Received: 16 June 2024 | Revised: 20 September 2024 | Accepted: 16 October 2024 | Published online: 2 January 2025

RESEARCH ARTICLE

Artificial Intelligence and Applications
2025, Vol. 3(1) 10–30
DOI: 10.47852/bonviewAIA42023630

BON VIEW PUBLISHING

A Methodology for Rapid Deployment of Diagnostic Applications for Human Gait & Posture Analysis and Exoskeleton Configurations, Combining AI with a Biomimetic Rig

Christos Kampouris[1,2,*] and Philip Azariadis[3]

[1]*Department of Product and Systems Design Engineering, University of the Aegean, Greece*

[2]*Department of Informatics and Computer Engineering, University of West Attica, Greece*

[3]*Department of Industrial Design & Production Engineering, University of West Attica, Greece*

Abstract: Research has shown that certain simple, yet valuable gait and posture diagnostic tests can be conducted without the involvement of human experts using optical detection software employing Artificial Intelligence (AI), either on par with some physicians or, in some cases, even better in terms of availability, productivity, ease, and cost. The purpose of the current paper is to propose a methodology that uses dedicated AI models for partial body frames instead of full-body pose trackers. To achieve this, we obtain training dataset (photos/videos) using a biomimetic rig instead of humans. In this research, we also study existing applications using already available AI pose trackers, their methodology, limitations regarding partial body views, and limitations in detecting mechanical devices and systems. Finally, to support our study, we present stand2squatAI_biorig software as an example of an AI-automated, diagnostic, real-time test that does not require physicians to complete the examination. We developed this software as a proof-of-concept to illustrate the proposed methodology and confirm our findings. Software applications developed in this manner can be used to study and diagnose various human conditions. The value of the proposed approach, which includes a biomimetic rig, is that it increases precision, reduces costs, and increases human safety; for example, during the application of mechanical aids while ensuring personal data privacy and overcoming ethical issues.

Keywords: diagnostic applications, gait and posture analysis, biomimetic rig, AI markerless motion capture, exoskeletons, prosthetics, personal data privacy, ethics issues

1. Introduction

Gait and posture analysis may offer invaluable diagnoses in numerous cases, ranging from congenital motorial problems to stroke rehabilitation assessments. Unfortunately, it is neither feasible, practical, nor viable to obtain laboratory-based measurements at the required frequency, especially for older adults, the unemployed, people lacking healthcare, rural, nomadic, and other significant portions of the population. However, technological progress has led to the miniaturization of electronics and the widespread availability of powerful processors, camera-equipped Internet-accessing smartphones and tablets. Specialized and potent artificial intelligence (AI) software tools based on computer-emulated deep neural networks may provide almost human perception to these smart devices, enabling them to optically detect human bodies, faces, and gestures. They may also recognize animals, flowers, or what else they are trained to do. These tools may rely on AI environments, frameworks, and ecosystems, such as TensorFlow [1], keras [2], Pytorch [3], H20.AI [4], OpenVino [5], Caffe [6], MXNet [7], PaddlePaddle [8], and ONNX [9]. Training is performed by presenting samples (e.g., photos, sketches, and videos) that are properly tagged or organized into enumerated sets. Beyond detection, AI tools may infer various properties such as the distance of a person from the camera, their age, even dimensions, volume, and location of their limbs.

AI pose tracking software, such as OpenPose [10], AlphaPose [11], Detectron [12], MediaPipe BlazePose [13], YOLOv7 [14], DeepLabCut [15], process photographs or video streams with one person or a crowd and produce lists with the 3D coordinates of human joints in each video frame. This capacity surpasses most common people's abilities and undoubtedly marks the era of nonhuman experts. Researchers have used this process to develop automatic gait and posture analysis tools such as Chen's sitting posture assessment [16], quantitative movement analysis [17], occupational posture evaluation [18], and Stanford's

*Corresponding author: Christos Kampouris, Department of Product and Systems Design Engineering, University of the Aegean and Department of Informatics and Computer Engineering, University of West Attica, Greece. Email: cab@aegean.gr

Sit2Stand [19]. Such tools are promising and guarantee that in the following years, health coverage will increase substantially while the related costs will diminish. The following paragraphs discuss existing diagnostic exercises, human pose trackers, and their limitations while roughly outlining the standard methodology used in the previously mentioned tools. Moreover, we focused on the precision, variance, anonymization, and ethical issues.

The main contribution of this study is the development of an alternative methodology that can be used to accelerate the progress of automated diagnostic tests, reduce associated costs, produce real-time diagnostic tools available to all, and further expand the applicability of AI optical tools in the study of mechanical joints, prosthetics, orthoses, exoskeletons, and rigs. As an example of applying this methodology, the sample software tool stand2squatAI_biorig was developed and presented. It works even with partial frames rather than full-body poses, surpassing the problems of latency, missing joints, and erroneous joint coordinate reporting, which are common in most pose trackers.

In the field of gait and posture analysis, diagnostic exercises play a significant role. One of the most common tests used by physicians to examine patients is the 5XSST (five repetitions sit-to-stand) test. Using a stopwatch, a physician measures the time it takes for the subject to stand up, from sitting on a 17″ height chair that normally rests against the wall, and sit down again, as quickly as possible, for 5 consecutive repetitions. The measured time is indicative of the person's fitness and is related to age, health (mental and motorial), medications, and other factors, such as body type, weight, and occupation. A variation is the 30 seconds-sit-to-stand (30CST) test, in which a physician measures the number of sit-to-stand repetitions achieved within 30 s.

Depending on the condition, numerous other tests may be suitable, including the Functional Reach Test [20], 6-min walk test [21], four square step test [22], timed up-and-go [23], sit-and-reach [24], passive-knee-extension-test, active/passive straight-leg-raise [25], or combined elevation test.

Physical tests of a similar philosophy are widely used, for example, from determining the most appropriate sports for young athletes, deciding whether a person has suffered a mild stroke, or judging if a driver is drunk, at least before the adoption of specialized breath-analyzing devices. The same or adapted tests may be applied to robots to assess their speed, dexterity, stamina, endurance, or fitness for use in a particular task. In fact, the more advanced and anthropomorphic a robot is, the better it is expected to score. Therefore, there are easy-to-perform motion tests that offer significant conclusions for patients, based on simple time measurements between certain poses. One of the targets of the current research is a methodology for developing such tests, employing AI for these simple time-measuring tasks that can be performed with a smartphone without the luxury of a physician. The structure of the paper is focused on the presentation of the proposed methodology: after the Introduction section, we present the Related Work (Section 2), analyze our Proposed Methodology (Section 3), and back our exegesis in the Experimental Results (Section 4). After the parathesis of Comparative Study (Section 5), Conclusions (Section 6) are drawn at the end.

2. Related Work

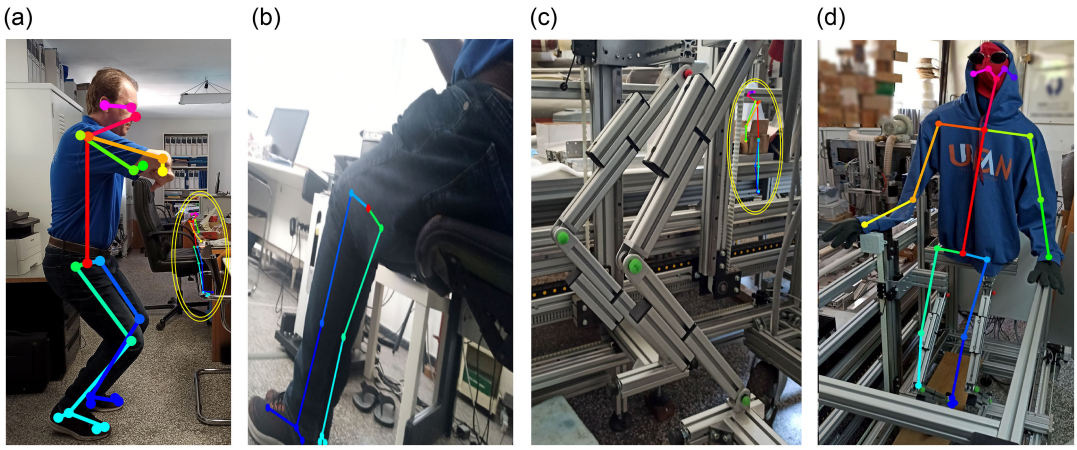
The most popular pose trackers or pose-estimators (as mentioned previously) employ computer-emulated neural networks to produce lists of 2D/3D joint coordinates from photographs or videos (Figure 2(a)), offering markerless motion capture [26]. They differ in speed, single/multi-person capability, number of reported keypoints including joints and features such as eyes and nose, locations of keypoints, detection precision [27, 28], ease of use, complexity of integration in other applications, and being proprietary or open. Some may perform better than others under specific applications, sports, or lighting conditions. Through further training, some may become more specialized in certain fields, such as crowded spaces, underwater activities, or yoga exercises. However, a few of the popular pose tracker issues (PTI) associated with the quality of measurements (Figures 2 and 3) are presented in Figure 1, which compares a typical AI pose tracker such as the popular OpenPose vs an alternative approach (stand2squatAI_biorig) which will be discussed in following sections.

Issues such as missed joints (PTI.b) and erroneous coordinates (PTI.c) are reasonable to expect and perhaps beyond the tracker creator's scope of intention. Experimentation and practice suggest that avoiding complicated multi-person scenes and close-up scenes significantly reduces missed joints and coordinates. Thus, whole-body frames are preferable for trackers to achieve a better

Figure 1
Popular pose tracker issues

	Description of Popular Pose Tracker Issues (PTI)	OpenPose	stand2squatAI biorig
PTI.a	Jitter in reported coordinates. With actual joints relatively static or slowly moving, estimated keypoints appear to jump around within or even out of the human body. This manifests as high-frequency noise and may be more common for pose estimators that identify keypoints in each frame separately, without trying to track the trails of those keypoints in subsequent frames.	× a video is required to demonstrate this issue	✓ many tests can be performed by classifying stances, without using joint coordinates
PTI.b	Missed joints due to occlusions e.g., hands behind the back poses, or due to partial frames (body parts being out of frame)	× see Figure 2.b	✓ partial frames are acceptable and preferable
PTI.c	Erroneous coordinates due to unusual stances, proximity to other humans or objects, small number of detected limbs, or unusual clothes/shoes/objects which the AI tracker has not been trained for.	× see Figures 2.b, 2.c	✓ partial body frames and focus on few limbs
PTI.d	Inability to identify mechanical or electromechanical and pneumatic systems such as artificial joints and exoskeletons.	× see Figures 2.c, 2.d	✓ mechanical parts recognized due to AI training dataset

Figure 2
Examples of issues encountered with OpenPose under challenging circumstances



performance. However, this has major consequences, as will be discussed in the following paragraph.

OpenPose excels at human posture recognition (Figure 2(a)). Nevertheless, it can be fooled by objects (see the yellow detail in Figure 2(a)). Using only partial frames results to erroneous coordinates (e.g., hip misplaced at knee in Figure 2(b)). Passive limb models on the rig are not recognized; instead, "background noise" from different objects results to false identification (Figure 2(c)). Figure 2(d) demonstrates that a clothed, upper-body mannequin and the "right perspective" (view angle) push OpenPose to identify a full human body, although it misplaces knees and ankles.

2.1. Shortfalls associated with full-body frames (FBFS)

1) FBFS.1: Loss of subject's anonymity during early training stage. The term "subject" refers to human volunteers that are used for the formation of a training dataset

Datasets for the AI training stage, such as photographs or videos of human subjects, cannot be secured. These data are necessary for the AI system to adjust the proper "synapse weights" that will enable the identification of features within the specific limits of each "class". The more the classes, synapses, and network depth, the more instances that can be identified by the AI, and the more data required for training. Preprocessing anonymizing methods may be applied, such as video editing to cover the eyes and other identification marks. Such solutions increase effort, time, and cost, but cannot secure neither the early data collection nor the storage stages.

Our proposal to infuse a biomimetic rig to mimic human stances during the training of the AI system, resolves all such issues, because no human subjects are involved in the training stage. All poses can be performed countless times by the rig for photo or video capture, for as long as the AI needs to mature.

2) FBFS.2: Loss of user's anonymity during tracking (usage) stage. The term "user" refers to a patient that would use the application for medical diagnosis reasons

If the AI system requires full-body coverage during video recording, then the patient's identity cannot be protected unless the subject wears a mask, covers birthmarks, tattoos, scars, and other characteristics. It can be argued that the tracking stage corresponds to the examination procedure in a doctor's office. However, a doctor is not necessary for an automated test because the test is intended to be initiated and performed by the human subject (the "application user"). Hence, all relevant privacy measures should be applied unless the participant is asked to disclose personal information voluntarily, which would reduce the test's adoption percentage in the patients' community.

Our proposed methodology, by focusing only on the joints under study, thus not exposing the entire human subject, resolves the anonymity issues compared to existing posers that require full-body coverage.

3) FBFS.3: Need for committees and supervision

To alleviate the previous (FBFS.1) and (FBFS.2), ethics committees must ensure that all measures are taken, to avoid mishandling of personal data; these measures are both costly and time consuming, in addition to the cost of assembling each committee. Unfortunately, various of the proposed countermeasures have failed, even in the best institutions, to public dismay [29–31]. It is an almost common belief that digital data cannot be protected if they become target of interest.

Therefore, our proposal does not record, store, or transfer to cloud/centralized mainframes any data. It conducts the developed test locally on the human's personal phone/computer and gives him the authority to further examine the results in any way that is convenient to him.

4) FBFS.4: Reduced resolution

Choosing whole-body frames over close-ups greatly reduces the available camera resolution, because the smaller the area under study, the smaller the useful resolution. This mishap occurs due to the inability of popular posers to identify certain features unless a full body is included. Furthermore, wide area homogenous lighting and "clear backgrounds" are difficult to achieve outside studios. Such factors affect the image quality, and hence, the locally effective resolution.

Thus, close-ups can be easier for users and are favorable for cameras and AI models.

5) FBFS.5: Reduced potential processing speed

Owing to increased neural network complexity, the larger the area and the greater the number of limbs and anatomical features to detect, the heavier the processing burden to deal with. Consequently, low-cost systems cannot achieve real-time operation. For example, a noncutting-edge desktop computer without GPUs may require approximately 20 s per frame to run OpenPose.

Our proof-of-concept software stand2squatAI_biorig, runs practically in real time, producing immediate results.

2.2. Addressing the full-body frames and other obstacles

To address the issue of common pose trackers versus partial human images, one approach would be to train AI to extrapolate the missing information [32], or to combine it with Bayesian inferencing, or to use Generative Adversarial Networks to produce plausible fill-ins [33], or even to compose a whole-body image by adding a photo of the missing parts as in S. Dvir's Partial-openpose. This photo may belong to the same or another human or be computer-generated. In general, it will need proper scaling, cropping, positioning, and blending to fit smoothly. Otherwise, the accuracy of the actual joints' coordinate estimation might be reduced. This approach would yield the best results for simple exercises without body turns. It could cure (FBFS.1), (FBFS.2), and (FBFS.3), but not (FBFS.4), because the actual partial image must be scaled down to occupy its normal area in the composed whole-body image.

Issue (PTI.d), the inability to identify mechanical systems, as illustrated in Figure 2(c) and (d) is clearly a matter of AI training. In Figure 3, note that although the background is purposely full of artificial structures and mechanical parts, OpenPose excludes all such "noise" and detects only human-like presence.

As seen in Figure 3(a), OpenPose recognizes the torso (side view) and hands of the mannequin, but not the artificial lower limbs. Putting tracksuit bottoms at the mechanical limbs (Figure 3(b)) helps identify thighs (femurs) but not tibia, or ankles. Placing shoes near the limbs (Figure 3(c)) helps identify ankles in the partial frame, but erroneously. A full-body frame, with clothes, enables OpenPose to identify all joints, as in Figure 3(d).

This weakness is important when it comes to detecting artificial limbs attached to patients or biomimetic rigs. It is feasible to further educate human-focused AI tools to include, e.g., mannequins and anthropomorphic robots, and this has been done to some extent, but with consequences:

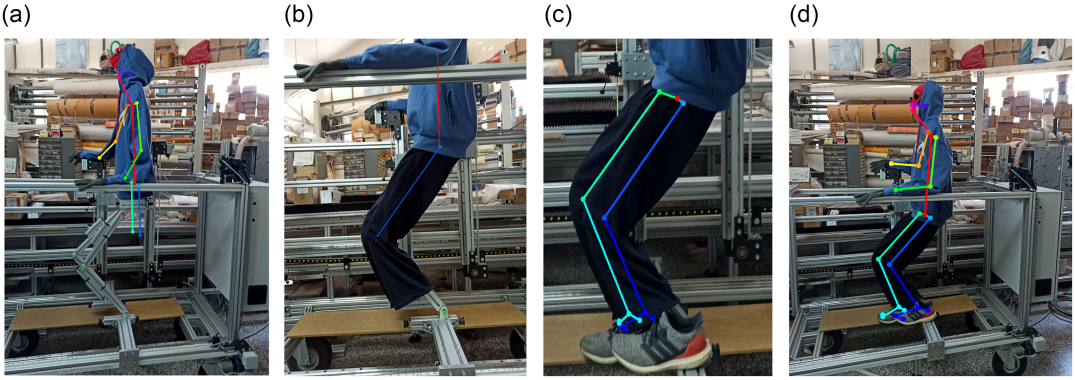
1) The more we ask an AI model to perceive, the more its complexity increases, requiring faster and more powerful computers, more training dataset, and more storage space, thus multiplying the effects of consequences (FBFS.4) and (FBFS.5).
2) There are no limitations regarding the design of artificial structures or choice of components, including the type and number of joints. Thus, a considerable effort to keep general use pose trackers constantly up to date would be required, and in vain.
3) Mechanical systems are typically developed by using a bottom-to-top approach. In other words, subsystems are designed, fabricated, tested, and improved until proper integration. This necessitates the availability of partial frame-focused evaluation systems because the full-body (e.g., a robot) is neither assembled nor operational until the latest project stages.

2.3. AI-automated diagnostic tools

Some diagnostic tests have been automated using AI and a smartphone or camera-equipped computer to capture the subject's motion and extract results without physician's intervention. Amazon's Halo Movement application [34] asks the subject to place a smartphone on the floor in front of them and execute 5 tests: single-leg stance, forward lunge, overhead squat, overhead reach, and feet-together squat. The application assesses mobility, stability, and posture, and reports "movement health scores" from 0 to 100. The app seized to function as of Aug. 1st, 2023, and uploaded data were deleted according to Amazon's announcement.

Stanford developed an automatic diagnostic tool for the 5XSST test [19], among others. An untrained individual is asked to record the subject performing 5 sit-to-stand repetitions from a 45° angle and upload the video to the cloud. OpenPose [10] is used to extract joints' coordinates. Using these coordinates, Stanford's algorithm identifies the standing stances and calculates motion durations and kinematic parameters, such as trunk angle, speed, and acceleration. It then generates a scientific report that, until recently, a specialized physician would be able to compile only with expensive motion capture equipment and dedicated software.

Figure 3
Using mannequins to force identification of mechanical limbs



The test results could be used to assess for example the subject's osteoarthritis status or mental health depending on underlying conditions. To use this tool in a nation-wide study, Stanford obtained digital informed consent from all 493 participants used in the tests.

2.4. Ethics

Whenever humans are involved, especially patients, ethics are of utmost importance. The following are three examples of good practice:

1) The Stanford University Institutional Review Board approved and oversaw the 5XSST [19] study protocol.
2) All nation-wide participants were asked for their consent.
3) To publish a human subject's photo in the paper by Boswell et al. [19], Stanford did not use a study participant or a student, but an actor who gave consent, even though the photograph is small, low-resolution, with superimposed joints' dots and connecting lines.

In general, training datasets may be generated by recording consenting human volunteers. Datasets from young and healthy individuals have been quite straightforward to obtain this way. However, diagnostic tests must be trained for the elderly and patients too, to be of substantial value. Besides ethical issues, practical problems are also significant. In addition, the advent of ChatGPT, Midjourney, and other text and art generating/processing AI tools which have been trained with "public" data from social media and corporate webpages, films, news broadcasts, to library and museum collections available on the internet, has had some serious effects: it raised the public awareness regarding AI training and laid the ground for litigations concerning intellectual property rights and privacy infringement.

AI is not to blame solely. In our age of information, data privacy, personal information leakage, and identity theft have been the subject of heated discussions and legislative actions worldwide. Unfortunately, as mentioned earlier, digital data are unprotected. Consequently, new automated diagnostic tools should gain the public's trust, to be widely used.

2.5. Common methodology

So far, the common methodology to automate a gait and posture analysis test using pose trackers can be roughly summarized as follows:

1) Choose the diagnostic test to automate.
2) Through initial trials, confirm that the task is feasible; that is, confirm that:

   2-1) the sequence can be video recorded from a fixed distance and angle,
   2-2) without occlusions that would cause unavailability of critical joint coordinates,
   2-3) with enough precision to extract meaningful results, etc.

3) Employ a pose tracker software to generate joint coordinates from the subjects' videos.
4) Write dedicated software to process the joints' coordinates lists.
5) Generate useful metrics for diagnosis or develop another dedicated AI model for this job.
6) Prepare some instructions for the subjects to follow.

Opportunities:

1) Smartphones and tablets are widely available.
2) Marker-less motion capture offers great ease plus user friendliness and foregoes the need for an expert to place markers.
3) Pose tracking software is available.
4) Universities and research institutions have ample manpower with software engineering skills to implement the dedicated joint coordinates processing and diagnosis inferencing.
5) The global population can greatly benefit from automated diagnostic tests.

Hence, the previous methodology will certainly produce remarkable results. Yet, there are some issues associated with the toolchain and, most notably, frame capture and processing. On one hand, optical detection tools endanger privacy. On the other hand, pose trackers have been optimized for producing coordinate lists, but this feature is still computationally intensive and may not be the only solution to a specific test's requirements. Our proposed methodology offers an alternative, as explained later.

2.6. The era of publicly available AI tools

Recently, several AI model-building environments and customizable AI tools have become publicly available, either free or with subscriptions, or other financial arrangements. For example, Google provides Teachable machine [35], based on the tensorflow.js library for machine learning, freely. It is an easy-to-use straightforward tool which can be trained to detect and classify images or sounds. The Teachable machine AI environment is based on deep learning on convolutional neural networks and relies on labeled data for training and making predictions [36]. It uses the "transfer learning" technique. With this approach, a MobileNet model has been pretrained with a vast amount of data and is now used in a more specific task, by replacing a minor part of it with a top layer based on the new data. This technique requires a smaller dataset and has less processing burden.

In this context, classification works as follows:

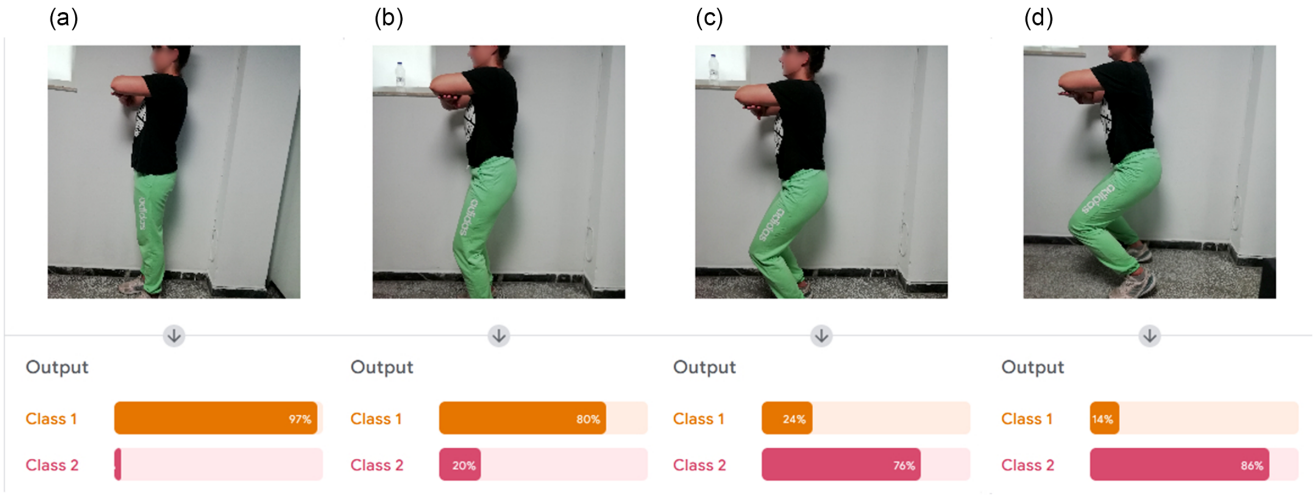
1) During model preparation, AI is presented with at least two classes (sets, collections) of photos. For example, one collection with photos of cats and another one of dogs. Each class has a number or tag.
2) Next, AI is asked to automatically train how to discern between classes.
3) From then on, when AI is shown an image, it will identify in which class it would fit best, reporting match scores. For example, if presented with a photo of a sheep, it might reply to cat:30% and dog:70%. The sum will always be 100% (Figure 4), and the image will always get classified (even if it is a lunchbox) because that is what the AI is asked to do. Naturally, neural networks cannot be 100% correct, especially if not trained adequately, or if, for example, the lighting conditions or viewing angles differ significantly from those in the training sets. In our case, Figure 4(a)–(d) demonstrate AI classification results for different squat depths, when asked to classify either standing (class 1) or shallow squats (class 2).

The trained model can be uploaded to a webpage accessible to everyone. Furthermore, it can be exported to programming environments such as TensorFlow [1], TensorFlow Lite, or p5js [37]. There it can be incorporated with the dedicated code which will use the AI classification results.

3. Proposed Methodology

A great role in the proposed methodology is held by the programmable Biomimetic Rig with four independent linear

Figure 4
Images from our AI setup (described in Section 4.3). Subject: Female, 170 cm. View: Sagittal plane, left knee. Full frame



motion axes. It can emulate walking cycles or other sequences, e.g., stand & walk, jog & run, squats, sitting & standing repetitive cycles, offering major advantages versus human subjects: memory, repeatability, stamina, adaptability, and controlled variability. The attached passive limbs are interchangeable and length adjustable in order to match different physiologies. Each has three rotating joints: hip, knee, angle, and can be mounted on the rig at various orientations [38]. As previous research has demonstrated, such a biomimetic rig [39], may be used for a multitude of tasks in studies mainly focusing on the lower extremities, such as:

1) Replicating human stances and gait cycles, in a precise manner, without intra-trial/inter-trial variance.
2) Studying and evaluating passive and active prosthetics and exoskeletons.
3) Adjusting dimensions/tuning parameters/assessing performance details of a rehabilitation device (e.g., a robotic walking assistant) according to a particular human patient's body and condition as emulated by the rig (before the rehabilitation device is attached to the patient).

Cases (1) – (3) take advantage of the fact that a human motion sequence recorded once, for example in a gait clinic, may be emulated by the biomimetic rig for numerous experiments under various circumstances and different instrumentation sets. In the scope of the general population, there are public datasets with joint coordinates obtained in labs with "golden standard" motion capture equipment [40], and datasets with kinematic data from wearable sensors [41], from force plates [42], etc. Thus, the datasets may be partial, containing precise information for some body segments, but not sufficient enough to interpolate for the whole body.

As discussed previously, common pose trackers:

1) Do not recognize mechanical systems, and these include the rig, too,
2) Do not work well with partial frames, and currently the rig does not include a full skeleton structure.

Hence, diagnostic tools using common pose trackers cannot be used with a rig. The rig may be fitted with clothes, to "disguise" its mechanical structure as in Figure 3, and an upper-body mannequin may be added on top. However, from a scientific perspective, adding a non-driven mannequin to achieve whole-body frames would just reduce the diagnostic tool resolution available for the lower extremities of interest. Consequently, another toolchain featuring a dedicated AI model should be built to apply diagnostic tools on rig-emulated ailments and orthotic setups. Prior to serving as an AI diagnosis subject, the rig can assist in the development and testing of the toolchain and the specific AI model by providing datasets (photos and video or kinematic data).

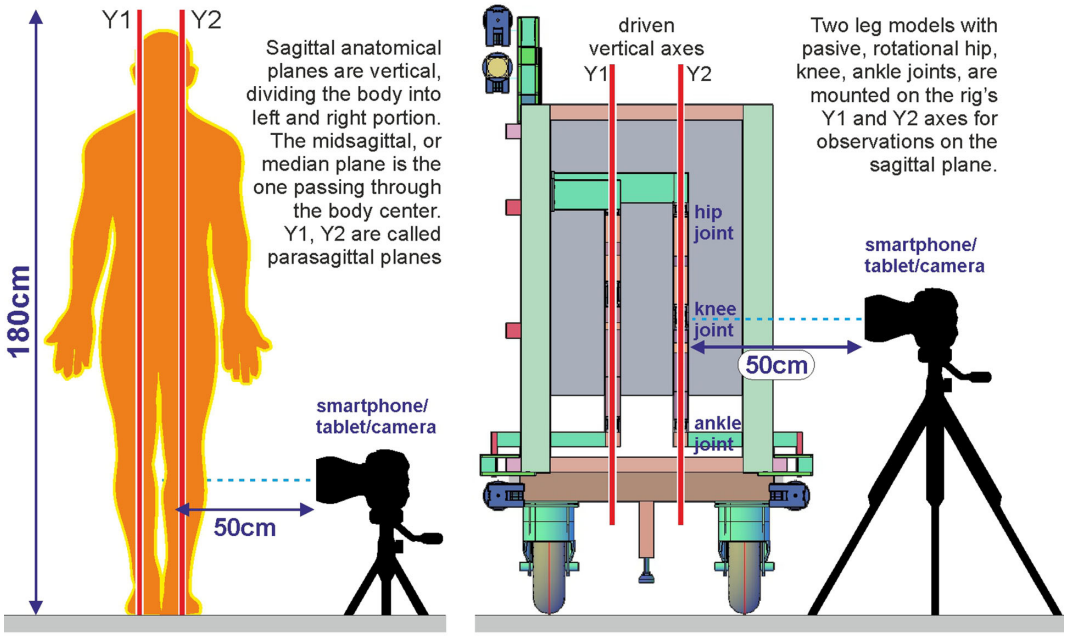
Therefore, employing a biomimetic rig to act as a training platform (Figure 5(b)) could minimize such issues. Using the rig to generate the training data is simple:

1) Put clothes on the rig (if the training dataset is for diagnostic tools also applicable to humans),
2) Instruct the rig to adopt certain stances (from public datasets or previously anonymized lab data),
3) Take a few photos or short videos.

Note that Figure 5 presents two alternative setups. Figure 5(a) shows the setup for partial human observation for the diagnostic test 5 x stand-to-shallow squat, described in Section 4. Figure 5(b) shows the biomimetic rig [38] and camera setup for emulating humans, in order to collect both training and testing datasets.

Steps (2) and (3) can be automated by connecting several static, remotely operated cameras to the rig or by connecting a single moving camera on a free axis of the rig (which will be programmed to move the camera at different spots) or with a camera on a robotic arm or a drone, even with a static camera and actuated mirrors (such as "galvo" optical heads). If the dataset requires videos of certain sequences, then, likewise, the rig can be instructed to perform and repeat the same sequence, while video is recorded as many times as necessary, with different clothes, shoes, lights, etc. To train the AI against clothing artifacts such as wrinkles, textures, and shading variances, clothes are necessary. In general, the effects of clothing are important for markerless motion capture. As for gait research institutes and clinics with "golden standard" equipment, markers are most often placed on the skin. Leggings and tights are not recommended because fabrics may slide over the skin as the limbs move. Clothes may not be used if training data are required for diagnostic tools that

Figure 5
Observation setups



are only applicable to the rig, or for tools running on radiology (X-ray) images, or other types of sensors. For every stance of interest, a few photos from slightly different angles and slightly different distances are necessary to train the AI model effectively. Too few photos may mislead the AI tool in detecting, for example, a certain fabric texture or shade rather than an anatomical feature. Even if the initial AI training proves inadequate, this is not an issue for the biomimetic rig: some poses or stances can be repeated, for example, with different garments, since the same data feed reproduces the exact same pose.

3.1. The proposed methodology for diagnostic tool automation

If full-body frames are not necessary for diagnosis purposes, an alternative approach can be followed with a dedicated AI model instead of a pose tracker. However, even if whole-body frames are necessary, it might still be beneficial to develop a dedicated AI model focusing, for example, on particular anatomical features instead of joint coordinates that could be provided by a pose tracker. Figure 6 compares the common "full-body + available pose tracker" methodology (left column) of paragraph 2.5, against the proposed "partial frame + dedicated AI model" methodology (right column). Similarities are extensive. Green color marks new stages; yellow differentiated ones; blue stages are identical.

Restrictions (2.1) – (2.3) are common in both cases. If a full-body frame setup satisfies them, the partial frame setup will also succeed if (2.0) holds. In general, it would be quite difficult to prepare an AI model that would outperform any of the available pose trackers unless willing to spend significant effort, time, and resources. However, for the case-by-case scenario, with a single diagnostic test examined each time, simple AI models may prove quite adequate. In (3.2), the terms quantitative and qualitative refer to AI deliverables. Qualitative results are, e.g., the labels "standing"/"squat" employed in our example (paragraph 8), or the labels "good posture"/"bad posture" [16]. A quantitative result could be "knee flexion = 63°", i.e., a numeric value for a certain variable. The term "better" in (3.3) refers to the personal information privacy aspect, as well as precision and partial data availability issues. In certain cases, it would even be possible to omit the dedicated software to process the AI model's classification results of step (4):

1) If the required diagnosis is qualitative and rather "intuitive" (e.g., sitting posture assessment), the AI model could generate it directly.
2) If the required diagnosis metrics are rather simple (e.g., minimum or maximum knee flexion angles), the AI model could infer them directly.

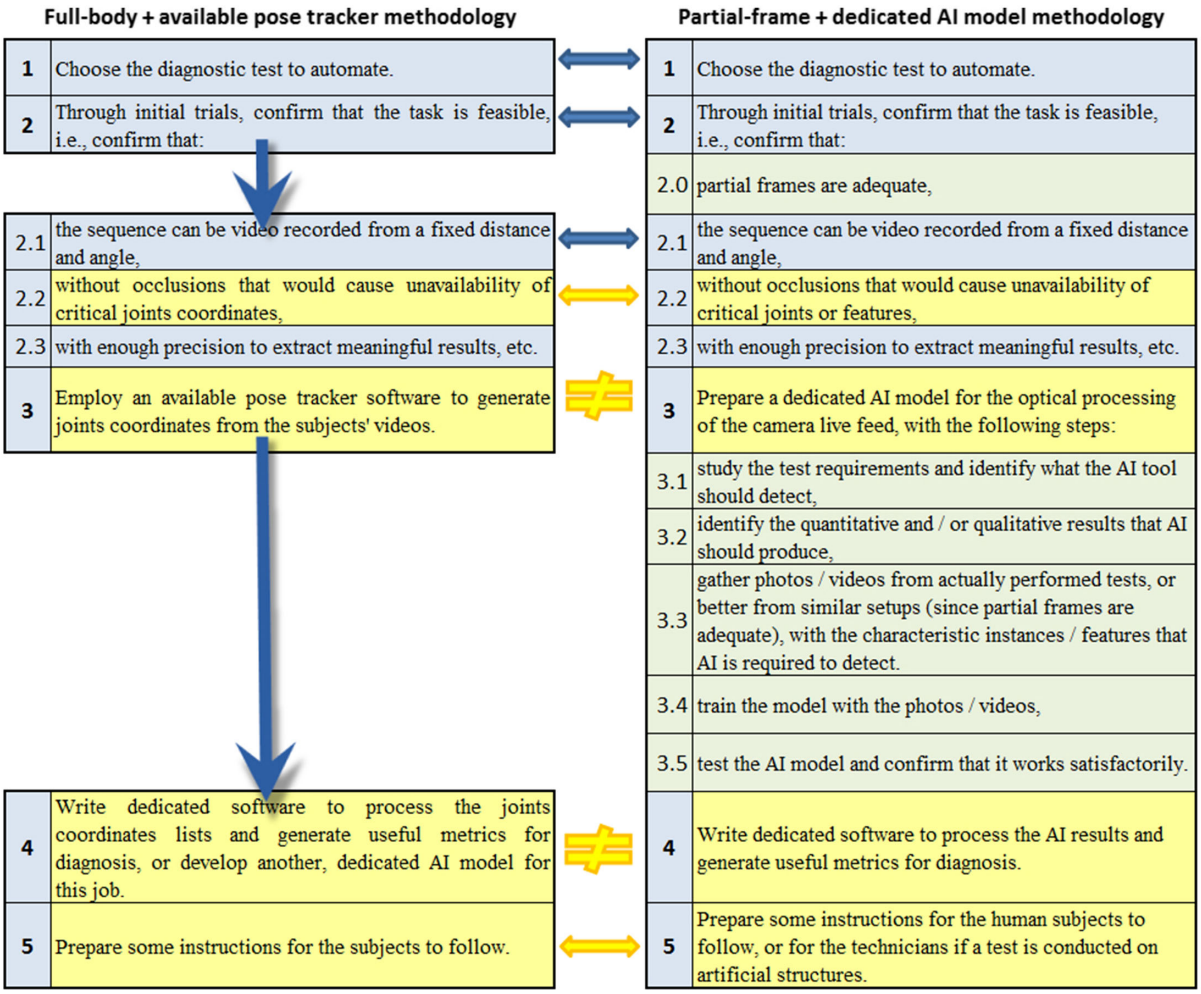
Alternatively, a posture assessment can be dealt for example, with:

1) a pose tracker + dedicated code, or
2) pose tracker + secondary AI model [17], or
3) pose tracker + dedicated code on dedicated hardware [43], or
4) special RGB-D camera (color and depth with included pose tracker) plus dedicated software [44]. Nevertheless, by omitting the pose tracker and adopting a partial frame approach, it would not be necessary to photograph the subject's head or face.

4. Experimental Results

To illustrate the proposed methodology, while showcasing the power of binary classification, we chose our diagnostic test called "5 x stand-to-shallow squat" for automation and developed the "stand2squatAI_biorig" software as a sample of a diagnostic tool. Conventionally, a physician would ask the subject to perform five consecutive shallow squats, measure the total exercise duration, and calculate the average cycle duration. This test can be automated, that is, performed by the purposely developed software

Figure 6
The differences between the common and our proposed methodology (on the right)



which incorporates an AI model to process the camera feed in real time and then calculate and report each cycle's estimated duration, average, minimum, maximum, and standard deviation. The AI model is trained and tested using close-up photos of the biomimetic rig performing the exercise, so that no human datum is used. The rest of this section discusses the "stand2squatAI_biorig", development process, according to the proposed methodology stages.

1) Methodology stage 1: We choose to automate 5 x stand-to-shallow squat diagnostic test

Test description: Measure the duration of five repetitions from standing straight to a shallow squat and back. The squat is shallow (knee flexion angle no more than about 90°), with a relatively straight back, so that the elderly can perform it. During exercise, the hands are folded on the chest, or touching the waist, provided that the knee is not occluded. The feet should be parallel and steady, always in contact with the ground without moving. The exercise is relatively mild. People who find it difficult or painful (e.g., osteoarthritis patients) may benefit from sliding their back on the wall or a suitable smooth surface. Alternatively, they could rest against a purposely built vertical sled, which could also limit the range of motion and prevent falls. Elderly patients and patients with kyphosis do not need to stand with their backs completely straight. The term "standing straight" refers more to the legs (knees), to the extent possible for each subject.

Necessary test tools: A physician would use a stopwatch to measure the total time and a calculator to divide the average cycle duration by 5.

Stance definitions (Figure 7)**:** When standing with fully extended knees, the knee flexion angle is 0° (included angle = 180°). A range of +5° (under-extension) to −5° (hyper-extension) is considered normal. At a shallow squat stance, knee flexion is approximately 90° or less.

2) Methodology stage 2: Confirmation of feasibility through initial trials

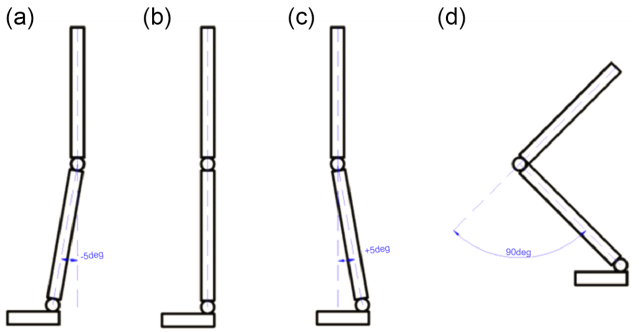
Stage 2.0: Partial frame adequacy

Only a knee close-up is necessary to measure the knee angle, or alternatively to detect whether the subject is standing or performing a squat. Partial frames are also desirable, as the system benefits from them over full-body frames.

Stage 2.1: Fixed distance and angle observability

A sagittal view with the camera at knee height, approximately 47 cm away for a subject approximately 170 cm tall, is adequate. The

Figure 7
(a) Hyper-extension −5deg, (b) straight standing stance 0deg, (c) under-extension +5deg, and (d) shallow squat ~90deg



camera distance is chosen so that one knee and legs are visible inside the frame both during squats and while standing, without having to move the camera. The feet, waist, torso, hands, and head need not be visible. In fact, the fewer parts visible, the easier it is for the AI tool to classify knees. The general objective is to constantly maintain the feature under detection inside the observation frame.

Stage 2.2: Occlusions check

The knee is not occluded from the selected viewpoint, provided that the subject does not wear long blouses. A pair of blue or black jeans or other comfortable trousers against a white or light-colored wall for increased contrast would be ideal.

Stage 2.3: Precision check

Initial trials confirm that precision is sufficient to extract meaningful results. In any case, the precision is expected to be approximately double that of a full-body frame arrangement due to the dimensions of the captured images.

Consequently, the task seems feasible.

3) Methodology stage 3: Preparation of the AI model for optical processing of the camera live feed

Stage 3.1: Diagnostic requirements

AI needs to detect either the knee flexion angle or the included angle in each video frame and relay the result to the algorithm that will measure cycle durations and extract the minimum, maximum, and average. However, to identify each exercise cycle, an increased knee angle resolution is not necessary. The algorithm used to measure the cycle duration could work with a resolution of 1o, 10o, or even 45o. The issue is analogous to the perception of a gray-scale image: nature affords countless shade variations; on screen, 256 steps from black to white seem quite smooth, yet a picture can be converted to 16 shades and remain fairly recognizable. Furthermore, a picture can be converted to black-and-white (two shades) and still be of significant value if there are image areas with sufficient contrast. For timing the exercise in the stand-to-squat test, it is necessary to discern 2 stances only. Therefore, AI needs to discern between the two classes corresponding to the two stances. Alternatively, if it was required, e.g., to continuously track the knee angle during the exercise, a 2-stance solution would be inappropriate. Naturally, the fewer the required training classes, the easier the AI's task, and the fewer the erroneous classifications to handle during/after frame capturing.

Stage 3.2: AI deliverables

Consequently, the AI model is required to label each video frame, either as "Class 1" or as "Class 2", with the latter one referring to the shallow squat.

Stage 3.3: Training dataset collection

Two photograph collections could be generated with the help of the rig or a consenting volunteer. One photo per collection is the least absolute. Three or four photos per collection could be sufficient for similar types of clothes. More photos per class do improve the outcome. In our example, approximately 20 photographs were used per class.

Stage 3.4: AI training

After uploading the photos in each class, training Teachable machine is automated, requiring only one action (Figure 8).

Stage 3.5: AI testing

Testing is immediate, with the help of a camera observing the rig movement or by uploading other rig photos that have or have not been used during training (Figure 9).

It should be emphasized that the training dataset was created with the use of the rig. The rig setup focused on knee flexion/extension angle (between femur and tibia) measurements. We photographed it at various stances, at various conditions, e.g., bare mechanism, clothed mechanism, from various viewpoints. Instead of relying on a publicly available but uncontrolled set of pictures, we chose to build our own dataset with the same camera setup and the same quality of data received. This way we achieved a dataset that is clear and correctly labeled right from the beginning. Secondly, we tried to have the two classes balanced, that is of the same magnitude of information for both. Finally, through a process of selection and omission, we reduced the dataset to be of high performance but of small size. This way the AI would perform in a real-time manner. Finally, class 1 includes 19 "standing stance" and class 2 includes 22 "shallow squat stance" color photos of 4160 × 3120 pixels resolution in jpg format. Note that Teachable machine downsamples the photographs internally.

4) Methodology stage 4: Creation of dedicated software to process the AI results and generate useful metrics to help diagnosis. To calculate the exercise cycle durations, some coding is necessary. The completed AI model was uploaded directly from Teachable machine [35], into the p5js editor [37], and the analysis algorithm was embedded in its sample (reference) code.

5) Methodology stage 5: Preparation of instructions for human subjects, or corresponding directives for the involved technicians, if the test is conducted on artificial structures. In our case, instructions have been included in the program's initial screen. Only 1 icon is required for operation (the "start/stop test"). The remaining icons regard the filtering parameter (Section 4.3, Classification filtering) and offer pseudo-data for experimentation if a camera is not available on the computer.

Closing on the application of the proposed methodology, we note the following: Due to the partial frame approach and close camera distance, self-recording is possible. Since the stand2squatAI_biorig measures the time between two clearly distinguished stances, we only need two classes to timestamp our tests to produce measurable results. The combined merits (Figure 10) of the partial frames and the few-class approach (i.e., 2-stance detection implementation) are numerous. Among these, we can emphasize on the following:

Figure 8

(a) Preparing the two class AI model on Teachable machine. (b) Testing the classification results for class 1 "standing stance" with photos that were not used during training stage. (c) Testing class 2 "shallow squat stance" classification

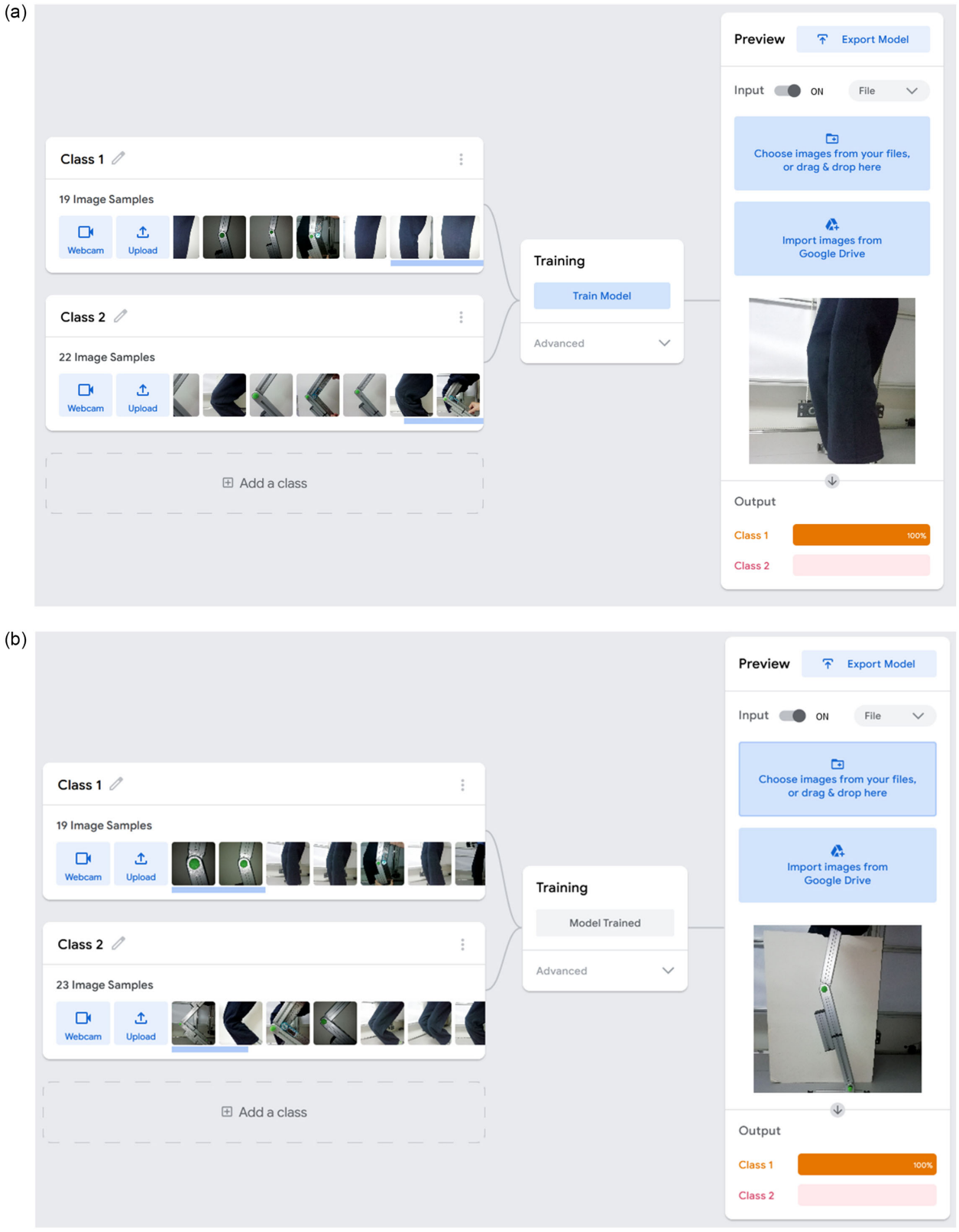
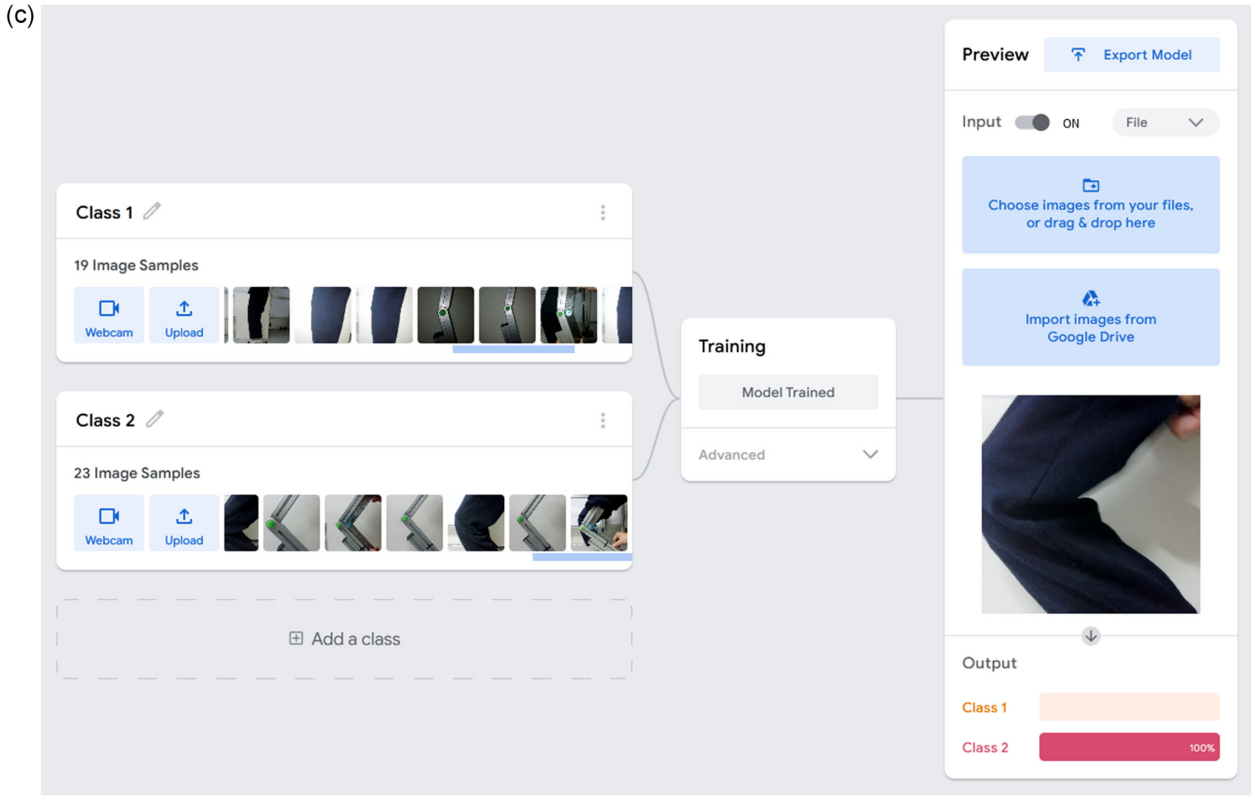


Figure 8
(Continue)



1) camera distance is not critical (Figure 10(a) and (b)), and a rigid fixture is not obligatory,
2) horizontal and vertical viewing angles (camera pan and tilt) are not critical (Figure 10(b)),
3) lens and perspective distortions are not critical, i.e., do not compromise accuracy significantly,
4) least possible training and testing dataset, minimizing cost, time, perplexity, and non-scientific complications,
5) lightest possible neural network for the AI tool, which may be particularly important if it is to be embedded in hardware or in a microcontroller (e.g., by using TensorFlow light).

The disadvantages, on the other hand, include the following:

1) Vulnerability to "jitter". This necessitates filtering, as explained in Section 4.3, Classification filtering,
2) Limitations in the extractable metrics. While two (2) stance detection is sufficient enough for timing an exercise, if it was required, for example, to calculate and graph velocities or accelerations against time, then knee angles or joints' coordinates with sufficient resolution would be required.

Concluding, the presented AI setup managed to identify artificial structures such as the biomimetic rig (Figure 8 preview), a goniometer (Figure 9), and worked exemplary accurately on detecting human subjects (males and females) at various observation angles and poses (Figures 9 and 10), even without training on human pictures. Instead, during the training stage of the AI, only photos taken from various poses of the biomimetic rig were used, with and without clothing on the mechanical limbs. Furthermore, human personal data are safeguarded during the "examination" procedure since camera tracking requires only partial images.

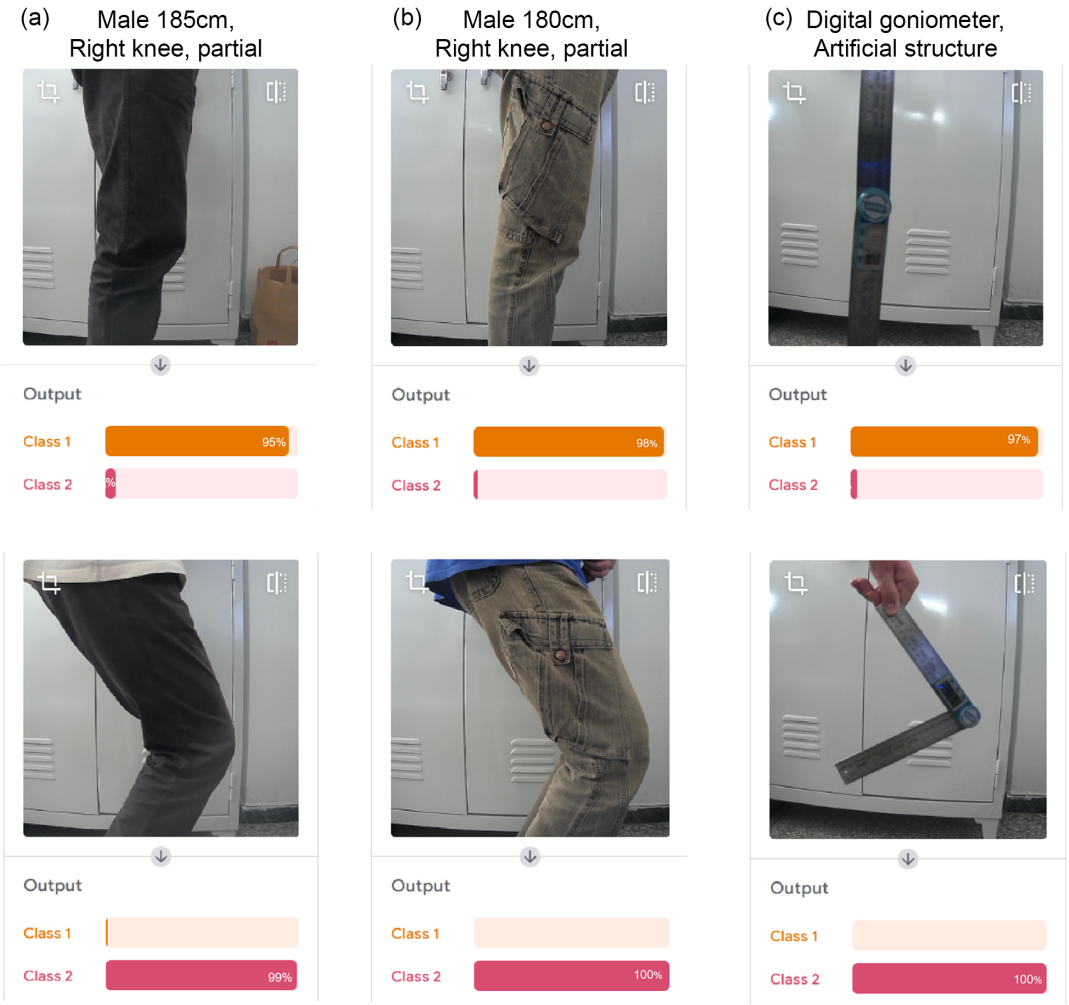
Regarding our trained-on-partial-frames AI, the recognition success rates are:

1) for the specified camera arrangement (sagittal view at knee level, approximately 50 cm away), the certainty levels reported by the AI model for standing and shallow squat stances exceed 95%. (See class percentages in Figures 9, 10 and 11).
2) for other camera arrangements (e.g., full-body view and non-trained AI), the certainty level is lower, as expected, for example, approximately 75%, as displayed in Figure 4.

However, the binary classification implementation will work while the certainty is above 51%. This translates to highly increased tolerance levels. Note that Figure 9(a) and (b) present the classification results for human subjects with clothes which differ in color, texture, and details. These clothes have not been used for training the AI. Figure 9(c) shows a goniometer which has not been used in training, either.

Regarding the "jitter" portion attributable to image capturing, and especially for partial frames, the AI model's efficiency may benefit from simple image improvement techniques (such as despeckling, converting to gray scale, contrast enhancing) and other more advanced (such as background removal, frame stabilization), according to each diagnostic test's specificities. These can be automatically applied before the classification procedure. The same or corresponding transformations might have to be applied to the training dataset, too. Conversely, AI training may benefit from "degrading" the dataset (e.g., by adding noise, lighting variations and artifacts, shadows), in an attempt to "ruggedize" the AI model against expected real-world conditions.

Figure 9
Classification results for various subjects, including a goniometer



In the following paragraphs, we introduce stand2squat AI_biorig. Basic information such as functions, calculations, and other details about the sample software are presented in Sections (4.1–4.3). The p5js environment offers ease of use and free universal access through any browser. That is why stand2squatAI_biorig has been developed with p5js, by building on the Teachable machine sample code (Figure 11) that accompanied the AI model.

4.1. Basic software functions

Figure 11 presents a screenshot of the application. The real-time camera input is shown on the top left corner. Underneath is the real-time classification result, the control buttons (icons), and the plots. The diagnostic test results appear on the right column. While the diagnostic tool runs, the embedded AI model classifies each camera frame, reporting the labels "Class 1" or "Class 2". These are transformed into "one" (standing) or "zero" (shallow squat) tags. While the test is being executed, that is, after the "start/stop diagnostic test" icon has been pressed, the 1 and 0 tags are stored in an array. Another array holds a millisecond timestamp for each corresponding frame. When the test has been completed, a filtering stage is required on the classification results to remove jitter. A detection stage follows, using a single pass "for" loop, to distinguish the transitions between stances as changes from 1 to 0, or from 0 to 1. When a transition is detected, a third array is filled with the duration of each stance. In this context, "jitter" is used as a general term, including numerous factors such as actual subject tremor, camera motion (especially when self-recording), camera autofocus, frame changes as the subject moves, lighting fluctuations, and AI false classifications.

Figure 12 presents an example of the stand-to-shallow squat test consisting of five repetitions. A simplified plot of the knee flexion angle appears in purple. The blue ribbon shows the AI tool classification results before the filtering stage. Notice the multiple erroneous transitions as thin bars. The blue columns show the estimated stances, correctly identified and timed after filtering. The gray columns show the exercise cycles, timed using the standing stances midpoints.

Two sequential stance transitions, that is, stand-to-squat plus squat-to-stand, which translate as $1 \rightarrow 0 \rightarrow 1$, constitute a full cycle. If the test begins and ends with the same stance, then:

$$\text{number of cycles} = \frac{\text{number of stances} - 1}{2} \quad \text{for odd number of stances} \tag{1}$$

Example: In Figure 11, the app has detected 7 terminal stances, which correspond to (7-1)/2 = 3 cycles.

Figure 10
Classification success with or without clothes, at various views and distances

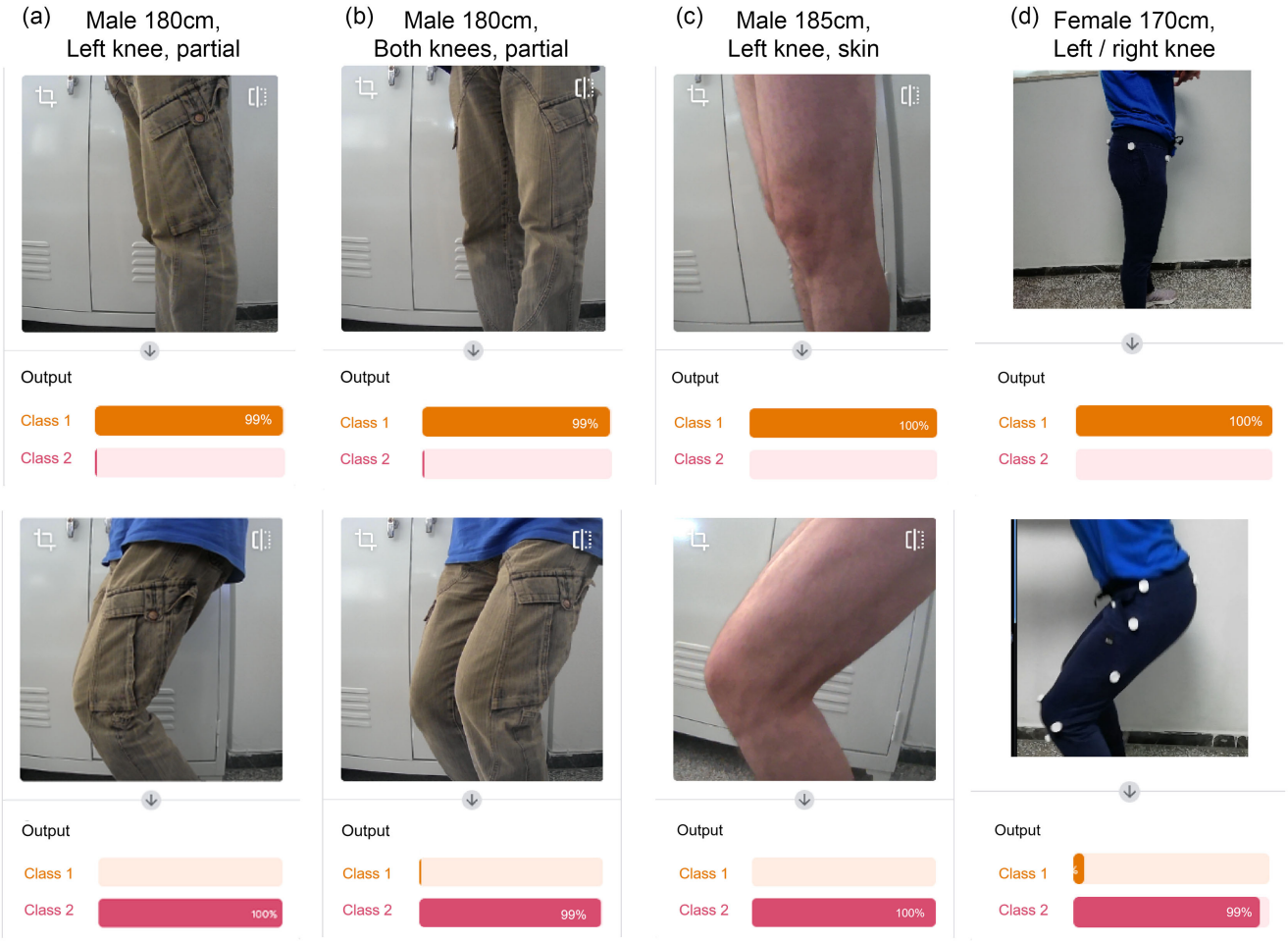
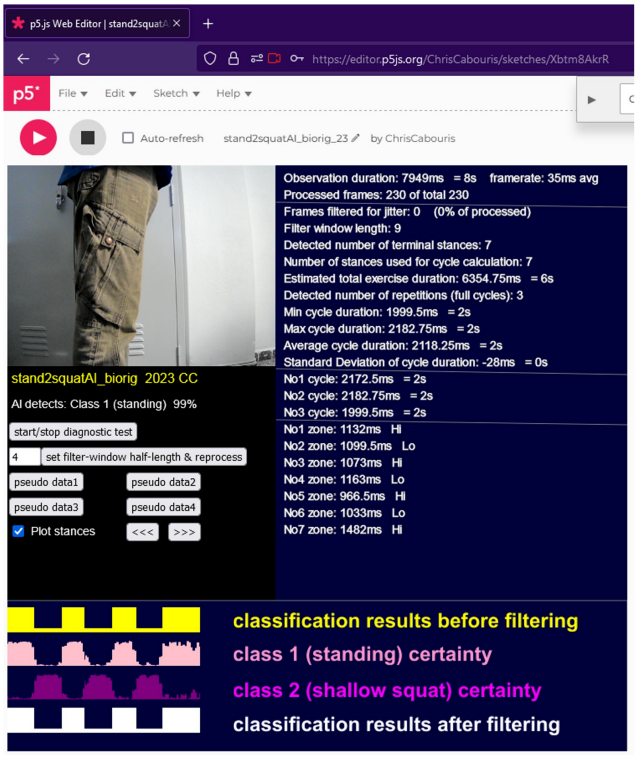


Figure 11
Screenshot: Stand2squatAI_biorig in p5js.org editor



However, if the test begins and ends with different stances, the last cycle is considered to be incomplete. The last stance is then discarded, and:

$$\text{number of cycles} = \frac{\text{number of stances} - 2}{2} \quad \text{for even number of stances} \tag{2}$$

Example: If the app detected 8 terminal stances, the eighth would be omitted, giving (8-2)/2 = 3 cycles as well.

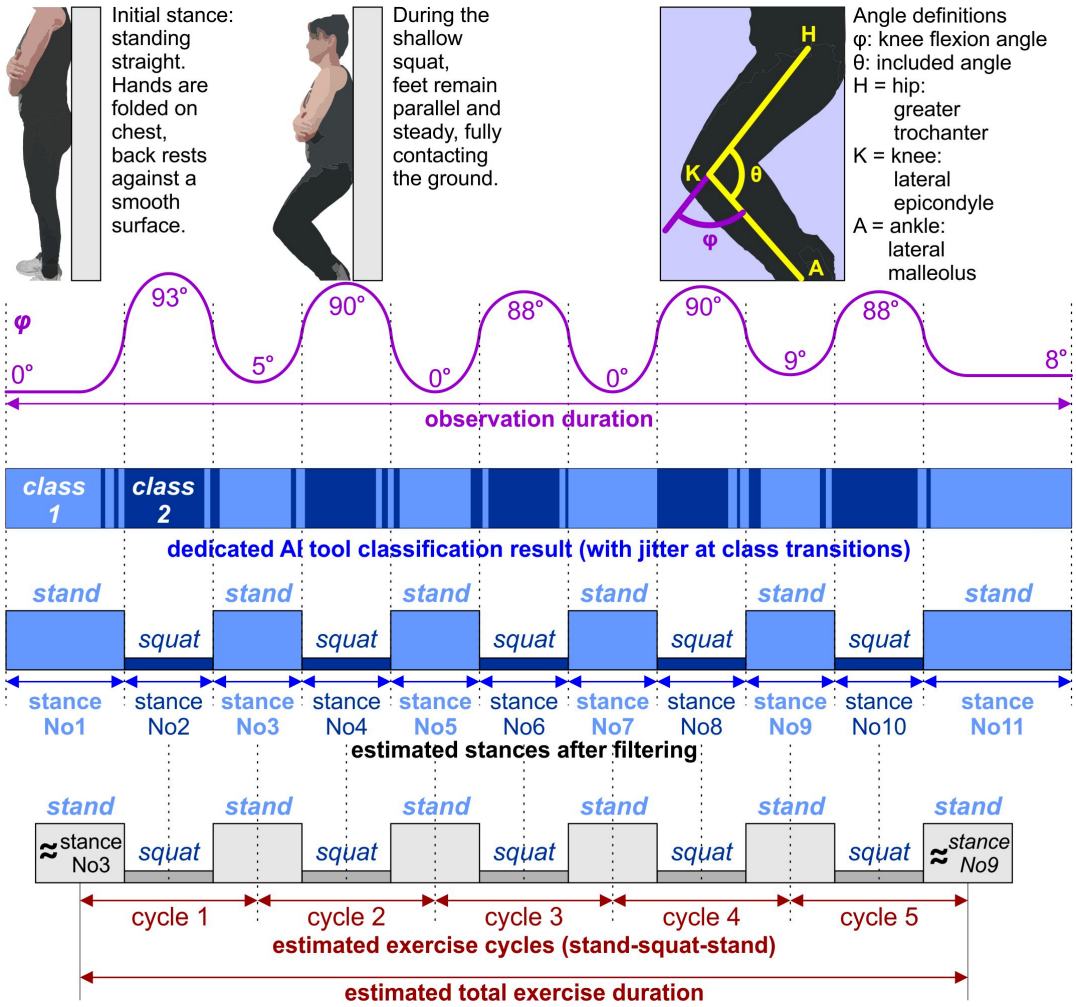
Three sequential stances (A, B, and C) are involved in each cycle (e.g., A: stand, B: squat, C: stand). Except for the first and last exercise cycles, each intermediate cycle is supposed to include the second half of stance A, the whole B, and the first half of C (Figure 12):

$$\text{intermediate cycle duration} = \frac{\text{stance A}}{2} + \text{stance B} + \frac{\text{stance C}}{2} \tag{3}$$

Example: In Figure 11, the app has detected 3 cycles (first, intermediate, and last). The intermediate cycle includes the second half of stance 3 (1073 ms), the whole stance 4 (1163 ms), and the first half of stance 5 (966.5 ms). Intermediate cycle duration would be $1073/2 + 1163 + 966/2 = 536.5 + 1163 + 483.25 = 2182.75$ ms, as displayed in Figure 11 No 2 cycle.

The first cycle usually seems longer because the first stance includes the lag from pressing the "start" button to the start of the exercise. This lag may be significant, particularly for self-recording.

Figure 12
Example of the five-repetition stand-to-shallow squat test



In other words, the absolute first stance duration may be misleading. Therefore, to improve the estimation for the first cycle, stance C is used twice instead of stance A as follows:

$$\text{first cycle duration} = \frac{\text{stance C}}{2} + \text{stance B} + \frac{\text{stance C}}{2} = \text{stance B} + \text{stance C} \tag{4}$$

Example: In Figure 11, the estimated first cycle includes the second half of stance 3 (1073 ms), the whole stance 2 (1099.5 ms), and the first half of stance 3 (1073 ms). Estimated first cycle duration would be 1073/2 + 1099.5 + 1073/2 = 1099.5 + 1073 = 2172.75 ms, as displayed in Figure 11 No 1 cycle. Without the correction, the first cycle duration would be 1132/2 + 1099.5 + 1073/2 = 566 + 1099.5 + 536.5 = 2202 ms.

Similarly, the last valid stance of the last cycle includes the lag after stopping the exercise until the stop button is pressed. Therefore:

$$\text{last cycle duration} = \frac{\text{stance A}}{2} + \text{stance B} + \frac{\text{stance A}}{2} = \text{stance A} + \text{stance B} \tag{5}$$

Example: In Figure 11, the estimated last cycle includes the second half of stance 5 (966.5 ms), the whole stance 6 (1033 ms), and the first half of stance 5 (966.5 ms). Estimated first cycle duration would be 966.5/2 + 1033 + 966.5/2 = 966.5 + 1033 = 1999.5 ms, as displayed in Figure 11 No 3 cycle. Without the correction, the last cycle duration would be 966.5/2 + 1033 + 1482/2 = 483.25 + 1033 + 741 = 2257.25 ms.

An array is filled with the duration of each cycle, and the metrics are then calculated and printed on the screen (Figure 11).

4.2. Metrics calculation

Stand2squatAI_biorig reports metrics as soon as the "start/stop" button is pressed for the second time, indicating the end of the observation stage of the diagnostic test.

$$total\ exercise\ duration = \sum cycle\ durations \tag{6}$$

$$Average\ cycle\ duration = avrg = \frac{total\ exercise\ duration}{number\ of\ cycles} \tag{7}$$

$$Mincycleduration = minimum(of\ allcycles'\ durations) \tag{8}$$

$$Maxcycleduration = maximum(of\ allcycles'\ durations) \tag{9}$$

$$cycle\ duration\ sample\ standard\ deviation = s = \sqrt{\frac{\sum (x-avrg)^2}{n-1}} \tag{10}$$

Where:

n = number of cycles,
avrg = average cycle duration,
x denotes each cycle duration

The total exercise duration is usually less than the "Observation duration," as shown in the first line of the results, because of discarding:

1) The last stance of the last incomplete cycle (if such exists).
2) Excess duration of the absolute first stance (compared to the third stance). The third stance is used to improve cycle duration estimation.
3) The excess duration of the last of the valid stances is similar to that in (b).

Conversely, if user presses the "start test" button after starting to move, the first stance may be smaller than the third, and the total reported exercise duration may appear longer than the "Observation duration". Likewise, if user presses the "stop test" button before finishing their motion!

Notes

1) Timing results are expressed in milliseconds (ms) and seconds (s).
2) If a subject performs more than five or fewer repetitions, the results would still be correct. Of course, the greater the number of repetitions, the better (regarding the average cycle duration precision).
3) If a subject performs the exercise starting with a squat, the results would again be correct. That is, cycle $0 \rightarrow 1 \rightarrow 0$ is also valid.
4) Frame rate affects precision; more in the case of young and athletic subjects who can move fast and less in the case of elders and patients. Usually, low-cost cameras and systems offer approximately 30 frames per second, which implies a nonconstant time lag of approximately 33 ms. The average frame rate and time lag are shown at the beginning of the results, below the "Observation duration" (Figure 11).

4.3. Classification filtering

A common problem in electronics is "zero-cross detection", which is required in order, for example, to synchronize the firing of power switches at the instance when the sinusoidal mains voltage (or current, in other cases) crosses the value of 0 V (or 0 A). The problems arise from the superimposed noise. The error (zero signal value minus superimposed noise) however small becomes significant when compared to nothing. Thus, a "naive" circuit may detect multiple zero crosses, before and after the "ideal" instance. The 2-class detection implementation presents similar "switching" issues. While minimum (standing) and maximum (shallow squat) knee flexion angles are safely detected, the middle values may fall into one class or the other. The AI tool will classify a subject's highest posture as "standing" even if the flexion angle is 10° instead of 0° because it differs significantly from a squat. Subtracting the certainties of the two classes (match scores) yields a large positive value. However, at the transition between stances, certainties become equal. Their difference is zero, causing vulnerability to noise and "switching" issues. These are expressed as "spikes" (very narrow stances, as if one jumped from squat-to-stand within milliseconds). These spikes would cause an overestimation of the number of stances and cycles and an underestimation of the average cycle duration. This would also ruin the reported minimum cycle duration, which, in this case, would be the duration of the narrowest spike, not corresponding to an actual cycle.

This issue can be addressed in several ways. Among them, by:

1) Increasing the number of classes. For example, a model with 3 classes: High (standing straight), Middle, Low (shallow squat). Then "spikes" will occur in the transitions H <-> M and M <-> L and can be easily grouped with the corresponding H or L class. H <-> L transitions are not expected, unless the frame rate is rather slow, or the AI tool inadequately trained.
2) Limiting the minimum acceptable stance duration. An unacceptably short duration can be grouped with the previous one or with the largest of its 2 neighbors, etc. In real-world situations, this easily applicable approach may require human intervention to select the most appropriate minimum limit, according to the equipment (camera frame rate, steady/handheld camera, lightning conditions, etc.,), the subject (fast/slow moving, with/without tremor) and other factors (clothes, background, etc.,).
3) Filtering. Practice shows that some sort of filtering is unavoidable. Simple filters will work satisfactorily, until noise levels become high.

stand2squatAI_biorig uses a centered, fixed-length, running-window, majority filter. That is, to determine the filtered value of a specific frame's stance (1/0), the algorithm sums the values of this frame with the values of w previous and w following frames. Thus:

$$window\ length = w + 1 + w = 2w + 1 \quad (11)$$

{window length is always odd}

Where:
w may be defined by the user

If the sum is > w (i.e., higher than 50%), then the result is 1; else, it is 0. The window is "mirrored" at the video edges: for the starting frames without w predecessors, the corresponding subsequent frames are used instead; likewise, for the ending frames. Small window lengths (e.g., 5) work very well for 0 or low "jitter". For significant jitter, the length must be increased (up to 99999), without approaching the total number of observed frames (it is automatically limited to 1/4 of the total number of frames). However, a large window is not a perfect solution, since it may, e.g., "trim" real cycles down to inexistence. User can select the window length, by defining w. When changing value, the results are extracted and presented again. This is feasible, because the arrays with the classifications of the last test are not altered until the next is observed. If the "plot zones" checkbox is ticked, each frame's classification is plotted at the bottom of the screen, before filtering (yellow ribbon) and after filtering (white ribbon). The 2-class certainties are plotted in pink and purple (Figure 11). So, the user may realize the effect of filtering and choose an appropriate value (Figures 13, 14 and 15). Figure 14 provides an example of successful filtering with w = 2 (window length = 2 * w + 1 = 5). On the contrary, Figure 15 demonstrates that an oversized filter (w = 4, window length = 9) results in missing a stance and an exercise cycle. The 4rth stance is lost because its duration is 4, i.e., small compared to the window length. In both Figures 14 and 15, IN denotes input in blue color. RW denotes the centered, fixed-length, running window, mirrored edges, and majority filter. OUT denotes the output stances in orange color.

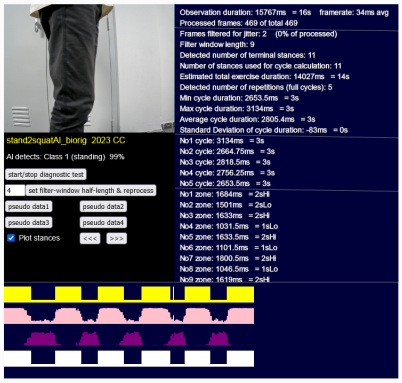
5. Comparative Study

After automating a diagnostic test through the proposed methodology, a fast accuracy check may be executed with the help of the biomimetic rig. In the example of stand2squatAI_biorig

Figure 13
Three examples of our single-camera diagnostic test performed on human subject

(a) Observed 11s, 5 cycles. 1.9s avg

(b) Observed 16s, 5 cycles, 2.8s avg



(c) Observed 34s, 9 cycles, 4.4s avg

Figure 14
Example of successful filtering with w = 2

IN:	1	1	1	1	1	0	1	0	0	0	1	1	0	1	1	0	0	0	1	0	1	1	1	1	1	0	0	0	0	1
RW:	5	5	5	4	4	3	2	1	2	2	2	3	4	3	2	2	2	1	2	3	4	4	5	4	3	2	1	1	1	1
OUT:	1	1	1	1	1	1	0	0	0	0	0	1	1	1	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0

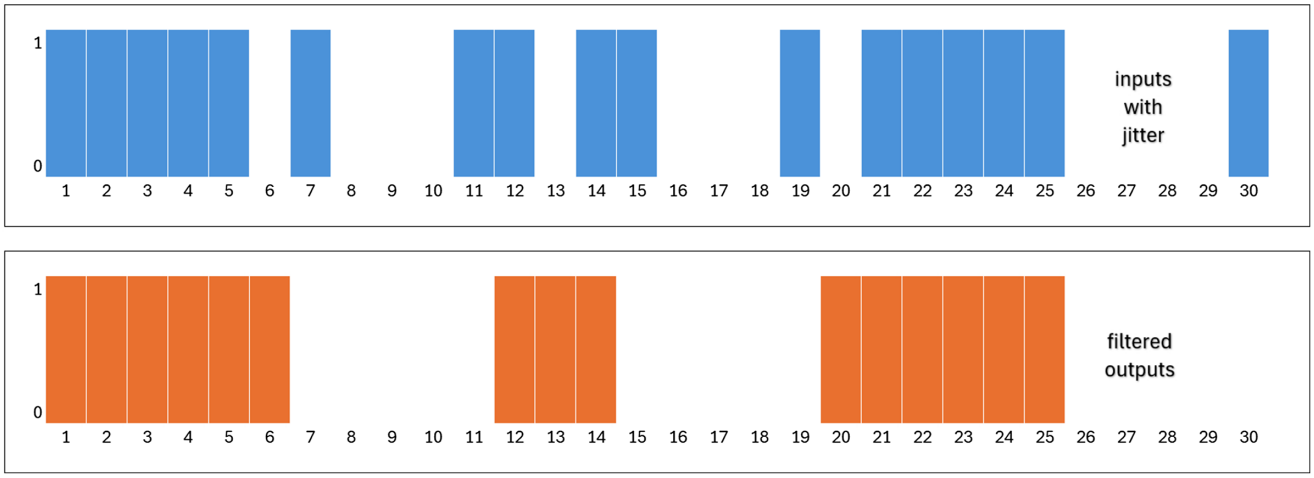


Figure 15
Too wide filter (w = 4, window length = 9) results to missing a stance and an exercise cycle

IN:	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0
RW:	7	7	6	5	4	3	3	3	3	4	5	5	5	5	5	5	5	5	5	5	6	6	6	6	5	4	3	2	1	0
OUT:	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0

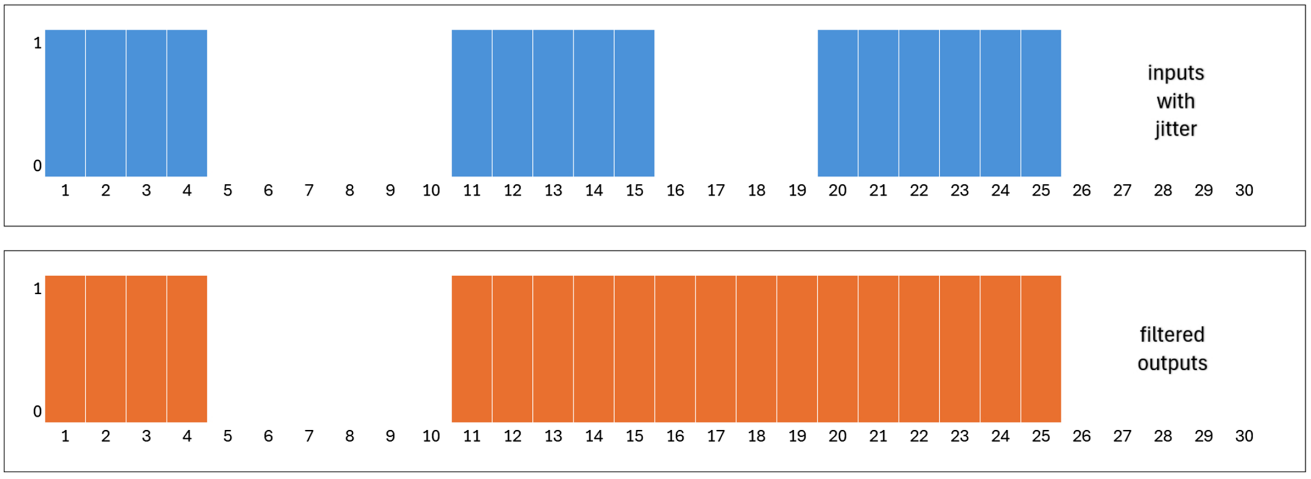


Figure 16
Snapshot after the rig executed test No 1

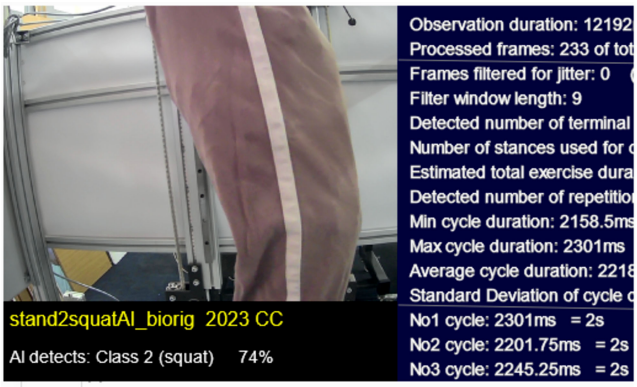
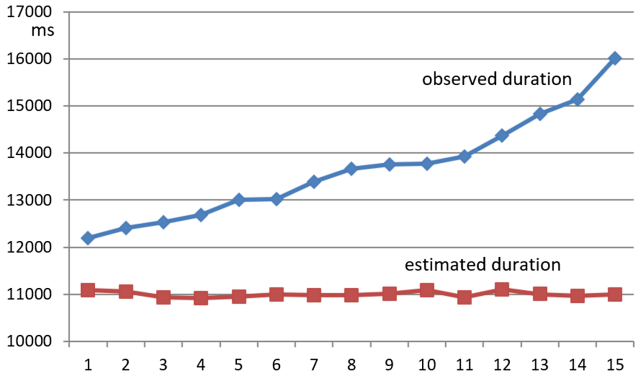


Figure 18
Observed vs estimated exercise durations



application, we programmed the rig to execute shallow squats, dressed it up with track pants (Figure 16) which had not been photographed for the training dataset. Then we purposely run the app on a typical Intel i7-4770 cpu @ 3.40 GHz pc, equipped with Nvidia GeForce GT620 graphics adapter and an off-the-shelf Foscam W21 FDH-1080p usb camera, to record 15 tests. Each test included 5 repetitions of standing to shallow squatting cycles. Some of the reported data are presented in Figure 17, with their order sorted according to observed exercise duration. Note that "zone Hi" corresponds to standing stance, while "zone Low" corresponds to shallow squat stance.

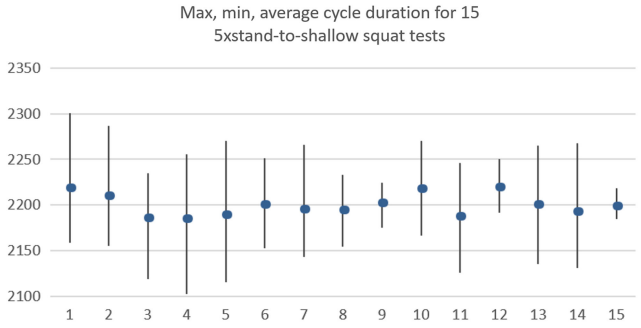
Figure 18 shows the observed vs estimated duration of each test. As previously noted, the observed duration depends upon the user's reaction time when starting and stopping the app. This uncontrollable jitter, which in this set of tests causes a variation of 16013–12192 = 3821 ms ~3.8 s, is rather high for a 11 s estimated overall time test. However, Equations (4) and (5) can handle such jitter. Thus, the estimated duration is quite consistent among tests as shown in Figure 18. Indeed, the remaining variation is 11098–10923 = 175 ms only, which represents 175/3821 = 4.6% of the observed duration variation.

As previously noted, the achieved frame rate is significant for the tests' accuracy. As shown in Figure 17 the average frame rate is not consistent among tests, with frame lag varying from 51 to 61 ms (19.6~16.4 fps). This depends on the system configuration. Figure 19 shows the effect of a slow and inconsistent frame rate. Here the estimated average varies from 2184.6 to 2219.6 ms = 35 ms or 35/2184.6 = 1.6%. Note that a high-performance system would easily reach 60 fps. Then the variation between the reported average, min, max cycle duration would be much smaller. Ideally, it should be average = min = max, because all exercises have been performed by the robotic rig executing the same motion pattern. Yet, the rig was dressed, and the cloth is expected to add variation [38] to the measurements.

Figure 17
Reports produced by stand2squatAI_biorig from 15 tests on the rig. Columns sorted by observation duration

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*
observation duration (ms)	12192	12407	12524	12678	13002	13026	13392	13661	13760	13778	13926	14379	14831	15140	16013
Frame lag (avg) ms	52	57	52	52	55	52	61	53	52	52	59	58	58	60	51
Processed frames	233	218	241	243	237	252	221	260	265	265	238	250	254	251	311
Frames filtered for jitter	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Filter window length	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
Detected number of terminal stances	11	11	11	11	11	11	11	11	11	11	12	11	11	11	11
Num. of stances used for cycle calculation	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
Estimated total exercise duration (ms)	11092	11048,8	10925,8	10923	10942,8	10999	10973,3	10972,5	11008,3	11086,3	10935,3	11098	11001,5	10963,8	10992,5
Detected number of reps (full cycles)	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Min cycle duration (ms)	2158,5	2154,75	2118,5	2102,5	2115,5	2152,75	2143,25	2154	2175	2166,5	2126	2191,75	2135,5	2131	2184,75
Max cycle duration (ms)	2301	2286,5	2235	2255,25	2270,25	2251	2265,75	2233	2224,25	2270	2246	2250,5	2264,75	2267,75	2218,25
Average cycle duration (ms)	2218,4	2209,75	2185,15	2184,6	2188,55	2199,8	2194,65	2194,5	2201,65	2217,25	2187,05	2219,6	2200,3	2192,75	2198,5
No1 cycle (ms)	2301	2286,5	2118,5	2102,5	2115,5	2217	2160	2154	2180,5	2270	2126	2233	2135,5	2238,5	2201,5
No2 cycle (ms)	2201,75	2193	2210,5	2175,5	2200,25	2152,75	2265,75	2233	2175	2254,25	2232	2191,75	2214,25	2131	2218,25
No3 cycle (ms)	2245,25	2154,5	2176,75	2255,25	2270,25	2194	2143,25	2176,75	2209	2166,5	2194,5	2211,25	2264,75	2267,75	2186,5
No4 cycle (ms)	2158,5	2211,5	2235	2175,75	2153,25	2184,25	2206,75	2176,25	2224,25	2193,5	2246,25	2211,5	2165	2147	2184,75
No5 cycle (ms)	2185,5	2203	2185	2214	2203,5	2251	2197,5	2232,5	2219,5	2202	2136,5	2250,5	2222	2179,5	2201,5
No1 zone (ms) Hi	1126	805,5	1573,5	1342,5	1301,5	1960	1807	1859	1476	1840,5	1940,5	2555	2912,5	3204	3395
No2 zone (ms) Low	1750,5	1776,5	1617,5	1637	1498	1667	1732	1569	1662	1653,5	1570	1658,5	1652	1703,5	1650,5
No3 zone (ms) Hi	550,5	510	501	465,5	617,5	550	428	585	518,5	616,5	556	574,5	483,5	535	551
No4 zone (ms) Low	1651	1645	1651	1667,5	1583	1652	1817	1614	1639	1618	1640,5	1567	1730	1617,5	1667,5
No5 zone (ms) Hi	551	586	618	550,5	617	451,5	469,5	653	553,5	656	627	675	485	492	550,5
No6 zone (ms) Low	1734,5	1636	1634	1737,5	1687	1734	1642,5	1599	1649	1513,5	1619	1578,5	1755	1743,5	1637
No7 zone (ms) Hi	470,5	451,5	467,5	485	549,5	468,5	532	502,5	566,5	650	524	590,5	534,5	556,5	548,5
No8 zone (ms) Low	1647,5	1728	1751	1702	1550,5	1649	1663	1663,5	1651,5	1568	1651,5	1568,5	1701,5	1642	1635,5
No9 zone (ms) Hi	551,5	515,5	500,5	462,5	656	602	555,5	523	579	601	665,5	695,5	392,5	453,5	550

Figure 19
Variation (error magnitude) in cycle duration estimates



Nevertheless, as seen in Figure 20, the stances jitter is quite uniform, and thus does not alter the appearance of the motion pattern. So, a diagnosis for the specific "patient" could be exported with a good confidence factor. The advantage of this quick accuracy assessment method is that it offers a benchmark for the expected results with the employed system configuration. Hence, it can be decided whether it is suitable for performing certain tests according to the required accuracy for diagnosing a specific ailment.

With the proposed methodology certain advantages arise:

1) Partial frame optical AI diagnostic tools can be developed quite fast and inexpensively.
2) Optical AI diagnostic tools can be used in humans, rigs, exoskeletons, prosthetics, orthotics.
3) Optical diagnostic tools can be run on recorded partial human motion data using a rig for emulation (with the human absent, without video available from the actual subject).
4) Affordable, precise, and consistent training and testing. For example, a rig facilitates multiple photos of various clothes, backgrounds, and perspectives at the same constant knee angle. Achieving this with a human subject would require motion capture systems with markers/stereotactic equipment/ special setups with ultrasonic, X-ray, or short-wave sensors/ tedious post-processing.
5) Human volunteers are not necessary for training the AI. Employing the biomimetic rig solves these ethical issues.
6) Similarly, volunteers are not required to test the AI and/or the finished diagnostic tool. In addition to ethics, the rig achieves reductions in costs and development time.
7) Processing speed. For example, running OpenPose on a noncutting-edge computer without GPUs requires several seconds per video frame. Hence, the exercise must be recorded first and the video processed later. On the other hand, a light, partial frame, dedicated AI model, like the presented, achieves "real-time" classification (certainly within milliseconds).
8) Diagnosis speed. The results appear on screen as soon as the test is completed (when the user presses the "stop test" icon).
9) Personal data privacy. No internet connection is necessary.
10) Ease of use. Users can see whether their stance is properly classified immediately prior to executing the test.
11) Simplicity. As illustrated, few-class models may perform adequately for certain diagnostic tests.

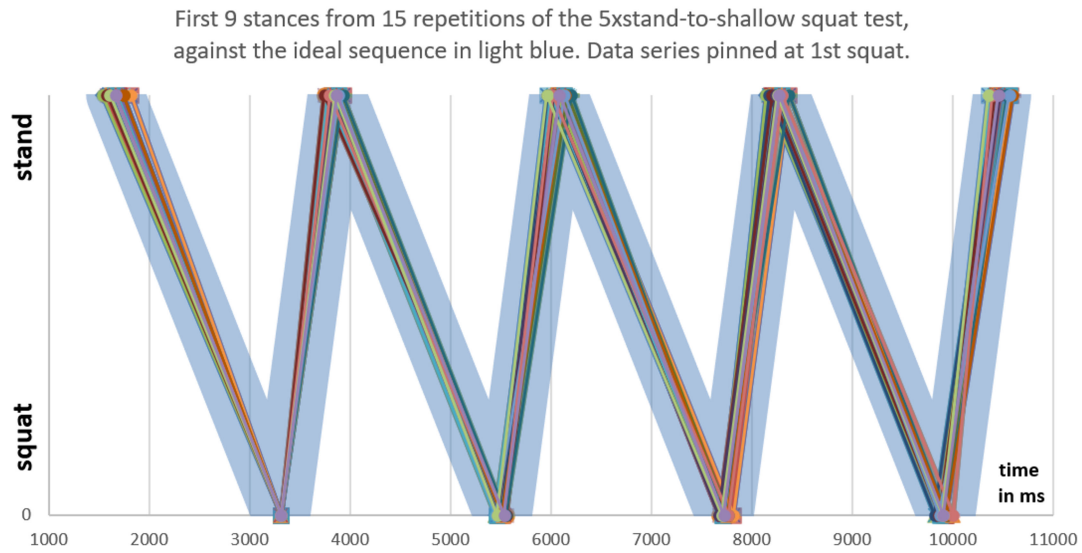
On the other hand, the disadvantages of few-class models include the following:

1) Vulnerability to "jitter". This necessitates filtering, as explained in Section 4.3 (classification filtering).
2) Limitations of extractable metrics. For example, two stances may be sufficient for timing an exercise but not for accurately calculating and graphing joint velocities or accelerations against time.

In the future, as processors become faster, the processing speed issue may lose importance, unless:

1) Fast cameras are necessary, according to the requirements of certain tests.
2) Multiple cameras are necessary (e.g., for joined frontal/sagittal plane studies, or for multiple close-up frames at various angles/regions or for multiple spectral imaging of the same area).

Figure 20
Apparent jitter in stance duration estimation after pinning the data series at first squat stance



3) Processor and computer shortages continue to plague the supply chain.

The proposed methodology and automated tools developed with its help can facilitate the assessment and fine-tuning of exoskeletons, prosthetics, and mechanical or medical aids. They can assist in the testing and evaluation of the combined human plus device system performance. Their significant advantage is the option to substitute the human subject with a biomimetic rig if original data from the subject are available from previous recordings. Future work will focus on these matters.

6. Conclusion

Automating gait and posture diagnostic tests with AI optical detection models and dedicated software on camera-equipped computers or smartphones is easy as illustrated with the proposed methodology. If applicable on the diagnostic test of interest, such automation may extend the reach of specialized healthcare services to wider population groups, as well as improve the services offered regarding orthotics, exoskeletons, prosthetics, and gait studies. The proposed methodology can assist in the automation while ensuring data privacy, thus maximizing the possibility of widespread public adoption and usage. The given example, stand2squatAI_biorig app, can be prepared quickly, efficiently, and cost-effectively, using free and open software tools.

Free, directly available to the general population, simple to use, AI tools for human diagnosis through open platforms, help in the Democratization of public health. The capability to perform self-diagnosis may play a basic role on early detection of health issues, let aside rehabilitation. It is of imperative importance to increase the number of people that have access to free or low-cost diagnostic solutions before contacting trained physicians which are of course indispensable in a proper healthcare organization.

An AI model's number of classes/complexity is to be determined both by the specific diagnostic test's requirements and by the developer's intentional choice between:

1) a simple and fast AI implementation, or
2) a complex and slow AI implementation.

For instance, while developing the example of stand2squatAI_biorig, the test's requirements allowed the choice of a simple and flexible AI implementation, which we favored to achieve a real-time response. Thus, we coupled it with a dedicated software to filter and generate useful metrics. Instead, we could have used a multi-class AI that detects and classifies the motion of the subject under study through all-in-between angles from zero knee flexion (-5^o) up to a shallow squat ($+95^o$) and combine it with a simpler post-processing software that would select the range of angles between the two stances to measure the cycle time.

To validate this diagnostic app, we performed a new set of experiments using the rig instead of volunteers. We created an algorithmic sit to squat cycle with fixed timing to be performed by the biomimetic rig for five consecutive times. We set the equipment to record the rig's motion 15 times, as if we had several volunteers. We included various initial delay responses, to check that our software algorithm is able to remove such jitter. With basic statistical analysis, we obtained an indication of the expected accuracy with the particular system configuration, which is quite promising for the implementation of such diagnostic apps.

We also discussed ethics issues that arise when patients' medical data are used without anonymization. Our proposed solution made use of an active biomimetic rig to accurately reproduce human poses and gait cycle. This was helpful at various stages of this research. Firstly, by using a rig instead of human subjects, we zero the personal data exposed to any system. Secondly, during the creation of the AI training dataset, the rig gave us the opportunity to set different viewing angles for the photos, for the exact same poses. It offered the chance to put various clothes on the mechanical limbs for the same stances. It repeated with ease any problematic stance during fine-tuning. Finally, it offered an as-detailed-as-needed dataset, excluding the human Intra and Inter-Subject variability [38] which up to now has been largely uncontrolled. The rig also discarded the burden of coping with a large number of human volunteers.

We would urge co-researchers to work on this methodology and automate gait and posture tests that require stance recognition or similar tests with neurological stimuli and reactions.

Acknowledgement

The authors would like to thank STAKAM S.A. for their continuous support.

Funding Support

This research was funded by STAKAM S.A. No specific grant was received from any funding agency in the public or not-for-profit sectors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The Teachable machine AI model used in this study is openly available in figshare at http://doi.org/10.6084/m9.figshare.27894876.

Author Contribution Statement

Christos Kampouris: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Philip Azariadis:** Supervision.

References

[1] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., . . . , & Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation*, 265–283.

[2] Nirupama, P., Lokesh, G., Veeranagaiah, C., & Bhanu Prakash, P. (2021). Automatic motorcyclist helmet rule violation detection using tensorflow and Keras in OpenCV. *International Journal of Advanced Research in Engineering and Technology*, *12*(4), 65–74.

[3] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., . . . , & Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. In *33rd Conference on Neural Information Processing Systems*, 1–12.

[4] Candel, A., & LeDell, E. (2016). *Deep learning with H2O*. USA: H2O.ai.

[5] Falch, L., & Lohan, K. S. (2024). Webcam-based gaze estimation for computer screen interaction. *Frontiers in Robotics and AI*, *11*, 1369566. https://doi.org/10.3389/frobt.2024.1369566

[6] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., ..., & Darrell, T. (2014). Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM International Conference on Multimedia*, 675–678. https://doi.org/10.1145/2647868.2654889

[7] Chen, T., Li, M., Li, Y., Lin, M., Wang, N., Wang, M., ..., & Zhang, Z. (2015). MXNet: A flexible and efficient machine learning library for heterogeneous distributed systems. *arXiv Preprint:1512.01274*.

[8] Ma, Y., Yu, D., Wu, T., & Wang, H. (2019). PaddlePaddle: An open-source deep learning platform from industrial practice. *Frontiers of Data & Computing*, *1*(1), 105–115. https://doi.org/10.11871/jfdc.issn.2096.742X.2019.01.011

[9] Palermo, M., Moccia, S., Migliorelli, L., Frontoni, E., & Santos, C. P. (2021). Real-time human pose estimation on a smart walker using convolutional neural networks. *Expert Systems with Applications*, *184*, 115498. https://doi.org/10.1016/j.eswa.2021.115498

[10] Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime multi-person 2D pose estimation using part affinity fields. In *2017 IEEE Conference on Computer Vision and Pattern Recognition*, 1302–1310. https://doi.org/10.1109/CVPR.2017.143

[11] Fang, H. S., Li, J., Tang, H., Xu, C., Zhu, H., Xiu, Y., ..., & Lu, C. (2023). AlphaPose: Whole-body regional multi-person pose estimation and tracking in real-time. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(6), 7157–7173. https://doi.org/10.1109/TPAMI.2022.3222784

[12] Hernández, Ó. G., Morell, V., Ramon, J. L., & Jara, C. A. (2021). Human pose detection for robotic-assisted and rehabilitation environments. *Applied Sciences*, *11*(9), 4183. https://doi.org/10.3390/app11094183

[13] Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., ..., & Grundmann, M. (2019). MediaPipe: A framework for building perception pipelines. *arXiv Preprint: 1906.08172*.

[14] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7464–7475. https://doi.org/10.1109/CVPR52729.2023.00721

[15] Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289. https://doi.org/10.1038/s41593-018-0209-y

[16] Chen, K. (2019). Sitting posture recognition based on openpose. *IOP Conference Series: Materials Science and Engineering*, *677*(3), 032057. https://doi.org/10.1088/1757-899X/677/3/032057

[17] Kidziński, Ł., Yang, B., Hicks, J. L., Rajagopal, A., Delp, S. L., & Schwartz, M. H. (2020). Deep neural networks enable quantitative movement analysis using single-camera videos. *Nature Communications*, *11*(1), 4054. https://doi.org/10.1038/s41467-020-17807-z

[18] Lin, P. C., Chen, Y. J., Chen, W. S., & Lee, Y. J. (2022). Automatic real-time occupational posture evaluation and select corresponding ergonomic assessments. *Scientific Reports*, *12*(1), 2139. https://doi.org/10.1038/s41598-022-05812-9

[19] Boswell, M. A., Kidziński, Ł., Hicks, J. L., Uhlrich, S. D., Falisse, A., & Delp, S. L. (2023). Smartphone videos of the sit-to-stand test predict osteoarthritis and health outcomes in a nationwide study. *npj Digital Medicine*, *6*(1), 32. https://doi.org/10.1038/s41746-023-00775-1

[20] Duncan, P. W., Weiner, D. K., Chandler, J., & Studenski, S. (1990). Functional reach: A new clinical measure of balance. *Journal of Gerontology*, *45*(6), M192–M197. https://doi.org/10.1093/geronj/45.6.M192

[21] Quinn, L., Khalil, H., Dawes, H., Fritz, N. E., Kegelmeyer, D., Kloos, A. D., ..., & Busse, M. (2013). Reliability and minimal detectable change of physical performance measures in individuals with pre-manifest and manifest Huntington disease. *Physical Therapy & Rehabilitation Journal*, *93*(7), 942–956. https://doi.org/10.2522/ptj.20130032

[22] Shearin, S. M., McCain, K. J., & Querry, R. (2020). Description of novel instrumented analysis of the four square step test with clinical application: A pilot study. *Gait & Posture*, *82*, 14–19. https://doi.org/10.1016/j.gaitpost.2020.08.119

[23] Beauchet, O., Fantino, B., Allali, G., Muir, S. W., Montero-Odasso, M., & Annweiler, C. (2011). Timed up and go test and risk of falls in older adults: A systematic review. *The Journal of Nutrition, Health & Aging*, *15*(10), 933–938. https://doi.org/10.1007/s12603-011-0062-0

[24] Zanevskyy, I., & Zanevska, L. (2017). Evaluation in the sit-and-reach flexibility test. *Journal of Testing and Evaluation*, *45*(2), 346–355. https://doi.org/10.1520/JTE20150298

[25] Neto, T., Jacobsohn, L., Carita, A. I., & Oliveira, R. (2015). Reliability of the active knee extension test and the straight leg raise test. *Journal of Sport Rehabilitation*, *24*(4), 1–12. https://doi.org/10.1123/jsr.2014-0220

[26] Avogaro, A., Cunico, F., Rosenhahn, B., & Setti, F. (2023). Markerless human pose estimation for biomedical applications: A survey. *Frontiers in Computer Science*, *5*, 1153160. https://doi.org/10.3389/fcomp.2023.1153160

[27] Needham, L., Evans, M., Cosker, D. P., Wade, L., McGuigan, P. M., Bilzon, J. L., & Colyer, S. L. (2021). The accuracy of several pose estimation methods for 3D joint centre localisation. *Scientific Reports*, *11*(1), 20673. https://doi.org/10.1038/s41598-021-00212-x

[28] Mehdizadeh, S., Nabavi, H., Sabo, A., Arora, T., Iaboni, A., & Taati, B. (2021). Concurrent validity of human pose tracking in video for measuring gait parameters in older adults: A preliminary analysis with multiple trackers, viewing angles, and walking directions. *Journal of NeuroEngineering and Rehabilitation*, *18*, 139. https://doi.org/10.1186/s12984-021-00933-0

[29] Acquisti, A. (2010). The economics of personal data and the economics of privacy. *Economics*, *11*, 24.

[30] Di Martino, M., Robyns, P., Weyts, W., Quax, P., Lamotte, W., & Andries, K. (2019). Personal information leakage by abusing the GDPR 'right of access'. In *Fifteenth Symposium on Usable Privacy and Security*, 371–385.

[31] Lee, W. W., Zankl, W., & Chang, H. (2016). An ethical approach to data privacy protection. *ISACA Journal*, *6*, 1–9.

[32] Rockwell, C., & Fouhey, D. F. (2020). Full-body awareness from partial observations. In *European Conference on Computer Vision*, 522–539. https://doi.org/10.1007/978-3-030-58520-4_31

[33] Tous, R., Nin, J., & Igual, L. (2023). Human pose completion in partial body camera shots. *Journal of Experimental & Theoretical Artificial Intelligence*, 1–11. https://doi.org/10.1080/0952813X.2023.2241575

[34] Fanton, M., Harari, Y., Giffhorn, M., Lynott, A., Alshan, E., Mendley, J., ..., & Jayaraman, A. (2022). Validation of Amazon Halo movement: A smartphone camera-based assessment of movement health. *npj Digital Medicine*, *5*(1), 134. https://doi.org/10.1038/s41746-022-00684-9

[35] Carney, M., Webster, B., Alvarado, I., Phillips, K., Howell, N., Griffith, J., ..., & Chen, A. (2020). Teachable machine: Approachable web-based tool for exploring machine learning classification. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–8. https://doi.org/10.1145/3334480.3382839

[36] Prasad, P. Y., Prasad, D., Malleswari, D. N., Shetty, M. N., & Gupta, N. (2022). Implementation of machine learning based Google teachable machine in early childhood education. *International Journal of Early Childhood Special Education*, *14*(3), 4132–4138. http://doi.org/10.9756/INT-JECSE/V14I1.221001

[37] Soon, W., & Cox, G. (2020). *Aesthetic programming: A handbook of software studies*. UK: Open Humanities Press.

[38] Kampouris, C., Azariadis, P., & Moulianitis, V. (2021). A methodology for assessing the impact of error components in gait analysis using closed-loop testing on a biomimetic rig. In *Novelties in Intelligent Digital Systems: Proceedings of the 1st International Conference*, 21–30. https://doi.org/10.3233/FAIA210071

[39] Kampouris, C., Azariadis, P., & Moulianitis, V. (2023). Introducing a biomimetic rig for simulating human gait cycles and its potential applications. In *Novel & Intelligent Digital Systems: Proceedings of the 3rd International Conference*, *1*, 152–163. https://doi.org/10.1007/978-3-031-44097-7_16

[40] Fukuchi, C. A., Fukuchi, R. K., & Duarte, M. (2018). A public dataset of overground and treadmill walking kinematics and kinetics in healthy individuals. *PeerJ*, *6*, e4640. https://doi.org/10.7717/peerj.4640

[41] Bovi, G., Rabuffetti, M., Mazzoleni, P., & Ferrarin, M. (2011). A multiple-task gait analysis approach: Kinematic, kinetic and EMG reference data for healthy young and adult subjects. *Gait & Posture*, *33*(1), 6–13. https://doi.org/10.1016/j.gaitpost.2010.08.009

[42] Mandery, C., Terlemez, Ö., Do, M., Vahrenkamp, N., & Asfour, T. (2015). The KIT whole-body human motion database. In *2015 International Conference on Advanced Robotics*, 329–336. https://doi.org/10.1109/ICAR.2015.7251476

[43] Piñero-Fuentes, E., Canas-Moreno, S., Rios-Navarro, A., Domínguez-Morales, M., Sevillano, J. L., & Linares-Barranco, A. (2021). A deep-learning based posture detection system for preventing telework-related musculoskeletal disorders. *Sensors*, *21*(15), 5236. https://doi.org/10.3390/s21155236

[44] Rocha, A. P., Choupina, H. M. P., Vilas-Boas, M. D. C., Fernandes, J. M., & Cunha, J. P. S. (2018). System for automatic gait analysis based on a single RGB-D camera. *PLoS One*, *13*(8), e0201728. https://doi.org/10.1371/journal.pone.0201728

How to Cite: Kampouris, C., & Azariadis, P. (2025). A Methodology for Rapid Deployment of Diagnostic Applications for Human Gait & Posture Analysis and Exoskeleton Configurations, Combining AI with a Biomimetic Rig. *Artificial Intelligence and Applications*, *3*(1), 10–30. https://doi.org/10.47852/bonviewAIA42023630