PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
UNIVERSITY OF TLEMCEN
FACULTY OF TECHNOLOGY

Electrical Engineering And Electronincs Department
Speciality: Systems Engineering
Major: Industrial Engineering

Master's Thesis

Artificial Intelligence designed for analysing human activity in a work environment

---

Author:

Mr. Mallek Abdelmalek

Presented before the jury on June 26th, 2023

PRESIDENT :	**M. H. Kahouadji**	MCB, University of Tlemcen
REVIEWER :	**M. M. Souier**	Pr, University of Tlemcen
EXPERT I2E :	**Mme. K A. Benachenhou**	MCB, University of Tlemcen
SOCIO-ECONOMIC PARTNER :	**M G C O. Merad**	Manager at SAFFEC sarl, Tlemcen
SUPERVISOR :	**Ms. W. Handouzi**	MCB, University of Tlemcen

2022/2023

Declaration of originality

In accordance with academic protocol, I assert that my thesis has not breached any copyright laws or proprietary rights, and that all incorporated concepts, methodologies, citations, and other works from external sources, whether published or not, have been properly cited and referenced, adhering to established conventions.

Moreover, I confirm that this document is an authentic replica of my thesis, encompassing any last modifications, authorized by both my thesis committee and the graduation service, and has not been presented as part of any other academic program at any other institution.

Acknowledgment

In the name of Allah, the Most Gracious, the Most Merciful, I begin this acknowledgment with a heartfelt expression of gratitude, thanks to Allah for blessing me with the opportunity and the strength to complete this master's degree. Without His blessings and mercy, this achievement would not have been possible.

I would also like to express my deep appreciation to my parents Mohammed and Nadera for their unwavering support and belief in me. Their patience and encouragement kept me motivated throughout this journey despite all the challenges and hard times. My brother Reda also deserves a special mention for being my source of inspiration and constant support.

I am extremely grateful to my supervisor, Wahida Handouzi, for her invaluable guidance and support. She not only provided me with academic knowledge but also treated me like a family member, offering emotional support and encouragement when I needed it the most. Her guidance was fundamental in shaping my research and bringing it to fruition.

I would like to express my deepest gratitude to my best binome, Dassine, for her unwavering support, dedication, and everything she gave me. Her help and advice were instrumental in completing this thesis. Without her support, I would not have been able to achieve this milestone.

Finally, Special thanks go to Dr. Zemirli for his invaluable orientations and project validation, which were critical in shaping my research and ensuring its quality.Also for Youcef Bekkai for his valuable assistance during my internship. His guidance provided me with invaluable practical experience, and his help in gathering the necessary resources for the final system testing was indispensable.

Dedication

I dedicate this thesis to the most beautiful soul that ever graced this earth - my beloved sister Amina may Allah have mercy on her with his vast mercy. Her presence in my life was a gift from Allah (SWT), and her untimely departure was a tragedy that I am still trying to come to terms with.

My sister was my guiding light, my biggest supporter, and my constant source of inspiration. Her unwavering belief in me and my abilities helped me overcome countless obstacles and achieve academic success beyond my wildest dreams. As I sit here, typing these words, I am filled with a deep sense of longing and nostalgia. I miss her infectious smile, her infectious laughter, and her infectious energy.

To my dear sister, I say: thank you for being the best sister a person could ever ask for. Your memory will continue to inspire me for the rest of my days, and I promise to make the most of the time I have left on this earth by working tirelessly to achieve my goals and make you proud.

May Allah grant you the highest place in Jannat Al Furdous and shower you with his infinite mercy and blessings.

Abstract

The rise of work-related musculoskeletal disorders (WMSDs) has become a major concern in various industries, leading to serious health problems and economic losses. Despite the automation of some manufacturing processes, manual tasks are still necessary and can pose ergonomic risks to workers. To address this issue, an AI-powered tool for ergonomic risk assessment has been developed. The tool successfully estimates the 3D human pose with a mean per joint position error (MPJPE) of 46.8 mm, using the Human3.6M dataset, and calculates the Rapid Entire Body Assessment (REBA) score in real time, providing a comprehensive assessment of ergonomic risk factors. Our approach has been validated by a specialist doctor in rehabilitation. The system employs a semi-supervised learning approach with a fully convolutional model based on dilated temporal convolution over 2D keypoints. The developed AI-powered tool provides immediate feedback, enabling enhanced actions for risk reduction. Case studies demonstrate the effectiveness of the approach for improving the accuracy and efficiency of ergonomic risk assessment in various industries.

Keywords:

REBA, WMSDs, Deep learning, temporal convolution, pose estimation.

Résumé

L'essor des troubles musculosquelettiques liés au travail (TMS) est devenu une préoccupation majeure au sein de diverses industries, entraînant des problèmes de santé graves et des pertes économiques. Malgré l'automatisation de certains processus de fabrication, les tâches manuelles demeurent nécessaires et peuvent présenter des risques ergonomiques pour les travailleurs. Afin de remédier à cette situation, un outil alimenté par l'intelligence artificielle (IA) destiné à l'évaluation des risques ergonomiques a été développé. Cet outil permet d'estimer avec succès la position tridimensionnelle du corps humain avec une erreur moyenne de position par articulation (MPJPE) de 46,8 mm, en utilisant l'ensemble de données Human3.6M, et de calculer en temps réel le score de l'Évaluation Rapide de Tout le Corps (REBA), offrant ainsi une évaluation globale des facteurs de risque ergonomique. Notre approche a reçu la validation d'un médecin spécialiste en rééducation. Le système adopte une approche d'apprentissage semi-supervisé avec un modèle entièrement convolutif basé sur une convolution temporelle dilatée appliquée aux points clés en deux dimensions. L'outil alimenté par l'IA développé fournit une rétroaction immédiate, permettant ainsi de renforcer les mesures visant à réduire les risques. Des études de cas démontrent l'efficacité de cette approche pour améliorer la précision et l'efficacité de l'évaluation des risques ergonomiques dans diverses industries.

Mots-clés :

REBA, TMS liés au travail, apprentissage profond, convolution temporelle, estimation de pose.

ملخص

تزايد الإصابات باضطرابات الجهاز العضلي الهيكلي المتعلقة بالعمل أصبح مصدر قلق رئيسي في مختلف الصناعات، حيث يؤدي إلى مشاكل صحية خطيرة وخسائر اقتصادية معتبرة. و على الرغم من أتمتة بعض عمليات التصنيع، الا أن المهام اليدوية لا تزال ضرورية والتي قد تشكل تهديدات متعلقة ببيئة العمل على العمال. لمعالجة هذا المشكل، قمنا بتطوير أداة تعمل بالذكاء الاصطناعي لتقييمها. هذه الأخيرة تقوم بتقدير وضعية الجسم البشري ثلاثية الأبعاد بنجاح بمعدل خطأ متوسط لكل مفصل يقدر بـ46.8 مم، باستعمال قاعدة البيانات Human 3.6M، و يحسب النظام درجة تقييم الجسم الكامل السريع لعوامل مخاطر الهندسة الإنسانية. تم التحقق من نهجنا من قبل طبيب متخصص في التأهيل. يستخدم النظام نهج التعلم شبه المشرف بنموذج مدمج بالكامل يستند الى التحويل الزمني الموسع للنقاط الرئيسية ثنائية الأبعاد، توفر الأداة المدعومة بالذكاء الاصطناعي إشعارًا فوريًا مما يمكن من اتخاذ إجراءات أفضل لتقليل المخاطر. وتوضح الدراسات الحالية فعالية هذه النهج في تحسين دقة وفعالية تقييم مخاطر الهندسية الإنسانية في مختلف الصناعات.

كلمات مفتاحية:

تقييم الجسم السريع الكامل، إصابات الجهاز العضلي الهيكلي المتعلقة بالعمل، التعلم العميق، التحول الزمني، تقدير الوضعية.

"O Deep Thought computer," he said, "the task we have designed you perform is this.
We want you to tell us..." he paused, "the Answer!"

-Douglas Adams,
"The Ultimate Hitchhicker's Guide to the Galaxy"

Table of Contents

List of Figures

List of Tables

Glossary

CNN	Convolutional neural network.
NN	Neural Network.
AI	Artificial Intelligence.
ML	Machine Learning.
DL	Deep Learning.
DNN	Deep neural network.
IMU	Inertial measurement unit.
FCN	Fully connected network.
MLP	Multi layer perception.
FPN	Feature pyramid network.
MAS	Motion analysis system.
MPJPE	Mean Per Joint Position Error.
HRC	Human-Robot collaboration.
HRI	Human-Robot interaction.
NIOSH	National Institute for Occupational Safety and Health.
MSD	Musculoskeletal disorders.
WMSD	Work-related musculoskeletal disorders.
REBA	Rapid entire body assessment.
RULA	Rapid upper limb assessment.
OWAS	Ovako Working Posture Assessment System.

General introdution

0.1 Motivation

Even though there has been significant progress made in the field of industrial automation over the course of the past several decades, manual labour is still necessary in many different fields. Workers who perform tasks that require physical exertion are at an increased risk of developing musculoskeletal disorders as a result of their jobs. Musculoskeletal problems caused by labour have become an issue of concern on a global scale. If it were simple to get a hold of and competently carried out, its risk assessment would be of great use to ergonomic job design and occupational safety measures. In developed countries, WMSDs are currently the largest source of sick leaves, work-related disability, and an overall loss in productivity without forgetting the increasing of social healthcare costs. In the European Union, they account for more than half of all work-related diseases and are responsible for more than 40% of all economic losses caused by occupational health and safety concerns [1]. In the United States of America, these conditions are responsible for more than 30% of all diseases and accidents that do not result in death [2]. In 1997, the National Institute for Occupational Safety and Health in the United States advised a set of actions to be included in every ergonomic program in order to avoid work-related musculoskeletal disorders. This was done as a reaction to a problem that had arisen. As a direct result of this suggestion and in light of the fact that it is of the extreme significance that ergonomic risk assessment be carried out in a way that is both effective and efficient, an abundance of methods and tools pertaining to ergonomic risk assessment have been developed over the course of the past few decades. These can be broken down into the following categories [3, 4, 5] :

- **Self-assessment:** where employees evaluate themselves using standardised forms.
- **Human observation:** when qualified personnel observe workers and make educated guesses about the angles at which they are standing, sitting, bending and moving either on-site and/or off-line video.
- **Direct measurement:** where anthropometric equipment and gadgets are worn by workers to collect data on their activities for the sake of ergonomic analysis.
- **Computer-based assessment:** where human body models are automatically generated from camera captures by specialised computer vision application, thereby giving systematic and objective model-based ergonomic measures [6].

Both self-assessment and human observation have subjective biases that make them inconsistent in the matter of result. At the moment, the human observation approach is the most common way for industrial workplaces to assess ergonomic risks. But even experienced ergonomists often make mistakes when making subjective category decisions. This is mostly because of poor visual conditions in the workplace, such as poor lighting, occlusions, and bad camera angles when taking videos or pictures [7]. Direct measurement and expert-based observation methods are also limited by the amount of time it takes to do the assessment and the technical knowledge the analysts need [8]. In the last few years, a number of disruptive technologies related to the "industry 4.0" paradigm have made it possible for ergonomics to be used in new and useful ways [9].In particular, artificial intelligence and technology breakthroughs are opening up novel paths for ergonomics thanks to the impact of automated data collecting and analysis on a new class of data-driven applications. In this group, we highlight computer vision systems like colour and depth devices, stereo cameras, and RGB colour cameras [10, 7, 11]. These AI-based methods can accurately find and analyse a person's posture and motion by automating the assessment process.

0.2 Objective

The objective of this research is to design and implement a cutting-edge system that leverages computer vision and deep learning to assess full-body postures in the workplace. The system is based on the Rapid Entire Body Assessment (REBA) posture assessment tool and aims to automate the process of ergonomic risk assessment.

The study focuses on the creation of a computer vision model that can accurately predict the human's body joint coordinates from images or digital videos of workers performing common occupational tasks such as lifting, pulling, machining, pushing and others.

The system calculates the angle between different body segments and computes relevant postural scores, which are then translated into REBA Grand scores that indicate the level of risk associated with a specific posture. The proposed system is validated through a statistical comparison with manual evaluations performed by two ergonomic experts, ensuring its reliability and accuracy. This method significantly reduces the time and resources required for REBA evaluations by eliminating the need for manual sampling and evaluation of posture from video recordings of workplace tasks.

In general, this research constitutes a significant advancement in the domain of ergonomics and occupational health by presenting a resolution to the difficulties linked with conventional ergonomic risk evaluation techniques. The findings of this study possess the capability to provide insights for forthcoming advancements in the field of ergonomic hasard evaluation and aid in the creation of novel and enhanced instruments and approaches for averting work-related musculoskeletal disorders.

0.3 Thesis structure

- **Introduction**: briefly introduce the motivation behind this research and state the objectives of the research.
- **Chapter 1**: this chapter defines and explains the main key concepts that this research relies on such as ergonomics, occupational health safety, risk assessment methods and artificial intelligence.
- **Chapter 2**: this chapter reviews the state of the art of techniques and methods that will be used to build the project and provides a comprehensive overview about it.

- **Chapter 3**: this chapter describes the research template design and methodology for this thesis besides all the steps involved in the AI-based system developing.
- **Chapter 4**: this chapter presents the major findings of the research, including the AI-based system performance. Also we discuss its advantages and limitations, as well as the improvements and finished by a general discussion.
- **General conclusion**: this final chapter summarises the main results and major contributions of our research, also discusses its implications for enhancing the work quality and occupational health and safety.

CHAPTER 1

GENERALITIES

1.1 Introduction

The 4th industrial revolution has brought about a whirlwind of technological advancements in the manufacturing sector, streamlining processes and altering the nature of human work. however, this progress has also introduced new health and safety hasards for workers besides their lack of productivity, as well as challenges to their existing skills and knowledge. in this chapter we will explore the main components and their impact on the conduct of this research.

1.2 Ergonomics

Ergonomics is an interdisciplinary field that deals with the performance of humans in their work environment, including their interaction with machines and the design of their workplace. It is derived from the Greek word ***"ergon"*** (meaning "work") and ***"nomos"*** (meaning "natural laws") by Wojciech Jastrzebowski, a polish engineer. Also known as the ***"human factor"*** , ergonomics aims to apply natural laws governing human work by studying the human body and its responses to internal and external forces (such as anthropometry and biomechanics), as well as work and environmental physiology, human behaviour in response to work, information processing and decision making (skill psychology), training and effort perception, and adapting equipment and devices for human use [12].

When it comes to ergonomics, there are two common approaches for reducing the risk of injury and improving overall health in workplace, as shown in Figure 1.1. The first approach, ***"fitting the task to the person"***, focuses on modifying the work environment and equipment to reduce physical strain [13] and improve task efficiency [14], in the other side, the second approach, ***"fitting the person to the task"***, involves selecting workers with appropriate physical and cognitive abilities or providing training to help them perform the job demand. In practice, both approaches are important and should be considered in combination to achieve optimal results for both the employee and the task.

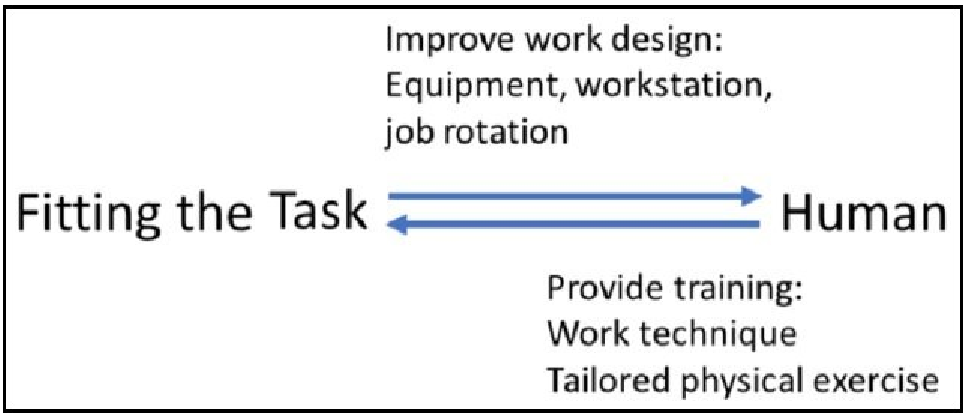


FIGURE 1.1: Illustration of two approaches in ergonomics. Based on [13, 15, 16, 17]

1.2.1 Scope of ergonomics

The purpose of ergonomics is to improve the health and capability of workers while simultaneously preserving their safety, comfort, and productivity [18]. During the early stages of this endeavor, the primary focus was on ensuring that human operators were capable of getting the most out of their respective pieces of machinery so as to maximise overall efficiency. Research on human performance was conducted by designers so that they may gain a better understanding of human capabilities, limitations, and responses to their surrounding environment. As time went on, Work-related satisfaction, quality of life, ergonomics, injury prevention, stress management, efficiency, and productivity were all found to be interconnected and intertwined [18]. It has become increasingly acknowledged that ergonomics plays an important role in cost containment, particularly in connection to absenteeism, retraining injured workers, medical expenses, liability insurance, and punitive damages. Today, ergonomics places a strong emphasis on ensuring the long-term health and safety of workers through the prevention and control of occupational injuries and illnesses, particularly in the area of cumulative trauma disorders and the associated costs. This can be done by keeping workers from getting hurt or sick on the job and taking care of those who do.

1.2.2 Ergonomics evolution

The focus on work-related musculoskeletal disorders and ergonomics has been around for centuries. In the early 20$^{\text{th}}$ century, Frederick Taylor and Frank and Lillian Gilbreth analysed human performance through task, motion, and time analysis, aiming to shorten cycle times for repetitive jobs such as bricklaying and manual material handling. Technology continued to evolve at an accelerating pace, and engineers and psychologists joined forces to understand how to design the human-machine interface. At the end of the 20$^{\text{th}}$ century, personal computers became a major focus in the field of ergonomics. In response, labour organisations in the United States pushed for the creation of ergonomic workplace standards to protect workers from these types of disorders.

However, some industries had successful ergonomic programs in place, while others pushed for the repeal of the OSHA ergonomic standard, which was eventually repealed by Congress in 2001. The implementation of ergonomic best practices has been shown to reduce workers' compensation costs and increase productivity, making it a cost-effective motivation for promoting safe work environments. As the world economy continues to become more global, many countries are adapting to incorporate ergonomic practices and standards. In America, ergonomics is a vital aspect of worker health and safety, and as part of the National Institute for Occupational Safety and Health's Total Worker Health initiative, ergonomic programs are integrated with health promotion activities, leading to greater effectiveness.

For example, combining an occupational health intervention that reduces the respiratory hasard with a smoking cessation program doubles the smoking quit rate. All four of the NIOSH Total Worker Health Centers emphasized strong ergonomic initiatives as a key component of their comprehensive approach to workplace health and safety [19].

1.2.3 Ergonomics across USA, Europe and Africa

While the USA and Europe have made significant strides in designing tasks, workspaces, and equipment that fit employees' physical capabilities and limitations, Africa lags behind in both implementation and awareness. The National Institute for Occupational Safety and Health (NIOSH) in the USA focuses on musculoskeletal disorders and provides recommendations and guidelines for various industries, while numerous industries and manufacturers in Europe apply ergonomic principles to their products and work environments, resulting in safer and more efficient workplaces in both regions. However, limited resources, a lack of awareness, and inadequate infrastructure contribute to the gap between Africa and the more developed regions. Additionally, the continent's diverse cultural, social, and economic contexts make it difficult to establish a unified approach to designing workspaces and equipment that consider employees' physical abilities and limitations. The absence of guidelines and standards tailored to the African context exacerbates the problem, resulting in a slower adoption of these practices. To bridge this gap, African countries must invest in research, education, and context-specific guidelines to improve the well-being and productivity of their workforce.

1.3 Occupational health

1.3.1 Musculoskeletal disorders (MSDs)

Musculoskeletal disorders, and particularly the sense of musculoskeletal-related pain, are among the many unanswered questions regarding a person's impression of ill health, which may or may not have a physiological cause. Msds are defined as the injuries and disorders to muscles, nerves, tendons, ligaments, joints, cartilage and spinal discs [20] and do not include injuries resulting from slips, trips, fall or accidents. Movements of the arms and hands that cause MSDs include flexing, extending, twisting, clenching, and reaching. As noted by (Lorusso et al, 2009) [21], numerous epidemiological studies have demonstrated that ergonomic factors and elements of work organisation significantly contribute to the onset of musculoskeletal disorders. Simple, routine movements are not inherently dangerous in typical work tasks. However, the repetition of these movements, often performed with force and at a high speed with little time for recovery, can make them hasardous in the workplace.

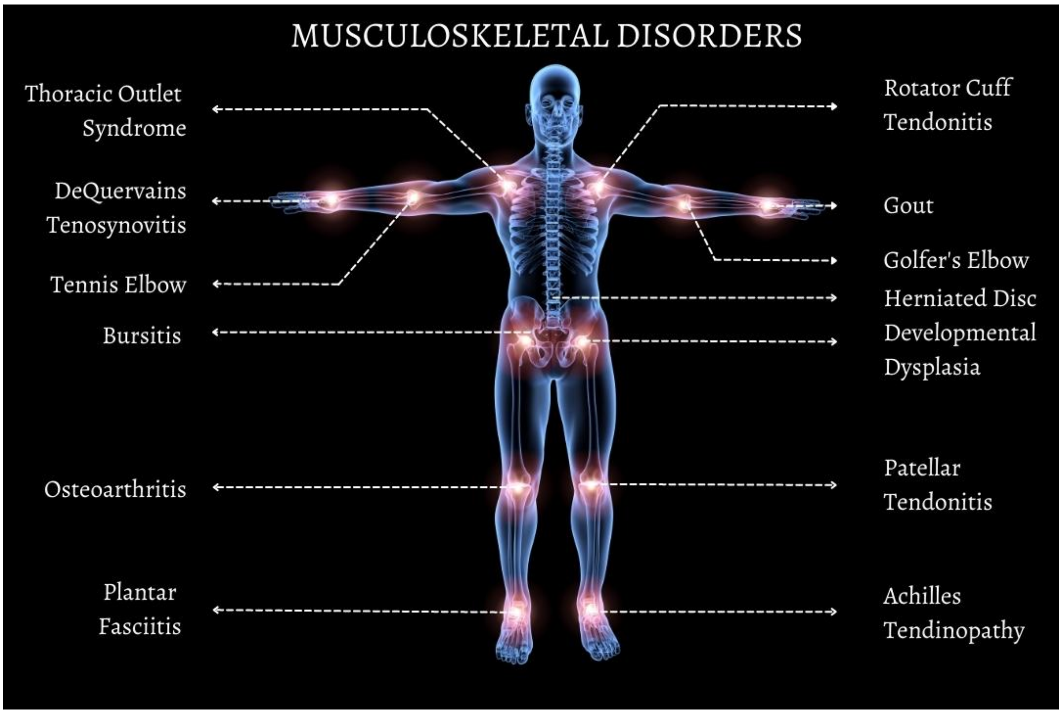


FIGURE 1.2: Common MSDs [22]

1.3.2 Diagnosis musculoskeletal disorders

MSDs have a correlation with certain types of work patterns [23]:

- Fixed or constrained body positions.
- Continual repetition of movements.
- Force concentrated on small parts of the body, such as the hand or wrist.
- A pace of work that does not allow sufficient recovery between movements

Musculoskeletal disorders are generally the result of the combined and interactive effects of various factors, including repetitiveness, forceful movements, and a lack of recovery time. Heat, cold, and vibration can also contribute to the development of MSDs. The areas of the human body that are most susceptible to these disorders include the lower back, upper back, neck, shoulders, knees, hipsthighs, elbows, anklesfeet, and wrists. But, the root causes of MSDs are complex and multi-faceted and can be examined from two key perspectives biomechanical physical or ergonomic factors and non-biomechanical factors [12]. The wear and tear of daily activities can cause damage to muscle tissue, and trauma to an area, such as sudden movements, automobile accidents, falls, fractures, sprains, dislocations, or direct blows to the muscle, can also lead to musculoskeletal pain. Musculoskeletal disorders can be seen in three stages: the early stage, the intermediate stage, and the late stage [24]. Each stage is linked to a certain time when pain shows up and this is illustrated in the figure 1.3 below:

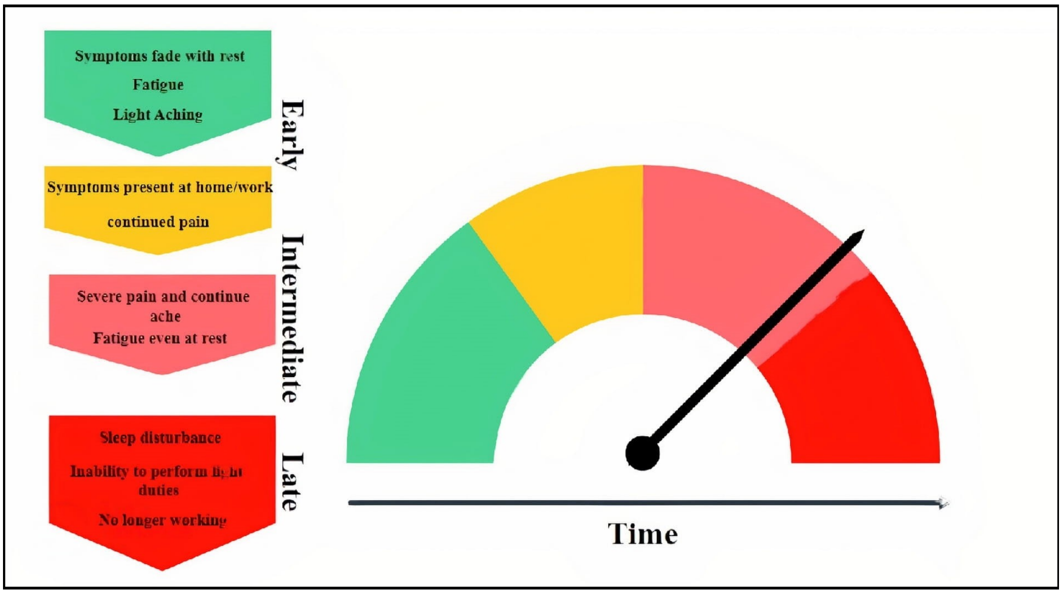


FIGURE 1.3: MSDs' stages development through time

1.3.3 Work-Related Musculoskeletal Disorders

Work-related musculoskeletal disorders (WMSDs) are injuries to workers' limbs that are caused or worsened by the working conditions in a workplace [25]. These disorders occur when there is a mismatch between the physical demands of the job and the physical ability of the worker's body. Some of the working conditions that can lead to WMSDs include regular heavy lifting, exposure to whole-body vibration, overhead work, working with the neck in a flexed position for long periods, and performing repetitive and forceful tasks [12]. The specific body parts affected by WMSDs depend on the nature of the tasks involved. For example, tasks that involve the use of the upper body may cause pain in the upper arm, lower arm, wrists, neck, and shoulders, while tasks involving the lower body may cause pain in the legs, trunk, and feet [23]. Some common examples of WMSDs include Carpal Tunnel Syndrome (Figure 1.4), Thoracic Outlet Syndrome (Figure 1.5), and Tendonitis (Figure 1.6).

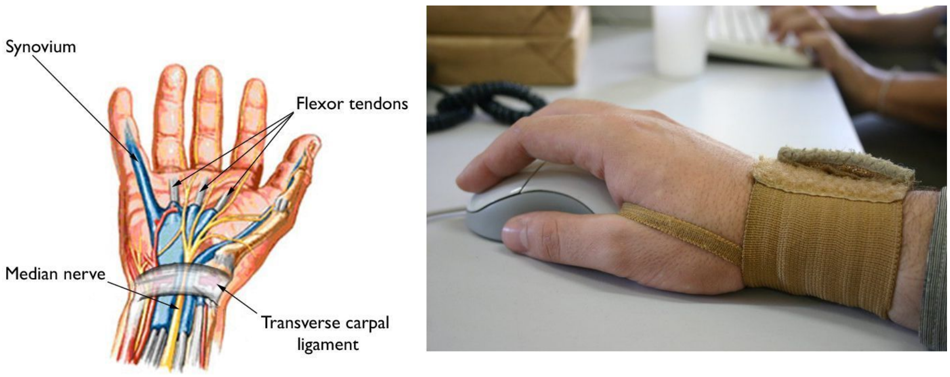


FIGURE 1.4: Carpal Tunnel Syndrome [24]

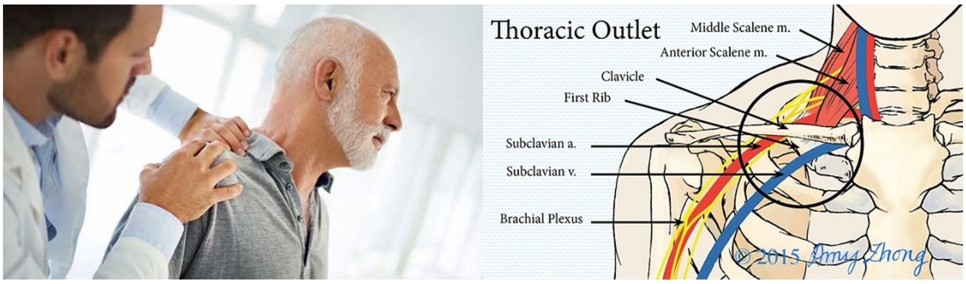


FIGURE 1.5: Thoracic Outlet Syndrome [26]

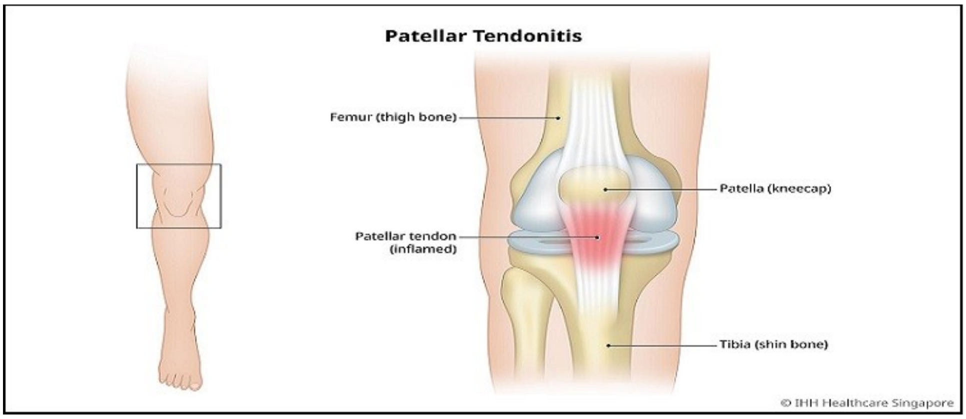


FIGURE 1.6: Patellar Tendonitis [27]

The connection between physical labour and its impact on health is influenced by numerous factors. It is noteworthy to mention that there are also several other hasards and their interactions not included in this model, such as organisational and psychosocial aspects. Work is determined by the tasks, workplace, tools, and schedules, referred to as the prescribed work [13]. Every individual is unique and has varying effects on how work tasks are actually carried out. Personal traits such as height, work methodology, and experience, as well as current personal conditions, play a significant role in the actual work activity. An internal physiological response, including muscular contractions and metabolic changes, will occur based on the activity performed and the individual's capability. This response can result in either fatigue and reduced health or sustained and improved health, depending on the length, frequency, and intensity level of the actual work activity [28]. By monitoring the actual work activity through techniques such as evaluating posture, force, and energy demand, researchers and practitioners can assess ergonomic hasards and enhance work design.

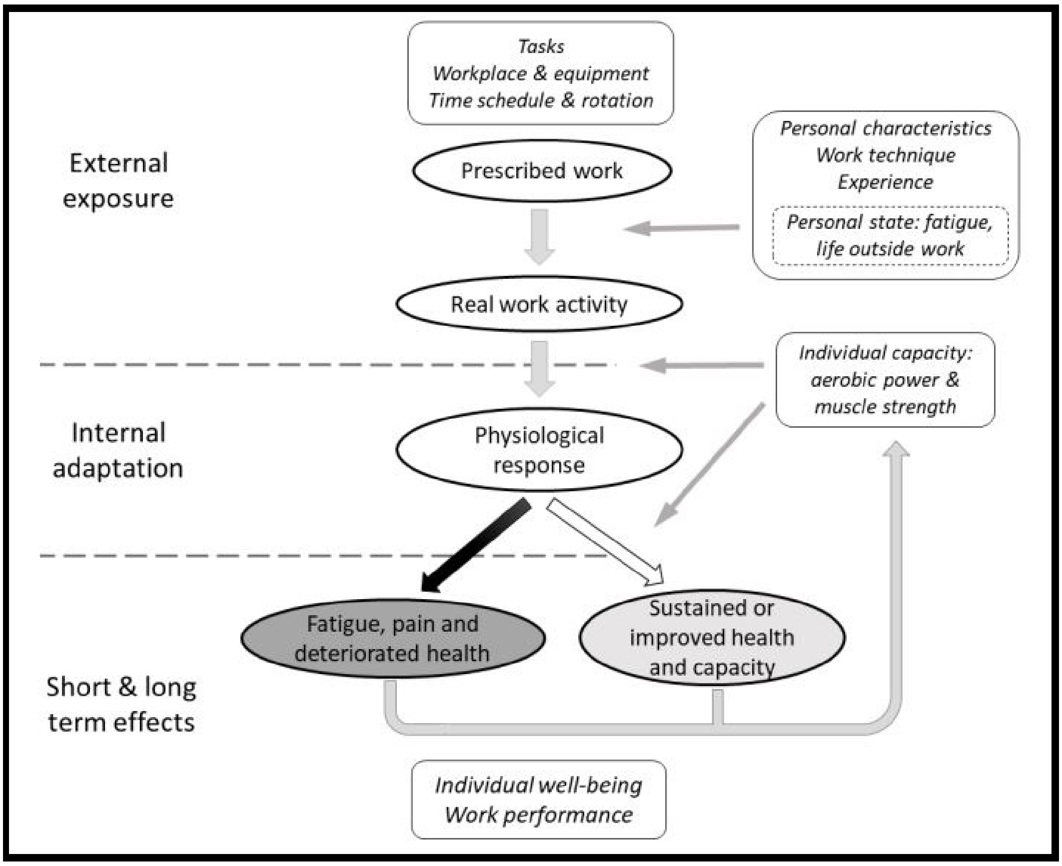


FIGURE 1.7: Model of relationships between physical work and its effects on health, with modifying factors. Model developed based on [28]

1.3.4 Prevention and Management

Preventive approaches to tackle WMSDs entail performing ergonomic risk assessments [29]. These assessments entail scrutinising potential risk factors, including exposure to detrimental biomechanical and psychosocial aspects of work and the workplace, with the objective of identifying noteworthy risks associated with musculoskeletal disorders. To gauge the degree of exposure to risk factors, each factor ought to be assessed based on three key quantitative attributes [30]:

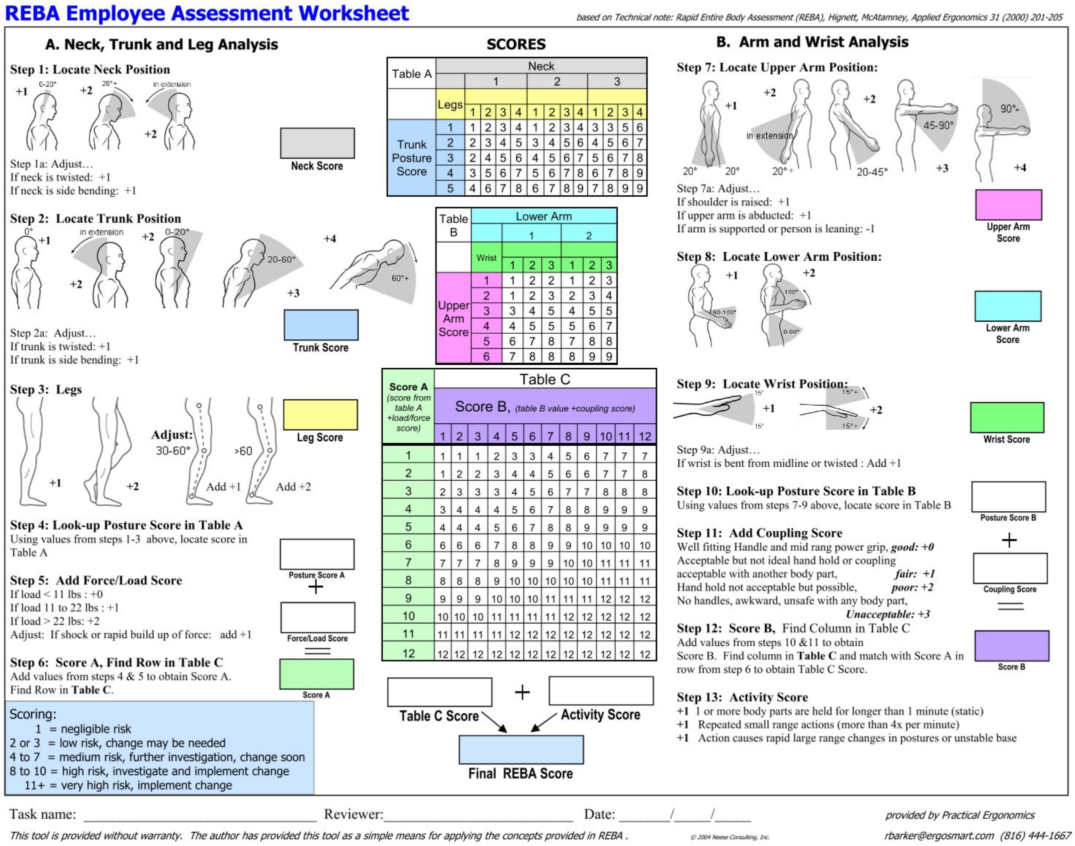
- **Intensity/amplitude:** The intensity, or "how much?", measures the size of the risk factor, such as the size of the joint angle in posture-related risk or the weight of the object handled in musculoskeletal load-related risk.
- **Frequency:** The frequency, or "how often?", measures the presence of the risk factor, including repetitiveness, number of work cycles, task variation, and occurrence of micropauses, taking into account the time needed for tissue recovery.
- **Duration:** The duration, or "how long?", measures the length of exposure to the risk factor.

However, although the possibility to automate systematic methods, still, the process ergonomists' experience-driven, despite the fact that AI-based assessment provide real-time and accurate result without the traditional pen-paper process that ends with subjective judgment.

1.4 Common Observational-based method for ergonomic risk assessment

1.4.1 REBA (Rapid Entire Body Assessment)

The method known as Rapid Entire Body Assessment (REBA) was created by Sue Hignett and Lynn McAtamney [31] at Nottingham Hospital in the year 2000. This approach was developed through the collaborative efforts of team consisting of ergonomists, physiotherapists, and nurses, who identified approximately 600 different working postures [31]. According to (Dohyung Kee, 2021) [32], REBA can be defined as a method or technique that enables a quick evaluation of a worker's neck, back, upper arms, forearms, wrists, and feet posture during any activity that may potentially result in musculoskeletal disorders (MSD), this means that it furnishes a system of scoring that evaluates the range of motion of specific body parts and the amount of force or load applied to the body during the task, additionally it's applicable to both the upper and lower body. The following worksheet shown in Figure 1.8 provides a framework for the assessment:



REBA Employee Assessment Worksheet

based on Technical note: Rapid Entire Body Assessment (REBA), Hignett, McAtamney, Applied Ergonomics 31 (2000) 201-205

A. Neck, Trunk and Leg Analysis

Step 1: Locate Neck Position

+1 0-20° +2 20°+ in extension +2

Step 1a: Adjust…
If neck is twisted: +1
If neck is side bending: +1

Neck Score

Step 2: Locate Trunk Position

0° +1 in extension +2 0-20° +2 20-60° +3 60°+ +4

Step 2a: Adjust…
If trunk is twisted: +1
If trunk is side bending: +1

Trunk Score

Step 3: Legs

+1 +2 Adjust: 30-60° Add +1 >60 Add +2

Leg Score

Step 4: Look-up Posture Score in Table A
Using values from steps 1-3 above, locate score in Table A

Posture Score A
+

Step 5: Add Force/Load Score
If load < 11 lbs : +0
If load 11 to 22 lbs : +1
If load > 22 lbs: +2
Adjust: If shock or rapid build up of force: add +1

Force/Load Score
=

Step 6: Score A, Find Row in Table C
Add values from steps 4 & 5 to obtain Score A.
Find Row in **Table C**.

Score A

Scoring:
1 = negligible risk
2 or 3 = low risk, change may be needed
4 to 7 = medium risk, further investigation, change soon
8 to 10 = high risk, investigate and implement change
11+ = very high risk, implement change

SCORES

Table A	Neck 1				Neck 2				Neck 3			
Trunk Posture Score / Legs	1	2	3	4	1	2	3	4	1	2	3	4
1	1	2	3	4	1	2	3	4	3	3	5	6
2	2	3	4	5	3	4	5	6	4	5	6	7
3	2	4	5	6	4	5	6	7	5	6	7	8
4	3	5	6	7	5	6	7	8	6	7	8	9
5	4	6	7	8	6	7	8	9	7	8	9	9

Table B	Lower Arm 1			Lower Arm 2		
Upper Arm Score / Wrist	1	2	3	1	2	3
1	1	2	2	1	2	3
2	1	2	3	2	3	4
3	3	4	5	4	5	5
4	4	5	5	5	6	7
5	6	7	8	7	8	8
6	7	8	8	8	9	9

Table C

Score A (score from table A +load/force score) \ Score B, (table B value +coupling score)	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	2	3	3	4	5	6	7	7	7
2	1	2	2	3	4	4	5	6	6	7	7	8
3	2	3	3	3	4	5	6	7	7	8	8	8
4	3	4	4	4	5	6	7	8	8	9	9	9
5	4	4	4	5	6	7	8	8	9	9	9	9
6	6	6	6	7	8	8	9	9	10	10	10	10
7	7	7	7	8	9	9	9	10	10	11	11	11
8	8	8	8	9	10	10	10	10	10	11	11	11
9	9	9	9	10	10	10	11	11	11	12	12	12
10	10	10	10	11	11	11	11	12	12	12	12	12
11	11	11	11	11	12	12	12	12	12	12	12	12
12	12	12	12	12	12	12	12	12	12	12	12	12

Table C Score + Activity Score

Final REBA Score

B. Arm and Wrist Analysis

Step 7: Locate Upper Arm Position:

20° 20° +1 in extension 20°+ +2 20-45° +2 45-90° +3 90°+ +4

Step 7a: Adjust…
If shoulder is raised: +1
If upper arm is abducted: +1
If arm is supported or person is leaning: -1

Upper Arm Score

Step 8: Locate Lower Arm Position:

+1 60-100° +2 100°+ 0-60°

Lower Arm Score

Step 9: Locate Wrist Position:

15° 15° +1 15°+ 15°+ +2

Step 9a: Adjust…
If wrist is bent from midline or twisted : Add +1

Wrist Score

Step 10: Look-up Posture Score in Table B
Using values from steps 7-9 above, locate score in Table B

Posture Score B
+

Step 11: Add Coupling Score
Well fitting Handle and mid rang power grip, ***good: +0***
Acceptable but not ideal hand hold or coupling acceptable with another body part, ***fair: +1***
Hand hold not acceptable but possible, ***poor: +2***
No handles, awkward, unsafe with any body part, ***Unacceptable: +3***

Coupling Score
=

Step 12: Score B, Find Column in Table C
Add values from steps 10 &11 to obtain Score B. Find column in **Table C** and match with Score A in row from step 6 to obtain Table C Score.

Score B

Step 13: Activity Score
+1 1 or more body parts are held for longer than 1 minute (static)
+1 Repeated small range actions (more than 4x per minute)
+1 Action causes rapid large range changes in postures or unstable base

Task name: ______________ Reviewer: ______________ Date: ____/____/____

provided by Practical Ergonomics

This tool is provided without warranty. The author has provided this tool as a simple means for applying the concepts provided in REBA.

© 2004 Neese Consulting, Inc.

rbarker@ergosmart.com (816) 444-1667

FIGURE 1.8: REBA worksheet

This method encompasses a series of actions such as monitoring labour activities, analysing body position, evaluating postures, calculating predetermined scores, determining the REBA score outcomes and promptly validating the action level as shown in the table 1.1:

Score	Level of MSD risk
1	Negligible risk, no action required
2-3	Low risk, change may be needed
4-7	Medium risk, further investigation, change soon
8-10	High risk, investigate and implement change
11+	Very high risk, implement change

TABLEAU 1.1: REBA score interpretation

The calculation of the REBA score involves using several tables (Table A, B, and C) and the REBA scoring format, which follows this methodology:

- Table A is used to obtain a score from group A (Trunk score, Neck score, and Legs score).
- Table B is used to obtain a score from group B (Upper arm score, Lower arm score, and Wrist score).
- Table C is used to determine a score based on the scores from Tables A and B.
- The final REBA score is obtained by summing the scores from Table C and the activity scoring.

To determine the risk level of musculoskeletal disorders (MSD), the scoring outcome will be compared to the interpretation table of the REBA score mentioned earlier. The following flowchart given by Hignett and McAtamney [31] illustrates the calculation methodology:

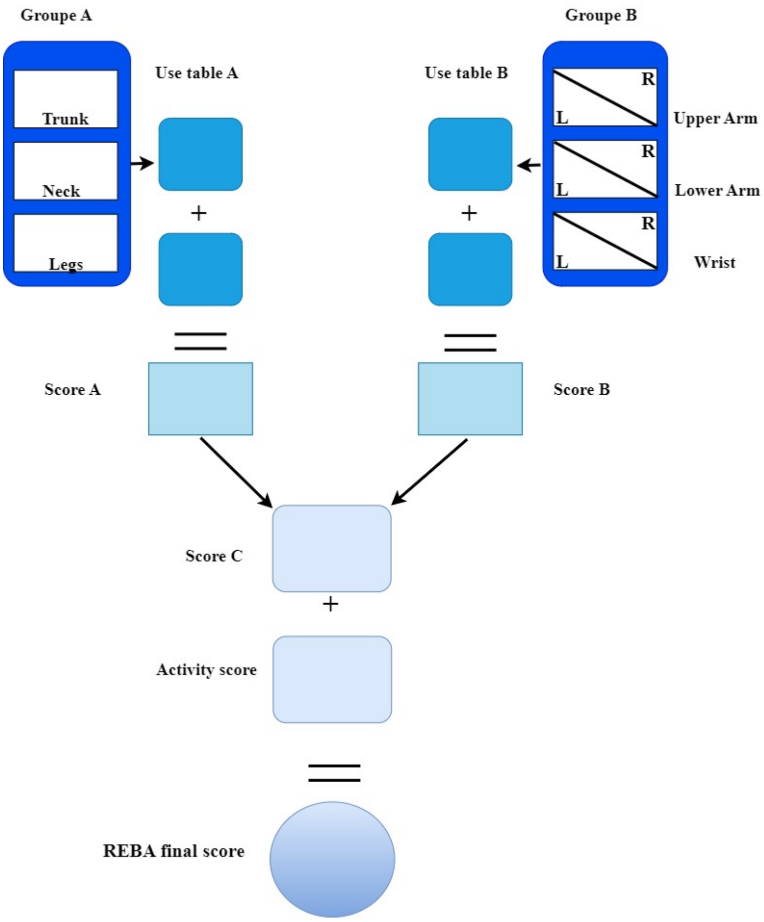


FIGURE 1.9: REBA score sheet methodology

1.4.2 RULA method (Rapid Upper Limb Assessment)

Rapid upper limb assessment, or RULA, is an observation-based method developed by L. McAtamney and E. N. Corlett [32] used to evaluate the posture of employees during work that involves the body's upper limb, where its aim is to identify potential work-related musculoskeletal disorders that may occur while doing the job and by assessing the body's posture, force, and repetition, and to provide recommendations in order to prevent and reduce the risk of occupational injury. The following figure (1.10) presents the RULA worksheet:

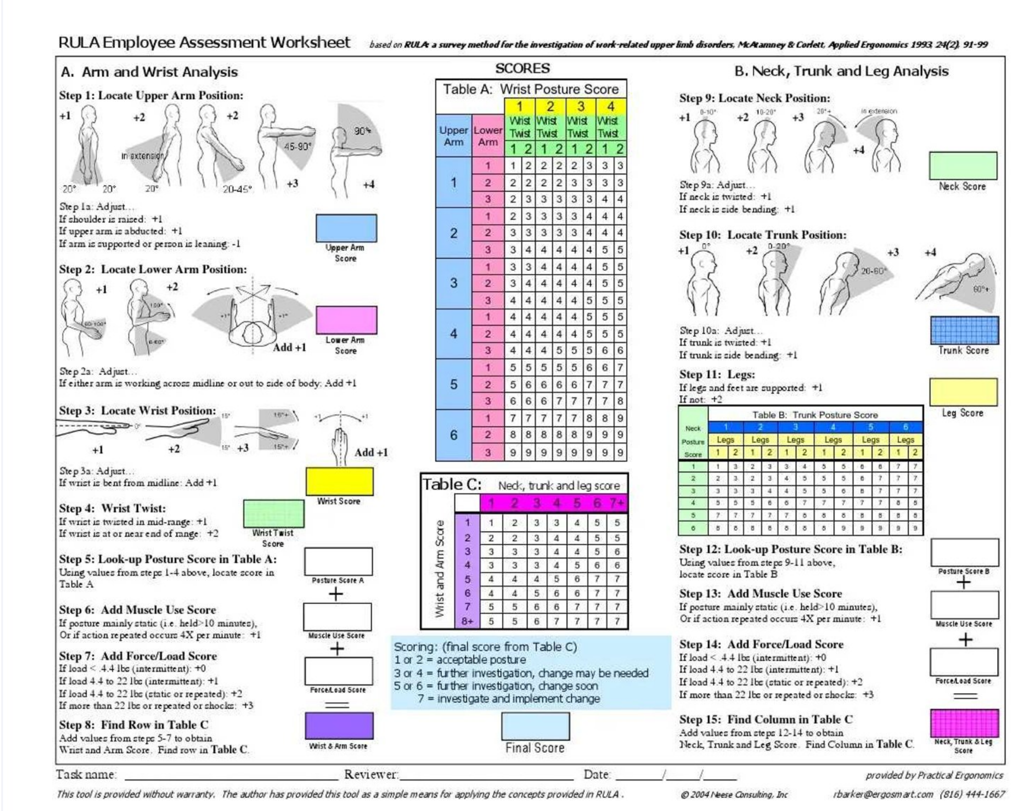


FIGURE 1.10: RULA worksheet

The calculation of RULA is done following these steps:

- Table A is used to obtain score from Group B which includes: Upper Arm, Lower Arm, Wrist and Wrist twist.
- Table B is used to obtain score from Group B which includes: Neck, Trunk and Legs.
- Calculate score C by summing Posture score A, muscle use score and force score.
- Calculate score D by summing Posture score B, muscle use and force score.

The figure 1.11 illustrates the RULA methodology:

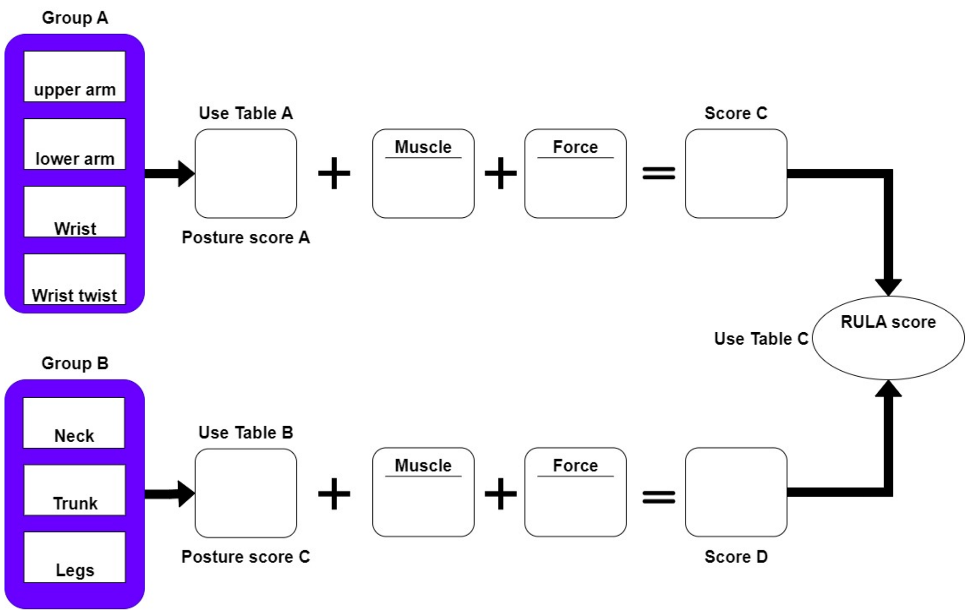


FIGURE 1.11: RULA score sheet methodology

The RULA final score is translated into four levels of action that specify the necessary intervention level to decrease the chances of physical loading injury to the worker as shown in the table 1.2 below:

Score	Level of MSDs risk
1-2	Acceptable working posture, no action required
3-4	Low risk, further investigation to change posture
5-6	Important risk, investigations to change posture
7+	Risk very high, posture changes required immediately

TABLEAU 1.2: RULA score interpretation

1.4.3 OWAS (Ovako Working Posture Analysis System)

OWAS, standing for "Ovako Working Posture Analysis System," is an observational technique that was developed by Ovako Oy [32], one of the largest steel bar and profiles manufacturers in Europe [33]. The method is based on assessments of work postures conducted in various divisions of the steel factory by 32 experienced steelworkers and international ergonomists. The main objective of this approach is to categorize work postures into one of 252 potential combinations, considering the back (4 categories), upper limb (3 categories), and lower posture (7 categories), as well as the weight of the load or the amount of force used (3 categories), as illustrated in the figure 1.12 below:

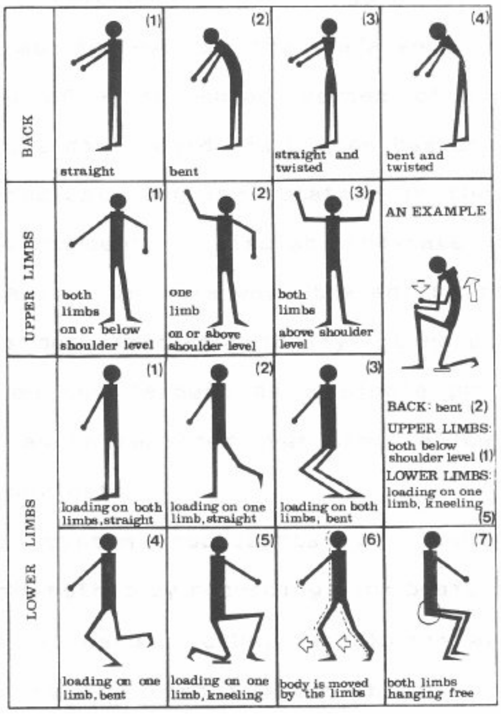


FIGURE 1.12: OWAS classification chart [33]

The classified work posture identified using OWAS is divided into four action categories, as shown in Table 3.3. These categories indicate the level of risk of injury associated with the posture and the priority for taking corrective actions, as described by (Lee & Han, 2013) [34].

OWAS Action Class	Interpretation
AC 01	Adequate postures, no special attention needed, except some cases
AC 02	Posture indicates some WMSDs risk, corrective measures need to be taken soon
AC 03	posture has serious harmful impact on musculoskeletal system, corrective measures must be taken as soon as possible
AC 04	Posture needs immediate changing

TABLEAU 1.3: Action class interpretation

1.4.4 Comparison of observation-based methods

The implementation of intervention programs for lowering exposure to WMSDs risk factors is the most well-known preventive measure, as stated in [35, 36, 37]. REBA, RULA, and OWAS are efficient and straightforward postural targeting methods for analysing worker postures and calculating the risk factor index. A number of research compared REBA vs. RULA vs. OWAS in order to extract significant insights for the aim of finding the best assessment method to apply in workplace under multiple contexts.
The general characteristics of the three methods which have been summarized by [38] are:

- Posture
- Force / External load
- Repetition motion
- Static posture
- Dynamic loading
- Coupling

The result fined by (Dohyung Kee, 2022) in [39] shows that the OWAS method can't be applied to assess the motion repetition, static posture, Dynamic loading and coupling, in the other hand, RULA and REBA assess the posture and Force/external load, as well as motion repetition and the effect of static posture. Meanwhile, in comparison to RULA, REBA has additional features to evaluate, namely dynamic loading, which is a quick shift in an unstable posture or basis, and coupling. During evaluation, the OWAS does not distinguish between the left and right upper extremities, whereas the RULA and REBA examine just the side thought to be under more stress. When it is difficult to determine which side is under more load, both sides are evaluated. The OWAS evaluates postural loads based on time sampling, whereas the REBA chooses and estimates the most frequent, protracted, or loaded postures [40]. Yet, it is typical for the OWAS and RULA to observe the most frequent, lengthy, or laden postures, as seen in the REBA. The three observational classification approaches include four or five action types or levels [31, 33, 41].

Although all those results, still the three methods do not account for the impact of recovery, vibration, duration, ambient condition, psychological and individual characteristics [39] which are known to influence the incidence of WMSDs. The table below provides a concise summary of the relatives strengths and limitations of the three methods [40].

Methods	**Strengths**	**limitations**
RULA	Fast and easy for use and assessment	The emphasis is on upper limb posture, coupling is excluded, difficulty of choosing which side to assess.
REBA	Fast and easy for use and assessment	The necessity to choose the right side to assess.
OWAS	Most efficient and user-friendly method for classifying leg posture in detail.	Exclusion of neck, elbow and wrist postures,the static posture and coupling are ignored.

TABLEAU 1.4: strengths and limitations of the three methods

1.5 Conclusion

In conclusion, we have discussed in this chapter the importance of ergonomics and the impact of musculoskeletal disorders (MSDs) and work-related musculoskeletal disorders (WMSDs) on employees' health and productivity, and different observational methods for ergonomic risk assessment were also presented. However, these methods are often time-consuming and subject to observer bias. In the next chapters, we will focus on the development of an automated system for ergonomic risk assessment using AI techniques.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In this chapter, we will explore the existing literature on the AI techniques that could be in ergonomic risk assessment, also, we'll analyse the advantages and limitations of these approaches and identify the research gaps that will be addressed in this study. Finally, we'll outline the research questions and objectives that will guide our investigation and provide a roadmap for the subsequent chapters of this thesis.

2.2 Artificial Intelligence impact on ergonomic risk assessment

The use of AI in ergonomic risk assessment can considerably increase the process's accuracy and efficiency because manual observations or surveys, which are frequently used in traditional ergonomic risk evaluations, can be time-consuming and error-prone. On the other hand, AI can aid in the automation of these processes by analysing data from a variety of sources, including motion sensors, wearable devices, and video recordings, to identify potential risks and estimate the risk of injury. This can lead to a more thorough and accurate examination of the workplace environment, improving overall workplace safety.AI can also assist in identifying patterns and trends in occupational injuries, providing insights into potential threats and allowing proactive steps to be implemented to prevent injuries. As an example, an AI-based system can evaluate data from injury reports to discover common variables such as repetitive motions or awkward postures, which can aid in identifying areas of the workplace that require attention. This can lead to the deployment of targeted measures to lower the risk of harm, such as work redesign or training

programs.
Evaluating, identifying, and assessing ergonomic risks for human operators in manufacturing is critical. Both physical and mental workloads, as well as fatigue, must be researched to determine the correlation between operator actions and labour events, and quantify the relationship between human work posture and the degree of ergonomic risk. Wearable gadgets, sensors, and videos are the primary data sources used to record operators' behaviours. Therefore, an operator model for ergonomic risk can be evaluated from the perspectives of physical activity assessment (through sensor-based and video data) and risk stratification, mental workload evaluation, and tiredness classification [42].

2.2.1 Sensor-based work assessment

A range of sensors and wearable devices have been introduced to gather data on posture and movement patterns. The work done by (Elena et al., 2021) [43] presents a systematic assessment of ergonomic wearable gadgets and how they can improve ergonomics. (Conforti et al, 2020) [44] used wireless inertial measurement devices (IMUs) and ML algorithms recognise posture patterns to quantify biomechanical risk in lifting load tasks, while (Hosseinian et al, 2019) [45] used random forest and SVM models to classify static and dynamic work activities from a chest accelerometer.

2.2.2 Video-based motion analysis

In addition to sensor data, cameras and movies have been widely employed to record human movement and gestures. Using computer vision and ML approaches, (Fernandez et al, 2020) [6] proposed a method to automatically compute rapid upper limb assessment (RULA) scores from digital video. Correspondingly, using a convolutional neural network (CNN)-based pose detector to infer 2D poses from photos and a deep neural network (DNN) to estimate RULA action levels, (Li et al, 2020) [46] introduced a real-time method for assessing postural risk factors linked with MSDs. Using ML for time and motion studies in laboratory simulations of timed repetitive activities for different hand activity level (HAL) and from footage of people completing 50 industrial tasks based on decision tree and the presented feature vector training methods, (Akkas et al, 2016) [47] and (Akkas et al, 2018) [48] developed the technique. In addition, a neural network-based motion analysis system (MAS) is created by (Bartolini et al, 2018) [49] to analyse the ergonomic risk in manual and assembly tasks by collecting operator movements and postures. In addition, a publication [50] presents an application utilising Microsoft Kinect [11] to utilise ML techniques such as AdaBoost [51] Trigger indication to detect lifting and lowering actions with real-time motion data capturing on the shop floor for ergonomic evaluations and risk assessment. In addition, statistical process control and data analytic techniques are utilised in this work [52] to develop a human motion analytic system to identify repetitive motion patterns and deviations from those patterns by collecting, transforming, storing, and analysing data from repetitive physical motions performed by manufacturing workers.

2.3 Human motion analysis

Throughout the centuries, our understanding of human movement has been contingent on the available human motion capture techniques [53]. In the last few decades, many 3D body posture systems have been made. Most of them fall into one of three main categories: direct measurement, observational methods, and the marker-less motion capture system, which is the focus of our thesis.

2.3.1 Marker-less capture system

Current computer vision and machine learning techniques have significant potential for marker-less human motion capture, but they have not been extensively researched for biomechanics applications, which require greater precision and resilience than other applications [54]. Recent research has mostly focused on monocular images and hard settings, like wild environments and figuring out where multiple people are [55]. Monocular images can be acquired by a single camera and are favoured for surveillance and entertainment purposes, but their performance is poor due to the ambiguous nature of 3D-2D projection. Self-occlusion is a significant contributor to this ambiguity, which can be resolved by employing numerous cameras. Hence, biomechanics applications often require many cameras to acquire multi-view images and increase pose estimation or tracking precision. Several researches have investigated computer vision and machine learning and offered marker-free approaches for biomechanical applications. Specifically, (Corazza et al, 2006)[56] and (Sandau et al, 2014) [57] have developed a generative technique for fitting a specified 3D body model to an eight-camera visual hull. The process of fitting is presented as an optimisation problem, and body segmentation and least-squares optimisation are utilised to estimate the joint centre positions. Even though these approaches are highly accurate, they rely heavily on background removal, which requires a controlled setting and lighting conditions. In addition, a huge number of cameras are required to create an accurate visible hull surface. Using training data, (Drori et al, 2017) [58] constructed a discriminative approach to directly map a monocular image to body position parameters from a monocular image. It is demonstrated that their system is capable of accurately calculating a cyclist's full-body kinematics and two-dimensional position. Unfortunately, their method's performance for 3D body pose estimation has not been evaluated. These researches illustrate the viability of computer vision and machine learning methodologies for biomechanics applications, however, their results have not been confirmed for additional biomechanical analysis, such as the calculation of joint force and moment.

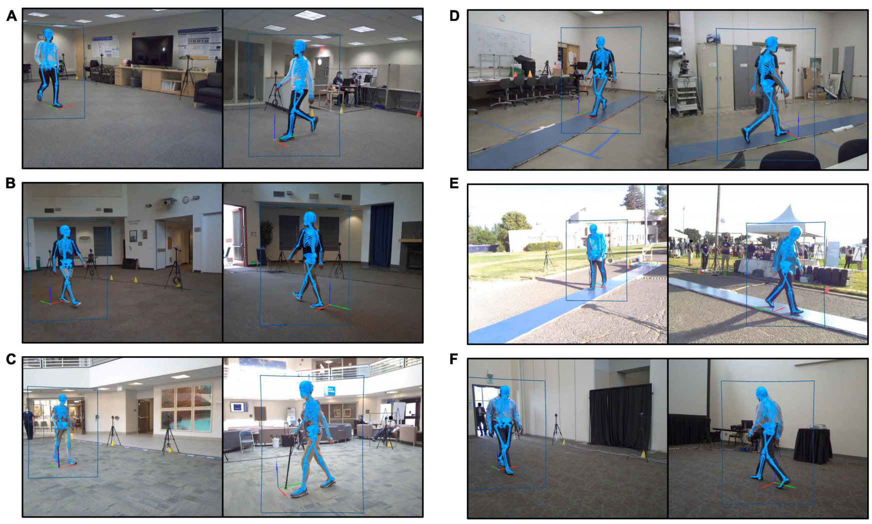


FIGURE 2.1: images by the marker-less motion capture system [59]

2.4 Human Pose Estimation approaches

2.4.1 The 2D human pose estimation

Various algorithms have been developed to determine the pixel coordinates (x_i, y_i) of each joint i in a static image, where these algorithms are typically classified into generative, discriminative and hybrid approaches.

a. Generative approaches

Back in 1980, Joseph O'Rourke and Norman I. Badler [52] proposed a method for 2D posture estimation in videos based on a top-down approach, where the human body is considered a connected tree model. Paper's [60] pictorial structures method introduced a deformable parts model in which the appearance of each component is treated independently, and pairwise interactions are modelled by spring-like connections between the parts. These developments led to the creation of broad frameworks for structural object detection [61], as well as significant advancements in 2D human pose estimation [62, 63]. Many of these generative approaches require storing a parametric model to learn spatial correlations between various body parts and use efficient graphical inference techniques for final pose estimation.

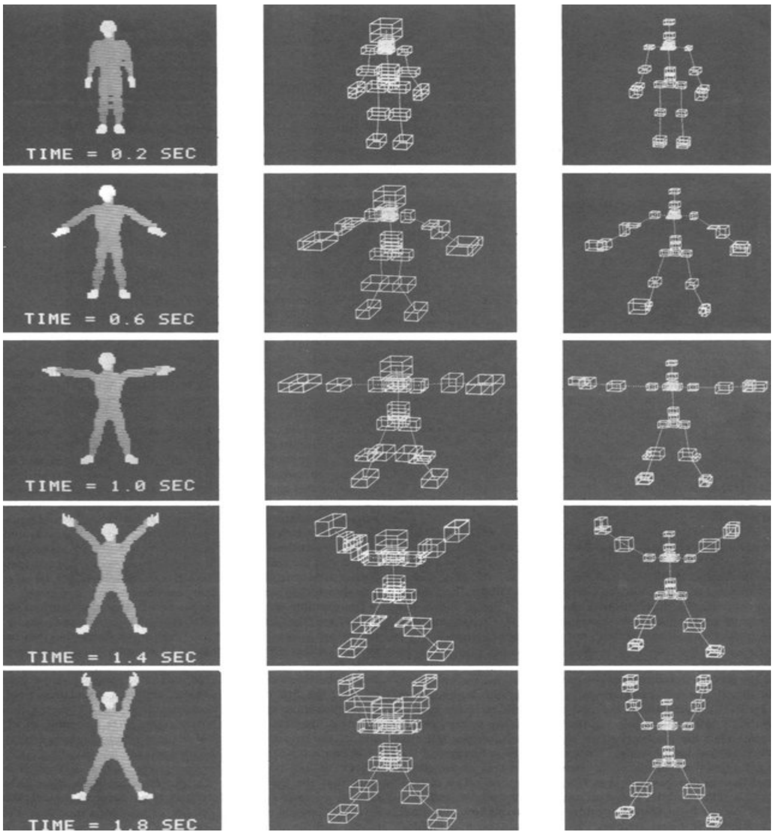


FIGURE 2.2: Tree-based model [52]

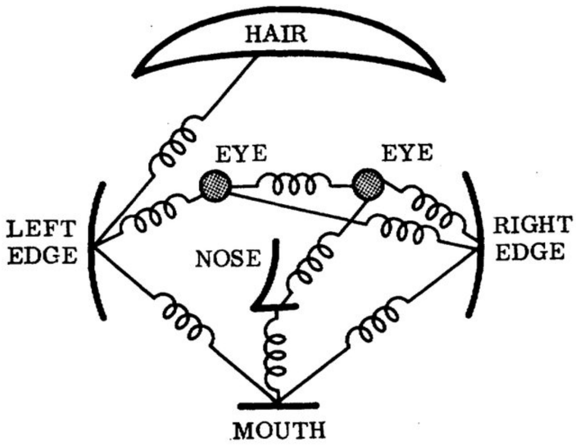


FIGURE 2.3: Pictorial structure model [60]

b. Discriminative / hybrid approaches

One of the most prominent contributions in 2D Human Pose Estimation was the hybrid approach where the human body is modelled as a flexible blend of elements [63]. Nonetheless, with recent advances in deep learning techniques since 2012, powerful discriminative Convolutional Neural Network (CNN) models proposed by (Toshev & Szegedy, 2014) [64], stacked hourglass networks by (Newell et al, 2016) [65], and OpenPose by (Cao et al, 2019) [66], have outperformed all previous approaches based on hand- engineered feature extraction. The majority of them are based on the notion of regressing heatmaps, which are a Gaussian distribution of joint positions in pixel space. Convolutional Pose Machines [67] and Stacked Hourglass methods were the first to employ multiple stacked networks with intermediate losses to address the issue of vanishing gradients. This allowed the networks to comprehend circumstances involving numerous instances of self and external occlusion.

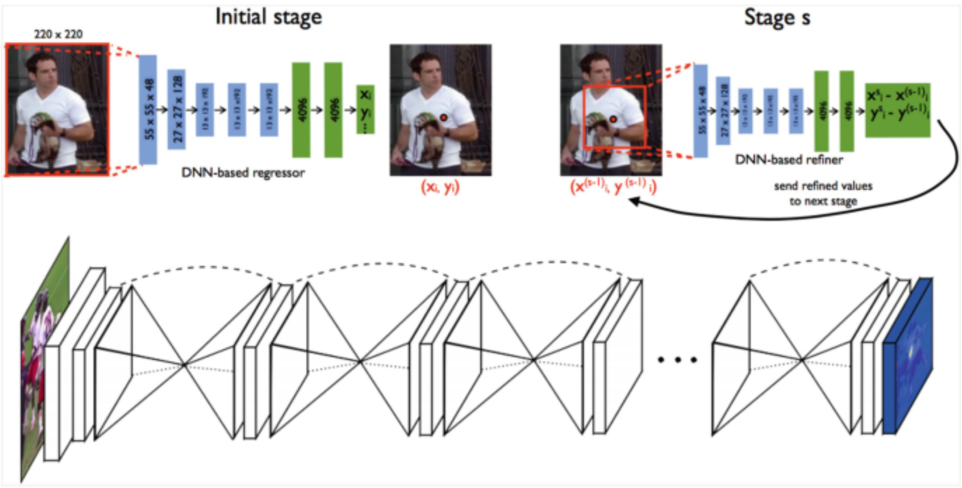


FIGURE 2.4: CNN architecture for 2D pose estimation (deep pose architecture is on top, stacked hourglass in bottom) [68]

2.4.2 The 3D human pose estimation

Since the Human 3.6M dataset [69] was published, a number of techniques as mentioned in [70], attempted to generate 3D posture estimates directly from monocular input. Two-stage methods such as [70] estimate 2D (pixel-space) poses from monocular images by first estimating 2D poses, secondly, 3D joints in a camera-relative frame are extracted from 2D pixels via constrained deep regression or by matching 2D poses with 2D projections produced from a library of existing 3D poses.

In order to predict a vector expressing 3D skeleton coordinates, a team published [71] a learnable fusion of network parameters predicting 2D joint locations and extracting 3D signals from the image. The work provided in [72] solves this challenge in a geometry-aware manner by dividing the method into two parts. An autoencoder first discovers a latent 3D representation of the person. The reconstructed 3D human is then utilised to learn 3D poses in a supervised environment. Unfortunately, all methods for estimating 3D joint coordinates from photos employ a root-relative (torso-relative) coordinate frame and cannot be applied to real-world 3D settings. Hence, all of these are learning the inherent properties of the camera and are not applicable to real-world applications.

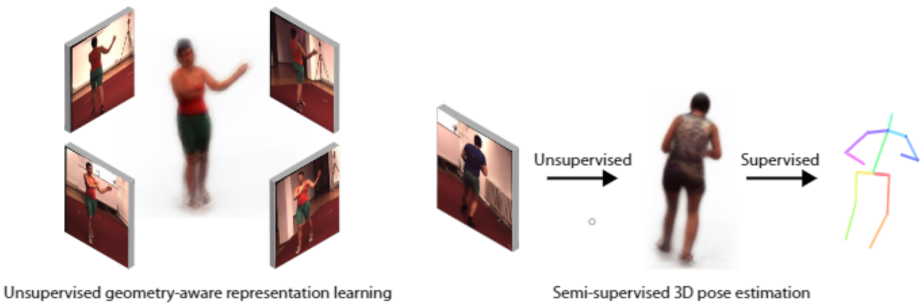


FIGURE 2.5: 3D pose estimation by geometry aware semi-supervised [72]

2.5 Gesture recognition

The automatic identification and analysis of human movement using visual input is an extensively researched field in computer vision due to both its intriguing scientific contribution and practical significance. Among the applications that could benefit from this technology, we can clearly distinguish control, navigation, and manipulation in both real and virtual environments, clinical diagnosis and monitoring, system engineering, multimodal human-computer interaction; and the serious games field. Gesture recognition can be defined as the process of extracting significant insight from human movements [68]. Research in gesture recognition has experienced explosive growth since 1990, when computers became capable of interacting in real-time and processing video streaming.In addition, it permits communication across great distances and in loud environments, which is particularly notable in a manufacturing environment because there are several noise sources that are difficult to eradicate and interfere with spoken communication. Thus, the gesture may play a crucial function in HRC and HRI in the industrial sector.

One way to classify gesture recognition involves considering the specific task to be performed, which can be either isolated or continuous. the first task involves classifying a pre-identified a gesture's starting and ending point, while the second task requires both detecting the starting and ending points of a gesture and classifying the gesture itself. Gesture recognition is a very complex process, especially when operating in the wild, and different challenges must be addressed like:

- **The temporal information encoding :** Temporal information is crucial since most gestures are dynamic, and temporal information can significantly alter the meaning of a sequence of frames [73].
- **Embedded vision-based system :** Since the system must function on embedded devices, we must employ lightweight models to accomplish this objective [74].

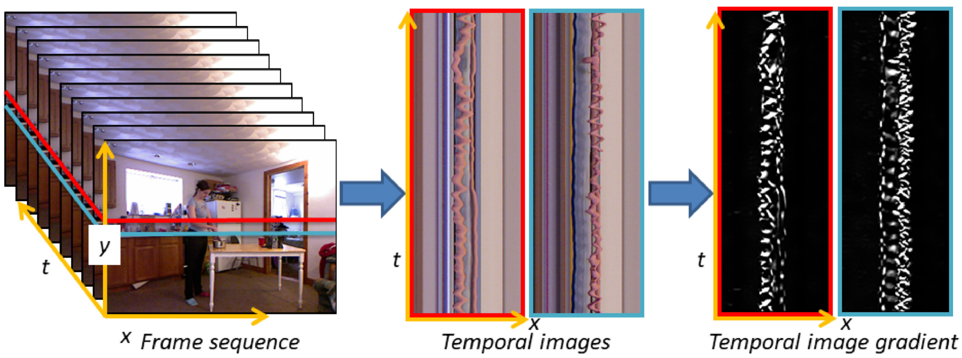


FIGURE 2.6: Temporal images representation

A vision-based system can be architected as shown in figure 2.7:

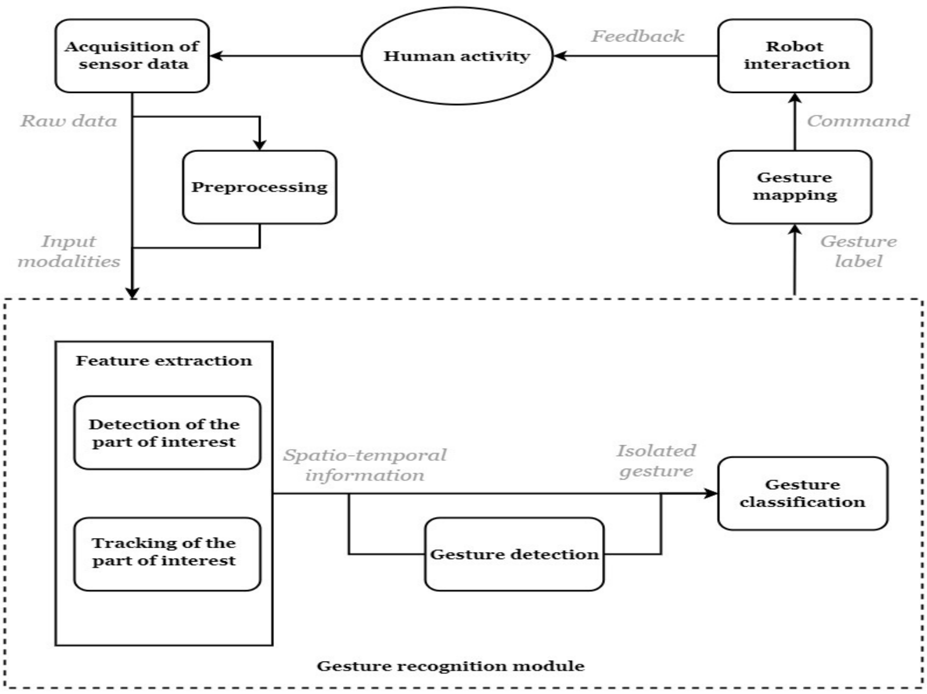


FIGURE 2.7: System workflow of gesture recognition

Figure 2.7 demonstrates that the Gesture Recognition module accepts one or more modalities as input. The modality denotes the data used as input to the module, which can operate with one or more of them (RGB data, depth data, skeleton data, and dense optical flow are the most often employed modalities). The first two modalities acquire raw data from conventional camera sensors. In contrast, the final two modalities are derived through the use of algorithms and provide more detailed information at the expense of computational expense. Normally, these modalities are not employed separately, but collectively, which necessitates combining the information from the various modalities. Fusion can occur at one or more of the following levels: data, features, and decision.
For human-robot interaction, one could have choose to use facial expressions or hand gestures recognition, but it is known that body signals are also crucial for attaining proper social interactions, since in some cases, the body can show better emotional expressions than the face [75], so they must be considered in the big picture.

2.6 Synthesis

The field of ergonomics and industrial environments has greatly benefited from the swift progress of machine learning and information technology, particularly since the establishment of the fourth industrial revolution. Thanks to the implementation of artificial intelligence technology, research in manufacturing ergonomics is shifting away from limited, population-based, and fixed analyses towards more personalised, dynamic, and comprehensive studies in real time.

In this chapter, we have explored several key concepts related to the use of artificial intelligence for the analysis of human activities under the ergonomic constraints of the industrial sector. First, we focused on the impact of artificial intelligence techniques on ergonomic risk assessment by highlighting the most relevant tools for data acquisition and discussing their different approaches. Next, in the human motion analysis section, we focused in particular on the marker-less capture system, which has gained popularity because of its ease of use and affordability. Finally, we explored human pose estimation approaches as a first phase, including 2D and 3D pose estimation methods. As we reviewed the advantages and disadvantages of each one of them, we also highlighted some of the limitations and challenges that researchers encounter when utilising them. In the second phase, we examined gesture recognition, which is a rapidly growing area of research that has numerous potential uses in industry 4.0, especially in augmented and virtual reality and human-robot interaction.

As a conclusion, this chapter provides a high-level overview of several key concepts in the broad application of artificial intelligence for human activity analysis. These tools and approaches can help to improve worker safety and health, particularly by avoiding the major issue of work-related musculoskeletal disorders and improving job quality, as well as having a positive socioeconomic impact.

CHAPTER 3

METHODOLOGY

3.1 Introduction

Our project aims to create an artificial intelligence that analyses human activity during work in order to prevent work-related illnesses by quantifying the ergonomic risk. In this chapter, we will first present the results of an investigation based on data concerning musculoskeletal disorders in a sample of the industry. Then, we will analyse the spread of MSDs among workers to estimate damage to the biomechanical system, in order to determine which ergonomic evaluation tool to automate. Next, we will specify all the necessary techniques, methods, and resources to create a new ergonomic evaluation system linked to observation methods and based on AI, as well as the development of another system dedicated to gesture recognition to naturalize human-robot interaction in an intuitive manner.
The development methodology is divided into three main sections. First, the "data preparation" section involves collecting and processing all the necessary data for result extraction, system training, and validation. Second, the second section is dedicated to constructing the neural network architecture that relies on the Pytorch framework for classification. Finally, the third section concerns the estimation of the posture score based on REBA and its validation.
In this chapter, we will detail each methodology step by step.

3.2 Ergonomic risk assessment tool selection

The observational methods mentioned earlier are widely used in general, but the manual task of conducting the assessment still has a lack of effectiveness due to multiple reasons like time consuming, subjective and biased judgment, assessment fitting to the task, etc. According to (Hita-Gutiérrez et al, 2020) [76] , the tools of

MSD assessment ca vary depending on multiple factors like the country, the culture, the companies' environment and how they are carrying them out, the working conditions and the acceptance of getting involved in the assessment itself, because of the purpose of ethnomethodological, the people tend to perform actions that are habitual and customary for them, but this can be changed under some circumstances [77]. In this work, using some data collected from 140 individual at SPA-ALZINC Ghazaouet [78] during an internship. The sample is composed of 31 female and 109 males, and the questions pertained to physical and work features, as well as pain experiences.
The major findings from an analysis job on the data are:

Workers who perform repetitive movement			Workers who exposed to Awkward postures			Workers who sustained to strained and staticposture		
LBP	UBP	LLP	LBP	UBP	LLP	LBP	UBP	LLP
100%	73%	50%	100%	82%	68%	41%	90,625%	86%

TABLEAU 3.1: Results of different pain feeling among workers. (LBP: lower back pain; UBP: upper back pain; LLP: lower limb pain)

Analysing those result, we can clearly notice that the low back pain is strongly correlated with the repetitive movement and the awkward postures, where the static posture causes the upper back pain as well as the lower limb pain, due to stress exerted on it. Based on that, we will adopt the REBA method for the automation process on account of the information mentioned in the literature review that is consistent with the insight extracted from the interviews with the workers as well as the data collected. Furthermore, the specialized physician's confirmation, after consultation, of the effectiveness of an AI-powered centralised system that evaluates workers' movements in real-time, especially in workplaces that require significant exertion, all to preserve their health by avoiding any MSD-causing factors.

3.3 Data preparation

This study uses 3D human pose estimation in video with temporal convolution and semi-supervised training [79], extending the Human 3.6M dataset [69], which is a large-scale dataset for 3D human pose estimation and one of the most used datasets for benchmarking 3D pose estimation methods. The dataset contains over 3.6 million 3D human poses captured in a controlled environment using a motion capture system composed of 4 cameras placed at different angles around the subject. Each camera performs 15 different actions (directions, discussion, eating, greeting, phoning, photo, posing, purchases, sitting, sitting down, smoking, waiting, walking, walk dog, and walk together.), resulting in a total of 150 sequences, where each sequence has a resolution of 1280 x 1024 pixels at a frame rate of 50 Hz (frames per second). The Human 3.6M dataset provides data in both parametrizations: the relative 3D joint position (R3DJP) and kinematic representation (KR), with a full skeleton containing the same number of joints (32) in both cases. The dataset can be found on the official website Human 3.6M Dataset and available for research and non-commercial purpose as mentioned in the Terms and conditions.

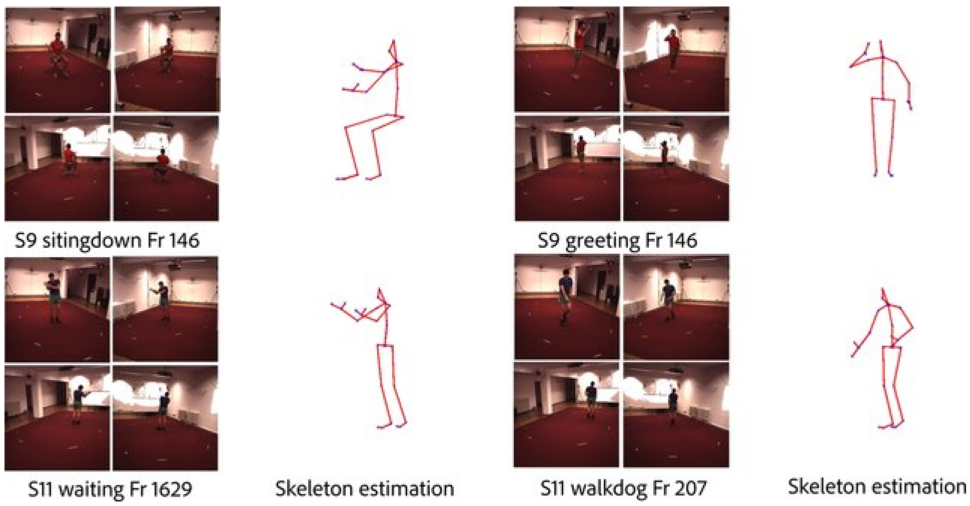


FIGURE 3.1: Representative visual results for pose estimation on Human 3.6M across four test sequences [80]

This study utilised images that had a visibility flag of either 1 or 2, and 17 key points were identified using the REBA notation for whole-body postural analysis. These key points included the nose, right and left eyes, shoulders, elbows, wrists, trunks, knees, ankles, and knuckles. A supervised model was trained using these images and key points. Figure 3.2 displays an image with the relevant body joints labelled.

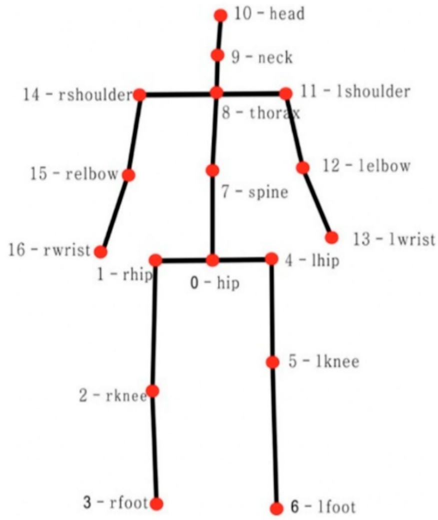


FIGURE 3.2: Full body annotated Keypoints

Moreover, this study also utilised a pre-trained neural network, specifically the keypoint rcnn R101 FPN 3x yaml, in conjunction with Detectron2 [81] to detect keypoints for 3D pose estimation. The keypoint rcnn R101 FPN 3x.yaml is an advanced neural network model that is employed for keypoint detection in images. This model is founded on a ResNet-101 backbone [82] with a feature pyramid network (FPN) and trained on the COCO dataset [83], which comprises over 330,000 images with more than 2.5 million object instances labelled with keypoints. The FPN facilitates the network to produce feature maps at various scales, which enhances the accuracy of keypoint detection. Throughout training, the model learns to forecast the positions of keypoints on the human body, such as the wrists, elbows, shoulders, hips, knees, and ankles. These keypoints can be employed to estimate the 3D pose of a person in an image or video. To enhance the accuracy of keypoint detection, data augmentation methods were employed during training. This involved rotating images at random angles and flipping them along their mid-vertical axis, which generated new training data and improved the generalisation capabilities of the model.

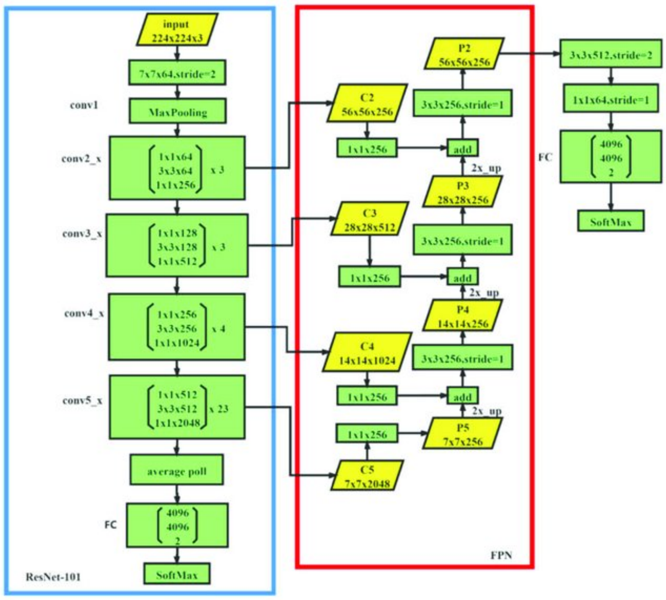


FIGURE 3.3: Final network structure of the ResNet101 + feature pyramid network (FPN) [84]

3.4 Network architecture

3.4.1 Temporal dilated convolutional model

The temporal convolution network is a new CNN-based model proposed by (Bai et al, 2018) [85] in the aim of adopting CNN for sequential modelling. In our study, the model we've adopted utilises fully convolutional architecture with residual connections to transform a 2D poses sequences as input through temporal convolution.

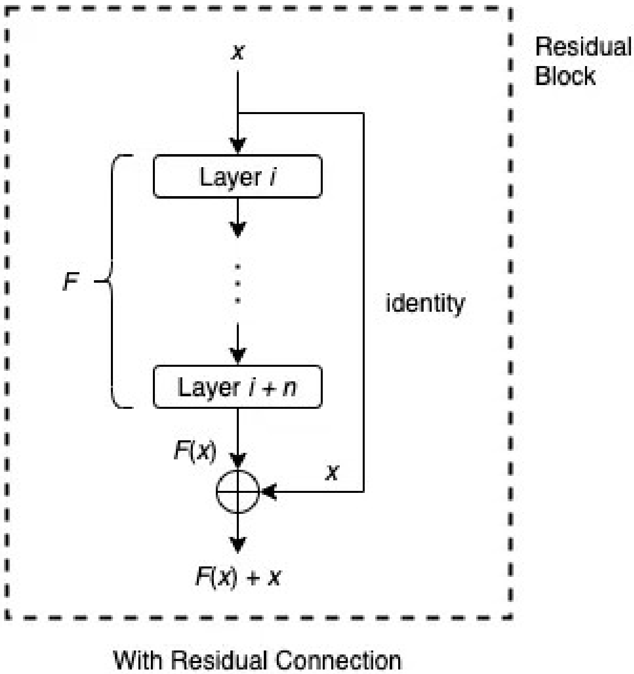


FIGURE 3.4: Residual bloc [86]

Convolutional models possess a key attribute in that they facilitate parallel processing across both the batch and time dimensions. In addition to the constant length of the gradient path connecting the input and output, regardless of the sequence length, our study reveals that a convolutional design effectively regulates the temporal receptive field, thereby conferring an edge in capturing temporal dependencies in 3D pose estimation. Furthermore, dilated convolutions are employed in order to capture extended temporal dependencies while utilising solely the lower layers within the 1D fully connected convolution network.
The incorporation of a dilation factor, denoted as l, within convolutional filters results in the establishment of a uniform spacing between neighbouring filter taps. This, in turn, leads to an increase in the scope of inputs encompassed by the nodes present in the hidden layers, thereby enabling the effective representation of extended historical and dependency relationships. A dilation factor of 1 results in a standard convolution, while in the other hand, longer dilation factors enable the output at the highest layer to encompass a wider range of inputs, thereby broadcasting the

receptive field of the convolutional network, this can be shown in this equation:

$$(F *_l K)(p) = \sum_{s+lt=p} F(s)K(t), \tag{3.4.1}$$

where $*_l$ is a dilated convolution, F(s) the input, k(t) the kernel and (F $*_l$ K)(p) is the output. The figure 3.5 illustrates an example of dilation factor of l=1 (left), l=2 (middle) and l=4 right.

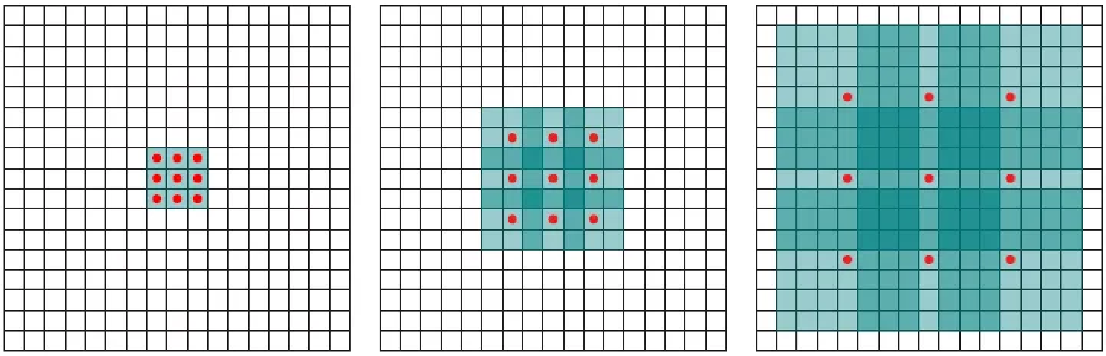


FIGURE 3.5: Example of dilation with multiple factor

3.4.2 Final architecture

Our final model for fully-convolutional 3D pose estimation comprises a series of interconnected layers and blocks. Specifically, the architecture is designed for a receptive field size of 243 frames, where B = 4 blocks and J = 17 joints. In this architecture, the convolutional layers are represented in green, where 2J, 3d1, 1024 denotes 2J input channels, kernels of size 3 with dilation 1, and 1024 output channels. For a sample 1-frame prediction, the tensor sizes are shown in parentheses, where (243, 34) represents 243 frames and 34 channels. The input layer takes the concatenated (x, y) coordinates of the J joints for each frame and applies a temporal convolution with kernel size W and C output channels. The input layer is followed by B ResNet-style blocks, which are surrounded by a skip-connection. Each block performs a 1D convolution with kernel size W and dilation factor D = W^B, followed by a convolution with kernel size 1. We apply batch normalisation, ReLU (rectified linear units) and dropout after each convolution, except the very last layer. The receptive field increases exponentially by a factor of W in each block, while the number of parameters increases linearly.

We set the filter hyperparameters W and D so that the receptive field for any output frame forms a tree that covers all input frames. The last layer of our model outputs a prediction of the 3D poses for all frames in the input sequence using both past and future data to exploit temporal information. Additionally, we explore the use of causal convolutions to evaluate real-time scenarios, where convolutions only have access to past frames [79]. Unlike convolutional image models that apply zero-padding to obtain as many outputs as inputs, we observed better results with unpadded convolutions. Instead, we padded the input sequence with replicas of the boundary frames to the left and right. Figure 4.6 illustrates an instantiation of our architecture for a receptive field size of 243 frames with B = 4 blocks. We set W = 3 with C = 1024 output channels for convolutional layers and use a dropout rate of p = 0.25.

Our proposed architecture demonstrates promising results for 3D pose estimation, especially for large receptive fields and the utilisation of temporal information. Nevertheless, further experiments and evaluations are necessary to assess its performance and potential applications in real-world scenarios.

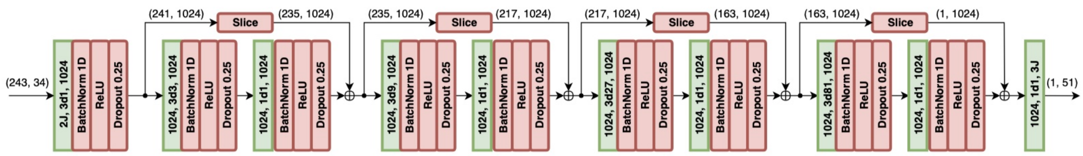


FIGURE 3.6: FCN 3D pose estimation architecture [79]

3.4.3 Training approach

In order to enhance and ensure accuracy progression in settings when there's limited labelled data available, we adopt the semi-supervised training method [79], which in the context of 3D human pose estimation is a type of machine learning that uses both labelled and unlabelled data in a co-training [87] methodology where the aim is to improve the model's prediction accuracy.

The task of estimating 3D human poses can be challenging and costly due to the difficulty in obtaining precise labels for extensive datasets. The method we are using to address this issue involving the utilisation of unannotated video data in conjunction with a readily accessible 2D keypoint detector to construct an autoencoder framework. The proposed methodology facilitates the estimation of 3D poses through the utilisation of input joint coordinates. This approach is advantageous as it is typically less complex and more cost-effective than obtaining complete ground-truth annotations. Additionally, the estimated poses can be projected back into their original 2D space. In the course of training, a loss function is employed to impose penalties on disparities between the initial input joint coordinates and those that are reconstructed by projecting backwards from the estimated poses. By adopting this approach, we facilitate our model to acquire knowledge not only on the influence of depth information on the overall posture but also on its correlation with each distinct frame. Enhancing the model's performance on tasks such as human pose estimation, which may pose difficulties or incur high costs in terms of precise labelling at a large scale, can be achieved through this approach. In general, the semi-supervised methodology presents a potentially effective resolution to address the obstacles associated with the estimation of 3D human pose.
Figure 3.7 illustrates the method mentioned above:

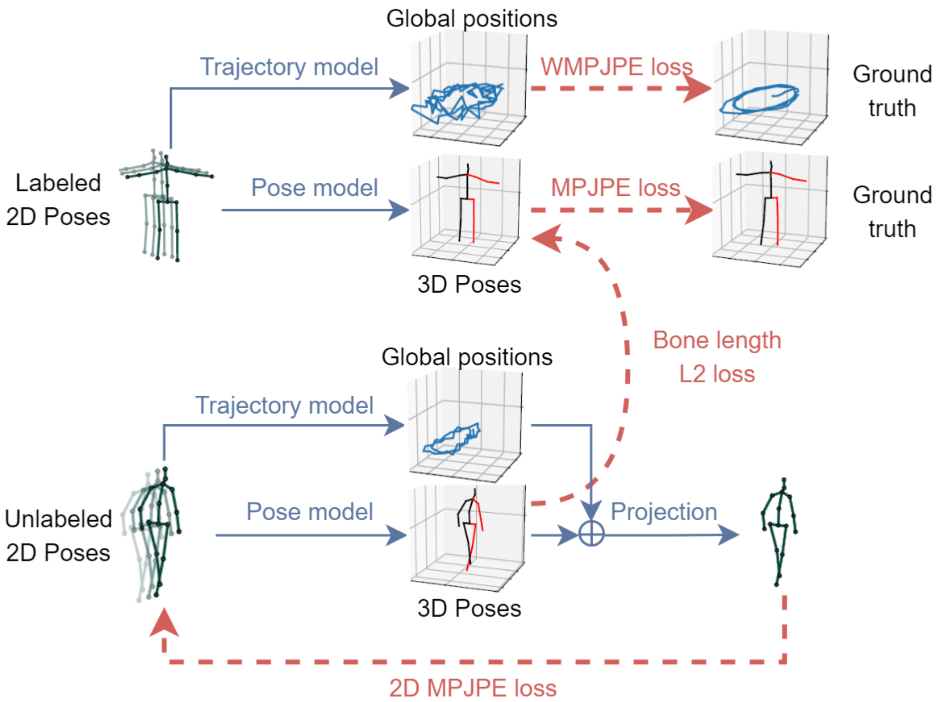


FIGURE 3.7: Semi-supervised training approach

The figure above represents our technique that involves integrating a supervised element with an unsupervised element, where this last functions as a regulator. The supervised loss is trained using ground truth 3D poses as the target for the labelled data; on the other hand, the unannotated data is utilised for the implementation of an autoencoder loss function, wherein the anticipated 3D poses are projected onto a 2D plane and subsequently assessed for coherence with the input. The 2D pose on the screen is influenced by both the trajectory and the location of all joints in relation to the root joint, due to the perspective projection. In the absence of global positioning, the subject will consistently be displayed at the centre of the screen with an unchanging level of magnification (constant scale). Consequently, a secondary network is trained with the purpose of performing regression on the individual's 3D trajectory. This trajectory is then utilised to accurately project the pose back to its 2D form. The network architecture employed for both pose and trajectory regression is identical; however, weight sharing is not implemented due to the observed adverse effects on training outcomes. The use of a weighted mean per-joint position error (WMPJPE) loss function is employed to optimise the trajectory due to the inverse relationship between the accuracy of the trajectory regression and the distance between the subject and the camera. The formula used for calculating the loss is:

$$E = \frac{1}{y_z} \|f(x) - y\|, \tag{3.4.2}$$

We give each sample a weight determined by the reciprocal of its ground-truth depth in camera space. Moreover, we do not require an accurate trajectory estimation for subjects at a distance because their 2D keypoints usually cluster together within a small region.

Our aim is to provide incentives for the prediction of plausible 3D poses, rather than mere replication of the input. In order to achieve this, we discovered that incorporating a soft constraint was a successful approach. This constraint involved roughly aligning the average bone lengths of the individuals in the unlabelled group with those of the labelled group, as demonstrated by the "Bone length L2 loss" in Figure 3.7.

3.4.4 Evaluation

Three evaluation protocols were employed in the assessment of the Human3.6M dataset:

- **Protocol #1 (MPJPE):** determines the mean per-joint position error in millimetres through the computation of the Euclidean distance between anticipated joint positions and actual joint positions.
- **Protocol #2 (P-MPJPE):** presents the error outcome subsequent to the alignment of the anticipated poses with the actual values in terms of translation, rotation, and scale.
- **Protocol #3 (N-MPJPE):** exclusively aligns the anticipated poses with the actual values in terms of scale. The aforementioned protocol has been employed in semi-supervised experiments and adheres to the methodology suggested in reference [72].

The Human3.6M dataset experiments utilised a 17-joint skeleton and a unified model was trained for all actions. The study involved conducting training on a sample of five subjects, specifically identified as S1, S5, S6, S7, and S8. Subsequently, the testing phase was carried out on two subjects, namely S9 and S11.

The final model will takes as input 2D keypoint sequences and will generate the convenient 3D pose estimation, the model will employ all the methods and techniques mentioned earlier in the sections above.

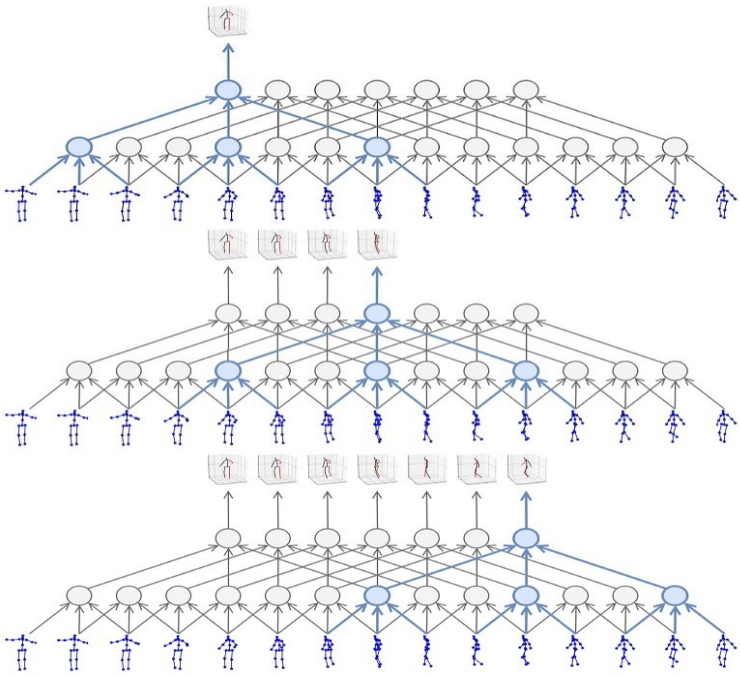


FIGURE 3.8: Temporal Convolution Model (Taking 2D keypoint sequences and outputs 3D pose estimation, from bottom to top)

3.5 REBA posture score estimation

We utilise Detectron2 [81] to obtain 2D kinematic joint locations and estimate REBA body posture scores. The model employs a combination of joint angles and force exerted on each joint. The Calculation of joint angles was performed through the utilisation of inverse kinematics. Additionally, the estimation of force was accomplished by employing deep learning algorithms that were based on joint position.

$$\theta_{ab} = \cos^{-1} \frac{\vec{a}\vec{b}}{|\vec{a}||\vec{b}|} \tag{3.5.1}$$

The equation above is used to calculate the angle made by two adjacent body segments (θ_{ab}), in the context of REBA ergonomic assessment, it is an important factor for assessing the risk of a task and identifying potential work-related musculoskeletal disorders. A and b are two vectors representing the adjacent body segments being analysed, where the vertical bars around the vectors denote their magnitude. To obtain the angle, we use the inverse cosine function ($\cos^{-1}$) of the resulting fraction and the output will be in radian convertible to degrees.

Angles are identified by the symbol θ and are labelled based on the limb involved, for example U representing the upper arm and La representing the lower arm, followed by a number indicating the type of movement. The movement types are numbered as follows: **flexion/extension** is 1, **twist** is 2, and **side-bending** is 3. **S** represents the score for a particular body part, and its subscripts identify the limb and the type of movement, based on the numerical labels described above. Some examples of REBA posture score calculations for each group are shown below, based on 2D coordinates of the body joints.

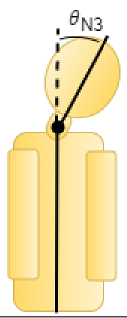
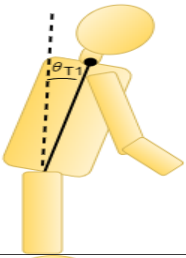
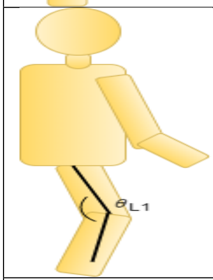
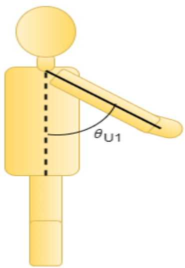
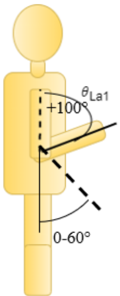
	$0° < \theta_N < 20°$: +1 $\theta_N > 20°$: +2 θ_{N1}: +2 (in extension) θ_{N2}: +1 to neck score θ_{N3}: +1 to neck score
	$\theta_T = 0°$: +1 $\theta_{T1} =$ +2 (in extension) $0° < \theta_T < 20°$: +2 $20° < \theta_T < 60°$: +3 $60° < \theta_T$: +4 θ_{T2}: +1 to trunk score θ_{T3}: +1 to trunk score
	Stand on two legs vertically: +1 Stand on one leg: +2 $30° < \theta_{Lg} < 60°$: +1 to legs score $\theta_{Lg} > 60°$: +2 to legs score
	$-20° < \theta_U < 20°$: +1 $\theta_U < -20°$: +2 $20° < \theta_U < 45°$: +2 $45° < \theta_U < 90°$: +3 $90 < \theta_U$: +4 Shoulder is raised: +1 to upper arm score Upper arm is abducted: +1 to upper arm score Person is leaning: -1 from upper arm score
	$60° < \theta_U < 100$: +1 $0 < \theta_U < 60$ or $\theta_U > 100°$: +2

TABLEAU 3.2: Example of REBA score for each group

3.6 Validation

The validity of the posture scores generated by the system has been established through their application to some commonly observed occupational workplace posture videos, (images extracted are shown in figure 3.9). The occupational health physician at SPA-ALZINC [78] conducted a comprehensive assessment of the postures, utilising the REBA worksheet to evaluate their suitability for the tasks at hand. Additionally, the joint angles were verified through the use of an online tool called ONLINE PROTRACTOR, which is a web-based measurement tool specifically designed for angles. The study adheres to [88] as a guiding principle for comprehending the proper techniques for documenting workplace postures to enhance precision and quality of analysis.

FIGURE 3.9: Example of workers' posture while performing occupational tasks

3.7 Conclusion

This chapter explores the multiple techniques and methods used to estimate the human 3D pose starting from the extracted 2D keypoints by Detectron2 and denoted by 17 points, which will be our model input structured in Section 3. Moreover, it will generate a skeleton-based output that will be used to calculate all the joint angles in order to calculate the REBA score frame by frame in real time.

CHAPTER 4

RESULTS

4.1 Introduction

This chapter presents a detailed account of all the experiments that were conducted during the course of this thesis. The first section describes the environment where our system was implemented and tested. The second section dives deeper into the results obtained from the semi-supervised training and compares it with supervised learning. And the last section, it includes 3D human pose estimation reconstruction and REBA score calculation using both weight generated.

4.2 Experimental setup

The training was conducted using the machine learning framework Pytorch [89] on Google Colaboratory [90]. Despite having access to a free GPU provided by Google Colab, along with a disk space of 78 GB and 12.7 GB of RAM, we had to purchase additional GPU computing units due to the large model size. To achieve this, we subscribed to Google Colab Pro+ for 50$, which gave us access to 500 units of A100 GPU with 40GB memory size, 1.55 GB/s bandwidth, 6912 CUDA cores, as well as 166 GB of disk space and 83.5 GB of RAM.

4.3 Trainings implementations

4.3.1 Semi-supervised training

This type of training for the 3D human pose estimation model was carried out with a vast range of hyperparameters. These included the selection of the dataset of interest, which was Human3.6M [69]. For the 2D keyoints within the ground truth, we used the ***cascade pyramid network*** fine-tuned on the 2D projection of Human 3.6M with two separated feature extractors, indexed ***dbb*** (Double Blackbone). The unlabelled subjects for the training are S2, S3, and S4, including all the actions. On the other side, the labelled subjects are S1, S5, S6, S7, and S8, while S9 and S11 were reserved for testing at the end of each epoch and in the final evaluation.

For the purpose of enhancing the model's accuracy, data-augmentation and test-time augmentation techniques were employed during the training, in which the first technique flipped poses horizontally to double the training dataset, while the second one also flipped the poses horizontally when testing the model. The model architecture was based on a ***(3 × 3 × 3)*** (receptive field of 27 frames) fully convolutional model. Due to some internet connection constraints (a larger model means less guarantee of finishing the training). where the first 3 is 3×1 convolutions in the foist layer, followed by two residual blocs with 3×1 convolutions.
Overall, all the hyperparameters and training arguments are shown in the table below:

Hyperparameters	**Value**	**Hyperparameters**	**Value**
Dataset	h36m	Warmup	1
Keypoints	cpn_ft_h36m_dbb	No evaluation during training	False
Subjects for training	S1, S5, S6, S7, S8	Dense connections	False
Subjects for testing	S9, S11	Disable optimisations	False
Actions	* (all actions)	Linear projection	False
Optimizer	Amsgrad[91]	Bone length term	True
Stride	1	No projection	False
Epochs	60	Visualization subject	None

Batch size	1024	Visualization action	None
Dropout	0.25	Visualization camera	0
Learning rate	0.001	Visualization video	None
Learning rate decay	0.95	Visualization skip	0
Data augmentation	True	Visualization output	None
Test-time augmentation	True	Visualization export	None
Architecture	3*3*3	Visualization bitrate	3000
Channels	1024	Visualization without ground truth	False
Subset	1	Visualization limit	-1
Downsample	1	Visualization downsample	1

TABLEAU 4.1: The semi-supervised training hyperparameters and arguments

4.3.2 Supervised training

On the contrary to semi-supervised training, supervised training means that we use the entire dataset as labelled, which includes all subjects from S1 to S8 for training the temporal dilated convolutional model. However, subjects 9 and 11 will be reserved for testing. The architecture for this training type has the same hyperparameters and arguments, as we will compare the model's performance using the three protocols mentioned earlier in Chapter 3.

4.4 Results

4.4.1 Reconstruction error on Human3.6M

a. Semi-supervised case

The table presents the results obtained from our convolutional model with $B = 2$ residual blocs and 27 input frames as receptive field:

Protocol / Action	MPJPE (mm)	P-MPJPE (mm)	N-MPJPE (mm)
Directions	44.5	34.9	42.67
Photo	58.42	44.53	56.51
Discussion	49.97	37.55	46.75
Eating	44.47	35.39	42.46
WalkDog	51.25	40.87	49.26
Purchases	44.36	34.36	42.51
Posing	42.29	36.33	45.08
Walking	35.37	28.2	34.37
Greeting	47.45	39.16	46.3
Phoning	51.52	39.79	49.06
Waiting	45.47	35.38	44.33
Sitting	57.67	46.31	55.17
Smoking	49.37	36.69	47.39
WalkTogether	38.65	31.93	36.69
SittingDown	65.6	52.69	62.67

TABLEAU 4.2: Semi-supervised results

After the results were obtained, we calculate the overall action-wise averages :

P1 average	P2 average	P3 average
48.6 mm	38.5 mm	46.8 mm

TABLEAU 4.3: Action-wise average with three protocols

b. Analysis

With a velocity error of 3.23 mm per joint, the obtained results suggest that the performance of the 3D human pose estimation model using semi-supervised learning is relatively decent. Looking at the results by action, we can see that the model performs better on certain actions than on others. Under protocol #1, the "Walk" action has the lowest error at 35.37 mm, whereas "SitD." has the maximum error at 65.6 mm. This indicates that the model may have difficulty estimating poses for certain actions, potentially due to the complexity of the poses involved or the variation in how individuals perform the action. Protocol #2 (P-MPJPE) has the smallest defect at 38.5 mm, followed by Protocol #3 (N-MPJPE) at 46.8 mm and Protocol #1 (MPJPE) at 48.6 mm. This indicates that using the predicted 3D pose as an input to a motion predictor can enhance the estimated pose's accuracy. As a whole, these results indicate that the semi-supervised learning approach used to train the 3D human pose estimation model is effective; however, there is still room for improvement, particularly for certain actions where the model's performance is less accurate, particularly when sub-optimal viewings conditions are considered.

c. Supervised case

Protocol / Action	MPJPE (mm)	P-MPJPE (mm)	N-MPJPE (mm)
Directions	46.42	35.02	43.67
Photo	57.63	43.66	55.41
Discussion	48.26	37.35	46.83
Eating	44.76	35.50	43.24
WalkDog	51.16	40.01	48.84
Purchases	45.02	34.39	43.29
Posing	46.79	35.72	45.28
Walking	36.02	27.58	34.40
Greeting	47.81	38.27	46.39
Phoning	50.63	38.38	48.26
Waiting	45.65	35.04	44.32
Sitting	57.34	45.64	55.29
Smoking	48.98	39.03	47.23
WalkTogether	38.78	31.49	36.64
SittingDown	55.03	42.53	52.58

TABLEAU 4.4: Supervised results

After the results were obtained, we calculate the overall action-wise averages :

P1 average	P2 average	P3 average
49.7 mm	38.9 mm	49.8 mm

TABLEAU 4.5: Action-wise average with three protocols

d. Analysis

In this attempt at training, we obtained a velocity error per-joint of 4.13 mm. While this value is relatively high under some application constraints, it is still acceptable to say that the model performed well during training. The results analysis reveals that the model performs better for actions that involve dynamic movement like walking, with lower MPJPE values of 36.02 mm and 27.58 mm P-MPJPE. On the other hand, we can clearly see a high error for actions that involve exposure to the camera from sagittal and transverse planes, such as taking a photo and sitting, suggesting that the model has difficulty accurately extracting 2D keypoints from such poses. This is also confirmed by the higher P-MPJPE values. Furthermore, the model's scale estimation accuracy needs improvement, as indicated by the higher N-MPJPE values, which confirms the issue with image planes.

4.4.2 Model's performance

a. Semi-supervised case

Figure 4.1 represents the mean distance between the predicted trajectories and ground truth over 60 epochs. The plot exhibits three distinct curves. The first curve, depicted in dashed blue, signifies the average distance between the predicted and ground truth trajectories for the training set, and evaluated on the labeled data. The second curve, illustrated in solid blue, represents the same assessment on the unlabeled data. Lastly, the third curve, portrayed in solid orange, denotes the same evaluation on the validation set.

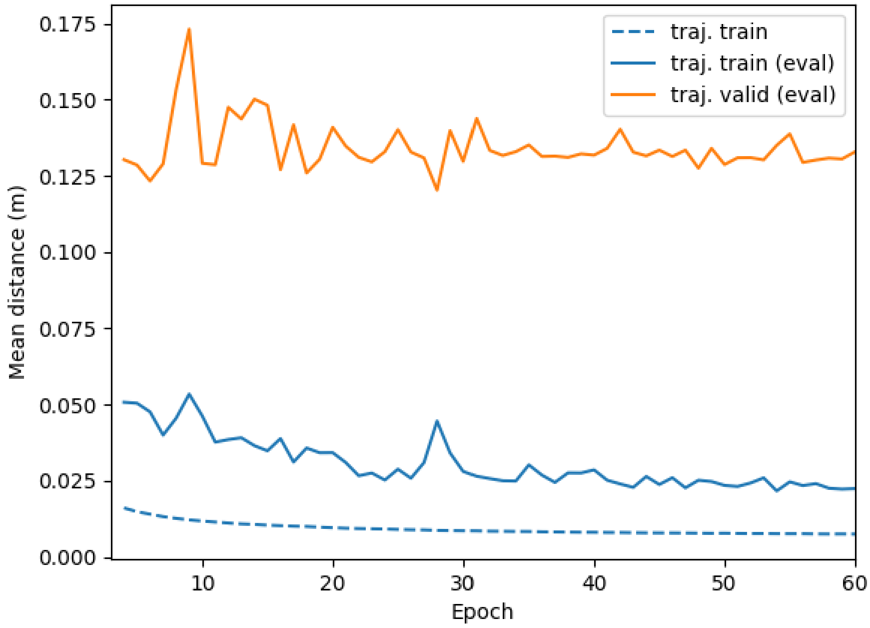


FIGURE 4.1: Mean distance between the predicted trajectories and ground truth

Figure 4.2 represents mean per-joint position error (MPJPE) in 2D. The plot displays four distinct curves. The first curve, depicted in blue, represents the MPJPE on the labeled training set, which is evaluated on the labeled data. The second curve, illustrated in dashed orange, represents the MPJPE on the unlabeled training set, evaluated on the labeled data. The third curve, represented by solid orange, also represents the evaluation on the unlabeled data. Finally, the fourth curve, depicted in green, represents the evaluation on the validation set.

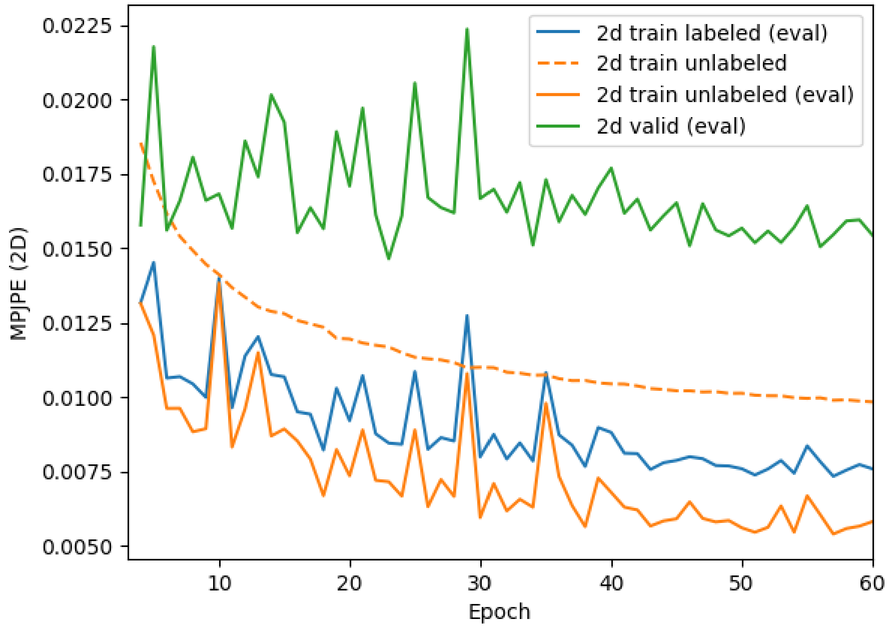


FIGURE 4.2: Mean per-joint position error in 2D

Figure 4.3 is representing the 3D loss between the predicted poses and the the ground truth 3D poses for the training set, as well as the evaluation training set and the validation set. The plot comprises of three distinct lines, each delineating a unique set. The blue dashed line depicts the MPJPE between the predicted and ground truth poses for the training set, while the blue solid line represents the MPJPE for the evaluated training set. The orange line, on the other hand, represents the MPJPE for the validation set.

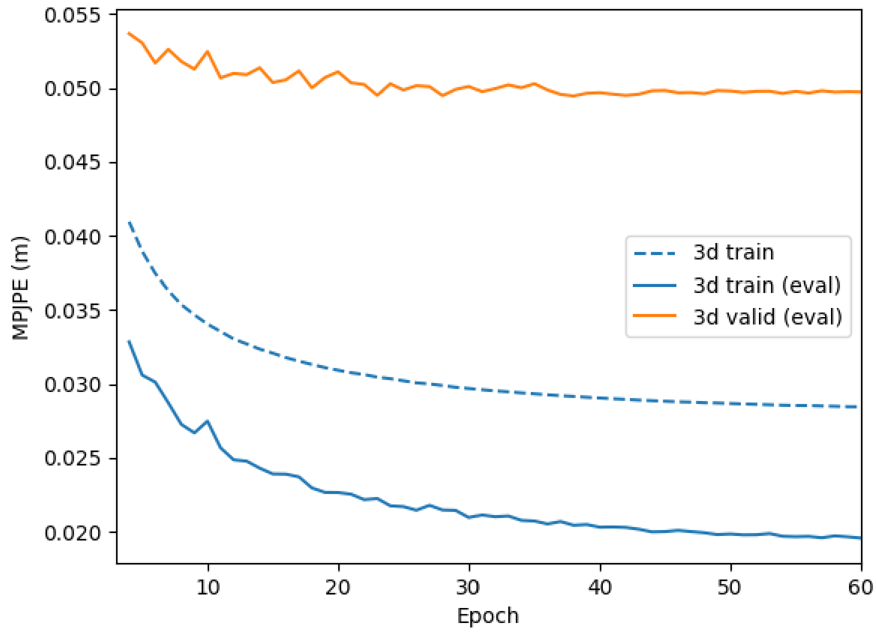


FIGURE 4.3: Mean per-joint position error in 3D

b. Analysis

In the mean distance representative figure (Figure 4.1), the training curves in blue are indicating that the model is improving over time, with the training loss decreasing quickly to reach 0,0058 m and the training validation reaching 0,026 m in around 50 epochs. The validation curve is significantly not improving with time due to some overfitting because there are only two unlabelled subjects unlabelled, which makes it relatively small compared to labelled subjects, which makes the model fit the labelled data too closely. In the second figure (Figure 4.2), the model did well even though the validation on 2D faced some overfitting troubles, but the data-augmentation regularisation worked well in that scenario, which means the back-projection to 2D worked correctly after regressing the 3D trajectory sample. The last result can be clearly seen in the 3D loss curves (Figure 4.3); we have stable model performance during the training as well as the validation, which cannot be considered overfitting because the threshold didn't get surpassed and the model is generalising well.

c. Supervised case

Fiure 4.4 presents the MPJPE between the predicted 3D poses and the ground truth poses for the training set, as well as the evaluation of the training set and the validation set. The figure has 3 curves, of which the blue dashed one represents the MPJPE error on the training set, the solid blue one illustrates the MPJPE error on the evaluated training set, and the orange one shows the MPJPE error on the validation set.

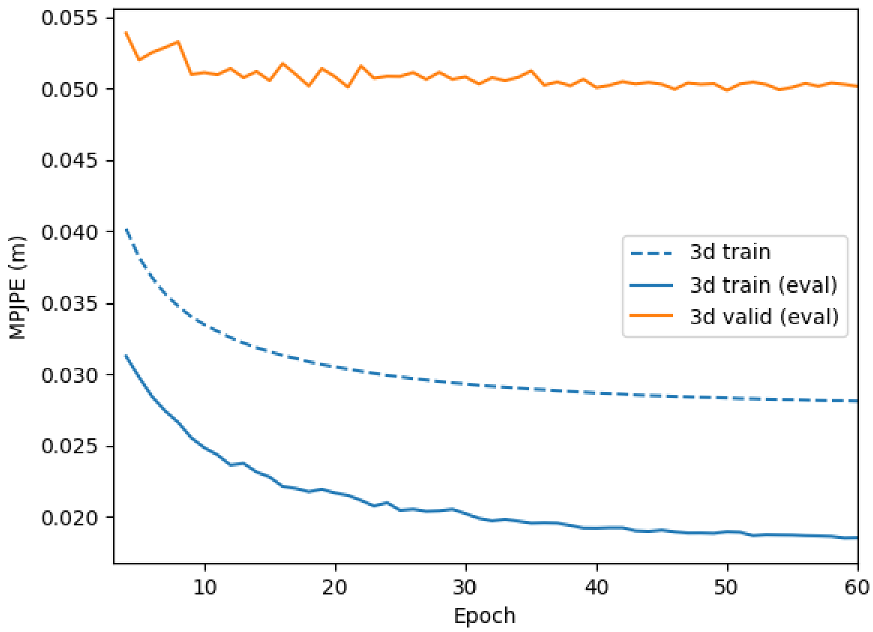


FIGURE 4.4: Mean per-joint position error in 3D

d. Analysis

The plot illustrates a consistent and steady decrease in the 3D train and 3D train (eval) curves over the course of the training process. This observation suggests that the model is acquiring knowledge and enhancing its performance on the training dataset as time progresses. In contrast, the 3D valid curve exhibits an alternative pattern, wherein the loss initially decreases but subsequently reaches a plateau or experiences a slight increase towards the conclusion of the training process. This implies that the model is exhibiting overfitting tendencies towards the training data and didn't make any effort from the 9$^{\text{th}}$ epoch, thereby indicating a lack of generalisation capability towards the validation data. However, the lack of improvement in the 3D valid curve suggests that the model may not be capturing important patterns or relationships in the validation set.

e. Discussion

Upon conducting experiments on semi-supervised and supervised training, an analysis of the results was performed, leading to the conclusion that the semi-supervised approach provides a superior estimation of the 3D human pose, as measured by the MPJPE protocol. The results of our analysis indicate that the semi-supervised approach exhibited a comparatively lower MPJPE value on the validation set as opposed to the supervised approach. The aforementioned observation implies that the semi-supervised methodology exhibits superior capacity for generalisation to novel data and for generating precise estimations of 3D human pose.
Furthermore, our discovery is congruent with prior studies that have exhibited the efficacy of semi-supervised learning in enhancing the proficiency of diverse computer vision assignments. It is noteworthy that the semi-supervised methodology necessitates a smaller quantity of labelled data in comparison to the supervised methodology. This is especially advantageous in situations where the acquisition of significant amounts of labelled data is arduous or costly.

4.5 Inference

In this section, we'll assess the effectiveness of the 3D human pose estimation model obtained from semi-supervised training. To achieve this, we'll test the model on arbitrary videos from different sources including manufacturing and analyse the predictions and reconstruction quality from multiple common views. Additionally, we'll calculate the REBA score in real-time on the output to evaluate the ergonomic risk level of the estimated pose. This evaluation will help us determine the accuracy and usefulness of the model in real-world scenarios and assess its potential for

practical applications.
The inference process will follow the steps illustrated in figure 4.5:

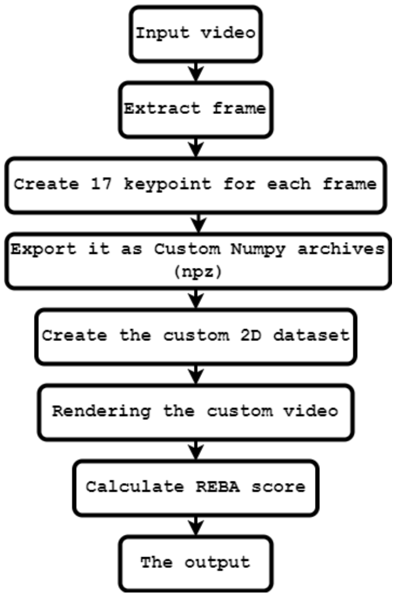


FIGURE 4.5: Inference process

The model testing can consider multiple conditions; in our case we focus on complex postures, sagittal plane views and noises. The reconstruction will display the right arm and the right foot in red and the rest of body in black skeleton based model.

a. Sample 01 (complex posture)

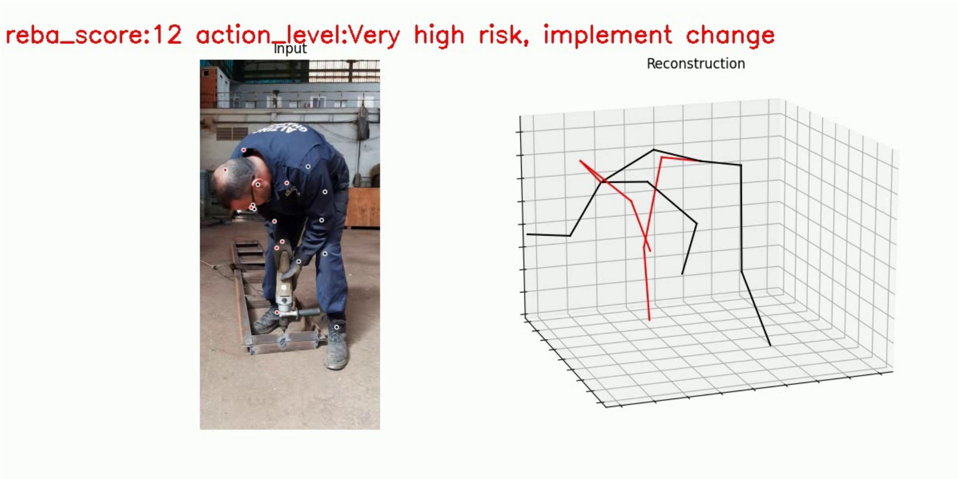


FIGURE 4.6: Sample 01 output

The model did well in mapping the 2D pose to 3D, even with the right arm and half of the right leg are hidden, the model has change the camera view while rendering to output the maximum accurate skeleton based 3D pose. The REBA score for this posture is 100 % correct after confirming the calculation manually using the REBA worksheet, as we can clearly see that the risk of having a work-related musculoskeletal disorder is incredibly high.

b. Sample 02 (complex posture)

FIGURE 4.7: Sample 02 output

In this capture, using 16 2D keypoints only, the model did better than the sample 01 in the rendering as we can see a perfect obvious skeleton based 3D pose. the REBA score is 92 % correct (manually, the score is 13), however, the action level stills the same and the posture needs to be change immediately.

c. Sample 03 (Sagittal plane)

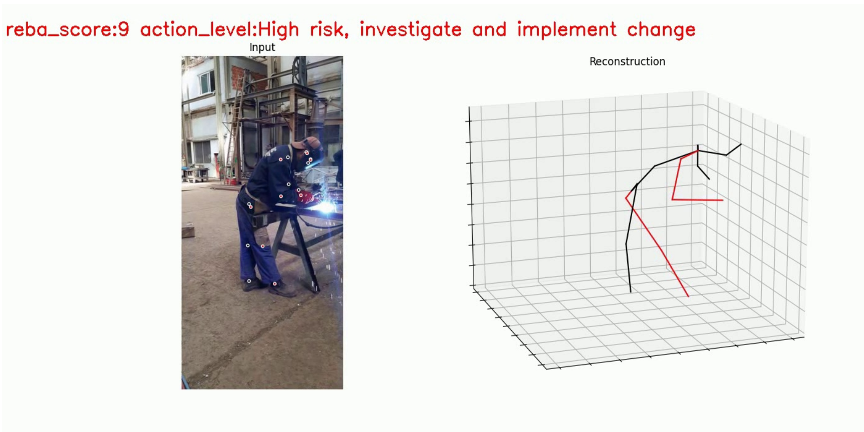


FIGURE 4.8: Sample 03 output

In this sample, the video capture was from sagittal view at 11 meters distance, we can clearly see some reconstruction error, there was a compromise between the right and left leg position, the model started a false prediction then rectified it quickly, this error is due to worker's pants unified color so the model has some trouble differentiating the right from the left. For the REBA score, the model calculated it 100% correctly.

d. Sample 04

FIGURE 4.9: Sample 04 output

In this sample, the model did well predicting the 3D pose, even with the capture angle is almost from the worker's back, the model didn't change the camera view to mirror view to get an more accurate reconstruction, unfortunately it wasn't fitting the 3D space, due to input video resolutions which were bigger than that of Human3.6M.

4.6 Prototyping

Our project is unique in the occupational field, as it contributes in many ways to the safety of workers while performing their tasks. This study is the first to estimate the body posture REBA score from an input video in real-time by implementing

Deep learning algorithm to automate the ergonomic risk assessment based on the observation. One of our project's outputs is the interpreted REBA score besides the action must be taken. Using such method, not only in the industrial sector, but even the healthcare sector occupants will be able to interpret patients movements with high accuracy, the thing that will reduce evaluation time and eliminate any human error. Additionally, our system which will be prototyped in its Beta version as a friendly-user web application allowing multiple digitisation services. While launching the web application, the interface shown in figure 4.10 will appear.

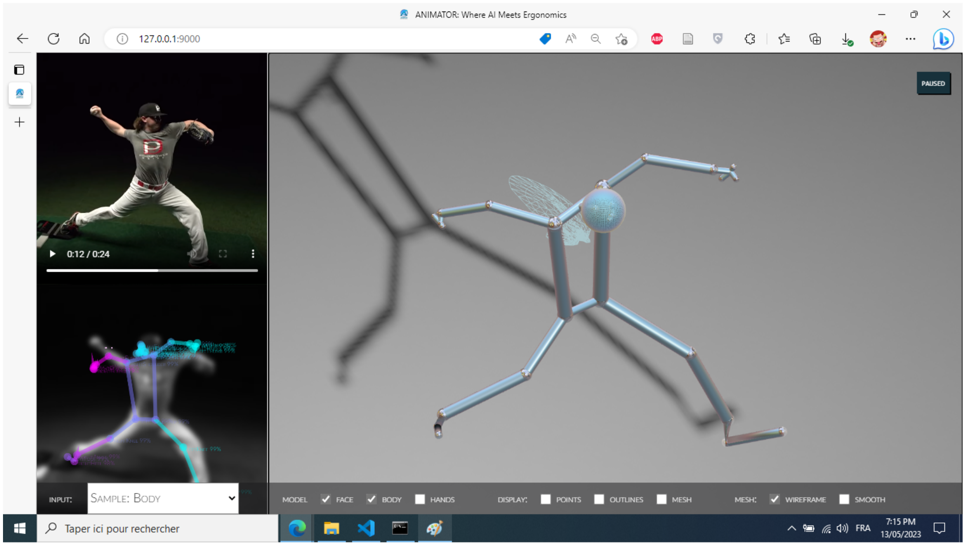


FIGURE 4.10: ANIMATOR user interface

In the left side, we find the input visualisation with both RGB and Gray scale besides the 2D estimated pose; the main content section is the rendering output and the control bar is situated in the bottom composed of:

- **Input:** a drop-down menu for choosing the file to process, we can find live stream, uploaded video and some test sample within the web application
- **Model:** a checkbox for choosing the pose estimation model whether it's face, body or hand.
- **Display:** a checkbox for choosing the output display type whether it's points, outlines or mesh.
- **Mesh:** a checkbox for choosing the mesh displaying type between wireframe or smooth.

4.7 General discussion

This study has demonstrated the effectiveness of semi-supervised learning in leveraging unlabelled video data to estimate human 3D pose, aided by temporal dilated convolution that is invariant to scale and visual noise. Furthermore, we have successfully utilised the model output to calculate the REBA score with an accuracy of over 92%. The model achieved a score of 48.6 mm mean per joint position error (MPJPE) between the predicted and ground truth, indicating good results, particularly considering the lack of reliance on any additional dataset. However, a 48.6 mm error still represents a relatively small average prediction error. As mentioned earlier, Overfitting occurs when a deep neural network fits the training data too closely, rendering it unable to generalise to new input data, which could potentially affect its performance. The architecture employed in the training demonstrated a mean absolute error (MAE) of 14 mm between the training and validation of unlabelled 2D data, suggesting a low likelihood of deep neural network overfitting. Overall, considering the architecture's configuration, training environment, and constraints, we have achieved a satisfactory outcome.

4.8 Conclusion

In this chapter, we presented the results obtained from training the temporal dilated convolutional network using both supervised and semi-supervised approaches. The results showed that the semi-supervised approach was more effective in generating a model for inferring input videos and obtaining real-time 3D pose estimations, along with calculating the REBA score. Therefore, we used the model obtained through the semi-supervised approach to carry out our inference tasks.

General conclusion

This study discussed a Deep Learning based technique to estimate the 3D pose of human from videos and calculating REBA score posture in real-time for the ergonomic context. This AI-based system is designed to maintain the occupational health of workers in the industrial sector, and it aims to advocate for a major change in the professional health law to recognise the Work-related musculoskeletal disorders in the first place, in the same time raise awareness among industrial occupant about the impact of false postures and false movements first on health, second on productivity and all the economic consequences that result especially the absenteeism rate and social compensation. The ergonomic application, despite its investment charge, it's a safe and profitable investment, in which its signs will appear in both social and economic environment. This result will pave the road towards a developed industrial and service sector. Our project has proven its effectiveness from different perspectives, such as contactless ergonomic assessment, movement data collection and corrective actions recommendations. As future work, we tend to implicate the Augmented reality in both industrial processes and ergonomic risk assessment, because the AR overlay the digital information onto the physical environment, which facilitates the manufacturing process understanding in parallel with providing an accurate and efficient way to identify and mitigate ergonomic hasards all virtually, especially that AR can enable real-time monitoring of workers' posture and movements.

Bibliography

[1] Stephen Bevan. Economic impact of musculoskeletal disorders (MSDs) on work in Europe. *Best Practice & Research Clinical Rheumatology*, 29(3):356–373, June 2015.

[2] Bureau of Labor Statistics. *Nonfatal occupational injuries and illnesses requiring days away from work*. United States Department of Labor Washington^ eDC DC, 2008.

[3] Daria Battini, Alessandro Persona, and Fabio Sgarbossa. Innovative real-time system to integrate ergonomic evaluations into warehouse design and management. *Computers & Industrial Engineering*, 77:1–10, November 2014.

[4] G. C. David. Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. *Occupational Medicine*, 55(3):190–199, May 2005.

[5] Nicolas Vignais, Fabien Bernard, Gérard Touvenot, and Jean-Claude Sagot. Physical risk factors identification based on body sensor network combined to videotaping. *Applied Ergonomics*, 65:410–417, November 2017.

[6] Manlio MassirisFernández, J. Álvaro Fernández, Juan M. Bajo, and Claudio A. Delrieux. Ergonomic risk assessment based on computer vision and machine learning. *Computers & Industrial Engineering*, 149:106816, November 2020.

[7] Pierre Plantard, Hubert P. H. Shum, Anne-Sophie Le Pierres, and Franck Multon. Validation of an ergonomic assessment method using Kinect data in real workplace conditions. *Applied Ergonomics*, 65:562–569, November 2017.

[8] Nipun D. Nath, Reza Akhavian, and Amir H. Behzadan. Ergonomic analysis of construction worker's body postures using wearable mobile sensors. *Applied Ergonomics*, 62:107–117, July 2017.

[9] B.A. Kadir, O. Broberg, and C.S.D. Conceição. Current research and future perspectives on human factors and ergonomics in Industry 4.0', Computers and Industrial Engineering. *Elsevier Ltd*, 137(2019):1–12, 2019.

[10] Jörg Krüger and The Duy Nguyen. Automated vision-based live ergonomics analysis in assembly operations. *CIRP Annals*, 64(1):9–12, January 2015.

[11] Xu Xu, Michelle Robertson, Karen B. Chen, Jia-hua Lin, and Raymond W. McGorry. Using the Microsoft Kinect™ to assess 3-D shoulder kinematics during computer use. *Applied Ergonomics*, 65:418–423, November 2017.

[12] Steven J. Scheer and Anil Mital. Ergonomics. *Archives of Physical Medicine and Rehabilitation*, 78(3, Supplement):S36–S45, March 1997.

[13] K. H. E. Kroemer and E. Grandjean. *Fitting The Task To The Human: A Textbook Of Occupational Ergonomics*. CRC Press, London, 5 edition, July 1997.

[14] Andreas Holtermann, Svend Erik Mathiassen, and Leon Straker. Promoting health and physical capacity during productive work: the Goldilocks Principle. *Scandinavian Journal of Work, Environment & Health*, 45(1):90–97, 2019.

[15] Stuart M. McGill. Evolving ergonomics? *Ergonomics*, 52(1):80–86, January 2009.

[16] A. Kilbom and J. Persson. Work technique and its consequences for musculoskeletal disorders. *Ergonomics*, 30(2):273–279, February 1987.

[17] Gisela Sjøgaard, Just Bendix Justesen, Mike Murray, Tina Dalager, and Karen Søgaard. A conceptual model for worksite intelligent physical exercise training–IPET–intervention for decreasing life style health risk indicators among employees: a randomized controlled trial. *BMC public health*, 14:652, June 2014.

[18] Donald E. Shrey and Michel Lacerte, editors. *Principles and practices of disability management in industry*. GR Press, Winter Park, Fla, 1995.

[19] Jack Dennerlein. Ergonomics and Musculoskeletal Issues. In *International Encyclopedia of Public Health*, pages 577–584. October 2016. Journal Abbreviation: International Encyclopedia of Public Health.

[20] Mike Maier. Work-related Musculoskeletal Disorders, Oregon, 1990-2000. January 2001.

[21] A. Lorusso, S. Bruno, and N. L'Abbate. [Musculoskeletal disorders among university student computer users]. *La Medicina Del Lavoro*, 100(1):29–34, 2009.

[22] Managing Editor. Know the Types of Musculoskeletal Disorders and How to Avoid Them, January 2023. Running Time: 163 Section: Personal Injury.

[23] Canadian Centre for Occupational Health and Safety Government of Canada. Work-related Musculoskeletal Disorders (WMSDs) : OSH Answers, February 2023. Last Modified: 2014-01-08.

[24] https://www.biron.com/en/education-center/neat-little-guide/musculoskeletal-disorders/ | Biron.

[25] European Agency for Safety and Health at Work. *OSH in figures : work-related musculoskeletal disorders in the EU : facts and figures*. Publications Office, LU, 2010.

[26] Thoracic Outlet Syndrome, August 2021.

[27] Patellar Tendonitis (Jumper's Knee) - Symptoms & Causes | Mount Elizabeth Hospitals.

[28] Karen Søgaard and Gisela Sjøgaard. Physiological Bases of Work Assessment. In John R. Wilson and Sarah Sharples, editors, *Evaluation of Human Work*, pages 419–445. CRC Press, London, New York, 2015.

[29] S Hignett. Intervention strategies to reduce musculoskeletal injuries associated with handling patients: a systematic review. *Occupational and Environmental Medicine*, 60(9):e6, September 2003.

[30] David J. Cochran. Work-Related Musculoskeletal Disorders: A Reference Book for Prevention by I. Kuorinka & L. Forcier 1995, 421 pp., $37.50 (pbk.) London: Taylor & Francis ISBN 0-7484-0132-6. *Ergonomics in Design*, 5(4):30–30, October 1997. Publisher: SAGE Publications Inc.

[31] Sue Hignett and Lynn McAtamney. Rapid Entire Body Assessment (REBA). *Applied Ergonomics*, 31(2):201–205, 2000.

[32] Dohyung Kee. Comparison of OWAS, RULA and REBA for assessing potential work-related musculoskeletal disorders. *International Journal of Industrial Ergonomics*, 83:103140, May 2021.

[33] Osmo Karhu, Pekka Kansi, and Iikka Kuorinka. Correcting working postures in industry: A practical method for analysis. *Applied Ergonomics*, 8(4):199–201, December 1977.

[34] Tzu-Hsien Lee and Chia-Shan Han. Analysis of Working Postures at a Construction Site Using the OWAS Method. *International Journal of Occupational Safety and Ergonomics*, 19(2):245–250, 2013. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10803548.2013.11076983.

[35] Alex Burdorf. The role of assessment of biomechanical exposure at the workplace in the prevention of musculoskeletal disorders. *Scandinavian Journal of Work, Environment & Health*, 36(1):1–2, January 2010.

[36] Marie-Ève Chiasson, Daniel Imbeau, Karine Aubry, and Alain Delisle. Comparing the results of eight methods used to evaluate risk factors associated with musculoskeletal disorders. *International Journal of Industrial Ergonomics*, 42(5):478–488, September 2012.

[37] Barbara Silverstein and Randy Clark. Interventions to reduce work-related musculoskeletal disorders. *Journal of Electromyography and Kinesiology*, 14(1):135–152, February 2004.

[38] Dohyung Kee. An empirical comparison of OWAS, RULA and REBA based on self-reported discomfort. *International journal of occupational safety and ergonomics: JOSE*, 26(2):285–295, June 2020.

[39] Dohyung Kee. Systematic Comparison of OWAS, RULA, and REBA Based on a Literature Review. *International Journal of Environmental Research and Public Health*, 19(1):595, January 2022.

[40] Esa-Pekka Takala, Irmeli Pehkonen, Mikael Forsman, Gert-Åke Hansson, Svend Erik Mathiassen, W. Patrick Neumann, Gisela Sjøgaard, Kaj Bo Veiersted, Rolf H. Westgaard, and Jörgen Winkel. Systematic evaluation of observational methods assessing biomechanical exposures at work. *Scandinavian Journal of Work, Environment & Health*, 36(1):3–24, January 2010.

[41] Lynn McAtamney. RULA: a survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics*, 24(2):91–99, April 1993. Publisher: Elsevier.

[42] Sujee Lee, Li Liu, Robert Radwin, and Jingshan Li. Machine Learning in Manufacturing Ergonomics: Recent Advances, Challenges, and Opportunities. *IEEE Robotics and Automation Letters*, 6(3):5745–5752, July 2021. Conference Name: IEEE Robotics and Automation Letters.

[43] Elena Stefana, Filippo Marciano, Diana Rossi, Paola Cocca, and Giuseppe Tomasoni. Wearable Devices for Ergonomics: A Systematic Literature Review. *Sensors*, 21(3):777, January 2021. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.

[44] Ilaria Conforti, Ilaria Mileti, Zaccaria Del Prete, and Eduardo Palermo. Measuring Biomechanical Risk in Lifting Load Tasks Through Wearable System and Machine-Learning Approach. *Sensors*, 20(6):1557, January 2020. Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.

[45] Seyed Mohammadreza Hosseinian, Yibo Zhu, Ranjana K. Mehta, Madhav Erraguntla, and Mark A. Lawley. Static and Dynamic Work Activity Classification from a Single Accelerometer: Implications for Ergonomic Assessment of Manual Handling Tasks. *IISE Transactions on Occupational Ergonomics and Human Factors*, 7(1):59–68, January 2019. Publisher: Taylor & Francis.

[46] Li Li, Tara Martin, and Xu Xu. A novel vision-based real-time method for evaluating postural risk factors associated with musculoskeletal disorders. *Applied Ergonomics*, 87:103138, September 2020.

[47] Oguz Akkas, Cheng-Hsien Lee, Yu Hen Hu, Thomas Y. Yen, and Robert G. Radwin. Measuring elemental time and duty cycle using automated video processing. *Ergonomics*, 59(11):1514–1525, November 2016.

[48] Oguz Akkas, Cheng Hsien Lee, Yu Hen Hu, Carisa Harris Adamson, David Rempel, and Robert G. Radwin. Measuring exertion time, duty cycle and hand activity level for industrial tasks using computer vision. *Ergonomics*, 60(12):1730–1738, December 2017. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00140139.2017.1346208.

[49] Marco Bortolini, Mauro Gamberi, Francesco Pilati, and Alberto Regattieri. Automatic assessment of the ergonomic risk for manual manufacturing and assembly activities through optical motion capture technology. *Procedia CIRP*, 72:81–86, January 2018.

[50] Chika Edith Mgbemena, John Oyekan, Ashutosh Tiwari, Yuchun Xu, Sarah Fletcher, Windo Hutabarat, and Vinayak Prabhu. Gesture Detection Towards Real-Time Ergonomic Analysis for Intelligent Automation Assistance. In Christopher Schlick and Stefan Trzcieliński, editors, *Advances in Ergonomics of Manufacturing: Managing the Enterprise of the Future*, Advances in Intelligent Systems and Computing, pages 217–228, Cham, 2016. Springer International Publishing.

[51] Vivek Phuloria. Machine learning simplified.

[52] Joseph O'Rourke and Norman I. Badler. Model-based image analysis of human motion using constraint propagation. *IEEE Transactions on Pattern Analysis and*

Machine Intelligence, PAMI-2(6):522–536, November 1980. Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.

[53] Cristian Ferrer, Josep Casas, Muhilan Par, and E Monte. Towards a Fair Evaluation of 3D Human Pose Estimation Algorithms. October 2009.

[54] Lars Mündermann, Stefano Corazza, and Thomas P Andriacchi. The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. *Journal of NeuroEngineering and Rehabilitation*, 3:6, March 2006.

[55] George Papandreou, Tyler Zhu, Nori Kanazawa, Alexander Toshev, Jonathan Tompson, Chris Bregler, and Kevin Murphy. Towards Accurate Multi-person Pose Estimation in the Wild, April 2017. arXiv:1701.01779 [cs].

[56] S. Corazza, L. Mündermann, A. M. Chaudhari, T. Demattio, C. Cobelli, and T. P. Andriacchi. A markerless motion capture system to study musculoskeletal biomechanics: visual hull and simulated annealing approach. *Annals of Biomedical Engineering*, 34(6):1019–1029, June 2006.

[57] Martin Sandau, Henrik Koblauch, Thomas B. Moeslund, Henrik Aanæs, Tine Alkjær, and Erik B. Simonsen. Markerless motion capture can provide reliable 3D gait kinematics in the sagittal and frontal plane. *Medical Engineering & Physics*, 36(9):1168–1175, September 2014.

[58] Ami Drory, Hongdong Li, and Richard Hartley. A learning-based markerless approach for full-body kinematics estimation in-natura from a single image. *Journal of Biomechanics*, 55:1–10, April 2017.

[59] Theresa E. McGuirk, Elliott S. Perry, Wandasun B. Sihanath, Sherveen Riazati, and Carolynn Patten. Feasibility of Markerless Motion Capture for Three-Dimensional Gait Assessment in Community Settings. *Frontiers in Human Neuroscience*, 16, 2022.

[60] M.A. Fischler and R.A. Elschlager. The Representation and Matching of Pictorial Structures. *IEEE Transactions on Computers*, C-22(1):67–92, January 1973. Conference Name: IEEE Transactions on Computers.

[61] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Pictorial Structures for Object Recognition. *International Journal of Computer Vision*, 61(1):55–79, January 2005.

[62] Learning to parse images of articulated bodies. The MIT Press, 2007. Book Title: Advances in Neural Information Processing Systems 19.

[63] Yi Yang and Deva Ramanan. Articulated Human Detection with Flexible Mixtures of Parts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12):2878–2890, December 2013. Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.

[64] Alexander Toshev and Christian Szegedy. DeepPose: Human Pose Estimation via Deep Neural Networks. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1653–1660, June 2014. arXiv:1312.4659 [cs].

[65] Alejandro Newell, Kaiyu Yang, and Jia Deng. Stacked Hourglass Networks for Human Pose Estimation, July 2016. arXiv:1603.06937 [cs].

[66] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, May 2019. arXiv:1812.08008 [cs].

[67] Shih-En Wei, Varun Ramakrishna, Takeo Kanade, and Yaser Sheikh. Convolutional Pose Machines, April 2016. arXiv:1602.00134 [cs].

[68] Leonid Sigal. Human Pose Estimation. In Katsushi Ikeuchi, editor, *Computer Vision: A Reference Guide*, pages 573–592. Springer International Publishing, Cham, 2021.

[69] Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7):1325–1339, July 2014. Publisher: IEEE Computer Society.

[70] Jinbao Wang, Shujie Tan, Xiantong Zhen, Shuo Xu, Feng Zheng, Zhenyu He, and Ling Shao. Deep 3D human pose estimation: A review. *Computer Vision and Image Understanding*, 210:103225, September 2021.

[71] Bugra Tekin, Pablo Márquez-Neila, Mathieu Salzmann, and Pascal Fua. Learning to Fuse 2D and 3D Image Cues for Monocular Body Pose Estimation, April 2017. arXiv:1611.05708 [cs].

[72] Helge Rhodin, Mathieu Salzmann, and Pascal Fua. Unsupervised Geometry-Aware Representation for 3D Human Pose Estimation, April 2018. arXiv:1804.01110 [cs].

[73] Kaoning Hu, Lijun Yin, and Tianyang Wang. Temporal Interframe Pattern Analysis for Static and Dynamic Hand Gesture Recognition. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 3422–3426, September 2019. ISSN: 2381-8549.

[74] Vision-Based Gesture Recognition | DigiKey.

[75] Beatrice de Gelder. Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3475–3484, December 2009. Publisher: Royal Society.

[76] Manuel Hita-Gutiérrez, Marta Gómez-Galán, Manuel Díaz-Pérez, and Ángel-Jesús Callejón-Ferre. An Overview of REBA Method Applications in the World. *International Journal of Environmental Research and Public Health*, 17(8):2635, January 2020. Number: 8 Publisher: Multidisciplinary Digital Publishing Institute.

[77] Sylvaine Tuncer. The Effects of Video Recording on Office Workers' Conduct, and the Validity of Video Data for the Study of Naturally-Occurring Interactions. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 17(3), September 2016. Number: 3.

[78] www.naltis.com. Qui sommes-nous ? Présentation ALZINC Production Zinc & dérivés.

[79] Dario Pavllo, Christoph Feichtenhofer, David Grangier, and Michael Auli. 3D human pose estimation in video with temporal convolutions and semi-supervised training, March 2019. arXiv:1811.11742 [cs].

[80] Matthew Trumble, Andrew Gilbert, Adrian Hilton, and John Collomosse. *Deep Autoencoder for Combined Human Pose Estimation and body Model Upscaling*. July 2018.

[81] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. Detectron2, 2019.

[82] Omar Elharrouss, Younes Akbari, Noor Almaadeed, and Somaya Al-Maadeed. Backbones-Review: Feature Extraction Networks for Deep Learning and Deep Reinforcement Learning Approaches, June 2022. arXiv:2206.08016 [cs].

[83] Jas. COCO-WholeBody, March 2023. original-date: 2020-07-12T07:40:15Z.

[84] Tao Feng, Jiange Liu, Xia Fang, Jie Wang, and Libin Zhou. A Double-Branch Surface Detection System for Armatures in Vibration Motors with Miniature Volume Based on ResNet-101 and FPN. *Sensors*, 20:2360, April 2020.

[85] Shaojie Bai, J. Zico Kolter, and Vladlen Koltun. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling, April 2018. arXiv:1803.01271 [cs].

[86] Wanshun Wong. What is Residual Connection?, December 2021.

[87] Yihao Zhang, Junhao Wen, Xibin Wang, and Zhuo Jiang. Semi-supervised learning combining co-training with active learning. *Expert Systems with Applications*, 41(5):2372–2378, April 2014.

[88] B.D. Lowe, P.L. Weir, and D.M. Andrews. Observation-based posture assessment: review of current practice and recommendations for improvement. Technical Report DHHS (NIOSH) Publication No. 2014–131, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Cincinnati, OH, 2014.

[89] PyTorch.

[90] Google Colaboratory.

[91] Sashank J. Reddi, Satyen Kale, and Sanjiv Kumar. On the convergence of adam and beyond, 2019.

[92] Young-Jin Kwon, Do-Hyun Kim, Byung-Chang Son, Kyoung-Ho Choi, Sungbok Kwak, and Taehong Kim. A Work-Related Musculoskeletal Disorders (WMSDs) Risk-Assessment System Using a Single-View Pose Estimation Model. *International Journal of Environmental Research and Public Health*, 19(16):9803, January 2022. Number: 16 Publisher: Multidisciplinary Digital Publishing Institute.

[93] MPII Human Pose Database.

[94] COCO - Common Objects in Context.

[95] Human Pose Estimation: Deep Learning Approach [2022 Guide].

[96] mlkit/android/vision-quickstart at master · googlesamples/mlkit.

[97] Panagiota Tsarouchi, Sotiris Makris, and George Chryssolouris. Human–robot interaction review and challenges on task planning and programming. *International Journal of Computer Integrated Manufacturing*, 29(8):916–931, August 2016. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/0951192X.2015.1130251.

[98] Assistive Gym v1.0, October 2022. original-date: 2019-09-26T03:42:20Z.

[99] Sara Santos, Duarte Folgado, João Rodrigues, Nafiseh Mollaei, Carlos Fujão, and Hugo Gamboa. Explaining the Ergonomic Assessment of Human Movement in Industrial Contexts. pages 79–88, January 2020.

[100] Md. Atiqur Rahman Ahad. *Motion History Images for Action Recognition and Understanding*. SpringerBriefs in Computer Science. Springer, London, 2013.

[101] J.K. Aggarwal and M.S. Ryoo. Human activity analysis: A review. *ACM Computing Surveys*, 43(3):1–43, April 2011.

[102] Alia Rifat, Pratiksha Pradip Pandao, and B. Shoban Babu. Machine Learning Model for Human Activity Analysis. *European Journal of Information Technologies and Computer Science*, 2(1):5–9, January 2022. Number: 1.

[103] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. June 2015.

[104] Romero Morais, Vuong Le, Truyen Tran, Budhaditya Saha, Moussa Mansour, and Svetha Venkatesh. Learning Regularity in Skeleton Trajectories for Anomaly Detection in Videos. March 2019.

[105] Waqas Sultani, Chen Chen, and Mubarak Shah. Real-world Anomaly Detection in Surveillance Videos, February 2019. arXiv:1801.04264 [cs].

[106] Matteo Zago, Ana Francisca Rozin Kleiner, and Peter Andreas Federolf. Editorial: Machine Learning Approaches to Human Movement Analysis. *Frontiers in Bioengineering and Biotechnology*, 8, 2021.

[107] Charu C. Aggarwal. *Outlier Analysis*. Springer, December 2016. Google-Books-ID: KyG1DQAAQBAJ.

[108] Anis Davoudi, Kumar Rohit Malhotra, Benjamin Shickel, Scott Siegel, Seth Williams, Matthew Ruppert, Emel Bihorac, Tezcan Ozrazgat-Baslanti, Patrick J. Tighe, Azra Bihorac, and Parisa Rashidi. Intelligent ICU for Autonomous Patient Monitoring Using Pervasive Sensing and Deep Learning. *Scientific Reports*, 9(1):8020, May 2019. Number: 1 Publisher: Nature Publishing Group.

[109] Marta Lorenzini, Wansoo Kim, Elena De Momi, and Arash Ajoudani. An Online Method to Detect and Locate an External Load on the Human Body with Applications in Ergonomics Assessment. *Sensors*, 20(16):4471, January 2020. Number: 16 Publisher: Multidisciplinary Digital Publishing Institute.

[110] Runyu L. Greene, Ming-Lun Lu, Menekse Salar Barim, Xuan Wang, Marie Hayden, Yu Hen Hu, and Robert G. Radwin. Estimating Trunk Angle Kinematics During Lifting Using a Computationally Efficient Computer Vision Method. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 64(3):482–498, May 2022.

[111] Ken Kato, Ryojun Ikeura, Shigeyoshi Tsutsumi, Soichiro Hayakawa, and Hideki Sawai. Analysis of human object lifting operation force using power assist device. *Journal of Physics: Conference Series*, 2107(1):012009, November 2021. Publisher: IOP Publishing.

[112] Lynn McAtamney and E. Nigel Corlett. RULA: a survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics*, 24(2):91–99, April 1993.

[113] Pierre Plantard, Hubert P. H. Shum, Anne-Sophie Le Pierres, and Franck Multon. Validation of an ergonomic assessment method using Kinect data in real workplace conditions. *Applied Ergonomics*, 65:562–569, November 2017.

[114] Hearst Acquires Work-Loss Data Institute (WLDI).

[115] Gerard Mcnulty. What is Carpal Tunnel Syndrome?, January 2019.

[116] Matti Viljanen, Antti Malmivaara, Jukka Uitti, Marjo Rinne, Pirjo Palmroos, and Pekka Laippala. Effectiveness of dynamic muscle training, relaxation training, or ordinary activity for chronic neck pain: randomised controlled trial. *BMJ (Clinical research ed.)*, 327(7413):475, August 2003.

[117] High-level expert group on artificial intelligence | Shaping Europe's digital future, February 2023.

[118] Deborah Petrat. Artificial intelligence in human factors and ergonomics: an overview of the current state of research. *Discover Artificial Intelligence*, 1(1):3, September 2021.

[119] Ritika Wason. Deep learning: Evolution and expansion. *Cognitive Systems Research*, 52:701–708, December 2018.

[120] Juergen Schmidhuber. Deep Learning in Neural Networks: An Overview. *Neural Networks*, 61:85–117, January 2015. arXiv:1404.7828 [cs].

[121] T S Huang. Computer Vision: Evolution and Promise.

[122] Shichao Jin, Qinghua Guo, Min Li, Qiuli Yang, Kexin Xu, Yuanzhen Ju, Jing Zhang, Jing Xuan, Yanjun Su, Qiang Xu, and Yu Liu. Application of deep learning in ecological resource research: Theories, methods, and challenges. *Science China Earth Science*, March 2020.

[123] A. Pastor López-Monroy and Jesús S. García-Salinas. Chapter 9 - Neural networks and deep learning. In Alejandro A. Torres-García, Carlos A. Reyes-García, Luis Villaseñor-Pineda, and Omar Mendoza-Montoya, editors, *Biosignal Processing and Classification Using Computational Learning and Intelligence*, pages 177–196. Academic Press, January 2022.

[124] Feedforward Neural Networks (FNN) - Deep Learning Wizard.

[125] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning, 2016.

[126] Yann LeCun, Bernhard Boser, John S. Denker, Donnie Henderson, Richard E. Howard, Wayne Hubbard, and Lawrence D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989. Publisher: MIT Press.

[127] Dian Palupi Restuputri. Metode REBA Untuk Pencegahan Musculoskeletal Disorder Tenaga Kerja. *Jurnal Teknik Industri*, 18(1):19–28, August 2017. Number: 1.

[128] Ni Luh Gede Aris Maytadewi Negara, I Dewa Putu Sutjana, and Ni Made Citra Aryani. Work Posture Analysis Using Rapid Entire Body Assessment on Workers at AR Tailor Denpasar. September 2021.

[129] Lynn McAtamney and E. Nigel Corlett. RULA: a survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics*, 24(2):91–99, April 1993.

[130] K. S. Lee, J. W. Yeom, H. M. Hur, J. M. Sohn, D. M. Oak, and S. G. Park. Comparisons of different evaluation tools for musculoskeletal disorders. *Fall Conferen Ergono Socie Korea*, 11:205–210, 2009.

[131] R. Divya and J. Dinesh Peter. Smart healthcare system-a brain-like computing approach for analyzing the performance of detectron2 and PoseNet models for anomalous action detection in aged people with movement impairments. *Complex & Intelligent Systems*, 8(4):3021–3040, August 2022.

[132] J. O'Rourke and N. I. Badler. Model-based image analysis of human motion using constraint propagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2(6):522–536, November 1980.

[133] Ching-Hang Chen and Deva Ramanan. 3D Human Pose Estimation = 2D Pose Estimation + Matching. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5759–5767, July 2017. ISSN: 1063-6919.

[134] Vito Modesto Manghisi, Markus Wilhelm, Antonello Uva, Bastian Engelmann, Michele Fiorentino, and Jan Schmitt. Towards gestured-based technologies for human-centred Smart Factories. *International Journal of Computer Integrated Manufacturing*, 36(1):110–127, January 2023. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/0951192X.2022.2121424.

[135] Annalisa Franco, Antonio Magnani, and Dario Maio. A multimodal approach for human activity recognition based on skeleton and RGB data. *Pattern Recognition Letters*, 131:293–299, March 2020.

[136] Sharath Chandra Akkaladevi, Matthias Propst, Michael Hofmann, Leopold Hiesmair, Markus Ikeda, Naresh Chowdary Chitturi, and Andreas Pichler. Programming-Free Approaches for Human–Robot Collaboration in Assembly Tasks. In Lihui Wang, Xi Vincent Wang, József Váncza, and Zsolt Kemény, editors, *Advanced Human-Robot Collaboration in Manufacturing*, pages 283–317. Springer International Publishing, Cham, 2021.

[137] Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7):1325–1339, July 2014. Publisher: IEEE Computer Society.

[138] Fisher Yu and Vladlen Koltun. Multi-Scale Context Aggregation by Dilated Convolutions, April 2016. arXiv:1511.07122 [cs].

[139] Izaak Van Crombrugge, Seppe Sels, Bart Ribbens, Gunther Steenackers, Rudi Penne, and Steve Vanlanduit. Accuracy Assessment of Joint Angles Estimated from 2D and 3D Camera Measurements. *Sensors (Basel, Switzerland)*, 22(5):1729, February 2022.

[140] Online Protractor.

APPENDIX A

Appendix

ANIMATOR
Where AI Meets Ergonomics

Startup Graduation Certificate Project

Under the Framework of ministerial decree number 1275

Team members:
Mr. Mallek Abdelmalek

Supervisor:
Mme. Handouzi Wahida

May 21, 2023

Contents

Chapter 1

Project presentation

1.1 Idea's concept

ANIMATOR is an AI-based tool designed for ergonomic risk assessment and digitisation to 3D environment. All started two years ago, when a massive data analysis project began by me to understand the actual situation of the Algerian industry, where the business problem was why aren't we developed?. The first analysis lead us to discover a huge gap of occupational health, where the well-being of workers wasn't a priority, which result on the task performance and productivity due to a lot of factors. The role of ANIMATOR is to assess the workers' postures and movements as well as analysing the feedback to ensure the safety and avoid the work-related musculoskeletal disorders which is a global health issue. Besides the industrial sector, ANIMATOR can be used as an Aid-rehabilitation tool, since it's AI-based, it can be fed by the patient data to output a customise training under the doctor supervision.

1.2 ANIMATOR's Values

The main ANIMATOR's values in its Beta version are:

- ANIMATOR will be the main key enhancing and guaranteeing the workers' occupational health by providing a tracking assessment in real-time besides a data-driven understanding of their behaviour, the thing that will make the tasks affectation more efficient.

- ANIMATOR will provide the insights for the suitable training for any type of work, which reduces the costs and burdens of traditional training, all of this through augmented reality.

- The physician therapists will be able to process more patients online without traveling.
- ANIMATOR will be a customised friendly-user tool.

1.3 ANIMATOR goals

The goals will vary depending on multiple aspects, in this section we'll present the main goals on the main three plans:

1.3.1 Operational plan

- Developing and enhancing the tool performance and the data collection.

1.3.2 Tactic plan

- Defining our "Purple Cow" to the agents by boosting our marketing plan, increasing our network and defining the pricing politics.
- Covering the main industrial zone in Tlemcen.

1.3.3 Strategic plan

- Launching ANIMATOR commercial version (Free trial demo + Premium access)
- Becoming a social partner.

1.4 Project planning

We're using Agile methodology to manage and complete the project. This approach allows us to be more flexible, adaptive and collaborative in our work, ensuring that we're delivering high-quality results that meet the need of our future agents. In figure 1.1 presents the Agile progress board, where the task are distributed to four main phase (**To do, Doing, Done, Deployment**).

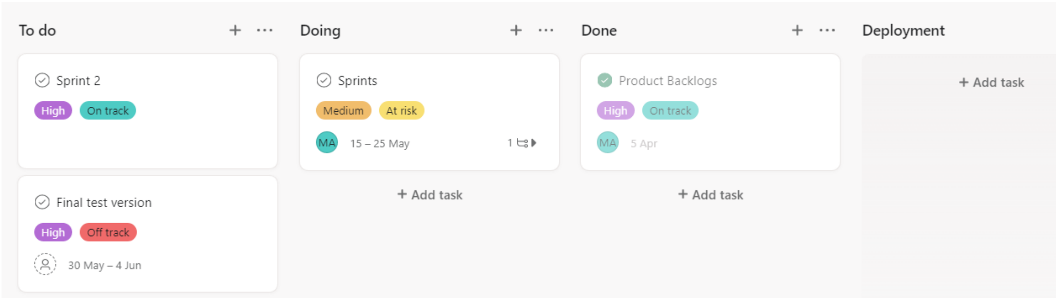


Figure 1.1: ANIMATOR progress board

Figure 1.2 represents the interactive dashboard of ANIMATOR task progress.

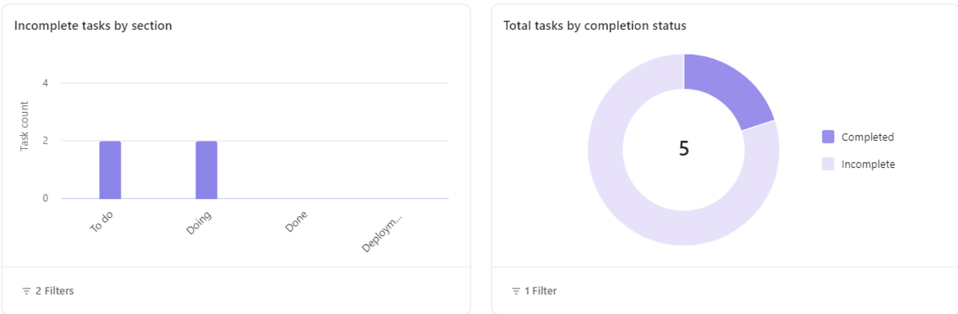


Figure 1.2: Animator progress dashboard

Chapter 2

Innovative aspects

2.1 Innovation in ANIMATOR

In concurrence with world's ergonomic leader such as Ergo-Plus[1] and others, ANIMATOR not only will be the first AI-based tool for ergonomic risk assessment in real time, but the first tool that using it we'll push to recognise the work-related musculoskeletal disorders as occupational injuries. Second, we'll lead ergonomics in Algeria's industrial field, besides all other services that ANIMATOR provides.

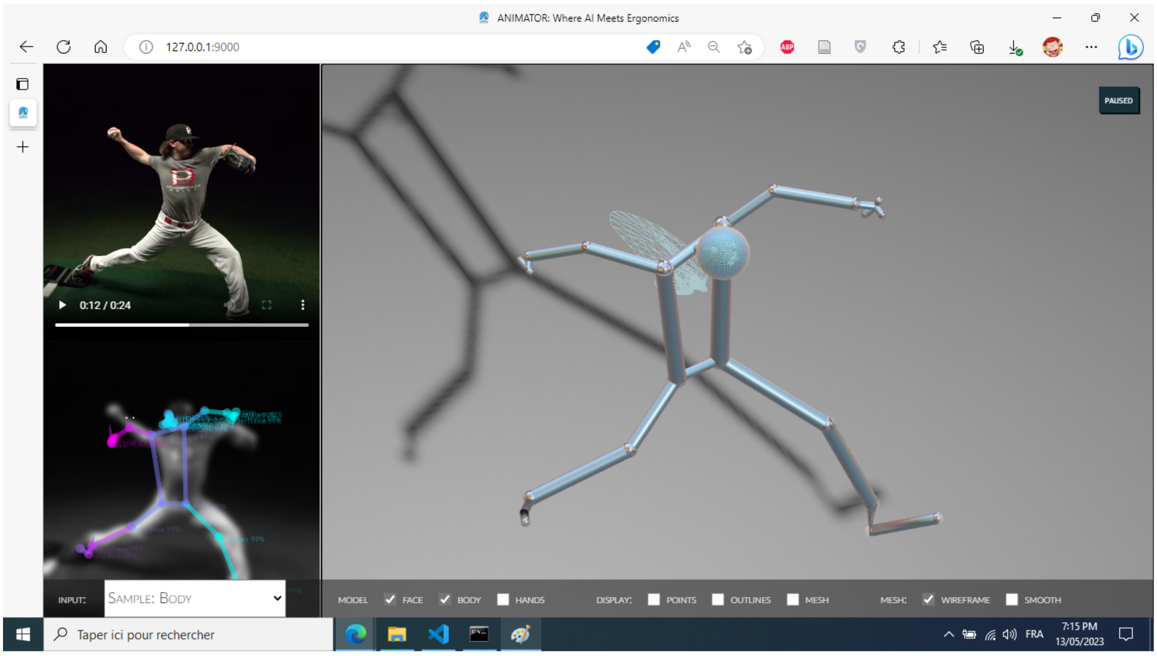


Figure 2.1: ANIMATOR user interface

[1]Ergo-plus can be accessed from this link Ergo-Plus

Chapter 3

Strategic market analysis

3.1 Market overview

Till nowadays, there isn't an ergonomic market or even ergonomic plan consideration in Algeria, another motivation for ANIMATOR. At the international level, the ergonomic market value can reach 26.7 Billion by 2027 according to *Markets and Markets*[1] website, and since Algeria is a country on the path of growth, soon or later it will be part of this market, especially the ISO certification is a requirement for the international market.

3.2 Competition market analysis

3.2.1 SWOT analysis

Strengths

- Unique and innovative service that provides real-time 3D human pose estimation and ergonomic risk assessment.
- Potential for digital twin construction.
- Ability to provide customised solution for various fitting task-person problems.
- Expertise in machine learning and computer vision besides the industrial engineering.

[1] Markets and Markets is a website designed for any market description with real data

Weaknesses

- New brand, new market creation.
- Weakness of infrastructure (internet and technology).
- limited initial funding.

Opportunities

- Growing market demand for technology-based solution especially in health-care, manufacturing and retail industries.
- Ability to expand into international market with similar needs.

Threats

- Social acceptance.

3.2.2 Porter's Five Forces analysis

Threats of new entrants

This threat is relatively low, since it's innovative and needs expertise in AI and ergonomics.

Bargaining power of suppliers

ANIMATOR may face challenges in terms of the availability of required hardware to develop and update.

Bargaining power of buyers

Relatively high or near high, since there are several similar software providing the same principle of service and functionalities. However, ANIMATOR is unique as it gives special features like the real-time ergonomic assessment and on demand digitisation.

Threat of substitutes

This threat is moderate as there are many other manual methods for ergonomic assessment, which is already new field in Algeria. However, ANIMATOR still advantageous.

Competitive rivalry

ANIMATOR is the 1st web application in its kind of functionality.

3.3 Marketing strategy

Our strategy will be an hybrid non traditional method. First, we'll push for the conscience of occupational health, make the industrial sector occupant aware of the harm of ergonomic risk and its impact on the economy. Then we'll start the event Marketing to showcase ANIMATOR's services and products.
The network will be built using social media, social partners (CNAS, CASNOS and Audit centers) and SEO (search engine optimisation). The pricing politics will start with freemium, which is free trial in which we'll provide the service of ergonomic risk assessment by impleenting it during 2 months. The first month is for data collection and processing, analysing the current work health and giving the recommendations to enhance the health. In the second month, we'll track and assess, comparing the the previous month and share the enhanced results through both visualisations and ground truth diagnosis, which will be seen in work productivity and workers' well being.
At that level, ANIMATOR will be a monthly subscription, and the prices will vary based on the enhancement level.

Chapter 4

Production plan

4.1 Development workflow

ANIMATOR is developed using System engineering, which is an interdisciplinary approach to designing, analysing and managing complex systems. The development cycle will follow this diagram below:

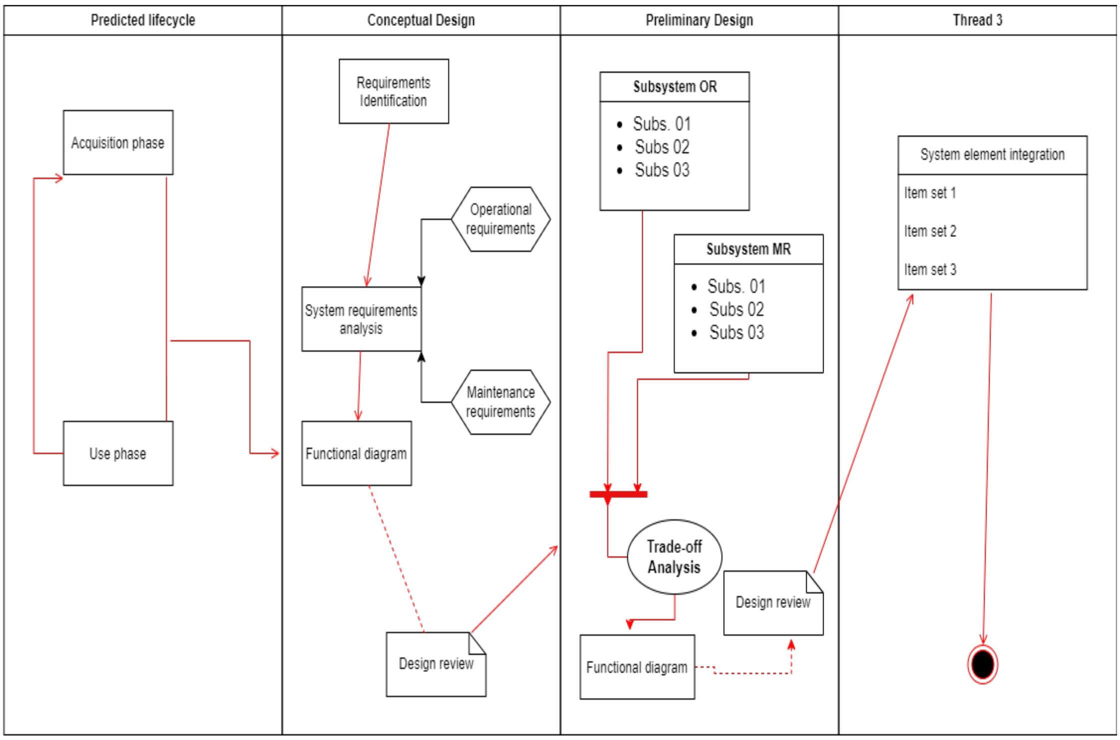


Figure 4.1: Development workflow under system engineering

4.2 Workforce

The integration of ANIMATOR will launch a new vision of the work quality and task completion in the Algerian industry by changing the classic decision to a data-driven decision. This unique leap is capable of creating a new chain of digital supply, mainly first by providing a new data mine from the industrial sector. The main job opportunities that will be created are in data analysis field in combination with Industrial engineering.

4.3 Key partners

Since ANIMATOR is a purely new technology Startup, we will focus gaining the major partners that can help us to release and spread our ideas and businesses:

- Tlemcen's faculty of technology
- Industrial partners
- CNAS and CASNOS
- Hospitals
- IBM CLOUD
- Microsoft Azure
- Hardware suppliers

Chapter 5

Prototype

Until this day, We have reached the capability of 3 AI-based services which are:

- Ergonomic risk assessment (using REBA score).
- 3D body pose estimation.
- face reconstruction.
- Hand pose estimation.

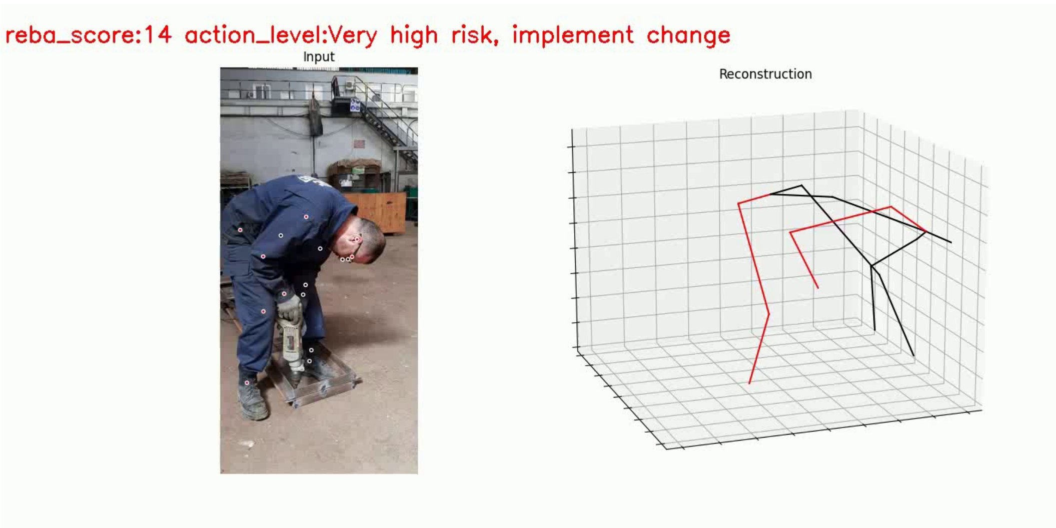


Figure 5.1: Ergonomic risk assesssment

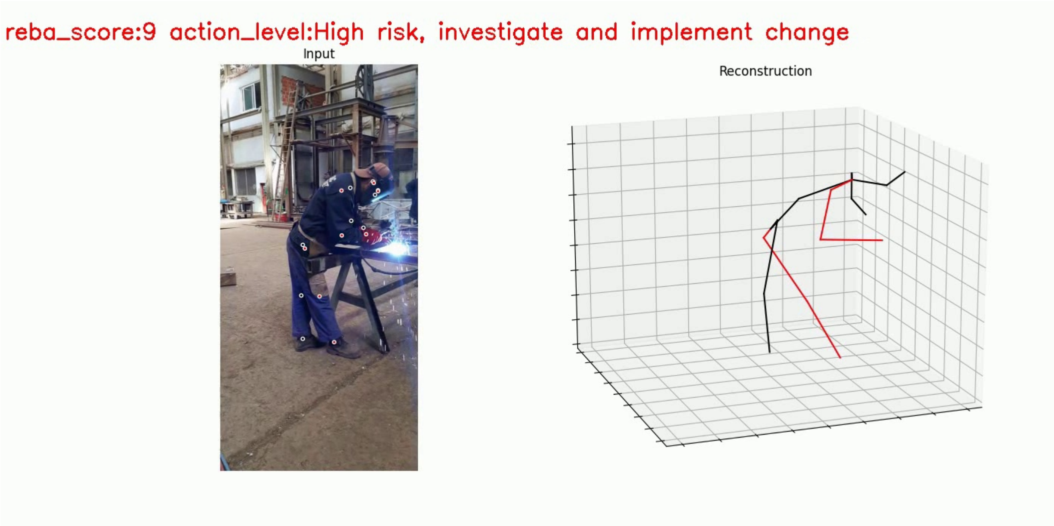


Figure 5.2: Ergonomic risk assesssment

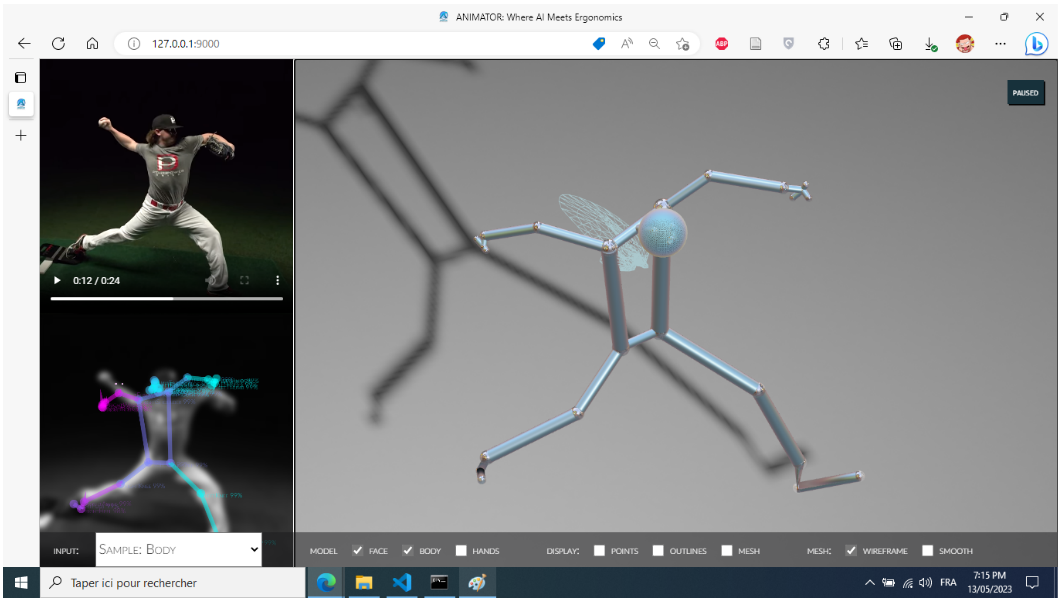


Figure 5.3: 3D body pose estimation

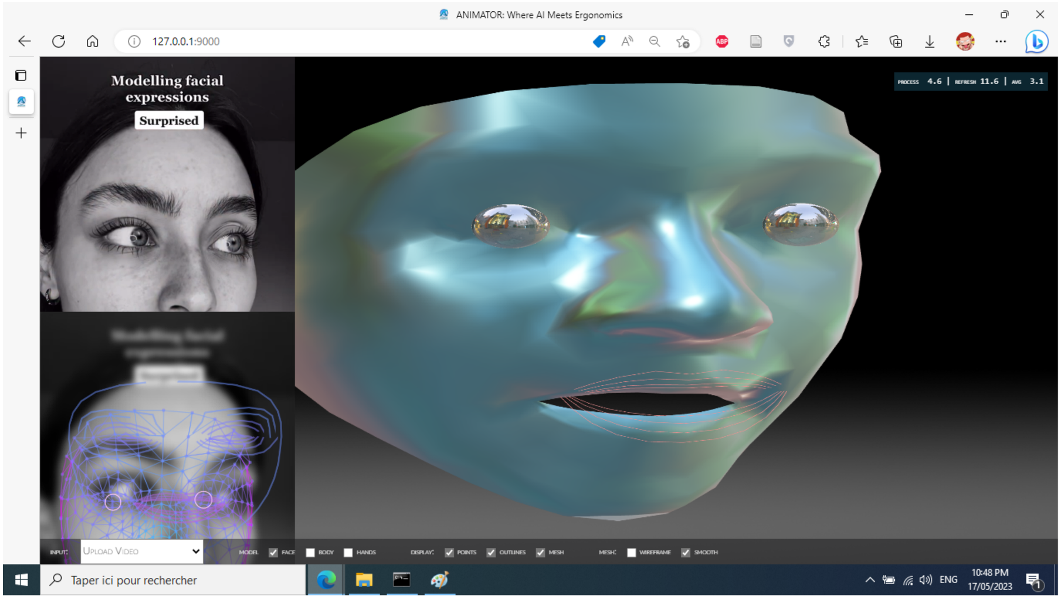


Figure 5.4: Face reconstruction-Smooth

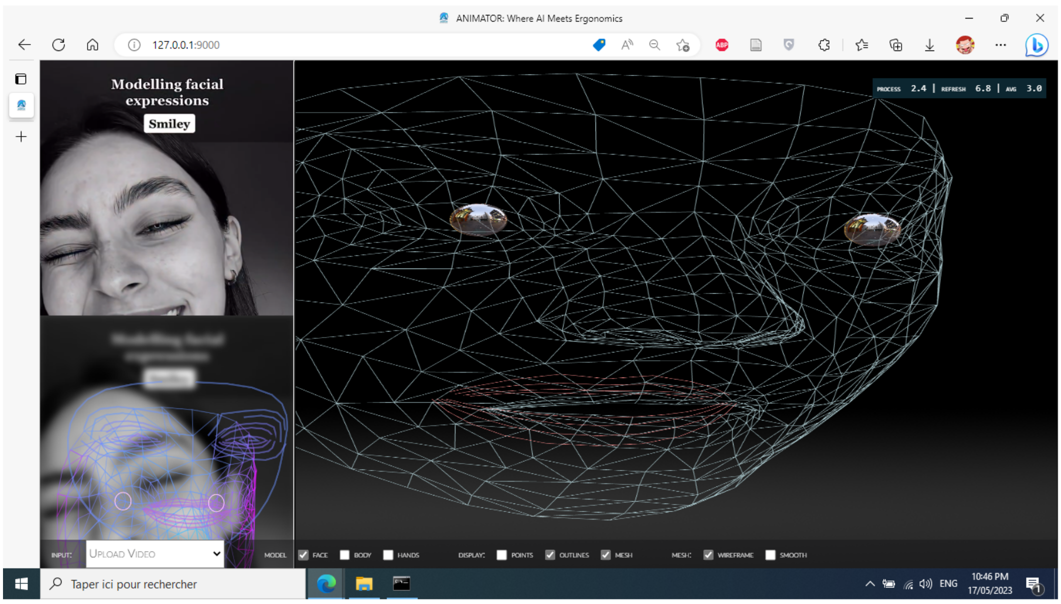


Figure 5.5: Face reconstruction-wireframe

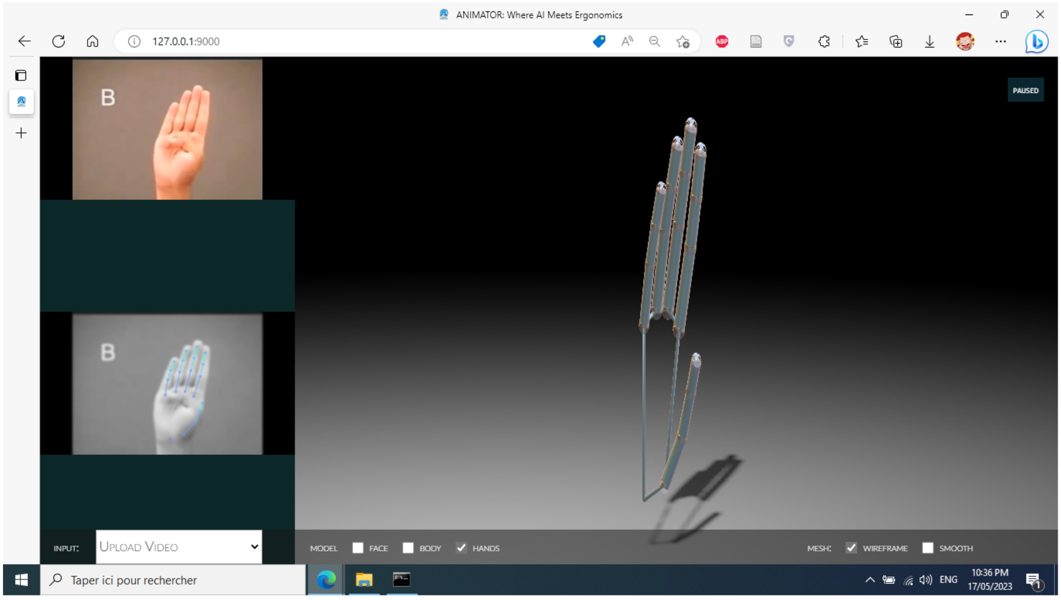


Figure 5.6: Hand pose estimation

Appendix B

Business Model Canvas

Business Model Canvas

Key Partners

- Tlemcen's faculty of technology.
- IBM Cloud.
- Microsoft azure.
- Industrial partners.
- CNAS/CASNOS.
- Hospitals.
- Hardware suppliers.
- Certification organisations.

Key Activities

- Research and development.
- Tailored ANIMATOR's services.
- Marketing and sales.
- Customer onboarding, training and support.
- Regular updates and enhancements.

Key Resources

- AI and Biomechanics expertise.
- Data science team.
- Ergonomists.
- Research laboratory on demand.

Value Propositions

- Real-time ergonomic risk assessment.
- Enhancing work and health quality.
- Provide simulated/AR solutions for training
- Innovative markerless human motion analysis.
- Telehealth-physical therapy.

Customer Relationships

- Personalised customer onboarding and training.
- Technical support using email, phone and chat.
- Proactive communication for updates.

5. Channels

- Online platforms.
- Direct sales.
- Partnership with CNAS/CASNOS and certification organisations.
- Content marketing using social media, TV and Podcasts.
- Events and tradeshows.

Customer Segments

- Manufacturing companies.
- Hospitals.
- physical therapy clinics.
- Professional sport's centers

Cost Structure

- Equipment costs.
- IT subscription costs.
- Third party costs

Revenue Streams

- Software As A service (SaaS)
- Consultancy services and Data-driven solutions sales.
- Subscription packages.
- Digitisation solutions sales (Digital twin).