

# **openTSST – An Open Web Platform for Large-Scale, Video-Based Motion Analysis During Acute Psychosocial Stress**

## **Bachelor's Thesis in Computer Science**

submitted  
by

Tobias Geßler

born 17.07.2001 in Pfaffenhofen an der Ilm

Written at

Machine Learning and Data Analytics Lab  
Department Artificial Intelligence in Biomedical Engineering  
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

in Cooperation with

Chair of Health Psychology  
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Advisors: Robert Richer M.Sc., Luca Abel M.Sc., Arne Küderle M.Sc.,  
Prof. Dr. Bjoern Eskofier,  
Prof. Dr. Nicolas Rohleder (*Chair of Health Psychology*)

Started: 15.03.2023

Finished: 14.08.2023



Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Die Richtlinien des Lehrstuhls für Bachelor- und Masterarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts.

Erlangen, den 14. August 2023



## Übersicht

Um potenzielle negative gesundheitliche Folgen abschwächen zu können, welche akuten Stressreaktionen zugrunde liegen, ist es von entscheidender Bedeutung, deren Mechanismen zu verstehen. Stressreaktionen werden in der Regel in einem Labor anhand von Biomarkern wie Cortisol und Alpha-Amylase im Speichel gemessen, deren Analyse komplexe Verfahren erfordert. Frühere Studien haben gezeigt, dass akuter psychosozialer Stress die Körperhaltung und -bewegung einer Person erheblich beeinflussen kann, was eine wertvolle Ergänzung zu den herkömmlichen Stressmarkern darstellt. Bewegungsanalyse wird üblicherweise mit speziellen Systemen durchgeführt, welche typischerweise optische oder inertielle Messeinheiten-basierte Verfahren nutzen. Diese Systeme sind jedoch nur begrenzt zugänglich und erfordern technisches Expertenwissen. Eine mögliche Alternative ist die videobasierte Bewegungsanalyse, da sie sowohl kostengünstig als auch genau ist.

In dieser Bachelorarbeit wird das *openTSST*-Framework vorgestellt, eine webbasierte Plattform, die speziell für die videobasierte Bewegungsanalyse mit Fokus auf akutem psychosozialen Stress entwickelt wurde. Die Plattform steht zur Verfügung unter <https://mad-opentsst.aibe.uni-erlangen.de>. Sie ermöglicht Forschern Videos hochzuladen, die während Stressprotokollen aufgenommen wurden. Die Videoverarbeitung und die Extraktion von Bewegungsmerkmalen werden automatisch in der Cloud durchgeführt, so dass für die Nutzung weder spezielle Hardware, Software oder Fachkenntnisse erforderlich sind. Die Plattform ist modular aufgebaut, um eine einfache Anpassung an neue Forschungserkenntnisse im Zusammenhang mit Bewegung und akuten Stress zu ermöglichen.

Um die Fähigkeiten der Plattform zu demonstrieren, wurde eine frühere Studie von Roos et al. [Roo19] erweitert, in der die Wirkung von Stresshabituation bei wiederholtem akutem psychosozialen Stress untersucht wurde. Unter Verwendung der Aufnahmen aus der Studie wurden Bewegungsparameter mithilfe der Plattform *openTSST* berechnet und geprüft, ob dort ebenfalls Habituationseffekte beobachtet werden können. Die Ergebnisse zeigten keine statistisch signifikanten Habituationseffekte bei den Bewegungsparametern, wie es bei den Cortisol Markern der Fall war. Nichtsdestotrotz zeigt dies die Vielseitigkeit der Plattform. Durch die Erleichterung des Zugangs zur bewegungsbasierten Stressanalyse kann *openTSST* die Verwendung von Bewegungsinformationen neben herkömmlichen Markern für akuten Stress vorantreiben und so zu einer ganzheitlicheren Bewertung der menschlichen Stressreaktion beitragen.

## Abstract

Understanding the underlying mechanisms of acute stress responses is essential to mitigating potential negative health outcomes. Stress responses are typically assessed in a laboratory setting using biomarkers such as cortisol and salivary alpha-amylase, which require complex laboratory procedures to analyse. Prior studies have shown that exposure to acute psychosocial stress can significantly influence an individual's body posture and movement, providing a valuable extension to traditional stress markers. Motion analysis is conventionally performed using optical or inertial measurement unit (IMU)-based motion capture systems; however, these systems come with limited accessibility and demand technical expertise. A possible alternative to these traditional methods is video-based motion analysis, as it is both cost-effective and remarkably accurate.

This Bachelor's thesis presents the *openTSST* framework, a web-based platform designed specifically for video-based motion analysis, with a particular focus on analysing acute psychosocial stress. Available at <https://mad-opentsst.aibe.uni-erlangen.de>, *openTSST* allows researchers to upload videos recorded during stress protocols. Video processing and feature extraction is performed automatically in the cloud, eliminating the need for specialised hardware, software, and expertise to use. The platform is designed in a highly modular way, ensuring seamless updates in line with emerging research insights on motion and acute stress.

To demonstrate the platform's capabilities, a former study by Roos et al. [Roo19] investigating the effect of stress habituation on repeated acute psychosocial stress was extended to assess whether habituation effects could also be observed in freezing-related motion parameters. Using the recordings from the study, motion parameters were computed via the *openTSST* platform. Results showed no statistically significant habituation effects within the motion parameters, as was the case with the saliva markers. Nonetheless, this showcases the platform's versatility, not only applicable to analyse current and future studies but also for retrospective analyses. By democratising access to motion-based stress analysis, *openTSST* can accelerate the incorporation of movement information alongside conventional markers for acute stress, thus contributing to a more holistic assessment of the human stress response.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Background</b>	<b>3</b>
2.1	Human Pose Estimation . . . . .	3
2.2	Video-Based Human Pose Estimation Tools . . . . .	4
<b>3</b>	<b>Related Work</b>	<b>7</b>
3.1	Influence of Internal States on Body Posture and Movement . . . . .	7
3.2	Applications Utilising Video-Based Motion Analysis . . . . .	8
<b>4</b>	<b>Methods</b>	<b>11</b>
4.1	Platform Requirements . . . . .	11
4.1.1	Functional Requirements . . . . .	11
4.1.2	Non-Functional Requirements . . . . .	12
4.2	Platform Architecture . . . . .	13
4.2.1	Processing Gateway . . . . .	14
4.2.2	Worker Cluster . . . . .	18
4.3	Data Processing Pipeline . . . . .	20
4.3.1	OpenPose . . . . .	21
4.3.2	Data Aggregation . . . . .	22
4.3.3	Data Preprocessing . . . . .	24
4.3.4	Feature Extraction . . . . .	26
4.4	Deployment . . . . .	27

<b>5</b>	<b>Evaluation</b>	<b>29</b>
5.1	Study Population . . . . .	29
5.2	Procedures . . . . .	30
5.3	Measurements . . . . .	31
5.3.1	Saliva Markers . . . . .	31
5.3.2	Video Recordings . . . . .	31
5.4	Data Processing . . . . .	32
5.4.1	Saliva Features . . . . .	32
5.4.2	Motion Features . . . . .	33
5.5	Statistics . . . . .	33
<b>6</b>	<b>Results</b>	<b>35</b>
6.1	Saliva Measures . . . . .	35
6.2	Motion Features . . . . .	37
6.2.1	Generic Features . . . . .	37
6.2.2	Expert Features . . . . .	38
<b>7</b>	<b>Discussion</b>	<b>39</b>
7.1	Study Evaluation . . . . .	39
7.2	openTSST Platform . . . . .	40
<b>8</b>	<b>Conclusion &amp; Outlook</b>	<b>43</b>
	<b>List of Figures</b>	<b>45</b>
	<b>List of Tables</b>	<b>47</b>
	<b>Bibliography</b>	<b>49</b>
<b>A</b>	<b>Additional Figures</b>	<b>63</b>
<b>B</b>	<b>Additional Tables</b>	<b>67</b>
<b>C</b>	<b>Acronyms</b>	<b>77</b>



# Chapter 1

## Introduction

Stress has become increasingly associated with adverse health outcomes [OCo21]. It is recognised as one of the leading contributing factors to long-term sickness and has, consequently, been the subject of extensive research [McE93]. Stress is a complex phenomenon that affects an individual in various ways, impacting both psychological and physiological states [Koo11]. At present, the Trier Social Stress Test (TSST) [Kir93] is the gold standard for inducing acute psychosocial stress in a laboratory setting [Dic04]. It involves a public speaking task and a challenging mental arithmetic task that are both performed in front of a panel of evaluators. Key elements of the stressful nature of the TSST are that it includes elements of uncontrollability and social-evaluative threat, which leads to strong responses of the two stress pathways, the sympathetic nervous system (SNS) and the hypothalamic-pituitary-adrenal (HPA) axis. This response is traditionally assessed using biomarkers such as cortisol and salivary alpha-amylase [Goo17; Nat09].

Previous research has shown that exposure to acute psychosocial stress can significantly influence an individual's body posture and movement [Abe22], thus providing a promising extension to existing markers for acute stress. Traditionally, motion analysis is performed using optical or IMU-based motion capturing. Although these techniques offer the best measurement quality, they come with high costs, have limited availability, and can interfere with natural behaviours during stress induction [Col18; Roe09]. A possible alternative is given by video-based motion analysis using tools like OpenPose [Cao19] or AlphaPose [Fan23]. They provide a cost-effective and straightforward way to extract and analyse motion patterns based solely on video data. Despite these advantages, video-based motion extraction and

analysis have not yet been widely adopted for motion analysis in psychology. This is likely due to the limited accessibility of those tools for non-technical researchers. In comparison to facial expression analysis, where tools such as *FaceReader* (Noldus, Wageningen, the Netherlands) provide a convenient solution [Lew14], there is currently a lack of out-of-the-box solutions for motion analysis.

The goal of this Bachelor's thesis is, therefore, to develop the *openTSST* framework, an easy-to-use tool for video-based motion analysis with a particular focus on acute psychosocial stress. The platform should be capable of performing an end-to-end extraction of meaningful motion parameters from video data to assess the influence of acute psychosocial stress on body posture and movements. The platform will be developed with a special focus on a modular architecture, enabling expansion and adaptation of the processing pipelines to meet evolving research requirements.

To demonstrate the capabilities of the *openTSST* framework, a proof-of-concept analysis will be performed on data from an existing study that investigated the effect of stress habituation on repeated acute psychosocial stress [Roo19]. The existing analysis of self-reports and saliva markers will be extended by investigating whether habituation effects can not only be observed in the cortisol response [Kir95], but also in the freezing-related motion parameters.

# Chapter 2

## Background

### 2.1 Human Pose Estimation

The process of pinpointing human keypoints or body parts from input data, such as images and videos, is commonly referred to as human pose estimation. This task has risen to prominence as a substantial challenge within the realm of computer vision and related fields [Che22]. That being said, advancements in motion capture technologies have revolutionised our interpretation and understanding of human movements. They have been employed in a variety of sectors, including sports science, animation, virtual reality, and healthcare [Bri18; Wan21b; Sco22]. The most notable motion capture techniques include optical, IMU-, and video-based motion capture.

Traditional optical motion capture systems operate by arranging a set of synchronised cameras around a designated capturing area. The subject is dressed in a specialised suit fitted with markers positioned on key joints and segments of interest [Gue05]. These systems are capable of pinpointing positions in 3D space with sub-millimetre precision, and they are considered the gold standard when it comes to motion capture [Eic16]. However, optical motion capture systems come with their own set of challenges. They are expensive, require a specialised environment, and have a time-consuming setup process [Gle02]. Furthermore, they can be intrusive and interfere with natural behaviour, as they require markers to be attached to individuals [Col18].

IMU-based motion capture systems are a promising alternative. They utilise a combination of accelerometers, gyroscopes, and, at times, magnetometers to measure body segment

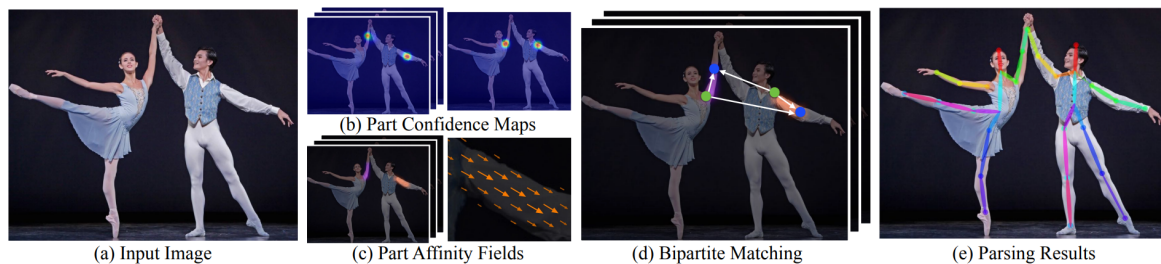
orientation and acceleration [Fil17]. IMU-based systems are portable and relatively inexpensive when compared to optical motion capture systems [Fil17]. They are particularly suitable for capturing movements in sports and outdoor activities. Yet, these systems have their drawbacks as well. They are notably prone to drift over time, which can lead to inaccuracies in motion data when captured over longer periods [Ahm13]. Moreover, for a full-body motion capture using IMUs, it is still necessary to attach a considerable number of sensors to the body [Roe09]. This can be intrusive for the individuals, thus also possibly interfering with natural movements. Even when considering sparse tracking systems [Hua18] that use fewer sensors to record full-body motion the need to strap sensors onto the body still presents some degree of intrusion.

On the other hand, video-based motion capture systems have recently seen a surge in popularity due to advancements in deep learning [Zhe23]. These systems can extract movement data directly from raw video data, making them easy and affordable to use. Additionally, they eliminate the need for markers, thereby avoiding interference with natural movements. Despite ongoing efforts by computer vision researchers to enhance the accuracy of deep learning-based pose estimation algorithms, the accuracy of video-based motion data is still a major concern [Nak20]. Furthermore, these systems can be affected by lighting conditions, and they may struggle with occlusions and complex movements.

## 2.2 Video-Based Human Pose Estimation Tools

State-of-the-art video-based human pose estimation tools are primarily reliant on deep learning-based architectures. They typically employ convolutional neural networks (CNNs) to extract robust features corresponding to each keypoint from input images. CNNs are capable of finding patterns and understanding structures without human supervision [Alz21]. These networks are trained using large, high-quality datasets like the common objects in context (COCO) keypoint detection dataset [Lin15] and the Max Planck Institute for Informatics (MPII) human pose dataset [Fan18].

Current video-based human pose estimation tools differ in their approach to modelling the human body [Wan21a]. The most common methodology employs a kinematic model – also referred to as a skeleton-based model – that is applicable to both 2D and 3D pose estimation. An alternative strategy found in 2D pose estimation employs planar or contour-



**Figure 2.1:** Illustration of the OpenPose processing pipeline [Cao19].

based models [Dub23]. To capture the most amount of details, 3D volumetric models can be employed to generate a detailed surface-based representation of the human body [Wan21a].

OpenPose, AlphaPose, DensePose, and OpenCap have emerged as prominent video-based pose estimation solutions due to their precise detection and inference capabilities, along with their widespread usage across various domains [Zhe23].

### OpenPose

OpenPose<sup>1</sup> was developed at Carnegie Mellon University as a real-time multi-person body, face, hand, and foot keypoint estimation tool [Cao19]. It first employs CNNs to predict confidence maps for body part detection and utilises part affinity fields to associate body parts with individuals in an image. Subsequently, a set of bipartite matching is performed to associate body part candidates, as illustrated in Figure 2.1. This technique allows for highly accurate real-time performance, regardless of the number of people in the image. The OpenPose library is open-source and freely available for non-commercial use. It supports a variety of hardware architectures, such as CUDA (Nvidia GPU), OpenCL (AMD GPU), and non-GPU (CPU-only) devices. OpenPose can provide up to 25 body and foot keypoint estimates, and it can be configured to additionally track up to 2x21 hand keypoints and 70 face keypoints. Images, video files, or cameras are supported as input [Cao19].

### AlphaPose

AlphaPose<sup>2</sup> is also an accurate multi-person pose estimation tool [Fan23]. In contrast to OpenPose, AlphaPose employs a top-down approach to multi-person pose estimation. This means that it first detects individuals inside an image, and then applies pose estimation to

<sup>1</sup><https://github.com/CMU-Perceptual-Computing-Lab/openpose>

<sup>2</sup><https://github.com/MVIG-SJTU/AlphaPose>

detect keypoints of each individual. AlphaPose is capable of detecting face, body, hand, and foot keypoints. It is compatible with the OpenPose keypoint output format, making it a convenient interoperable tool [Fan23].

### **DensePose**

DensePose<sup>3</sup> is a surface-based pose estimation tool [Gül18]. DensePose is built upon a variant of Mask R-CNN, a straightforward and flexible framework for object instance segmentation [He18]. It is a part of Detectron2 [Wu19], which was developed by Facebook’s artificial intelligence research team and it is capable of mapping pixels of RGB images to a 3D surface-based representation of the human body. It has been trained using the DensePose-COCO dataset, which provides the annotations for 50,000 image-to-surface correspondences [Gül18].

### **OpenCap**

OpenCap<sup>4</sup> is an open-source platform designed to compute movement dynamics using videos recorded from a minimum of two smartphones [Uhl22]. After setting up the smartphones on tripods, the platform guides users through a camera calibration process using a checkerboard. OpenCap takes less than five minutes to set up and is then ready for data collection and visualisation of 3D kinematics. There is no need for specialised hardware, software, or expertise, as the platform utilises automated cloud-based processing. OpenCap is capable of accurately predicting dynamic measures, including muscle activations, joint loads, and joint moments. These data points can potentially be used to screen for injury or disease risk, evaluate intervention efficacy, and inform rehabilitation decisions [Uhl22].

---

<sup>3</sup><https://github.com/facebookresearch/detectron2/tree/main/projects/DensePose>

<sup>4</sup><https://github.com/stanfordnmb/opencv-core>

# Chapter 3

## Related Work

### 3.1 Influence of Internal States on Body Posture and Movement

Human body posture and movements can offer valuable insights into the physical and psychological state of an individual [Wal98; Dae12; Bal00]. Emotional information, for instance, is expressed through several channels, including facial expressions, body motion, and posture [Gel15]. Humans naturally interpret a variety of emotions, from joy and happiness to sadness and anger, simply by observing movements [De 06]. Atkinson et al. demonstrated this ability using dynamic and static body expressions, visualised through point-light displays. Participants could reliably differentiate between emotions such as anger, disgust, fear, happiness, and sadness [Atk04].

Additionally, Vrij et al. examined how public self-consciousness and behavioural control influence hand movements during deception. They hypothesised that individuals with high self-consciousness and greater behavioural control would display fewer hand movements while lying. They conducted a study involving 56 participants who were interviewed twice, once truthfully and once deceptively. The results revealed a connection between the participant's personality traits and non-verbal behaviours during deception, emphasising the influence of internal states on physical behaviours [Vri97]. Furthermore, van der Zee et al. utilised a motion capture suit to record full-body movements, with the aim of detecting deceit. They discovered that the sum of joint displacements allowed them to distinguish whether interviewees were lying or telling the truth with 74% accuracy [Zee19].

When exposed to acute stress, humans often exhibit a significant decrease in bodily movements, commonly referred to as freezing [Bra04; Roe17]. Freezing is a well-studied phenomenon and is a common defensive response in animals facing threats from predators [Löw15; Eil05]. Roelofs et al. found that social threat cues could induce freezing behaviour in humans. In their study, 50 female participants stood on a stabilometric force platform while being exposed to various viewing cues. Angry faces notably reduced the participants' body sway [Roe10]. Haagenars et al. observed similar behaviour in response to unpleasant films [Hag14]. Doumas et al. found that exposure to social evaluative threat leads to systematic changes in postural control [Dou18].

The impact of acute psychosocial stress on body posture and movement could be a valuable extension to the current biomarkers for acute stress detection. Abel et al. conducted a study to measure body posture and movement using an IMU-based system during acute psychosocial stress. They gathered data from 41 participants who underwent the TSST and the friendly Trier Social Stress Test (f-TSST) – a friendly version of the TSST that does not activate the HPA axis in healthy individuals [Wie13] – as a control condition in a randomised order on two consecutive days. They computed various motion features from the IMU recorded data. The results indicated that, compared to the control condition, participants exhibited defensive freezing behaviour during acute stress, evident in the reduced head, hand, chest, and overall body motion [Abe22]. Following a similar protocol, Richer et al. conducted a study with 21 participants. They developed a machine learning pipeline, capable of detecting acute stress with an accuracy of 90%, based on the captured IMU data [Ric23].

## 3.2 Applications Utilising Video-Based Motion Analysis

Video-based motion analysis is utilised in a broad spectrum of applications due to the low barrier of entry and recent advancements in computer vision. Video-based human pose estimation tools are widely available and provide high accuracy (Section 2.2).

*FaceReader* (Noldus, Wageningen, the Netherlands) is an application used to decode facial expressions automatically, providing insights into emotional reactions and an objective assessment of emotions. The software uses deep learning-based algorithms to find a person's face, which is subsequently mapped using almost 500 keypoints. By utilising artificial neural networks, the emotions expressed are then classified [Tec23]. The accuracy of *FaceReader* has been validated in several studies. Skiendziel et al. investigated the correlation between



*FaceReader* and the manual facial action coding system (FACS), which scores for 20 action units. FACS enables the objective description of facial expressions based on individual muscle activations. On average, 80% of the expressions assessed by *FaceReader* were correctly classified [Ski19]. Chia-Yin Yu and Chih-Hsiang Ko utilised *FaceReader* to recognise consumer emotions towards graphic styles and found that *FaceReader* could be an effective tool to evaluate consumer emotion in the field of design research [Yu17]. Terzis et al. employed *FaceReader* to measure emotions during self-assessment tests with an efficacy of over 87% [Ter10]. Further studies have demonstrated the capabilities of *FaceReader* to assess emotion in response to tourism advertisements and the perceived healthiness of consumer products [Had19; Pic21].

In the field of sports analysis, video-based methods have become a crucial training tool for coaches, giving them a better understanding of all aspects of an athlete's performance [Ran20; Bar08; Wil08]. As well, training quality often depends on the correct execution of exercises, which can be improved using automated systems [Gar13]. For instance, human pose estimation in combination with artificial intelligence is capable of providing personalised athletic training experiences and of detecting postural misalignments of an athlete's body that need to be corrected [Wan19; Par22]. Movement analysis is also an integral part of physical therapy practice; video-based motion analysis is employed by sports physical therapists to enhance assessment and treatment outcomes [Hen21; Yan18].

Video-based motion analysis has found widespread use in the field of human gait analysis [Sin18]. The technology enables applications such as spinal deformity classification [Sum23] and Parkinson's disease diagnosis by examining bilateral differences between left and right arm swings [Abe21]. Several low-cost video-based tools for clinical gait analysis have already been developed [Sod09; Vis19; DAn20]. Stenum et al. have proposed a versatile workflow that leverages video-based pose estimation to perform gait analyses using videos recorded from multiple perspectives. OpenPose is used to detect keypoints of healthy adults walking overground, and a variety of spatiotemporal and kinematic gait parameters are calculated. This workflow allows researchers to perform quantitative gait analysis using only household devices and computer vision. This marks a major enhancement in the accessibility of clinical gait analysis, which traditionally requires complex measurement setups [Ste21]. Boswell et al. have introduced an approach for self-guided, smartphone-assisted quantitative motion analysis using the five-repetition sit-to-stand test. This experiment was conducted by participants in their own homes across 35 US states, where they recorded themselves with

their personal smartphones sitting and then standing five times. By employing OpenPose to perform human pose estimation, the researchers successfully extracted motion parameters from smartphone videos. They found that these parameters significantly correlated with various factors, including diagnoses of osteoarthritis, indicators of physical and mental health, body mass index, age, race, and ethnicity. This not only demonstrates the feasibility of at-home movement analysis, but also allows for affordable, objective, and digital outcome metrics for large-scale, nationwide studies [Bos23].

# Chapter 4

## Methods

### 4.1 Platform Requirements

The *openTSST* platform is intended to be capable of performing an end-to-end extraction of meaningful motion parameters from video data to assess the influence of acute psychosocial stress on body posture and movements. The platform shall offer a web interface for users to upload videos they recorded during TSST. These videos should then be processed automatically in the cloud. Once the processing is completed, the researcher shall be notified and be able to access the results. A thorough requirement analysis was carried out to ensure the *openTSST* platform is designed to meet researchers' needs and quality expectations. This section outlines the identified functional and non-functional requirements for the platform.

#### 4.1.1 Functional Requirements

The following functional requirements define what the platform must do, and what its features and functions are:

- The *openTSST* platform should provide a web interface that is intuitive and user-friendly, allowing even non-technical researchers to interact with the platform with ease, thus lowering the barrier to entry. Furthermore, the web interface needs to provide a comprehensive usage guide as part of the platform's documentation. This guide shall provide clear instructions on effectively utilising the platform, along with troubleshooting strategies to resolve commonly encountered issues.

- TSST studies typically involve multiple study participants and may encompass several trials per participant, such as tests conducted on consecutive days [Roo19] or variations of the TSST protocol like the f-TSST [Wie13]. To spare researchers the repetitive work of individually uploading all these videos, the platform must be designed to implement a batch processing mode capable of simultaneously uploading multiple videos. Moreover, given that processing can be time-consuming, particularly for large batches of data, the platform must be able to execute the computation asynchronously. The researcher should be able to track the progress of running processing tasks in real-time. Once processing is finished, a mechanism should be in place to notify the researcher that the results are ready, along with a direct link for accessing them.
- The TSST encompasses multiple phases, therefore the platform must provide a mechanism to partition videos into multiple segments, each to be processed individually. Furthermore, it should be possible to trim the beginning and end of each video.
- The platform shall be usable in its full functionality without the necessity of an account. Processing tasks shall be accessible via a unique identifier, ensuring a secure way to access tasks and results without an account. Furthermore, it shall be possible for administrators, authenticated via a login screen, to access an overview of all processing tasks. Admins shall also have the capability to remove old processing tasks and examine their results.
- Given that each study's processing requirements may differ, the platform should support the ability to individually enable or disable processing options beyond the basic computation performed for every task. This not only conserves resources but also enables faster processing.

### 4.1.2 Non-Functional Requirements

The following non-functional requirements specify criteria in regard to the architecture and operation of the platform:

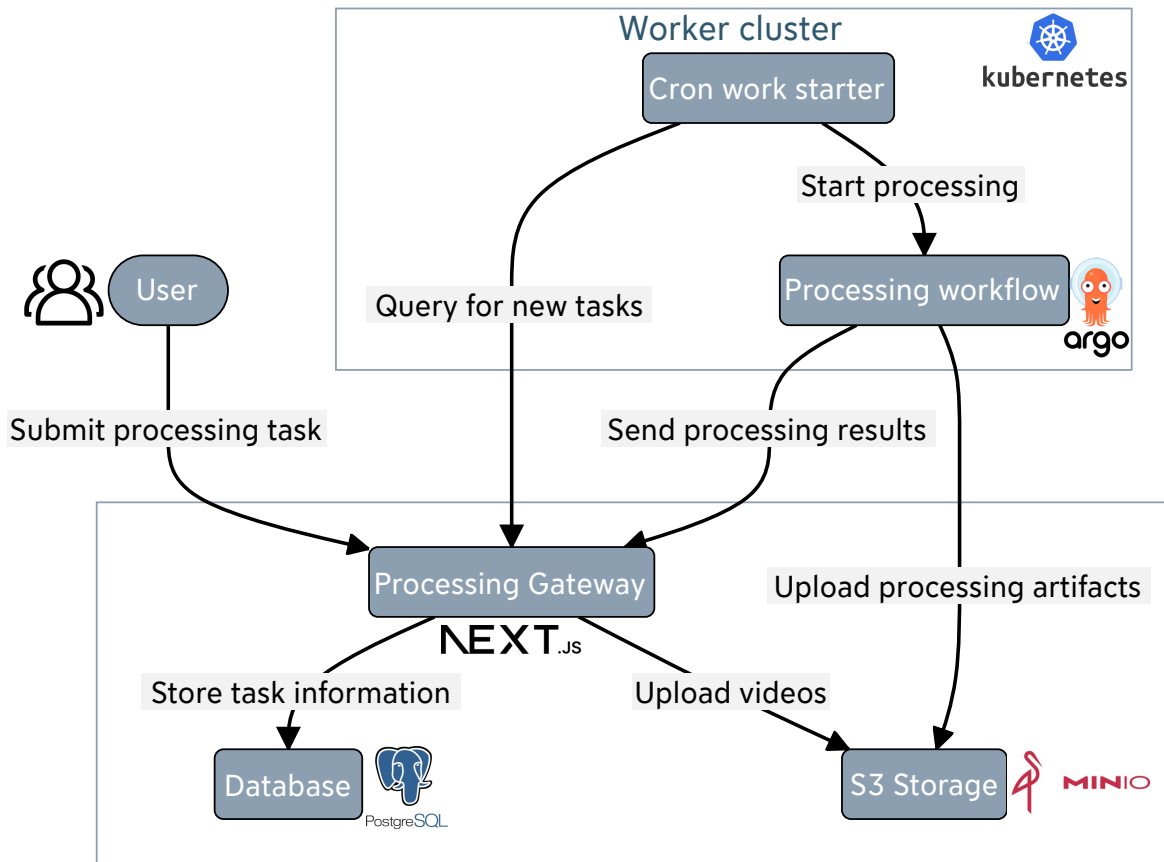
- The *openTSST* platform's web interface must be responsive to user interactions. Due to the high computational load of processing tasks, the system must schedule jobs with minimal overhead, and, even during high load, the user's interactions shall remain

fluent. Furthermore, the platform shall be designed to accommodate a growing number of researchers. The platform must be able to handle growing demands without any compromise in performance. If all available computational resources are in use, the platform shall remain responsive and all additional jobs shall be queued until additional resources are available.

- Given the ever-evolving needs of researchers, the platform's data processing pipeline must feature a modular design. This will enable future adaptations to the data processing pipeline with minimal alterations to the fundamental platform structure.
- Moreover, the codebase should conform to best coding practices and standards, and it should include comprehensive instructions to aid new developers in the project. This will minimise the effort for pipeline modifications, bug fixes, and general enhancements made by new contributors.
- Furthermore, it is important to protect sensitive user data and research findings. The security measures to be employed include but are not limited to secure data transmission and the implementation of robust access control mechanisms.

## 4.2 Platform Architecture

Designing systems that can handle computationally intensive workloads on a large scale is a challenge, so the underlying platform architecture needs to be thought through carefully. The *openTSST* platform is divided into two primary components. The first component is the *processing gateway* that provides both a web interface for the user as well as the business logic for coordinating processing tasks. The second component is the *worker cluster*, which is specifically designed to process compute-intensive workloads. Video files and processing artefacts are stored in a simple storage service (S3) instance. A comprehensive schematic representation of the platform architecture can be found in Figure 4.1.



**Figure 4.1:** High-level *openTSST* platform architecture overview.

### 4.2.1 Processing Gateway

The primary responsibility of the *processing gateway* is to offer a web-based interface for user interaction with the platform. It also manages all logic for scheduling processing workflows, sending user notifications, and storing persistent workflow states. The gateway is built in TypeScript using the Next.js<sup>1</sup> framework, a powerful framework for the development of full-stack web applications. The integration of web interface (*frontend*) and business logic (*backend*) within a single codebase significantly accelerates the development process and enhances its efficiency. To facilitate the exchange of data between the *frontend* and the *backend*, the library tRPC<sup>2</sup> is employed. This fully typesafe library allows the construction of application programming interfaces (APIs) without schemas or code generation.

<sup>1</sup><https://nextjs.org/>

<sup>2</sup><https://github.com/trpc/trpc>

**Table 4.1:** Processing gateway environment variables overview.

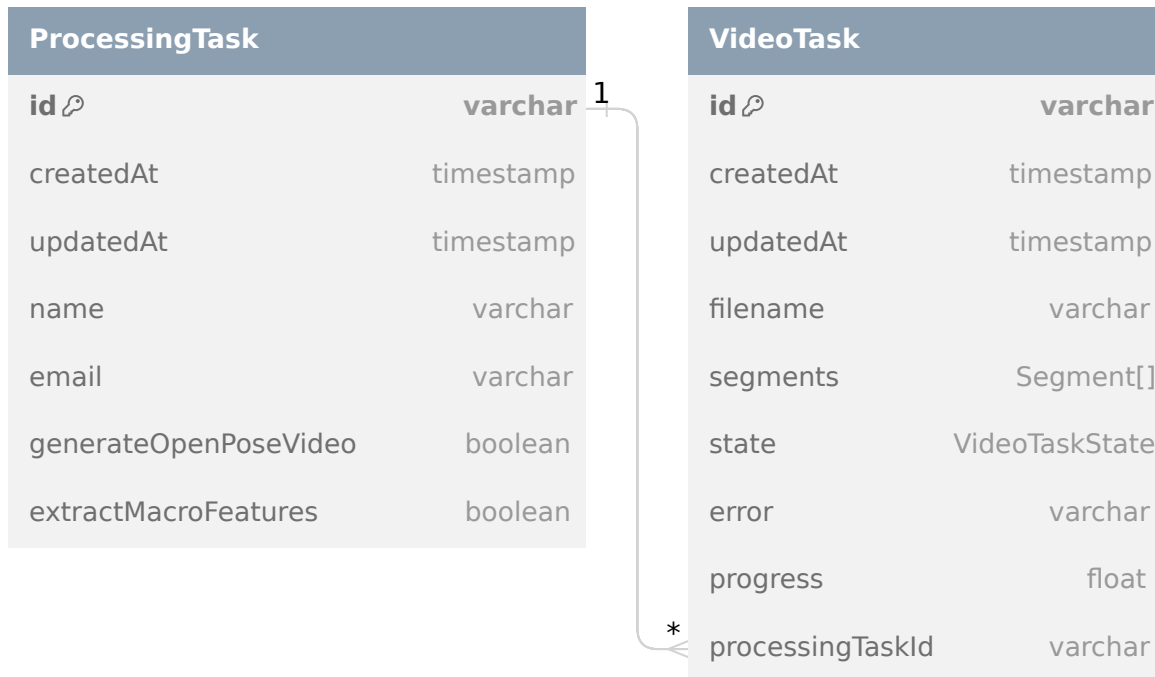
Name	Explanation
DATABASE_URL	PostgreSQL database URL
S3_REGION	S3 Region
S3_ACCESS_KEY	S3 Access Key
S3_SECRET_KEY	S3 Secret Key
NEXT_PUBLIC_S3_ENDPOINT	S3 Endpoint
NEXT_PUBLIC_S3_BUCKET	S3 Bucket
EMAIL_ADDRESS	<i>openTSST</i> Email Address
EMAIL_PASSWORD	<i>openTSST</i> Email Password
DEPLOY_URL	Deployment URL
INTERNAL_API_KEY	Internal API Key
ADMIN_PASSWORD	Admin Password
TOKEN_SECRET	JWT Token Secret

To inject configuration parameters as well as sensitive information – such as access tokens – into the application, environment variables are employed. Table 4.1 provides a comprehensive list of all required environment variables. Environment variables that begin with the prefix `NEXT_PUBLIC_` are also exposed on the client side.

Users can submit TSST videos in the form of a processing task. A processing task includes information such as name, email, processing options, and a batch of videos. Each video file within the batch triggers the creation of an individual video task, associated with the processing task. The processing status of each video task is stored in the dedicated state field. While the task is being processed, its real-time progress is tracked through the progress field. This granular approach is crucial as internal processing occurs at the level of individual video tasks and not at the level of aggregated processing tasks. The system stores persistent processing task information in a PostgreSQL<sup>3</sup> database. An illustrative representation of the database schema is provided in Figure 4.2. The *processing gateway* utilises Prisma<sup>4</sup> as a tool for managing migrations and orchestrating the object-relational mapping (ORM) layer.

<sup>3</sup><https://www.postgresql.org/>

<sup>4</sup><https://www.prisma.io/>



**Figure 4.2:** Processing gateway database models.

The web interface is constructed using React<sup>5</sup>, a popular JavaScript library for creating dynamic web interfaces. The use of Tailwind CSS<sup>6</sup> in combination with Flowbite<sup>7</sup> – an open-source library with over 600 Tailwind components – covers the aesthetic and intuitive styling requirements of the web interface. Moreover, to adapt to user preferences, a dark mode has been integrated. Screenshots of selected interfaces can be found in Appendix A.

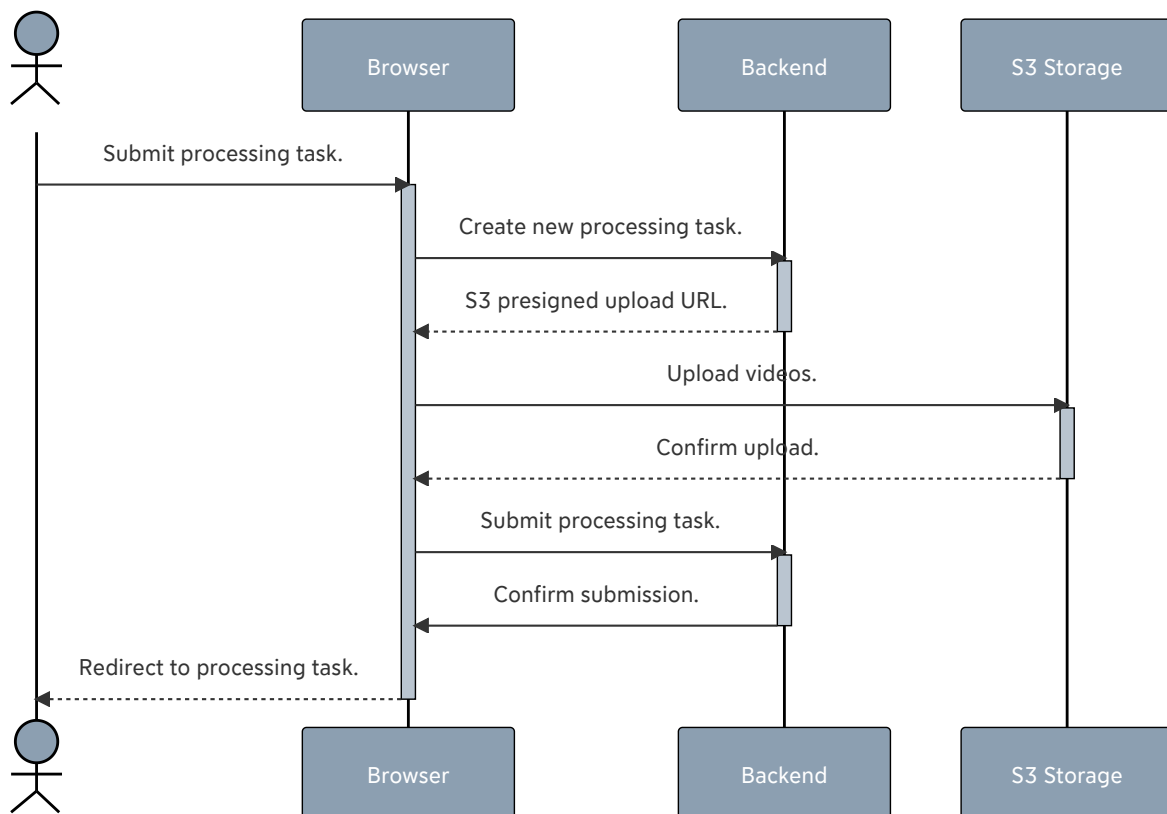
The web interface itself consists of three primary pages. The first page is the processing task submission page, which enables users to create new processing tasks. Users can assign an alias for the task and an email address for receiving updates. Additionally, optional processing steps can be selected via the available checkboxes. Users can then either select or drag and drop all study videos intended for processing into a designated upload area. Upon selecting each video, the video segmentation interface opens. This interface permits users to partition and trim the video into multiple segments through a tailored video timeline interface.

<sup>5</sup><https://react.dev/>

<sup>6</sup><https://tailwindcss.com/>

<sup>7</sup><https://flowbite.com/>





**Figure 4.3:** Processing task submission flow.

Upon submission, the initial information, which includes the task alias, email, processing options, video filenames, and segmentation data, is transmitted to the *backend* server. This information is then stored in the database, and a unique identifier for the task, along with a pre-signed URL for S3 storage, is generated and returned. This pre-signed URL enables the client to upload files to a pre-determined S3 bucket, only allowing access to files that begin with a specific key prefix. In this context, the processing task ID serves as the prefix. File uploads are restricted to a maximum file size of 1 GB per file. Once the browser receives the pre-signed URL, it triggers the simultaneous upload of all study videos. Upon completion of the upload, the client dispatches a submission request to the *backend* server, thereby marking the processing task as ready for processing. Throughout this process, users are presented with an informative loading screen that provides progress bars and updates. Once the task has been successfully submitted, the user is redirected to the detail page of the newly created processing task. Figure 4.3 further illustrates the workflow of this submission process.

The second page, the processing details page, enables users to monitor the status of their processing tasks and view the results, should they already be available. Users also receive an overview of all selected processing options. For each submitted video associated with the task, they can view processing artefacts, such as the extracted movement features, the OpenPose pose estimation output as pandas<sup>8</sup> DataFrame, and the rendered OpenPose overlay video. If the results are not yet available, the page presents the current status of the processing task. If the processing has already started, the user can monitor the progression of the processing steps in real-time. In the event of an error, the corresponding error message is displayed. The system periodically updates the page to load new information as it becomes available.

The last page of the website features an admin panel that enables administrators to have an overview of all the created processing tasks, and, if necessary, to delete them. JSON web tokens (JWTs) are used to authenticate admin users. Users can initiate a login request via the web interface. If the provided credentials are valid, the *backend* responds with a signed JWT, which is stored in the browser's local storage. This token is transmitted as a Bearer Token [Jon12] in the "Authorization" header to authenticate requests to the *backend* server. Logout functionality, which removes the JWT from the browser's local storage, is provided.

Lastly, the interface features a landing page that provides a short description of the platform and a "Getting Started" guide. Beyond this, a comprehensive user guide is available. This guide explains the platform's capabilities in more detail and offers instructions to users for optimal utilisation of the platform. It also includes troubleshooting advice to address frequently encountered issues.

Furthermore, the *processing gateway* offers internal API endpoints for facilitating communication with the *worker cluster*. These endpoints enable the cluster to query for new video tasks, as well as to update the status and process of existing tasks. The internal API endpoints are safeguarded via an internal API key, transmitted as a Bearer Token [Jon12].

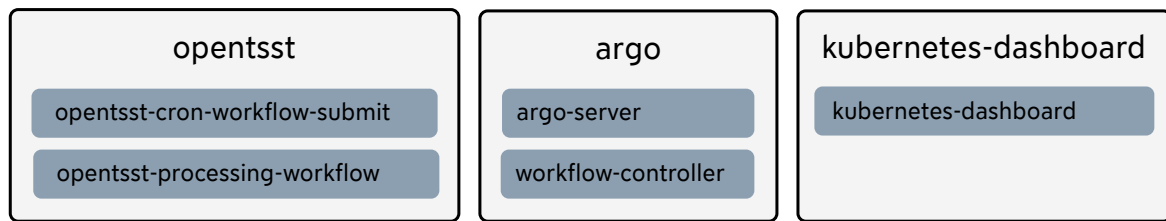
## 4.2.2 Worker Cluster

The *worker cluster*, responsible for processing the actual video tasks, operates within a Kubernetes<sup>9</sup> environment. Kubernetes is a declarative container orchestration framework designed for deploying and scaling containerised applications. It has a large and rapidly

---

<sup>8</sup><https://pandas.pydata.org/>

<sup>9</sup><https://kubernetes.io/>



**Figure 4.4:** Illustration of selected Kubernetes deployments and workflows in their respective namespaces in the worker cluster.

growing ecosystem, and it offers a wide variety of services, support, and tools [Kub23]. Kubernetes is also portable, allowing for local deployments as well as for deployments on all major cloud providers. This is especially useful for the *worker cluster*, as it can be deployed both on a local machine for testing purposes, but also on a cloud provider like Amazon Web Services (AWS) for production deployments.

Certain compute-intensive tasks executed by the *worker cluster* require graphics processing unit (GPU) resources. Therefore, the *worker cluster* must be deployed on a GPU-equipped machine. As such, the Nvidia device plugin<sup>10</sup> has been configured to enable Kubernetes deployments on GPU-equipped machines. This plugin enables the usage of GPU resources in Kubernetes, allowing containers to request GPU resources. The Nvidia device plugin is configured as a daemonset, ensuring that each node in the cluster has access to GPU resources.

The heart of the *worker cluster* is the *data processing pipeline* (Section 4.3), constructed based on Argo workflows<sup>11</sup>. Argo workflows is an open-source container-native workflow engine for orchestrating parallel jobs on Kubernetes, making it an ideal choice for running compute-intensive workflows. Especially when deployed on a cloud Kubernetes engine like AWS Elastic Kubernetes Service, compute resources can scale up or down depending on the current workload, which enables high throughput while maintaining cost-effectiveness. A high-level overview of the Kubernetes cluster configuration is illustrated in Figure 4.4.

To initiate new data processing workflows for a video task, a cron job is configured in the *worker cluster* to query the *processing gateway* every 60 seconds. The job queries all video tasks submitted to the *processing gateway* by a user and transmutes their processing state to “Started” in one transaction, thereby avoiding race conditions. A new Argo workflow that runs the data processing pipeline is then started for each task. The processing options

<sup>10</sup><https://github.com/NVIDIA/k8s-device-plugin>

<sup>11</sup><https://argoproj.github.io/argo-workflows/>

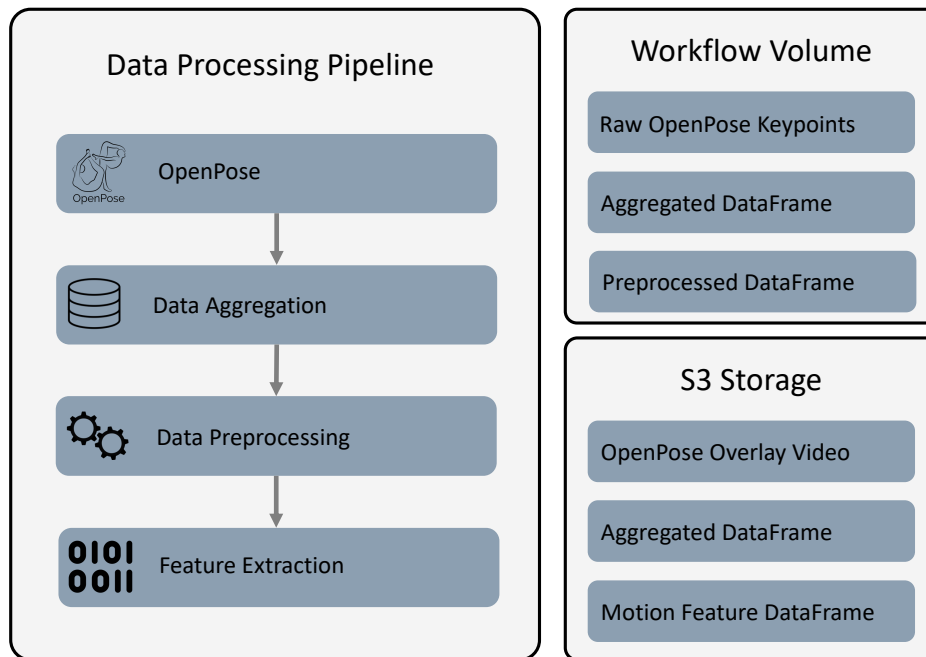
submitted by the user are passed into the workflow as parameters. Any errors encountered during this process are reported back to the *processing gateway*. This pull-only approach allows deployments of the *worker cluster* behind firewalls without the necessity of opening inbound ports.

Monitoring the *worker cluster* is facilitated by the Argo user interface (UI) and the Kubernetes dashboard. The Argo UI provides an in-depth overview of both ongoing and past workflows. It allows for the inspection of specific tasks, their log file and real-time tracking of workflow statuses. Additionally, the Kubernetes dashboard service is configured to inspect the Kubernetes cluster, providing a low-level overview of all resources inside the cluster and conveying useful information about the health of Argo system pods. However, port forwarding is necessary to access the Argo UI or the Kubernetes dashboard, as these services are not exposed via an ingress – a Kubernetes resource to expose routes from outside the cluster to services within the cluster – for security reasons.

### 4.3 Data Processing Pipeline

The data processing pipeline is composed of multiple steps running inside an Argo workflow. This architecture was chosen to enable a highly modular processing pipeline design. All processing steps running in the pipeline are executed in a container, allowing for greater flexibility. While a data processing workflow is active, a temporary shared workflow volume is created, which can be accessed by each container running in the pipeline. This volume is used to share processing artefacts with other steps in the workflow. Argo workflow also enables artefact upload to a configured S3 provider.

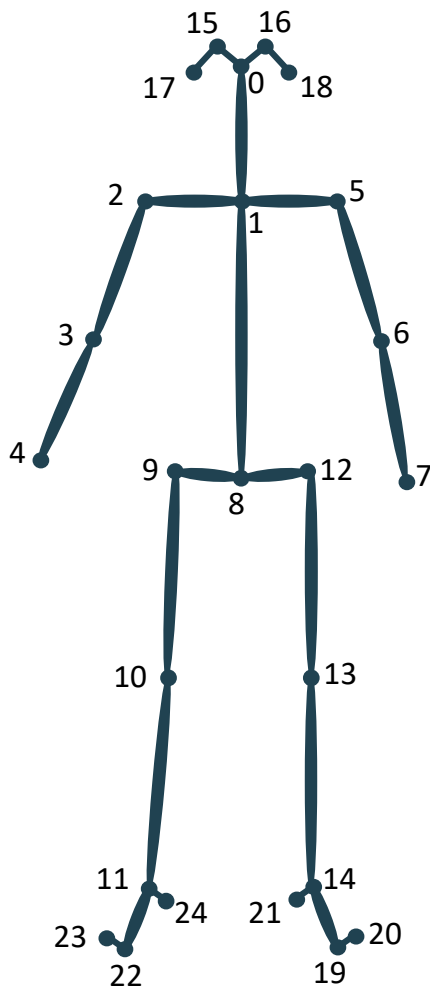
At the start of the pipeline, a video analysis container is run to extract information, such as frames per second and video duration, from the video file in order to facilitate subsequent processing steps. A visualisation of the main data processing pipeline steps, along with the generated artefacts, is provided in Figure 4.5. The processing is mainly based on a master's project by Liv Herzer. Following each pipeline step, an update regarding the video task's progress is transmitted to the *processing gateway*. Moreover, while the OpenPose container is running, the progress of the OpenPose motion extraction is reported at five-second intervals. Upon completion of the pipeline, the exit handler notifies the *processing gateway* and either reports success or relays error messages.



**Figure 4.5:** Illustration of the main data processing steps, along with generated artefacts.

### 4.3.1 OpenPose

OpenPose has been chosen to perform pose estimation, as it delivers some of the highest accuracy, is simple to use, and is flexible. OpenPose also allows the rendering of a skeleton overlay video as well as the generation of JavaScript object notation (JSON) keypoint output. The OpenPose output format is supported by a wide variety of tools. An overview of the keypoints extracted by OpenPose can be found in Figure 4.6. OpenPose processing is executed in a custom Docker container, equipped with a precompiled version of OpenPose. The system initially downloads the raw video, which the user uploaded to S3, to the local container into a temporary directory. Subsequently, OpenPose is utilised to conduct motion analysis on the raw video. The task saves the resulting keypoints in JSON format to the shared workflow volume. Additionally, if the “generateOpenPoseVideo” workflow parameter specifies it, a video with an overlaid OpenPose skeleton is rendered. The task then uploads the rendered video to S3 for user access. To expedite OpenPose processing significantly, the OpenPose processing operates on a GPU. Depending on the number of available GPUs, a semaphore within the Argo workflow template constrains the number of OpenPose processing tasks that can run concurrently.



No.	Keypoint
0	Nose
1	Neck
2	Right Shoulder
3	Right Elbow
4	Right Wrist
5	Left Shoulder
6	Left Elbow
7	Left Wrist
8	Mid Hip
9	Right Hip
10	Right Knee
11	Right Ankle
12	Left Hip
13	Left Knee
14	Left Ankle
15	Right Eye
16	Left Eye
17	Right Ear
18	Left Ear
19	Left Big Toe
20	Left Small Toe
21	Left Heel
22	Right Big Toe
23	Right Small Toe
24	Right Heel

**Figure 4.6:** OpenPose BODY\_25 body part mapping.

### 4.3.2 Data Aggregation

This step aggregates keypoint data within a Python-based Docker container. The OpenPose keypoints, generated in the prior stage, are fetched from the shared workflow volume and transformed into a single Pandas DataFrame. This DataFrame provides a structured and intuitive representation of the OpenPose output data, thereby simplifying successive data analysis or machine learning tasks. The resulting DataFrame is stored as a CSV file on the shared workflow volume for accessibility in later steps and is also uploaded to the S3 for user access. The data aggregation algorithm is outlined as follows.

### 1. JSON Data Extraction

Each JSON keypoint file in the specified directory is processed such that the frame number is parsed from the filename and the keypoint data is extracted from the JSON structure. Files lacking any keypoint data are omitted. Files that contain data for more than a single detected individual are flagged and the corresponding keypoints are disregarded (set to NaN).

### 2. Keypoint Processing

The extracted keypoints are restructured into a two-dimensional array of the shape  $(n, 3)$ , representing  $x$ ,  $y$ , and their confidence values respectively. Keypoints with a confidence score of 0, signifying an uncertainty in detection, have their  $x$  and  $y$  coordinates substituted with NaN.

### 3. DataFrame Construction

The processed keypoints are subsequently transformed into a pandas DataFrame with a hierarchical column index structure. The columns are orchestrated at two levels: the first corresponds to the body part, while the second specifies the axis ( $x$ -coordinate,  $y$ -coordinate, and confidence score). An extra column is added, containing a flag indicative of the presence of more than one person in the frame. The DataFrame is indexed by the frame number.

### 4. DataFrame Aggregation and Reorganisation

DataFrames for all frames are concatenated into a singular DataFrame and sorted by frame number and body part. The 25 body part integer indices are then mapped to their associated human body keypoint labels such as “Nose”, “Neck”, “RightShoulder”, etc. Subsequently, the column level “channel” is appended to the DataFrame columns. The column levels of the DataFrame are rearranged to place the “channel” level between the “body\_part” and “axis” levels. Initially, each column’s value for the level channel is set to “pos”.

### 4.3.3 Data Preprocessing

Following the initial aggregation of data, further preprocessing of the DataFrame is carried out. This stage includes data cleaning, normalisation, and transformation. Upon completion, the DataFrame is subsequently stored again as a CSV file within the shared workflow volume.

#### 1. Data Interpolation and Filtering

For instances where OpenPose fails to accurately detect the participant’s poses for a brief period, linear interpolation is applied to estimate missing keypoints. Any data gaps exceeding half a second in duration are eliminated (set to NaN). Additionally, to reduce the impact of high-frequency, low-magnitude noise in the OpenPose data, a 6 Hz, fifth-order, zero-lag, low-pass Butterworth filter is applied. The resulting DataFrame is denoted as  $df_{inter}$ .

#### 2. Normalisation by Body Height

The interpolated data is then normalised based on the participant’s body height. This normalisation is essential to compensate for variations in camera orientations across different trials, thereby ensuring the comparability of measurements. First, the body height is computed as the 95th percentile of the vertical distance between the keypoints “Nose” ( $y_{upper}$ ) and “RightBigToe” ( $y_{lower}$ ) throughout the data sequence (Equation 4.1). This calculation method mitigates the effects of occasional erroneous measurements. Subsequently, every coordinate in the DataFrame is divided by this height factor to achieve normalisation according to Equation 4.2. In this normalised DataFrame, the channel is consequently renamed from “pos” to “pos\_norm”.

$$height = \text{quantile}(y_{lower} - y_{upper}, q = 0.95) \quad (4.1)$$

$$df_{norm} = \frac{df_{inter}}{height} \quad (4.2)$$

#### 3. Relative Position to Neck Calculation

The position of each body part relative to the neck is determined by subtracting the neck’s coordinates from the coordinates of each body part in the normalised DataFrame (Equation 4.3).



$$df_{rel} = df_{norm} - df_{neck} \quad (4.3)$$

Where  $df_{neck}$  denotes the coordinates for the “Neck” body part. The channel attribute in this DataFrame is renamed from “pos” to “pos\_rel2Neck”.

#### 4. Wrist Relative to Elbow Position

The position of each wrist in relation to its corresponding elbow is calculated. This process is analogous to the previous step, but it specifically targets the “Wrist” body part relative to the “Elbow” body part for both the “Left” and “Right” sides.

#### 5. Data Concatenation

The resulting DataFrames from the preceding steps (interpolated, normalised, body parts relative to neck, and wrists relative to elbows) are concatenated column-wise to form a single DataFrame. The columns of this DataFrame are then restructured using a hierarchical column structure, with the top level named “data\_format”, and the other levels named “body\_part”, “channel”, and “axis”. The newly introduced “data\_format” level is set to “open\_pose” for every column.

#### 6. Velocity computation

Velocities for each channel of the concatenated DataFrame are computed. Three kinds of velocities are computed: the velocities in the x- and y-dimensions, along with a 2-dimensional velocity.

The 2-dimensional velocity is determined by applying the Euclidean norm to the difference in position, which is then divided by the elapsed time. This computation effectively yields the velocity measured in units of pixels per frame. Velocities for the x and y components are calculated similarly, leveraging the respective differences in the x and y coordinates of the position.

Finally, all “pos” strings in the original channel names are substituted with “vel” and are added to the original DataFrame.

### 4.3.4 Feature Extraction

Following the preprocessing of data, a variety of features are calculated, similar to previous work by Abel et al. [Abe22]. Features are computed for individual body parts. Additionally, several body parts are grouped to reflect movements of specific body areas (Table 4.2). The calculated features can be divided into two categories: generic features and expert features.

Generic features do not require prior domain knowledge. They include basic statistical measures such as mean and standard deviation. Signal characteristics like entropy are also included in generic features (Table 4.3). In contrast, expert features are based on prior knowledge. They are designed specifically to represent movement patterns that have been identified in previous studies or observed during data collection.

**Table 4.2:** Definition of body part groups. {L/R} denotes the left and right sides.

Group	Body Parts
Trunk	{L/R} Hip, {L/R} Shoulder, Neck
Upper Extremities	{L/R} Shoulder, {L/R} Elbow, {L/R} Wrist
Lower Extremities	{L/R} Knee, {L/R} Ankle, {L/R} BigToe, {L/R} SmallToe, {L/R} Heel
Total Body	All body parts

**Table 4.3:** Overview of computed generic features.

Name	Short Form
Mean	mean
Standard deviation (SD)	std
Entropy	entropy
Absolute Energy	abs_energy
Mean Crossing	mean_crossing
Zero Crossing	zero_crossing
FFT Aggregated Centroid	fft_aggregated_centroid
FFT Aggregated Kurtosis	fft_aggregated_kurtosis
FFT Aggregated Skew	fft_aggregated_skew
FFT Aggregated Variance	fft_aggregated_variance

## 4.4 Deployment

The *openTSST* platform is deployed on servers operated by the *Machine Learning and Data Analytics Lab*. The deployment is accessible at <https://mad-opentsst.aibe.uni-erlangen.de>. A MinIO<sup>12</sup> S3 server with a private bucket for the *openTSST* deployment is utilised to store raw video files and processing artefacts. It is accessible exclusively within the university's private network. Consequently, users are required to connect to the university's virtual private network (VPN) to utilise the *openTSST* platform. The platform's *processing gateway* is deployed on a web server using a docker-compose setup. This setup includes a PostgreSQL database to store persistent processing task information and the Next.js server. Additionally, the *worker cluster* is deployed on a workstation equipped with an Nvidia Titan Xp GPU, as it is responsible for executing computationally intensive tasks. Similar to the MinIO server, the *worker cluster* is located within the university's private network.

---

<sup>12</sup><https://github.com/minio/minio>



# Chapter 5

## Evaluation

To demonstrate the *openTSST* platform’s capabilities, a proof-of-concept analysis was performed on a previously recorded study [Roo19]. Participants were asked to take part in the TSST on two consecutive test days.

### 5.1 Study Population

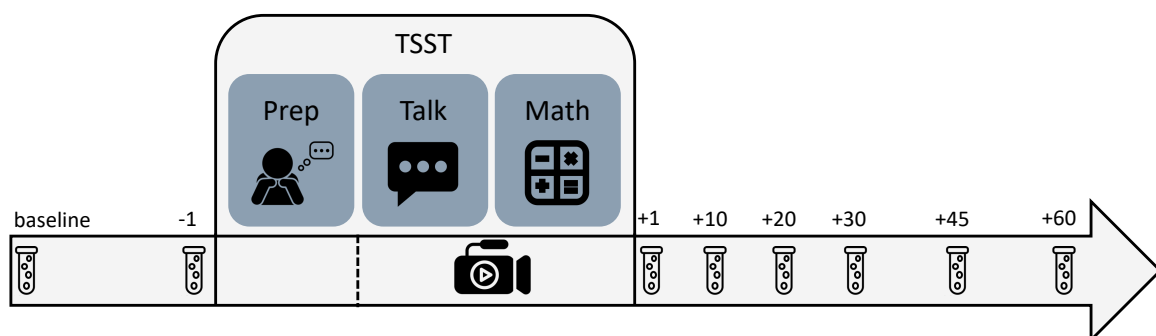
A total of 160 participants, comprised of 108 females and 52 males, participated in this study. Table 5.1 presents a summary of the participants’ demographic and anthropometric data. Participants were recruited via print and multi-media advertising. An online survey tool was utilised for the screening of potential participants. To be considered eligible for the study, participants had to meet the following criteria: age between 18 and 35 years, Body Mass Index between 18 and 30 kg/m<sup>2</sup>, non-smoking, not dependent on alcohol or illicit drugs, absence of medication use, and no major psychiatric disorders or symptoms evident during the screening. As compensation for participation in the study, participants received 45 EUR for their time.

**Table 5.1:** Demographic and anthropometric data of the participants. Mean  $\pm$  SD.

	Age [years]	Weight [kg]	Height [cm]	BMI [kg/m <sup>2</sup> ]
<b>Female</b>	25.14 $\pm$ 7.03	62.34 $\pm$ 8.56	167.55 $\pm$ 5.87	22.17 $\pm$ 2.49
<b>Male</b>	25.81 $\pm$ 5.85	76.25 $\pm$ 10.71	181.73 $\pm$ 7.72	23.04 $\pm$ 2.53
<b>Total</b>	25.36 $\pm$ 6.66	66.86 $\pm$ 11.35	172.16 $\pm$ 9.31	22.45 $\pm$ 2.53

## 5.2 Procedures

Participants were scheduled to visit the laboratory on two consecutive test days. In order to control for the impact of circadian cortisol rhythm, the study was conducted between the hours of 14:00 and 18:00. Upon arrival, baseline saliva samples were collected. After a rest period of 45 minutes, an additional saliva sample was collected and participants were introduced to the TSST [Kir93], which is recognised as the gold standard method for inducing psychosocial stress in a laboratory setting [Dic04]. During the TSST, participants are escorted into a testing room where they are met with a committee composed of one male and one female member. This committee serves as the audience for evaluating the participant's performance. The TSST comprises two main tasks: public speaking and a verbal arithmetic task. Initially, participants are given five minutes to prepare a speech to persuade the committee members that they are the ideal candidate for their dream job. Subsequently, participants are instructed to deliver this five-minute speech in front of the committee. Notably, the committee refrains from providing any form of verbal or nonverbal encouragement or feedback during the TSST, which enhances the stressfulness of the situation. Following the speech, participants are asked to participate in a serial subtraction task for an additional five minutes. They are instructed to subtract a specific number from a given starting number (17 from 2043 on day 1 and 13 from 2011 on day 2) as accurately and swiftly as possible. If participants make an error, they are instructed to start over from the beginning. Both the speech and arithmetic portions of the TSST were video-recorded. Furthermore, saliva samples were collected at +1, +10, +20, +30, +45, and +60 minutes after the TSST [Roo19]. The procedure of a single test day is illustrated in Figure 5.1.



**Figure 5.1:** Overview of the study procedure of a single test day.

## **5.3 Measurements**

### **5.3.1 Saliva Markers**

The response of the HPA axis was assessed using salivary cortisol concentrations. Salivary samples were collected using Salivettes (Sarstedt, Nümbrecht, Germany) and preserved for subsequent evaluation at -30 °C post-laboratory sessions. The salivary samples were then centrifuged at 2000g for ten minutes at 20 °C and cortisol concentrations were evaluated in duplicate using a chemiluminescence immunoassay (CLIA, IBL International, Hamburg, Germany). The intra-assay and inter-assay CVs were below 10% [Roo19].

### **5.3.2 Video Recordings**

In accordance with the TSST protocol, video data was captured during the speech and arithmetic phases of the TSST procedure. The data was recorded with a sampling rate of 30 Hz and stored as MP4 files. The recordings were trimmed to only include the relevant test duration. The framing of video captures was aligned to fully encapsulate the upper body including the hip of the participant. Lower extremities were not visible in the recordings. Individuals were positioned in front of a white wall with their respective participant number visible in the background.

## 5.4 Data Processing

### 5.4.1 Saliva Features

To evaluate the activity of the HPA axis, four saliva features were calculated from the raw cortisol samples. In the following,  $t_i$  denotes the time at each measurement point, while  $S_i$  represents the cortisol level at the corresponding point. The initial saliva sample ( $S_0$ ) was excluded from further analysis, as its primary purpose was for baseline comparisons and potential participant exclusion from the study. The raw samples ( $S_1$ - $S_7$ ) were then processed to compute four features:  $AUC_G$ ,  $AUC_I$ ,  $\Delta c_{max}$ , and  $a_{14}$ .

$AUC_G$  and  $AUC_I$  are variations of the area under the curve ( $AUC$ ), as proposed by Pruessner et al. [Pru03]. The  $AUC_G$ , which denotes the area under the curve with respect to the ground, was calculated as follows:

$$AUC_G = \sum_{i=1}^6 \frac{(S_{i+1} + S_i) \cdot (t_{i+1} - t_i)}{2} \quad (5.1)$$

Additionally, the  $AUC_I$ , representing the area under the curve with respect to the increase, was computed as follows:

$$AUC_I = \sum_{i=1}^6 \frac{(S_{i+1} + S_i) \cdot (t_{i+1} - t_i)}{2} - S_1 \cdot (t_7 - t_1) \quad (5.2)$$

Furthermore, the maximum increase in cortisol ( $\Delta c_{max}$ ) was quantified as the difference between the peak cortisol level following the TSST and the cortisol level recorded prior to the TSST ( $S_1$ ):

$$\Delta c_{max} = \max\{S_2, \dots, S_7\} - S_1 \quad (5.3)$$

Finally,  $a_{14}$ , the slope of the cortisol level from  $S_1$  to  $S_4$ , was calculated as:

$$a_{14} = \frac{S_4 - S_1}{t_4 - t_1} \quad (5.4)$$



### 5.4.2 Motion Features

Motion features were computed as described in Section 4.3. Given that the available video data solely depicts the upper body, the normalisation strategy described in Step 2 was revised to normalise between the body parts “Nose” and “MidHip”. Measures from the lower body parts were consequently disregarded.

## 5.5 Statistics

Initially, a statistical analysis was conducted to confirm the cortisol habituation effect. Subsequent to this, the extracted motion features underwent statistical evaluations to determine if the habituation effect was discernible in the motion features as well.

A test for normal distribution via the Shapiro-Wilk test unveiled nonconformity for certain features. Thus, Wilcoxon signed-rank tests were executed on all features, with the condition as a between-variable, as all participants were subjected to both conditions (day 1 and day 2). The statistical analysis was performed using the Python package *biopsykit* [Ric21], based on *pingouin* [Val18]. The significance level was set at  $\alpha = 0.05$  and effect sizes were reported as Hedge’s  $g$ . In an attempt to correct for the multiple comparisons problem, Bonferroni corrections were employed across all tests. All figures and tables utilise the following notation to denote statistical significance: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .



# Chapter 6

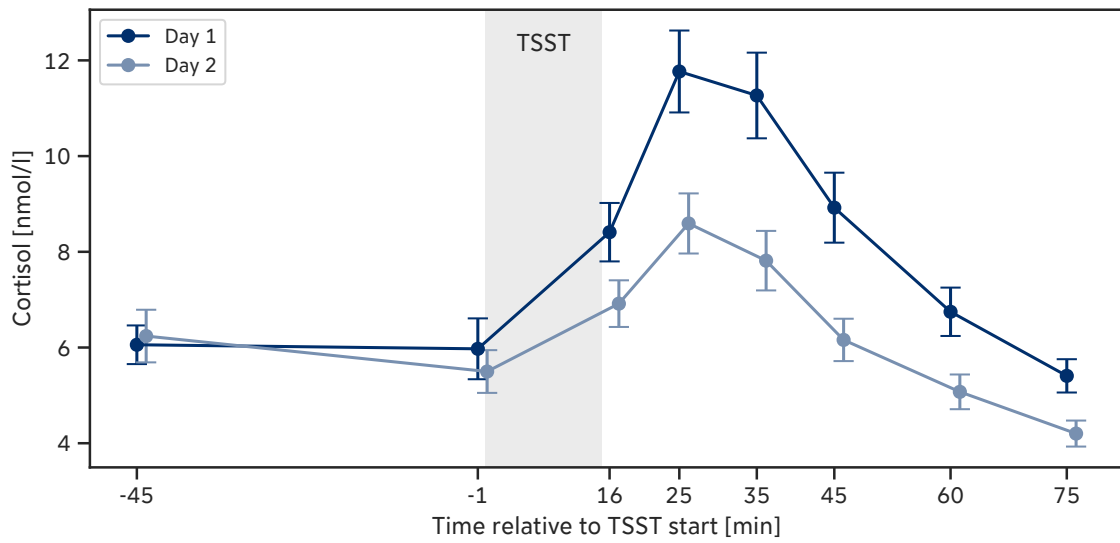
## Results

This chapter presents the results obtained from the analysis of salivary measures and motion features. From the initial cohort of  $N = 160$  participants, a total of 51 were eliminated due to the absence of video recordings from one of the test days or both test days. Additionally, misalignments in camera framing led to computational errors in the feature extraction, resulting in the additional exclusion of 29 participants. Consequently, the participant count for the subsequent statistical analysis was reduced to  $N = 80$ , comprised of 49 females and 31 males.

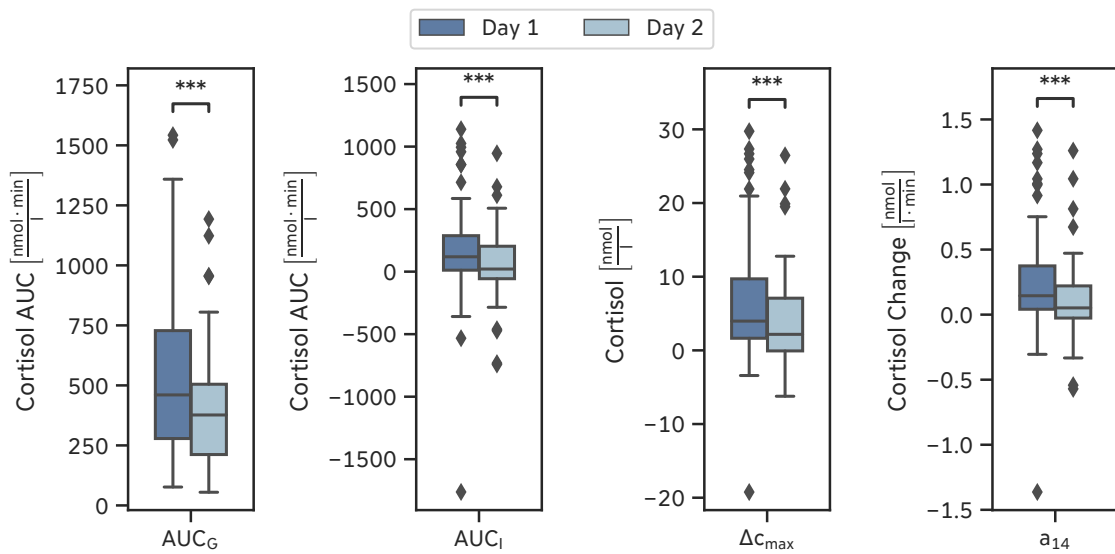
### 6.1 Saliva Measures

Exposure to the TSST on the first day precipitated an approximate doubling of cortisol levels subsequent to the TSST, as seen in Figure 6.1. In particular, an increase of 97% in mean cortisol levels was noted when comparing the measurement taken immediately prior to the start of the TSST ( $S_1$ ) with the peak measurement ( $S_3$ ) collected 10 minutes after the TSST. The second day showed a comparatively modest 56% increase in mean cortisol levels from  $S_1$  to  $S_3$ , suggesting that cortisol response habituation had occurred. These findings are in agreement with the prior analysis conducted on this study by Roos et al. [Roo19].

Furthermore, the statistical analysis of the calculated cortisol features revealed a significant difference across all computed features between first and second TSST exposure. The results of this statistical analysis are presented in Table 6.1. The largest effect sizes were identified for  $AUC_G$ , with an effect size of  $g = 0.489$ , and  $a_{14}$ , with an effect size of  $g = 0.408$ . All computed cortisol features for each day are plotted in Figure 6.2.



**Figure 6.1:** Cortisol response to the TSST: Mean  $\pm$  Standard error of cortisol samples across all participants for each test day.



**Figure 6.2:** Cortisol response to the TSST: Mean  $\pm$  Standard error of cortisol features across all participants for each test day.

**Table 6.1:** Statistical analysis of cortisol feature analysis.

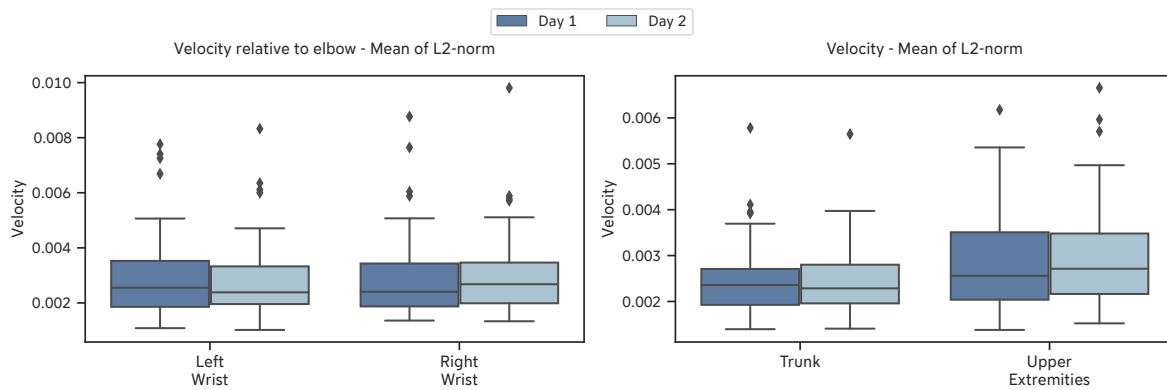
Saliva Feature	W	p	Hedges' g
$AUC_G$	408	<0.001***	0.489
$AUC_I$	671	<0.001***	0.355
$\Delta c_{max}$	665	<0.001***	0.384
$a_{14}$	521	<0.001***	0.408

## 6.2 Motion Features

Due to computational errors and insufficient data, 13 expert features have been disregarded from the analysis. This section presents the results derived from the remaining 224 motion features, comprising 172 generic and 52 expert features. Given the number of features, only a selected subset of results is presented in this section. A comprehensive overview of the statistical analysis results encompassing all features can be found in Appendix B.

### 6.2.1 Generic Features

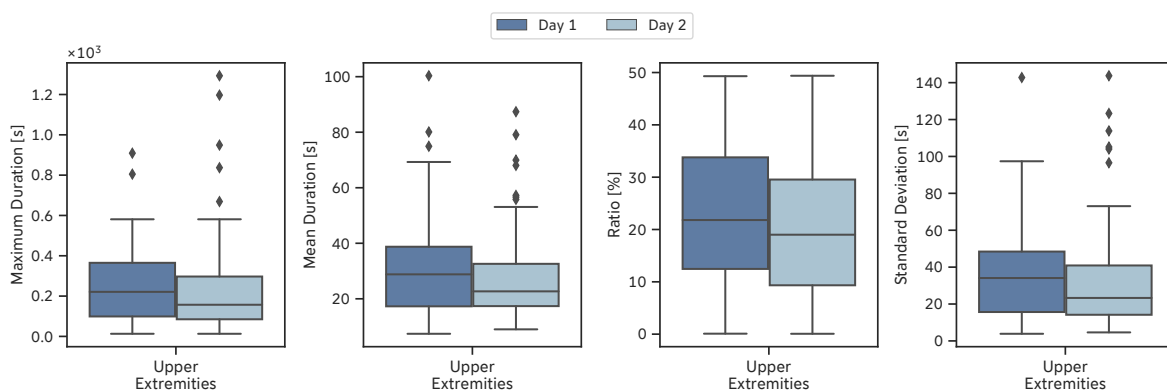
From the 172 generic features, no significant differences were found between data from day one and day two. Various generic features quantify the hand movement during the TSST, including the mean  $vel\_norm$  for both the left and right wrists. As depicted in Figure 6.3, there was no statistically significant increase in hand movement from day one to day two in the form of increased mean normalised velocity. Additionally, an increase in the mean velocity of body part groups, such as the Upper Extremities and Trunk – a potential indicator of freezing behaviour on the first day – was also not found. There were no considerable variations between the two test days to suggest freezing behaviour. The comprehensive statistical evaluation of the calculated generic features is available in Table B.1.



**Figure 6.3:** Plot of selected generic motion features.

## 6.2.2 Expert Features

The statistical analysis did not identify any statistically significant differences between the data from the two days for any of the expert features. A complete overview of the statistical results for all expert features is provided in Table B.2. Static periods observed during the TSST, which can be an indicator of freezing behaviour, did not show longer mean and maximum durations. An example of the static periods feature for the Upper Extremities is illustrated in Figure 6.4. The static periods identified during the first day did not show significantly prolonged mean and maximum durations when compared to the second day.



**Figure 6.4:** Plot of selected expert motion features.

# Chapter 7

## Discussion

### 7.1 Study Evaluation

Roos et al. conducted a study to investigate the effect of stress habituation on repeated acute psychosocial stress [Roo19]. In this thesis, this investigation was extended by analysing whether habituation could be observed in freezing-related motion parameters. Consistent with the finds of Roos et al., salivary measurements demonstrated habituation effects across all computed parameters derived from the raw cortisol values; however, there was no notable habituation effect in the freezing-related motion features. This suggests that, although repeated participation in the TSST leads to a habituation effect on cortisol secretion, it may not similarly influence motion-related behaviours. It is plausible that the intrinsic stressfulness of the TSST continues to trigger defensive freezing, even after multiple exposures.

Prior research by Abel et al. and Richer et al. explored defensive freezing behaviours during acute psychosocial stress [Abe22; Ric23]. As a comparative baseline, they employed the f-TSST [Wie13]. Their results show a considerable increase in defensive freezing motion parameters when participants were exposed to the TSST compared to the f-TSST. In contrast to these studies, both test days of the current evaluation only utilised the TSST, so no such baseline was available.

A limitation of the study evaluation is the restricted field of view of the captured videos, encompassing only the participants' upper body. Although a comprehensive set of motion features covering the upper bodies was computed, some features were inevitably omitted due to this constraint. In contrast, previous research had motion data covering the entire body.

Nevertheless, the lack of motion habituation presents an interesting area for exploration. Future research should investigate motion habituation in the context of repeated acute psychosocial stress further.

## 7.2 openTSST Platform

One main goal of the *openTSST* platform was to improve accessibility to video-based motion analysis during acute psychosocial stress. Incorporating motion-based stress analysis along with traditional stress markers can provide a more holistic assessment of the human response to acute stress. The platform presented in this thesis successfully demonstrates end-to-end motion parameter extraction during acute psychosocial stress.

Compared to optical or IMU-based motion capture techniques, the video-based approach is cost-effective and simple to set up. No specialised hardware is required and video recordings are already part of the TSST protocol. The accuracy of video-based motion analysis has also reached remarkable accuracy and is thus sufficient for motion parameter extraction. Moreover, since no markers are needed, there is no interference with natural human behaviour, unlike other methods. Nonetheless, the complex operation of video-based pose estimation tools can be a barrier to broader adoption. These tools demand technical expertise and a GPU-equipped machine for efficient processing. The *openTSST* platform bridges this gap, simplifying the user experience and handling pose estimation behind the scenes using its GPU-equipped worker cluster.

Additionally, the unprocessed OpenPose output does not yield any direct value to the researcher. Advanced data processing and feature extraction are required for meaningful psychosocial stress analysis. This is a complex process, requiring substantial technical knowledge, as the pose estimation data processing is performed using complex Python scripts. Executing these scripts requires a local Python environment with multiple installed packages and adaptation to the available data, making the setup time-consuming. To lower these barriers, *openTSST* provides researchers with a user-friendly, web-based interface that facilitates motion parameter extraction from videos, concealing technical details. The platform achieves this by utilising its custom cloud-based processing pipeline. All necessary tools reside there, allowing for fully automated processing. Researchers receive the aggregated OpenPose pose estimation output, converted into a CSV file for subsequent tasks, in addition to the extracted motion features. All the results are downloadable via the web platform.



As research into motion behaviour during acute stress is constantly evolving, the platform features a highly modular data processing pipeline design, ensuring adaptability to emerging research needs. It provides a robust foundation for motion feature extraction, adapting to align with the needs of future studies.

One limitation of the current processing pipeline is that it computes a predefined set of motion features. Two aspects of this are potentially suboptimal. First, most features are based on normalised data. The DataFrame is normalised to participants' body heights using specific body keypoints. Video content capturing only the upper body might result in a failed normalisation. Second, the expert features that are computed require a configured threshold. Although preconfigured, the threshold may not be optimal for all recorded input data as it might vary based on the recordings. Moreover, the *openTSST* platform's present deployment restricts its operation to within the university's own virtual private network. Before a public rollout, it is essential to address legal aspects, such as the inclusion of an imprint and General Data Protection Regulation (GDPR) compliance. Potential abusive usage also requires consideration, and appropriate technical measures, such as robust firewall protections, should be implemented.

Nevertheless, the *openTSST* platform serves as a robust foundation for video-based motion parameter extraction during acute psychosocial stress, holding the potential to increase the adoption of body posture and movement information to assess acute stress reactions.



# Chapter 8

## Conclusion & Outlook

This Bachelor's thesis presents *openTSST*, a web-based platform designed for large-scale, video-based motion analysis during acute psychosocial stress. While the TSST is the gold standard for inducing acute psychosocial stress in a laboratory setting, analyses of these tests have historically focused on traditional stress markers such as cortisol and inflammatory markers. Available at <https://mad-opentsst.aibe.uni-erlangen.de>, the platform empowers researchers to extend their analyses of TSST studies by including body posture and movement features without requiring extensive technical proficiency. Built on a highly modular architecture, the *openTSST* platform is designed to handle high computational loads and can accommodate a growing number of researchers. The platform is also capable of being adapted and expanded to address evolving research needs.

*openTSST* addresses the traditional challenges of cost, time, and technical proficiency typically required in the analysis of psychosocial stress through motion parameters. In contrast to the cumbersome nature of optical or IMU-based motion capture, video-based motion capture is a more cost-effective and less time-consuming option. As video recording is already a part of the TSST protocol, there is no need for additional equipment. This allows *openTSST* to not only be able to analyse current and upcoming studies but also to be able to analyse studies previously conducted where video data are available. By facilitating large-scale, affordable, and easy-to-use motion-based analysis of acute psychosocial stress, *openTSST* holds the potential to significantly accelerate the adoption of motion-based stress analysis in the field.

The capabilities of the *openTSST* platform were demonstrated by extending a previous study by Roos et al., which examined the effect of stress habituation on repeated acute psychosocial stress [Roo19]. The analysis was extended beyond self-report and saliva markers to assess whether habituation effects could also be observed in freezing-related motion parameters. Video recordings from the original study were processed using *openTSST* to extract motion features. However, the statistical analysis did not reveal the expected habituation effect in the extracted motion features, as it did with the saliva markers. These findings underscore the need for further research into the effects of motion habituation in the TSST across successive test days.

The highly modular design of *openTSST* serves as a robust foundation for future advancements in the field of video-based motion analysis of acute psychosocial stress. Future research could focus on refining and expanding the platform's data processing pipeline. Its flexible architecture provides an ideal basis for such enhancements, including but not limited to the implementation of more sophisticated data processing techniques or the integration of additional motion feature extraction algorithms. Moreover, the platform could be extended to feature different types of analyses besides just motion-based analysis. For instance, incorporating facial expression and voice analysis using tools such as OpenFace<sup>1</sup> and OpenDBM<sup>2</sup> could be valuable extensions. Another promising inclusion would be remote photoplethysmography, a technique that can detect blood volume changes from observations of a person's face or skin. By bundling all these methods together, the platform could evolve into an all-in-one solution for contactless stress analysis.

---

<sup>1</sup><https://github.com/TadasBaltrusaitis/OpenFace>

<sup>2</sup>[https://github.com/AiCure/open\\_dbm](https://github.com/AiCure/open_dbm)

# List of Figures

2.1	Illustration of the OpenPose processing pipeline [Cao19]. . . . .	5
4.1	High-level <i>openTSST</i> platform architecture overview. . . . .	14
4.2	Processing gateway database models. . . . .	16
4.3	Processing task submission flow. . . . .	17
4.4	Illustration of selected Kubernetes deployments and workflows in their respective namespaces in the worker cluster. . . . .	19
4.5	Illustration of the main data processing steps, along with generated artefacts.	21
4.6	OpenPose BODY_25 body part mapping. . . . .	22
5.1	Overview of the study procedure of a single test day. . . . .	30
6.1	Cortisol response to the TSST: Mean $\pm$ Standard error of cortisol samples across all participants for each test day. . . . .	36
6.2	Cortisol response to the TSST: Mean $\pm$ Standard error of cortisol features across all participants for each test day. . . . .	36
6.3	Plot of selected generic motion features. . . . .	38
6.4	Plot of selected expert motion features. . . . .	38
A.1	<i>openTSST</i> processing task submission interface. . . . .	63
A.2	<i>openTSST</i> video segmentation popup. . . . .	64
A.3	<i>openTSST</i> processing task result overview. . . . .	65



# List of Tables

4.1	Processing gateway environment variables overview. . . . .	15
4.2	Definition of body part groups. {L/R} denotes the left and right sides. . . . .	26
4.3	Overview of computed generic features. . . . .	26
5.1	Demographic and anthropometric data of the participants. Mean $\pm$ SD. . . . .	29
6.1	Statistical analysis of cortisol feature analysis. . . . .	37
B.1	Statistical results of all generic features. . . . .	68
B.2	Statistical results of all static periods expert features. . . . .	74





# Bibliography

- [Abe21] Kenta Abe, Ken-Ichi Tabei, Keita Matsuura, Kazuyuki Kobayashi, and Tomoyuki Ohkubo. “OpenPose-based Gait Analysis System For Parkinson’s Disease Patients From Arm Swing Data”. In: *2021 International Conference on Advanced Mechatronic Systems (ICAMechS)*. Tokyo, Japan: IEEE, Dec. 2021, pp. 61–65. ISBN: 978-1-66541-752-5. DOI: 10.1109/ICAMechS54019.2021.9661562. URL: <https://ieeexplore.ieee.org/document/9661562/> (visited on 08/12/2023).
- [Abe22] Luca Abel. “Machine Learning-Based Detection of Acute Psychosocial Stress from Dynamic Movements”. MA thesis. Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2022.
- [Ahm13] Norhafizan Ahmad, Raja Ariffin Raja Ghazilla, Nazirah M Khairi, and Vijayabaskar Kasi. “Reviews on various inertial measurement unit (IMU) sensor applications”. In: *International Journal of Signal Processing Systems* 1.2 (2013). Publisher: EJournal Publishing, pp. 256–262.
- [Alz21] Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, and Laith Farhan. “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions”. en. In: *Journal of Big Data* 8.1 (Mar. 2021), p. 53. ISSN: 2196-1115. DOI: 10.1186/s40537-021-00444-8. URL: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00444-8> (visited on 07/19/2023).
- [Atk04] Anthony P Atkinson, Winand H Dittrich, Andrew J Gemmell, and Andrew W Young. “Emotion Perception from Dynamic and Static Body Expressions in Point-Light and Full-Light Displays”. en. In: *Perception* 33.6 (June 2004),

- pp. 717–746. ISSN: 0301-0066, 1468-4233. DOI: 10.1068/p5096. URL: <http://journals.sagepub.com/doi/10.1068/p5096> (visited on 08/08/2023).
- [Bal00] Gene Ball and Jack Breese. “Relating Personality and Behavior: Posture and Gestures”. In: *Affective Interactions*. Ed. by Gerhard Goos, Juris Hartmanis, Jan Van Leeuwen, and Ana Paiva. Vol. 1814. Series Title: Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 196–203. DOI: 10.1007/10720296\_14. URL: [http://link.springer.com/10.1007/10720296\\_14](http://link.springer.com/10.1007/10720296_14) (visited on 08/08/2023).
- [Bar08] Sian Barris and Chris Button. “A Review of Vision-Based Motion Analysis in Sport.” en. In: *Sports Medicine* 38.12 (2008), pp. 1025–1043. ISSN: 0112-1642. DOI: 10.2165/00007256-200838120-00006. URL: <http://link.springer.com/10.2165/00007256-200838120-00006> (visited on 08/12/2023).
- [Bos23] Melissa A. Boswell, Łukasz Kidziński, Jennifer L. Hicks, Scott D. Uhrich, Antoine Falisse, and Scott L. Delp. “Smartphone videos of the sit-to-stand test predict osteoarthritis and health outcomes in a nationwide study”. en. In: *npj Digital Medicine* 6.1 (Mar. 2023), p. 32. ISSN: 2398-6352. DOI: 10.1038/s41746-023-00775-1. URL: <https://www.nature.com/articles/s41746-023-00775-1> (visited on 03/06/2023).
- [Bra04] H. Stefan Bracha. “Freeze, Flight, Fight, Fright, Faint: Adaptationist Perspectives on the Acute Stress Response Spectrum”. en. In: *CNS Spectrums* 9.9 (Sept. 2004), pp. 679–685. ISSN: 1092-8529, 2165-6509. DOI: 10.1017/S1092852900001954. URL: [https://www.cambridge.org/core/product/identifier/S1092852900001954/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S1092852900001954/type/journal_article) (visited on 08/09/2023).
- [Bri18] Michel S. Brink and Koen A.P.M. Lemmink. “Performance analysis in elite football: all in the game?” en. In: *Science and Medicine in Football* 2.4 (Oct. 2018), pp. 253–254. ISSN: 2473-3938, 2473-4446. DOI: 10.1080/24733938.2018.1532659. URL: <https://www.tandfonline.com/doi/full/10.1080/24733938.2018.1532659> (visited on 07/15/2023).
- [Cao19] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. *Open-Pose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields*.

- arXiv:1812.08008 [cs]. May 2019. URL: <http://arxiv.org/abs/1812.08008> (visited on 07/09/2023).
- [Che22] Haoming Chen, Runyang Feng, Sifan Wu, Hao Xu, Fengcheng Zhou, and Zhen-guang Liu. *2D Human Pose Estimation: A Survey*. arXiv:2204.07370 [cs]. Apr. 2022. URL: <http://arxiv.org/abs/2204.07370> (visited on 07/15/2023).
- [Col18] Steffi L. Colyer, Murray Evans, Darren P. Cosker, and Aki I. T. Salo. “A Re-view of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless Sys-tem”. en. In: *Sports Medicine - Open* 4.1 (Dec. 2018), p. 24. ISSN: 2199-1170, 2198-9761. DOI: 10.1186/s40798-018-0139-y. URL: <https://sportsmedicine-open.springeropen.com/articles/10.1186/s40798-018-0139-y> (visited on 07/19/2023).
- [Dae12] Nele Dael, Marcello Mortillaro, and Klaus R Scherer. “Emotion expression in body action and posture”. In: *Emotion* 12.5 (Oct. 2012). ZSCC: 0000303, pp. 1085–1101. ISSN: 1528-3542. DOI: 10.1037/a0025737. URL: <http://dx.doi.org/10.1037/a0025737>.
- [DAn20] Erika D’Antonio, Juri Taborri, Eduardo Palermo, Stefano Rossi, and Fabrizio Patane. “A markerless system for gait analysis based on OpenPose library”. In: *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. Dubrovnik, Croatia: IEEE, May 2020, pp. 1–6. ISBN: 978-1-72814-460-3. DOI: 10.1109/I2MTC43012.2020.9128918. URL: <https://ieeexplore.ieee.org/document/9128918/> (visited on 08/12/2023).
- [De 06] Beatrice De Gelder. “Towards the neurobiology of emotional body language”. en. In: *Nature Reviews Neuroscience* 7.3 (Mar. 2006), pp. 242–249. ISSN: 1471-003X, 1471-0048. DOI: 10.1038/nrn1872. URL: <https://www.nature.com/articles/nrn1872> (visited on 08/08/2023).
- [Dic04] Sally S. Dickerson and Margaret E. Kemeny. “Acute stressors and cortisol re-sponses: A theoretical integration and synthesis of laboratory research”. In: *Psychological Bulletin* 130.3 (2004). ISBN: 0033-2909 (Print)\n0033-2909 (Link-ing), pp. 355–391. ISSN: 00332909. DOI: 10.1037/0033-2909.130.3.355. URL: 10.1037/0033-2909.130.3.355.

- [Dou18] Michail Doumas, Kinga Morsanyi, and William R. Young. “Cognitively and socially induced stress affects postural control”. en. In: *Experimental Brain Research* 236.1 (Jan. 2018). tex.ids: Doumas2018, pp. 305–314. issn: 0014-4819, 1432-1106. doi: 10.1007/s00221-017-5128-8. url: <http://link.springer.com/10.1007/s00221-017-5128-8> (visited on 10/17/2020).
- [Dub23] Shradha Dubey and Manish Dixit. “A comprehensive survey on human pose estimation approaches”. en. In: *Multimedia Systems* 29.1 (Feb. 2023), pp. 167–195. issn: 0942-4962, 1432-1882. doi: 10.1007/s00530-022-00980-0. url: <https://link.springer.com/10.1007/s00530-022-00980-0> (visited on 08/13/2023).
- [Eic16] Patric Eichelberger, Matteo Ferraro, Ursina Minder, Trevor Denton, Angela Blasimann, Fabian Krause, and Heiner Baur. “Analysis of accuracy in optical motion capture – A protocol for laboratory setup evaluation”. en. In: *Journal of Biomechanics* 49.10 (July 2016), pp. 2085–2088. issn: 00219290. doi: 10.1016/j.jbiomech.2016.05.007. url: <https://linkinghub.elsevier.com/retrieve/pii/S0021929016305681> (visited on 07/16/2023).
- [Eil05] David Eilam. “Die hard: A blend of freezing and fleeing as a dynamic defense—implications for the control of defensive behavior”. en. In: *Neuroscience & Biobehavioral Reviews* 29.8 (2005), pp. 1181–1191. issn: 01497634. doi: 10.1016/j.neubiorev.2005.03.027. url: <https://linkinghub.elsevier.com/retrieve/pii/S0149763405000643> (visited on 08/08/2023).
- [Fan18] Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai, and Cewu Lu. *RMPE: Regional Multi-person Pose Estimation*. arXiv:1612.00137 [cs]. Feb. 2018. url: <http://arxiv.org/abs/1612.00137> (visited on 07/09/2023).
- [Fan23] Hao-Shu Fang, Jiefeng Li, Hongyang Tang, Chao Xu, Haoyi Zhu, Yuliang Xiu, Yong-Lu Li, and Cewu Lu. “AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time”. en. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45.6 (June 2023), pp. 7157–7173. issn: 0162-8828, 2160-9292, 1939-3539. doi: 10.1109/TPAMI.2022.3222784. url: <https://ieeexplore.ieee.org/document/9954214/> (visited on 05/09/2023).
- [Fil17] Alessandro Filipposchi, Norbert Schmitz, Markus Miezal, Gabriele Bleser, Emanuele Ruffaldi, and Didier Stricker. “Survey of Motion Tracking Methods

- Based on Inertial Sensors: A Focus on Upper Limb Human Motion”. en. In: *Sensors* 17.6 (June 2017), p. 1257. ISSN: 1424-8220. DOI: 10.3390/s17061257. URL: <http://www.mdpi.com/1424-8220/17/6/1257> (visited on 07/16/2023).
- [Gar13] John Garhammer and Harvey Newton. “Applied Video Analysis for Coaches: Weightlifting Examples”. en. In: *International Journal of Sports Science & Coaching* 8.3 (Sept. 2013), pp. 581–594. ISSN: 1747-9541, 2048-397X. DOI: 10.1260/1747-9541.8.3.581. URL: <http://journals.sagepub.com/doi/10.1260/1747-9541.8.3.581> (visited on 08/12/2023).
- [Gel15] B. de Gelder, A.W. de Borst, and R. Watson. “The perception of emotion in body expressions: Emotional body perception”. en. In: *Wiley Interdisciplinary Reviews: Cognitive Science* 6.2 (Mar. 2015). tex.ids: DeGelder2015, pp. 149–158. ISSN: 19395078. DOI: 10.1002/wcs.1335. URL: 10.1002/wcs.1335 (visited on 10/18/2020).
- [Gle02] M. Gleicher and N. Ferrier. “Evaluating video-based motion capture”. In: *Proceedings of Computer Animation 2002 (CA 2002)*. 2002, pp. 75–80. DOI: 10.1109/CA.2002.1017510.
- [Goo17] William K. Goodman, Johanna Janson, and Jutta M. Wolf. “Meta-analytical assessment of the effects of protocol variations on cortisol responses to the Trier Social Stress Test”. en. In: *Psychoneuroendocrinology* 80 (June 2017), pp. 26–35. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2017.02.030. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453016309702> (visited on 08/09/2023).
- [Gue05] Gutemberg Guerra Filho. “Optical Motion Capture: Theory and Implementation.” In: *RITA* 12 (Jan. 2005), pp. 61–90.
- [Gül18] Rıza Alp Güler, Natalia Neverova, and Iasonas Kokkinos. *DensePose: Dense Human Pose Estimation In The Wild*. arXiv:1802.00434 [cs]. Feb. 2018. URL: <http://arxiv.org/abs/1802.00434> (visited on 07/20/2023).
- [Had19] Arghavan Hadinejad, Brent D. Moyle, Noel Scott, and Anna Kralj. “Emotional responses to tourism advertisements: the application of FaceReader™”. en. In: *Tourism Recreation Research* 44.1 (Jan. 2019), pp. 131–135. ISSN: 0250-8281, 2320-0308. DOI: 10.1080/02508281.2018.1505228. URL: <https://www.>

- tandfonline.com/doi/full/10.1080/02508281.2018.1505228 (visited on 08/12/2023).
- [Hag14] Muriel A. Hagenaaers, Karin Roelofs, and John F. Stins. “Human freezing in response to affective films”. en. In: *Anxiety, Stress, & Coping* 27.1 (Jan. 2014), pp. 27–37. ISSN: 1061-5806, 1477-2205. DOI: 10.1080/10615806.2013.809420. URL: <http://www.tandfonline.com/doi/abs/10.1080/10615806.2013.809420> (visited on 10/17/2020).
- [He18] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. *Mask R-CNN*. arXiv:1703.06870 [cs]. Jan. 2018. URL: <http://arxiv.org/abs/1703.06870> (visited on 07/09/2023).
- [Hen21] Craig P. Hensley, Erin M. Lenihan, Kyle Pratt, Aayush Shah, Erin O’Donnell, Ping-Chen Nee, Jungwha Lee, Amy Yang, and Alison H. Chang. “Patterns of video-based motion analysis use among sports physical therapists”. en. In: *Physical Therapy in Sport* 50 (July 2021), pp. 159–165. ISSN: 1466853X. DOI: 10.1016/j.ptsp.2021.05.003. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1466853X21000845> (visited on 08/12/2023).
- [Hua18] Yinghao Huang, Manuel Kaufmann, Emre Aksan, Michael J. Black, Otmar Hilliges, and Gerard Pons-Moll. “Deep inertial poser: learning to reconstruct human pose from sparse inertial measurements in real time”. en. In: *ACM Transactions on Graphics* 37.6 (Dec. 2018), pp. 1–15. ISSN: 0730-0301, 1557-7368. DOI: 10.1145/3272127.3275108. URL: <https://dl.acm.org/doi/10.1145/3272127.3275108> (visited on 07/16/2023).
- [Jon12] Michael B. Jones and Dick Hardt. *The OAuth 2.0 Authorization Framework: Bearer Token Usage*. Published: RFC 6750. Oct. 2012. DOI: 10.17487/RFC6750. URL: <https://www.rfc-editor.org/info/rfc6750>.
- [Kir93] Clemens Kirschbaum, Karl-Martin Pirke, and Dirk H. Hellhammer. “The ‘Trier Social Stress Test’ – A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting”. In: *Neuropsychobiology*. Vol. 28. Issue: 1-2 ISSN: 0302282X. 1993, pp. 76–81. ISBN: 978-0-87421-656-1. DOI: 10.1159/000119004.

- [Kir95] Clemens Kirschbaum, Jens C. Prussner, Arthur A. Stone, Ilona Federenko, Jens Gaab, Doris Lintz, Nicole Schommer, and Dirk H. Hellhammer. “Persistent High Cortisol Responses to Repeated Psychological Stress in a Subpopulation of Healthy Men.” en. In: *Psychosomatic Medicine* 57.5 (1995), pp. 468–474. ISSN: 0033-3174. DOI: 10.1097/00006842-199509000-00009. URL: <http://journals.lww.com/00006842-199509000-00009> (visited on 08/09/2023).
- [Koo11] J.M. Koolhaas, A. Bartolomucci, B. Buwalda, S.F. De Boer, G. Flügge, S.M. Korte, P. Meerlo, R. Murison, B. Olivier, P. Palanza, G. Richter-Levin, A. Sgoifo, T. Steimer, O. Stiedl, G. Van Dijk, M. Wöhr, and E. Fuchs. “Stress revisited: A critical evaluation of the stress concept”. en. In: *Neuroscience & Biobehavioral Reviews* 35.5 (Apr. 2011), pp. 1291–1301. ISSN: 01497634. DOI: 10.1016/j.neubiorev.2011.02.003. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0149763411000224> (visited on 08/04/2023).
- [Kub23] Kubernetes. *Kubernetes*. 2023. URL: <https://kubernetes.io/docs/concepts/overview/>.
- [Lew14] Peter Lewinski, Tim M. Den Uyl, and Crystal Butler. “Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader.” en. In: *Journal of Neuroscience, Psychology, and Economics* 7.4 (Dec. 2014), pp. 227–236. ISSN: 2151-318X, 1937-321X. DOI: 10.1037/npe0000028. URL: <http://doi.apa.org/getdoi.cfm?doi=10.1037/npe0000028> (visited on 07/11/2023).
- [Lin15] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. *Microsoft COCO: Common Objects in Context*. \_eprint: 1405.0312. 2015.
- [Löw15] Andreas Löw, Mathias Weymar, and Alfons O. Hamm. “When Threat Is Near, Get Out of Here: Dynamics of Defensive Behavior During Freezing and Active Avoidance”. en. In: *Psychological Science* 26.11 (Nov. 2015), pp. 1706–1716. ISSN: 0956-7976, 1467-9280. DOI: 10.1177/0956797615597332. URL: <http://journals.sagepub.com/doi/10.1177/0956797615597332> (visited on 08/08/2023).
- [McE93] Bruce S. McEwen. “Stress and the Individual: Mechanisms Leading to Disease”. en. In: *Archives of Internal Medicine* 153.18 (Sept. 1993), p. 2093. ISSN:

- 0003-9926. DOI: 10.1001/archinte.1993.00410180039004. URL: <http://archinte.jamanetwork.com/article.aspx?doi=10.1001/archinte.1993.00410180039004> (visited on 08/09/2023).
- [Nak20] Nobuyasu Nakano, Tetsuro Sakura, Kazuhiro Ueda, Leon Omura, Arata Kimura, Yoichi Iino, Senshi Fukashiro, and Shinsuke Yoshioka. “Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras”. In: *Frontiers in Sports and Active Living* 2 (May 2020), p. 50. ISSN: 2624-9367. DOI: 10.3389/fspor.2020.00050. URL: <https://www.frontiersin.org/article/10.3389/fspor.2020.00050/full> (visited on 07/16/2023).
- [Nat09] U.M. Nater and N. Rohleder. “Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: Current state of research”. en. In: *Psychoneuroendocrinology* 34.4 (May 2009), pp. 486–496. ISSN: 03064530. DOI: 10.1016/j.psyneuen.2009.01.014. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453009000328> (visited on 08/09/2023).
- [OC021] Daryl B. O’Connor, Julian F. Thayer, and Kavita Vedhara. “Stress and Health: A Review of Psychobiological Processes”. en. In: *Annual Review of Psychology* 72.1 (Jan. 2021), pp. 663–688. ISSN: 0066-4308, 1545-2085. DOI: 10.1146/annurev-psych-062520-122331. URL: <https://www.annualreviews.org/doi/10.1146/annurev-psych-062520-122331> (visited on 07/11/2023).
- [Par22] Antonis Pardos, Melina Tziomaka, Andreas Menychtas, and Ilias Maglogiannis. “Automated Posture Analysis for the Assessment of Sports Exercises”. en. In: *Proceedings of the 12th Hellenic Conference on Artificial Intelligence*. Corfu Greece: ACM, Sept. 2022, pp. 1–9. ISBN: 978-1-4503-9597-7. DOI: 10.1145/3549737.3549784. URL: <https://dl.acm.org/doi/10.1145/3549737.3549784> (visited on 08/12/2023).
- [Pic21] Marco Pichierri, Alessandro M. Peluso, Giovanni Pino, and Gianluigi Guido. “Health claims’ text clarity, perceived healthiness of extra-virgin olive oil, and arousal: An experiment using FaceReader”. en. In: *Trends in Food Science & Technology* 116 (Oct. 2021), pp. 1186–1194. ISSN: 09242244. DOI: 10.1016/j.tifs.2021.05.032. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0924224421003551> (visited on 08/12/2023).



- [Pru03] Jens C. Pruessner, Clemens Kirschbaum, Gunther Meinlschmid, and Dirk H Hellhammer. “Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change”. en. In: *Psychoneuroendocrinology* 28.7 (Oct. 2003), pp. 916–931. ISSN: 03064530. DOI: 10.1016/S0306-4530(02)00108-7. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0306453002001087> (visited on 07/26/2023).
- [Ran20] Keerthana Rangasamy, Muhammad Amir As’ari, Nur Azmina Rahmad, Nurul Fathiah Ghazali, and Saharudin Ismail. “Deep learning in sport video analysis: a review”. In: *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 18.4 (Aug. 2020), p. 1926. ISSN: 2302-9293, 1693-6930. DOI: 10.12928/telkomnika.v18i4.14730. URL: <http://telkomnika.uad.ac.id/index.php/TELKOMNIKA/article/view/14730> (visited on 08/12/2023).
- [Ric21] Robert Richer, Arne Küderle, Martin Ullrich, Nicolas Rohleder, and Bjoern Eskofier. “BioPsyKit: A Python package for the analysis of biopsychological data”. In: *Journal of Open Source Software* 6.66 (Oct. 2021), p. 3702. ISSN: 2475-9066. DOI: 10.21105/joss.03702. URL: <https://joss.theoj.org/papers/10.21105/joss.03702> (visited on 07/31/2023).
- [Ric23] Robert Richer, Veronika Koch, Victoria Müller, Arne Küderle, Luca Abel, Nicolas Rohleder, and Bjoern M. Eskofier. “StressPose - Detection of Acute Psychosocial Stress from Body Posture and Movements”. In: (2023).
- [Roe09] Daniel Roetenberg, Henk Luinge, and Per Slycke. “Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors”. In: *Xsens Motion Technol. BV Tech. Rep.* 3 (Jan. 2009).
- [Roe10] Karin Roelofs, Muriel A. Hagens, and John Stins. “Facing Freeze: Social Threat Induces Bodily Freeze in Humans”. en. In: *Psychological Science* 21.11 (Nov. 2010). tex.ids: Roelofs2010, pp. 1575–1581. ISSN: 0956-7976, 1467-9280. DOI: 10.1177/0956797610384746. URL: 10.1177/0956797610384746 (visited on 10/18/2020).
- [Roe17] Karin Roelofs. “Freeze for action: neurobiological mechanisms in animal and human freezing”. en. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.1718 (Apr. 2017). tex.ids= Roelofs2017 publisher: Royal

- Society, p. 20160206. ISSN: 0962-8436, 1471-2970. DOI: 10.1098/rstb.2016.0206. URL: <https://royalsocietypublishing.org/doi/10.1098/rstb.2016.0206> (visited on 01/01/2023).
- [Roo19] Lydia G. Roos, Johanna Janson, Sarah C. Sturmbauer, Jeanette M. Bennett, and Nicolas Rohleder. “Higher trait reappraisal predicts stronger HPA axis habituation to repeated stress”. eng. In: *Psychoneuroendocrinology* 101 (Mar. 2019), pp. 12–18. ISSN: 1873-3360. DOI: 10.1016/j.psyneuen.2018.10.018.
- [Sco22] Bradley Scott, Martin Seyres, Fraser Philp, Edward K. Chadwick, and Dimitra Blana. “Healthcare applications of single camera markerless motion capture: a scoping review”. In: *PeerJ* 10 (May 2022), e13517. ISSN: 2167-8359. DOI: 10.7717/peerj.13517. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9148557/> (visited on 07/13/2023).
- [Sin18] Jasvinder Pal Singh, Sanjeev Jain, Sakshi Arora, and Uday Pratap Singh. “Vision-Based Gait Recognition: A Survey”. In: *IEEE Access* 6 (2018), pp. 70497–70527. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2018.2879896. URL: <https://ieeexplore.ieee.org/document/8528404/> (visited on 08/12/2023).
- [Ski19] Tanja Skiendziel, Andreas G. Rösch, and Oliver C. Schultheiss. “Assessing the convergent validity between the automated emotion recognition software Noldus FaceReader 7 and Facial Action Coding System Scoring”. en. In: *PLOS ONE* 14.10 (Oct. 2019). Ed. by Jan De Houwer, e0223905. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0223905. URL: <https://dx.plos.org/10.1371/journal.pone.0223905> (visited on 07/21/2023).
- [Sod09] P. Soda, A. Carta, D. Formica, and E. Guglielmelli. “A low-cost video-based tool for clinical gait analysis”. In: *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Minneapolis, MN: IEEE, Sept. 2009, pp. 3979–3982. DOI: 10.1109/IEMBS.2009.5333623. URL: <http://ieeexplore.ieee.org/document/5333623/> (visited on 08/12/2023).
- [Ste21] Jan Stenum, Cristina Rossi, and Ryan T. Roemmich. “Two-dimensional video-based analysis of human gait using pose estimation”. en. In: *PLOS Computational Biology* 17.4 (Apr. 2021). Ed. by Dina Schneidman-Duhovny, e1008935. ISSN:

- 1553-7358. DOI: 10.1371/journal.pcbi.1008935. URL: <https://dx.plos.org/10.1371/journal.pcbi.1008935> (visited on 07/20/2023).
- [Sum23] Himanshu Kumar Suman and Tanmay Tulsidas Verlekar. “Video-Based Gait Analysis for Spinal Deformity”. en. In: *Computer Vision – ECCV 2022 Workshops*. Ed. by Leonid Karlinsky, Tomer Michaeli, and Ko Nishino. Vol. 13805. Series Title: Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2023, pp. 278–288. DOI: 10.1007/978-3-031-25072-9\_18. URL: [https://link.springer.com/10.1007/978-3-031-25072-9\\_18](https://link.springer.com/10.1007/978-3-031-25072-9_18) (visited on 08/12/2023).
- [Tec23] Noldus Information Technology. *FaceReader: A robust automated system for facial expressions analysis*. 2023. URL: <https://www.noldus.com/facereader>.
- [Ter10] Vasileios Terzis, Christos N. Moridis, and Anastasios A. Economides. “Measuring instant emotions during a self-assessment test: the use of FaceReader”. en. In: *Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research*. Eindhoven The Netherlands: ACM, Aug. 2010, pp. 1–4. ISBN: 978-1-60558-926-8. DOI: 10.1145/1931344.1931362. URL: <https://dl.acm.org/doi/10.1145/1931344.1931362> (visited on 08/12/2023).
- [Uhl22] Scott D. Uhlich, Antoine Falisse, Łukasz Kidziński, Julie Muccini, Michael Ko, Akshay S. Chaudhari, Jennifer L. Hicks, and Scott L. Delp. *OpenCap: 3D human movement dynamics from smartphone videos*. en. preprint. July 2022. DOI: 10.1101/2022.07.07.499061. URL: <http://biorxiv.org/lookup/doi/10.1101/2022.07.07.499061> (visited on 05/09/2023).
- [Val18] Raphael Vallat. “Pingouin: statistics in Python”. In: *Journal of Open Source Software* 3.31 (Nov. 2018), p. 1026. ISSN: 2475-9066. DOI: 10.21105/joss.01026. URL: <http://joss.theoj.org/papers/10.21105/joss.01026> (visited on 07/31/2023).
- [Vis19] Aditya Viswakumar, Venkateswaran Rajagopalan, Tathagata Ray, and Chandu Parimi. “Human Gait Analysis Using OpenPose”. In: *2019 Fifth International Conference on Image Information Processing (ICIIP)*. Shimla, India: IEEE, Nov. 2019, pp. 310–314. ISBN: 978-1-72810-899-5. DOI: 10.1109/ICIIP47207.2019.8985781. URL: <https://ieeexplore.ieee.org/document/8985781/> (visited on 08/12/2023).

- [Vri97] Aldert Vrij, Lucy Akehurst, and Paul Morris. “Individual Differences in Hand Movements During Deception”. In: *Journal of Nonverbal Behavior* 21.2 (1997), pp. 87–102. ISSN: 01915886. DOI: 10.1023/A:1024951902752. URL: <http://link.springer.com/10.1023/A:1024951902752> (visited on 07/26/2023).
- [Wal98] Harald G Wallbott. “Bodily expression of emotion”. In: *European journal of social psychology* 28.6 (1998). Publisher: Wiley Online Library, pp. 879–896.
- [Wan19] Jianbo Wang, Kai Qiu, Houwen Peng, Jianlong Fu, and Jianke Zhu. “AI Coach: Deep Human Pose Estimation and Analysis for Personalized Athletic Training Assistance”. en. In: *Proceedings of the 27th ACM International Conference on Multimedia*. Nice France: ACM, Oct. 2019, pp. 374–382. ISBN: 978-1-4503-6889-6. DOI: 10.1145/3343031.3350910. URL: <https://dl.acm.org/doi/10.1145/3343031.3350910> (visited on 08/12/2023).
- [Wan21a] Jinbao Wang, Shujie Tan, Xiantong Zhen, Shuo Xu, Feng Zheng, Zhenyu He, and Ling Shao. “Deep 3D human pose estimation: A review”. en. In: *Computer Vision and Image Understanding* 210 (Sept. 2021), p. 103225. ISSN: 10773142. DOI: 10.1016/j.cviu.2021.103225. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1077314221000692> (visited on 07/20/2023).
- [Wan21b] Yuqing Wang. “Physical Education Teaching in Colleges and Universities Assisted by Virtual Reality Technology Based on Artificial Intelligence”. en. In: *Mathematical Problems in Engineering* 2021 (Apr. 2021). Ed. by Sang-Bing Tsai, pp. 1–11. ISSN: 1563-5147, 1024-123X. DOI: 10.1155/2021/5582716. URL: <https://www.hindawi.com/journals/mpe/2021/5582716/> (visited on 07/15/2023).
- [Wie13] Uta S. Wiemers, Daniela Schoofs, and Oliver T. Wolf. “A friendly version of the Trier Social Stress Test does not activate the HPA axis in healthy men and women”. en. In: *Stress* 16.2 (Mar. 2013), pp. 254–260. ISSN: 1025-3890, 1607-8888. DOI: 10.3109/10253890.2012.714427. URL: 10.3109/10253890.2012.714427 (visited on 10/17/2020).
- [Wil08] Barry D. Wilson. “Development in video technology for coaching”. en. In: *Sports Technology* 1.1 (Jan. 2008), pp. 34–40. ISSN: 1934-6182, 1934-6190. DOI: 10.

- 1080/19346182.2008.9648449. URL: <http://www.tandfonline.com/doi/full/10.1080/19346182.2008.9648449> (visited on 08/12/2023).
- [Wu19] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. *Detectron2*. 2019. URL: <https://github.com/facebookresearch/detectron2>.
- [Yan18] Zhichao Yang, Mohammad H. Rafiei, Alexis Hall, Caroline Thomas, Hali A. Midtlien, Alexander Hasselbach, Hojjat Adeli, and Lynne V. Gauthier. “A Novel Methodology for Extracting and Evaluating Therapeutic Movements in Game-Based Motion Capture Rehabilitation Systems”. en. In: *Journal of Medical Systems* 42.12 (Dec. 2018), p. 255. ISSN: 0148-5598, 1573-689X. DOI: 10.1007/s10916-018-1113-4. URL: <http://link.springer.com/10.1007/s10916-018-1113-4> (visited on 08/12/2023).
- [Yu17] Chia-Yin Yu and Chih-Hsiang Ko. “Applying FaceReader to Recognize Consumer Emotions in Graphic Styles”. en. In: *Procedia CIRP* 60 (2017), pp. 104–109. ISSN: 22128271. DOI: 10.1016/j.procir.2017.01.014. URL: <https://linkinghub.elsevier.com/retrieve/pii/S221282711730015X> (visited on 07/17/2023).
- [Zee19] Sophie van der Zee, Ronald Poppe, Paul J Taylor, and Ross Anderson. “To freeze or not to freeze: A culture-sensitive motion capture approach to detecting deceit”. In: *PLoS One* 14.4 (Apr. 2019), e0215000. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0215000. URL: <http://dx.doi.org/10.1371/journal.pone.0215000>.
- [Zhe23] Ce Zheng, Wenhan Wu, Chen Chen, Taojiannan Yang, Sijie Zhu, Ju Shen, Nasser Kehtarnavaz, and Mubarak Shah. “Deep Learning-Based Human Pose Estimation: A Survey”. In: *ACM Comput. Surv.* (June 2023). Place: New York, NY, USA Publisher: Association for Computing Machinery. ISSN: 0360-0300. DOI: 10.1145/3603618. URL: <https://doi.org/10.1145/3603618>.



# Appendix A

## Additional Figures

### Submit a processing task

Task name

Your email


You will get notified via email when your processing task has been completed.

Processing options

Generate OpenPose Video  
Generate a video with the OpenPose skeleton overlay.

Extract Macro Features  
Extract empkins macro features from the video.

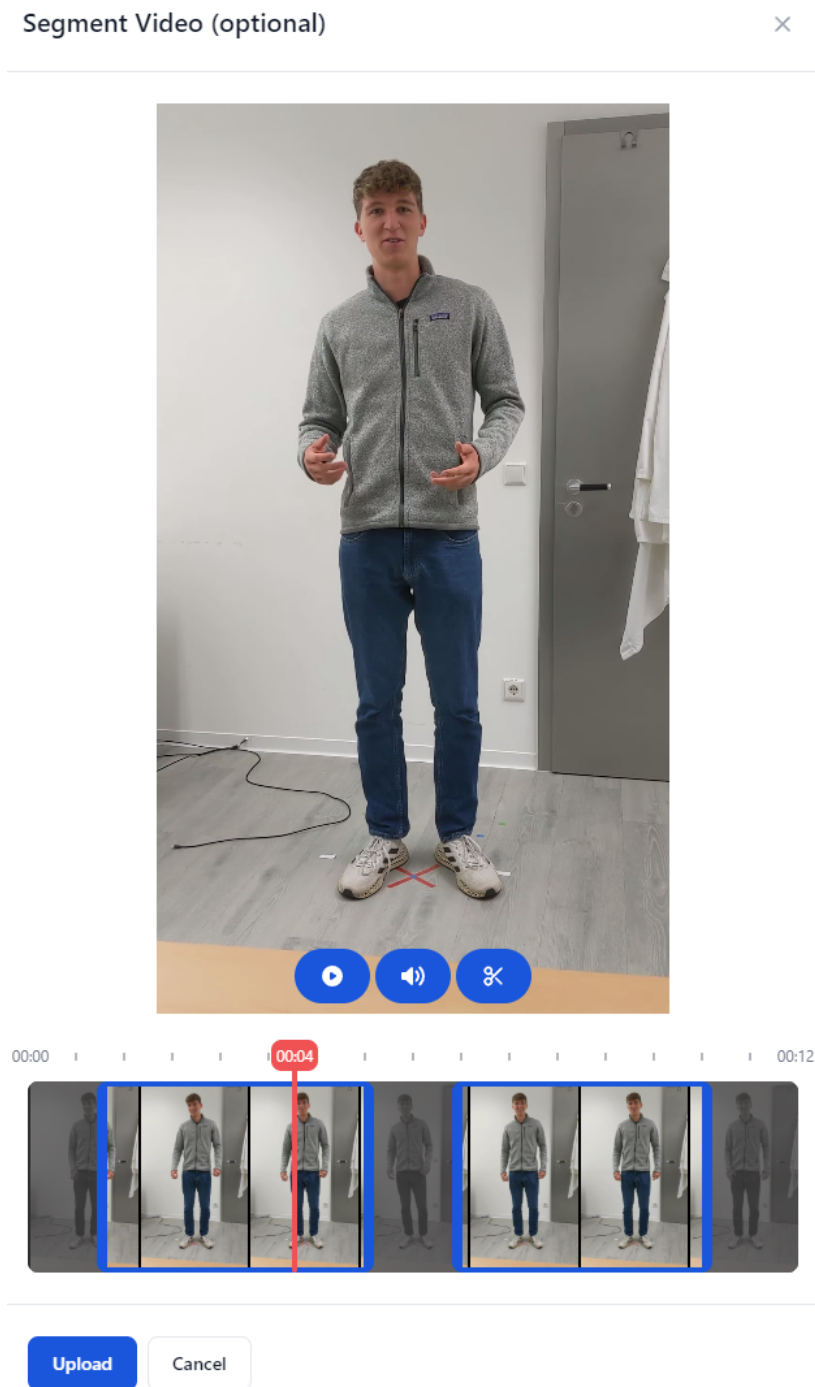
Uploaded videos



Upload a video file or drag and drop  
Only MP4 files up to 1GB are allowed.

[Submit](#)

**Figure A.1:** *openTSST* processing task submission interface.



**Figure A.2:** *openTSST* video segmentation popup.



## TSST Study

11 minutes ago

Email: tobias.gatzler@fhnw.de

### Processing options

Generate openpose video:

Extract macro features:

PXL\_20230808\_123218138.mp4

**i** Info: The video processing has been completed. You can view the results here.

### Configured segments:

SEGMENT	START	END
1	00:00	00:12

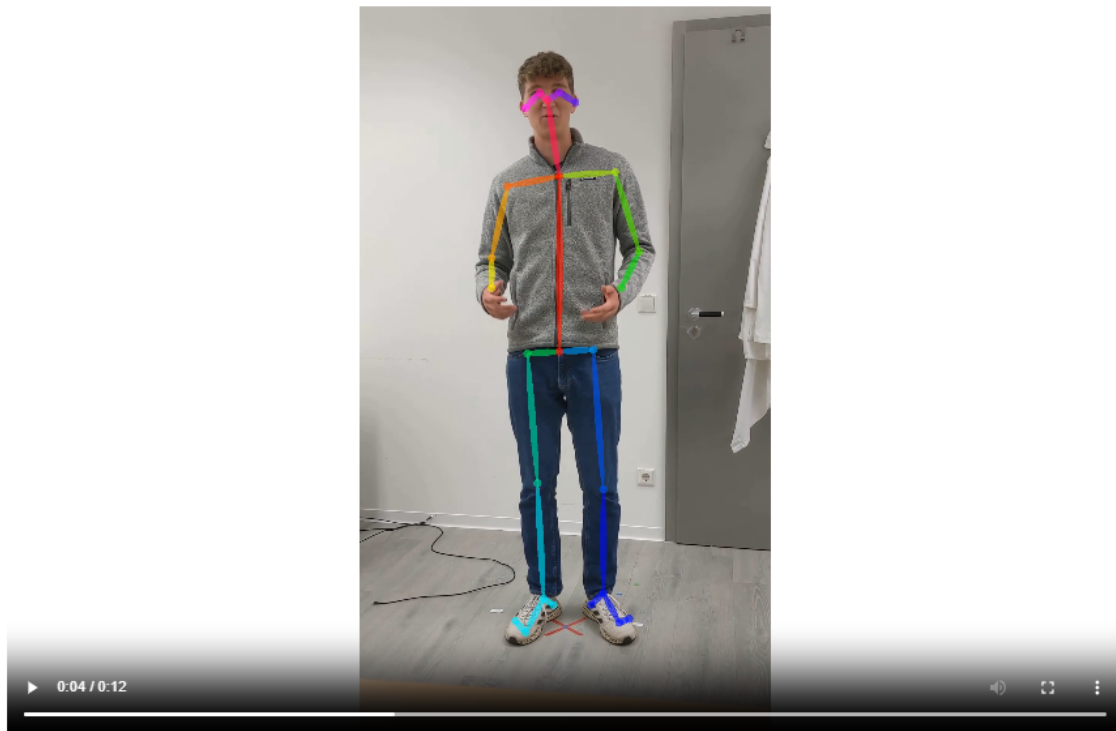
### Macro Features:

[Download openpose macro features csv](#)

### OpenPose Aggregated Data:

[Download openpose dataframe csv](#)

### OpenPose overlaid video:



**Figure A.3:** *openTSST* processing task result overview.



# **Appendix B**

## **Additional Tables**

**Table B.1:** Statistical results of all generic features.

Body Part	Channel	Metric	W	p	Hedges' g
Left Elbow	pos_norm	SD	1189	>0.999	0.105
		Entropy	1423	>0.999	0.119
	pos_norm_rel2neck	FFT Centroid	1509	>0.999	0.031
		FFT Skewness	1562	>0.999	-0.115
		SD	1013	0.812	0.247
		Entropy	1503	>0.999	0.101
	vel_norm	FFT Kurtosis	1533	>0.999	-0.128
		FFT Variance	1585	>0.999	-0.002
		Abs. Energy	1533	>0.999	-0.119
		Mean	1557	>0.999	-0.105
	vel_norm_rel2neck	Mean Crossings	1435.500	>0.999	-0.068
		Abs. Energy	1313	>0.999	-0.234
		Mean	1547	>0.999	-0.111
		SD	1326	>0.999	-0.232
	vel_norm_rel2neck_2d	Mean Crossings	1605	>0.999	0.007
		FFT Centroid	1504	>0.999	0.063
		FFT Skewness	1292	>0.999	0.157
		FFT Kurtosis	1386	>0.999	-0.109
Left Wrist	pos_norm	FFT Variance	1463	>0.999	0.064
		SD	1197	>0.999	0.144
	pos_norm_rel2leftelbow	Entropy	1352	>0.999	0.064
		FFT Centroid	1452	>0.999	0.085
		FFT Skewness	1591	>0.999	0.002
		SD	1089	>0.999	0.199
	pos_norm_rel2neck	Entropy	1548	>0.999	-0.052
		FFT Kurtosis	1616	>0.999	0.052
		FFT Variance	1374	>0.999	0.103
		FFT Centroid	1456	>0.999	0.120

		FFT Skewness	1398	>0.999	-0.133
		SD	997	0.634	0.245
		Entropy	1580	>0.999	-0.046
		FFT Kurtosis	1458	>0.999	-0.093
		FFT Variance	1455	>0.999	0.114
	vel	SD	1293	>0.999	0.183
	vel_norm	Abs. Energy	1421	>0.999	0.068
		Mean	1572	>0.999	0.051
		Mean Crossings	1339	>0.999	-0.136
	vel_norm_rel2leftelbow	Abs. Energy	1600	>0.999	0.022
		Mean	1544	>0.999	0.034
		SD	1515	>0.999	0.080
		Mean Crossings	1374	>0.999	-0.151
	vel_norm_rel2leftelbow_2d	FFT Centroid	1554	>0.999	-0.003
		FFT Skewness	1115	>0.999	0.302
		Zero Crossings	1235.500	>0.999	-0.212
		FFT Kurtosis	1581	>0.999	-0.094
		FFT Variance	1547	>0.999	-0.002
	vel_norm_rel2neck	Abs. Energy	1525	>0.999	0.004
		Mean	1537	>0.999	0.042
		SD	1597	>0.999	0.018
		Mean Crossings	1450	>0.999	-0.104
	vel_norm_rel2neck_2d	FFT Centroid	1498	>0.999	0.021
		FFT Skewness	1167	>0.999	0.228
		Zero Crossings	1233.500	>0.999	-0.212
		FFT Kurtosis	1587	>0.999	-0.073
		FFT Variance	1520	>0.999	0.023
Neck	pos_norm	SD	1317	>0.999	-0.240
	vel	SD	1504	>0.999	-0.103
	vel_norm	Mean	1515	>0.999	-0.100
		Mean Crossings	1581	>0.999	0.038
Nose	pos_norm	SD	1586	>0.999	-0.154

		Entropy	1299	>0.999	0.198
	pos_norm_rel2neck	FFT Centroid	1517	>0.999	0.064
		FFT Skewness	1567	>0.999	0.014
		SD	1080	>0.999	0.223
		Entropy	1336	>0.999	0.153
		FFT Kurtosis	1534	>0.999	0.032
		FFT Variance	1529	>0.999	0.054
	vel	SD	1607	>0.999	-0.144
	vel_norm	Abs. Energy	1600	>0.999	-0.288
		Mean	1607	>0.999	-0.049
		Mean Crossings	1559	>0.999	0.076
	vel_norm_rel2neck	Abs. Energy	1572	>0.999	-0.129
		Mean	1573	>0.999	-0.038
		SD	1608	>0.999	-0.088
		Mean Crossings	1579	>0.999	0.059
	vel_norm_rel2neck_2d	FFT Centroid	1454	>0.999	0.121
		FFT Skewness	1508	>0.999	-0.001
		Zero Crossings	1112.500	>0.999	-0.258
		FFT Kurtosis	1411	>0.999	0.019
		FFT Variance	1370	>0.999	0.120
Right Elbow	pos_norm	SD	1527	>0.999	-0.063
		Entropy	1570	>0.999	0.053
	pos_norm_rel2neck	FFT Centroid	1564	>0.999	-0.013
		FFT Skewness	1401	>0.999	-0.140
		SD	889	0.103	0.263
		Entropy	1616	>0.999	0.043
		FFT Kurtosis	1335	>0.999	-0.187
		FFT Variance	1529	>0.999	0.001
	vel	SD	1407	>0.999	-0.156
	vel_norm	Abs. Energy	1384	>0.999	-0.193
		Mean	1499	>0.999	-0.132
		Mean Crossings	1541.500	>0.999	0.010

	vel_norm_rel2neck	Abs. Energy	1237	>0.999	-0.214
		Mean	1445	>0.999	-0.133
		SD	1206	>0.999	-0.287
		Mean Crossings	1572.500	>0.999	0.064
	vel_norm_rel2neck_2d	FFT Centroid	1482	>0.999	0.077
		FFT Skewness	1473	>0.999	0.034
		FFT Kurtosis	1570	>0.999	-0.049
		FFT Variance	1461	>0.999	0.066
Right Wrist	pos_norm	SD	1368	>0.999	0.038
		Entropy	1525	>0.999	0.095
	pos_norm_rel2neck	FFT Centroid	1473	>0.999	0.004
		FFT Skewness	1538	>0.999	-0.163
		SD	1126	>0.999	0.178
		Entropy	1534	>0.999	0.052
		FFT Kurtosis	1483	>0.999	-0.219
		FFT Variance	1467	>0.999	0.007
	pos_norm_rel2rightelbow	FFT Centroid	1392	>0.999	0.011
		FFT Skewness	1488	>0.999	-0.069
		SD	1403	>0.999	0.130
		Entropy	1505	>0.999	0.076
		FFT Kurtosis	1435	>0.999	-0.104
		FFT Variance	1420	>0.999	0.003
	vel	SD	1602	>0.999	-0.055
	vel_norm	Abs. Energy	1584	>0.999	-0.154
		Mean	1461	>0.999	-0.138
		Mean Crossings	1567	>0.999	0.044
	vel_norm_rel2neck	Abs. Energy	1588	>0.999	-0.188
		Mean	1437	>0.999	-0.135
		SD	1562	>0.999	-0.198
		Mean Crossings	1543.500	>0.999	0.076
	vel_norm_rel2neck_2d	FFT Centroid	1402	>0.999	0.090
		FFT Skewness	1475	>0.999	0.072

		Zero Crossings	1113	>0.999	-0.268
		FFT Kurtosis	1489	>0.999	-0.153
		FFT Variance	1395	>0.999	0.109
	vel_norm_rel2rightelbow	Abs. Energy	1596	>0.999	-0.153
		Mean	1556	>0.999	-0.127
		SD	1613	>0.999	-0.136
		Mean Crossings	1513	>0.999	0.062
	vel_norm_rel2rightelbow_2d	FFT Centroid	1464	>0.999	0.067
		FFT Skewness	1428	>0.999	0.128
		Zero Crossings	1081.500	>0.999	-0.268
		FFT Kurtosis	1426	>0.999	-0.136
		FFT Variance	1434	>0.999	0.075
Trunk	pos_norm	SD	1359	>0.999	-0.189
		Entropy	1389	>0.999	0.027
	pos_norm_rel2neck	FFT Centroid	1460	>0.999	0.004
		FFT Skewness	1419	>0.999	-0.046
		SD	1424	>0.999	0.133
		Entropy	1539	>0.999	-0.054
		FFT Kurtosis	1428	>0.999	0.013
		FFT Variance	1472	>0.999	-0.018
	vel	SD	1598	>0.999	-0.080
	vel_norm	Abs. Energy	1591	>0.999	-0.082
		Mean	1590	>0.999	0.006
	vel_norm_rel2neck	Abs. Energy	1497	>0.999	-0.166
		Mean	1585	>0.999	0.053
		SD	1547	>0.999	-0.166
	vel_norm_rel2neck_2d	FFT Centroid	1471	>0.999	0.083
		FFT Skewness	1550	>0.999	-0.092
FFT Kurtosis		1475	>0.999	0.118	
FFT Variance		1461	>0.999	0.060	
Upper Extremities	pos_norm	SD	1555	>0.999	-0.087
		Entropy	1458	>0.999	0.108



pos_norm_rel2neck	FFT Centroid	1455	>0.999	0.026
	FFT Skewness	1423	>0.999	-0.174
	SD	1135	>0.999	0.216
	Entropy	1506	>0.999	0.080
	FFT Kurtosis	1402	>0.999	-0.215
	FFT Variance	1396	>0.999	0.025
vel	SD	1582	>0.999	-0.067
vel_norm	Abs. Energy	1613	>0.999	-0.117
	Mean	1484	>0.999	-0.102
vel_norm_rel2neck	Abs. Energy	1366	>0.999	-0.175
	Mean	1475	>0.999	-0.100
	SD	1375	>0.999	-0.216
vel_norm_rel2neck_2d	FFT Centroid	1509	>0.999	0.078
	FFT Skewness	1354	>0.999	0.125
	FFT Kurtosis	1528	>0.999	-0.047
	FFT Variance	1524	>0.999	0.069

---

**Table B.2:** Statistical results of all static periods expert features.

Body Part	Channel	Metric	W	p	Hedges' g
Left Wrist	vel_norm_rel2leftelbow	Counts per Minute	1509	>0.999	-0.119
		Mean Duration (s)	1129	>0.999	-0.228
		Ratio (%)	1528	>0.999	-0.119
		Std. Duration (s)	1065.500	>0.999	-0.002
Left Wrist & Right Wrist	Static Periods	Counts per Minute	1283	>0.999	0.163
		Mean Duration (s)	1577.500	>0.999	-0.043
		Ratio (%)	1356	>0.999	0.130
		Std. Duration (s)	1426	>0.999	0.048
	vel_norm_rel2neck	Counts per Minute	1466	>0.999	0.091
		Mean Duration (s)	1394.500	>0.999	-0.006
		Ratio (%)	1494	>0.999	0.050
		Std. Duration (s)	1337	>0.999	-0.042
Mid Hip	vel_norm	Counts per Minute	1409	>0.999	-0.105
		Mean Duration (s)	829	>0.999	-0.379
		Ratio (%)	1279	>0.999	-0.158
		Std. Duration (s)	793	>0.999	-0.254
	vel_norm_rel2neck	Counts per Minute	1363	>0.999	-0.122
		Mean Duration (s)	660	>0.999	-0.358
		Ratio (%)	1244	>0.999	-0.197
		Std. Duration (s)	606	>0.999	-0.277
Nose	vel_norm	Counts per Minute	1605	>0.999	-0.011
		Max. Duration (s)	1190	>0.999	0.055
		Mean Duration (s)	1293	>0.999	0.112
		Ratio (%)	1511	>0.999	0.024
		Std. Duration (s)	1316	>0.999	0.119
	vel_norm_rel2neck	Counts per Minute	1588	>0.999	-0.044
		Mean Duration (s)	1415.500	>0.999	-0.143
		Ratio (%)	1585	>0.999	-0.029

		Std. Duration (s)	1230.500	>0.999	-0.206		
Right Wrist	vel_norm_rel2rightelbow	Counts per Minute	1416	>0.999	-0.248		
		Mean Duration (s)	1143	>0.999	-0.256		
		Ratio (%)	1395	>0.999	-0.270		
Trunk	vel_norm	Std. Duration (s)	1129	>0.999	-0.256		
		Counts per Minute	1525	>0.999	-0.060		
		Max. Duration (s)	1039	>0.999	-0.100		
		Mean Duration (s)	1468	>0.999	-0.105		
		Ratio (%)	1548	>0.999	-0.082		
	vel_norm_rel2neck	Std. Duration (s)	1434	>0.999	-0.051		
		Counts per Minute	1532	>0.999	-0.087		
		Mean Duration (s)	1371	>0.999	-0.139		
		Ratio (%)	1548	>0.999	-0.100		
		Std. Duration (s)	1453	>0.999	-0.131		
		Upper Extremities	vel_norm	Counts per Minute	1275	>0.999	0.183
				Max. Duration (s)	1042	>0.999	0.024
Mean Duration (s)	1287			>0.999	0.091		
Ratio (%)	1251			>0.999	0.160		
vel_norm_rel2neck	Std. Duration (s)		1251	>0.999	0.077		
	Counts per Minute		1593	>0.999	0.055		
	Max. Duration (s)		1045	>0.999	0.031		
	Mean Duration (s)		1190	>0.999	0.199		
	Ratio (%)		1160	>0.999	0.203		
		Std. Duration (s)	1158	>0.999	0.131		



# Appendix C

## Acronyms

**TSST** Trier Social Stress Test

**f-TSST** friendly Trier Social Stress Test

**GPU** graphics processing unit

**S3** simple storage service

**VPN** virtual private network

**JSON** JavaScript object notation

**SD** standard deviation

**IMU** inertial measurement unit

**JWT** JSON web token

**UI** user interface

**CNN** convolutional neural network

**FACS** facial action coding system

**ORM** object-relational mapping

**SNS** sympathetic nervous system

**HPA** hypothalamic-pituitary-adrenal

**COCO** common objects in context

**MPII** Max Planck Institute for Informatics

**AWS** Amazon Web Services

**GDPR** General Data Protection Regulation

**API** application programming interface