OFFICIAL

Video Machine Learning in Australian High Performance Sport

Insights Report for Performance Support Practitioners

Prepared by

Dr Christopher Papic

November 2023

Australian Sports Commission Acknowledgement of Country

The Australian Sports Commission (ASC) acknowledges the Traditional Custodians of the lands where its offices are located, the Ngunnawal people and recognise any other people or families with connection to the lands of the ACT and region, the Wurundjeri Woi-wurrung people of the Kulin Nation, the people of the Yugambeh Nation and the Gadigal people of the Eora Nation.

The ASC extends this acknowledgment to all the Traditional Custodians of the lands and First Nations Peoples throughout Australia and would like to pay its respects to all Elders past, present and future.

The ASC recognises the outstanding contribution that Aboriginal and Torres Strait Islander peoples make to society and sport in Australia and celebrates the power of sport to promote reconciliation and reduce inequality.

Executive Summary

Background

Video Machine Learning (VML) is a rapidly evolving field of computer vision that has the potential to automate video-based analysis within the daily performance environment (DPE). VML has practical implications for the timely provision of performance data within the training or competition environment.

Purpose

The aims of this Insight Report on VML use in the Australian High Performance (HP) sport system were to:

- Provide a current snapshot of performance support practitioners' perspectives and use of video analysis techniques and/or VML to analyse sport performance in the DPE;
- Provide an overview of VML applications, challenges, and best practices in HP sport;
- Develop recommendations for best practice implementation of this technology in HP sport.

Methodology

A movement science focus group developed an online survey for performance support practitioners on their use of video analysis (both qualitative and quantitative), as well as their current and planned VML use and their perspectives on barriers for integrating VML within the DPE. An external project lead, with specialist expertise in VML in HP sport, was appointed to develop the insight report. The project lead was responsible for assisting with survey design; summarising the survey responses; providing relevant information on the technology, challenges and barriers to integration, as related to HP sport; and providing summary recommendations for future use of VML in the DPE.

Support Practitioner Survey Findings

- The large majority of performance support practitioner survey responders use video analysis to support the coach and athlete.
- Less than one third currently use VML within the DPE, with most of these responders using sport-specific applications (such as Sparta2 for swimming).
- More than three quarters would like to use VML across various scenarios in the DPE to expedite performance data feedback to the coach and athletes.
- VML integration barriers included communication between practitioners across the network; accuracy and reliability concerns; time constraints; expertise; access to resources; and 'buy-in' from coaches and athletes.

Key Insight Report Recommendations

- Specifically trained VML models will produce better results for a given activity, athlete and testing environment than general VML models.
- The measurement technology (either current or VML approach) should reflect the sensitivity required for the performance parameters being assessed (Quality Assurance).
- Ensure model training datasets are representative of the activity and athletes being investigated (especially for use with some Paralympic athletes and atypical movement patterns) or VML outputs will be sub-optimal.
- Provide practitioner professional development and access to technical expertise for long-term implementation of this technology in the DPE.
- Develop VML best practice guidelines.

Contents

Executive Summary 3

Project overview 5

1. Glossary of Relevant Video Machine Learning and Data Analysis Terms 6

2. Summary of insights report findings 7

3. Introduction to VML 8

4. VML Applications in Sport 10

4.1 Performance analysis 10

4.2 Biomechanical analysis 12

5. Challenges and best practices 18

5.1 Data capture and modelling 18

5.2 Ecological validity of models 21

6. VML insights survey 23

6.1 Video machine learning use 24

6.2 Practitioner-defined challenges to VML 26

7. Summary of VML Insights Report Recommendations 28

8. References 31

Appendix A: VML insights survey questions 34

Project overview

Performance support practitioners (practitioners) throughout the Australian High Performance (HP) sport system are tasked with supporting coaches and athletes in skill development, performance analysis, technique analysis, and decision-making by providing evidence-based recommendations. Technical analysis in training and competitive environments frequently includes video analysis and evaluation of task performance, through qualitative and quantitative means. The balance between deciding what and when to test with vision-based analyses is an ongoing challenge for practitioners. The 'gold-standard' of optical motion capture in a lab can provide accurate kinematic data, but can be time-consuming and resource intensive, and lacks ecological validity which often leads to the outcome data not translating to on-field movement characteristics (Di Paolo et al., 2023). Whilst vision-based analysis in applied training and competition environments replicates the demands of the sport, feasible methods to obtain timely, valid, and reliable quantitative data are sometimes limited.

Video Machine Learning (VML) is the creation and use of models that have learned from a dataset to recognise distinct image features in video frames. The nature of this technology allows access to a large amount of information and analysis possibilities for investigating human movement with rapid processing times. Due to the complexity of these data processing, it is necessary to understand the limitations and usability of various VML technology before choosing an application appropriate for the performance question. The rapidly changing nature of computer and video technology, particularly in sporting applications, requires transparent external considerations around measures, limitations, and considerations for its application. Given the substantial impact that VML could play within the Australian HP Sport system, practitioners could benefit from expanding their knowledge about the processes underpinning VML and its practical applications to ensure a more informed use of this technology.

At present, VML technology use is primarily in the exploration phase throughout the Australian HP sport system. There may be documents and guidelines written by individuals, however, there are no global references to guide decision making, and the resources that do exist may be outdated and provide conflicting advice. In most cases the resources that do exist are unlikely to have been peer reviewed by experts within the field. It is proposed that a new set of guidelines be developed to provide up-to-date evidence-based recommendations around various VML software and processes for practitioners working in HP sport. To inform these recommendations, we require a greater understanding of the applications of VML in HP sport and perspectives of practitioners who are currently using video-based technologies to evaluate athlete performance.

The aims of this insights report were to:

(i) Provide a current snapshot of Performance Support Practitioners' perspectives and use when using video analysis techniques and/or VML to analyse sport performance in the DPE;
(ii) Provide an overview of VML including applications, challenges, and best practices of this technology in HP sport;
(iii) Develop initial recommendations for best practice implementation of this technology in HP sport.

1. Glossary of Relevant Video Machine Learning and Data Analysis Terms

Video machine learning (VML) = Application of machine learning techniques and algorithms to analyse and process video data for a variety of applications, such as, event recognition, object/point recognition, and human pose estimation.

Human pose estimation (HPE) = Estimation of the positions and orientations of key body joints and segments in video frames.

VML Software = are applications or software packages (either open source or commercial) that perform aspects of VML processing such as pose estimation. Common examples, as related to use with sporting applications, are DeepLabCut; Open Pose; BlazePose; AlphaPose and OpenCap. The Australian Institute of Sport have developed their own VML software called Pipelines.

VML Training dataset = Collection of annotated (real or synthesised) video or images (sometimes with complementary data such as inertial sensors outputs; 3D body scans; motion capture information etc.) that are used to teach a VML model how to recognise and classify specific features within videos. Common examples of VML datasets are COCO; HumanEva; Human3.6; TNT15; Halpe-FullBody; AGORA, ASPset-510; PIFU and ICON.

ASPset-510 = Australian Sports Pose Dataset that includes videos with 3D pose annotations across a variety of dynamic sports-related actions in applied environments.

Common Objects in Context (COCO) = Widely used dataset in VML which includes a comprehensive collection of annotated images and videos of people.

VML Training = The process where a computer 'learns' from examples of annotated videos (training dataset) and then adjusts its internal settings (see "weights" below) based on those examples over many training iterations (e.g., 100s of thousands). This process helps the computer learn to recognise patterns and make classifications in new, unseen videos.

Transfer learning = A method of using a pre-trained model on a large dataset as a starting point to train a new model for a specific video analysis task.

Ground truth = Reliable and objective reference data (e.g., annotated video frames) that is used to validate VML models.

Convolutional neural network (CNN) = type of deep learning model designed for analysing and processing video data. It uses specialised layers called convolutional layers to automatically extract and learn features from input data, which is effective for tasks like image object recognition and event classification.

Weights = Numerical values that determine how important the features from input data are to the final output classification. Adjusting these weights is a fundamental aspect of VML model training and fine-tuning to improve the accuracy and performance of the model.

Red-Green-Blue (RGB) = Primary colour channels of Red, Green, and Blue used to create and display video images. Each pixel in a video frame is defined by its specific R-G-B values, which determine the colour and appearance of the pixel.

Functional Data Analysis (FDA) = statistical methods used to analyse continuous, waveform, or time series data where these data are modelled as functions.

Computer Vision (CV) = Computer vision is concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images. It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding.

2. Summary of insights report findings

Performance Support Practitioner Survey findings

- Video analysis techniques to support athletes and coaches are used frequently across a broad range of sports and settings (indoor and outdoor), during both training and competition.
- Video analysis is primarily utilised for within- and post-session qualitative feedback and post-session quantitative feedback.
- Only one in four practitioners provide quantitative feedback as part of video analysis of a given task within a training or competition session.
- Barriers identified for using video analysis techniques include:
    - Time/labour constraints,
    - Limited access to resources, and
    - Lack of expertise.
- Practitioners have a good understanding of VML principles and the potential benefits of implementing this technology in HP sport.
- Nearly one third of practitioners are currently implementing and/or testing VML technology in sport.
- Around three quarters of practitioners are planning to implement or are interested in implementing VML technology in the future. As current applications and plans for VML across the network varied in terms of sports, settings, and the software being used, there is a need for transparent guidelines for evaluating their validity and reliability.
- Practitioner identified challenges for implementing VML are:
    - communication between practitioners across the network,
    - accuracy and reliability of models,
    - time constraints,
    - expertise, access to resources (e.g., appropriate computer hardware), and
    - ‘buy-in’ from coaches and athletes.

Applications of VML

- VML has the potential to substantially improve the speed of data analytics in a non-invasive manner, allowing for more timely and representative performance feedback to coaches/athletes within training and competitive environments.
- There are two main applications of VML technology typically associated with performance support servicing in the Australian HP sport system:
    - performance analysis involving event and/or action recognition, and
    - biomechanical analysis, such as, pose estimation and derivation of athlete kinematics.
- From a quality assurance and VML applicability perspective, understanding what represents a meaningful change for the end-user and the required sensitivity of the quantitative outputs is essential to determine if VML is an appropriate analysis methodology.

VML Technology Considerations

- Whilst bespoke, opensource, and commercial VML approaches have had some published levels of validation, their reported accuracy varies and application to sport-specific environments, tasks, and athletes will further influence model accuracy.
- Quality of data-in will improve quality of data-out of a model; practitioners will need to consider how video data are captured, the size of the datasets, and whether the datasets are athlete and sport-specific and translatable for the given task.
- Open source and commercially available software with user friendly interfaces and customisable performance analytics may be more feasible for practitioners to implement across the Australian HP sport system than developing bespoke solutions.

3. Introduction to VML

Machine learning is the creation and implementation of computer algorithms that have learned from a dataset to classify or estimate an outcome. Machine learning techniques in sport research are gaining momentum and are used across a variety of applications, such as, sport injury prediction (Rossi et al., 2018), the detection of movement insufficiencies in injured athletes (Richter et al., 2021), and the estimation of kinetic outcomes (e.g., ground reaction forces, joint moments) from motion capture data (Kipp et al., 2018; Mundt et al., 2020). VML refers to the arm of machine learning algorithms that are specific to image recognition in video. This can include classifying specific events or actions during sport (e.g., a goal attempt in soccer or foot strike detection for automated analysis) (Rangasamy et al., 2020), player tracking (Vats et al., 2022), identifying objects or landmarks based on their visual characteristics (Mathis et al., 2018), and/or estimating joint centres and body segments based on biomechanically-informed datasets; Human Pose Estimation (HPE) (Badiola-Bengoa & Mendez-Zorrilla, 2021).

VML models are deep neural networks which are often in the form of a convolutional neural network (CNN), a highly efficient algorithm for image processing (Yamashita et al., 2018). CNNs are a class of neural networks that are, in essence, representations of neurons within a human brain and are inspired by the organisation of our visual cortex. These models use image feature learning where the probability of a match is calculated, known as a "weight", between the Red-Green-Blue (RGB) pixel characteristics for a region of the video frame, known as the "input", and the RGB characteristics of the region surrounding the object, referred to as the user-defined "ground truth". CNNs include several layered networks that perform different calculations and manipulations of the image before classifying the object (Figure 1).

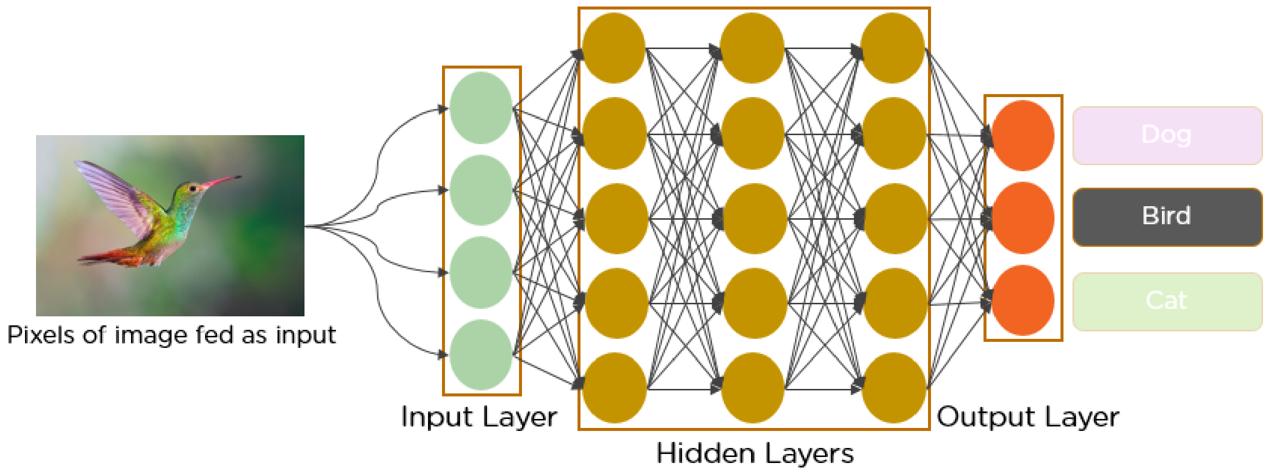

Figure 1. Bar chart Visual representation of a convolution neural network for image feature recognition. (https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/)

Consider a CNN that has undergone training to identify the centre of a cricket ball in a collection of video frames. The first layer of the network scans small sections of the entire image systematically (across each RGB channel) extracting basic features of the image into arrays of pixel values. The arrays from this first layer then taken through a range of other layers (e.g., convolution layers) that reduce their size, but increase their complexity – making it more effective at detecting subtle differences in image features (i.e., the cricket ball travelling across the field of view). The final "fully connected layer" classifies the object in the image, along with a "loss" value – how accurate is the classification compared with the ground truth. The output in this example could be the *x*- and *y*-axis pixel coordinates of the estimated centre of the cricket ball in each video frame. Training of CNNs involves many iterations of the process detailed above (e.g., 100s of thousands) where the weights in these layers are continually optimised to minimise loss and improve the accuracy of the final model.

When developing VML models, the process of transfer learning is often used to significantly reduce the size of the training dataset and the time taken to train the overall network. Transfer learning uses a set of weights previously trained to identify image characteristics in a very large image database as a starting point for a new model (Mathis et al., 2018). Transfer learning involves updating the pre-trained weights by comparing the input with the ground truth in new images (e.g., a dataset of manually labelled body landmarks on an athlete). As an analogy for this type of learning, imagine a person is interested in bird watching; they already have the capacity to recognise birds in the world, but are now shown select species of birds in a forest by an expert (think of this as the ‘ground truth’). With practice they become more efficient at spotting and classifying these species of bird (except that VML models, compared with the person, can move, and visually scan the forest at lightning speed).

Given the speed in which VML software can detect features in large video datasets, these tools have been identified to have the potential of being implemented for performance and biomechanical analyses within HP sport. These tools can potentially be used to automatically process videos of athletes to expedite player tracking information, biomechanical variables and key time events. The potential outcome can include the ability to derive performance visuals (e.g., video tracking overlays and figures) which may be used to support feedback to athletes and coaches (Figure 2).

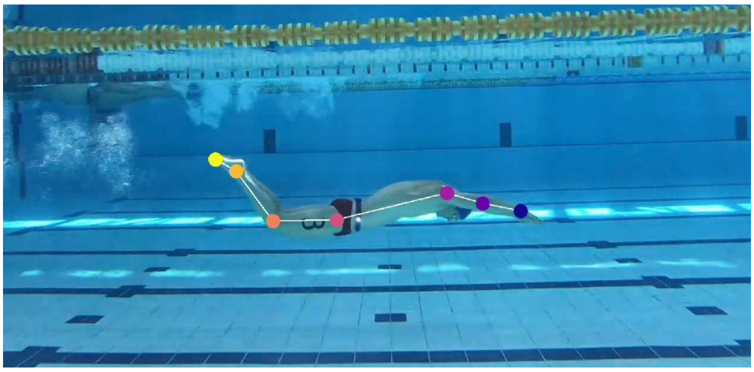

Figure 2. Example of VML software DeepLabCut labelled video output.

4. VML Applications in Sport

4.1 Performance analysis

VML methods can be used analyse performance outcomes, often in competition or during team-based matches where athletes are tracked across a field, court, or swimming pool. Accurately detecting and tracking players in competition or team-based sports provides crucial training and competition insights and is a crucial component of the role of a performance analyst. Professional sporting codes around the world (such as basketball, football and tennis) have also started using VML approaches for player tracking information. VML approaches have the significant advantage over other tracking modalities in that they allow for the tracking of both teams (or opponents) concurrently, providing potentially crucial information on team and individual tactics and strategies. This information can also potentially be used in the future to develop further novel team tactics and strategies using other machine learning approaches. This approach has already been employed by Tennis Australia (in an exploratory analysis) to leverage generative machine learning process from prior player tracking data to develop strategies against opponents.

Highlighting the recognition of the need for quality control and accuracy of the optical player tracking system output, the International governing body for football (FIFA) outlines a comprehensive process for determining accuracy levels against gold standard measurement techniques (FIFA, 2022). Processes for assessment includes comparative analysis of player and ball movement (within a 30x30m test area covered by Vicon motion cameras) using a series of agility tests and specific football movements. Accuracy of straight-line running velocity is compared using laser tracking and a field survey with quantity surveying equipment is used to provide comparative absolute position relative to position within the pitch. While stadium/court level player tracking systems using VML approaches and in-situ calibrated systems are not a feature within Australia currently, there is potential that this could become viable in future years.

In conjunction with player tracking, VML typically within HP sport scenarios would also incorporate event recognition. These event recognition functions can use image feature, player tracking, and/or pose estimation methods to differentiate specific sports events. For example, a Soccer Video Scene and Event Dataset was developed to include 100 short videos of key game events such as corners, free-kicks, goals, and penalties (Hong et al., 2018). The location of clusters of players on the field generally differs between each of these events, and thus, represents different image features in the video frame that can be extracted by the VML model. Several CNN models, utilising transfer learning methods, were trained on the dataset of these events and their known classification. These models had capacity to automatically detect these events in video of an elite level soccer game, based on player positions on the field. Accuracy varied from 74-89% for the two best performing models, where reduced accuracy was observed when the number of event types included in the dataset increased from four to six. A limitation of this type of modelling is that similar player positions on the field can occur between two or more event classifications, such as a free kick located near the corner of the field and a conventional corner. The importance of appropriate feature selection methods (e.g., the number of event classifications included in a dataset) to minimise classification errors is discussed further in this document. Another application of event recognition models included a dataset of 200 short video sequences of successful soccer goals from a variety of settings and camera angles and 200 sequences of non-goals, with a variety of events such as near-miss shots sourced from YouTube.com and classified by the researchers (Tsagkatakis et al., 2017). CNN models, again utilising transfer learning, were then trained to classify goal events in soccer games with one type of VML algorithm showing near perfect event detection accuracy (98%). Event recognition models can reduce time spent encoding the frequency of key game analysis events throughout a game by practitioners and can expedite review of these events when providing feedback to athletes and coaches.

As an example, within the Australian HP system, Swimming Australia has provided race analysis for its targeted swimmers in benchmark events over the last 25 years to assist coaches/athletes to breakdown components of the athletes' racing performance. Key skill-based and pacing metrics are provided over each lap of the race to gain insights into areas for improvement and for comparative analyses. Traditionally this has been conducted with a significant human resource cost with analysts individually annotating races post-race to develop these metrics. Since 2017, Swimming Australia has led the development of a semi-automated system called Sparta 2 which has incorporated a VML module for tracking of all swimmers in the race (Elipot, 2019; Hall et al., 2021). This system uses both an individual tracking function relative to a pool calibration procedure with embedded event recognition routines to determine functions such as identifying individual stroke cycles. The VML module consists of a 2-step CNN. The first step processes the entire video frame (resolution: 4K; bitrate: 150mb/s) and aims to approximately detect the swimmers' position in the image. Small image crops are then produced around each head and processed to accurately identify the centre of each swimmer's head (second step). This 2-step approach gives the best trade-off between detection accuracy and computation speed. Once the swimmers' head are localised in the image, a final series of convolutions is applied to perform the stroke detection. To be able to implement this VML module, Swimming Australia built a very large training dataset which was entirely manually annotated by trained operators. This dataset contained approximately 900k instances representing the centre of a swimmer's head racing the 4 strokes and in very diverse conditions (12 different pools, indoor and outdoor; international or domestic environments; different camera angles, lenses and focal lengths etc.). This work was performed over an 18-month period.

The Lead Biomechanist at Swimming Australia (Dr. Marc Elipot) led a project to determine the reliability of the VML performance metric outputs of Sparta 2 when compared to trained operators. Results showed that the Sparta 2 VML system produced substantially reduced typical error rates compared to manual annotation across most of the metrics when analysed at a 95% confidence interval. While not a validation process, in that there was no ground truth comparison, the VML system was shown to be substantially more reliable. Variations in the manual annotation were attributed to human error and calibration errors in metrics associated with manually using lane rope buoys to determine criterion distances.

Event recognition can also include classifying the type of action(s) performed by an athlete. For example, a VML model was trained on a dataset of tennis actions comprising video from 55 different subjects performing multiple repetitions of distinct tennis shots (Gourgari et al., 2013). This model was able to detect several key tennis actions, such as backhands, forehands, and services with an accuracy of 88.16% (Vinyes Mora & Knottenbelt, 2017). Likewise, Rezaei and Wu (2022) utilised a layered VML approach of several algorithms to classify when the action of heading occurred during soccer. Figure 3 summarises the key stages of their VML modelling approach. They utilised a pre-existing soccer video dataset, (Soccer DB - (Jiang et al., 2020)), which included over 60,000 annotated images of the ball. Automatic detection of the ball was developed through a pre-trained object-based detection algorithm that was applied to the training dataset to identify the location of the ball (Figure 3). The frames were then automatically cropped around the ball, known as the ‘bounding box’, and the cropped images were then applied to another deep learning algorithm known as a temporal shift module, for spatiotemporal feature aggregation - trained to classify when the ball met the player’s head. The model ‘DeepImpact’ correctly identified 92.9% of headers in video of five new soccer games, which included 546 headers and 60,018 non-header actions. These recognition models can have a variety of implications for practitioners, with injury monitoring being one such application. In this example, there is emerging evidence that repeated sub-concussive impacts associated with soccer heading may pose a risk for cumulative brain changes (Lipton et al., 2013). A model like DeepImpact could be used to track overall team header exposure across training and competition, and subsequent models have been developed to include individual player tracking and monitoring.

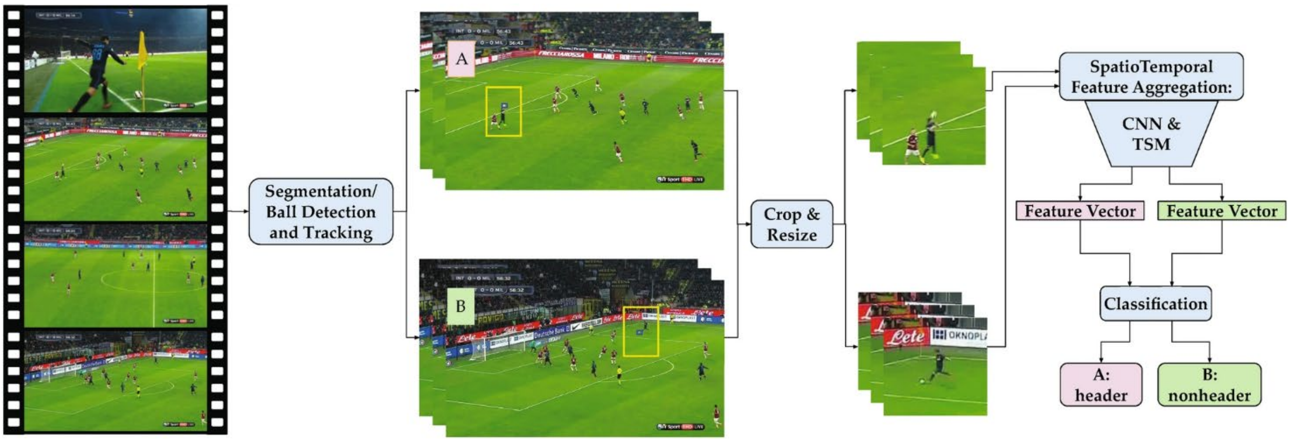

Figure 3. Header detection VML model summary. Reprinted from “Automated soccer head impact exposure tracking using video and deep learning” by A. Rezaei, 2022, Scientific Reports, 12, p3.

4.2 Biomechanical analysis

To preserve ecological validity when analysing an athlete's performance and movement patterns, the ability to quantify critical variables within a training session or competition in a non-invasive manner is paramount. The caveat to this is that the results must be accurate and reliable within the bounds of the sensitivities required to detect a meaningful change or discriminate between performance levels. There is significant potential for VML technologies to expedite video-based biomechanical analysis. However, the limitations of technologies need to be understood and addressed based on the end use-case. Included in the consideration is an understanding of the movement pattern to be analysed and therefore the resulting methodologies to be employed (e.g., single versus multi-camera; 2D versus pseudo-3D approximations; marker-based tracking versus estimated joint centres).

For bespoke applications, a simple marker-based system can be appropriate to provide key tracking information. To illustrate an example in this area, in cycling, tracking the approximated hip position with a marker-based system using image-recognition techniques has been used within the Australian HP system previously as a method to investigate and quantify the effectiveness of sprint starting techniques. Free and open-source video analysis programs (such as Tracker – a video analysis and physics modelling program - https://physlets.org/tracker/) automatically track identified markers in a pre-calibrated space to allow for rapid performance-based feedback to be provided to athletes within a session, thereby reinforcing skill modification interventions.

While the previous cycling example is representative of a standard image-recognition computer vision (CV) approach rather that technically a VML system, these marker-based systems can expand the tracking complexity through the additional use of VML processes. Point recognition VML models are suitable for translational movements with small user defined targets (Zhang & Yang, 2021) which can include marker-based systems. The open-source VML software DeepLabCut (Mathis et al., 2018) was originally developed for tracking animal behaviour based on user defined landmarks in video and offers a higher level of customisation. This software has been applied in sport research to reduce time associated with manual digitisation of body landmarks. As an example, a simple marker system (black painted circles) was applied to the skin over key body landmarks (e.g., shoulder, hip, knee) on competitive swimmers and video was captured of underwater glide trials to evaluate hydrodynamic outcomes (Papic et al., 2021). A neural network model consisting of 400 annotated frames of glide video was trained to detect these landmarks and digitised new glide videos at an average rate of 54 landmarks per second. A calibration procedure and subsequent kinematic analyses were then performed to derive 2D kinematic outcomes, such as, glide factor (an estimate of the swimmer's hydrodynamic resistance), angles, and velocities.

Despite their popularities and the outstanding results reported within CV domain, VML is still relatively new technique for quantitative motion analysis within the biomechanics discipline. The majority of existing VML techniques or technologies are derived from human data that are prepared and annotated by untrained personals, or by those who have limited background in the mechanics of the human body, shape, joints and functional movement patterns. Within the biomechanics domain, VML technologies can be used to approximate an underlying skeletal structure of a person using one of two main general methodologies. Over the last decade much of this development and research has been associated with the Human Pose Estimation (HPE) VML processing. More recently, and of interest to the biomechanics discipline, is the development work utilising shape analysis and incorporation of complementary technologies.

The shape analysis techniques are a promising area of development that revolves around the identification and segmentation of body shape and body segments. Included in these technologies are the inclusion of complementary information (termed 2.xD features within the Computer Vision field) such as normal and depth maps extracted from standard 2D image views. These allow extra layers of information to assist in the derivation of 3D shape, 3D features, 2D and 3D segments, body lengths and structural hierarchy. All of this information then provides extra detail and/or constraints that can be applied during the VML training process. The significant benefit of this methodology is that it has the scope to provide alignment to a more comprehensive and representative anatomical framework with the ability to quantify the rotation of segments that can be missing when using a single joint centre estimation point. This then has the potential to provide potentially greater representative movement pattern information.

While computationally and technically more complex than current pose estimation models, the potential benefit of this type of solution is to derive a more accurate biomechanically and anatomically informed outcome to counter the limitations in pose estimation models being developed in the CV domain. Less computationally expensive variant can also be facilitated under certain conditions, especially in cases where 3D surface representations of the human shapes are not required. Recent work published in this space using synthetically and semi-synthetically generated data ((Patel et al., 2021; Saito et al., 2019; Xiu et al., 2022) using the AGORA, PIFU, ICON databases and technologies respectively) indicate that by instilling additional 2.xD features with anatomically and biomechanically relevant information, the VML methodology will result in more accurate and robust VML models with biomechanically meaningful outputs.

The second methodology centres around the development of human pose estimation (HPE) models through the localisation of annotated joint centre or body landmark approximations and this has become a rapidly growing area of research in the VML domain with a seemingly endless array of applications. HPE models differ from point recognition models as they estimate the centre of the joint and segments of the body rather than a specific landmark or target (e.g., the location of the marker positioned on the athlete). There are numerous opensource software for marker-less HPE such as OpenPose (Cao et al., 2017), BlazePose (Bazarevsky et al., 2020), AlphaPose (Fang et al., 2022), and OpenCap (Uhlrich et al., 2022) which include models that were trained on a variety of datasets and learning algorithms. The Common Objects in Context (COCO) dataset, for instance, includes over 60,000 annotated images of people (Lin et al., 2014). An exemplar learning algorithm for HPE are Part Affinity Fields (PAFs) which are used in the OpenPose model and comprises of confidence maps and 2D vector fields that encode the location and orientation of body segments throughout the image (Cao et al., 2017). This method is a 'bottom-up' approach for detecting pose and is common in HPE software (e.g., 'DeeperCut' (Insafutdinov et al., 2016)). These algorithms first detect all the human joints in an image and then assemble poses for everyone, linking segments for each person, as illustrated in Figure 4. Camera calibration procedures can then be applied to 2D coordinate data from multi-camera HPE to reconstruct pseudo-3D poses using triangulation methods (Köykkä et al., 2022). There has also been some opensource models that have a 3D variant (eg. 3D OpenPose). Concurrently, there has been recent additional attempts to develop specific 3D pose datasets incorporating syntheised information from animation sources, or small datasets with limited movement patterns from indoor marker-based system (eg. Human3.6M, HumanEva, Halpe-FullBody, AGORA).

Figure 4. Example of 2D multi-person HPE. Reprinted from "DeeperCut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model" by E. Insafutdinov et al., 2016, Proceedings of the Computer Vision–ECCV: 14th European Conference, Amsterdam, p35.

In either 2D or 3D pose estimation approaches, the main objectives were to train ML models to be able to predict the 2D or 3D ground truth labels, or representations of the labels (e.g., heatmaps). This is performed subject to loss/cost function that is defined to minimise the mean squared error (or similar) between the predicted results and the ground truth. For more information summarising the different approaches used in 2D and 3D pose estimation and the limitations, advantages and disadvantages of each, the following articles are suggested (Dubey & Dixit, 2023; Needham et al., 2021; Wang et al., 2021; Zheng et al., 2023).

It is important to note that most of these 2D and 3D HPE models do not contain the ability to provide individual segment orientation analysis (which is similar to the limitation of standard traditional '3D' digitising systems like APAS or Peak), as the segments are only defined as a vector between scalar points (approximated joint centres). 2D images are only able to provide 2D model outputs with reference to a (X,Y) cartesian coordinate system and the majority of VML publicly available pose estimation models and databases incorporate this type of information. This can be thought of as equivalent to manual 2D video field-based analysis undertaken by performance support practitioners through 2D digitising programs like SiliconCoach or Kinovea. However, as trained performance support practitioners performing 2D manual digitising, there is underpinning quality assurance processes and knowledge that is assumed to be applied. This involves being cognizant of the constraints that this type of testing involves (such as camera position relative to plane of action and that it is not possible to accurately provide 2D joint angles when movement of those segments is outside this plane). Coupled with this is an anatomically reference training process that practitioners undergo to assist with the identification of the functional joint centre locations through a range of motion about each joint. VML approaches are not able to discriminate these features and will report their derived joint angle value irrespective of these limitations. It is therefore essential that the performance support practitioners take on a quality assurance role to provide a manual hygiene oversight when utilising these technologies.

True 3D outputs from 3D HPE models would require segment orientation, which necessitates a minimum of 3 markers per rigid segment being accurately identified (as per the data collection and subsequent modelling pipeline used in marker-based systems like Vicon). The 3D HPE models are termed pseudo-3D in that they only have a derived (X,Y,Z) with respect to a global coordinate system. From a technical perspective, while the global space is 3D, the human model remains essentially 2D given the true 3D pose (3D position and orientation) of each rigid segment is not able to be determined. The use of a triangulation approach from the output of multiple cameras in HPE also makes these outputs susceptible to errors in estimating joint centre from a given camera view. Therefore, high accuracy of 2D pose estimation of each estimated joint position from each camera view is required for adequate 3D reconstruction (Kitamura et al., 2022). Number and optimisation of camera position relevant to the movement pattern being analysed, in addition to the accuracy in determining the intrinsic and extrinsic camera and volume calibration variables, will also influence the accuracy of the pseudo-3D outputs.

Opensource VML software has been applied to a variety of sport-specific actions and settings for kinematic analyses. For example, OpenPose has been applied to running (Van Hooren et al., 2023) and cycling (Serrancolí et al., 2020), BlazePose to javelin (Köykkä et al., 2022), and AlphaPose to badmington (Ding et al., 2022). As a sample lab-based comparison of these three Opensource VML software (OpenPose v1.6.0; AlphaPose v0.3.0; and DeepLabCut’s pre-trained human pose model v2.1.7) against outputs from a 3D full body marker-based motion capture system (Qualysis), comparative trials of walking, running and countermovement jumps were performed (Needham et al., 2021). The researchers reported a large systematic difference (of between 30-50mm) in the approximated hip and knee joints; while much smaller variance was found in the ankle joint (1-15mm) which varied depending on the activity (see Figure 5). Where there are large systemic differences in approximated joint centres, errors are compounded when differentiating these data for velocity analysis, further highlighting the need for training datasets that incorporate biomechanically informed joint centre annotations. OpenPose and AlphaPose demonstrated comparable performance to one another and outperformed DeepLabCut. The researchers suggested that the position differences from the gold standard 3D marker-based motion analysis was likely due to mislabelling of ground truth data in the training datasets. This highlights that a simple comparison of HPE outputs to a gold-standard marker-based system is problematic. The standard HPE datasets are not constrained to any real functional joint centre, while traditional marker-based system incorporates significant modelling processes to determine the functional joint centre throughout a segment range of motion.

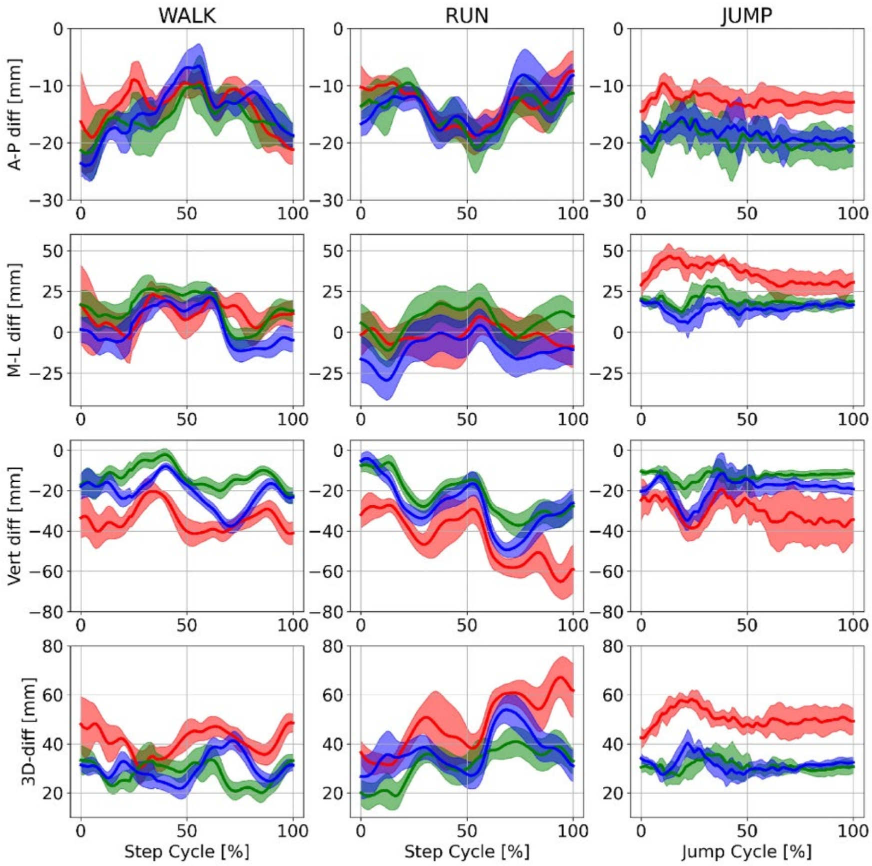

Figure 5. Example mean (± SD) differences between marker based and markerless trajectories for the right hip joint centre trajectories of a single participant (P10) during walking (left), running (centre) and jumping (right). Marker based trajectories (black), OpenPose (green), AlphaPose (blue) and DeepLabCut (red). Row 1 = anterior–posterior differences. Row 2 = medial–lateral. Row 3 = superior–inferior differences. Row 4 = 3D Euclidean differences. Reprinted from "The accuracy of several pose estimation methods for 3D joint centre localisation" by L. Needham et al., 2021, Scientific Repots 11(1), 20673.

A prominent opensource VML software package for using HPE to estimate both the 3D kinematics and 3D kinetics is OpenCap, which has been in development at Stanford University since 2019. OpenCap comprises an iOS application, a web application, and cloud computing. The web application enables users to record videos simultaneously on 2 or more iOS devices and to visualise the resulting estimated 3D kinematics. Figure 6 illustrates an overview of the OpenCap data collection and analysis methods used to estimate 3D kinematics and kinetics. Using cloud-based computing, 2D keypoints are extracted from multi-view videos using either OpenPose or HRNet algorithms, with these algorithms reportedly used due to their performance and inclusion of foot keypoints (Uhlrich et al., 2023). Pseudo-3D keypoints are computed by triangulating these synchronised 2D keypoints. These pseudo-3D keypoints are converted into a more comprehensive 3D anatomical marker set using a recurrent neural network (LSTM) trained on motion capture data. 3D kinematics are then estimated from marker trajectories using deep learning models and inverse kinematics in OpenSim. Finally, kinetic measures are estimated using muscle-driven dynamic simulations that track 3D kinematics.

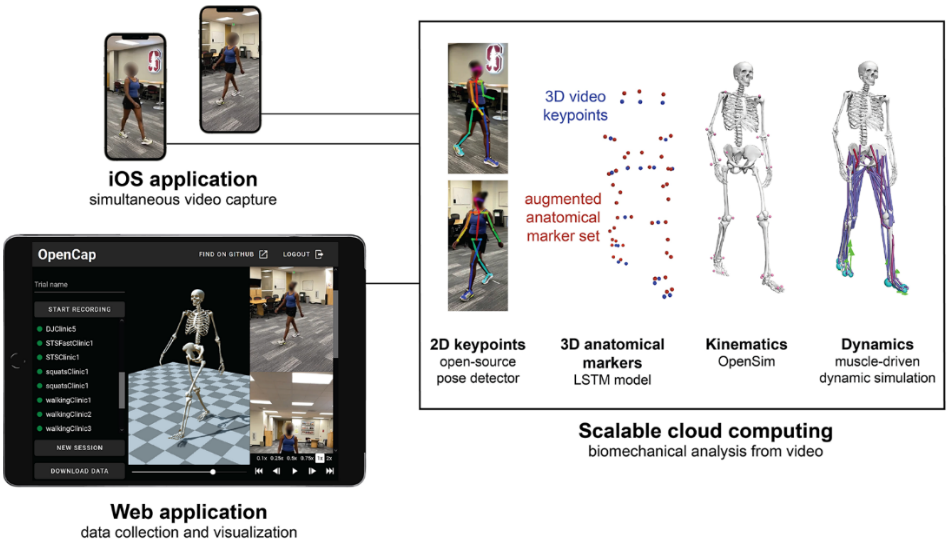

Figure 6. Simplified overview of OpenCap data collection and analysis pipeline for estimating 3D kinematics and kinetics. Reprinted from "OpenCap: Human movement dynamics from smartphone videos" by S. D. Uhlrich et al., 2023, PLOS Computational Biology, 19(10), e1011462.

In a recently published article, the OpenCap was compared to results from an 8-camera marker-based motion analysis system and force plate analysis using 4 separate simple dynamic activities: walking; squats; sit-to-stand and drop jumps (Uhlrich et al., 2023). For the laboratory-based ground truth comparison, the 3D kinematics and kinetics were calculated from the measured marker and forceplate data using OpenSim with the same modelling and simulations pipeline used in the OpenCap process. Results shown in Table 1 show the results of the comparative analysis for 10 healthy adults. Errors for each activity were averaged over trials and participants (n = 10), and the reported mean is an average over activities and degrees of freedom (six for pelvis position and orientation [kinematics only], three for the lumbar, three per hip, one per knee, and two per ankle). Kinematic and joint moment errors are presented as the mean and range over the degrees of freedom, and kinetic errors are additionally presented as the MAE as a percentage of the range.

Table 1. Mean absolute error (MAE) in kinematics and kinetics from OpenCap compared to laboratory-based motion capture and force plates. Reprinted from "OpenCap: Human movement dynamics from smartphone videos" by S. D. Uhlrich et al., 2023, PLOS Computational Biology, 19(10), e1011462.

Kinematics (MAE)	Walking	Squat	Sit-to-stand	Drop jump	Mean
Rotations (n = 18) [°]	4.1 (2.3–6.6)	4.1 (1.8–7.2)	4.7 (1.7–10.3)	5.1 (2.3–8.6)	4.5
Translations (n = 3) [mm]	12.3 (6.8–19.6)	12.3 (5.8–18.4)	13.2 (5–20.3)	11.5 (6.3–16.5)	12.3
Ground reaction forces (MAE)					
Vertical [%BW]	8.2 (7.5%)	6.4 (20.0%)	5.7 (13.4%)	25.2 (13.8%)	11.4 (13.7%)
Anterior-posterior [%BW]	2.1 (6.7%)	1.3 (37.5%)	1.9 (31.0%)	8.9 (17.3%)	3.5 (23.1%)
Medio-lateral [%BW]	1.1 (17.1%)	5.7 (85.4%)	3.2 (110.5%)	5.3 (29%)	3.8 (60.5%)
Joint moments (MAE)					
All degrees of freedom (n = 15) [%BW*ht]	0.75 (0.20–1.32, 19%)	0.97 (0.11–1.93, 45%)	0.68 (0.13–1.09, 60%)	2.50 (1.15–5.90, 25%)	1.22 (37%)

The authors also highlighted that the outcome of OpenCap analysis can be influenced by both environmental and experimental factors. To maximise the accuracy of the results the authors recommended following best practice processes when collecting video data of the activity. These included recommendations on the subject clothing and background; placement of cameras to minimise occlusions; adjusting to common subject entry points in the fields of view; and positioning the cameras at an optimal distance from the subject and calibration locations.

In addition to these Opensource VML software, commercially available software for marker-less tracking is now available, such as, Theia3D (https://www.vicon.com/software/markerless/). Theia3D software is easy to use but relies on a multicamera setup. When assessed against marker-based motion capture for walking, squatting, and forward hopping Theia3D was accurate for sagittal plane knee flexion (1-3°) but showed greater error in other joints and planes of motion (e.g., 22° for hip flexion-extension during squatting, 8° for ankle flexion-extension during hopping, and 8.29° for ankle rotation) (Ito et al., 2022). The advantage of these software, however, is that they also include supporting analysis software. In the case of Theia3D, Vicon

ProCalc 1.5 can be used to generate joint data and visuals, or to export results into Visual3D, Python or MATLAB for further biomechanical analyses.

While understanding the accuracy of reliability of VML approaches relative to a gold-standard is appropriate, it is also important to highlight that the end use-case of VML is often a field-based testing scenario. In this case, it is essential to understand how the VML approach would compare to current scenario-based testing options. Typically, this would be a 2D video analysis protocol aligned with identified critical performance variables. As such, validation of the current technology is also appropriate to determine if the potential advantages of VML, through its decreased analysis time, is an appropriate substitution based on the performance requirements. To highlight this, compared to a gold standard marker-based opto-reflective system (Vicon); '3D' video digitising of the elbow angle was performed during a cricket bowling action using a mechanical arm moving through a known range of motion (Elliott et al., 2007). Results showed that RMS error for the marker-based system was 0.6°, with the '3D' manual digitising at 2.3°. However, in real-world tests, with shirts occluding parts of the segment and joint centre, the RMS error was reported as significantly higher. This highlights the need for validity testing of both the traditional analysis protocol and the proposed VML process to happen in-situ and concurrently.

With a view to utilising VML approaches to automate simple kinematic analyses in biomechanical training or competition analysis, the ability to incorporate accurate automated key frame detection would be a crucial first step in the provision of representative VML output data. An example of this could be in the provision of key performance-based parameters such as the degree of knee bend during eccentric loading of the take-off leg during long jump or high jump take-off stance phase. The ability to have an accurate foot strike detection algorithm within the VML pipeline would be critical. Likewise, effective algorithms for toe-off detection would also be crucial in determining accurate stance phase time.

Preliminary results from Dr. Mundt from UWA (within current AIS research project on the 'A Biomechanically Informed Pose Estimation Model for Australian High Performance Sport, 2023) has highlighted the inefficiencies in current algorithms to define key events from the simplistic pose estimation outputs to the level of accuracy that would be required for critical variable outcome feedback. To highlight this, the knee flexion/extension angle during a series of moderate running trials was quantified at the frame prior (based on 50Hz filming) and after for their defined foot strike and foot flat conditions. Substantial performance outcome differences were highlighted based on the time point selected (see Figure 7), reinforcing the importance of accurate and representative key frame detections within the VML processes.

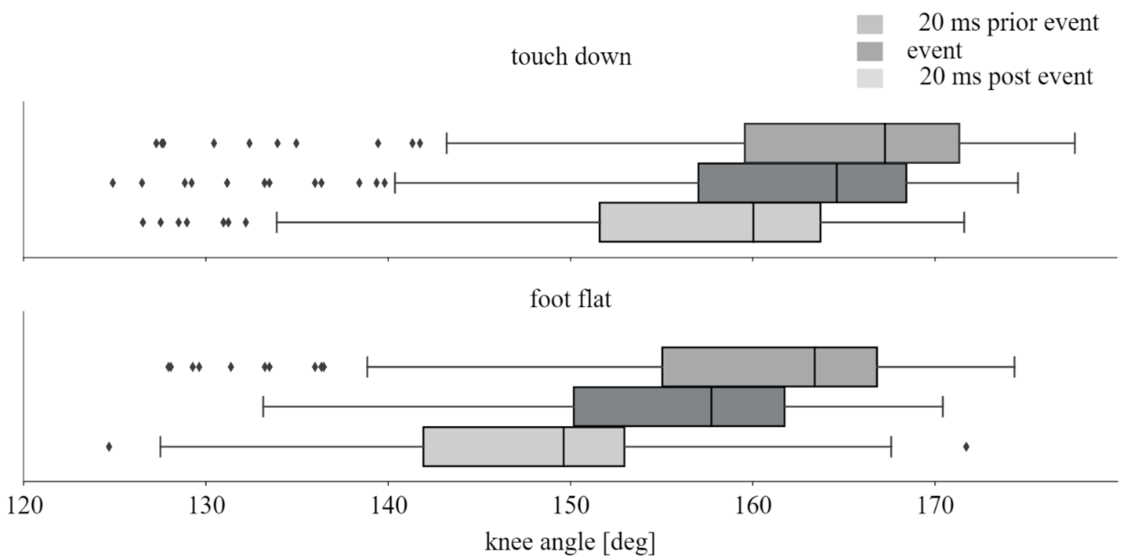

Figure 7. Average knee joint calculation through variances in foot strike detection and foot flat detection for moderate speed running in OpenPose.

It is essential to recognise that any reported error analysis or validation needs to be considered with situational context and applying referenced error rates is not sufficient to ensure that the resulting performance outcomes are within the degree of accuracy and reliability that can be applied to a field-based setting. Validation of the action of interest in the environment (or range of environments) that the VML will be used in will be essential to determine validity and reliability in that sports-specific and environment scenario.

5. Challenges and best practices

VML technology shows significant promise for the future of sport science, however, in its current form it is not without limitations. Recall the bird watching analogy from earlier in this report. Now imagine the 'trained birdwatcher' is placed in a new environment; it is foggy, the trees are denser and different colour, and there are several species of birds that are like those that they had originally been trained to identify. The capacity to locate and classify bird species is now limited and may lead to false positive classifications. When the margin for winning and losing in sport is so small, the accuracy of performance outcomes derived by VML needs to be appropriate as it may lead to the provision of incorrect or misleading information. The potential risks escalate when this information is used to inform changes to an athlete's program, strategy or technique as the long-term effect of wrong information in these cases can then be significant. Understanding challenges associated with VML can help inform best practices when implementing this technology in Australian HP sport.

5.1 Data capture and modelling

Figure 8 illustrates several key domains that need to be considered when developing a robust VML model. Data capture quality and specificity of the dataset are paramount. Unlike other experimental designs, a priori sample size calculations for VML datasets aren't applicable (Richter et al., 2021). Ensuring training datasets are large enough and specific are essential for ecological validity of the model; "quality-in equals quality-out". Many opensource and commercially available VML models are trained on general populations performing non-sport specific actions and therefore estimated joint centres are not based on biomechanically informed datasets (Rapczyński et al., 2021). Coupled with this is the issue that many of the data sets are mass-annotated by people without any specific training in defining anatomical joint centres or landmarks, leading to problems in joint centre approximations resulting from movement of a segment about a joint through its full range of motion. High quality sport-specific datasets are generally small-scale but could be artificially enlarged by leveraging on pre-annotated motion capture datasets (Mundt, 2023). However, given that these datasets are generally derived from laboratory settings and include markers in the video, their ecological validity in applied settings without makers comes into question (Nibali et al., 2021). The Australian Sports Pose Dataset (ASPset-510) was developed with this limitation in mind and consists of video clips with pseudo-3D pose annotations across a variety of dynamic sports-related actions in applied environments (Nibali et al., 2021). Updating pre-trained models with additional sports-specific data is another method to improve specificity (Kitamura et al., 2022). Regardless of what the dataset consists of, whether the model is ecologically valid for a given task won't be known until it assessed in the setting of interest, with the athletes of interest, and performing the tasks of interest.

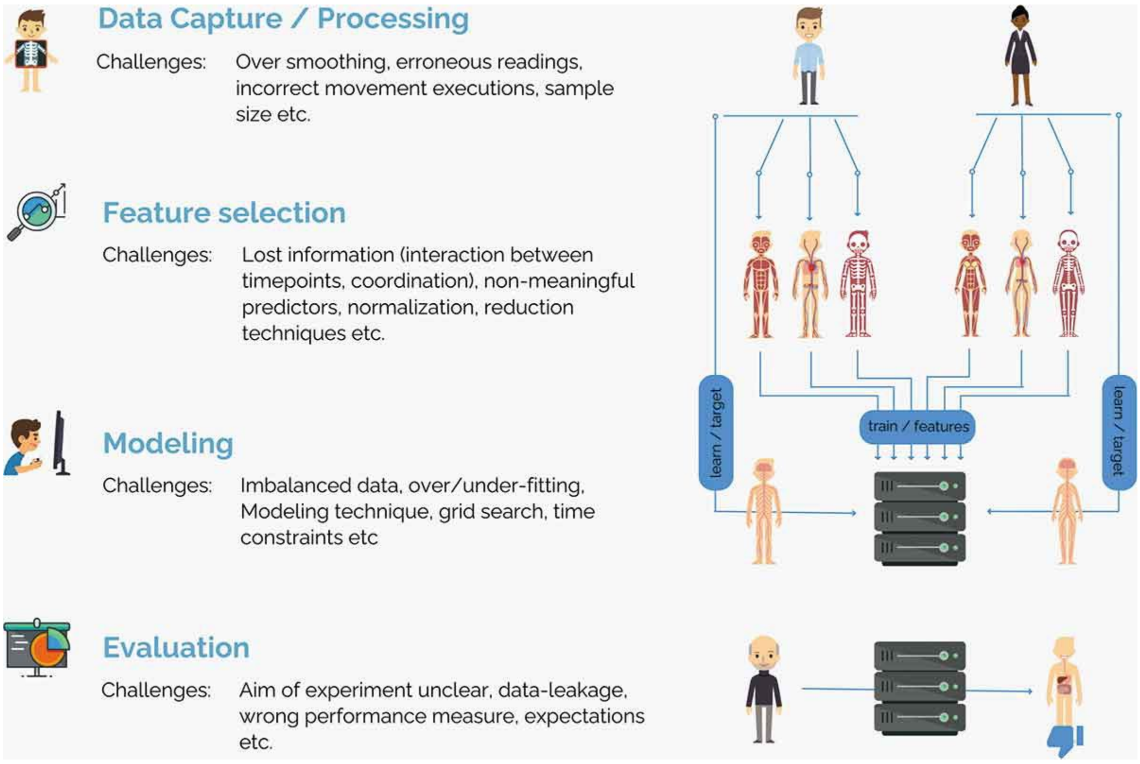

Figure 8. Challenges associated with feature-based machine learning. Reprinted from "Machine learning in sports science: challenges and opportunities" by C. Richter, 2021, Sports Biomechanics, p2.

Whilst bespoke, open source, and commercial VML approaches have been validated and published in the extant literature, their accuracy varies, and even sport-specific models will be influenced by environmental factors. For instance, if the model is trained on annotated images of athletes performing javelin throwing during night training sessions under artificial lighting, accuracy of this model under midday sun may be questionable. For the dataset to be specific it needs to include variability in environmental factors (e.g., day/night, rain, cloud, sun) and athletes. The actions of the sport will also need to be reflected in the dataset, for example, OpenPose was trained on people in normal orientations but was found to mislabel the legs as arms, and vice versa, during extreme poses such as a handstand (Figure 9) (Kitamura et al., 2022). For para-athletes, sport-specific VML models require development and specificity of these models to different classifications of athletes needs to be well considered. Models "can only estimate data well that is similar to those they are trained on" (Mundt, 2023, p. 17).

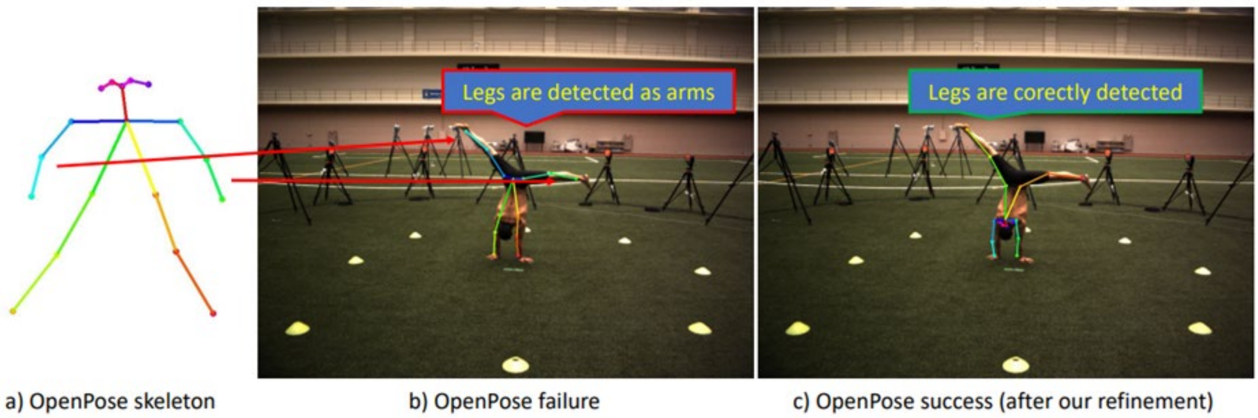

Figure 9. Exemplar of incorrect HPE (b) due to training on a non-specific dataset and subsequent correct HPE (c) with refinement of the model with sports-specific data. Reprinted from "Refining OpenPose with a new sports dataset for robust 2D pose estimation" by T. Kitamura, 2022, Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, p672.

Data capture techniques and the subsequent dataset quality is also dependent upon hardware specifications. Accuracy of point recognition models trained in DeepLabCut were compared with gold-standard 3D marker-based motion capture system in running (Van Hooren et al., 2023) and manual digitisation in swimming (Papic et al., 2021). No significant differences in sagittal plane hip angles were found with the gold-standard approach when running at 2.78ms^{-1}, but accuracy was reduced at 3.33ms^{-1} due to motion blur associated with a 50Hz camera capture rate (Van Hooren et al., 2023). Errors in glide kinematic outcomes were within the bounds of intra-rater manual digitisation (95%CIs of five repeated manual digitisation attempts on separate days), however, relative error increased at higher underwater glide velocities due to motion blur associated with the 30Hz camera capture rate and poor lighting conditions (Papic et al., 2021). As with manual digitising, when the target shape is obscured or blurred in video, point recognition models are susceptible to digitisation error (Zhang & Yang, 2021). To limit image distortion and ensure appropriate capture quality for a training dataset one must consider appropriate camera specifications for the given task. This should be reflected in the performance practitioner's skill set in following best practice for field-based motion analysis. As with any filming associated with quantitative analysis through any manual digitising process, optimal camera set-up is required to deliver the best possible outcome for a given VML methodology. The appropriate setting of shutter speed, as an example, will minimise any of the motion blur reported in the research presented above provided that the lighting conditions are appropriate. Likewise, the selection of an appropriate filming frame rate to complement the activity being analysed will ensure that key event detection is maximised. However, this is dependent on whether the ML algorithms have been trained on a sufficient frame rate to inform the final model.

For object recognition models targets of interest need to be clearly visible in most frames, and where obscured, appropriate error thresholds need to be selected to remove incorrectly labelled points. This can be automated in the processing phase to include interpolation of missing data, with a recognition that quality control and validation of the sports specific scenario being investigated will be crucial to ensure data hygiene is appropriate. Point recognition models are more appropriate for single camera 2D biomechanical analyses in applied settings, especially for actions that are primarily in a single plane of motion. In general, fixed camera setups with pre-calibrated capture areas can improve data capture quality and optimise accuracy of biomechanical outcomes derived from VML techniques within the limitations that have been discussed earlier in the report.

The majority of VML studies use single camera setups which reduce human setup requirements and computational effort. However, detailed features may be missed due to occlusion of landmarks and body segments. Multi-camera approaches for event recognition models and HPE provide multiple viewpoints and reduce occlusion but increase computational requirements (Cust et al., 2019; Zhang & Yang, 2021) and might not be feasible for applications in training and competition environments. There is little literature to inform optimal camera positions within sport specific analyses in a multi-camera VML approach and there will likely be a considerable process of optimisation in this regard to minimise resultant errors or reduced reliability. When considering temporary field-based testing scenarios, this has the potential to affect inter-testing validity and reliability and should be considered within any validation process. When developing a bespoke VML application a common barrier to VML implementation is access to appropriate hardware to train complex models. VML software require graphics processing units (GPUs) with an appropriate level of Video Random Access Memory (VRAM) and overall high computing power to train models on large datasets for many iterations, which can be costly and time consuming. Cloud-based computing power can be leveraged to train VML models using platforms such as GoogleColaboratory (i.e., virtual GPUs) but may present with challenges around regional data storage and privacy policies.

Best practice for model development also involves determining what features are selected in the dataset and how meaningful they are (Figure 8). Determining appropriate feature selection is more applicable to VML event recognition models. For models to be interpretable, a small selection of key features will be more beneficial than training a model on all outcomes as this can increase classification errors (Richter et al., 2021). For example, the previously mentioned soccer event detection model had greater accuracy for detecting four events (corner, free-kick, long view with no events, close-up view) compared with the addition of penalty and goal events into the model (Hong et al., 2018).

5.2 Ecological validity of models

As detailed throughout this section, a 'one size fits all' VML software will likely not be applicable for all practitioners and settings. The application of models in Australian HP sport will need to consider the usability of different VML models and include situational evaluations of their accuracy to ensure ecological validity. Evaluation of model accuracy compared with the ground truth should be conducted on performance and/or biomechanical outcomes of interest. To begin this process an 'acceptable error' threshold needs to be established based on the desired outcome. This threshold will require consultation between end users (practitioner, coach, and athletes) and is likely to be fluid. For example, accuracy of detecting headers in soccer using DeepImpact was 92.9% for a large number of assessed actions (Rezaei & Wu, 2022), which may be appropriate for injury monitoring. However, missing 7% of goal attempts in a hockey competition when reviewing game footage or 7% of derived shoulder angular velocities during cricket bowling with high relative error may be inappropriate for the desired outcome.

This reinforces the requirement of the performance support practitioner and the coach to have robust discussions on what the expectations are of the testing methodology sensitivities (irrespective of either the traditional testing protocol or the proposed VML approach) for the performance outcome results. This entails a further understanding of what represents a 'meaningful change' in terms of evaluating a training intervention and/or a 'discriminant ability' to be able to confidently discriminate between levels of expertise for a given critical variable. If, when conducting validation processes, the practitioner cannot be confident of providing results (through either traditional field-based analytics or a VML approach) to within the sensitivity required, then further refinement is necessary. For VML, this will likely involve a trial-and-error approach with evaluation of model accuracy and subsequent modification to i) the dataset; ii) modelling procedures; iii) calibration procedures; and/or iv) data processing techniques (e.g., reconstruction, interpolation, filtering, and adjustments).

Merely observing root-mean-square error of reconstructed landmark positions derived from a model doesn't consider accuracy across certain phases of the action, between athlete differences, and amplified effects of digitising error on biomechanical outcomes (Winter, 2009). Biomechanical analyses will often involve complex time-series data of an athlete (e.g., joint angles or velocities over a running cycle). Statistical Parametric Mapping (SPM) and Functional Data Analysis techniques such as functional principal components analysis (*f*PCA) are effective when evaluating biomechanical waveforms under different conditions and their use by sports science researchers and practitioners is increasing due to availability of opensource software (Warmenhoven et al., 2019; Warmenhoven et al., 2018). Practitioners could consider normalising biomechanical outcomes for a task of interest (e.g., athletics sprint start, rowing cycle, breaststroke pullout, triple jump) to time and using time-series statistical techniques to evaluate the ecological validity of a VML model. For example, Hooren et al. (2023) used SPM and time normalised RMSE analyses to evaluate the accuracy of OpenPose compared with 3D motion capture (Figure 10). This approach allowed the researchers to observe possible limitations of the model during certain phases of the running cycle and whether the model could be generalised to all participants. No statistical differences were found in lower limb joint angles between the two approaches at this running velocity, however, there was greater variance in ankle joint angles between methods and errors were higher for some participants for certain joint angles (e.g., see the individual light-blue line in the knee angle RMSE figure below). In practice, if these errors were above the pre-established 'acceptable error' threshold it may be decided that the model was not generalisable to several of the athletes.

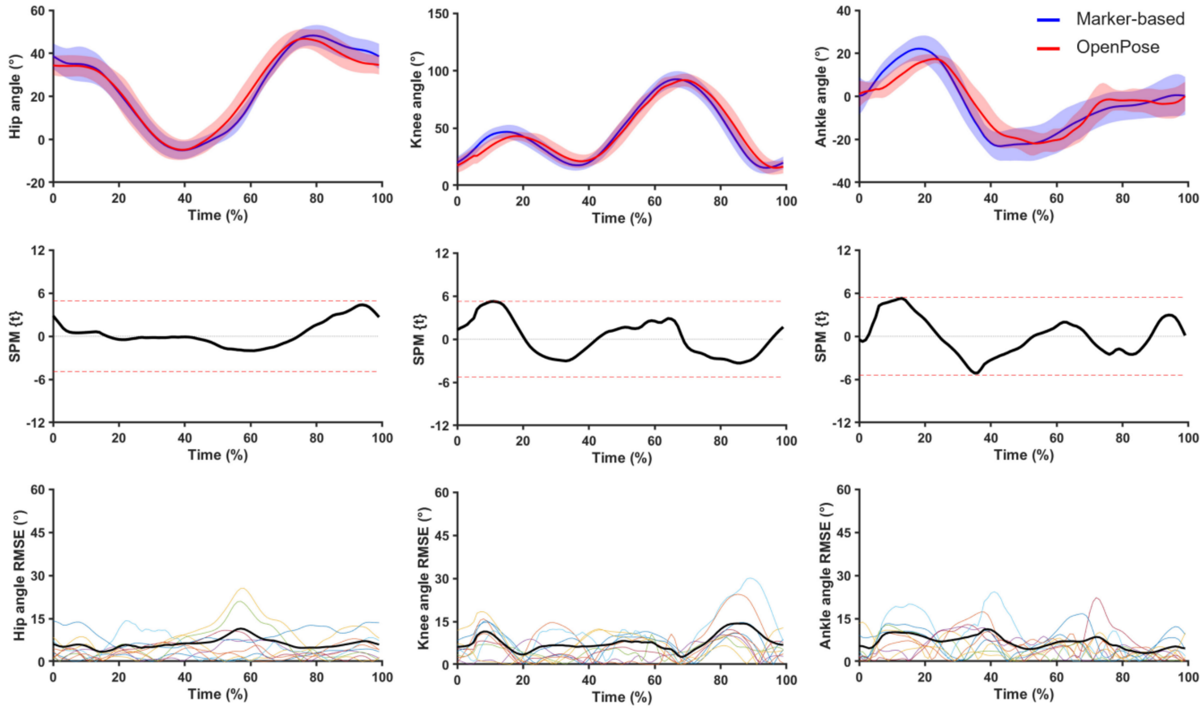

Figure 10. Exemplar of model validation procedures (OpenPose vs marker-based 3D motion capture). Top: marker-based (blue) and marker-less motion capture (OpenPose, red) sagittal-plane joint angles at 2.78ms $^{-1}$ for the hip (left column), knee (middle column), and ankle (right column). Middle: corresponding paired samples t-test statistic from statistical parameter mapping in relation to the critical threshold (red dashed line). Bottom: root-squared error of the difference between marker-based and marker-less (OpenPose) motion capture for each joint. Coloured lines represent individual root squared errors; the thick bold line represents the root-mean squared error averaged over all participants per time point. Reprinted from "The accuracy of marker-less motion capture combined with computer vision techniques for measuring running kinematics" by B. V. Hooren, 2023, Scandinavian Journal of Medicine & Science In Sports, 33, p972.

Despite these challenges we are moving into a world where VML software is proliferating, and sport-specific datasets will continue to be developed and made available. This has been reflected by the rapid uptake on VML processes within technology applications provided by commercial companies. While outside of the scope of this Insight Report, there are other significant aspects of VML technologies that can apply to the athlete's daily performance environment (DPE) including areas like 'gamification' which can help drive athlete engagement in training sessions if structured well. From a scientific point of view though, many of these commercial products demonstrate little or no validation processes that can be applied to provide a level of quality assurance.

The prospect of accurate and feasible VML technologies to provide timely feedback to athletes and coaches and inform evidence-based decision making is enticing for HP sport. Given the wide array of VML approaches, bespoke software that allows input of pose and calibration data from multiple VML software and user-selected outcomes of interest would be advantageous. However, until this time the Australian HP sport system will need to leverage the wide array of expertise across practitioners and researchers to collaborate on the development and implementation of VML technology. This will require transparency across the network, including an understanding of what is currently being performed, what are the limitations of these applications, and data sharing.

6. VML insights survey

There were 36 performance support practitioners across a range of NIN and sport specific affiliations (e.g., AusCycling, Swimming Australian, Paralympics Australia, Paddle Australia), predominantly within the Movement Science disciplines, that voluntarily completed the VML insights survey (Appendix A). Almost all responders (97%) conduct video analysis with athletes they work with, across 16 listed individual and team aquatic, track, and field-based sports. Across these practitioners, the top three sports they were involved in were swimming, cycling, and rowing. Video analyses were performed both in training and competition environments, in indoor and outdoor venues. The prevalence of video analysis applications utilised by the practitioners is summarised in Figure 111. Most practitioners used video analysis to provide in-session and post-session qualitative feedback (training or competition), and quantitative feedback post-session. The most frequent response from practitioners included all three of these application types. Only one in four practitioners provided quantitative feedback to athletes and coaches within-session, highlighting the potential impact that VML technology may have.

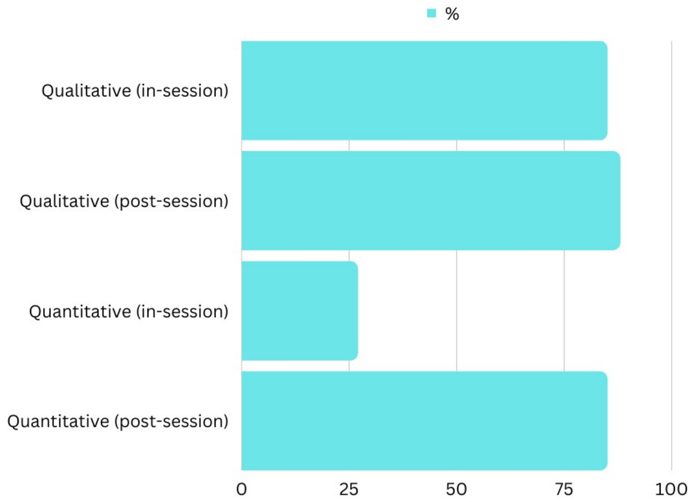

Figure 11. Prevalence of video analysis applications by practitioners in high performance sport

Of the 50% of practitioners who digitise videos to derive kinematic outcomes, Kinovea was the most used software (78%), with a mix of other software used by a small proportion of practitioners: CLOGGS, Tracker, Dartfish, SiliconCoach, VidMark, KPASS, and other bespoke software. For the remaining 50% of responders, barriers to using video analysis techniques included time/labour constraints, access to resources, and expertise (Figure 12).

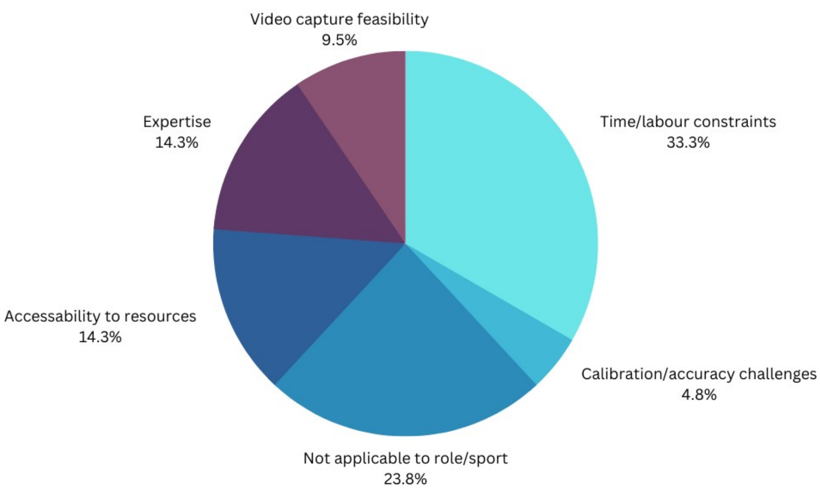

Figure 12. Practitioner-defined barriers to digitising video data.

6.1 Video machine learning use

Practitioners had a good understanding of what VML entails and the potential benefits of this technology for sport. Responses centred around several key themes illustrated in Figure 13.

Figure 13. Word cloud of key themes detailed in response to "briefly describe what video machine learning means to you?"

Time efficient video processing, with implications for feedback to athletes, was a frequent theme associated with VML techniques in the responses.

> Practitioner responses
>
> "The ability to enhance video capture and analysis processes which allows more time for the practitioner to be present to interpret the data and provide feedback vs time spent analysing and digitising footage."
>
> "VML utilises artificial intelligence to automate, augment, and/or estimate video digitisation and analysis processes. It has the potential to facilitate capture and rapid analysis of athlete biomechanics in sporting environments without encumbering the athlete."
>
> "Machine learning means automating simple but long and laborious processes so that Sport Scientists can deliver useful data more quickly and focus their time on the analysis and interpretation of the data."

It was also highlighted that there is a need for representative datasets to derive valid kinematic data (a concept that was discussed in this report as a potential barrier for practitioners when implementing VML methods).

> Practitioner response
>
> "Using a collection of videos of the same athletic poses or movements as inputs in order to train an algorithm (or multiple algorithms) to track and process kinematic data related to the overall performance."

Only 28% of the practitioners currently use and/or are testing VML technologies. These practitioners are implementing three main technologies as summarised in Figure 14. Use of whole-body pose-estimation software appears to be limited across the network (e.g., Theia3D), rather, practitioners are implementing software to automate specific time events (e.g., swimming stroke cycles) and point recognition of key body landmarks (e.g., head, hip, knee, ankle).

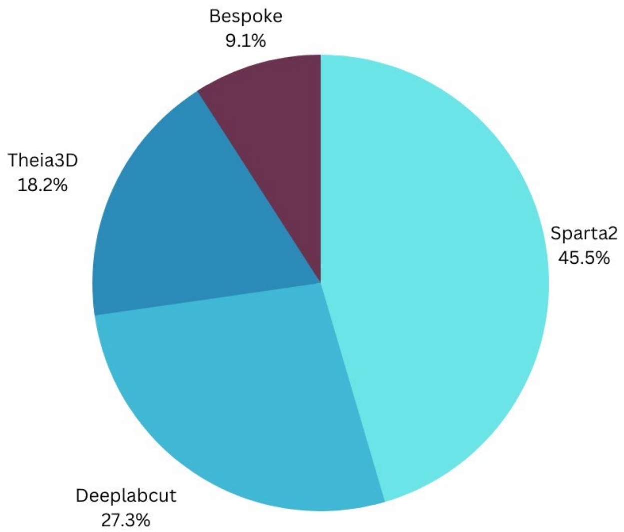

Figure 14. Prevalence of software use across practitioners utilising video machine learning in sport.

Established VML protocols across the network are limited. Those that are established centre on guidance for setup and calibration, as well as ongoing analysis and data validation procedures. Examples of protocols that are followed include:

- Bespoke SPARTA2 protocol within the swimming sports science network which includes standards of calibration, waterline correction protocols, and rules regarding the manual addition of kick counts and breakouts to analysis.
- Bespoke Theia3D calibration and setup protocol for multicamera analysis – it was indicated that specific documentation around naming conventions and post processing methods is required.
- Elipot, M. (2019). A new paradigm to do and understand the race analyses in swimming: The application of convolutional neural networks. ISBS Proceedings Archive, 37(1), 455.
- Hall, A., Victor, B., He, Z., Langer, M., Elipot, M., Nibali, A., & Morgan, S. (2021). The detection, tracking, and temporal action localisation of swimmers for automated analysis. Neural Computing and Applications, 33, 7205-7223.

6.2 Practitioner-defined challenges to VML

One distinct challenge for VML is that access to validated sport-specific models is limited and these models are only optimised for data they have been trained on, as discussed earlier in this report. Consequently, an 'accurate' model for a given task and environment doesn't necessarily mean that it will be robust and translatable to other athletes in a similar, but different, setting.

Practitioner responses

“Limited research on the validity and reliability of the kinematic data that is produced using machine learning systems in sport.”

“…substantial libraries of manually digitised sport data are required for sport-specific VML applications.”

“The current limits of the video learning sets being used to train the algorithms (i.e. lack of video on certain sports and/or of para-athletes).”

Due to the fast-paced demands of keeping up with the artificial intelligence race, there are concerns around implementing VML software before taking the necessary steps to evaluate its ecological validity for each purpose. This is especially the case in field-based testing scenarios with constantly changing environments with new calibrations needing to be continually performed, begging the question of their practicality. Not understanding the limitations and accuracy of models may lead to data misuse and misinterpretation. This was a prevalent theme in practitioner responses. Personal knowledge and technical expertise of VML methods was another practitioner-defined challenge to implementing VML. Evaluating validity and reliability of these systems will require time, practitioner training, technical proficiency, and adequate communication between members of the HP sport system.

Practitioner responses

“Too high a reliance placed on machine learning and automation means that little things may get missed and if you simply trust the process and trust the numbers or automation, things may not be interpreted in the right manner.”

“Risk of potential misuse by coaches, athletes or practitioners who do not understand how the data is generated and its associated accuracy and reliability.”

Appropriate “buy in” from all stakeholders needs to be considered if VML technology is going to be effectively implemented in Australian HP sport. Proverbially speaking, VML false starts with implementation will need to occur in training environments (and be clearly articulated as such) rather that in the real race. Failure to do so increases the risk of practitioners, coaches, and athletes not trusting the technology going forward. Consultation with end-users should occur at all stages of testing and implementation of this technology.

Practitioner responses

“Coach understanding and trust of new technology.”

“…coach-led support for and capacity to implement and drive engagement.”

“…can only deploy when fully ready for the athletes / coaches.”

Practitioner responses reflected common challenges associated with VML which were consistent with those described in Section 7. Despite these challenges, 77% of the practitioners have commenced or have expressed desire for implementing VML technology, highlighting that “if the technology can reproduce the same level of accuracy as humans the possibilities for scaling data capturing is much greater than humans” (practitioner response). Current plans centre around sport test cases using 2D fixed setups to expedite kinematic data. Practitioners that are interested in exploring this technology further noted that individual training and access to transparent guidelines on developing appropriate VML models is needed. Furthermore, resources should include time effective workflows for video capture, analysis, and athlete feedback.

7. Summary of VML Insights Report Recommendations

1. VML technology holds significant implications for expediting analytics but requires situational context and validity assessment to maximise effectiveness in the DPE.

- VML technology can be used for event and action recognition, point recognition (e.g., landmarks, objects), shape analysis and HPE to expedite analysis but practitioners will need to ensure the VML technology is 'fit-for-purpose'. Practitioners will need to assess and validate suitability based on required environments, performance outcomes and measurement sensitivities.

2. "Quality-in equals quality-out": Ecological validity of VML models is dependent on data capture quality and specificity of training datasets.

- Camera specifications, setup and following best practice field-based filming processes need to be considered when capturing data to minimise image blurring and occlusion of the desired task and/or landmarks.
- Fixed camera setups with pre-calibrated spaces are recommended as they have the potential to optimise accuracy and feasibility of VML technology within the limitations and constraints that are inherent to the technology and specific VML processes.
- Datasets need to include variability in environments, athletes, and actions that are translatable to the desired sport and outcomes of interest for all types of VML technologies.
- Para-athlete VML datasets are an area for ongoing development and specificity of these models to different classifications of athletes needs to be well considered.
- Pre-annotated datasets (e.g. 3D motion capture) can be leveraged to train VML models for a distinct task provided they are task specific.
- The development of machine learning algorithms to increase the accuracy of key event detection is crucial if an automated VML approach is to be incorporated.

3. Access to VML expertise will be required as bespoke modelling methods are not going to be feasible for all practitioners to implement.

- Refinement of pre-existing models with sport-specific data can improve model accuracy for a given task, however, these methods require specialised expertise in modelling techniques and will need to be considered when planning to integrate VML into the DPE.
- Collaboration between researchers and practitioners with multidisciplinary expertise could be used to support the development of bespoke models for certain sports.

4. Using opensource and commercially available VML software is a feasible approach for implementing this technology. However, a 'one size fits all' VML software will not be applicable to all situtations in Australian HP sport.

- Development of activity and environment specific trained models are likely to be required to ensure higher fidelity data output compared to generalised models.
- Opensource and commercially available VML software using generalised models will be easier to implement than bespoke trained models, however, their translatability to different situations will vary (see recommendation 5).

5. Ecological validity of VML models need to be established before implementation with end users.

- When evaluating ecological validity, an 'acceptable error' threshold needs to be established based on the desired outcome. This threshold will require consultation between end users (practitioners, coach, and athletes) and is likely to be fluid and analysis dependent.
- A 'trial-and-error' approach will need to occur in testing phases of VML technology to achieve accuracy within the desired threshold. This will require evaluation of model accuracy and subsequent modification to i) the dataset; ii) modelling procedures; iii) calibration procedures; and/or iv) data processing techniques (e.g., reconstruction, interpolation, filtering, and adjustments).
- Ecological validity should be evaluated on the derived outcomes of interest, rather than just simple accuracy measures (eg. RMSE of estimated joint centres). If time-series biomechanical outcomes are of interest, statistical approaches such as FDA and SPM can be used to effectively evaluate validity of the model.

6. The development of user-friendly software that allows input of pose and calibration data from multiple VML software and user-selected analyses.

- Software that allows input of coordinate and calibration data from multiple VML software and user-selection of data processing, analysis, and outcomes of interest would be advantageous for feasibility of this technology in Australian HP sport.

7. Practitioner professional development and ongoing technological support by experts is required for long-term implementation of this technology.

- Challenges will arise when implementing this technology and practitioners will require technical support to address them. Time constraints will be a significant factor that will limit practitioners seeking out solutions to these challenges independently. It is recommended that these areas be addressed as these may pose as a significant barrier to implementing VML.

8. For successful implementation, communication and transparency is required.

- Until VML technology is accurate in all required sports scenarios and easy to implement, it is recommended that the Australian HP sport system utilise the wide array of expertise across practitioners and researchers to collaborate on the development and implementation of VML technology.
- It is recommended that communication and transparency in the use of VML across the network is highlighted as a key action. This should include an understanding of what is currently being performed, what are the limitations of these applications, and appropriate sharing of data.

- Communication will need to include collaboration with end users at all stages of developing and implementing this technology to maximise the outcomes of the use of this technology.

9. VML best practice guidelines are required.

- These guidelines will aim to provide contemporary evidence-based recommendations around various VML software and workflow processes for performance support practitioners to implement VML. These should also include best practice guidelines for optimising data collection and interpretation, with a recognition that these guidelines will have situational context.
- It is recommended that consideration be given to support practitioner training related to the implementation of these best practice guidelines.

8. References

Badiola-Bengoa, A., & Mendez-Zorrilla, A. (2021). A systematic review of the application of camera-based human pose estimation in the field of sport and physical exercise. *Sensors, 21*(18), 5996.

Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). Blazepose: On-device real-time body pose tracking. *arXiv preprint arXiv:2006.10204*.

Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. Proceedings of the IEEE conference on computer vision and pattern recognition,

Cust, E. E., Sweeting, A. J., Ball, K., & Robertson, S. (2019). Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance. *Journal of sports sciences, 37*(5), 568-600.

Di Paolo, S., Nijmeijer, E., Bragonzoni, L., Dingshoff, E., Gokeler, A., & Benjaminse, A. (2023). Comparing lab and field agility kinematics in young talented female football players: Implications for ACL injury prevention. *European Journal of Sport Science, 23*(5), 859-868.

Ding, N., Takeda, K., & Fujii, K. (2022). Deep reinforcement learning in a racket sport for player evaluation with technical and tactical contexts. *IEEE access, 10*, 54764-54772.

Dubey, S., & Dixit, M. (2023). A comprehensive survey on human pose estimation approaches. *Multimedia Systems, 29*(1), 167-195.

Elipot, M. (2019). A new paradigm to do and understand the race analyses in swimming: The application of convolutional neural networks. *ISBS Proceedings Archive, 37*(1), 455.

Elliott, B. C., Alderson, J. A., & Denver, E. R. (2007). System and modelling errors in motion analysis: Implications for the measurement of the elbow angle in cricket bowling. *Journal of Biomechanics, 40*(12), 2679-2685.

Fang, H.-S., Li, J., Tang, H., Xu, C., Zhu, H., Xiu, Y., Li, Y.-L., & Lu, C. (2022). Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

FIFA. (2022). *Handbook of Test Methods for EPTS devices Version 3*. Fédération Internationale de Football Association.

Gourgari, S., Goudelis, G., Karpouzis, K., & Kollias, S. (2013). Thetis: Three dimensional tennis shots a human action dataset. Proceedings of the IEEE conference on computer vision and pattern recognition workshops,

Hall, A., Victor, B., He, Z., Langer, M., Elipot, M., Nibali, A., & Morgan, S. (2021). The detection, tracking, and temporal action localisation of swimmers for automated analysis. *Neural Computing and Applications, 33*, 7205-7223.

Hong, Y., Ling, C., & Ye, Z. (2018). End-to-end soccer video scene and event classification with deep transfer learning. 2018 International Conference on Intelligent Systems and Computer Vision (ISCV),

Insafutdinov, E., Pishchulin, L., Andres, B., Andriluka, M., & Schiele, B. (2016). Deepercut: A deeper, stronger, and faster multi-person pose estimation model. Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VI 14,

Ito, N., Sigurðsson, H. B., Seymore, K. D., Arhos, E. K., Buchanan, T. S., Snyder-Mackler, L., & Grävare Silbernagel, K. (2022, 2022/10/01/). Markerless motion capture: What clinician-scientists need to know right now. *JSAMS Plus, 1*, 100001. https://doi.org/https://doi.org/10.1016/j.jsampl.2022.100001

Jiang, Y., Cui, K., Chen, L., Wang, C., & Xu, C. (2020). Soccerdb: A large-scale database for comprehensive video understanding. Proceedings of the 3rd International Workshop on Multimedia Content Analysis in Sports,

Kipp, K., Giordanelli, M., & Geiser, C. (2018). Predicting net joint moments during a weightlifting exercise with a neural network model. *Journal of Biomechanics, 74*, 225-229.

Kitamura, T., Teshima, H., Thomas, D., & Kawasaki, H. (2022). Refining OpenPose with a new sports dataset for robust 2D pose estimation. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision,

Köykkä, M., Vohlakari, K., Joutsen, T., Vänttinen, T., Piironen, P., Rantalainen, T., & Cronin, N. (2022). Markerless video-based estimation of 3D approach velocity in the javelin throw. *ISBS Proceedings Archive, 40*(1), 355.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13,

Lipton, M. L., Kim, N., Zimmerman, M. E., Kim, M., Stewart, W. F., Branch, C. A., & Lipton, R. B. (2013). Soccer heading is associated with white matter microstructural and cognitive abnormalities. *Radiology, 268*(3), 850-857.

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.

Mundt, M. (2023). Bridging the lab-to-field gap using machine learning: a narrative review. *Sports Biomechanics*, 1-20.

Mundt, M., Koeppe, A., David, S., Bamer, F., Potthast, W., & Markert, B. (2020). Prediction of ground reaction force and joint moments based on optical motion capture data during gait. *Medical Engineering & Physics, 86*, 29-34.

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.

Nibali, A., Millward, J., He, Z., & Morgan, S. (2021). ASPset: An outdoor sports pose video dataset with 3D keypoint annotations. *Image and Vision Computing, 111*, 104196.

Papic, C., Sanders, R. H., Naemi, R., Elipot, M., & Andersen, J. (2021). Improving data acquisition speed and accuracy in sport using neural networks. *Journal of sports sciences, 39*(5), 513-522.

Patel, P., Huang, C.-H. P., Tesch, J., Hoffmann, D. T., Tripathi, S., & Black, M. J. (2021). AGORA: Avatars in geography optimized for regression analysis. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,

Rangasamy, K., As'ari, M. A., Rahmad, N. A., Ghazali, N. F., & Ismail, S. (2020). Deep learning in sport video analysis: a review. *TELKOMNIKA (Telecommunication Computing Electronics and Control), 18*(4), 1926-1933.

Rapczyński, M., Werner, P., Handrich, S., & Al-Hamadi, A. (2021). A baseline for cross-database 3d human pose estimation. *Sensors, 21*(11), 3769.

Rezaei, A., & Wu, L. C. (2022, 2022/06/03). Automated soccer head impact exposure tracking using video and deep learning. *Scientific Reports, 12*(1), 9282. https://doi.org/10.1038/s41598-022-13220-2

Richter, C., O'Reilly, M., & Delahunt, E. (2021). Machine learning in sports science: challenges and opportunities. *Sports Biomechanics*, 1-7.

Rossi, A., Pappalardo, L., Cintia, P., Iaia, F. M., Fernández, J., & Medina, D. (2018). Effective injury forecasting in soccer with GPS training data and machine learning. *PloS one, 13*(7), e0201264.

Saito, S., Huang, Z., Natsume, R., Morishima, S., Kanazawa, A., & Li, H. (2019). Pifu: Pixel-aligned implicit function for high-resolution clothed human digitization. Proceedings of the IEEE/CVF international conference on computer vision,

Serrancolí, G., Bogatikov, P., Huix, J. P., Barberà, A. F., Egea, A. J. S., Ribé, J. T., Kanaan-Izquierdo, S., & Susín, A. (2020). Marker-less monitoring protocol to analyze biomechanical joint metrics during pedaling. *IEEE access, 8*, 122782-122790.

Tsagkatakis, G., Jaber, M., & Tsakalides, P. (2017). Goal!! event detection in sports video. *Electronic Imaging, 2017*(16), 15-20.

Uhlrich, S. D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A. S., Hicks, J. L., & Delp, S. L. (2022). OpenCap: 3D human movement dynamics from smartphone videos. *BioRxiv*, 2022.2007. 2007.499061.

Uhlrich, S. D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A. S., Hicks, J. L., & Delp, S. L. (2023). OpenCap: Human movement dynamics from smartphone videos. *PLOS Computational Biology, 19*(10), e1011462.

Van Hooren, B., Pecasse, N., Meijer, K., & Essers, J. M. N. (2023). The accuracy of markerless motion capture combined with computer vision techniques for measuring running kinematics. *Scandinavian Journal of Medicine & Science in Sports, 33*(6), 966-978.

Vats, K., Fani, M., Clausi, D. A., & Zelek, J. S. (2022). Evaluating deep tracking models for player tracking in broadcast ice hockey video. *arXiv preprint arXiv:2205.10949*.

Vinyes Mora, S., & Knottenbelt, W. J. (2017). Deep learning for domain-specific action recognition in tennis. Proceedings of the IEEE conference on computer vision and pattern recognition workshops,

Wang, J., Tan, S., Zhen, X., Xu, S., Zheng, F., He, Z., & Shao, L. (2021). Deep 3D human pose estimation: A review. *Computer Vision and Image Understanding, 210*, 103225.

Warmenhoven, J., Cobley, S., Draper, C., Harrison, A., Bargary, N., & Smith, R. (2019, 2019/05/04). Considerations for the use of functional principal components analysis in sports biomechanics: examples from on-water rowing. *Sports Biomechanics, 18*(3), 317-341. https://doi.org/10.1080/14763141.2017.1392594

Warmenhoven, J., Harrison, A., Robinson, M. A., Vanrenterghem, J., Bargary, N., Smith, R., Cobley, S., Draper, C., Donnelly, C., & Pataky, T. (2018). A force profile analysis comparison between functional data analysis, statistical parametric mapping and statistical non-parametric mapping in on-water single sculling. *Journal of Science and Medicine in Sport, 21*(10), 1100-1105.

Winter, D. A. (2009). *Biomechanics and motor control of human movement*. John wiley & sons.

Xiu, Y., Yang, J., Tzionas, D., & Black, M. J. (2022). Icon: Implicit clothed humans obtained from normals. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018, 2018/08/01). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging, 9*(4), 611-629. https://doi.org/10.1007/s13244-018-0639-9

Zhang, X., & Yang, F.-Q. (2021, 2021/03/22). Machine Learning-Based Multitarget Tracking of Motion in Sports Video. *Complexity, 2021*, 5533884. https://doi.org/10.1155/2021/5533884

Zheng, C., Wu, W., Chen, C., Yang, T., Zhu, S., Shen, J., Kehtarnavaz, N., & Shah, M. (2023). Deep learning-based human pose estimation: A survey. *ACM Computing Surveys, 56*(1), 1-37.

Appendix A: VML insights survey questions

1. Do you currently use any form of video analysis with the athletes or sports you work with?

   If yes, please detail the sports you currently use video analysis with and the settings in which you use them in (e.g., gym, indoor pool, outdoor field):

2. What are your applications of video analysis techniques?

   Options: Qualitative analysis/feedback within-session (e.g., technical feedback); Qualitative analysis/feedback (post-session/later date); Quantitative analysis/feedback within-session (e.g., kinematic variables); Quantitative analysis/feedback post-session/later date.

3. Do you digitise video of your athletes to perform kinematic analyses.

   If yes: please detail what software you use to perform digitisation and analyses in.

   If no: please detail any limiting factors for digitising video and performing subsequent analyses.

4. Briefly describe what video machine learning means to you.

5. Do you currently use video machine learning technology for sporting analysis?

   Please elaborate on the software you use and the sports you currently use this technology with.

6. Do you have established protocols for your video machine learning testing - either guidelines from published research or self-established protocols?

   If yes: please provide further information.

7. Are there plans to work with video machine learning technology in the near future for sporting analysis?

   Please elaborate on your plans.

8. If you do not currently work with video machine learning software would you be interested in exploring how these tools could be applicable to you?

9. In your opinion and/or experience, are there any barriers to using VML in high-performance sport.

Australian Government
Australian Sports Commission

AIS

ausport.gov.au

Leverrier Street Bruce ACT 2617
PO Box 176 Belconnen ACT 2616
+61 2 6214 1111