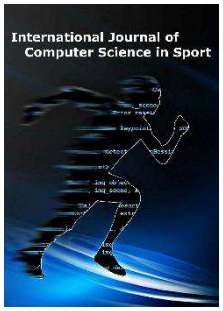
International Journal of Computer Science in Sport

Volume 23, Issue 1, 2024

Journal homepage: http://iacss.org/index.php?id=30

sciendo

DOI: 10.2478/ijcss-2024-0007



The application of Machine and Deep Learning for technique and skill analysis in swing and team sport-specific movement: A systematic review

Chloe Leddy[1], Richard Bolger[2], Paul J. Byrne[3], Sharon Kinsella[3], Lilibeth Zambrano[1]

[1]South East Technological University (Kilkenny Road Campus), Department of Aerospace & Mechanical Engineering, Kilkenny Rd, Carlow;

[2]South East Technological University (Cork Road Campus), Department of Sport & Exercise Science, Waterford, Ireland;

[3]South East Technological University (Kilkenny Road Campus), Department of Health & Sport Sciences, Kilkenny Road, Carlow, Ireland

Abstract

There is an ever-present need to objectively measure and analyze sports motion for the determination of correct patterns of motion for skill execution. Developments in performance analysis technologies such as inertial measuring units (IMUs) have resulted in enormous data generation. However, these advances present challenges in analysis, interpretation, and transformation of data into useful information. Artificial intelligence (AI) systems can process and analyze large amounts of data quickly and efficiently through classification techniques. This study aimed to systematically review the literature on Machine Learning (ML) and Deep Learning (DL) methods applied to IMU data inputs for evaluating techniques or skills in individual swing and team sports. Electronic database searches (IEEE Xplore, PubMed, Scopus, and Google Scholar) were conducted and aligned with the PRISMA statement and guidelines. A total of 26 articles were included in the review. The Support Vector Machine (SVM) was identified as the most utilized model, as per 7 studies. A deep learning approach was reported in 6 studies, in the form of a Convolutional Neural Network (CNN) architecture. The in-depth analysis highlighted varying methodologies across all sports inclusive of device specifications, data preprocessing techniques and model performance evaluation. This review highlights that each step of the ML modeling process is iterative and should be based on the specific characteristics of the movement being analyzed.

KEYWORDS: INERTIAL SENSORS, MACHINE LEARNING, NEURAL NETWOKS, PERFORMANCE ANALYSIS

Introduction

Observation and analysis of sport motion are crucial in understanding the proficiency of technical aspects within the sport. Objective assessment of athletic performance is conducted by means of tracking kinematic, temporal, and dynamic parameters of movement (Yeo & Park, 2020). The application of artificial intelligence (AI) to sports motion data has been shown to be beneficial in the realm of performance assessment and analysis (Dwyer, Kempe, & Knobbe, 2022). The use of AI techniques in sports motion analysis refers to the recognition and classification of patterns and skills (Bonidia et al., 2018). With sufficient data, machine learning (ML) models can identify patterns in performance data leading to objective and standardized performance evaluations (Vrigkas, Nikou, & Kakadiaris, 2015) and predictions of emergent behavior (Bonidia et al., 2018). These automated processes can assist coaches in making complex decisions for targeted training and injury prevention strategies (Richter, O'Reilly & Delahunt, 2021). For sport technique analysis and performance prediction conducted in a field-based context, inertial measuring unit (IMU) sensors are commonly used as inputs for the recognition and classification of complex sport actions (Alanen, Räisänen, Benson, & Pasanen, 2021; Camomilla, Bergamini, Fantozzi, & Vannozzi, 2018). Sport motion analysis conducted by IMU systems results in substantial amounts of data that may not be explicitly labeled according to the detected activities and generally contains noise and fluctuations in the signal. This data requires labor-intensive cleaning and filtering to remove noise and artifacts (Moran et al., 2015), a process that may be prone to human error. The benefits of AI for biomechanical data management are focused on the ability of ML and DL models to handle large amounts of data (Aresta et al., 2022), and efficiently classify data to find patterns of information contributing to the enhanced knowledge about an athlete's performance parameters and predisposition to injury (Claudino et al., 2019; Van Eetvelde, Mendonça, Ley, Seil, & Tischer, 2021).

Performance analysis tests in sports should exhibit robustness in terms of validity, reliability, and sensitivity, while the protocol should simulate the actual sporting performance as closely as possible (Currell & Jeukendrup, 2008). As such, the ecological validity of sports performance research can be enhanced using markerless motion capture (Baca, Dabnichki, Hu, Kornfeind, & Exel, 2022). Traditional marker-based performance analysis methods are the preferred method of biomechanical analysis, but there are limitations associated with gold-standard laboratory-based methods (Colyer, Evans, Cosker, & Salo, 2018). The gathering of data that requires specific equipment and expertise makes this method of movement assessment less portable and accessible outside of a biomechanics laboratory (Roggio et al., 2021). Athletic performance assessment occurring on the sports field ensures data analytics accurately represent the player's technique accurately and the normal standards that are tailored to the athletes' requirements. Continuous and automated classification of sports motion in situ is achievable using body-mounted IMUs (Ahmadi et al., 2014; Preatoni, Nodari, & Lopomo, 2020), coupled with the application of an ML or DL model. Sport movement analysis conducted by means of body-mounted IMUs spans a variety of applications such as quantifying skill level (Wang, Zhao, Chan, & Li, 2018), analyzing technique (Kim & Kim, 2017) and detecting and categorizing movement (Salman, Qaisar, & Qamar, 2017). Wireless IMUs contain a combination of accelerometers, gyroscopes and more commonly the addition of magnetometers (measuring magnetic field) (Ahmad, Ghazilla, Khairi, & Kasi, 2013). By fusing the information measuring accelerometer and gyroscope data from these sensors and converting them to orientation and angular velocity, valuable information regarding movement can be obtained (Kok, Hol, & Schön, 2017).

Machine learning (ML) is a subdivision of artificial intelligence (AI) that focuses on the ability of computers to improve and learn from the data that they are fed (Alzubi, Nayyar, & Kumar, 2018). A recent systematic review gives a comprehensive overview and description of AI techniques and methods that are commonly utilized in the prediction of performance in team

sports and reported that artificial neural networks were most utilized (Claudino et al., 2019). This is analogous to the systematic review by Cust, Sweeting, Ball, and Robertson (2019), reporting that Convolutional Neural Network (CNN) and Support Vector Machine (SVM) were the most utilized algorithms for sport-specific movement recognition. Deep Neural Network (DNN) model architectures are built on multiple layers of representations, and they have revolutionized many areas of machine intelligence, enabling superhuman accuracy for challenging time series classification and image recognition tasks (Mathis, Schneider, Lauer, & Mathis, 2020). There is continuous nonlinear transformation performed at each layer, decomposing objects into key points with semantic meaning leading to very complex representations of the data. DNNs capable of high-level representations in the deep layer automatically generate optimal features from input data and identify hidden patterns in the datasets, and thus have become popular in the field of activity recognition (Chmait & Westerbeck, 2021; Wang, Chen, Hao, Peng, & Hu, 2019). The advantage of CNNs over other classification algorithms is that a CNN model learns the best features to represent the objects in the images (Chai, Zeng, Li, & Ngai, 2021). Hence, DL models built with a robust CNN architecture can automatically extract kinematic parameters and detect sporting events from spatiotemporal visual semantic data (Zecha, Einfalt, Eggert, & Lienhart, 2018). However, there are challenges associated with utilizing CNN architectures for feature extraction and classification in sport motion analysis. A CNN model requires an extensively large database of labeled instances, as compared to ML counterparts, and accurately labelling large datasets, particularly in video instances is a considerable challenge (Islam, Nooruddin, Karray, & Muhammad, 2022). Sport motion analysis often requires datasets to be split into distinct classes, in instances where different categorical variables are the output variable such as predicting performance based on skill level (Wang et al., 2018). SVM performs well in high-dimensional datasets and can discriminate features well and classify categorical labels into distinct classes (Angerbauer, Palmanshofer, Selinger, & Kurz, 2021). However, SVM may struggle to generalize new scenarios and requires large amounts of labeled data to train.

Human activity recognition (HAR) is concerned with the automatic recognition of specific movements or actions and has significantly contributed to the field of sports movement detection and recognition (Host & Ivašić-Kos, 2022). HAR is a challenging time-series classification task that involves a framework of data acquisition, pre-processing, model training, and model performance evaluation (Gupta et al., 2022). HAR systems based on sensing technologies have two major components in the pipeline which are feature extraction and classification (Ann & Theng, 2014). The two approaches are commonly coupled with ML techniques, where the spatiotemporal parameters of motion are automatically predicted using a set of features extracted from the IMU signals (Zhou et al., 2020). Pre-processing techniques for time-series data such as filtering, frequency analysis and machine learning have been applied to maintain the shape of the signal and to increase detection accuracy (Jalal, Quaid, Tahir, & Kim, 2020; Worsey, Espinosa, Shepherd, & Thiel, 2019). Removal of interference noise or miscellaneous signal fluctuations and low-frequency components can be achieved by applying low-pass and high-pass filters such as the Butterworth filter (Yin, Chen, Miao, Jiang, & Chen, 2021) and the Kalman Filter (Batool, Jalal, & Kim, 2020). Feature extraction involves reducing the dimensionality of the dataset by selecting the most relevant features for increased model accuracy and data representation (Zhu, San-Segundo, & Pardo, 2017). The most common methods of feature extraction in ML algorithms for human movement biomechanics include Principal Component Analysis (PCA), Information Theory, Fast Fourier Transforms (FFTs), and Wavelet Decomposition (Halilaj et al., 2018). The generated features are typically categorized into three types: time domain, frequency domain, and time-frequency domain features. Feature selection methods are techniques used to select the most discriminative and informative features by further reducing data dimensionality and computation time (Chandrashekar & Sahin, 2014).

Feature selection methods commonly used in HAR include Stepwise (forward and backward) feature selection, Hill climbing, and Informal selection (Halilaj et al., 2018). After signal processing is complete, the extracted features are used as input to the selected ML or DL model for activity recognition and classification. ML methods implement a scientific principle of 'trial and error', the process requires constant evaluation and refinement based on the model's prediction and loss of data (Jung, 2022, p.113). Evaluation metrics are used to assess the performance of classification and regression models and to avoid the bias-variance trade-off phenomena (i.e., overfitting or underfitting of the training data). Precision, specificity, sensitivity (also referred to as 'recall'), accuracy, and a confusion matrix are some of the most common evaluation metrics which are utilized in determining the performance of an ML model for sport motion detection and recognition tasks (Cust et al., 2019).

The aim of this study was to systematically review the literature on ML and DL methods applied to IMU data inputs for evaluating techniques or skills in individual swing and team sports. Previous reviews have been conducted on the application of AI for motion analysis in sports (Bonidia et al., 2018; Halilaj et al., 2018). Sensor-based reviews have examined IMUs for quantifying sport-specific movement (Alanen et al., 2021; Chambers, Gabbett, Cole, & Beard, 2015) sports performance evaluation (Camomilla et al., 2018), and real-time kinematic analysis in sport (Rana & Mittal, 2021). However, no study to date has aimed to systematically review studies that have used ML or DL techniques to quantify sport motion patterns via body-mounted IMUs inputs to evaluate skill in individual swing and team sports. There is a particular emphasis on ML methods inclusive of data processing and modeling in addition to sensor device specifications. Given the developments in the field of wearable technologies and AI for sport performance analysis, this review is necessary to understand the research area and to provide insight into potential research and applications in the field. This study intends to inform researchers of the current state of the application of AI for the recognition and classification of sports motion for technique evaluation in team and individual swing sports. Specifically, the research questions are: (1) What are the ML and DL methods that have been applied to IMU data in studies conducting technique and skill analysis in team and individual swing sports? (2) What are the current ML and DL frameworks for sports activity recognition inclusive of signal processing, feature extraction, and model performance evaluation techniques? (3) What are the most utilized sensor modalities in sports activity recognition inclusive of device specification, sampling rate, and sensor placement?

Methods

Search Strategy

This review was written in alignment with the Preferred Reported Item for Systematic Review and Meta-Analysis (PRISMA; Liberati et al., 2009) statement and guidelines. The different stages of the review protocol are identified in Figure 1. The search terms were created using advanced searches of keywords and phrases on the following electronic databases; IEEE Xplore, PubMed, Scopus, and Google Scholar. These terms are detailed in Table 1. The advanced searches were filtered down to include studies between January 2005 and January 2022 as no relevant studies were identified prior to this, and this period will be more reflective of the current applications of ML and DL for sport-specific movement recognition and classification. The search strategy was piloted and refined before being executed for the final selection of research articles to be included in this review. Once the selection of studies from the databases searches was finalized, snowballing was used to identify additional studies (Wohlin, 2014). The inclusion and exclusion criteria for this review are detailed in Table 2.

Table 1 Keyword Search Terms for Each Respective Database

IEEE Xplore
(((IMU OR sensor OR inertial sensor) AND (machine learning OR artificial intelligence OR data mining OR predictive modelling OR deep learning) AND (sport* OR athlete* OR player*)) AND (recognition OR classification OR detection)) AND (movement OR skill)
Scopus
((((“IMU” OR “sensor” OR “inertial sensor”) AND (“machine learning” OR “artificial intelligence” OR “data mining” OR “predictive modelling” OR “deep learning”) AND (“sport*” OR “athlete*” OR “player*”)) AND (“recognition” OR “classification” OR “detection”)) AND (“movement” OR “skill”))
PubMed
(((“IMU” OR “sensor” OR “inertial sensor”) AND (“machine learning” OR “artificial intelligence” OR “data mining” OR “predictive modelling” OR “deep learning”) AND (“sport*” OR “athlete*” OR “player*”)) AND (“recognition” OR “classification” OR “detection”)) AND (“movement” OR “skill”)
Google Scholar
(((IMU OR sensor OR inertial sensor) AND (machine learning OR artificial intelligence OR data mining OR predictive modelling OR deep learning) AND (sport* OR athlete* OR player*)) AND (recognition OR classification OR detection)) AND (movement OR skill)

Eligibility Criteria and Selection Process

Table 2. Inclusion and Exclusion Criteria

Inclusion Criteria	**Exclusion Criteria**
The movement was examined for sport-specific technique and/or skill purposes.	Does not investigate sports performance for the purpose of technique and/or skill detection and the classification associated with it.
Clearly defined which AI techniques or algorithms are being used, and how they’re tested and evaluated.	Data was gathered using smartphone-based sensors and/or not attached to fixed points on the subject’s body.
Individual swing and team sports.	Carried out movement analysis solely for everyday activity purposes e.g., walking, and sitting. Carried out movement analysis solely for clinical purposes.
Primary research is published as full-length studies in a peer-reviewed journal.	Studies were published before 2005.
The studies were written in English.	Review studies.

Methodological Quality Assessment

Studies were assessed for risk of bias using the evaluation criteria as described by Claudino et al (2019) are outlined in Table 3. Scores were allocated based on how well each criterion was met, up to a maximum possible score of 8 (low risk of bias). Studies with a score of 4 or lower were deemed a high risk of bias and were therefore excluded from this study.

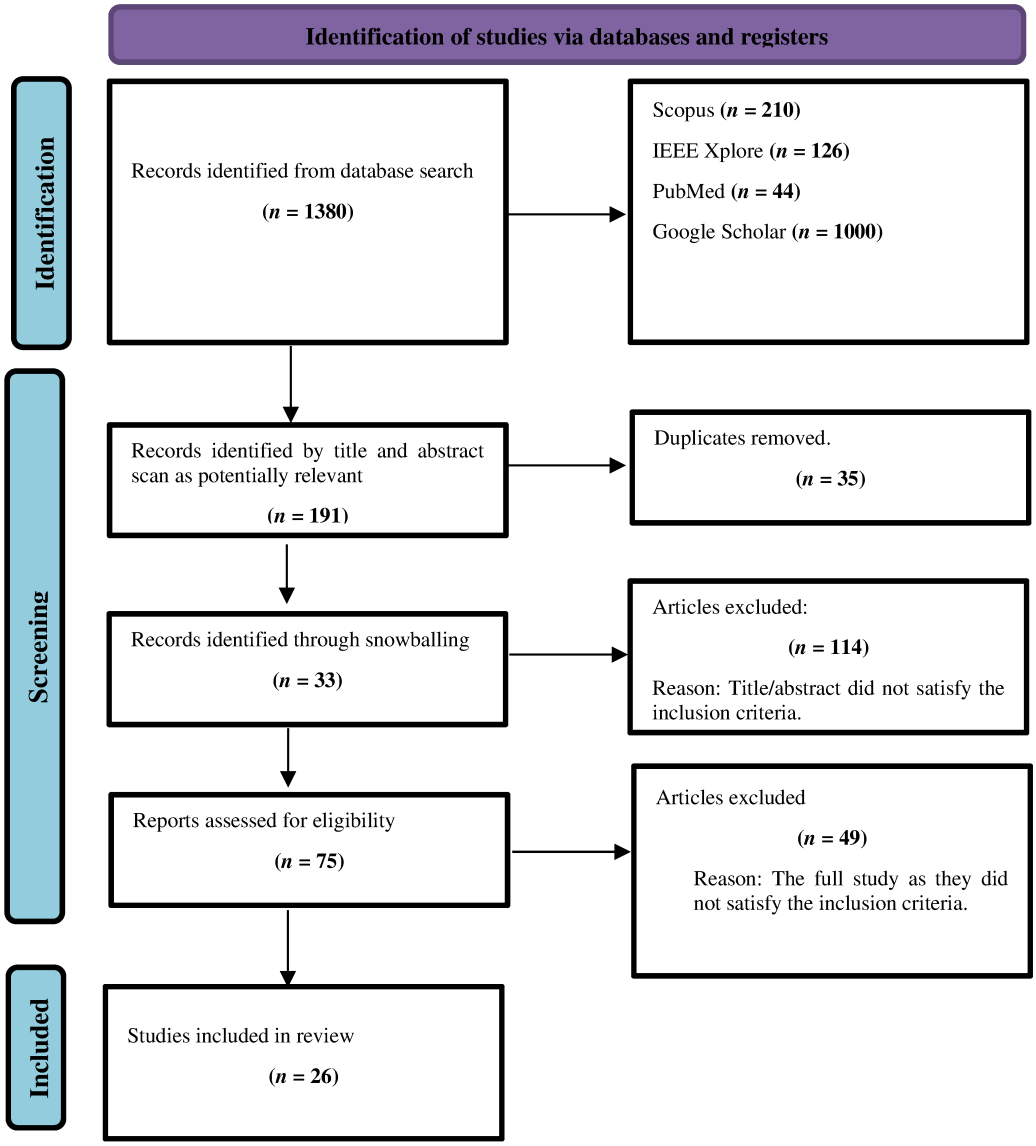
Table 3. Risk of bias evaluation criteria

			Scoring		
Criteria		**Definition**	0	1	2
A	Peer-reviewed	A study published in a peer-reviewed journal	No	Yes	-
B	Real-world approach	The targeted technique/skill being analyzed was described. The approach was performed with real results/data of athletes.	No	Yes	-
C	Population defined	Age, gender, sport, and level of experience were described.	No	Partly	Yes
D	Experimental design	The experimental design of the study was described and replicable. (The IMU or Video specification was described.)	No	Partly	Yes
E	Artificial Intelligence	The AI techniques were described including the method(s) of model evaluation and results.	No	Partly	Yes

Data Extraction

A total of 13 study characteristic parameters were extracted from the 26 research studies including the title, author, year of publication, participant details (number of participants, gender, age, experience, and/or skill level), sport technique/skill targeted, device (sensor or camera) specification, associated sampling rate and/or resolution and FPS, sensor placement, the filter used for sensor fusion, ML or DL model used, features extracted, model evaluation method and model performance metric. For this review, the sport-specific movement was defined as technique or skill associated with training or gameplay for team sports and individual swing sports that require a non-periodic and dynamic range of complex movements (e.g., change of direction, running, jumping, swinging, jumping). Sports where the body is in a fixed position or where the full range of movements are restricted by equipment associated with the sport have been excluded e.g., skiing, canoeing, and yoga.

Figure 1. PRISMA flow diagram used in the reviewing process



Results

In the scope of this systematic review, an initial 1380 studies were identified through the database search, and an additional 33 studies were identified through reference scanning. A total of 26 studies were included in this systematic review due to a strict inclusion criterion. An analysis of the reviewed studies in relation to the specific research questions and characteristics of these studies is provided in Table 4 using the reviewing methodology previously described by (Wickramasinghe, 2022).

Table 4. Studies Reviewed and the Research Questions Addressed

N	Reference	RQ1	RQ2	RQ3	N	Reference	RQ1	RQ2	RQ3
1.	Anand et al. (2017)	Y	Y	Y	14.	McGrath et al. (2021)	Y	Y	Y
2.	Bo (2022)	Y	Y	N	15.	Reilly et al. (2021)	Y	Y	Y
3.	Brzostowski and Szwach (2018)	Y	Y	Y	16.	Salim et al. (2020)	Y	Y	Y
4.	Cust et al. (2021)	Y	Y	Y	17.	Salman et al. (2017)	Y	Y	Y
5.	Fan et al. (2021)	Y	Y	Y	18.	Shahar at al. (2020)	Y	Y	Y
6.	Ganser et al. (2021)	Y	Y	Y	19.	Smith and Bedford (2020)	Y	Y	Y
7.	Gençoğlu and Gümüş (2020)	Y	Y	Y	20.	Steels et al. (2020)	Y	Y	Y
8.	Hollaus et al. 2020	Y	Y	Y	21.	Tan & Xie (2021)	Y	Y	Y
9.	Jowitt et al. (2020)	Y	Y	Y	22.	Van den Tillaar et al. (2021)	Y	Y	Y
10.	Kautz et al. (2017)	Y	Y	Y	23.	Wang et al. (2016)	Y	Y	Y
11.	Khan et al. (2017)	Y	Y	Y	24	Wang et al. (2018)	Y	Y	N
12.	Kim and Park (2020)	Y	Y	Y	25.	Whiteside et al. (2017)	Y	Y	Y
13.	McGrath et al. (2019)	Y	Y	Y	26.	Zhang et al. (2021)	Y	Y	Y

Experimental Design

A variety of team sports (18 studies) and individual swing sports (8 studies) were investigated for the identification and classification of key performance parameters relating to sport-specific technique and skill as shown in Figure 2. The sports investigated in the reported studies included American football (n = 1), Australian Rules Football (n = 1), badminton (n = 2), basketball (n = 2), cricket (n = 5), golf (n = 1), handball (n = 2), field hockey (n = 1), netball (n = 1), soccer (n = 2), tennis (n = 3), table tennis (n = 1), volleyball (n = 3). One study carried out shot detection, classification, and kinematic analysis of shot types in swing sports tennis, badminton, and squash (Anand, Sharma, Srivastava, Kaligounder, & Prakash, 2017).

Across the studies reported, activity recognition and classification are carried out by traditional supervised ML models in 17 studies and DL models in 8 studies, as illustrated in Figure 3. One study conducted ensemble modeling by combining three traditional ML models (Decision Tree (DT), Random Forest (RF), and SVM) for the automatic classification of swing motion in table tennis (Tan & Xie, 2021). In the studies reviewed studies, a classification approach is proposed encompassing the two major classes of classification: binary and multi-class. In binary classification, the given dataset is assigned exactly one of two classes. In multi-class classification, the given dataset is assigned to several classes established by the predetermined classification rules (Jha, Dave, & Madan, 2019). There are instances of binary classification in this review, for example, a CNN was used to identify unique features of IMU signal data and classify the data into a catch (success) or drop (fail) in American Football (Hollaus, Stabinger, Mehrle, & Raschner, 2020). In tennis, multiclass classification was used to detect and classify stroke type (Brzostowski & Szwach, 2018; Whiteside, Cant, Connolly, & Reid, 2017), and shot type (Anand et al., 2017; Ganser, Hollaus, & Stabinger 2021). A CNN architecture was used to differentiate types of shooting postures in basketball (Fan, Bi, Wang, Zhang, & Sun, 2021), an SVM model was used to classify throw type and approach in handball (van den Tillaar, Bhandurge, & Stewart, 2021) and to analyze shots in cricket (Khan, Nicholson, & Plötz, 2017).

Specific locomotive activities related to skill fluency were reported for classification models, for example, an SVM predicted skill level in volleyball (Wang et al., 2018), and netball (Smith & Bedford, 2020). Linear Discriminant Analysis (LDA) was used in the recognition and classification of 7 volleyball activities (Salim et al., 2020). LDA assumes that all classes are

linearly separable and work well in distinguishing multiple activity classes if the features extracted exhibit clear boundaries (Vaibhaw, Sarraf, & Pattnaik, 2020). Feature ranking algorithms identify kinematic features related to sport technique and classify accelerometer-derived features to predict throwing velocity and accuracy in handball (Gençoğlu & Gümüş, 2020) and detection of change of direction (COD > 45°) in soccer (Reilly, Morgan, Czanner, & Robinson, 2021).

The pooled sample consisted of 493 participants, representing the cumulative sample size reported in all 26 studies). Among the 22 studies that reported gender distribution, 74% (n = 283) were male participants, while 26% (n = 102) were female participants. The total number of participants reported ranged from 2 participants (Steels et al., 2020) to 70 participants (Anand et al., 2017). Sports experience level was reported in 18 out of the included studies, with variations in reporting across different sports. Grade level was specified in cricket (McGrath, Neville, Stewart, & Cronin, 2019), Australian football (Cust, Sweeting, Ball, & Robertson, 2021), tennis (Whiteside et al., 2017) and soccer (Reilly et al., 2021). Other studies reported the amateur professionalism in basketball (Hollaus et al., 2020), badminton (Steels et al., 2020), cricket (Jowitt, Durussel, Brandon, & King, 2020; Khan et al., 2017; McGrath, Neville, Stewart, Clinning, & Cronin, J, 2021), golf (Kim & Park, 2020), handball (van den Tillaar et al., 2021), netball (Smith & Bedford, 2020), tennis (Brzostowski & Szwach 2018; Ganser et al., 2021), volleyball (Kautz et al., 2017; Wang et al., 2018), and years of experience was reported in basketball (Fan et al., 2021). Wang, Guo, and Zhao (2016) specified that the badminton players were either members of the university club or team.

Figure 2. The distribution of sports investigated by studies reported in this review.

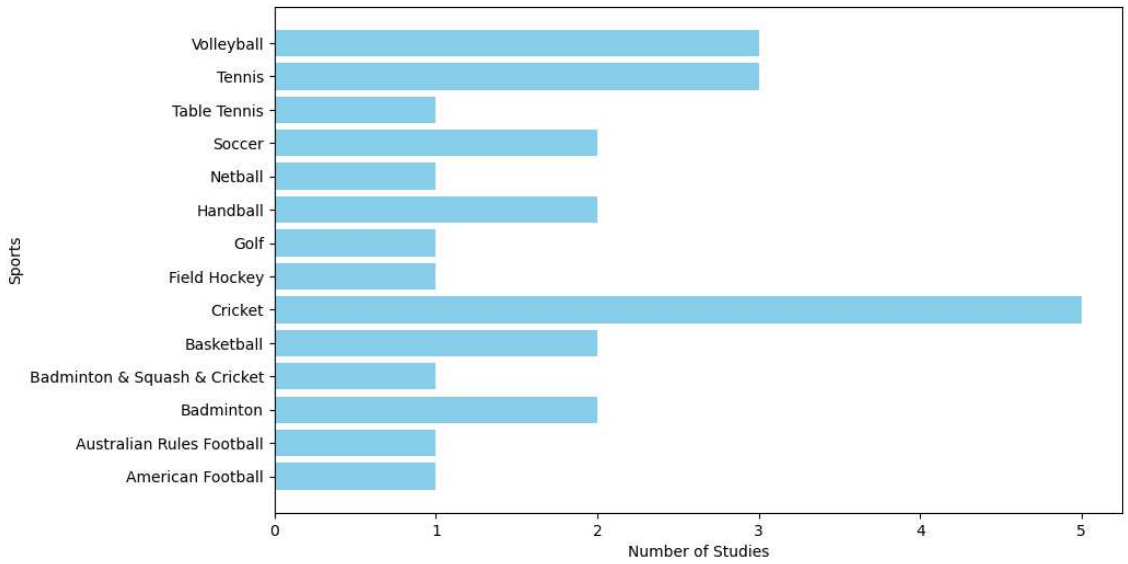
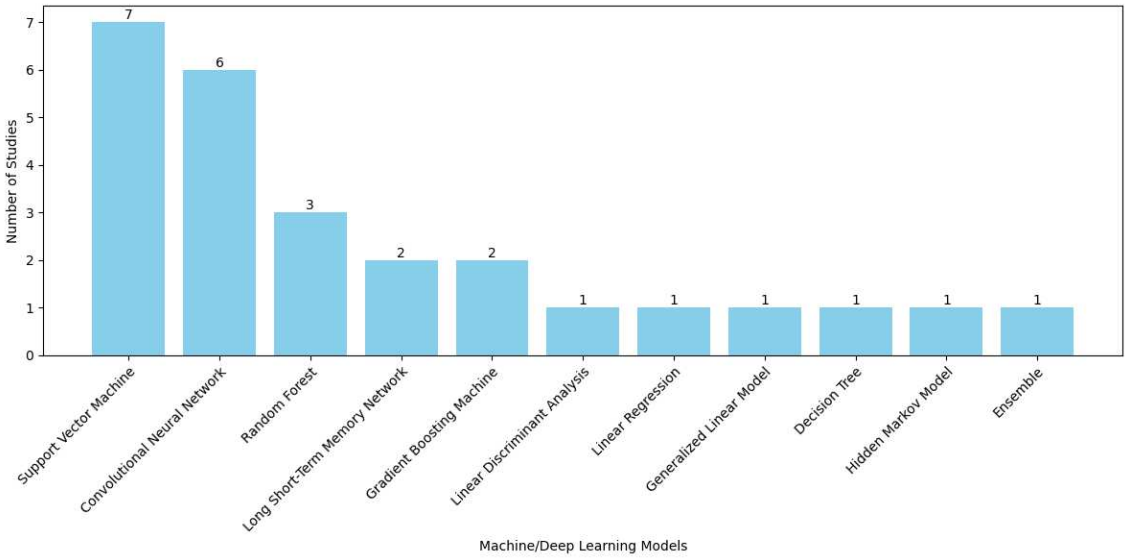


Figure 3. The distribution of machine and deep learning models investigated by studies reported in this review.



All datasets reported in this review are original datasets consisting of IMU-derived signal features that were collected through primary research involving the participants reported in the studies. IMU-based datasets for feature learning varied from 28,583 instances (Whiteside et al., 2017) to 100 instances (Wang et al., 2018) for ML training and validation.

Data Acquisition and Device Specifications

IMUs were employed to detect and capture sports motions in all the studies reviewed, and the details are presented in Table 5. The dataset includes a diverse selection of IMUs, including 23 commercially available IMU systems, with an additional three that are custom-adapted for experimental purposes (Hollaus et al., 2020; Salman et al., 2017; Wang et al., 2018). However, it is noteworthy that one study within the review did not publish adequate information regarding their sensing module (Bo, 2022). The number of body-mounted sensors used ranged from one to four, respectively. The sensor placement varied throughout the studies under review, with the most frequently reported location being the wrist (n = 16 studies), upper back between the shoulder blades (n = 4 studies), upper arm (n = 4 studies), left and right thigh (n = 1 study), left and right leg (n = 1 study), foot (n = 2), ankle (n =2 studies), waist (n = 2 studies), chest (n =1 study), pelvis (n = 1) and head (n = 1).

Commercial purpose-built vests that held the IMU device in place were reported in 4 studies (Jowitt et al., 2020; McGrath et al., 2019; McGrath et al., 2021; Smith & Bedford, 2020). Sufficient information regarding the system architecture including the gyroscope, accelerometer, and magnetometer range (◦/s) and sampling rate (Hz) was included in 70% of studies. The ranges for accelerometers varied from ± 2 G - ± 16 G, for gyroscopes it was from 250 ◦/s - ± 2000 ◦/s, and for magnetometers, it was from 1000 µT - ± 4900 µT. The highest sampling frequency reported was 1125 Hz (van den Tillaar et al., 2021), while the lowest was 20 Hz (Zhang, Lyu, Zhang, & Yang, 2021). A high sampling rate of the IMU is an important factor during high angular velocities such as overarm throwing and swinging (Gençoğlu & Gümüş, 2020).

Table 5. Data Acquisition Specification, IMU Placement and Sampling Frequency in Reported Studies (NR = not reported)

N	Reference	Number of IMU	IMU Placement	Sampling Frequency	Dataset Size
1.	Anand et al. (2017)	1	Dominant wrist	100 Hz	15600
2.	Bo (2022)	5	Left and right thigh, left and right thigh, and torso	50 Hz	NR
3.	Brzostowski and Szwach (2018)	1	Dominant wrist	100 Hz	2415
4.	Cust et al. (2021)	1	Ankle	500 Hz	587
5.	Fan et al. (2021)	2	Dominant wrist and foot	100 Hz	12177
6.	Ganser et al. (2021)	1	Dominant wrist	1000 Hz	5682
7.	Gençoğlu & Gümüş (2020)	1	Dominant wrist	500 Hz	2307
8.	Hollaus et al. (2020)	2	Wrists	NR	759
9.	Jowitt et al. (2020)	1	Between shoulder blades	100 Hz	782
10.	Kautz et al. (2017)	1	Right wrist	39 Hz	12180
11.	Khan et al. (2017)	4	Distal tibias and wrists	100 Hz	974
12.	Kim and Park (2020)	3	Left wrist, head, and pelvis	200 Hz	389
13.	McGrath et al. (2019)	1	Between shoulder blades	250 Hz	401
14.	McGrath et al. (2021)	1	Between shoulder blades	250 Hz	1555
15.	Reilly et al. (2021)	1	Chest	100 Hz	20000
16.	Salim et al. (2020)	2	Wrists	NR	990
17.	Salman et al. (2017)	3	Upper arm, forearm, and wrist	150 Hz	300
18.	Shahar et al. (2020)	4	Right and left wrist, chest and waist	200 Hz	NR
19.	Smith and Bedford (2020)	1	Between shoulder blades	100 Hz	5585
20.	Steels et al. (2020)	3	Rackets Grip, wrist, and upper arm	100 Hz	6639
21.	Tan and Xie (2021)	1	Dominant Wrist	200 Hz	500
22.	Van den Tillaar et al. (2021)	1	Distal dorsal side of dominant arm	1125 Hz	1020
23.	Wang et al. (2016)	4	Right and left wrist, right ankle and waist	100 Hz	NR
24.	Wang et al. (2018)	1	Right wrist	NR	100
25.	Whiteside et al. (2017)		Forearm	500 Hz	28583
26.	Zhang et al. (2021)	1	Dominant foot	20 Hz	960

Data Pre-processing and Feature Extraction

Filtering of acceleration signals to increase recognition and classification accuracy was reported in 16 studies. The application of either a low or high-pass filter is common in biomechanical data analysis prior to inverse kinematic procedures (Crenna, Rossi, & Berardengo, 2021). The choice of filtering and cutoff frequency will depend on the characteristics of the signal and the goal of the analysis. Low-pass filtering was applied in the time-domain or frequency-domain, and in some instances both domains to reduce the influence of noise fluctuations on the accelerometer data. The use of low-pass filtering included the Butterworth 1 Hz low-pass filter (McGrath et al., 2019), Butterworth 5 Hz third-order low-pass filter (Jowitt et al., 2020), Butterworth 8 Hz low-pass filter (Kautz et al., 2017), Butterworth 10 Hz tenth low-pass filter (Kim & Park, 2020), Butterworth 20 Hz second order filter (Gençoğlu & Gümüş, 2020), a one-dimensional Gaussian filtering method (Tan & Xie, 2021), Finite Impulse Response (FIR) filter used in conjunction with a Kaiser Window Function (Ganser et al., 2021), Simple Moving Average (SMA) algorithm (Brzostowski & Szwach, 2018), 50-point moving average technique (Cust et al., 2021; Whiteside et al., 2017) and a 3-point filter moving average (Wang et al., 2018). High-pass filters are used to remove corrupted accelerometer signals and low-frequency information, these include: a fourth-order 1 Hz Butterworth high-pass filter (McGrath et al., 2021), Kalman Filter (Zhang et al., 2021), and one study used a low-pass Savitzky-Golay filter and a 0.5 Hz high-pass filter, applied to the acceleration data (Smith & Bedford, 2020). Linear interpolation was used for signal resampling to address variation in the sampling frequencies in 4 studies (Ganser et al., 2021; Hollaus et al. 2020; van den Tillaar et al., 2021; Whiteside et al., 2017).

Detection methods to segment the motion states were reported in 22 studies. Window segmentation which involves splitting the sensor signals into distinct windows of time (measured in seconds), was the most common method of activity detection (Bo, 2022; Brzostowski & Szwach, 2018; Cust et al., 2021; Jowitt et al., 2020; Khan et al., 2017; Kim and Park, 2020; McGrath et al., 2019; McGrath et al. 2021; Reilly et al., 2021; Salim et al., 2020; Salman et al., 2017; Shahar, Ghazali, As'ari, & Swee, 2020). Smith & Bedford, 2020; Steels et al., 2020; Tan & Xie, 2021; van den Tillaar et al., 2021; Whiteside et al., 2017; Zhang et al., 2021). A thresholding technique provides a hard cut-off point for activity signals and is often used to identify specific actions or patterns in the signal such as accelerations or jerks associated with striking a ball, for example. A combination of window segmentation and thresholding was reported in 4 studies (Anand et al., 2017; Ganser et al., 2021; Kautz et al., 2017; Wang et al., 2016).

Further feature extraction methods are used to reduce the dimensionality of the dataset and prepare it for ML modeling for improved accuracy and were reported in 5 studies. Power spectral density analysis using a Fourier Transform was reported to convert the time series from the time domain to the frequency domain (Kim & Park, 2020). Khan et al. (2017) also reported the use of the Fourier Transform in combination with signal energy, signal entropy and total squared jerk calculations to comprehend the complexity of shots on each axis. Mel-frequency cepstral coefficients (MFCC) were used to extract frequency-based features and involved PCA for dimensionality reduction and an FFT in the MFCC calculation (Brzostowski & Szwach, 2018).

Principal Component Analysis (PCA) was also used to reduce the dimensionality and compare variance in the data, and with three principal components the different skill levels of volleyball players were visible resulting in increased classification accuracy of the SVM model (Wang et al., 2018). Dynamic Time Warping (DTW) was used in the window segmentation of shot analysis in individual swing sports (Anand et al., 2017).

To distinguish the participant's activity or specific motion states from the sensor data, extracted

features are used as inputs to the ML algorithm. Detection methods such as window segmentation enable feature values of each data segment to be calculated, and feature vectors are then constructed. The signals are split into windows of a fixed time, and this approach is particularly suited to real-time applications such as in situ sports analysis due to the lack of pre-processing determinants required (Banos, Galvez, Damas, Pomares, & Rojas, 2014). Most studies derived features from acceleration data in the X, Y, and Z axes. The specific features extracted were reported according to the signal processing technique used in 20 studies.

Time domain features were reported, such as the mean value of the acceleration along the X, Y, and Z axes, standard deviation, skewness, and kurtosis within the window (Anand et al., 2017; Cust et al., 2021; Jowitt et al., 2020; Kim & Park, 2020; Reilly et al., 2021; Salim et al., 2020; Salman et al., 2017; Shahar et al., 2020; van den Tillaar et al., 2021; Whiteside et al., 2017; Zhang et al., 2021). Studies reported a combination of features extracted from both the time and frequency domains, utilizing two or more categories from these three options (Brzostowski & Szwach, 2018; Kautz et al., 2017; Khan et al., 2017; McGrath et al., 2019; McGrath et al., 2021; Smith & Bedford, 2020; Wang et al., 2016; Wang et al., 2018).

Tan and Xie (2021) describe twenty-two features which are extracted from time domain, inter-axial and additional features to comprehensively describe the acceleration in the X, Y, and Z axes for recognition of athletes' limb movement in table tennis. Gençoğlu and Gümüş (2020) describe the use of ML algorithm of inertial sensors without the use of features as a reference method for prediction of throwing velocity in volleyball. DL algorithms such as CNN and Recurrent Neural Network (RNN) leverage the power of automatic complex feature learning through backpropagation and the use of multiple layers such as convolutional and maximum pooling layers (Bo, 2022; Fan et al., 2020; Ganser et al., 2021; Hollaus et al., 2020; Steels et al., 2020). The types of features that which were used as input to a CNN algorithm were described as the course of linear acceleration and angular velocity for each stroke in badminton (Steels et al., 2020).

Modeling and Performance Evaluation

In all 26 papers, the outcome variable was the recognition of a sport-specific movement and an associated analysis of that motion by an ML or DL model. SVMs were reported in 7 papers (Khan et al., 2017; McGrath et al., 2019; Salman et al., 2017; Shahar et al., 2020; Smith & Bedford, 2020; Wang et al., 2018; Whiteside et al., 2017). RF was reported in 3 studies, for the detection and classification of key events in Australian football, cricket, and soccer respectively (Cust et al., 2021; Jowitt et al., 2020; Reilly et al., 2021). Two studies reported the use of a gradient boosting machine (GBM) to classify throw type (circular or whip-like) and throw approach type (running, standing, or jumping) from the signal features (van den Tillaar et al., 2021) and to predict bowling release speed and perceived intensity zone in cricket (McGrath et al., 2021).

A multilayer DT recognizer was used in soccer players to identify the joint movement of athletes (Zhang et al., 2021). Logistic regression was used to detect and classify tennis stroke types (forehand, backhand and serve) to assist in tennis technique analysis (Brzostowski & Szwach, 2018). One study took an ensemble learning approach by combining DT, RF, SVM, and CNN to create an integrated learning algorithm to effectively distinguish four types of swings in table tennis (Tan & Xie, 2021). The use of two-layer Hidden Markov Model (HMM) was used in badminton training to accurately recognize and categorize 14 distinct badminton strokes based on accelerometer data (Wang et al., 2016).

The use of a CNN model was reported in 6 studies (Fan et al., 2021; Ganser et al., 2021; Hollaus et al., 2020; Kautz et al., 2017; Kim & Park, 2020; Steels et al., 2020). Utilising Long Short-

Term Memory Networks (LSTM) Anand et al. (2018) combined CNN and Bidirectional LSTM neural network architectures to develop a novel sports analytics system for comprehensive analysis of swing sports. Similarly, Bo (2022) implemented a Multilayer Parallel LSTM algorithm for the recognition of basketball player activities achieving 94% recognition accuracy.

In studies that used two or more ML classifiers for comparative or evaluation purposes, only the best-performing algorithm in each study was selected for reporting in this review. Evaluation metrics were used in all studies to assess the predictive performance of the model. Accuracy, which is given as the percentage ratio of correct predictions against all predictions was the most common evaluation metric as per 18 studies. Reported accuracy varied from 99% (Steels et al., 2020) to 78.17% (Salim et al., 2020). Accuracy was calculated through a variety of methods such as calculating the Mean Absolute Error (MAE), Area Under the Curve (AUC), and F1 score.

Cross-validation techniques to split the data into training and testing subsets was a common technique employed to facilitate model evaluation. Various folds (4 to 10) of the k-Fold cross-validation fitting procedure was used to facilitate the evaluation of the model's performance by dividing the dataset into K-equal-sized subsets to provide a robust evaluation (Brzostowski & Szwach, 2018; Shahar et al., 2020; Smith & Bedford, 2020; Wang et al., 2016; Whiteside et al., 2017). K-fold cross-validation was reported in various studies to fine-tune the hyperparameters to maximize the performance of the model on the validation set (Bo, 2022; Jowitt et al., 2020; McGrath et al., 2019; McGrath et al., 2021).

Repeated k-Fold cross-validation, which involves repeating the cross-validation procedure multiple times and reporting the mean results across all folds was reported as a technique to facilitate model performance evaluation (Reilly et al., 2021) and fine-tuning of hyperparameters (Wang et al., 2018). Stratified k-fold cross-validation is an extension of the k-fold cross-validation technique that ensures that each k-fold maintains the same distribution as the original dataset and was reported in 3 studies in this review (Cust et al., 2021; Ganser et al., 2021; Hollaus et al., 2020).

Leave-One-Out-Cross-Validation (LOOCV) which can measure accuracy, precision and recall was used to evaluate the aggregated performance of the models, in LOOCV the model is iteratively trained and tested on each individual data point (Brzostowski & Szwach, 2018; Kautz et al., 2017; Khan et al., 2017; Kim & Park, 2020; McGrath et al., 2021; Salim et al., 2020; Salman et al., 2017; Smith & Bedford, 2020; van den Tillaar et al., 2021).

Precision as a measure of quality, and recall as a measure of quantity, were frequently used in combination with the F1-score (as the harmonic mean between precision and recall) to evaluate the performance of the model (Anand et al 2017; Cust et al., 2021; Fan et al., 2021; Ganser et al., 2021; Kautz et al., 2017; McGrath et al., 2019; Salim et al., 2020; Salman et al., 2017; Steels et al., 2020; van den Tillaar et al., 2021; Wang et al., 2018). Absolute Error (MAE) of the bowling release speed in cricket was used in conjunction with the F-1 score of a GBM to determine accuracy (McGrath et al., 2021). The area under the receiver operating characteristic curve (AUROC) was used to determine sensitivity and specificity in 3 studies (Cust et al., 2021; Jowitt et al., 2020; Reilly et al., 2021). A confusion matrix was used in 8 studies to visualize the predictive performance of the model (Ganser et al., 2021; Hollaus et al., 2020; McGrath et al., 2019; McGrath et al., 2021; Smith & Bedford, 2020; Steels et al., 2020; Tan & Xie, 2021; Whiteside et al., 2017). In Table 6, a synopsis and noteworthy remarks for all studies investigated are provided. This table contributes to a comprehensive understanding of the research landscape concerning the application of AI to IMU data inputs for swing and team sport-specific recognition and classification.

Table 6. *S*ynopsis and Noteworthy Remarks of Studies Investigated

Reference	**Sport Technique or Skill Targeted**	**Synopsis**	**Noteworthy Findings**
Anand et al. 2017	Shot detection and classification of badminton, tennis, and squash.	Built on the innovative concept of Jerk-effective for precise calculation of the point of impact in a shot sensor signal, coupled with a neural network architecture (CNN and BLSTM) for shot classification, this study provides a thorough analysis of swing sports.	In this study, a novel approach to shot type classification was introduced, by leveraging advanced feature engineering and exploring the correlations between physical motion and sensor data. This resulted in enhanced accuracy of shot classification and gave valuable insights into the relationship between physical and sensor measurements.
Bo, (2022)	Recognition of 6 basketball activities.	A multisensory architecture is utilized in combination with a Multilayer Parallel LSTM for activity recognition achieving an overall accuracy of 94.77%.	The addition of DQN (Deep Q-Network) to adaptively control the sampling frequency of the sensing module reduces energy consumption by 76% while maintaining a recognition accuracy of 90%.
Brzostowski and Szwach (2018)	Automatic classification of stroke type in tennis	This study proposed a tennis technique analysis system using data from a wrist-worn accelerometer and a K-NN and logistic regression classifiers. The main functionalities are tennis stroke type (forehand, backhand and serve) detection and classification to assist in the analysis of tennis serve in training.	*Effect of window size:* When the window size is increased to 80 samples, the classification performance of k-NN and Logistic Regression classifiers are also increased.
Cust et al. (2021)	Recognition and classification of 4 Australian Rules Football (AF) kick types.	Ankle-mounted IMUs are utilized in combination with supervised ML models to create semi-automated training kick detection and type at varying distances (10, 20, and 30 meters) classification.	The limited range of the accelerometers ($\pm$ 16 g), and gyroscopes ($\pm$ 2000°/s) used in the study resulted in saturation in the signals of kicks due to high velocities and forceful impacts. This highlights the importance of carefully aligning the device specifications and capabilities with the specific activity to be analyzed.
Fan et al. (2021)	Basketball shooting posture.	An IMU sensor fusion framework dedicated to basketball shooting postures is designed to collect shooting posture data and perform sensor data fusion and event detection using the Visual Geometry Group Network 16 (VGG16) deep learning model based on CNN architecture.	The use of several convolutional units to extract frequency and spatial features of the input data. The max pooling layer is mainly used for reducing feature dimensionality.

Ganser et al. (2021)	Automatic classification of shot types in tennis.	Complex tennis analysis such as the indication of success rates per shot type can assist in training prescription and real-time biomechanical performance analysis. This paper proposes a model which classifies tennis shot types into three and five-shot types using a fully CNN.	This study found that, with the utilization of a rich and homogeneous dataset, a deep neural network approach achieved high accuracies in classifying tennis shots. Using a combination of filtered angular velocity (ωy) and jerk (j) proved effective in achieving a reliable detection rate.
Gençoğlu and Gümüş (2020)	Kinematic parameters associated with throwing accuracies in handball.	The aim of this study was to predict throwing velocity in handballs by using inertial measuring units and machine learning.	The findings indicate a high sampling frequency should be considered when assessing activities that are associated with high angular velocities such as overarm throwing.
Hollaus et al. (2020)	American Football ball catches; success and failure.	CNN network identifies unique features in the data sequences and classifies the data as a catch or drop.	Magnetometer and accelerometer sensor data seem to be of the highest significance for increasing the mean accuracy of the neural network for the classification of catch and jump activities in American Football.
Jowitt et al. (2020)	Cricket bowling action workload.	This study aimed to develop and independently validate a robust machine-learning algorithm that utilizes GPS, accelerometer, and gyroscope data for the automatic detection of cricket bowling deliveries during both training and match scenarios.	In this study, the trade-off between sensitivity and specificity are discussed. The selection process for evaluation prioritizes as many true positive cases (of specified testing conditions) as possible. This approach is inclusive as it ensures that fewer positive cases are missed, but it also increases the risk of more negative included cases being included, which do not meet the specific criteria.
Kautz et al. (2017)	Monitoring and classification of 9 volleyball activities.	A novel monitoring system based on deep CNN for sensor-based activity recognition in volleyball is introduced. The proposed CNN model is compared with 5 traditional ML models and outperforms by a substantial margin (16%) in terms of accuracy and computation time.	The reported performance evaluation metrics include a high recall rate of 99%, indicating the CNN model's robust ability to identify the majority of major events. However, the precision rate of 34.8% suggests that the model may detect a higher number of false positives. The balanced accuracy of 90% suggests the model exhibits both sensitivity and specificity. This study's evaluation approach highlights the importance of including several metrics for comprehensiveness.
Khan et al. (2017)	Shot analysis and skill classification in cricket.	This study presents a hierarchical shot analysis system that creates meaningful visualizations of key performance parameters for the	The use of Empirical Cumulative Distribution Function (ECDF) is used to visualize the distribution of sensor datapoints and gives insight to the

		assessment of batting quality in cricket.	characteristics and overall distribution of the data.
Kim and Park (2020)	Golf swing analysis.	In this study golf swing, segmentation was conducted using inertial measuring units and motion capture cameras, with the application of machine learning to estimate and divide golf swings.	The findings of this study show that CNN and BLSTM methods offer distinct advantages of extracting and learning "hidden" features. These features are typically difficult to extract using heuristic approaches and have the additional benefit of considering multiple signals simultaneously.
McGrath et al. (2019)	Cricket fast bowling action detection.	The fast-bowling workload in cricket has an established relationship with performance. An IMU in combination with an SVM approach was used for event detection in cricket bowling.	This study highlights that not all models will perform well on all problems. The SVM model had the highest accuracy (and F1 score) when trained using data sampled at 250 Hz but, had the lowest accuracy when trained on down sampled 25 Hz data. Features derived from the gyroscope sensor showed high importance for accurately detecting bowling events and non-events..
McGrath et al. (2021)	Cricket bowling release speed and perceived intensity.	This study investigated whether an IMU placed at the thoracic spine and the application of a machine learning model could accurately predict bowling release speed (BRS) and perceived intensity zone (PIZ) and investigated the accuracy at a range of different sensor sampling frequencies.	The Extreme Gradient Boosting (XGB) model performed best in this study. BRS with a mean absolute error (MAE) of 3.61 km/h and PIZ with an F score of 88% was measured using the XGB model. Features derived from the gyroscope during the pre-delivery and delivery phases were significant.
Reilly et al. (2021)	Automatic detection of change of direction (COD > 45°) in soccer.	This study aims to develop an automated classification model utilizing multi-sensor data (GPS, acceleration, gyroscopic) to detect changes of direction (COD) greater than 45°. Video analysis was employed to establish ground truth for the occurrence of COD events exceeding 45°.	Among the variables considered the GPS-COD angle variable demonstrates the highest level of importance in correspondence with the most powerful RF model.
Salim et al. (2020)	Automatic recognition of specific volleyball actions.	Employing linear discriminant analysis to automatically recognize specific types of volleyball actions (underhand serve, overhead pass, serve, forearm pass, one-hand pass, smash, and block) based on motion data captured by IMU sensors strapped onto the wrist of 8 volleyball players.	The dominant hand plays an important role in action type classification, as the UAR for the dominant hand was 67.78% for the dominant hand compared to 61.79% for the non-dominant hand.

Salman et al. (2017)	Cricket bowling action legality analysis.	This paper examined the movement of a bowling arm in the correct rotation and precise angle using sensing technology. The action profiles are then used to simplify the complex biomechanical movement of the bowling arm.	Leveraging sensor data from the upper arm, forearm, and wrist in cricket yields remarkably high accuracy, particularly in capturing high-impact accelerations.
Shahar et al. (2020)	Classification of 6 field hockey activities (passing, drag flick, drive, dribbling, receiving and tackles).	This study focuses on the influence of sensor numbers and placement on the recognition accuracy for field hockey activities.	The choice of sensor number and placement and their impact on classification accuracy is discussed. Notably, the combination of all four sensors located at the chest, waist, right wrist, and left wrist, in combination with the use of a SVM model showed favorable accuracy (96.7%) for activity recognition in field hockey.
Smith and Bedford (2020)	Automatic classification of movement types, change of direction and intensity in netball.	Automatically classify locomotion in elite Netball matches using IMU and machine learning models where subjects perform rapid changes in movement type, direction, and intensity.	The peak accuracy was attained by incorporating both accelerometer and gyroscope features, mapped across both time and frequency domains.
Steels et al. (2020)	Recognition and classification of badminton stroke types.	Automatic recognition of 7 important seven specific badminton strokes using inertial measuring unit sensor data and a CNN model.	By utilizing only accelerometer data the novel CNN introduced in this study is capable of differentiating activities with a recognition rate of 86% at a sampling frequency of 50 Hz. With the addition of gyroscope data, the precision significantly improves to 99%. This surpasses the results achieved by a traditional CNN, with precision rates of 79% and 88% respectively.
Tan and Xie (2021)	Automatic classification of swing motion types in table tennis.	This paper analyzes table tennis swing motions—forehand attack, backhand push, forehand rub, and backhand rub—using three-axis accelerometer data. Three machine learning algorithms (DT, RF, SVM) are employed for automatic classification, and a novel classification algorithm (CNN) based on integration is designed.	The inclusion of residual convolution in the design of time series network prevents performance degradation and ensures the preservation of training outcomes.

Van den Tillaar et al. (2021)	Handball throw type classification and estimation of ball velocity.	The signal features served as inputs for the classification model, distinguishing throw type (circular and whip) and throw approach types (running, standing, or jumping). The models also measured the ball throw velocity (m/s) using a radar gun.	The main finding of this study is that by utilizing IMU data in conjunction with an SVM model, it was possible to predict peak ball velocity with a margin of error of 1.1 m/s, and classify approach and throw types with an accuracy of 79-87%.
Wang et al. (2016)	Classification of 14 badminton stroke types.	A window segmentation method for stroke detection is proposed in combination with a HMM to classify badminton strokes. Suitable sensor locations for badminton stroke recognition are discussed.	A new window segmentation method with containing stroke points (WCSP) is described, 1-second windows enable the selection of relevant sample points around badminton stroke events, creating observation windows with specific lengths, and ensuring that stroke points are positioned at the center of these windows. The recognition accuracy of the sensor placed at the right wrist is highest (81.63%) for recognition of badminton stroke types.
Wang et al. (2018)	Assessing skill level based on spike actions in Volleyball.	Quantifying athletic performance by tracking player motion and assessing skill level using a wrist-mounted IMU and a high-speed camera for ground truth.	Principal Component Analysis (PCA) was used to reduce computational extraction load and to increase skill assessment.
Whiteside et al. (2017)	Monitoring hitting load associated with shot types in Tennis.	Automatically classifying tennis stroke types and tracking their frequency is crucial for training prescription. This study classified 9 shot types using wrist-worn IMU features from 19 athletes during 66 training sessions, validated with ground truth from recorded video footage.	The top-performing SVM (cubic kernel) achieved precision and recall rates above 97.5% for all shots. Using wrist-mounted IMU signal features, three stroke types can be classified with 97.4% accuracy using a machine learning classifier
Zhang et al. (2021)	Automatic recognition of football player activities during a training session.	A sensor module collects data from acceleration and angular velocity sensors, which is then analyzed by extracting the corresponding feature vector.	The authors recommend the following framework for identifying complex sports activities (excluding walking and stationary actions) using a multilayer decision tree recognizer: 1. Extract acceleration sensor data. 2. Calculate acceleration amplitude. 3. Obtain trough value, peak value, standard deviation, and covariance of acceleration and angular velocity. 4. Use a second layer of recognition for detailed analysis of football players' joint movements during training.

Discussion

Model Selection for Recognition and Classification of Sport Movement

This study aimed to systematically review the literature on ML and DL methods applied to IMU data inputs for evaluating techniques or skills in individual swing and team sports. The outcome variables in all studies were the recognition of a sport-specific movement and an associated analysis of that motion for technique and skill analysis by an ML model. AI can assist with practical issues in training and coaching by processing and analyzing large amounts of data proficiently and identifying patterns of interest related to performance and this can assist in the formation of targeted training and injury prevention programs (Claudino et al., 2019).

The most applied ML model was the SVM (n = 7) due to its ability to discriminate features in large and homogenous datasets, and its repeated accuracy and low computation power requirements for classification and regression tasks. SVM is efficient at classification where there is a clear margin of separation between classes in the feature space (Misra & Li, 2020). Therefore, SVM may be best suited to sport recognition tasks where the activity classes are well-defined and separable, for example, recognition of different stroke types in tennis (Whiteside et al., 2017). Wang et al. (2018) demonstrated the ability of an SVM to classify three different skill levels of volleyball players based on spike actions and achieved a high accuracy of 94%. The use of different kernel functions and the number of features influence the model performance, Shahar et al. (2020) found that the use of a Cubic SVM has the highest accuracy (96.7%) compared to the use of a Fine Gaussian SVM (52.1%) in the classification of 6 field hockey activities measured by 4 IMUs. This corresponds with the results by Preatoni et al. (2020) which examined the use of SVM with several different kernel functions for the classification of 4 functional fitness exercises during a continuous workout and reported that SVM with a cubic kernel based on data from 0.6 second windows with a 10% overlap had the highest accuracy of 97.8%. Radial Basis Function (RBF) kernel might be more suited to activities that are not as well separated due to its ability to capture complex non-linear relationships (Komang, Surya, & Ratna, 2019).

Halilaj et al. (2018) explored the application of ML in human biomechanics and identified model parameters as coefficients in linear regression, weights in artificial neural networks, and support vectors in SVMs. One study found that the best classification accuracy of k-NN for the vector size is equal to 25 features, while the best classification accuracy for the Logistic Regression-based classifier, for the feature vector is equal to 50 (Brzostowski & Szwach, 2018).

DL approaches were commonly reported, where a CNN architecture formed the basis of DL models reported in studies (n = 6). CNN is traditionally considered to be an image-processing algorithm widely adapted for computer vision tasks (Raj & Kos, 2023) however neural network models have shown propitious results in processing time-series data (Abdel-Basset et al., 2021). CNN has also shown favorable results in human activity recognition tasks due to its ability to extract the features automatically from unprocessed data and is outperforming traditional classifier models which require extensive feature engineering and are computationally expensive (Ramamurthy & Roy, 2018).

CNN architectures are built on multiple layers of convolutional operations, at each layer convolutional kernels are applied to extract deep features from the sensor data, this data is then sent to pooling layers to reduce spatial dimensions of the feature maps and reveal underlying patterns (Habibi Aghdam & Jahani Heravi, 2017). CNN is particularly good at inferring long-term repeated activities and automatically learning the deep features contained in the recursive pattern. The results of studies reported in this review suggest that CNN may be preferable in assessing complex spatial and temporal data over long periods of time due to the ability of CNN

to automatically learn hierarchical representations. However, in scenarios where computational efficiency and robustness to noise are a priority, SVM and HMM may be advantageous.

A sequential CNN was used to train its signal parameters on general information of a broad range of catches and drops in a wide range of conditions rather than defining specific actions (Hollaus et al., 2020). CNN is equipped to handle more complex activity recognition tasks involving intricate spatial patterns and diverse movement patterns of the upper and lower limbs, such as recognition of different shooting postures in basketball (Fan et al., 2021). With only accelerometer data, a novel CNN was able to distinguish nine badminton activities with 86% precision when using a sampling frequency of 50 Hz, the addition of gyroscope data leads to 99% precision (Steels et al., 2020).

DL offers many advantages in sensor-based activity recognition compared to traditional supervised ML models such as SVM, HMM which rely on heuristic and hand-crafted features requiring human expertise and domain knowledge. DL models can automatically extract complex features and can exploit unlabeled samples for data training, this is particularly advantageous when attempting to monitor athletic movement in real-life or live scenarios (Wang et al., 2019).

The use of traditional supervised ML models inclusive of DT (Zhang et al., 2021), RF (Cust et al., 2021; Reilly et al., 2021), HMM (Wang et al., 2016), GBM (van den Tillaar et al., 2021) and Logistic Regression (Brzostowski & Szwach, 2018) were reported in this review with high accuracies highlighting the effectiveness of traditional ML models for sport activity recognition tasks.

The selection of an appropriate ML or DL model is crucial because assumptions made about the model's performance potential may prove inaccurate if it cannot perform well with a given dataset. The "no free lunch" (NFL) theorem for supervised learning (Wolpert & Macready, 1997) explains that no one ML algorithm is universally the best fit for all classification problems. The best framework will depend on the specific characteristics of the task being performed, such as the device specification (sampling frequency), the dataset characteristics (sample size, number of classes, complexity of the features and the distribution of the data), and the performance evaluation metrics. The implications of the NFL theorem were shown in studies reported in this review. In studies which analyzed more than one ML model for performance comparison, it was found that the best-performing algorithm depended on the characteristics of the datasets. For tennis stroke detection and classification, a k-NN model outperformed LR for datasets of professional players, but LR performed best for datasets for amateur players, and datasets combining both amateur and professional players (Brzostowski & Szwach, 2018). Several ML algorithms were applied to accelerometer and gyroscope data to automatically classify locomotion in netball, SVM outperformed all other models inclusive of an ensemble, ANN, linear discriminate, decision tree and k-NN model (Smith & Bedford, 2020).

There are other key considerations in selecting an ML model based on the type of data, number of features, and data distribution. A CNN model is commonly implemented due to its ability to extract patterns and representations with precision and accuracy. However, the classification accuracy of CNN models increases when the size of the training data increases, and as such neural networks tend to perform better on large homogenous datasets, such as those often seen in studies implementing computer vision and human pose estimation models. When the number of features is small, regression models are often used as they perform well in cases where the dataset is linearly separable. The Akaike information criterion (AIC) (Akaike, 1973) and the Bayesian information criterion (BIC) (Schwarz, 1978) are statistical tools for comparative evaluation among time series models and can be used in model selection based on the relative quality of statistical models for a given set of data.

Data Acquisition and Device Specifications

Experimental setup varied throughout all studies as outlined in Table 5. All datasets reported in this review were primary datasets captured by IMU inputs from respective study participants in their natural sporting environment. Wang et al (2018) used an SVM model to assess skill level based on spike actions in volleyball and gathered just 100 datasets from 10 participants. Although these researchers achieved a high level of accuracy (94%), the results of this study may not be as dependable as studies which utilized larger datasets. In tennis shot classification, a DNN achieved the greatest accuracies when a rich, homogenous dataset was used (Ganser et al., 2021).

Varying use of single and multiple body-mounted sensors was reported in the methodology. Placing multiple sensors at key joint locations can lead to increased accuracy of event detection and model prediction (Serpush et al., 2022), but it is also associated with increased computational loads, pre-processing demands, and modeling complexity. The influence of sensor combination and location for activity recognition in field hockey was examined (Shahar et al., 2020) and found that optimal outcome was achieved through the fusion of sensors on the chest, waist, right and left wrist which lead to 96.7% accuracy of a cubic SVM.

The choice of sensor placement and combination is task-specific and must be considered with relevance to capturing critical movements and minimization of signal interference. For example, when analyzing kicking technique in Australian rules football, one sensor placed at the ankle was sufficient (Cust et al., 2021), whereas the analysis of 6 different basketball activities involving upper and lower limbs and different movement phases would be indicative of requiring more sensors (Bo, 2022). One study reported that the highest accuracy of SVM for predicting the biomechanics of a bowling arm was achieved when joint sensor data from three sensors were combined (Salman et al., 2017). For individual swing sports where there is a greater emphasis on swing motion and club trajectory (Matsumoto et al., 2020), the use of just one sensor located at the wrist (Ganser et al., 2021) and arm (Whiteside et al., 2017) produced favorable accuracy.

In studies where more than one sensor was used, certain sensor locations which were proximate to the desired activity accelerations were significant, for example, the acceleration magnitude of the right wrist was used to determine a threshold to detect strokes in badminton (Wang et al., 2016). Out of the 15 studies in this review which use only one body-mounted IMU sensor favorable model performance accuracies were reported ranging from 80% (Reilly et al., 2021) to 96% (Ganser et al., 2021). An optimal sensor placement will be unobtrusive in nature while preserving prespecified performance criteria for a given task (Chaves, Barontini, Mendes, Compán, & Lourenço, 2023).

Sampling frequencies between 100 - 499 Hz are typically reported in studies conducting motion analysis in sports (Hribernik, Umek, Tomažič, & Kos, 2022). The influence of sampling frequency on the performance of a CNN model for badminton stroke recognition was examined (Steels et al., 2020), and reported a stable trend between 25 Hz and 100 Hz, but a significant drop was observed at 12.5 Hz. Overall, these authors recommended a sampling frequency of 25 Hz to maintain a balance between computational load and accuracy of predictions (Steels et al., 2020). An in-depth tutorial on the use of body-worn inertial sensors for human activity recognition has been previously published (Bulling, Blanke, & Schiele, 2014).

Device specifications including the gyroscope, accelerometer, and magnetometer range (◦/s) and sampling rate (Hz) should be reported when using IMUs to obtain sensor-derived parameters for sports performance evaluation (Chambers et al., 2015), and a rationale given for the chosen sampling frequency to give context to the reader (Gómez-Carmona, Rojas-Valverde, Rico-González, Ibáñez, & Pino-Ortega, 2020). Methodological differences were shown to be a

limitation in determining the validity and reliability of different IMUs for analysis of sprinting (Macadam, Cronin, Neville, & Diewald, 2019). Moreover, methodological specifications predetermined by the manufacturer are likely to be divergent from the research conditions in which sports movement analysis is conducted (Van der Kruk & Reijne, 2018).

Data Preprocessing and Feature Extraction

Data pre-processing is a crucial stage in the HAR framework, and a clearly defined and well-orchestrated pre-processing of data contributes to the outcome of the ML and DL algorithm performance. For activity recognition tasks involving signals, filtering is applied in the data preprocessing stage to denoise sensor data (Davila, Cretu, & Zaremba, 2017). The studies in this review report cutoff frequencies ranging from 0.5 Hz to 20 Hz. Low-pass filtering was primarily applied to the time-domain as a smoothing function and to remove noise from the signal (Gjoreski & Gams, 2011). A recent review on the challenges in sensor-based HAR found that the Butterworth (low-pass filter) had the best performance for real-time applications, but the waveform delay can be significant depending on the cut-off frequency (Das Antar, Ahmed, & Ahad, 2019).

High-pass filtering was reported in 3 studies and was used to remove unwanted low-frequency signal components while attaining higher-frequency components. High-pass filtering may be more suited to activity recognition tasks involving sudden changes in movement due to its ability to remove the gravitational component from the signal and leave only the dynamic acceleration from the detected activity, for example in the detection of COD in netball a 0.5 Hz high pass filter was used (Smith & Bedford, 2020). When choosing which filter to apply there are many important considerations including the cutoff frequency and the trade-off between noise reduction and signal distortion, which could lead to a loss of important information (de Cheveigné & Nelken, 2019).

After filtering the signal, data detection techniques such as windowing are applied to extract relevant features from the signal (Lara & Labrador, 2013). A range of window sizes were used in the studies ranging from 10 seconds (McGrath et al., 2019; McGrath et al., 2021) to 0.5 seconds (Jowitt et al., 2020; Salim et al., 2020; Whiteside et al., 2017). The choice of window size depends on the complexity of the classification problem and the specific characteristics of the activity being analyzed, the recognition of more dynamic and complex activities should undergo a more advanced segmentation process due to the non-cyclical nature of recognition (Banos et al., 2014). In handball, a 6 second window was used to capture the entire throwing motion in combination with a wide range of signal features allowing for detailed extraction of kinematic information associated with throwing accuracies (van den Tillaar et al., 2021). A 0.5 second window was used to capture more rapid movements associated with volleyball action recognition (Salim et al., 2020). One study in this review reported that when the window size is increased the classification performance of k-NN and Logistic Regression classifiers are also increased (Brzostowski & Szwach, 2018).

Feature engineering is the transformation of sensor data into features suitable for analysing and it is typically used to derive lower-dimensional representations of the data when datasets are high-dimensional. The features that are necessary to classify the activities of interest were calculated from the distinct and overlapping time frame windows. Signal features derived from the X, Y, and Z axes of the accelerometer and gyroscope showed the highest importance for accurately detecting events and increasing the mean accuracy of the ML model performance in studies reported in this review (Gençoğlu & Gümüş, 2020; McGrath et al., 2019; McGrath et al., 2021; Reilly et al., 2021; van den Tillaar et al., 2021). Whereas magnetometer and accelerometer sensor data seem to be of the highest significance for increasing the mean accuracy of the neural

network for the classification of catch and jump activities in American Football (Hollaus et al., 2020). The addition of gyroscope data increased the precision of a CNN for distinguishing activities in badminton from 86% to 99% when gyroscope data was added (Steels et al., 2020). A study on the influence of feature selection methods on exercise classification with IMU data (O'Reilly et al., 2017) showed that optimization of classification and feature selection techniques resulted in an increase in the F1-Score of between 1-8%. There are several important factors to consider when selecting an ML model for feature selection such as the type of data, the amount of data, correlation between features and the complexity of the task (Chen, Dewi, Huang, & Caraka, 2020).

The most reported features were extracted from the time domain (e.g., maximum, minimum, mean, median, standard deviation, kurtosis, and skewness). Time-domain features are frequently used due to their ability to reduce complexity and effectiveness in capturing important and interpretable characteristics of the signal (Dargie, 2009). Features extracted from the frequency domain such as mean and median frequency, spectral entropy, spectral energy, and magnitude of Fourier transform coefficients) were also reported in this review (Smith & Bedford, 2020; Wang et al., 2016).

Time-domain features generally capture the overall dynamics and shape of the signal and prove to be most efficacious in detecting activities from accelerometer signals (Erdaş, Atasoy, Açıcı, & Oğul, 2016), and this may be enough to define an activity class (Chong, Tjurin, Niemelä, Jämsä & Farrahi, 2021). Rosati, Balestra, and Knaflitz (2018) compared feature sets for activity recognition and reported that the time-domain feature set allowed for easier interpretation of complex biomechanical changes. However, the addition of frequency-domain features can reveal underlying oscillations or spectral patterns (Tran, Guan, Pilantanakitti, & Cohen, 2014) and can lead to a more comprehensive overview of the activity.

In the automatic classification of rapid motion in Netball, a tree discriminant SVM performed best when the model was trained with time-domain features, and this was attributed to direct comparisons of peak accelerations between take-off and landing which characterize the jumping activities class (Smith & Bedford, 2020). A tree discriminant SVM is superior in identifying low entropy regions and classifies opposite class instances close to decision boundaries (Chau, Li, & Yu, 2014), and time-domain features can help to reveal where data points from different activities overlap or are in proximity. Time domain features derived from the accelerometer y-axis and gyroscope x-axis were the most significant in predicting approach type, throw type, and ball velocity in experienced handball players (van den Tillaar et al., 2021).

For HAR problems, such as those associated with sport motion analysis, high-dimensional datasets are becoming increasingly common and may result in increased model learning time and complexity of the data processing (Ray, Alshouiliy, & Agrawal, 2021). PCA is a technique that is used to reduce the dimensionality of large datasets, by creating new uncorrelated variables information loss is minimized while the variance in the dataset is maximized (Jolliffe & Cadima, 2016). PCA is reported in two studies for dimensionality reduction (Brzostowski & Szwach, 2018) and classification of skill level (Wang et al., 2018).

The Fourier Transform (FT) is the foundation of discrete signal analysis due to its ability to deal with linear time-invariant operators, FT deals with the transform for continuous-time, infinite-duration signals, and for discrete-time, finite-duration signals (Discrete Fourier Transform). By FT, a continuous-time signal or discrete sequence can be uniquely represented by its equally spaced sample values. FFT was used in the Mel-Frequency Cepstral Coefficients (MFCC) calculation for feature extraction of the time domain, frequency domain, and time-frequency domain features (Brzostowski & Szwach, 2018). Discrete Wavelet Transform (DWT) is the most used method in the filtering of time-frequency. A key feature of this wavelet transform is its

temporal resolution which captures both frequency and location information. There has been a much-documented success in using DWT for extracting discriminative features for the purpose of activity classification (Ahmadi et al., 2014; Mitchell, Monaghan, & O'Connor, 2013; Preece, Goulermas, Kenney, & Howard, 2009). DTW is used in time-series analysis tasks for measuring similarity or optimal alignment between two temporal (time-dependent) sequences (Müller, 2007). DTW was reported in one study for shot detection in swing sports (Anand et al., 2017).

Model Performance Evaluation

Evaluation metrics are used to determine the overall performance and generalization accuracy of the ML model and are generally conducted at the end of the ML pipeline. The most common metric of evaluation reported in studies in this review was accuracy, which represents the number of correct predictions made as a ratio to all predictions made. The extensive use of accuracy as the main evaluation metric for ML and DL algorithms is evident, but it is not necessarily the most appropriate in all situations, especially in cases where the dataset is unbalanced (Khan, Azal, Chaudhari, & Chandra, 2024). In such cases, additional metrics like precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC-ROC) may provide a more comprehensive evaluation of the model's performance (Jeni, Cohn, & De La Torre, 2013).

Included in the findings of this review are confusion matrices, which are visual representations of prediction results on a classification problem and illustrate how they compare to their true classes (Al-jabery, Obafemi-Ajayi, Olbricht, & Wunsch 2020). Additionally, precision (Brzostowski & Szwach, 2018; Kautz et al., 2017), recall (Whiteside et al., 2017), and F1 score (Wang et al., 2018) were other commonly reported measures in studies throughout this review. Precision equals the number of true positives divided by the sum of true positives and false positives, while recall is determined by dividing the number of true positives divided by the sum of true positives and false negatives. The F1-score (also known as f-score or f-measure), represents the harmonic mean between recall and precision values (Vakili, Ghamsari, & Rezaei, 2020). Sensitivity (also known as Recall) and specificity were metrics that are commonly utilized together, as they are inversely related to one another: as sensitivity increases, specificity tends to decrease and vice versa. The specificity metric is used to measure the fraction of negative patterns that are correctly classified (Hossin & Sulaiman, 2015).

The Receiver Operating Curve (ROC) and Area Under the Curve (AUC) were used to evaluate the classifier performance and to discriminate an optimal solution during the classification training. The AUC value exhibits the overall ranking performance of the classifier by denoting the overall sum of positive and negative examples respectively (Hossin & Sulaiman, 2015). Although the performance of AUC is high for the evaluation and discrimination of the classifier, the computational cost is high, especially when generating a solution for a multiclass problem. Mean Absolute Error (MAE) was used to calculate the distance between predicted and ground truth and was reported in this review in 2 studies.

The training and testing process is critical in the success of an ML model in terms of robustness, accuracy, and generalizability (Uçar, Nour, Sindi & Polat, 2020). Splitting the data into training and testing data subsets using techniques such as k-fold cross-validation and LOOCV (Bragança, Colonna, Oliveira& Souto, 2022) can facilitate model evaluation and the optimization or fine-tuning of hyperparameters (Elgeldawi, Sayed, Galal, & Zaki, 2021).

Configuration variables that are intrinsic to the model are often referred to as model parameters and are estimated based on training data. The choice of which cross-validation method is used for the training and testing process will depend on the size and distribution of the dataset. LOOCV may be more suited to datasets which are difficult to classify, and this is because LOOCV can hierarchically rank the models based on performance, additionally, LOOCV is

likely to underestimate the accuracy of a model when the dataset is simple to classify (Pardakhti, Mandal, Ma, & Yang, 2021).

A 10-fold LOOCV was used in the multilevel classification of movement type, COD, and intensity in netball (Smith & Bedford, 2020) where the dataset was likely to exhibit complexity, high dimensionality, and class imbalance. k-Fold cross-validation may be more suited to datasets that are larger in size and representative distributions. The process involves dividing the datasets into k-disjoint subsets (folds) of equal size (Refaeilzadeh, Tang, Liu, 2009) which allows for a more comprehensive evaluation of the model's performance across the various folds of data. Stratified k-Fold cross-validation is useful for handling imbalanced datasets (Szeghalmy & Fazekas, 2023).

When using k-Fold cross-validation the choice of detection methods and particularly overlapping windows is of importance, because data samples are shared across folds which may result in increased dependencies between testing and training sets, this would violate the assumptions of k-Fold cross-validation and could lead to overestimation of the classification performance (Dehghani, Glatard, & Shihab, 2019). Insufficient training data may reduce the model's ability to attain a proper distribution trend and may lead to poor generalizability to new data (overfitting) in neural networks and failing to capture the underlying pattern (underfitting) in SVM models (Kulsoom et al., 2022).

Conclusion

The findings of this systematic review indicate that employing ML and DL algorithms for processing the large volume of data resulting from IMU devices can lead to accurate and objective assessment of individual swing and team sport performance. In addressing the research questions, SVMs formed the base of the methodology, followed by CNNs in combination with different model architectures. In sports performance analysis, individual swing sports tend to involve repetitive high-impact movements, whereas team sports often exhibit non-linear complex patterns of movement. CNNs have shown promising results for processing sequential time-series data, as evidenced by the application in analyzing individual swing sport movements as per three studies in this review.

All steps of the experimental setup such as data pre-processing steps, feature engineering, and sampling frequency of the IMU devices should be selected in alignment with the targeted sport movement characteristics. In scenarios where specific movements are classified through multilevel classification, such as stroke type classification for monitoring hitting load and skill assessment based on specific actions, various experimental considerations are suggested throughout the studies in this review. These considerations include the use of a low-pass such as Butterworth to isolate frequency components of the signal, employing a sliding window to segment continuous data, the extraction of frequency-domain and time-domain features using techniques such as Fourier Transform, and utilizing k-Fold cross-validation to assess the generalizability of the model. F1-score, precision and recall assessing the model's performance of classifying different activities classes e.g. shot type. This methodology may be particularly useful in individual swing sports (e.g. tennis and badminton) where there may be a focus on identifying characteristics of repetitive swing motion, or in team sports where there is a focus on classifying actions according to skill or quality.

It is challenging to provide definitive methodological recommendations regarding the use of AI for analyzing movements in individual swing sports or team sports because each of these considerations is intricately tied to the specific task at hand, or the movement to be analyzed. While future researchers can gain some perspective from studies reported in this review

regarding device specifications (sampling frequency, sensor combination and location) signal processing techniques (filtering, feature extraction methods, feature selection, detection methods such as window conditions and thresholding) and model evaluation metrics, careful consideration should be taken based on specific requirements of the task.

Each step of the ML pipeline should be approached iteratively, considering the knowledge gained from the proceeding steps. For instance, the choice of appropriate filter would depend on the characteristics of the signal following data collection, and dimensionality reduction should be performed once the structure of the dataset is understood. Conducting various window conditions, resampling the signal frequency at different rates, applying multiple machine learning models, and utilizing a range of performance evaluation metrics presents as a thorough method for assessing the effects of methodological conditions on the movement data as highlighted by studies reported in this review.

References

Abdel-Basset, M., Hawash, H., Chakrabortty, R. K., Ryan, M., Elhoseny, M., & Song, H. (2021). STDeepHAR: Deep Learning Model for Human Activity Recognition in IoHT Applications. *IEEE Internet of Things Journal, 8*(6), 4969–4979. https://doi.org/10.1109/JIOT.2020.3033430

Ahmad, N., Ghazilla, R.A., Khairi, N.M., & Kasi, V. (2013). Reviews on Various Inertial Measurement Unit (IMU) Sensor Applications. *International Journal of Signal Processing Systems, 1*(2), 256-262. https://doi.org/10.12720/ijsps.1.2.256-262

Ahmadi, A., Mitchell, E., & Destelle, F., Gowing, M., O'Connor, N., Richter, C., & Moran, K. (2014). Automatic Activity Classification and Movement Assessment During a Sports Training Session Using Wearable Inertial Sensors. *Proceedings. 11th International Conference on Wearable and Implantable Body Sensor Networks,* 98-103. https://doi.org/10.1109/BSN.2014.29

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & B. F. Csaki (Eds.), *Second International Symposium on Information Theory, 267–281.* Academiai Kiado: Budapest. https://doi.org/10.1007/978-1-4612-0919-5_38

Alanen, A., Räisänen, A., Benson, L., & Pasanen, K. (2021). The use of inertial measurement units for analyzing change of direction movement in sports: A scoping review. *International Journal of Sports Science & Coaching, 16*(6)*,* 1332 - 1353. https://doi.org/10.1177/17479541211003064

Al-jabery, K. K., Obafemi-Ajayi, T., Olbricht, G. R., & Wunsch II, D. C. (2020). Data analysis and machine learning tools in MATLAB and Python. In K. K. Al-jabery, T. Obafemi-Ajayi, G. R. Olbricht, & D. C. Wunsch II (Eds.), *Computational Learning Approaches to Data Analytics in Biomedical Applications* (pp. 231-290). Academic Press. ISBN 9780128144824. https://doi.org/10.1016/B978-0-12-814482-4.00009-7

Alzubi, J.A., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. *Journal of Physics: Conference Series,* 1142. https://doi.org/10.1088/1742-6596/1142/1/012012

Anand, A., Sharma, M., Srivastava, R., Kaligounder, L., & Prakash, D. (2017). Wearable Motion Sensor Based Analysis of Swing Sports. *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2017, 261-267. https://doi.org/10.1109/ICMLA.2017.0-149

Angerbauer, S., Palmanshofer, A., Selinger, S., & Kurz, M. (2021). Comparing Human Activity Recognition Models Based on Complexity and Resource Usage. *Applied Sciences*, *11*(18), 8473. https://doi.org/10.3390/app11188473

Ann, O. C. & Theng, L.B. (2014). Human activity recognition: A review. *IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014)*, 389-393. https://doi.org/10.1109/ICCSCE.2014.7072750

Aresta, S., Bortone, I., Bottiglione, F., Di Noia, T., Di Sciascio, E., Lofù, D., Musci, M., Narducci, F., Pazienza, A., Sardone, R., & Sorino, P. (2022). Combining Biomechanical Features and Machine Learning Approaches to Identify Fencers' Levels for Training Support. *Applied Sciences*, *12*(23), 12350. https://doi.org/10.3390/app122312350

Baca, A., Dabnichki, P., Hu, C. W., Kornfeind, P., & Exel, J. (2022). Ubiquitous Computing in Sports and Physical Activity-Recent Trends and Developments. *Sensors*, *22*(21), 8370. https://doi.org/10.3390/s22218370

Banos, O., Galvez, J. M., Damas, M., Pomares, H., & Rojas, I. (2014). Window Size Impact in Human Activity Recognition. *Sensors*, *14*(4):6474-6499. https://doi.org/10.3390/s140406474

Batool, M., Jalal, A., & Kim, K. (2020). Telemonitoring of Daily Activity Using Accelerometer and Gyroscope in Smart Home Environments. *Journal of Electrical Engineering and Technology, 15*, 2801–2809. https://doi.org/10.1007/s42835-020-00554-y

Bo, Y. (2022). A reinforcement learning-based basketball player activity recognition method using multisensors. *Mobile Information Systems, 2022*, 1-9. https://doi.org/10.1155/2022/6820073

Bonidia, R.P., Rodrigues, L.A., Avila-Santos, A.P. Sanches, D.S., Brancher, J.D. & Mustapha, A. (2018). Computational Intelligence in Sport: A Systematic Literature Review. *Advances in Human-Computer Interaction, 2018.* https://doi.org/10.1155/2018/3426178

Bragança, H., Colonna, J. G., Oliveira, H. A. B. F., & Souto, E. (2022). How Validation Methodology Influences Human Activity Recognition Mobile Systems. *Sensors (Basel, Switzerland)*, *22*(6), 2360. https://doi.org/10.3390/s22062360

Brzostowski, K., & Szwach, P. (2018). Data fusion in ubiquitous sports training: Methodology and application, *Wireless Communications and Mobile Computing*, *2018*, 1-14. https://doi.org/10.1155/2018/8180296

Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys, 46*(3), 1-33. https://doi.org/10.1145/2499621

Camomilla, V., Bergamini, E., Fantozzi, S., & Vannozzi, G. (2018). Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: A systematic review. *Sensors*, *18*(3), 873. https://doi.org/10.3390/s18030873

Chai, J., Zeng, H., Li, A., & Ngai, E. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications, 6*, 100134. https://doi.org/10.1016/j.mlwa.2021.100134

Chambers, R., Gabbett, T. J., Cole, M. H., & Beard, A. (2015). The Use of Wearable Microsensors to Quantify Sport-Specific Movements. *Sports medicine (Auckland, N.Z.)*, *45*(7), 1065–1081. https://doi.org/10.1007/s40279-015-0332-9

Chandrashekar, G. & Sahin, F. (2014) A Survey on Feature Selection Methods. *Computers and Electrical Engineering*, *40*(1), 16-28. https://doi.org/10.1016/j.compeleceng.2013.11.024

Chau, A.L., Li, X., & Yu, W. (2014). Support vector machine classification for large datasets using decision tree and Fisher linear discriminant. *Future Generation Computer Systems, 36*, 57-65.

Chaves, E., Barontini, A., Mendes, N., Compán, V., & Lourenço, P. B. (2023). Methodologies and challenges for optimal sensor placement in historical masonry buildings. *Sensors*, *23*(23), 9304. https://doi.org/10.3390/s23239304

Chen, R.C., Dewi, C., Huang, S., & Caraka, R.E. (2020). Selecting critical features for data classification based on machine learning methods. *Journal of Big Data, 7*, 1-26. https://doi.org/10.1186/s40537-020-00327-4

Chmait, N., & Westerbeek, H. (2021). Artificial Intelligence and Machine Learning in Sport Research: An Introduction for Non-data Scientists. *Frontiers in sports and active living*, *3*, 682287. https://doi.org/10.3389/fspor.2021.682287

Chong, J., Tjurin, P., Niemelä, M., Jämsä, T., & Farrahi, V. (2021). Machine-learning models for activity class prediction: A comparative study of feature selection and classification algorithms. *Gait & Posture*, *89*, 45-53. https://doi.org/10.1016/j.gaitpost.2021.06.017

Claudino, J. G., Capanema, D. O., de Souza, T. V., Serrão, J. C., Machado Pereira, A. C., & Nassis, G. P. (2019). Current Approaches to the Use of Artificial Intelligence for Injury Risk Assessment and Performance Prediction in Team Sports: a Systematic Review. *Sports medicine - open*, *5*(1), 28. https://doi.org/10.1186/s40798-019-0202-3

Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. (2018). A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System. *Sports medicine - open*, *4*(1), 24. https://doi.org/10.1186/s40798-018-0139-y

Crenna, F., Rossi, G. B., & Berardengo, M. (2021). Filtering Biomechanical Signals in Movement Analysis. *Sensors*, 21(13), 4580. https://doi.org/10.3390/s21134580

Currell, K., & Jeukendrup, A. E. (2008). Validity, reliability and sensitivity of measures of sporting performance. *Sports medicine (Auckland, N.Z.)*, *38*(4), 297–316. https://doi.org/10.2165/00007256-200838040-00003

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. https://doi.org/10.1080/02640414.2018.1521769

Cust, E. E., Sweeting, A. J., Ball, K., & Robertson, S. (2021). Classification of Australian football kick types in-situation via ankle-mounted inertial measurement units. *Journal of sports sciences*, *39*(12), 1330–1338. https://doi.org/10.1080/02640414.2020.1868678

Dargie, W. (2009). Analysis of Time and Frequency Domain Features of Accelerometer Measurements. *In 2009 Proceedings of 18th International Conference on Computer*

Communications and Networks (pp. 1-6). San Francisco, CA, USA. https://doi.org/10.1109/ICCCN.2009.5235366

Das Antar, A., Ahmed, M., & Ahad, M. A. R. (2019). Challenges in Sensor-based Human Activity Recognition and a Comparative Analysis of Benchmark Datasets: A Review. *In 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR) (pp. 134-139).* Spokane, WA, USA. https://doi.org/10.1109/ICIEV.2019.8858508

Davila, J. C., Cretu, A. M., & Zaremba, M. (2017). Wearable Sensor Data Classification for Human Activity Recognition Based on an Iterative Learning Framework. *Sensors*, *17*(6), 1287. https://doi.org/10.3390/s17061287

de Cheveigné, A., & Nelken, I. (2019). Filters: When, Why, and How (Not) to Use Them. *Neuron*, *102*(2), 280-293. https://doi.org/10.1016/j.neuron.2019.02.039

Dehghani, A., Glatard, T., & Shihab, E. (2019). Subject Cross Validation in Human Activity Recognition. *ArXiv, abs/1904.02666*. https://doi.org/10.48550/arXiv.1904.02666

Dwyer, D. B., Kempe, M., & Knobbe, A. (2022). Editorial: Using Artificial Intelligence to Enhance Sport Performance. *Frontiers in sports and active living*, *4*, 886730. https://doi.org/10.3389/fspor.2022.886730

Elgeldawi, E., Sayed, A., Galal, A. R., & Zaki, A. M. (2021). Hyperparameter Tuning for Machine Learning Algorithms Used for Arabic Sentiment Analysis. *Informatics*, *8*(4), 79. https://doi.org/10.3390/informatics8040079

Erdaş, Ç. B., Atasoy, I., Açıcı, K., & Oğul, H. (2016). Integrating Features for Accelerometer-based Activity Recognition. *Procedia Computer Science*, *98*, 522-527. https://doi.org/10.1016/j.procs.2016.09.070

Fan, J., Bi, S., Wang, G., Zhang, L., & Sun, S. (2021). Sensor Fusion Basketball Shooting Posture Recognition System Based on CNN. *Journal of Sensors,* 1-16. https://doi.org/10.1155/2021/6664776Ganser, A., Hollaus, B., & Stabinger, S. (2021). Classification of Tennis Shots with a Neural Network Approach. *Sensors (Basel, Switzerland)*, *21*(17), 5703. https://doi.org/10.3390/s21175703

Gençoğlu, C., & Gümüş, H. (2020). Standing Handball Throwing Velocity Estimation with a Single Wrist-Mounted Inertial Sensor. *Annals of Applied Sport Science, 8*, 0-0. https://doi.org/10.29252/AASSJOURNAL.893

Gómez-Carmona, C.D., Rojas-Valverde, D., Rico-González, M., Ibáñez, S.J., & Pino-Ortega, J. (2020). What is the most suitable sampling frequency to register accelerometry-based workload? A case study in soccer. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology, 235*, 114 - 121. https://doi.org/10.1177/1754337120972516

Gjoreski, H., & Gams, M. (2011). Accelerometer data preparation for activity recognition. *In Proceedings of the International Multiconference Information Society, Ljubljana, Slovenia*, 1014.

Gupta, N., Gupta, S. K., Pathak, R. K., Jain, V., Rashidi, P., & Suri, J. S. (2022). Human activity recognition in artificial intelligence framework: a narrative review. *Artificial intelligence review*, *55*, 4755 – 4808. https://doi.org/10.1007/s10462-021-10116-x

Habibi Aghdam, H., Jahani Heravi, E. (2017). Convolutional Neural Networks. In: Guide to Convolutional Neural Networks. 85-130. *Springer, Cham.* https://doi.org/10.1007/978-3-319-57550-6_3

Halilaj, E., Rajagopal, A., Fiterau, M., Hicks, J. L., Hastie, T. J., & Delp, S. L. (2018). Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities. *Journal of biomechanics*, *81*, 1–11. https://doi.org/10.1016/j.jbiomech.2018.09.009

Hribernik, M., Umek, A., Tomažič, S., & Kos, A. (2022). Review of Real-Time Biomechanical Feedback Systems in Sport and Rehabilitation. *Sensors*, *22*(8), 3006. https://doi.org/10.3390/s22083006

Host, K., & Ivašić-Kos, M. (2022). An overview of Human Action Recognition in sports based on Computer Vision. *Heliyon*, *8*(6), e09633. https://doi.org/10.1016/j.heliyon.2022.e09633

Hossin, M., & Sulaiman, M. N. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process, 5*(2), 01-11. https://doi.org/10.5121/ijdkp.2015.5201

Hollaus, B., Stabinger, S., Mehrle, A., & Raschner, C. (2020). Using Wearable Sensors and a Convolutional Neural Network for Catch Detection in American Football. *Sensors (Basel, Switzerland)*, *20*(23), 6722. https://doi.org/10.3390/s20236722

Islam, M. M., Nooruddin, S., Karray, F., & Muhammad, G. (2022). Human activity recognition using tools of convolutional neural networks: A state of the art review, data sets, challenges, and future prospects. *Computers in Biology and Medicine*, *149*. https://doi.org/10.1016/j.compbiomed.2022.106060

Jalal, A., Quaid, M. A. K., Tahir, S. B. u. d., & Kim, K. (2020). A Study of Accelerometer and Gyroscope Measurements in Physical Life-Log Activities Detection Systems. *Sensors (Basel, Switzerland)*, *20*(22), 6670. https://doi.org/10.3390/s20226670

Jeni, L. A., Cohn, J. F., & De La Torre, F. (2013). Facing Imbalanced Data Recommendations for the Use of Performance Metrics. *International Conference on Affective Computing and Intelligent Interaction and workshops : [proceedings]. ACII (Conference), 2013,* 245–251. https://doi.org/10.1109/ACII.2013.47

Jha, A., Dave, M., & Madan, S. (2019) Comparison of Binary Class and Multi-Class Classifier Using Different Data Mining Classification Techniques, *Proceedings of International Conference on Advancements in Computing & Management (ICACM) 2019.* http://doi.org/10.2139/ssrn.3464211

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, *374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Jowitt, H. K., Durussel, J., Brandon, R., & King, M. (2020). Auto detecting deliveries in elite cricket fast bowlers using microsensors and machine learning. *Journal of sports sciences*, *38*(7), 767–772. https://doi.org/10.1080/02640414.2020.1734308

Jung, A. (2022) *Machine Learning: The Basics*, (1st ed) Springer, Singapore (Chapter 1).

Kautz, T., Groh, B.H., Hannink, J., Jensen, U., Strubberg, H., & Eskofier, B.M. (2017). Activity recognition in beach volleyball using a Deep Convolutional Neural

Network. *Data Mining and Knowledge Discovery, 31*, 1678 - 1705. https://doi.org/10.1007/s10618-017-0495-0

Khan, A., Nicholson, J., & Plötz, T. (2017). Activity Recognition for Quality Assessment of Batting Shots in Cricket using a Hierarchical Representation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 1*(3), 1 - 31. https://doi.org/10.1145/3130927

Khan, A., Azal, M., Chaudhari, O., & Chandra, R. (2024). A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation. *Expert Systems with Applications, 244*, 122778. https://doi.org/10.1016/j.eswa.2023.122778

Kim, M., & Park, S. (2020). Golf Swing Segmentation from a Single IMU Using Machine Learning. *Sensors*, *20*(16):4466. https://doi.org/10.3390/s20164466

Kim, W., & Kim, M. (2017). Sports motion analysis system using wearable sensors and video cameras. *2017 International Conference on Information and Communication Technology Convergence (ICTC)*, 1089-1091. https://doi.org/10.1109/ICTC.2017.8190863

Kok, M., Hol, J., & Schön, T. (2017). Using Inertial Sensors for Position and Orientation Estimation. *Foundations and Trends in Signal Processing, 11*(1-2),1-153. https://doi.org/10.1561/9781680833577

Komang, M.G., Surya, M.N., & Ratna, A.N. (2019). Human activity recognition using skeleton data and support vector machine. *Journal of Physics: Conference Series, 1192*. https://doi.org/10.1088/1742-6596/1192/1/012044

Kulsoom, F., Narejo, S., Mehmood, Z., Chaudhry, H. N., Butt, A., & Bashir, A. K. (2022). A review of machine learning-based human activity recognition for diverse applications. *Neural Computing & Applications, 34*, 18289–18324. https://doi.org/10.1007/s00521-022-07665-9

Lara, O. D., & Labrador, M. A. (2013). A Survey on Human Activity Recognition using Wearable Sensors. *IEEE Communications Surveys & Tutorials*, *15*(3), 1192-1209. https://doi.org/10.1109/SURV.2012.110112.00192

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine*, *6*(7), e1000100. https://doi.org/10.1371/journal.pmed.1000100

Macadam, P., Cronin, J., Neville, J., & Diewald, S. (2019). Quantification of the validity and reliability of sprint performance metrics computed using inertial sensors: A systematic review. *Gait & posture*, *73*, 26–38. https://doi.org/10.1016/j.gaitpost.2019.07.123

Mathis, A., Schneider, S., Lauer, J., & Mathis, M. W. (2020). A Primer on Motion Capture with Deep Learning: Principles, Pitfalls, and Perspectives. *Neuron*, *108*(1), 44–65. https://doi.org/10.1016/j.neuron.2020.09.017

Matsumoto, K., Tsujiuchi, N., Ito, A., Kobayashi, H., Ueda, M., & Okazaki, K. (2020). Proposal of Golf Swing Analysis Method Using Singular Value Decomposition. Proceedings, *49*(1), 91. https://doi.org/10.3390/proceedings2020049091

McGrath, J., Neville, J., Stewart, T., Clinning H. & Cronin, J. (2021). Can an inertial measurement unit (IMU) in combination with machine learning measure fast bowling speed and perceived intensity in cricket? *Journal of Sports Sciences, 39*(12), 1402-1409. https://doi.org/10.1080/02640414.2021.1876312

McGrath, J.W., Neville, J., Stewart, T. & Cronin., J. (2019). Cricket fast bowling detection in a training setting using an inertial measurement unit and machine learning*, Journal of Sports Sciences*, *37*(11), 1220-1226. https://doi.org/10.1080/02640414.2018.1553270

Misra, S., & Li, H. (2020). Chapter 9 - Noninvasive fracture characterization based on the classification of sonic wave travel times. *In S. Misra, H. Li, & J. He (Eds.), Machine Learning for Subsurface Characterization (pp. 243-287).* Gulf Professional Publishing. ISBN 9780128177365. https://doi.org/10.1016/B978-0-12-817736-5.00009-0

Mitchell, E., Monaghan, D., & O'Connor, N. E. (2013). Classification of sporting activities using smartphone accelerometers. *Sensors (Basel, Switzerland)*, *13*(4), 5317–5337. https://doi.org/10.3390/s130405317

Moran, K., Ahmadi, A., Richter, C., Mitchell, E., Kavanagh, J., O'Connor, N. (2015). Automatic Detection, Extraction, and Analysis of Landing During a Training Session, Using a Wearable Sensor System. *Procedia Engineering*, *112,* 184-189. https://doi.org/10.1016/j.proeng.2015.07.197

Müller, M. (2007). Dynamic Time Warping. In R. Baeza-Yates & B. Ribeiro-Neto (Eds.), Information Retrieval for Music and Motion (pp. 69 - 84). https://doi.org/10.1007/978-3-540-74048-3_4

O'Reilly, M.A., Johnston, W., Buckley, C., Whelan, D.F., & Caulfield, B.M. (2017). The influence of feature selection methods on exercise classification with inertial measurement units. *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 193-196. https://doi.org/10.1109/BSN.2017.7936039

Pardakhti, M., Mandal, N., Ma, A.W., & Yang, Q. (2021). Practical Active Learning with Model Selection for Small Data. *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 1647-1653. https://doi.org/10.1109/ICMLA52953.2021.00263

Preatoni, E., Nodari, S., & Lopomo, N. F. (2020). Supervised Machine Learning Applied to Wearable Sensor Data Can Accurately Classify Functional Fitness Exercises Within a Continuous Workout. *Frontiers in bioengineering and biotechnology*, *8*, 664. https://doi.org/10.3389/fbioe.2020.00664

Preece, S. J., Goulermas, J. Y., Kenney, L. P., & Howard, D. (2009). A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. *IEEE transactions on bio-medical engineering*, *56*(3), 871–879. https://doi.org/10.1109/TBME.2008.2006190

Raj, R., & Kos, A. (2023). An improved human activity recognition technique based on convolutional neural network. *Scientific Reports*, *13*, 22581. https://doi.org/10.1038/s41598-023-49739-1

Ramamurthy, S.R., & Roy, N. (2018). Recent trends in machine learning for human activity recognition—A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8*. https://doi.org/10.1002/widm.1254

Rana, M. & Mittal, V (2021). Wearable Sensors for Real-Time Kinematics Analysis in Sports: A Review, in *IEEE Sensors Journal*, *21*(2), 1187-1207. https://doi.org/10.1109/JSEN.2020.3019016

Ray, S., Alshouiliy, K., & Agrawal, D. P. (2021). Dimensionality Reduction for Human Activity Recognition Using Google Colab. *Information*, *12*(1), 6. https://doi.org/10.3390/info12010006

Refaeilzadeh, P., Tang, L., Liu, H. (2009). Cross-Validation. In: LIU, L., ÖZSU, M.T. (eds) Encyclopedia of Database Systems. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-39940-9_565

Reilly, B., Morgan, O, Czanner, G. & Robinson, M.A. (2021). Automated Classification of Changes of Direction in Soccer Using Inertial Measurement Units. *Sensors*, *21*(14), 4625. https://doi.org/10.3390/s21144625

Richter, C., O'Reilly, M., & Delahunt, E. (2021). Machine learning in sports science: challenges and opportunities. *Sports biomechanics*, 1–7. Advance online publication. https://doi.org/10.1080/14763141.2021.1910334

Roggio, F., Ravalli, S., Maugeri, G., Bianco, A., Palma, A., Di Rosa, M., & Musumeci, G. (2021). Technological advancements in the analysis of human motion and posture management through digital devices. *World Journal of Orthopedics*, *12*(7), 467–484. https://doi.org/10.5312/wjo.v12.i7.467

Rosati, S., Balestra, G., & Knaflitz, M. (2018). Comparison of Different Sets of Features for Human Activity Recognition by Wearable Sensors. *Sensors (Basel, Switzerland)*, *18*(12), 4189. https://doi.org/10.3390/s18124189

Salim, F. A., Haider, F., Postma, D. B. W., van Delden, R., Reidsma, D., Luz, S., & van Beijnum, B. J. F. (2020). Towards Automatic Modeling of Volleyball Players' Behavior for Analysis, Feedback, and Hybrid Training. *Journal for the Measurement of Physical Behaviour*, *3*(4), 323-330. https://doi.org/10.1123/jmpb.2020-0012

Salman, M., Qaisar, S.B., & Qamar, A.M. (2017). Classification and legality analysis of bowling action in the game of cricket. *Data Mining and Knowledge Discovery, 31*, 1706-1734. https://doi.org/10.1007/s10618-017-0511-4

Schwarz, G. (1978). Estimating the Dimensions of a Model. *The Annals of Statistics,* 6(2), 461-464.

Serpush, F., Menhaj, M. B., Masoumi, B., & Karasfi, B. (2022). Wearable Sensor-Based Human Activity Recognition in the Smart Healthcare System. *Computational intelligence and neuroscience*, *2022*, 1391906. https://doi.org/10.1155/2022/1391906

Shahar, N., Ghazali, N. F., As'ari, M. A., & Swee, T. T. (2020). Wearable inertial sensor for human activity recognition in field hockey: Influence of sensor combination and sensor location. *In: 2nd Joint International Conference on Emerging Computing Technology and Sports, JICETS 2019*, 25 - 27 November 2019, Bandung, Indonesia. https://doi.org/10.1088/1742-6596/1529/2/022015

Smith, P., & Bedford, A. (2020). Automatic Classification of Locomotion in Sport: A Case Study from Elite Netball. *International Journal of Computer Science in Sport*, *19*(2), 1-20. https://doi.org/10.2478/ijcss-2020-0007

Steels, T., Van Herbruggen, B., Fontaine, J., De Pessemier, T., Plets, D., & De Poorter, E. (2020). Badminton Activity Recognition Using Accelerometer Data. *Sensors*, *20*(17), 4685. https://doi.org/10.3390/s20174685

Szeghalmy, S., & Fazekas, A. (2023). A Comparative Study of the Use of Stratified Cross-Validation and Distribution-Balanced Stratified Cross-Validation in Imbalanced Learning. *Sensors, 23*(4), 2333. https://doi.org/10.3390/s23042333

Tan, F., & Xie, X. (2021). Recognition Technology of Athlete's Limb Movement Combined Based on the Integrated Learning Algorithm. *Journal of Sensors, 2021*, 1-9. https://doi.org/10.1155/2021/3057557

Tran, A., Guan, J., Pilantanakitti, T., & Cohen, P.R. (2014). Action Recognition in the Frequency Domain. *ArXiv, abs/1409.0908*. https://doi.org/10.48550/arXiv.1409.0908

Uçar, M. K., Nour, M., Sindi, H., & Polat, K. (2020). The Effect of Training and Testing Process on Machine Learning in Biomedical Datasets. *Mathematical Problems in Engineering*, *2020*, 2836236. https://doi.org/10.1155/2020/2836236

Vaibhaw, Sarraf, J., & Pattnaik, P.K. (2020). Chapter 2 - Brain–computer interfaces and their applications. In V.E. Balas, V.K. Solanki, & R. Kumar (Eds.), An Industrial IoT Approach for Pharmaceutical Industry Growth (pp. 31-54). Academic Press. ISBN 9780128213261. https://doi.org/10.1016/B978-0-12-821326-1.00002-4

Vakili, M., Ghamsari, M.K., & Rezaei, M. (2020). Performance Analysis and Comparison of Machine and Deep Learning Algorithms for IoT Data Classification. *ArXiv, abs/2001.09636*. https://doi.org/10.48550/arXiv.2001.09636

van den Tillaar, R., Bhandurge, S., & Stewart, T. (2021). Can Machine Learning with IMUs Be Used to Detect Different Throws and Estimate Ball Velocity in Team Handball? *Sensors*, *21*(7), 2288. https://doi.org/10.3390/s21072288

van der Kruk., E, & Reijne., M. M. (2018) Accuracy of human motion capture systems for sport applications; state-of-the-art review. *European Journal of Sport Science, 18*(6), 806-819. https://doi.org/10.1080/17461391.2018.1463397

Van Eetvelde, H., Mendonça, L. D., Ley, C., Seil, R., & Tischer, T. (2021). Machine learning methods in sport injury prediction and prevention: a systematic review. *Journal of experimental orthopaedics*, *8*(1), 27. https://doi.org/10.1186/s40634-021-00346-x

Vrigkas, M., Nikou, C., & Kakadiaris, I. A. (2015). A Review of Human Activity Recognition Methods. *Frontiers in Robotics and AI*, 2(28). https://doi.org/10.3389/frobt.2015.00028

Wang, Z., Guo, M., & Zhao, C. (2016). Badminton stroke recognition based on body sensor networks. IEEE Transactions on Human-Machine Systems, 46(5), 769-775. https://doi.org/10.1109/THMS.2016.2571265

Wang, Y., Zhao, Y., Chan, R.H., & Li, W.J. (2018). Volleyball Skill Assessment Using a Single Wearable Micro Inertial Measurement Unit at Wrist. *IEEE Access, 6*, 13758-13765. https://doi.org/10.1109/ACCESS.2018.2792220

Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep Learning for Sensor-based Activity Recognition: A Survey. *ArXiv, abs/1707.03502*. https://doi.org/10.1016/j.patrec.2018.02.010

Whiteside, D., Cant, O., Connolly, M., & Reid, M. (2017). Monitoring Hitting Load in Tennis Using Inertial Sensors and Machine Learning. *International journal of sports*

physiology and performance, *12*(9), 1212–1217. https://doi.org/10.1123/ijspp.2016-0683

Wickramasinghe, I. (2022). Applications of Machine Learning in cricket: A systematic review. *Machine Learning with Applications*, *10*, 10043510. https://doi.org/10.1016/J.MLWA.2022.100435

Wohlin, C. (2014). Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering. *in Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering. London, England, United Kingdom: Association for Computing Machinery*, 1–10. https://doi.org/10.1145/2601248.2601268

Wolpert, D. H., & Macready, W.G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1), 67-82. https://doi.org/10.1109/4235.585893

Worsey, M. T., Espinosa, H. G., Shepherd, J. B., & Thiel, D. V. (2019). Inertial Sensors for Performance Analysis in Combat Sports: A Systematic Review. *Sports (Basel, Switzerland)*, *7*(1), 28. https://doi.org/10.3390/sports7010028

Yeo, S. S., & Park, G. Y. (2020). Accuracy Verification of Spatio-Temporal and Kinematic Parameters for Gait Using Inertial Measurement Unit System. *Sensors (Basel, Switzerland)*, *20*(5), 1343. https://doi.org/10.3390/s20051343

Yin, C., Chen, J., Miao, X., Jiang, H., & Chen, D. (2021). Device-Free Human Activity Recognition with Low-Resolution Infrared Array Sensor Using Long Short-Term Memory Neural Network. *Sensors, 21*(10), 3551. https://doi.org/10.3390/s21103551

Zecha, D., Einfalt, M., Eggert C. & Lienhart, R. (2018) Kinematic Pose Rectification for Performance Analysis and Retrieval in Sports, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1872-18728. https://doi.org/10.1109/CVPRW.2018.00232

Zhang, B., Lyu, M., Zhang, L. & Yang. W. (2021). Artificial Intelligence-Based Joint Movement Estimation Method for Football Players in Sports Training. *Mobile Information Systems,* Vol. 2021. https://doi.org/10.1155/2021/9956482

Zhou, L., Fischer, E., Tunca, C., Brahms, C.M., Ersoy, C., Granacher, U., & Arnrich, B. (2020). How We Found Our IMU: Guidelines to IMU Selection and a Comparison of Seven IMUs for Pervasive Healthcare Applications. *Sensors*, *20*(15), 4090. https://doi.org/10.3390/s20154090

Zhu, J., San-Segundo, R. & Pardo, J. (2017). Feature extraction for robust physical activity recognition. *Human-centric Computing and Information Sciences*, *7*(1). https://doi.org/10.1186/s13673-017-0097-2