Abstract

We present and share\(^1\) a foundational dataset of multi-angle video recordings of scripted athletic movements to enable the development of computer vision research applications that evaluate and identify lower-body injury risk. The focus of the dataset is female athletes, who are at a substantially increased risk of anterior cruciate ligament (ACL) injury and are therefore a top priority for sports science. In our study, varsity and club sport athletes perform two assessment movements (the countermovement jump and the drop jump). These jump tasks are used ubiquitously in sports medicine research to characterize athleticism and to identify risk factors that indicate ACL injury propensity. The novelty of the dataset centers on (i) the type of movement data (purposeful, evaluative movements that need to be tracked with a high degree of precision), (ii) our generalized collection method that can be replicated with ease by non-experts, and (iii) the amount of data collected (we collected data from 55 division one (D1) female athletes performing 3 – 5 iterations of each jumps, for a total of 480 jumps). Data from each camera was manually aligned and a fully automated pipeline was built to extract knee information from athletes. Ideally, any athlete or researcher will be able to easily replicate our setup and assemble a compatible and complementary dataset to propel the development and assessment of injury propensity models.

1. Introduction

Athletic injuries are often thought of as explainable events; for example, “he’s an older athlete, he shouldn’t have been playing at that age”, or “she didn’t know her limits, she shouldn’t have pushed herself like that”. In actuality,
injury is a common occurrence across a broad spectrum of athletes, and the fine line between healthy progression and overexertion in training is often blurred.

Assessing the true cause of injury is difficult due to the fact that injuries are also frequently attributed to simple bad luck. Often a single bad landing or incident of player contact will be cited as the root of the injury; however, studies have shown that certain individuals are indeed more susceptible to non-contact anterior cruciate ligament (ACL) injuries than others. In college sports, women on basketball and soccer teams experience ACL injuries almost three times as often as men do, despite having comparable athletic exposure [1, 19]. Since college athletes are generally very experienced in their sport and have trained themselves to reach a high level of athleticism, it is surprising that one demographic is affected so much more dramatically than others. Furthermore, ACL injuries are especially debilitating injuries, and often involve significant recovery time. Once an athlete first injures an ACL, the risk of subsequent injury also increases [9, 5]. In extreme cases, the ACL may be so damaged that it is replaced by a tendon from another part of the body. While these tendons can be stronger, they are not particularly suited to function as ACLs and often limit athletes for the rest of their lives. ACL injuries can quickly end an athlete’s career, and many athletes are considered lucky if they only lose a season due to an ACL injury. However, knowing the effects and likelihood of ACL injuries is only useful information if it aids in their avoidance.

Fortunately, ACL injury prevention risk is is a well studied area [19, 5], and many programs have been proposed and validated to reduce ACL injury risk before it occurs. This is promising work; however, not every athlete is at risk for ACL injury, and such programs cut into the already limited time of athletes, especially those who are also students. There are several considerations for how to best identify injury risk in athletics, such as athlete tracking during live performance. In the case of ACL injuries, risk is best identified by professionals using simple movements such as counter-movement jumps and drop jumps. These evaluations are simple, short, and informative. Using knowledge of these movements, trainers are able to identify which athletes are at risk, as well as the factors contributing to such risk. This frees non-at-risk athletes to focus on their own needs while allowing injury-reduction programs to direct their focus exclusively and effectively to those at risk.

Of course, even with simple tests, identifying at-risk athletes is difficult in practice. Skilled athletic trainers need to allot time in order to monitor athletes and assess risk based on identified risk studies [5], and they often have many tasks and large teams to oversee. Additionally, the amount of institutional support for athletic training can vary widely within the collegiate athletics space. Outside of college and professional sports, many active people do not have access to trainers at all.

We believe that computer vision has advanced far enough that these risk tests can be conducted by trained computer vision algorithms with consumer grade equipment. As a first step toward this goal, we have collected video data from athletes performing a drop jump, which is often used to evaluate ACL injury risk [5]. Our data was collected by placing three consumer grade cameras in a semicircle around an athlete to allow researchers to perform 3D or 4D reconstructions. We intentionally left the camera placement inexact, requiring built models to generalize to variations in setups, which increase the likelihood of a model’s applicability to unseen data. Ideally, any non-expert could create an approximation of our setup and data collection technique using their own cameras, achieving the ability to assess themselves for ACL injury risk. To our knowledge, this is the first dataset of its kind.

By focusing our initial collection on athletes within a high-risk demographic at the start of our research, we target those who are more likely to show evidence of risk. In this way, we can more quickly obtain a wide amount of data pertaining directly to those who are susceptible to ACL injuries. This also allows our research to have a more substantial impact by aiding the group of athletes that is most affected by these debilitating injuries.

In the future, we hope to design, implement, and deploy an evaluative system that can identify ACL risk at or beyond the level of a trained athletic trainer. Ideally, this work will motivate further research into the use of computer vision as an assessment tool in athletics. As an initial start, we have developed a pipeline to process videos from collection to knee angle estimation using Carnegie Mellon University’s OpenPose software [20], which is free and accessible to anyone. This groundwork will allow us to begin modeling the specific movements needed to assess risk and will lead to a multitude of research into injury identification and individualized prevention techniques accessible to any and all athletes.

2. Related Work in Sports Medicine and Computer Vision

In this section we first provide a brief introduction to evaluative movements and their uses for identify injury risk identification and fitness assessments in the field of sports medicine. Next, we discuss cases of computer-aided risk evaluation. Finally, we discuss the current state of human body tracking in the computer vision community.

2.1. Sports Medicine

In the field of sports medicine there are two evaluative movements that are often used to assess athletes and identify injury risk: the countermovement jump and the drop
jump. In this section we will provide an introduction to each movement and explain how the movement has been used in the field of sports medicine.

2.1.1 The Countermovement Jump

The countermovement jump is a simple task that provides reliable insight into athlete fitness [7]. The jump consists of three steps, all performed without pausing. The athlete begins the jump with their arms held out directly in front of them, perpendicular to their body. The athlete then squats so that the knees are at 90 degree angles, simultaneously arcing their arms downward, past their legs, and behind their back, until their arms are again parallel to the floor. They then immediately reverse the arc with their arms, swinging them forward to propel the jump upward, fully extending their legs and using the combined momentum to jump into the air. Finally, the athlete lands as close to their jumping off point as possible; they allow their knees to bend and to absorb the impact of the jump. The full procedure is shown from left to right in 2.

The countermovement jump has a long history of use as an evaluative tool for athletes [14, 3, 4]. It has been used to measure an athlete’s physiological adaptation to their power training, or to assess the level of fatigue [4]. These studies include statistical analyses that focus on estimation of power, force, velocity, and displacement variables. Recently, sports medicine research has expanded its approach to include temporal jump information, considering the means of variables as well as maxima from an assessment [3, 21]. Ideally, these same variables can be tracked automatically using video data and body tracking models, leading to individualized insight into the effects of an athlete’s training program.

2.1.2 Drop Jump

The drop jump begins from an elevated platform. The athlete steps forward off the edge of the platform; it is important to ensure that this stepping motion follows the horizontal plane of the initial platform, rather than a typical downward step. As the athlete drops to the ground, they should land with their knees slightly bent. In a continuous motion, the knees should continue bending with the arms naturally extending behind the athlete as their torso leans forward. Immediately following this position, the athlete should jump, swinging their arms forward and using the combined momentum of the drop and their crouching position in order to propel themselves upward and extend their legs. As in the countermovement jump, the athlete should make an effort to land with their feet in the same position they occupied in the crouching position.

As with the countermovement jump, the drop jump has a strong history as an evaluative tool, and is often used in conjunction with the countermovement jump for a fuller picture of athleticism [21, 11, 6]. However, the jump landing task also has a long history of use for ACL evaluations in particular. Hewett et al. [5] identified joint angles and joint loads of 205 female athletes performing a drop jump and monitored which athletes sustained ACL injuries. It was found that knee angle, ground reaction force, and speed were all predictive of ACL injury propensity. This work, which featured a large number of participants, is a strong motivator for our collected data, as we were able to ensure that our video recordings adequately captured the movement areas relevant to analysis. Additionally, future work will consider the risk variables identified in this study when considering which model variables need to be fine-tune models and how to assign an at-risk rating to a participant.

2.2. Computer-driven risk assessment and human judges with the LESS protocol

The most similar data collection effort to our research is the LESS collection [13], which used two cameras positioned at 90 degree angles (one in the front, one to the right) to record 2691 participants performing a drop jump. In conjunction, electromagnetic tracking was used to obtain high-precision, gold standard tracking data, such as knee angles and hip angles, during the jump. Human judges then used the video data to annotate each participant’s jump in accordance with the LESS system, a pre-defined list of common jump errors. The human-judgment-based LESS system was found to be as reliable as the gold standard electromagnetic tracking. This study showed that trained human judges are able to identify and assess jump risk themselves, without the need for gold standard ground truth. While these results are a great boon for evaluations, since electromagnetic tracking is not widely available, the LESS system still requires experts to be trained and available to perform evaluations. We believe that a feasible, more inclusive alternative is to train computer vision models to accurately work on widely available, consumer grade video recordings.

While our current focus is on the collection of data from D1 athletes, we believe that models need to be evaluated on a broad representation of the population. Unfortunately, the LESS study did not result in a public dataset of jump videos which could be used to train and evaluate computer vision models.

2.2.1 Computer vision for body tracking, 3D reconstruction, and 4D reconstruction

In recent years, computer vision has seen well-known advances across a large array of disciplines due to deep learning, and human body tracking is no exception. Deep networks have been trained that track human movement performing unique actions in 3D [15] and across dynamic...
Figure 2. The countermovement jump. After positioning themselves standing upright with the arms outstretched in front of them, parallel to the ground, (left) the athlete begins the countermovement jump by arcing their arms downward while simultaneously bending their knees (center) until their arms are behind their body, again parallel to the ground, and their knees are bent at 90 degree angles. From this position, the (right) the athlete reverses the arc, throwing their arms into the air above them to maximize their momentum while extending their legs and performing the jump.

Figure 3. The drop jump (pictures described from left to right): 1) an athlete positions themselves on a box in preparation for the drop jump 2) an athlete begins to steps forward in order to allow themselves to drop to the ground 3) the athlete lands and immediately begins to jump straight up 4) the athlete performs the jump.

scenes [12]. Deep learning methods also allow for reconstruction of 4D shapes through the aid of temporal redundancy and multiview perspectives [10, 17], and identification of human actions in a multitude of contexts [23]. Considering the wealth of advances made, in which people are tracked whilst performing complex movements [8, 18], models may be nearing accuracies that would allow automatic evaluation of injury risk, guided by the findings in sports medicine research. If said models can be utilized with consumer grade recording equipment, they could provide health guidance and prevent injuries for athletes without access to trained professionals or rare and expensive equipment.

3. Data Set Collection Method

All data was collected at standard athletic facilities of a division 1 (D1) university in the Midwestern USA. Participants were limited to females who participated on an Olympic or varsity athletic team. Teams included soccer, basketball, rugby, rowing, swimming, cross country, fencing, and track. A human subjects data collection protocol was developed and approved by the University’s Human Subjects IRB and all data was obtained under the protocol.

Data collection was held in locker rooms, a large gym area, and a facilities entryway. For each collection, a 24 inch box was placed against a wall and two force plates were placed in front of the box for the drop jump. A black
backdrop was placed along the wall directly behind the participants. Three Sony a6300 cameras were set up in a semi-circle around the force plate, positioned approximately 45 degrees from the camera at the center. Each camera’s settings were matched among all cameras used in the recordings.

Data was collected as follows: participants were first instructed to sign a human subjects data collection consent form. They then were asked to go to a dressing room to put on a pair of textured tights, allowing participants’ legs to be more easily identified and tracked by texture, and providing a clearer representation of the knees, ankles, and hips. Participants then stood on two force plates which were calibrated to their weight. As they stood on the force plate, three cameras on separate stands were focused on the participant’s legs. Cameras were placed in an approximate semi-circle around participants and were routinely adjusted between participants to ensure the participant was captured with high detail. After the cameras were focused on the participant legs, their focus was locked and a checkered stereo calibration screen was held in front of the participant and shown to all three cameras. After calibration, participants were instructed on how to perform the countermovement jump. They then performed between three and five countermovement jumps. After these tasks were complete, participants were asked to step off of, and then return to, the force plate so the force plates could be re-calibrated. The participants then proceeded to step onto a box placed behind the force plates, where they were instructed on how to perform the drop jump. They performed three to five drop jump movements as their last task. Data collection took approximately six minutes per participant.

The majority of our data was obtained by recording one participant per video, recording at 120 frames per second. We also collected three data collection sessions with 15 participants captured at 30 frames per second. It was not always possible to collect accurate force plate data due to equipment limitations.

4. Data processing and preliminary results

We constructed a simple pipeline to process the data and ensure that it was collected correctly, as well as to obtain preliminary results as a proof of concept. We first describe our data processing pipeline and then describe our preliminary findings.

4.1. Data processing pipeline

In this pipeline, we process the data from post-collection to keypoint identification. The greatest barrier to completely automatic processing was the lack of time-synchronized capture. Ideally, future work will employ synchronized cameras, or perhaps camera information from only one angle will be sufficient to diagnose injury risk.

4.1.1 Temporal alignment

Data from each camera was temporally aligned using Final Cut Pro, although it is important to note that this step should be feasible across a range of editing software. We aligned each camera and cut any footage that was not present on
 Nonetheless, other uses of the evaluative movements re-

Table 2. We evaluated OpenPose’s ability to accurately identify lower extremity joints (hips, knees, ankles) at keyframes of evaluative jumps. Keyframes were defined as the lowest point of the athlete during landing, which is most indicative of risk of ACL injury [5]. For each keyframe, an evaluator reported their confidence in OpenPose’s joint estimation on a scale from one to five, with one being certain OpenPose was incorrect, three being uncertain, and five being certain in the validity of its estimation. We evaluated a subset of 32 landings from each camera over three participants. Our analysis showed OpenPose appeared less stable on the left camera, which we attribute to a need to set the camera up closer to the participants due to the collection space constraints.

all three cameras. Alignment points were chosen based on uniqueness and ease of viewing from all angles. Typically, videos were aligned on the moment when the athlete’s feet left the ground on the first countermovement jump.

4.2. Initial results

We pursued two avenues to obtain initial results for assessment. We first performed stereo calibration using OpenCV [2] and then used this information to perform key-point matching between videos.

Second, we used OpenPose [20] 2D skeletal matching matching to obtain initial results. We used OpenPose’s default parameters and evaluated how accurately participants’ joints could be tracked during both the countermovement and drop jumps.

OpenPose performed impressively for an off-the-shelf algorithm. Some weaknesses of OpenPose arose when participants jumped, and parts of their body were obfuscated from the frame. However, OpenPose was able to re-identify the participant immediately as they returned to frame, allowing key point estimates to be made at landing time (an essential moment for ACL injury risk identification [5]). Nonetheless, other uses of the evaluative movements re-

quire the participant to be tracked throughout the full movement [14, 22, 3]. We encountered two key problems when employing OpenPose that made it difficult to work with the data. First, by default, OpenPose is equipped to handle a large number of onscreen participants. Often, non-participants were identified and tracked by OpenPose despite being far in the distance, making key-point results more difficult to work with. In other cases, workout equipment was detected as human and skeletonized. Second, we noticed that OpenPose estimation varied frame-to-frame, which causes us to question if it is stable enough to identify needed joints and angles with enough precision to properly evaluate athletes.

We further investigated OpenPose’s tracking by performing a quantitative analysis of participants’ landings [5]. First, for each jump a participant performed, we identified one key frame as the lowest point in the landing. We then processed a subset of participants through OpenPose. We manually examined OpenPose’s joint estimates at each key frame from each camera’s perspective. Each estimation was rated between one and five, with one being confident that OpenPose’s joint estimates were incorrect, three being uncertain, and five being completely confident that OpenPose’s joint estimation was correct. Overall, we found OpenPose to have promising results. Likely, the joint estimation models should be transfer learned on this dataset to improve accuracy.

4.3. Limitations

We believe this dataset is the first of its kind and therefore enables new research on an important athletic health problem. However, there are several limitations to this initial work that we hope to overcome through the collection of additional data. First, we collected data only from female athletes, because of their higher general risk of ACL injury than male athletes. However, male athletes also suffer ACL injuries. In the future, we believe it will be prudent to collect additional data from male athletes. Secondly, our data focused solely on D1 athletes from a single school. In the future, we hope to collect more participants from outside of the pool of athletic performers. However, there are several immediate benefits to beginning our collection with this subset of the population: athletes are more likely to be in good shape, be aware of a physical evaluation such as ours, and be trained on how to proceed. Additionally, they were given the option to receive an evaluation for ACL injury risk during data collection, which allowed us to increase our amount of data while benefiting the athletes. Thirdly, while there is an abundance of D1 athletes on the campus where our data was collected, many other places do not have access to such a vast amount of top athletic performers. Because of this, the collection of hard-to-obtain data was a priority which we plan to supplement, and hope will be
supplemented by others, with future data collections. Additionally, we collected data both from athletes who were well versed in the jumps they performed and athletes who had never performed the jumps before. We believe this will provide an acceptable variety in data to ensure that we are not collecting data exclusively from experts. Nonetheless, we acknowledge that most of our athletes were college aged and in good shape. Future collections will need to focus on data collection that is more representative of everyday athletes, since they would most benefit from off-the-shelf self-evaluation due to a lack of access to athletic trainers and coaches. An example of this type of study would involve the collection of data from local running, biking, or other intramural sports clubs. Fourthly, while data was collected in different locations across campus (providing the data with variety that built models will need to adapt to), we do not believe that we have collected data from all possible location and lighting types. In particular, we used a single type of camera for data collection, and while we believe that future data will be made robust to this limitation, this approach warrants further investigation. Fifthly, while we have established out-of-the-box baselines to ensure that the data can be processed, we would not consider these baselines to be sufficient for replacing the LESS method. Identifying and validating a model to replace the LESS method will be required for future work. Finally, we have limited our data collection to the stated two variations of evaluative movements. These movements are ubiquitous in sports medicine literature, but we plan to create a simple way to provide detailed information on them and their importance to those who would like to participate in ACL risk evaluation.

5. Conclusion

We have collected a large dataset composed of footage of athletes performing multiple iterations of evaluative jumps. These recordings were captured using consumer grade cameras in varied environments. We believe these recordings, coupled with recent work on body modeling and 3D reconstruction, will allow researchers to build models to accurately track key movement features, such as joint angle and jump height, in order to evaluate participants for ACL injury risk. ACL injury can incapacitate an athlete for a significant amount of time and creates risk for repeat injuries. Currently, risk assessment is only available to athletes who have expert coaches and trainers dedicated to survey their risk and provide training regimes to mitigate possible injury. Ideally, evaluative models created using this data will be usable for self-assessment, circumventing the need to rely on a personal trainer for assessment. For athletes without coaches, the potential is even more extreme, as it allows them the opportunity to be evaluated with an expert system at a low cost without needing access to either a high-end system or a low-end gym. While data collected for this dataset was collected in high resolution video format, such video should be considered the highest-end possibility for collection, which can be easily downsampled to obtain synthetic approximations of even lower-grade cameras. If researchers train models across these conditions we expect that they have the potential to maintain their integrity, even with data captured from cheaper web cameras. Throughout this paper, we have discussed various ideas for how researchers could utilize our data for future work. We plan to continue collecting additional data and augmenting our released data over time. While the data we have released
thus far is a strong starting point, we believe that future collections will allow us to incorporate even more variability into the data collection process, allowing models to have increased generalizability. At this time we did not collect gold standard measurements from high-end modeling systems. We made this decision in order to increase the feasibility of performing collections in a variety of locations and environments. Often, high-end systems require specific setups which are not portable. However, we plan to release summaries of the expected ranges for important measures (such as knee angle) as a guide for model builders without a strong background in the sports medicine field. These will allow those without access to high-end modeling systems to measure and understand the ranges of measures associated with increased ACL injury risk. In consideration of our access to trained professionals, we will have experts annotate the videos with LESS protocol [13] judgments in order to further highlight ACL injury risk characteristics. Thus far we have not collected expert annotations because ground truth data from human body tracking should be easily obtainable from the data sources, and are equally as indicative of ACL risk. Nevertheless, as we expand the dataset with more athletes we believe the LESS scoring data may be a useful shortcut for gut-checking modeling results. We note that the dataset benefits from the constrained nature of the evaluative tests, and if successful we plan to incorporate more challenging evaluations into our collection, such as reactive agility tests [16]. Finally, we hope to collect data using a variety of different cameras in order to test how different models are able to perform with different cameras. While much of this variation can be synthetically generated through either down-sampling image quality or camera mapping, we believe that the physical usage of different cameras will create a solid benchmark for such manipulations. We hope the computer vision community recognizes the importance of this dataset and joins us in our effort to keep athletes of all backgrounds healthy by preemptively identifying potential injuries.

References


