influence adherence? *JMIR Mhealth Uhealth*, 6(3):e47, Mar 2018.
- [16] Luca Arrotta, Gabriele Civitarese, and Claudio Bettini. Dexar: Deep explainable sensor-based activity recognition in smart-home environments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(1):1–30, 2022.
- [17] Letícia Avellar, Carlos Stefano Filho, Gabriel Delgado, Anselmo Frizera, Eduardo Rocon, and Arnaldo Leal Junior. Ai-enabled photonic smart garment for movement analysis. *Scientific Reports*, 12:4067, 03 2022.
- [18] Boris Babic, Sara Gerke, Theodoros Evgeniou, and I Glenn Cohen. Beware explanations from ai in health care. *Science*, 373(6552):284–286, 2021.

- [19] Michael J Bade and Jennifer E Stevens-Lapsley. Early high-intensity rehabilitation following total knee arthroplasty improves outcomes. *journal of orthopaedic & sports physical therapy*, 41(12):932–941, 2011.
- [20] Shaojie Bai, J Zico Kolter, and Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018.
- [21] Mirza Mansoor Baig, Hamid GholamHosseini, Aasia A Moqem, Farhaan Mirza, and Maria Lindén. A systematic review of wearable patient monitoring systems—current challenges and opportunities for clinical adoption. *Journal of medical systems*, 41:1–9, 2017.
- [22] Guha Balakrishnan, Amy Zhao, Adrian V Dalca, Fredo Durand, and John Guttag. Synthesizing images of humans in unseen poses. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8340–8348, 2018.
- [23] Shane Barratt and Rishi Sharma. A note on the inception score. *arXiv preprint arXiv:1801.01973*, 2018.
- [24] Ali Barzegar Khanghah, Geoff Fernie, and Atena Roshan Fekr. Design and validation of vision-based exercise biofeedback for tele-rehabilitation. *Sensors*, 23(3):1206, 2023.
- [25] Sandra F Bassett and Harry Prapavessis. Home-based physical therapy intervention with adherence-enhancing strategies versus clinic-based management for patients with ankle sprains. *Physical therapy*, 87(9):1132–1143, 2007.
- [26] Luckshman Bavan, Karl Surmacz, David Beard, Stephen Mellon, and Jonathan Rees. Adherence monitoring of rehabilitation exercise with inertial sensors: A clinical validation study. *Gait & Posture*, 70:211–217, 2019.
- [27] Tom L Beauchamp and James F Childress. *Principles of biomedical ethics*. Oxford University Press, USA, 2001.
- [28] Gillinder Bedi, Facundo Carrillo, Guillermo A Cecchi, Diego Fernández Slezak, Mariano Sigman, Natália B Mota, Sidarta Ribeiro, Daniel C Javitt, Mauro Copelli, and Cheryl M Corcoran. Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophrenia*, 1(1):1–7, 2015.
- [29] Mehdi Benchoufi and Philippe Ravaud. Blockchain technology for improving clinical research quality. *Trials*, 18(1):1–5, 2017.

- [30] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828, 2013.
- [31] Victor C Espinoza Bernal, Shivayogi V Hiremath, Bethany Wolf, Brooke Riley, Rochelle J Mendonca, and Michelle J Johnson. Classifying and tracking rehabilitation interventions through machine-learning algorithms in individuals with stroke. *Journal of Rehabilitation and Assistive Technologies Engineering*, 8:20556683211044640, 2021. PMID: 34646574.
- [32] Johanna Theresia Biebl, Marzena Rykala, Maximilian Strobel, Pawandeep Kaur Bollinger, Bernhard Ulm, Eduard Kraft, Stephan Huber, and Andreas Lorenz. App-based feedback for rehabilitation exercise correction in patients with knee or hip osteoarthritis: Prospective cohort study. *J Med Internet Res*, 23(7):e26658, Jul 2021.
- [33] Ali Borji. Pros and cons of gan evaluation measures. *Computer vision and image understanding*, 179:41–65, 2019.
- [34] Issam Boukhenoufa, Xiaojun Zhai, Victor Utti, Jo Jackson, and Klaus D. McDonald-Maier. Wearable sensors and machine learning in post-stroke rehabilitation assessment: A systematic review. *Biomedical Signal Processing and Control*, 71:103197, 2022.
- [35] Hendrio Bragança, Juan G Colonna, Horácio ABF Oliveira, and Eduardo Souto. How validation methodology influences human activity recognition mobile systems. *Sensors*, 22(6):2360, 2022.
- [36] Matthias Braun, Patrik Hummel, Susanne Beck, and Peter Dabrock. Primer on an ethics of ai-based decision support systems in the clinic. *Journal of medical ethics*, 47(12):e3–e3, 2021.
- [37] Louise Brennan, Antonio Bevilacqua, Tahar Kechadi, and Brian Caulfield. Segmentation of shoulder rehabilitation exercises for single and multiple inertial sensor systems. *Journal of Rehabilitation and Assistive Technologies Engineering*, 7:2055668320915377, 2020. PMID: 32913661.
- [38] R Brown. Challenges to successful ai implementation in healthcare. *Data Science Central*, 2022.
- [39] David Burns, Philip Boyer, Helen Razmjou, Robin Richards, and Cari Whyne. Adherence patterns and dose response of physiotherapy for rotator cuff pathology: Longitudinal cohort study. *JMIR Rehabil Assist Technol*, 8(1):e21374, Mar 2021.

- [40] David M Burns, Nathan Leung, Michael Hardisty, Cari M Whyne, Patrick Henry, and Stewart McLachlin. Shoulder physiotherapy exercise recognition: machine learning the inertial signals from a smartwatch. *Physiological measurement*, 39(7):075007, 2018.
- [41] Chris Burtin, Beatrix Clerckx, Christophe Robbeets, Patrick Ferdinande, Daniel Langer, Thierry Troosters, Greet Hermans, Marc Decramer, and Rik Gosselink. Early exercise in critically ill patients enhances short-term functional recovery. *Critical care medicine*, 37(9):2499–2505, 2009.
- [42] Jonathan Bush. How ai is taking the scut work out of health care. *Harvard Business Review*, 5, 2018.
- [43] Ting Cao, Mohammad Ali Armin, Simon Denman, Lars Petersson, and David Ahmedt-Aristizabal. In-bed human pose estimation from unseen and privacy-preserving image domains. In *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, pages 1–5. IEEE, 2022.
- [44] Marianna Capecci, Maria Gabriella Ceravolo, Francesco Ferracuti, Sabrina Iarlori, Andrea Monteriu, Luca Romeo, and Federica Verdini. The kimore dataset: Kinematic assessment of movement and clinical scores for remote monitoring of physical rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(7):1436–1448, 2019.
- [45] Valeria Caso and Antonio Federico. No lockdown for neurological diseases during covid19 pandemic infection, 2020.
- [46] Simon Caton and Christian Haas. Fairness in machine learning: A survey. *ACM Computing Surveys*, 2020.
- [47] Guillermo Cecchi. With ai, our words will be a window into our mental health, 2017.
- [48] Sang Hoon Chae, Yushin Kim, Kyoung-Soub Lee, and Hyung-Soon Park. Development and clinical evaluation of a web-based upper limb home rehabilitation system using a smartwatch and machine learning model for chronic stroke survivors: Prospective comparative study. *JMIR Mhealth Uhealth*, 8(7):e17216, Jul 2020.
- [49] Baoyu Chen, Yi Zhang, Hongchen Tan, Baocai Yin, and Xiuping Liu. Pman: Progressive multi-attention network for human pose transfer. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(1):302–314, 2021.

- [50] Chen Chen, Roozbeh Jafari, and Nasser Kehtarnavaz. Utd-mhad: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor. In *2015 IEEE International conference on image processing (ICIP)*, pages 168–172. IEEE, 2015.
- [51] An-Ti Chiang, Qi Chen, Yao Wang, and Mei R. Fu. Kinect-based in-home exercise system for lymphatic health and lymphedema intervention. *IEEE Journal of Translational Engineering in Health and Medicine*, 6:1–13, 2018.
- [52] Swathikan Chidambaram, Yathukulan Maheswaran, Kian Patel, Viknesh Sounderajah, Daniel A Hashimoto, Kenneth Patrick Seastedt, Alison H McGregor, Sheraz R Markar, and Ara Darzi. Using artificial intelligence-enhanced sensing and wearable technology in sports medicine and performance optimisation. *Sensors*, 22(18):6920, 2022.
- [53] A Chłon-Dominczak et al. Population ageing in europe: Facts, implications and policies research and innovation. *European Commission*. Available at: https://ec.europa.eu/research/social-sciences/pdf/policy_reviews/kina26426enc.pdf (Accessed: 30 October 2019), 2014.
- [54] Terese L Chmielewski, Steven Z George, Susan M Tillman, Michael W Moser, Trevor A Lentz, Peter A Indelicato, Troy N Trumble, Jonathan J Shuster, Flavia M Cicuttini, and Christiaan Leeuwenburgh. Low-versus high-intensity plyometric exercise during rehabilitation after anterior cruciate ligament reconstruction. *The American journal of sports medicine*, 44(3):609–617, 2016.
- [55] François Chollet et al. Keras. <https://github.com/fchollet/keras>, 2015.
- [56] Kah Meng Chong. Privacy-preserving healthcare informatics: a review. *ITM Web of Conferences*, 36:04005, 01 2021.
- [57] Chun-Yang Chou, Ding-Yang Hsu, and Chun-Hung Chou. Predicting the onset of diabetes with machine learning methods. *Journal of Personalized Medicine*, 13(3):406, 2023.
- [58] Avishek Choudhury and Onur Asan. Impact of accountability, training, and human factors on the use of artificial intelligence in healthcare: Exploring the perceptions of healthcare practitioners in the us. *Human Factors in Healthcare*, 2:100021, 2022.
- [59] Jen-Li Chung, Lee-Yeng Ong, and Meng-Chew Leow. Comparative analysis of skeleton-based human pose estimation. *Future Internet*, 14(12), 2022.

- [60] Micheal Clark, Scott Lucett, et al. *NASM essentials of corrective exercise training*. Lippincott Williams & Wilkins, 2010.
- [61] Mark Coeckelbergh. E-care as craftsmanship: virtuous work, skilled engagement, and information technology in health care. *Medicine, health care and philosophy*, 16:807–816, 2013.
- [62] Ana Rita Córias, Min Hun Lee, and Alexandre Bernardino. A low-cost virtual coach for 2d video-based compensation assessment of upper extremity rehabilitation exercises. *Journal of NeuroEngineering and Rehabilitation*, 19(1):1–16, 2022.
- [63] European Commission, Directorate-General for Economic, Financial Affairs, and Economic Policy Committee. *The 2015 ageing report : economic and budgetary projections for the 28 EU Member States (2013-2060)*. Publications Office, 2015.
- [64] Imre Csiszar. A geometric interpretation of darroch and ratcliff’s generalized iterative scaling. *The Annals of Statistics*, pages 1409–1413, 1989.
- [65] Chris Culnane, Benjamin IP Rubinstein, and Vanessa Teague. Health data in an open world. *arXiv preprint arXiv:1712.05627*, 2017.
- [66] Janis J. Daly, Jessica P. McCabe, John Holcomb, Michelle Monkiewicz, Jennifer Gansen, and Svetlana Pundik. Long-dose intensive therapy is necessary for strong, clinically significant, upper limb functional gains and retained gains in severe/moderate chronic stroke. *Neurorehabilitation and Neural Repair*, 33(7):523–537, 2019. PMID: 31131743.
- [67] Thomas Davergne, Antsa Rakotozafiarison, Hervé Servy, and Laure Gossec. Wearable activity trackers in the management of rheumatic diseases: Where are we in 2020? *Sensors*, 20(17), 2020.
- [68] Simon De Lusignan, S-T Liaw, Paul Krause, Vasa Curcin, M Tristan Vicente, Georgios Michalakidis, Lars Agreus, Peter Leysen, Nicola Shaw, and Kumara Mendis. Key concepts to assess the readiness of data for international research: Data quality, lineage and provenance, extraction and processing errors, traceability, and curation. *Yearbook of medical informatics*, 20(01):112–120, 2011.
- [69] Sara García de Villa, David Casillas-Pérez, Ana Jiménez-Martín, and Juan Jesús García-Domínguez. Simultaneous exercise recognition and evaluation in prescribed routines: Approach to virtual coaches. *Expert Systems with Applications*, 199:116990, 2022.

- [70] Bappaditya Debnath, Mary O'Brien, Motonori Yamaguchi, and Ardhendu Behera. A review of computer vision-based approaches for physical rehabilitation and assessment. *Multimedia Systems*, 28, 02 2022.
- [71] Nour El Houda Dehimi and Zakaria Tolba. Attention mechanisms in deep learning: Towards explainable artificial intelligence. In *2024 6th International Conference on Pattern Analysis and Intelligent Systems (PAIS)*, pages 1–7. IEEE, 2024.
- [72] Gisela Miranda Difini, Marcio Garcia Martins, and Jorge Luis Victória Barbosa. Human pose estimation for training assistance: a systematic literature review. In *Proceedings of the Brazilian Symposium on Multimedia and the Web*, pages 189–196, 2021.
- [73] Gulustan Dogan, Sinem Sena Ertas, and İremnaz Cay. Human activity recognition using convolutional neural networks. In *2021 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1–5. IEEE, 2021.
- [74] Elham Dolatabadi, Ying Xuan Zhi, Bing Ye, Marge Coahran, Giorgia Lupinacci, Alex Mihailidis, Rosalie Wang, and Babak Taati. The toronto rehab stroke pose dataset to detect compensation during stroke rehabilitation therapy. In *Proceedings of the 11th EAI international conference on pervasive computing technologies for healthcare*, pages 375–381, 2017.
- [75] Nathalie Doriot and Xuguang Wang. Effects of age and gender on maximum voluntary range of motion of the upper body joints. *Ergonomics*, 49(3):269–281, 2006.
- [76] Zolnamar Dorjsembe, Sodtavilan Odonchimed, and Furen Xiao. Three-dimensional medical image synthesis with denoising diffusion probabilistic models. In *Medical Imaging with Deep Learning*, 2022.
- [77] Juan M Durán. Computer simulations in science and engineering. 2018.
- [78] David Eatock, Caterina Francesca Guidi, Veronika Kunz, Vasileios Margaras, Nicole Scholz, Gaby Umbach, and Ionel Zamfir. Demographic outlook in the european union 2017. 2017.
- [79] Shabnam Sadeghi Esfahlani, Hassan Shirvani, Javaid Butt, Iraj Mirzaee, and Karim Sadeghi Esfahlani. Machine learning role in clinical decision-making: neuro-rehabilitation video game. *Expert Systems with Applications*, 201:117165, 2022.

- [80] Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639):115–118, 2017.
- [81] Alessandro Fabris, Stefano Messina, Gianmaria Silvello, and Gian Antonio Susto. Algorithmic fairness datasets: the story so far. *Data Mining and Knowledge Discovery*, 36(6):2074–2152, 2022.
- [82] Ahmed Fadhil, Yunlong Wang, and Harald Reiterer. Assistive conversational agent for health coaching: A validation study. *Methods of Information in Medicine*, 58, 05 2019.
- [83] Anna Ferrari, Daniela Micucci, Marco Mobilio, and Paolo Napoletano. On the personalization of classification models for human activity recognition. *IEEE Access*, 8:32066–32079, 2020.
- [84] Anna Ferrari, Daniela Micucci, Marco Mobilio, and Paolo Napoletano. Personalization in human activity recognition. *arXiv preprint arXiv:2009.00268*, 2020.
- [85] Bruno Ferreira, Pedro M. Ferreira, Gil Pinheiro, Nelson Figueiredo, Filipe Carvalho, Paulo Menezes, and Jorge Batista. Deep learning approaches for workout repetition counting and validation. *Pattern Recognition Letters*, 151:259–266, 2021.
- [86] Peter Filzmoser and Valentin Todorov. Review of robust multivariate statistical methods in high dimension. *Analytica chimica acta*, 705(1-2):2–14, 2011.
- [87] Peter Filzmoser and Valentin Todorov. Robust tools for the imperfect world. *Information Sciences*, 245:4–20, 2013.
- [88] Michele Fiori, Davide Mor, Gabriele Civitaresse, and Claudio Bettini. Gnn-xar: A graph neural network for explainable activity recognition in smart homes. *arXiv preprint arXiv:2502.17999*, 2025.
- [89] Luciano Floridi and Josh Cowls. A unified framework of five principles for ai in society. *Machine learning and the city: Applications in architecture and urban design*, pages 535–545, 2022.
- [90] Paris: Organisation for Economic Cooperation and Development. ecommendation of the council on artificial intelligence (oecd legal instruments. oecd/ legal/o449), 2019.
- [91] Mélanie Bourassa Forcier, Hortense Gallois, Siobhan Mullan, and Yann Joly. Integrating artificial intelligence into health care through data access: can the gdpr act as a beacon for policymakers? *Journal of Law and the Biosciences*, 6(1):317–335, 2019.

- [92] Maurice Fréchet. Sur la distance de deux lois de probabilité. In *Annales de l'ISUP*, volume 6, pages 183–198, 1957.
- [93] Alyson Gamble. Artificial intelligence and mobile apps for mental healthcare: a social informatics perspective. *Aslib Journal of Information Management*, 72(4):509–523, 2020.
- [94] Yiwen Gao, He Li, and Yan Luo. An empirical study of wearable technology acceptance in healthcare. *Industrial Management & Data Systems*, 115(9):1704–1723, 2015.
- [95] Lynne V Gauthier, Chelsea Kane, Alexandra Borstad, Nancy Strahl, Gitendra Uswatte, Edward Taub, David Morris, Alli Hall, Melissa Arakelian, and Victor Mark. Video game rehabilitation for outpatient stroke (vigorous): protocol for a multi-center comparative effectiveness trial of in-home gamified constraint-induced movement therapy for rehabilitation of chronic upper extremity hemiparesis. *BMC neurology*, 17(1):1–18, 2017.
- [96] Niayesh Gharaei, Waidah Ismail, Crina Grosan, and Rimuljo Hendradi. Optimizing the setting of medical interactive rehabilitation assistant platform to improve the performance of the patients: A case study. *Artificial Intelligence in Medicine*, 120:102151, 2021.
- [97] Davoud Gholamiangonabadi, Nikita Kiselov, and Katarina Grolinger. Deep neural networks for human activity recognition with wearable sensors: Leave-one-subject-out cross-validation for model selection. *Ieee Access*, 8:133982–133994, 2020.
- [98] P Ghosh. Ai early diagnosis could save heart and cancer patients. *Science Correspondent. BBC News*, 2018.
- [99] Christian Giang, Elvira Pirondini, Nawal Kinany, Camilla Pierella, Alessandro Panarese, Martina Coscia, Jenifer Miehlebradt, Cécile Magnin, Pierre Nicolo, Adrian Guggisberg, and Silvestro Micera. Motor improvement estimation and task adaptation for personalized robot-aided therapy: a feasibility study. *BioMedical Engineering OnLine*, 19, 05 2020.
- [100] Judy Wawira Gichoya, Liam G McCoy, Leo Anthony Celi, and Marzyeh Ghassemi. Equity in essence: a call for operationalising fairness in machine learning for healthcare. *BMJ health & care informatics*, 28(1), 2021.
- [101] Oonagh M Giggins, Ulrik McCarthy Persson, and Brian Caulfield. Biofeedback in rehabilitation. *Journal of neuroengineering and rehabilitation*, 10(1):1–11, 2013.

- [102] Francesca Gimigliano and Stefano Negrini. The world health organization” rehabilitation 2030: a call for action”. *European Journal of Physical and Rehabilitation Medicine*, 53(2):155–168, 2017.
- [103] Felipe Giuste, Wenqi Shi, Yuanda Zhu, Tarun Naren, Monica Isgut, Ying Sha, Li Tong, Mitali Gupte, and May D Wang. Explainable artificial intelligence methods in combating pandemics: A systematic review. *IEEE Reviews in Biomedical Engineering*, 2022.
- [104] Caroline Lubick Goldzweig, Greg Orshansky, Neil M Paige, Ali Alexander Towfigh, David A Haggstrom, Isomi Miake-Lye, Jessica M Beroes, and Paul G Shekelle. Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. *Annals of internal medicine*, 159(10):677–687, 2013.
- [105] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Advances in neural information processing systems. *Curran Associates, Inc*, 27:2672–2680, 2014.
- [106] Susan Dorr Goold and Mack Lipkin Jr. The doctor–patient relationship: challenges, opportunities, and strategies. *Journal of general internal medicine*, 14(Suppl 1):S26, 1999.
- [107] Thomas Grote and Philipp Berens. On the ethics of algorithmic decision-making in healthcare. *Journal of medical ethics*, 46(3):205–211, 2020.
- [108] Anjan Gudigar, Udupi Raghavendra, Sneha Nayak, Chui Ping Ooi, Wai Yee Chan, Mokshagna Rohit Gangavarapu, Chinmay Dharmik, Jyothi Samanth, Nahrizul Adib Kadri, Khairunnisa Hasikin, et al. Role of artificial intelligence in covid-19 detection. *Sensors*, 21(23):8045, 2021.
- [109] WHO Guideline. Recommendations on digital interventions for health system strengthening. *World Health Organization*, pages 2020–10, 2019.
- [110] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville. Improved training of wasserstein gans. *Advances in neural information processing systems*, 30, 2017.
- [111] Guan N Guo, Jitendra Jonnagaddala, Sanjay Farshid, Vojtech Huser, Christian Reich, and Siaw-Teng Liaw. Comparison of the cohort selection performance of australian medicines terminology to anatomical therapeutic chemical mappings. *Journal of the American Medical Informatics Association*, 26(11):1237–1246, 2019.

- [112] Ibrahim Habli, Tom Lawton, and Zoe Porter. Artificial intelligence in health care: accountability and safety. *Bulletin of the World Health Organization*, 98(4):251, 2020.
- [113] Massinissa Hamidi and Aomar Osmani. Human activity recognition: A dynamic inductive bias selection perspective. *Sensors*, 21(21):7278, 2021.
- [114] Jiawei Han, Jian Pei, and Micheline Kamber. *Data mining: concepts and techniques*. Elsevier, 2011.
- [115] Kaiming He, X. Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.
- [116] Jeffrey Heaton. Ian goodfellow, yoshua bengio, and aaron courville: Deep learning: The mit press, 2016, 800 pp, isbn: 0262035618. *Genetic Programming and Evolvable Machines*, 19, 10 2017.
- [117] Thomas Hellsten, Jonny Karlsson, Muhammed Shamsuzzaman, and Göran Pulkkis. The potential of computer vision-based marker-less human motion analysis for rehabilitation. *Rehabilitation Process and Outcome*, 10:11795727211022330, 2021. PMID: 34987303.
- [118] Francisco Herrera, F. Charte, Antonio Rivera Rivas, and María José Del Jesus. *Multilabel Classification*. 01 2016.
- [119] Roschelle A Heuberger. The frailty syndrome: a comprehensive review. *Journal of nutrition in gerontology and geriatrics*, 30(4):315–368, 2011.
- [120] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.
- [121] Susan Hillier and Gakeemah Inglis-Jassiem. Rehabilitation for community-dwelling people with stroke: Home or centre based? a systematic review. *International Journal of Stroke*, 5(3):178–186, 2010. PMID: 20536615.
- [122] Geoffrey E Hinton and Sam Roweis. Stochastic neighbor embedding. *Advances in neural information processing systems*, 15, 2002.
- [123] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

- [124] Jeeyoung Hong, Jeongeun Kim, Suk Wha Kim, and Hyoun-Joong Kong. Effects of home-based tele-exercise on sarcopenia among community-dwelling elderly adults: Body composition and functional fitness. *Experimental gerontology*, 87:33–39, 2017.
- [125] A House of Lords Select Committee et al. Ai in the uk: ready, willing and able. *House of Lords*, 36, 2018.
- [126] He Huang, Steven L Wolf, and Jiping He. Recent developments in biofeedback for neuromotor rehabilitation. *Journal of neuroengineering and rehabilitation*, 3(1):1–12, 2006.
- [127] Paul Humphreys. The philosophical novelty of computer simulation methods. *Synthese*, 169:615–626, 2009.
- [128] Hiroshi Inoue. Data augmentation by pairing samples for images classification. *arXiv preprint arXiv:1801.02929*, 2018.
- [129] Brian Kenji Iwana. On mini-batch training with varying length time series. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4483–4487. IEEE, 2022.
- [130] Kirsten Jack, Sionnadh Mairi McLean, Jennifer Klaber Moffett, and Eric Gardiner. Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. *Manual therapy*, 15(3):220–228, 2010.
- [131] Mohd Javaid, Abid Haleem, and Ravi Pratap Singh. Chatgpt for healthcare services: An emerging stage for an innovative perspective. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, 3(1):100105, 2023.
- [132] M Jayamohan and S Yuvaraj. A novel human action recognition using grad-cam visualization with gated recurrent units. *Neural Computing and Applications*, pages 1–16, 2025.
- [133] Diane U Jette, Reg L Warren, and Christopher Wirtalla. The relation between therapy intensity and outcomes of rehabilitation in skilled nursing facilities. *Archives of physical medicine and rehabilitation*, 86(3):373–379, 2005.
- [134] Zhen Jia, Zhang Zhang, Liang Wang, and Tieniu Tan. Human image generation: A comprehensive survey. *arXiv preprint arXiv:2212.08896*, 2022.

- [135] Fei Jiang, Yong Jiang, Hui Zhi, Yi Dong, Hao Li, Sufeng Ma, Yilong Wang, Qiang Dong, Haipeng Shen, and Yongjun Wang. Artificial intelligence in healthcare: past, present and future. *BMJ*, 2:svn-2017, 06 2017.
- [136] Liangxiao Jiang, Lungan Zhang, Liangjun Yu, and Dianhong Wang. Class-specific attribute weighted naive bayes. *Pattern recognition*, 88:321–330, 2019.
- [137] Mike Jones, George Collier, David J. Reinkensmeyer, Frank DeRuyter, John Dzivak, Daniel Zondervan, and John Morris. Big data analytics and sensor-enhanced activity management to improve effectiveness and efficiency of outpatient medical rehabilitation. *International Journal of Environmental Research and Public Health*, 17(3), 2020.
- [138] Rachel E Jordan, Peymane Adab, and K K Cheng. Covid-19: risk factors for severe disease and death. *BMJ*, 368, 2020.
- [139] Seung-kook Jun, Suren Kumar, Xiaobo Zhou, Daniel K Ramsey, and Venkat N Krovi. Automation for individualization of kinect-based quantitative progressive exercise regimen. In *2013 IEEE International Conference on Automation Science and Engineering (CASE)*, pages 243–248. IEEE, 2013.
- [140] Christoph M Kanzler, Mike D Rinderknecht, Anne Schwarz, Ilse Lamers, Cynthia Gagnon, Jeremia P O Held, Peter Feys, Andreas R Luft, Roger Gassert, and Olivier Lambercy. A data-driven framework for selecting and validating digital health metrics: use-case in neurological sensorimotor impairments. *n p j Digital Medicine*, 3:80, 2020.
- [141] Kimmo Karkkainen and Jungseock Joo. Fairface: Face attribute dataset for balanced race, gender, and age for bias measurement and mitigation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 1548–1558, 2021.
- [142] Arnab Karmakar and Deepak Mishra. A robust pose transformational gan for pose guided person image synthesis. In *Computer Vision, Pattern Recognition, Image Processing, and Graphics: 7th National Conference, NCVPRIPG 2019, Hubballi, India, December 22–24, 2019, Revised Selected Papers 7*, pages 89–99. Springer, 2020.
- [143] Justin Keogh, Jason Lake, and Paul Swinton. Practical applications of biomechanical principles in resistance training: Moments and moment arms. *Journal of Fitness Research*, 2:39–48, 01 2013.

- [144] Ali Barzegar Khanghah, Geoff Fernie, and Atena Roshan Fekr. A novel approach to tele-rehabilitation: Implementing a biofeedback system using machine learning algorithms. *Machine Learning with Applications*, 14:100499, 2023.
- [145] Varada Vivek Khanna, Krishnaraj Chadaga, Niranjana Sampathila, Srikanth Prabhu, Rajagopala Chadaga, and Shashikiran Umakanth. Diagnosing covid-19 using artificial intelligence: A comprehensive review. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 11(1):25, 2022.
- [146] Jayoung Kim, Alan S Campbell, and Joseph Wang. Wearable non-invasive epidermal glucose sensors: A review. *Talanta*, 177:163–170, 2018.
- [147] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- [148] Daniel N Kluttz and Deirdre K Mulligan. Automated decision support technologies and the legal profession. *Berkeley Technology Law Journal*, 34(3):853–890, 2019.
- [149] Ravi Komatireddy, Anang Chokshi, Jeanna Basnett, Michael Casale, Daniel Goble, and Tiffany Shubert. Quality and quantity of rehabilitation exercises delivered by a 3-d motion controlled camera: A pilot study. *International journal of physical medicine & rehabilitation*, 2(4), 2014.
- [150] Heli Koskimäki. Avoiding bias in classification accuracy-a case study for activity recognition. In *2015 IEEE symposium series on computational intelligence*, pages 301–306. IEEE, 2015.
- [151] Sotiris Kotsiantis. Decision trees: A recent overview. *Artificial Intelligence Review*, pages 1–23, 04 2013.
- [152] William Kraemer, Kent Adams, E. Cafarelli, Gary Dudley, Cathryn Dooly, M. Feigenbaum, Steven Fleck, Barry Franklin, Andrew Fry, Jay Hoffman, Robert Newton, Jeffrey Potteiger, Michael Stone, Nicholas Ratamess, and Travis Triplett-McBride. Position stand: Progression models in resistance training for healthy adultsamerican college of sports medicinemed sci sports exerc20023436438011828249. *Medicine and science in sports and exercise*, 34:364–80, 02 2002.
- [153] Steven Krahn, Kurt Gerdes, and Herbert Sutter. Department of energy technology readiness assessments-process guide and training plan. Technical report, DEPARTMENT OF ENERGY WASHINGTON DC, 2008.

- [154] Koushik Sivarama Krishnan and Karthik Sivarama Krishnan. Vision transformer based covid-19 detection using chest x-rays. In *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*, pages 644–648. IEEE, 2021.
- [155] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25, pages 200–209. Curran Associates, Inc., 2012.
- [156] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.
- [157] Pranjal Kumar, Siddhartha Chauhan, and Lalit Kumar Awasthi. Human activity recognition (har) using deep learning: Review, methodologies, progress and future research directions. *Archives of Computational Methods in Engineering*, 31(1):179–219, 2024.
- [158] Yogesh Kumar, Apeksha Koul, Ruchi Singla, and Muhammad Fazal Ijaz. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *Journal of ambient intelligence and humanized computing*, pages 1–28, 2022.
- [159] Craig Kuziemy, Anthony Maeder, Oommen John, Shashi Gogia, Arindam Basu, Sushil Meher, and Márcia Ito. Role of artificial intelligence within the telehealth domain: Official 2019 yearbook contribution by the members of imia telehealth working group. *Yearbook of Medical Informatics*, 28, 04 2019.
- [160] Sofoklis Kyriazakos, Hannes Schlieter, Kai Gand, Massimo Caprino, Massimo Corbo, Peppino Tropea, Elda Judica, Irma Sterpi, Stefan Busnatu, Patrick Philipp, et al. A novel virtual coaching system based on personalized clinical pathways for rehabilitation of older adults—requirements and implementation plan of the vcare project. *Frontiers in Digital Health*, 2:546562, 2020.
- [161] Frank La Rue. Report of the special rapporteur on the promotion and protection of the right to freedom of opinion and expression. *Human Rights Council*, 16, 2011.
- [162] Federico Carmelo La Vattiata. Ai-based medical devices: the applicable law in the european union. *BioLaw Journal-Rivista di BioDiritto*, (4):412–437, 2022.

- [163] Kenneth Lai, Svetlana N Yanushkevich, Vlad Shmerko, and Ming Hou. Capturing causality and bias in human action recognition. *Pattern Recognition Letters*, 147:164–171, 2021.
- [164] Olivier Lambercy, Rea Lehner, Karen Chua, Seng Kwee Wee, Deshan Kumar Rajeswaran, Christopher Wee Keong Kuah, Wei Tech Ang, Phyllis Liang, Domenico Campolo, Asif Hussain, et al. Neurorehabilitation from a distance: can intelligent technology support decentralized access to quality therapy? *Frontiers in Robotics and AI*, 8:612415, 2021.
- [165] Peter Langhorne, Julie Bernhardt, and Gert Kwakkel. Stroke rehabilitation. *The Lancet*, 377(9778):1693–1702, 2011.
- [166] Audrey Laporte, Ruth Croxford, and Peter C Coyte. Can a publicly funded home care system successfully allocate service based on perceived need rather than socioeconomic status? a canadian experience. *Health & social care in the community*, 15(2):108–119, 2007.
- [167] Elizabeth J Layman. Ethical issues and the electronic health record. *The health care manager*, 27(2):165–176, 2008.
- [168] Colin Lea, Michael D Flynn, Rene Vidal, Austin Reiter, and Gregory D Hager. Temporal convolutional networks for action segmentation and detection. In *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 156–165, 2017.
- [169] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.
- [170] Min Lee, Daniel Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez i Badia. *Enabling AI and Robotic Coaches for Physical Rehabilitation Therapy: Iterative Design and Evaluation with Therapists and Post-Stroke Survivors*. 06 2021.
- [171] Min Hun Lee, Daniel P Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez i Badia. Learning to assess the quality of stroke rehabilitation exercises. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*, pages 218–228, 2019.
- [172] Sunghoon I. Lee, Catherine P. Adans-Dester, Matteo Grimaldi, Ariel V. Dowling, Peter C. Horak, Randie M. Black-Schaffer, Paolo Bonato, and Joseph T. Gwin. Enabling

- stroke rehabilitation in home and community settings: A wearable sensor-based approach for upper-limb motor training. *IEEE Journal of Translational Engineering in Health and Medicine*, 6:1–11, 2018.
- [173] Grégoire Lefebvre, Samuel Berlemont, Franck Mamalet, and Christophe Garcia. Blstm-rnn based 3d gesture classification. In *ICANN*, pages 381–388, 2013.
- [174] Daniel Leightley, Moi Hoon Yap, Jessica Coulson, Yoann Barnouin, and Jamie S McPhee. Benchmarking human motion analysis using kinect one: An open source dataset. In *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, pages 1–7. IEEE, 2015.
- [175] Wanqing Li, Zhengyou Zhang, and Zicheng Liu. Action recognition based on a bag of 3d points. In *2010 IEEE computer society conference on computer vision and pattern recognition-workshops*, pages 9–14. IEEE, 2010.
- [176] Yalin Liao, Aleksandar Vakanski, and Min Xian. A deep learning framework for assessing physical rehabilitation exercises. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(2):468–477, 2020.
- [177] Yalin Liao, Aleksandar Vakanski, Min Xian, David Paul, and Russell Baker. A review of computational approaches for evaluation of rehabilitation exercises. *Computers in Biology and Medicine*, 119:103687, 2020.
- [178] Siaw-Teng Liaw, Harshana Liyanage, Craig Kuziemsy, Amanda L Terry, Richard Schreiber, Jitendra Jonnagaddala, and Simon de Lusignan. Ethical use of electronic health record data and artificial intelligence: recommendations of the primary care informatics working group of the international medical informatics association. *Yearbook of Medical Informatics*, 29(01):051–057, 2020.
- [179] Siaw-Teng Liaw, Gawaine Powell-Davies, Christopher Pearce, Helena Britt, Lisa McGlynn, and Mark Fort Harris. Optimising the use of observational electronic health record data: Current issues, evolving opportunities, strategies and scope for collaboration. *Australian Family Physician*, 45(3):153–156, 2016.
- [180] Jonathan Lin and Dana Kulic. Segmenting human motion for automated rehabilitation exercise analysis. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2012:2881–4, 08 2012.

- [181] Steven Y Lin, Megan R Mahoney, and Christine A Sinsky. Ten ways artificial intelligence will transform primary care. *Journal of general internal medicine*, 34:1626–1630, 2019.
- [182] Ji Liu, Heshan Liu, Mang-Tik Chiu, Yu-Wing Tai, and Chi-Keung Tang. Pose-guided high-resolution appearance transfer via progressive training. *arXiv preprint arXiv:2008.11898*, 2020.
- [183] Kun Liu, Wu Liu, Chuang Gan, Mingkui Tan, and Huadong Ma. T-c3d: Temporal convolutional 3d network for real-time action recognition. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.
- [184] Meichen Liu, Kejun Wang, Ruihang Ji, Shuzhi Sam Ge, and Jing Chen. Pose transfer generation with semantic parsing attention network for person re-identification. *Knowledge-Based Systems*, 223:107024, 2021.
- [185] Ran Liu, Miye Wang, Tao Zheng, Rui Zhang, Nan Li, Zhongxiu Chen, Hongmei Yan, and Qingke Shi. An artificial intelligence-based risk prediction model of myocardial infarction. *BMC bioinformatics*, 23(1):1–17, 2022.
- [186] Jeffrey W Lockhart and Gary M Weiss. Limitations with activity recognition methodology & data sets. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, pages 747–756, 2014.
- [187] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.
- [188] Deborah Lupton. M-health and health promotion: The digital cyborg and surveillance society. *Social Theory & Health*, 10:229–244, 2012.
- [189] David D Luxton and Laurel D Riek. Artificial intelligence and robotics in rehabilitation. 2019.
- [190] Liqian Ma, Xu Jia, Qianru Sun, Bernt Schiele, Tinne Tuytelaars, and Luc Van Gool. Pose guided person image generation. *Advances in neural information processing systems*, 30, 2017.
- [191] Paweł Maciejasz, Jörg Eschweiler, Kurt Gerlach-Hahn, Arne Jansen-Troy, and Steffen Leonhardt. A survey on robotic devices for upper limb rehabilitation. *Journal of neuroengineering and rehabilitation*, 11(1):1–29, 2014.

- [192] Alexander Macintosh, Nicolas Vignais, Eric Desailly, Elaine Biddiss, and Vincent Vigneron. A classification and calibration procedure for gesture specific home-based therapy exercise in young people with cerebral palsy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:144–155, 2021.
- [193] Farah Magrabi, Elske Ammenwerth, Jytte Brender McNair, Nicolet F De Keizer, Hannele Hyppönen, Pirkko Nykänen, Michael Rigby, Philip J Scott, Tuulikki Vehko, Zoie Shui-Yee Wong, et al. Artificial intelligence in clinical decision support: challenges for evaluating ai and practical implications. *Yearbook of medical informatics*, 28(01):128–134, 2019.
- [194] Sidi Ahmed Mahmoudi, Otmane Amel, Sédrick Stassin, Margot Liagre, Mohamed Benkedadra, and Matei Mancas. A review and comparative study of explainable deep learning models applied on action recognition in real time. *Electronics*, 12(9):2027, 2023.
- [195] Sumit Majumder, Tapas Mondal, and M. Jamal Deen. Wearable sensors for remote health monitoring. *Sensors*, 17(1), 2017.
- [196] Martina Mancini, Mahmoud El-Gohary, Sean Pearson, James McNames, Heather Schlueter, John G Nutt, Laurie A King, and Fay B Horak. Continuous monitoring of turning in parkinson’s disease: rehabilitation potential. *NeuroRehabilitation*, 37(1):3, 2015.
- [197] Naveen Kumar Mangal and Anil Kumar Tiwari. A review of the evolution of scientific literature on technology-assisted approaches using rgb-d sensors for musculoskeletal health monitoring. *Computers in Biology and Medicine*, 132:104316, 2021.
- [198] Gary Marcus. Deep learning: A critical appraisal. *arXiv preprint arXiv:1801.00631*, 2018.
- [199] Nicolas Martelli, Déborah Eskenazy, Carole Déan, Judith Pineau, Patrice Prognon, Gilles Chatellier, Marc Sapoval, and Olivier Pellerin. New european regulation for medical devices: what is changing? *Cardiovascular and interventional radiology*, 42:1272–1278, 2019.
- [200] Ester Martinez-Martin and Miguel Cazorla. A socially assistive robot for elderly exercise promotion. *IEEE Access*, 7:75515–75529, 2019.

- [201] Jayson S Marwaha, Adam B Landman, Gabriel A Brat, Todd Dunn, and William J Gordon. Deploying digital health tools within large, complex health systems: key considerations for adoption and implementation. *NPJ digital medicine*, 5(1):13, 2022.
- [202] Emanuele Marzetti, Riccardo Calvani, Matteo Tosato, Matteo Cesari, Mauro Di Bari, Antonio Cherubini, Marianna Broccatelli, Giulia Saveria, Mariaelena D’Elia, Marco Pahor, et al. Physical activity and exercise as countermeasures to physical frailty and sarcopenia. *Aging clinical and experimental research*, 29:35–42, 2017.
- [203] Nancy E. Mayo. Stroke rehabilitation at home. *Stroke*, 47(6):1685–1691, 2016.
- [204] Nancy E. Mayo, Sharon Wood-Dauphinee, Robert Côté, David Gayton, Joseph Carlton, Joanne Buttery, and Robyn Tamblyn. There’s no place like home. *Stroke*, 31(5):1016–1023, 2000.
- [205] John McCarthy, Marvin L Minsky, Nathaniel Rochester, and Claude E Shannon. A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, 27(4):12–12, 2006.
- [206] Jonathan C Mcleod, Tanner Stokes, and Stuart M Phillips. Resistance exercise training as a primary countermeasure to age-related chronic disease. *Frontiers in Physiology*, 10:645, 2019.
- [207] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning. *ACM computing surveys (CSUR)*, 54(6):1–35, 2021.
- [208] Ciro Mennella, Susanna Alloisio, Antonio Novellino, and Federica Viti. Characteristics and applications of technology-aided hand functional assessment: A systematic review. *Sensors*, 22(1), 2022.
- [209] Ciro Mennella, Massimo Esposito, Giuseppe De Pietro, and Umberto Maniscalco. Promoting fairness in activity recognition algorithms for patient’s monitoring and evaluation systems in healthcare. *Computers in Biology and Medicine*, 179:108826, 2024.
- [210] Ciro Mennella, Umberto Maniscalco, Giuseppe De Pietro, and Massimo Esposito. A deep learning system to monitor and assess rehabilitation exercises in home-based remote and unsupervised conditions. *Computers in Biology and Medicine*, page 107485, 2023.

- [211] Ciro Mennella, Umberto Maniscalco, Giuseppe De Pietro, and Massimo Esposito. Generating a novel synthetic dataset for rehabilitation exercises using pose-guided conditioned diffusion models: A quantitative and qualitative evaluation. *Computers in Biology and Medicine*, page 107665, 2023.
- [212] Ciro Mennella, Umberto Maniscalco, Giuseppe De Pietro, and Massimo Esposito. The role of artificial intelligence in future rehabilitation services: a systematic literature review. *IEEE Access*, 2023.
- [213] Ciro Mennella, Umberto Maniscalco, Giuseppe De Pietro, and Massimo Esposito. Ethical and regulatory challenges of ai technologies in healthcare: A narrative review. *Heliyon*, 2024.
- [214] Ciro Mennella, Umberto Maniscalco, Giuseppe De Pietro, and Massimo Esposito. Modeling rehabilitation dataset to implement effective ai assistive systems. *Discover Artificial Intelligence*, 4(1):43, 2024.
- [215] Antonio Migliore. On the new regulation of medical devices in europe. *Expert review of medical devices*, 14(12):921–923, 2017.
- [216] Dubravko Miljković. Brief review of self-organizing maps. In *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 1061–1066, 2017.
- [217] Kristine K Miller, Rebecca E Porter, Erin DeBaun-Sprague, Marieke Van Puymbroeck, and Arlene A Schmid. Exercise after stroke: patient adherence and beliefs after discharge from rehabilitation. *Topics in stroke rehabilitation*, 24(2):142–148, 2017.
- [218] Alina Miron, Nouredin Sadawi, Waidah Ismail, Hafez Hussain, and Crina Grosan. Intellirehabds (irds)—a dataset of physical rehabilitation movements. *Data*, 6(5):46, 2021.
- [219] Martin Missmann, Katrin Gollner, Andrea Schroll, Michael Pirchl, Vincent Grote, and Michael J Fischer. Impact of different isokinetic movement patterns on shoulder rehabilitation outcome. *International Journal of Environmental Research and Public Health*, 19(17):10623, 2022.
- [220] Thomas Baltzer Moeslund and Erik Granum. A survey of computer vision-based human motion capture. *Comput. Vis. Image Underst.*, 81:231–268, 2001.

- [221] Puria Azadi Moghadam, Sanne Van Dalen, Karina C Martin, Jochen Lennerz, Stephen Yip, Hossein Farahani, and Ali Bashashati. A morphology focused diffusion probabilistic model for synthesis of histopathology images. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2000–2009, 2023.
- [222] Matteo Morando, Serena Ponte, Elisa Ferrara, and Silvana Dellepiane. Definition of motion and biophysical indicators for home-based rehabilitation through serious games. *Information*, 9(5), 2018.
- [223] Heimo Müller and Andreas Holzinger. Kandinsky patterns. *Artificial intelligence*, 300:103546, 2021.
- [224] Tom Nadarzynski, Oliver Miles, Aimee Cowie, and Damien Ridge. Acceptability of artificial intelligence (ai)-led chatbot services in healthcare: A mixed-methods study. *Digital health*, 5:2055207619871808, 2019.
- [225] Nilah Ravi Nair, Lena Schmid, Fernando Moya Rueda, Markus Pauly, Gernot A Fink, and Christopher Reining. Dataset bias in human activity recognition. *arXiv preprint arXiv:2301.10161*, 2023.
- [226] V. Nangalia, David Prytherch, and G. Smith. Health technology assessment review: remote monitoring of vital signs - current status and future challenges. *Critical Care*, 14(5):233, September 2010. Projects: CHMI. Health informatics. Clinical outcome modelling.
- [227] Aravind Natarajan, Hao-Wei Su, and Conor Heneghan. Assessment of physiological signs associated with covid-19 measured using wearable devices. *NPJ digital medicine*, 3(1):156, 2020.
- [228] Miriam E Nelson, W Jack Rejeski, Steven N Blair, Pamela W Duncan, James O Judge, Abby C King, Carol A Macera, and Carmen Castaneda-Sceppa. Physical activity and public health in older adults: recommendation from the american college of sports medicine and the american heart association. *Circulation*, 116(9):1094, 2007.
- [229] Natalia Neverova, Riza Alp Guler, and Iasonas Kokkinos. Dense pose transfer. In *Proceedings of the European conference on computer vision (ECCV)*, pages 123–138, 2018.
- [230] Andrew Y. Ng and Michael I. Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. In *Advances in Neural Information*

- Processing Systems 14 (NIPS 2001)*, NIPS'01, page 841–848, Cambridge, MA, USA, 2001. MIT Press.
- [231] Sarah Nicholson, Falko F. Sniehotta, Frederike van Wijck, Carolyn A. Greig, Marie Johnston, Marion E. T. McMurdo, Martin Dennis, and Gillian E Mead. A systematic review of perceived barriers and motivators to physical activity after stroke. *International Journal of Stroke*, 8(5):357–364, 2013. PMID: 22974010.
- [232] Augustus Odena, Christopher Olah, and Jonathon Shlens. Conditional image synthesis with auxiliary classifier gans. In *International conference on machine learning*, pages 2642–2651. PMLR, 2017.
- [233] Council of Europe. Responsibility and ai. <https://rm.coe.int/responsability-and-ai-en/168097d9c5>, 2019.
- [234] Erienne V Olesh, Sergiy Yakovenko, and Valeriya Gritsenko. Automated assessment of upper extremity movement impairment due to stroke. *PloS one*, 9(8):e104487, 2014.
- [235] The Swedish National Council on Medical Ethics. In brief – artificial intelligence in healthcare. <https://smer.se/wp-content/uploads/2020/06/smer-2020-2-in-brief-artificial-intelligence-in-healthcare.pdf>, 2020.
- [236] Luca Oneto and Silvia Chiappa. *Fairness in Machine Learning*, page 155–196. Springer International Publishing, 2020.
- [237] Ohad Oren, Bernard J Gersh, and Deepak L Bhatt. Artificial intelligence in medical imaging: switching from radiographic pathological data to clinically meaningful endpoints. *The Lancet Digital Health*, 2(9):e486–e488, 2020.
- [238] World Health Organization. International classification of functioning, disability and health : Icf, 2001.
- [239] World Health Organization. *Ageism in artificial intelligence for health: WHO policy brief*. World Health Organization, 2022.
- [240] World Health Organization et al. World report on ageing and health. 2015.
- [241] World Health Organization et al. Framework on integrated, people-centred health services. *Geneva: World Health Organization*, 2019, 2016.

- [242] World Health Organization et al. *Integrated care for older people: guidelines on community-level interventions to manage declines in intrinsic capacity*. World Health Organization, 2017.
- [243] World Health Organization et al. Ethics and governance of artificial intelligence for health: Who guidance. 2021.
- [244] Kai Packhäuser, Lukas Folle, Florian Thamm, and Andreas Maier. Generation of anonymous chest radiographs using latent diffusion models for training thoracic abnormality classification systems. In *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, pages 1–5. IEEE, 2023.
- [245] Adeline Paiement, Lili Tao, Sion Hannuna, Massimo Camplani, Dima Damen, and Majid Mirmehdi. Online quality assessment of human movement from skeleton data. In *British Machine Vision Conference*, pages 153–166. BMVA press, 2014.
- [246] Evan V Papa, Xiaoyang Dong, and Mahdi Hassan. Resistance training for activity limitations in older adults with skeletal muscle function deficits: a systematic review. *Clinical interventions in aging*, pages 955–961, 2017.
- [247] Aakash Parmar, Rakesh Katariya, and Vatsal Patel. *A Review on Random Forest: An Ensemble Classifier*, pages 758–763. 01 2019.
- [248] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Edouard Duchesnay, and Gilles Louppe. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12, 01 2012.
- [249] Ana Pereira, Duarte Folgado, Ricardo Cotrim, and Inês Sousa. Physiotherapy exercises evaluation using a combined approach based on semg and wearable inertial sensors. In *Biosignals*, pages 73–82, 2019.
- [250] Luis Perez and Jason Wang. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*, 2017.
- [251] Dana Pessach and Erez Shmueli. Algorithmic fairness. In *Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook*, pages 867–886. Springer, 2023.

- [252] Walter HL Pinaya, Petru-Daniel Tudosiu, Jessica Dafflon, Pedro F Da Costa, Virginia Fernandez, Parashkev Nachev, Sebastien Ourselin, and M Jorge Cardoso. Brain imaging generation with latent diffusion models. In *MICCAI Workshop on Deep Generative Models*, pages 117–126. Springer, 2022.
- [253] Michael O. Powell, Aviv Elor, Ash Robbins, Sri Kurniawan, and Mircea Teodorescu. Predictive shoulder kinematics of rehabilitation exercises through immersive virtual reality. *IEEE Access*, 10:25621–25632, 2022.
- [254] Ghanashyama Prabhu, Noel E. O’Connor, and Kieran Moran. Recognition and repetition counting for local muscular endurance exercises in exercise-based rehabilitation: A comparative study using artificial intelligence models. *Sensors*, 20(17), 2020.
- [255] Fabian Prasser, Helmut Spengler, Raffael Bild, Johanna Eicher, and Klaus A Kuhn. Privacy-enhancing etl-processes for biomedical data. *International journal of medical informatics*, 126:72–81, 2019.
- [256] S Phani Praveen, Parvathaneni Naga Srinivasu, Jana Shafi, Marcin Wozniak, and Muhammad Fazal Ijaz. Resnet-32 and fastai for diagnoses of ductal carcinoma from 2d tissue slides. *Scientific Reports*, 12(1):20804, 2022.
- [257] Martine T E Puts, Samar Toubasi, Melissa K Andrew, Maureen C Ashe, Jenny Ploeg, Esther Atkinson, Ana Patricia Ayala, Angelique Roy, Miriam Rodríguez Monforte, Howard Bergman, et al. Interventions to prevent or reduce the level of frailty in community-dwelling older adults: a scoping review of the literature and international policies. *Age and ageing*, 46(3):383–392, 2017.
- [258] Xuebin Qin, Zichen Zhang, Chenyang Huang, Masood Dehghan, Osmar R Zaiane, and Martin Jagersand. U2-net: Going deeper with nested u-structure for salient object detection. *Pattern recognition*, 106:107404, 2020.
- [259] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PMLR, 2021.
- [260] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*, 2017.

- [261] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. Zero-shot text-to-image generation. In *International Conference on Machine Learning*, pages 8821–8831. PMLR, 2021.
- [262] Sandeep Reddy, Sonia Allan, Simon Coghlan, and Paul Cooper. A governance model for the application of ai in health care. *Journal of the American Medical Informatics Association*, 27(3):491–497, 2020.
- [263] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.
- [264] Qaiser Riaz, Anna Vögele, Björn Krüger, and Andreas Weber. One small step for a man: Estimation of gender, age and height from recordings of one step by a single inertial sensor. *Sensors*, 15(12):31999–32019, 2015.
- [265] Marco Tulio Ribeiro, Sameer Singh, and Carlos 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*, pages 1135–1144, 2016.
- [266] Thanassis Rikakis, Aisling Kelliher, Jinwoo Choi, Jia-Bin Huang, Kris Kitani, Setor Zilevu, and Steven L Wolf. Semi-automated home-based therapy for the upper extremity of stroke survivors. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*, pages 249–256, 2018.
- [267] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10684–10695, 2022.
- [268] Joachim Roski, BA Hamilton, W Chapman, J Heffner, R Trivedi, G Del Fiol, R Kukafka, P Bleicher, H Estiri, J Klann, et al. How artificial intelligence is changing health and health care, 2019.
- [269] J Ross, C Webb, and F Rahman. Artificial intelligence in healthcare. *London: Academy of Medical Royal Colleges*, 2019.
- [270] Stuart J Russell. *Artificial intelligence a modern approach*. Pearson Education, Inc., 2010.

- [271] Yves Rybarczyk, Clément Cointe, Tiago Gonçalves, Vitor Minhoto, Jan Kleine Deters, Santiago Villarreal, Arián Aladro Gonzalo, Jonathan Baldeón, and Danilo Esparza. On the use of natural user interfaces in physical rehabilitation: a web-based application for patients with hip prosthesis. *Journal of Science and Technology of the Arts*, 10(2):15–24, May 2018.
- [272] Nouredin Sadawi, Alina Miron, Waidah Ismail, Hafez Hussain, and Crina Grosan. Gesture correctness estimation with deep neural networks and rough path descriptors. In *2019 International Conference on Data Mining Workshops (ICDMW)*, pages 595–602. IEEE, 2019.
- [273] Avinab Saha, Shashank Gupta, Sravan Kumar Ankireddy, Karl Chahine, and Joydeep Ghosh. Exploring explainability in video action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8176–8181, 2024.
- [274] Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily L Denton, Kamyar Ghasemipour, Raphael Gontijo Lopes, Burcu Karagol Ayan, Tim Salimans, et al. Photorealistic text-to-image diffusion models with deep language understanding. *Advances in Neural Information Processing Systems*, 35:36479–36494, 2022.
- [275] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. *Advances in neural information processing systems*, 29, 2016.
- [276] Arthur L Samuel. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 44(1.2):206–226, 2000.
- [277] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018.
- [278] Hussein Sarwat, Hassan Sarwat, Shady A. Maged, Tamer H. Emara, Ahmed M. El-bokl, and Mohammed Ibrahim Awad. Design of a data glove for assessment of hand performance using supervised machine learning. *Sensors*, 21(21), 2021.
- [279] Marietje Schaake. The european commission’s artificial intelligence act. *Stanford Univeristy Human-Centered Artificial Intelligence (BHAI), Stanford, Canada*, pages 2021–06, 2021.

- [280] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural networks*, 61:85–117, 2015.
- [281] Silvana Secinaro, Davide Calandra, Aurelio Secinaro, Vivek Muthurangu, and Paolo Biancone. The role of artificial intelligence in healthcare: a structured literature review. *BMC medical informatics and decision making*, 21:1–23, 2021.
- [282] Viktor Seib, Benjamin Lange, and Stefan Wirtz. Mixing real and synthetic data to enhance neural network training—a review of current approaches. *arXiv preprint arXiv:2007.08781*, 2020.
- [283] London: National Health Service. The topol review: Preparing the healthcare workforce to deliver the digital future. <https://topol.hee.nhs.uk/>, 2019.
- [284] Tong Sha, Wei Zhang, Tong Shen, Zhoujun Li, and Tao Mei. Deep person generation: A survey from the perspective of face, pose, and cloth synthesis. *ACM Computing Surveys*, 55(12):1–37, 2023.
- [285] Elham Shabaninia, Hossein Nezamabadi-pour, and Fatemeh Shafizadegan. Transformers in action recognition: A review on temporal modeling. *arXiv preprint arXiv:2302.01921*, 2022.
- [286] Thanveer Shaik, Xiaohui Tao, Niall Higgins, Lin Li, Raj Gururajan, Xujuan Zhou, and U Rajendra Acharya. Remote patient monitoring using artificial intelligence: Current state, applications, and challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 13(2):e1485, 2023.
- [287] Ashish Sharma, Inna W Lin, Adam S Miner, David C Atkins, and Tim Althoff. Human–ai collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nature Machine Intelligence*, 5(1):46–57, 2023.
- [288] Catherine Sherrington, Anne Tiedemann, Nicola Fairhall, Jacqueline CT Close, and Stephen R Lord. Exercise to prevent falls in older adults: an updated meta-analysis and best practice recommendations. *New South Wales public health bulletin*, 22(4):78–83, 2011.
- [289] Xingjian Shi, Zhoung Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 28, 2015.

- [290] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019.
- [291] Daniel Siewiorek, Asim Smailagic, and Anind Dey. Architecture and applications of virtual coaches. *Proceedings of the IEEE*, 100(8):2472–2488, 2012.
- [292] Patrice Y Simard, David Steinkraus, John C Platt, et al. Best practices for convolutional neural networks applied to visual document analysis. In *Icdar*, volume 3. Edinburgh, 2003.
- [293] Andrew K Simpson, Debduot Biswas, John W Emerson, Brandon D Lawrence, and Jonathan N Grauer. Quantifying the effects of age, gender, degeneration, and adjacent level degeneration on cervical spine range of motion using multivariate analyses. *Spine*, 33(2):183–186, 2008.
- [294] Christine A Sinsky, Rachel Willard-Grace, Andrew M Schutzbank, Thomas A Sinsky, David Margolius, and Thomas Bodenheimer. In search of joy in practice: a report of 23 high-functioning primary care practices. *The Annals of Family Medicine*, 11(3):272–278, 2013.
- [295] Diego Soliño-Fernandez, Alexander Ding, Esteban Bayro-Kaiser, and Eric L Ding. Willingness to adopt wearable devices with behavioral and economic incentives by health insurance wellness programs: results of a us cross-sectional survey with multiple consumer health vignettes. *BMC Public Health*, 19:1–8, 2019.
- [296] Andreas Stöckl. Evaluating a synthetic image dataset generated with stable diffusion. *arXiv preprint arXiv:2211.01777*, 2022.
- [297] Gábor J. Székely, Maria L. Rizzo, and Nail K. Bakirov. Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, 35(6):2769 – 2794, 2007.
- [298] Konstantin Tachkov, Antal Zemplenyi, Maria Kamusheva, Maria Dimitrova, Pekka Siirtola, Johan Pontén, Bertalan Nemeth, Zoltan Kalo, and Guenka Petrova. Barriers to use artificial intelligence methodologies in health technology assessment in central and east european countries. *Frontiers in Public Health*, 10:921226, 2022.
- [299] Karolina Talar, Alejandro Hernández-Belmonte, Tomas Vetrovsky, Michal Steffl, Ewa Kałamacka, and Javier Courel-Ibáñez. Benefits of resistance training in early and late stages of frailty and sarcopenia: a systematic review and meta-analysis of randomized controlled studies. *Journal of clinical medicine*, 10(8):1630, 2021.

-
- [300] Kashvi Taunk, Sanjukta De, Srishti Verma, and Aleena Swetapadma. A brief review of nearest neighbor algorithm for learning and classification. In *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, pages 1255–1260, 2019.
- [301] Portia E Taylor, Gustavo JM Almeida, Jessica K Hodgins, and Takeo Kanade. Multi-label classification for the analysis of human motion quality. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2214–2218. IEEE, 2012.
- [302] Lucas Theis, Aäron van den Oord, and Matthias Bethge. A note on the evaluation of generative models. *arXiv preprint arXiv:1511.01844*, 2015.
- [303] D. Theodoros and Trevor Russell. Telerehabilitation: current perspectives. *Studies in health technology and informatics*, 131:191–209, 02 2008.
- [304] Maya Thomas, Sanjana Murali, Benjamin Scott S Simpson, Alex Freeman, Alex Kirkham, Daniel Kelly, Hayley C Whitaker, Yi Zhao, Mark Emberton, and Joseph M Norris. Use of artificial intelligence in the detection of primary prostate cancer in multiparametric mri with its clinical outcomes: a protocol for a systematic review and meta-analysis. *BMJ open*, 13(8):e074009, 2023.
- [305] Eric J Topol. High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine*, 25(1):44–56, 2019.
- [306] Brandon Trabucco, Kyle Doherty, Max Gurinas, and Ruslan Salakhutdinov. Effective data augmentation with diffusion models. *arXiv preprint arXiv:2302.07944*, 2023.
- [307] Nola D Tracy, John C Young, and Robert L Mason. Multivariate control charts for individual observations. *Journal of quality technology*, 24(2):88–95, 1992.
- [308] Viet-Thi Tran, Carolina Riveros, and Philippe Ravaud. Patients’ views of wearable devices and ai in healthcare: findings from the compare e-cohort. *npj digital medicine* 2 (53), 2019.
- [309] Peppino Tropea, Hannes Schlieter, Irma Sterpi, Elda Judica, Kai Gand, Massimo Caprino, Iñigo Gabilondo, Juan Gómez-Esteban, Stefan Busnatu, Crina Sinescu, Sofoklis Kyriazakos, Sadia Anwar, and Massimo Corbo. Scoping review on virtual coaching in medical care: Rehabilitation, the great absentee (preprint). *Journal of Medical Internet Research*, 21, 11 2018.

- [310] Lorainne Tudor Car, Dhakshenya Ardhithy Dhinakaran, Bhone Myint Kyaw, Tobias Kowatsch, Shafiq Joty, Yin-Leng Theng, and Rifat Atun. Conversational agents in health care: Scoping review and conceptual analysis. *J Med Internet Res*, 22(8):e17158, Aug 2020.
- [311] Andrea Turolla, Giacomo Rossettini, Antonello Viceconti, Alvisa Palese, and Tommaso Geri. Musculoskeletal physical therapy during the covid-19 pandemic: is telerehabilitation the answer? *Physical therapy*, 100(8):1260–1264, 2020.
- [312] Aleksandar Vakanski, Jake Ferguson, and Stephen Lee. Mathematical modeling and evaluation of human motions in physical therapy using mixture density neural networks. *Journal of Physiotherapy & Physical Rehabilitation*, 1, 10 2016.
- [313] Aleksandar Vakanski, Hyung-pil Jun, David Paul, and Russell Baker. A data set of human body movements for physical rehabilitation exercises. *Data*, 3(1):2, 2018.
- [314] Benjamin Van Giffen, Dennis Herhausen, and Tobias Fahse. Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods. *Journal of Business Research*, 144:93–106, 2022.
- [315] R PS Van Peppen, Gert Kwakkel, Sharon Wood-Dauphinee, H JM Hendriks, Ph J Van der Wees, and Joost Dekker. The impact of physical therapy on functional outcomes after stroke: what’s the evidence? *Clinical rehabilitation*, 18(8):833–862, 2004.
- [316] Nelleke C van Wouwe, Diksha Mohanty, Anushree Lingaiah, Scott A Wylie, and Kathrin LaFaver. Impaired action control in patients with functional movement disorders. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 32(1):73–78, 2020.
- [317] B Vellas and S Sourdet. Prevention of frailty in aging. *J Frailty Aging*, 6(4):174–7, 2017.
- [318] Patrick von Platen, Suraj Patil, Anton Lozhkov, Pedro Cuenca, Nathan Lambert, Kashif Rasul, Mishig Davaadorj, and Thomas Wolf. Diffusers: State-of-the-art diffusion models, 2022.
- [319] Elise Vonstad, Kerstin Bach, Beatrix Vereijken, Xiaomeng Su, and Jan Nilsen. Performance of machine learning models in estimation of ground reaction forces during balance exergaming. *Journal of NeuroEngineering and Rehabilitation*, 19, 02 2022.
- [320] Elise Klæbo Vonstad, Xiaomeng Su, Beatrix Vereijken, Kerstin Bach, and Jan Harald Nilsen. Comparison of a deep learning-based pose estimation system to marker-based and kinect systems in exergaming for balance training. *Sensors*, 20(23), 2020.

-
- [321] Ronny Votel and Na Li. Next-generation pose detection with movenet and tensorflow.js. <http://blog.tensorflow.org/2021/05/next-generation-pose-detection-with-movenet-and-tensorflowjs.html>, 2017. 2021.05.17.
- [322] Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H Beck. Deep learning for identifying metastatic breast cancer. *arXiv preprint arXiv:1606.05718*, 2016.
- [323] Zhengwei Wang, Qi She, and Tomas E Ward. Generative adversarial networks in computer vision: A survey and taxonomy. *ACM Computing Surveys (CSUR)*, 54(2):1–38, 2021.
- [324] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.
- [325] Nick S Ward, Fran Brander, and Kate Kelly. Intensive upper limb neurorehabilitation in chronic stroke: outcomes from the queen square programme. *Journal of Neurology, Neurosurgery & Psychiatry*, 90(5):498–506, 2019.
- [326] Wenchuan Wei, Carter McElroy, and Sujit Dey. Towards on-demand virtual physical therapist: Machine learning-based patient action understanding, assessment and task recommendation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(9):1824–1835, 2019.
- [327] Wenchuan Wei, Carter Mcelroy, and Sujit Dey. Using sensors and deep learning to enable on-demand balance evaluation for effective physical therapy. *IEEE Access*, 8:99889–99899, 2020.
- [328] Switzerland) WHO Group Consultation on Health Telematics (1997: Geneva. A health telematics policy in support of who’s health-for-all strategy for global health development : report of the who group consultation on health telematics, 11-16 december, geneva, 1997, 1998.
- [329] Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, et al. The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9, 2016.

- [330] Chengming Xu, Yanwei Fu, Chao Wen, Ye Pan, Yu-Gang Jiang, and Xiangyang Xue. Pose-guided person image synthesis in the non-iconic views. *IEEE Transactions on Image Processing*, 29:9060–9072, 2020.
- [331] Rikiya Yamashita, Mizuho Nishio, Richard K. G. Do, and Kaori Togashi. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, 9:611–629, 2018.
- [332] Lingbo Yang, Pan Wang, Chang Liu, Zhanning Gao, Peiran Ren, Xinfeng Zhang, Shanshe Wang, Siwei Ma, Xiansheng Hua, and Wen Gao. Towards fine-grained human pose transfer with detail replenishing network. *IEEE Transactions on Image Processing*, 30:2422–2435, 2021.
- [333] Karen Yeung. A study of the implications of advanced digital technologies (including ai systems) for the concept of responsibility within a human rights framework. *MSI-AUT (2018)*, 5, 2018.
- [334] Su-Zi Yoo, Mi-Hyun No, Jun-Won Heo, Dong-Ho Park, Ju-Hee Kang, So Hun Kim, and Hyo-Bum Kwak. Role of exercise in age-related sarcopenia. *Journal of exercise rehabilitation*, 14(4):551, 2018.
- [335] Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015.
- [336] Lvmin Zhang and Maneesh Agrawala. Adding conditional control to text-to-image diffusion models, 2023.
- [337] Tianxiang Zhang, Michelle Karg, Jonathan Feng-Shun Lin, Dana Kulic, and Gentiane Venture. Imu based single stride identification of humans. In *2013 IEEE RO-MAN*, pages 220–225. IEEE, 2013.
- [338] Ying Xuan Zhi, Michelle Lukasik, Michael H Li, Elham Dolatabadi, Rosalie H Wang, and Babak Taati. Automatic detection of compensation during robotic stroke rehabilitation therapy. *IEEE journal of translational engineering in health and medicine*, 6:1–7, 2017.