Biomedical Signal Processing and Control 71 (2022) 103197

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Wearable sensors and machine learning in post-stroke rehabilitation assessment: A systematic review

Issam Boukhennoufa [a], Xiaojun Zhai [a,*], Victor Utti [b], Jo Jackson [b], Klaus D. McDonald-Maier [a]

[a] School of Computer Science and Electronic Engineering, University of Essex, United Kingdom
[b] School of Sport, Rehabilitation and Exercise Sciences, University of Essex, United Kingdom

A R T I C L E I N F O

Keywords:
Stroke rehabilitation
Wearable sensors
Machine learning
Feature engineering
Systematic review

A B S T R A C T

A cerebrovascular accident or stroke is the second commonest cause of death in the world. If it is not fatal, it can result in paralysis, sensory impairment and significant disability. Rehabilitation plays an important role to help survivors relearn lost skills and assist them to regain independence and thus ameliorate their quality of life. With the development of technology, researchers have come up with new solutions to assist clinicians in monitoring and assessing their patients; as well as making physiotherapy available to all. The objective of this review is to assess the recent developments made in the field of post-stroke rehabilitation using wearable devices for data collection and machine learning algorithms for the exercises' evaluation. To do so, PRISMA guidelines for systematic reviews were followed. Scopus, Lens, PubMed, ScienceDirect and Microsoft academic were electronically searched. Peer-reviewed papers using sensors in post-stroke rehabilitation were included, for the period between 2015 to August 2021. Thirty-three publications that used wearable sensors for patients' assessment were included. Based on that, we have proposed a taxonomy that divided the assessment systems into three categories namely activity recognition, movement classification, and clinical assessment emulation. Moreover, The most commonly employed sensors as well as the most targeted body–limbs, outcome measures, and study designs are reviewed, in addition to the examination of the machine learning approaches starting from the feature engineering to the classification done. Finally, limitations and potential study directions in the field are presented.

1. Introduction

Worldwide, there are more than 13.7 million episodes of stroke each year, with a quarter of the over 25 population experiencing it in their lifetime [1]. A stroke is a brain attack that occurs when blood flow is cut off to a part of the brain, subsequently resulting in the death of brain cells [2,3]. There are three main types of stroke [4]: Transient Ischemic Attack (TIA) [5], ischemic stroke [6], and hemorrhagic stroke [7].

1. TIA is caused by a temporary interruption to the blood supply to the brain and may result in no lasting neurological deficit, it is considered to be a precursor and warning of a future stroke.
2. Ischemic stroke which is estimated at 87 per cent of strokes [8], occurs when a blood vessel supplying blood to the brain is obstructed.
3. Hemorrhagic stroke happens when a blood vessel ruptures [9].

Brain damage caused by stroke - if not deadly - will influence how the body functions including instigating temporary or permanent paralysis [10,11]. Subsequently, some stroke survivors will make a quick recovery, while others will need help and more time to recuperate, and relearn skills they lost [12,13].

To speed up the process of recovery, and to regain their independence, post-stroke patients ought to engage in physical therapy or rehabilitation [14,15]. The conventional approach is for physical therapists to evaluate physical activities of patients through visual observation, clinical impression, or tests and measures [16–18]. Rehabilitation activities might include:

- Motor skill exercises: to ameliorate the strength of the muscles and body coordination [19].
- Mobility training: in order to relearn functional activities including walking which may include the use of, mobility aids, such as walkers, wheelchairs and canes to help support the body's weight [20].

* Corresponding author.
E-mail address: xzhai@essex.ac.uk (X. Zhai).

https://doi.org/10.1016/j.bspc.2021.103197
Received 30 June 2021; Received in revised form 28 August 2021; Accepted 30 September 2021
1746-8094/

- Constraint-induced rehabilitation or forced-use therapy: to improve limb function, where the patients practise using the affected limb while the unaffected one is held still [21].
- Active or passive Range Of Motion (ROM): to help patients regain the ROM of the affected body joints [22].

However, this approach presents many limitations [23], indeed the availability of therapy may be limited and the patients need regular consultations in order to achieve their goals [24], moreover the additional expense of public and private transport from and to hospitals are an additional burden to the patients' finances [25]. Also, transportation to hospitals may cause discomfort and pain to post-stroke patients who lack the mobility and energy to leave their houses and periodically visit their doctors for training sessions [26]. Besides, doctors and therapists are overwhelmed with the workload with sessions lasting more than half an hour - on average - with a cadence of many sessions per week [27].

To tackle these issues, researchers have developed applications to assess rehabilitation outcomes using novel technologies namely "wearable sensors" [28], which provide a high level of portability and low price giving researchers and therapists a plethora of possibilities and solutions [29]. Indeed, wearable sensors allow patients to execute their exercises at home relieving them of the drain of transportation. Subsequently, several types of sensing devices are used in applications extending from monitoring subjects' physiologic responses like Electromyography (EMG) [30], Electrocardiogram (ECG) [31], or glucose level in the blood [32] to evaluating kinematics of the individuals: gait, ROM, balance using Inertial Measurement Units (IMU) [33]. These sensors are employed in conjunction with clinical tests and outcome measures, such as sit-to-stand [34], Timed Up and Go (TUG) [35] to give an objective assessment and monitoring of the patient condition [36].

Besides, the breakthrough in Machine Learning (ML) that provide outstanding performance tasks that used to require a lot of knowledge and time to model [37], as well as the tremendous advances made in processing system technologies that made the ML computing possible have given researchers more tools and resources to handle and process the data collected from the sensors and hence permitting a more accurate and quicker assessment [38].

The objective of this paper is to assess the progress made in the domain of stroke rehabilitation and to make a status report of the different technological developments in smart upper and lower limb recovery, with the objective to answer the following questions:

- What are the different aims of the post-stroke rehabilitation systems?
- What wearable sensing devices are more used?
- What are the most common outcome measures and the targeted sensors' placements?
- What are the different study designs followed by the researcher in this field?
- Which ML algorithms and feature engineering techniques were more used?
- What limitations and challenges are encountered by researchers and what are the possible direction to take in this field of study?

In the following section, we introduce the review method used within this study, talk about the procedure for the selection of the papers and present the results of the selection. After that, we give a discussion about the different included papers by surveying the different wearable sensors used, the outcome measures, the types of the assessment systems and the different algorithms. Then, we present the different limitations and challenges encountered in the post-stroke rehabilitation to finally give some tips on potential direction to take to have more effective systems.

2. Review method

2.1. Literature search strategy

A literature search was undertaken using the five following databases: Lens, PubMed, Scopus, ScienceDirect Microsoft academic. Works dealing with a variation of the following aspects were included: Stroke, body part, rehabilitation, sensor, system type, algorithm, and wearable systems. Title and abstract keywords and their synonyms were employed in several combinations for every database with the help of 2dSearch [39] to convert from the different database search-syntaxes. Articles published from January 2015 to August 2021 (The period has been chosen arbitrarily to evaluate recent trends) were reviewed. This search includes English-written peer-reviewed journal papers and conference-proceeding articles only. The search query including the used search terms is listed in Table 1.

2.2. Study selection

The process of selecting articles consisted of following the steps introduced by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [40]. An electronic computerised-search was performed for the last decade (January 2015 - August 2021). Using reference manager software, duplicates were removed, the remaining works were screened by their titles and abstracts. After that, the selected papers were fully read and filtered out using the inclusion/exclusion criteria given in the following. When authors published numerous studies on the same research topic, only the most recent one was retained.

2.2.1. Inclusion criteria

aaa

1. Only web-available journal articles or conference papers were considered.
2. Works published in the period January 2015 and August 2021.
3. The system is intended for lower/upper limb rehabilitation assessment.
4. Works using wearable sensors only for the data collection.
5. Works using machine learning algorithms for the assessment.

2.2.2. Exclusion criteria

aaa

1. Reviews, magazine or book chapter papers.
2. Non-English written articles.

Table 1
Our search query.

Parameter	Search query
Stroke	("stroke " or "post-stroke ") And
Body part	(Lower-body OR upper-body OR "upper body" OR Lower-extremity OR "lower extremity" OR upper-extremity OR "upper extremity" OR Lower-limb OR "lower limb" OR "upper-limb " OR "lower body" OR upper limb) And
Rehabilitation	("rehabilitation" OR "telerehabilitation"OR "physical therapy"OR"telemedicine"OR "neuro-rehabilitation"OR "Motor-recovery") And
Sensor	(" Wearable sensor" OR "wearable device "OR " wearable sensing device "OR "wearable detector "OR "IMU "OR "EMG ""Accelerometer ") And
Type	(" assessment"OR "monitoring "OR " quantification "or "evaluation") And
Algorithm	("machine learning"OR "intelligent system "OR "Deep learning "OR "classification ") and
Undesired results	(Not (" Heat-stroke" or "heat" or "stroke detection" or "robot" or "stroke detection"))

3. Stroke detection systems.
4. Previous works of the same author on the same topic (only most recent is considered).
5. Robotic systems or exoskeleton based systems (considered to be obtrusive and already covered in other review papers).
6. Non-wearable sensors based systems like cameras and radars.

2.3. Results

Initially, an overall 530 articles were identified using our search query, 303 duplicates were excluded either using a reference manager software or manually. The resulting 227 papers were screened based on their titles and abstracts, 82 were selected for full-text evaluation based on our inclusion/exclusion criteria. A total of 33 papers fitted the conditions and were finally retained. Fig. 1 shows the flow of information through the different search phases of this systematic review.

3. Discussion

Study characteristics related to the wearable sensor used and its placement, the monitored exercises, the participants, the selected features and the ML algorithm used and the classification performance for the included papers are presented in Table 2. The studies are divided into three categories based on the assessment type namely activity recognition, movement classification and clinical assessment emulation (explained bellow). After that a more in-depth discussion on each topic is done separately with a quantitative comparison done at the end of this section.

3.1. Assessment systems and outcome measures

In the post-stroke rehabilitation, and based on the reviewed papers we have introduced a new taxonomy gleaned from which we classified the assessments systems in post-stroke rehabilitation. Subsequently, we distinguished three assessment approaches depending on the system's aim: activity recognition, movement classification and clinical assessment emulation.

3.1.1. Activity recognition

Are systems which aim to identify specific movements of rehabilitation of the patients and differentiate between them for record and monitoring purposes [41–51], in this category researchers monitored Activities of Daily Living (ADL) [75] and they most frequently covered detecting general activities like standing, sitting, lying, standing up, sitting down [42,44,47,48,50], performing kitchen tasks like making a drink, chopping food [42] and other routine activities like making the bed, reading and lacing shoes [48], folding, sweeping and brushing teeth [46,48,49]. Other researchers covered activities for specific body parts like recognising different hand gestures [41], arm gestures [43] and some exercises to strengthen shoulders, and arms [48].

3.1.2. Movement classification

The system objective is to classify well and poorly executed tasks [52–61,63,64], to do so many approaches were followed. Some researchers implemented systems to distinguish between normal and abnormal gaits for lower-limb rehabilitation [56,59,60], in which participants executed 10 m walks. Other researchers assessed the execution of ADLs [54,57] like different kitchen related activities or routine bedroom tasks. Moreover, Lee et.al [52] utilised exercises that belong to popular batteries of tests like Fugi Mayer assessment (FMA) [76,77], and

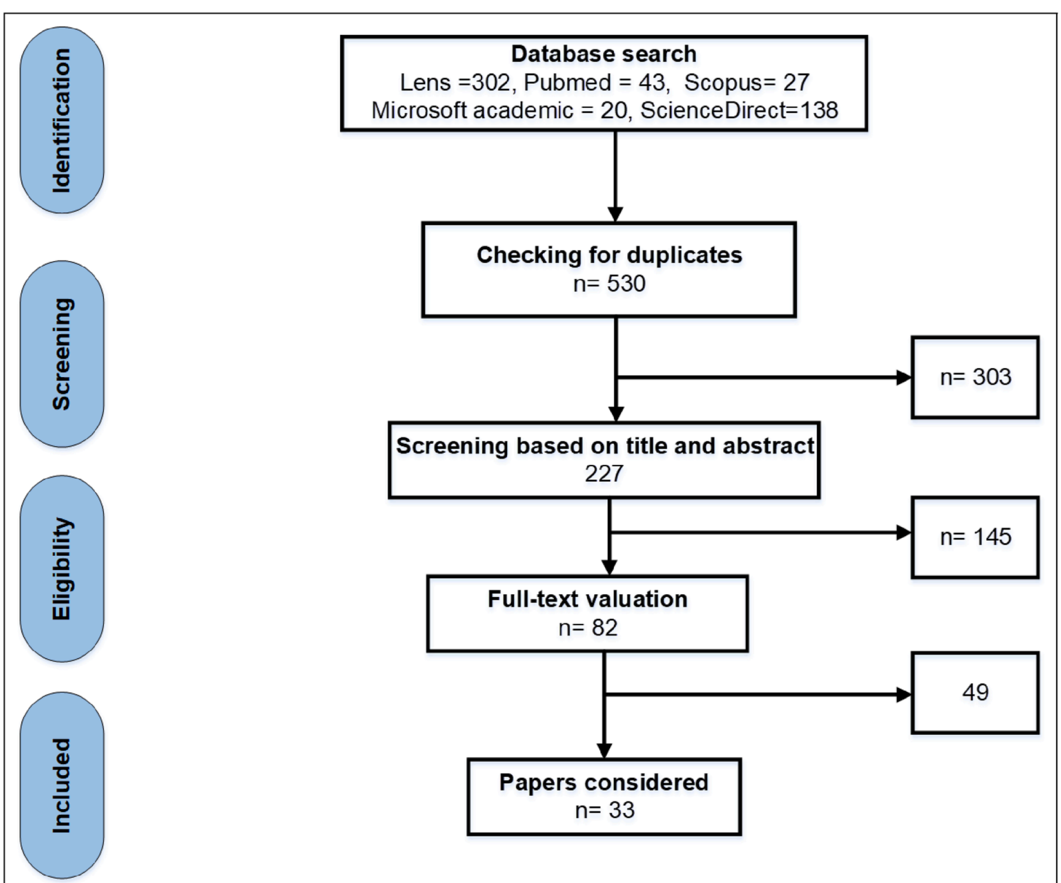

Fig. 1. Review flow chart based on the PRISMA guidelines.

Table 2
Summary of the reviewed papers.

Paper	Sensor/limb	Exercise	Participants	Feature	ML method	Best performance
			Activity recognition			
[41]	EMG/ Forearm	9 different hand gestures	3 AB[1]	Time domain features corresponding to variance, waveform length, root mean square, zero-crossing and auto-regressive coefficients. PCA was then applied to the resulting 56 features vector and the three first components that contributed with 95.86% of the overall information were utilised.	MLP and SVM	Accuracy of 96.25%
[42]	IMU, level sensor/ Hand	ADLs (Walking, standing, sitting, up/down, drinking)	15 AB	Discrete cosine transform on the segmented time series data signals to extract frequency domain features and regroup the energy in the low frequency coefficients.	SVM, MLP	Accuracy of more than 92% for SVM
[43]	Accelerometer/ Arm	ADLs (20 arm movements)	10 SP [2]	The data was segmented to time windows and down sampled and the normalised magnitude of the acceleration was used. The different segments are then labeled according to the activity. Two different configurations were used for the participants: naturalistic data where patients are in their houses and 97.89% on semi-naturalistic data where patients are in labs	CNN	Accuracies of 88.87% on the naturalistic data and 97.89% on the semi-naturalistic data respectively.
[44]	IMU/ Right-front hip	ADLs (41 mobility tasks)	15 AB, 17 El [3], 12 SP	Extracted a number of 76 time series features, relief-F, correlation-based feature selection and fast correlation based filter were then used to select the most relevant features.	Bayes, SVM and RT	Variant for different tasks
[45]	IMU/ Hand and hips	Bilateral shoulder flexion with both hands interlocked; wall push exercise; active scapular exercise; and towel slide exercise.	23 SP	Raw data from gyroscope, accelerometer and the combination of both to Recognise and record the type and frequency of the rehabilitation exercises.	CNN	99.9%
[46]	IMU/ Wrist	ADLs (Doing the laundry, performing kitchen tasks, shopping related tasks, and making the bed.)	10 AB, 10 SP	Extracted overall and axial means, overall and axial variances, entropy, minima, and maxima. The feature vectors were then compressed using PCA to reduce the high from 11 to 3 columns.	K-Means, KNN, RF, SVM, RBF SVM	RF accuracy 83%
[47]	IMU, barometer/ Waist	ADLs (Sitting, Lying, Standing, Stairs Up, Stairs Down, and Walking)	30SP	Features included statistical measures of the sensor signal, its derivatives, and the frequency domain (mean, range,skewness…etc)	RF	Trained on stroke activity achieved 75%
[48]	IMU, barometer/ Sternum	ADLs (sitting, standing, walking, lying,sit-to-stand, stand-to-sit,walking up and down the stairs, taking the elevator, washing hands, eating, pouring and drinking water, sleeping, shoe lacing, reading the newspaper..etc)	12SP	Different algorithms developed from previous researches by detecting transitional phases for different ADLs	Hierarchical Fuzzy Inference System	70.3%
[49]	IMU/ Wrist, arm	ADLs (Chopping food, vacuuming, sweeping, spreading jam or butter, folding laundry, eating, brushing teeth…etc)	11SP	Time series features (mean, standard deviation, autocorrelation, and slope) and frequency domain features (not mentioned)	DT, RF, SVM, and eXtreme Gradient Boosting (XGBoost)	82%
[50]	IMU/ Waist	ADLs (Walking, walking up, walking down, sit to stand, stand to sit, laying)	30AB	Segmented data were encoded into images using GMAF technique	Different CNN models	VGG16 98.53%
[51]	IMU, sEMG/ Wrist, arms forearms, legs, ankles	ADLs (Walking, tooth brush, gace washing, drinking)	9SP, 14AB	Noise was filtered through Butterworth filter and band-pass, then data were normalised, then 28 different time-series features were extracted RMS, meqn, variance…etc	SVM (linear and rbf), Adaboost, KNN, RF, DT, KNN	SVM-rbf 82.47%.
			Movement classification			

(*continued on next page*)

Table 2 (*continued*)

Paper	Sensor/limb	Exercise	Participants	Feature	ML method	Best performance
[52]	IMU/ Wrist	Motor tasks associated with the FMA	20 SP, 10 El	Applied the minimal-redundancy maximal-relevance algorithm on the minimum, maximum, range, mean, standard deviation, RMS values, and the number of zero crossings of the time-series data.	LR, RF	87% and 84.3% successively
[53]	Accelerometer/ Finger and wrist	Estimate amount of hand use	18 AB	Extracted multiple time-series features (mean, inter-quartile range, minimum and maximum, root mean square of the acceleration time-series, standard deviation, ratio of the energy at the dominant frequency to the entire signal, energy of the time-series, skewness, kurtosis, and signal entropy) then a correlation based feature selection was utilised to identify the most relevant features.	SVR	0.11 RMSE
[54]	IMU/ Arms and chest	ADLs (washing the face, applying deodorant, combing the hair, donning and doing glasses, preparing and eating a slice of bread, …etc)	48 SP	Raw data to measure functional primitive	CNN	70 %
[55]	IMU and pressure sensors/ Legs and feet	extension and abduction of the legs, sit-to-stand, gait and Bipodaal Bridge	NA	Extracted 64 features consisting of mean and the variance for the different sensing nodes	TB, RT, hyper-plane, MLP	MLP reported the best F-measure with 97.9%
[56]	IMU/ Shanks	Gait	15 SP	Used a total of 18 features consisting of Hidden Markov Model (Log-likelihood, EL model, Log-likelihood PS model, Log-likelihood, HD model, Difference between log-likelihoods given EL and PS models, Difference between log-likelihoods given EL and HD models, Difference between log-likelihoods given PS and HD models) time (Mean value Evaluated, Standard deviation, Variance, Maximum, Minimum, Range) and frequency domain features (Power at first dominant frequency (P1), Power at second dominant frequency, First dominant frequency, Second dominant frequency, Total power (PT), P1/PT).	SVM	LOSO cross validation and an accuracy of 90.5%
[57]	IMU/ index and finger	9 different ADLs: resting, eating, pouring water, drinking, brushing,folding towel, grasp bottle, grasp brush, and grasp towel.	10AB, 12SP	A low-pass fourth-order Butterworth filter was applied to all the signals to remove the tremor noise. A high pass fourth-order Butterworth filter was implemented for frequency analysis to eliminate the continuous component of the signal. then data was normalised then different features were extracted: skewness, average, RMS, jerk…etc	SVM, ANN	ANN 99.9% for a dataset containing both SP and AB
[58]	IMU/ Wrist, arm, sternum	Uni-manual tasks, bi-manual asymmetric tasks, bi-manual symmetric tasks all performed with dominant and non-dominant hand	20SP, 20AB	Classifier Attribute Evaluator, ReliefF, Info Gain Attribute Evaluator and Gain Ratio Attribute were used to select the most relevant features then Root Mean Square, Mean, Signal Magnitude Area, Signal Vector Magnitude, Energy, Entropy, FFTPeak, and Standard Deviation were then selected.	Bayes,SMO, IBk, KStar, Multiclass Classifier, Bagging, DT, J48 and RF	RF 85%
[59]	IMU/ Lower back, both sides of the thigh, shank, foot	10 m gait	11SP, 9NDP[4]	Data were filtered with a fourth-order bi-directional Butterworth band-pass filter, then minimal peak distance and minimal peak height were applied to the resulting data. after that different gaits parameters were computed.	RF, Adaboost, DT, Gaussian naive bayes, MLP	The shank placement DT 89.13%

(*continued on next page*)

Table 2 (*continued*)

Paper	Sensor/limb	Exercise	Participants	Feature	ML method	Best performance
[60]	IMU/ shank	10 m gait	8SP,7AB	Data were normalised and labeled different gait phases	MLP	99.35%
[61]	IMU, EMG, temperature/ Arms and chest	Flexor synergy, shoulder flexion hand to lumbar pronation and supination	NA	Employed Empirical mode decomposition [62] to partition the times series data into the three first intrinsic mode functions. The mean values and standard deviations of these components are used in conjunction with mean values and standard deviations, entropy and energy of the motion signals as features for large joint actions.	AdaBoost	Accuracy of 99.25%.
[63]	IMU,/ Arms, forearms, thighs	ADLs (e.g. walking, walking up/downstairs, arm and leg flexion/extension, arm rotation, writing, using phone, drinking)	NA	walking-related gait parameters (stride duration, cadence and stride count)	unspecified regression models	
[64]	IMU,/ Arm, forearms, hand	Flexion/extension of the elbow, supination/pronation of the forearm, extension/flexion of the wrist	13AB, 13SP	Linear interpolation was done to synchronise data, then data was normalised	Different KNN models, Different SVM models, Fine tree	Fine KNN 98.5%
			Clinical assessment emulation			
[65]	Accelerometer/ Wrist and the sternum	tasks associated with WFT	34 AB	Segmented time-series data	RF	correlation with therapists scores R^2=0.97
[66]	IMU/ Forearm	ADLs (Doing the laundry, Performing kitchen activities, Shopping, Making the bed.)	10 AB, 10 SP	Extracted entropy, mean, and variance-based measures	Tree based	Accuracy of 88%
[67]	IMU/ Arms and chest	A battery of activities from WMFT	16 SP	Derived the time-series magnitudes of displacement, velocity, acceleration, and jerk to extract multiple time series features i.e.. minimum, maximum, and mean values, root mean-square value, ratio of the magnitude of the dominant frequency and total signal energy, jerk, skewness, signal entropy, kurtosis, correlation coefficients computed for different axes, and duration of the data segments.	RF	0.38 RMSE
[68]	Accelerometer, flex/ Shoulder, elbow, wrist, finger	Seven different exercises based on the short FMA	24SP	Raw sensor data was denoised with 5 point smooth method, and AMP, MEAN, RMS, JERK, and ApEn were extracted and then RRelief algorithm was applied to find the optimal features for each exercise.	ELM,SVM	SVM 92.2%
[69]	IMU/ Sternum, arms, wrist, elbow	Synergy, out of synergy, combination of synergies, wrist/hand function and fine motor coordination	8SP	RMS, mean, entropy, dominant frequency.	DT, Bagging Forest	Bagging Forest reported lowest RMSE
[70]	Accelerometer/ arms, shanks	ADLs (Exercises from the Oxford Grading Motor Scale)	4SP	Gravity component was removed from the norm of the acceleration data then the mean, max, mean, normalized average rectified jerk, powers and frequencies of FFT	SVM	82%
[71]	Accelerometer/ Wrist	ADLs (Exercises from the Oxford Grading Motor Scale)	59SP	signal vector magnitude is computed by substracting the gravity effect from the acceleration, then DWT to extract wavelet coefficients, normalised Sum of Absolute value of DWT coefficients is used as features	LMGP, lSVM, rbf SVM, mlp	LMGP reached RMSE 3.12 for Chronic) and 5.75 for acute
[72]	IMU/ Wrist	grabbing a cube and moving it for an ARAT assesment	34SP	Raw data from IMUs	Matching pursuit	Accuracy of 95 percent
[73]	IMU/ Wrist,sternum	continuous, random, voluntary upper-limb movements spanning the entire range of active motion	23SP	Zero-Crossing Decomposition applied on gravity free acceleration, resulting data is normalised to engineer different features	unspecified regression model	R^2 value of 0.985
[74]	IMU/ Wrist, and feet	Stretch and hold their arms for 20 s, and lift and stretch their left or right leg	15SP	Features related to the degree of drift of the limbs	Ensemble algorithm and SVM	Accuracy of 83.3% for SVM

[1]AB: Able-Bodied; [2]SP: Stroke patients; 3El: Elderly; 4NDP: Neurologically disordered participants

extension or flexion of elbow and flexibility movements of shoulders [55,61].

Movement classification englobes as well systems that quantify limb use in order to classify the tasks, in [58], Miller et.al distinguished between uni-manual and bi-manual tasks using both dominant and non-dominant activities while Liu et.al [53] estimated the amount of the affected hand use compared to the unaffected hand. In [63] Derungs extracted digital biomarkers consisting of convergence points physical activity and functional ROM to investigate the affected and less-affected body side. Whereas, Balestra et.al [64], identified different executed tasks in order to count the number of repetitions and determine a correlation with the degree of severity of stroke.

3.1.3. Clinical assessment emulation

In this category, systems that aim to quantify the level of correctness in executing the prescribed exercises are identified. Researchers achieved this by using popular post-stroke assessment scoring systems [65–74]:

FMA variants are the most commonly used batteries of tests from the included works [68,69,73], it comprises five domains namely motor functioning, balance, sensation, joint functioning and joint pain in both upper and lower extremities rehabilitation. Scale items are scored on the basis of ability to complete the item using a 3-point ordinal scale where a score of 0 means the incapacity to perform, 1 a partial performance and 2 a full performance of the task. The total possible score is 226 divided into 100 points for motor functioning, 14 for balance, 24 for sensation while joint functioning joint pain have 44 points each. Other variants of this assessment were used such as the short FMA used in developed in [77] which includes less exercises than the original.

Wolf Motor Function Test (WMFT) [78,79] is an upper-limb assessment system through timed and functional tasks, the most popular form consists of 17 items in which 6 involve timed functional tasks, 2 are measures of strength, while the remaining consist of analysing the quality of movement quality when performing various activities. It uses a scaling system that ranges from 0 that signifies Does (i.e no attempt with the limb being tested) to 5 that signifies the attempt was made with a normal-appearing movement. Two included studies used the WMFT [65,67,73].

Action Research Arm Test (ARAT) [80] is a 19-item observational measures for upper-limb post-stroke assessment. Items comprising the ARAT are categorised into four subscales namely grasp, grip, pinch and gross movement. Task performance is rated on a 4-point scale, ranging from 0 (no movement) to 3 (movement performed normally). Two of the included works used the ARAT system [66,72]

Oxford Grading Motor-Scale (OGM) [81] used in a single study included [70], it evaluates the muscle strength of the rehabilitated patient and can help diagnose problems in which weakness plays a role. It is not proper to stroke rehabilitation and targets both upper and lower extremities. According to the OGM scale, muscle strength is graded from 0 to 5 where 0 implies no muscle contraction and 5 equals movement through a full range against full resistance. Performing OGM requires knowledge of muscle anatomy so that the joints can be positioned correctly as well as the tendon and muscle palpated in order to make a judgement on how much muscle action can be made on the patient.

Chedoke Arm and Hand Activity Inventory (CAHAI) [82], it is an upper-limb post-stroke clinical assessment method that evaluates functional ability. The original CAHAI involved 13 functional items that and incorporates a range of movements and grasps that reflect stages of motor recovery following stroke. The clinician will score based on patient's performance at a scale from 1 that implies a weak performance to 7 that shows complete independence. From the included works Chen et. al used the CAHAI in [71].

National Institutes of Health Stroke Scale (NIHSS) is a 15-item neurologic scale used to assess the effect of acute cerebral infarction on different levels of consciousness, language, neglect, visual field-loss, extraocular movement, motor strength, ataxia, dysarthria, and sensory loss. Scores range from 0 to 42, with higher scores indicating greater severity. A single included paper [64] used this assessment system. Fig. 2 shows our study taxonomy and the different categories.

3.2. Wearable sensors

Over the past few years, effort has been put into developing unobtrusive, effective and objective motion-modeling systems, taking advantage of the progress made in the sensor technology which became more compact and more power-efficient [83]. All the included works utilised IMUs for the data acquisition [42–50,52–61,65–70,63,71–74,64]. IMUs are devices that combine linear acceleration from accelerometer and the angular turning rates from gyroscopes [84]. IMUs were chosen for their portability and for their low-costs, but also because they provide accurate modeling of the participant motion. Some studies used individual accelerometers [43,45,53,54,65,67,68,70,71] or gyroscopes [45] while the rest used their combination to give more detailed information. Moreover, IMUs were coupled with different sensors to acquire more information: a barometric pressure sensor to detect changes in altitude [47,48], insole pressure sensors in [55] to measure the force exercised by the feet while performing the activities, flex sensors to measure the amount of deflection or bending while griping objects [68], liquid level detectors in a cup [42] to measure drinking activity and EMG sensors [61,51] to measure the activity of the muscles that can translate as strength. Only a single study did not use IMUs and employed EMG sensors only [41].

3.3. Sensors' placements

Placement of the sensing technology on the body has shown a heterogeneous distribution linked to the different nature of the employed technology and to the purpose for which the monitoring system was designed. Systems that focused on upper-limb rehabilitation used more frequently the wrists [46,49,52,53,58,65,68,69,71–74] in twelve studies, arms [43,49,54,58,61,67,69,70,63,64] in ten studies, four studies used forearm [41,66,63,64], three studies used fingers [53,57,68], and hands [42,45,64], and a single study used elbow and shoulder [68]. These placements were targeted to monitor activities that involved using hands. By contrast, systems that focused on lower limbs for activities that involved walking utilised more frequently the chest in eight studies [48,54,58,61,65,67,69,73] the shank in four works [56,59,60,70], thighs [55,59,63], and feet [55,59,74] in three studies, while two works targeted either the hip [44,45], or the waist [47,50], and finally a single study used the lower back [59]. Fig. 3 shows the targeted placements reviewed in the different studies.

3.4. Study designs and populations

In the included works, different study settings were explored. More commonly it was in controlled environments [85] like labs and hospitals were patients are under direct supervision of the researchers and therapists. Other studies used semi-naturalistic environments [86] where a home-environment are replicated in the labs e.g participants performing their exercises in a kitchen environment under supervision of researchers. Other studies monitored participants in an outpatient home environment [63]. For the study population, many works recruited stroke survivors with different degrees of severity after getting ethical approvals [43–49,52,54,56–60,66–70,63,71–74,64]. Some of them undertook a cohort study by combining them with Able Bodied (AB) participants [43–46,52,57,58,60,67,64] elderly [52] and neurologically disordered patients [59]. Other works only used AB [41,42,50], while some studies did not specify or did not use participants [61,53]. For the number of participants it varied from 4 SP [70] to 59 SP [71].

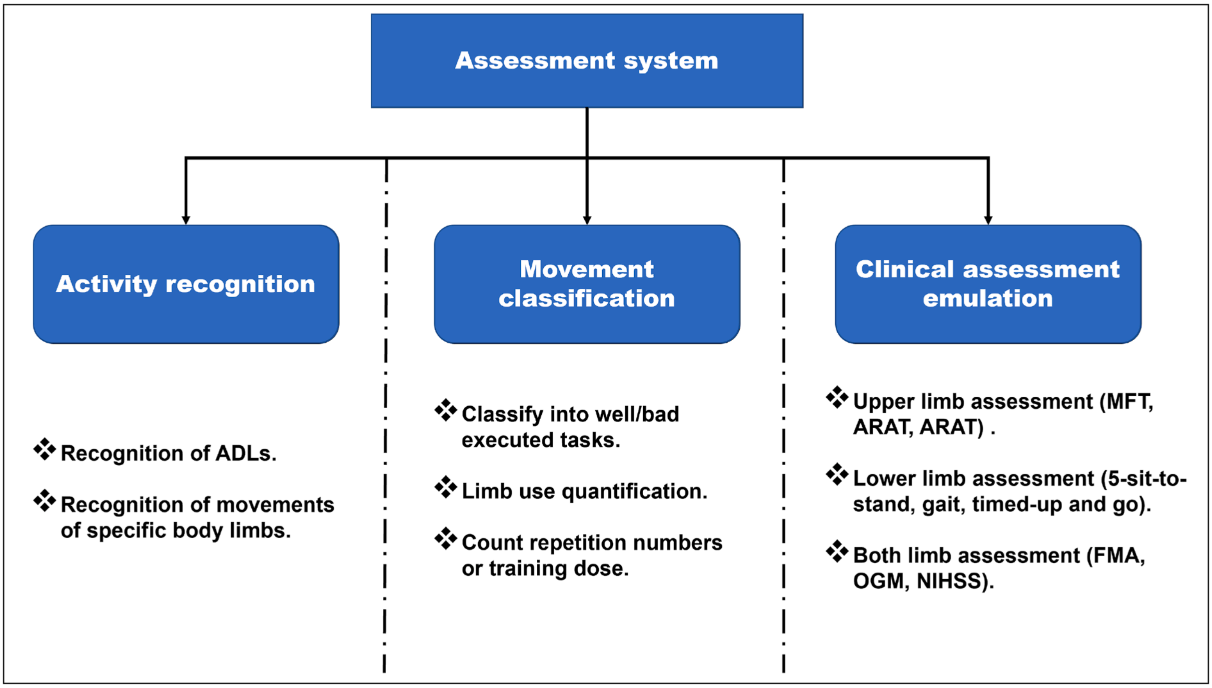

Fig. 2. Proposed taxonomy for post-stroke systems' classification.

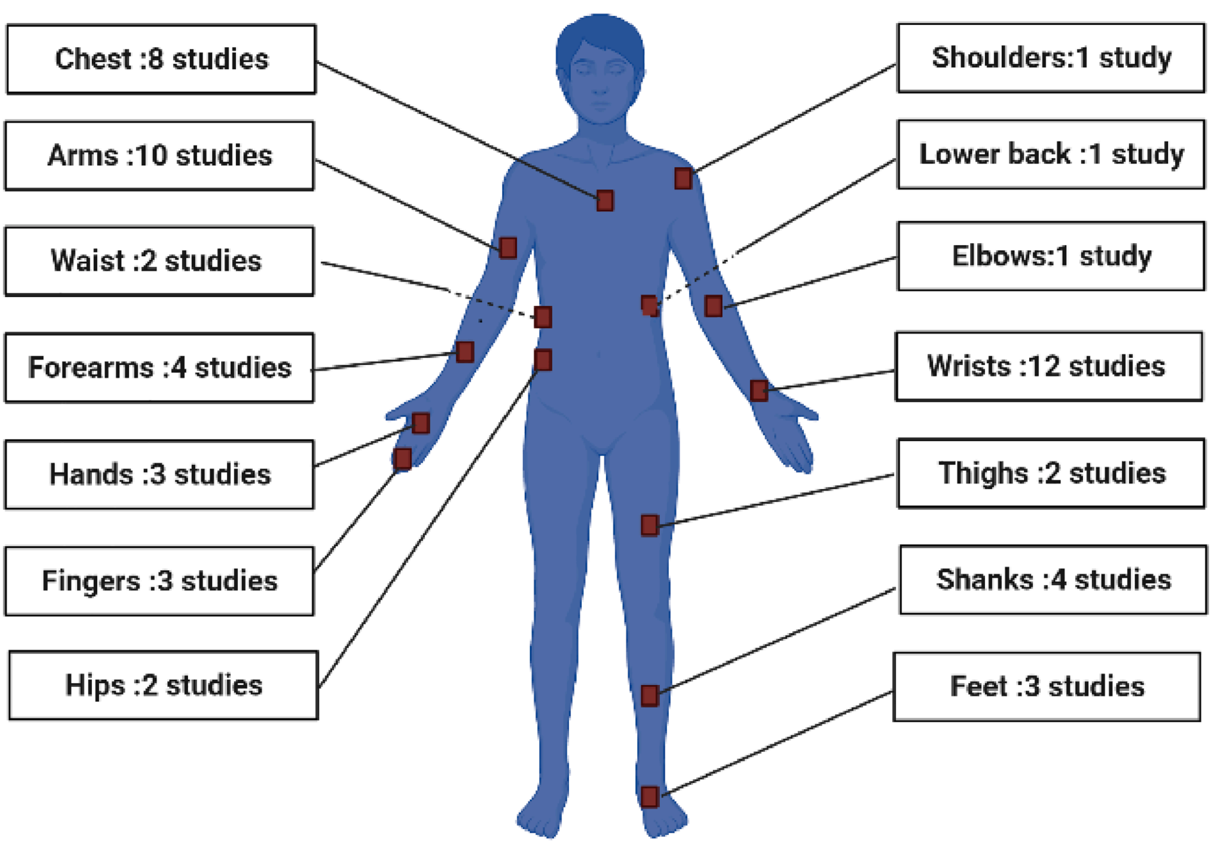

Fig. 3. Sensors' placements in the included studies.

3.5. Pre-processing and feature engineering

Feature engineering is the process of creating features from raw data to improve the accuracy of a system [87]. Some sophisticated ML algorithms i.e deep learning don't require features and can learn to find similarities and differences in raw data automatically [88].Prior to selecting features, pre-processing is undertaken on the data in order to make it ready for the feature study.

According to the different included papers, for filtering unwanted data, designed modules have usually applied threshold-based methods to filter sensor data [43,44] or used different statistical tools to interpolate the missing data points [64]. Moreover to filter frequency based noise, in frequency domain, other methods are applied such as power spectral density (PSD) [70,56] Fast Fourier Transforms (FFT) [70,58], as well as designing different filtering to remove the fluctuations in sensor signals. For example, in [44,57] noise and unwanted information is filtered out by a low-pass fourth-order Butterworth filter, after that a high pass fourth-order Butterworth filter was implemented for frequency analysis to eliminate the continuous component of the signal. In [59] Hsu et.al filtered data with a fourth-order bi-directional Butterworth band-pass filter.

Moreover When dealing with accelerometer data, gravity is usually removed from the acceleration as done in [71,73,64] by computing the magnitude of acceleration $a(t)$ and subtracting 1. $a(t) = \sqrt{a_x^2(t) + a_y^2(t) + a_z^2(t)}$, in which a_x, a_y, a_z is the acceleration along the x, y, z. The gravity effect can be removed by $VM = |a - 1|$. Compared with raw acceleration triaxial data, VM is insensitive to the gravity effect. In addition, using multiple data sources and thus different sampling rates requires data to be synchronised, in order to have the same time basis, this has been done by first identifying segments from timestamps and then using linear interpolation as in [64] or padding with zero [61] on the lower frequency data source.

Since the data collected from wearable sensors is time-series, it should be structured in order to be studied. Time series segmentation can be considered either as a pre-processing step for variety of data mining tasks or as trend analysis techniques. It is also considered as a discretisation problem [89]. A fixed length window is used to segment a time series into sub-sequences and the time series is then represented by the primitive shape patterns that are formed [90]. Segmentation was used by all the papers included herein with time windows varying from 2 s to 10 s depending on the monitored activities. For deep learning algorithms data after these steps is ready to be fed [45,54,60,65].

By contrast, conventional ML algorithms require further data processing and features that most describe the activities are selected and extracted. In most of the included papers, feature engineering is handcrafted based on the authors' knowledge of the human movements. Time-domain based features [91] were the most commonly used approach, numerous works extracted Root Mean Square (RMS), mean of time windows, variances, correlatiion between different axes and features, minima and maxima, skewness and other related features [41,44,46,47,49,52,53,55,57,58,61,66–70,73,74]. Some studies coupled it with frequency domain features by converting the data segments using Discrete Cosine Transform (DCT) [42] and extract energy and frequency related features or using Fourier transform [58,70] and extract frequency components. In [71],Lee et.all used the Discrete Wavelet Transform (DWT) representation to extract wavelet coefficients and then computed their normalised sSum of absolute value. Whereas, in [73] zero-crossing decomposition is applied on the gravity free acceleration data, to then extract relevant features.

On an other hand, Boukhennoufa et.al in [50] encoded time series windows into Gramian Angular Field [92] images and fed them into some popular computer vision algorithms. Studies involving post-stroke rehabilitation require usually many sensors with multiple axes, this could yield to huge numbers of features and cause systems to over-fit. To remedy to this, dimensional reduction technique were used. Dimensionality reduction refers to techniques for reducing the number of input variables in training data by projecting the data to a lower dimensional subspace which captures the essence of the data. Multiple techniques were used in the studies included here. Yang et.al, and Tran et.al [41,46] used a technique called Principal Component Analysis (PCA) that transforms data into fewer dimensions. keeping the three first components only allowed Yang et.al to keep 95.86% of the overall information stored in 56 feature vectors while it allowed Tran to keep 99% of the information, reducing it from the 11 feature vectors. Other studies [44,58] employed Relief-F that takes a filter-method approach to feature selection to keep only the most relevant features.

3.6. Machine learning

ML is an application of AI that provides systems with the ability to automatically learn and improve from experience without being explicitly programmed to do so using the features selected before. Depending on weather to incorporate the outcomes, ML algorithms can be divided into two major categories: unsupervised learning and supervised learning. Unsupervised learning is well known for feature extraction, while supervised learning is suitable for predictive modelling through building some relationships between the patient traits and the outcome of interest [93]. All the papers included used supervised ML algorithms.

Support Vector Machines (SVM) were the most used classifier [94,42,44,46,49,53,56,57,68,70,74], it was used mainly for classification problems in activity recognition but also in regression problems for clinical assessments where participants are given a clinical score [65–71,73,74]. The reasons for choosing SVM variants is their good-generalisation ability for sequential data structures [95] and datasets that are not too large. This has been the case in most of the reviewed paper as recruiting post-stroke patients is not an easy task. Moreover, SVM has different kernel types allowing to deal both with linear and non-linear problems.

Random Forrest (RF) and more specifically Random Trees (DT) were also massively employed [44,46,47,49,55,60,61,65,67,69]. DT is one of the commonest oldest ML algorithm, it models it decision logics to outcomes in a tree-like architecture. Its easiness to interpret as well as its rapidity to learn made it popular to use in the tele-rehabilitation domain and especially in multi-class activity recognition problems. The reason for that is when going through the tree for a classification sample, the outcomes of all tests at each node will provide relevant information to infer about its class. RF were less used than the former [46,47,52,58], the reason is it is an ensemble of RT making it more prone to over-fitting. It is only used when the available dataset is relatively large.

Artificial Neural networks (ANNs) were also a common choice among researchers for post-stroke rehabilitation assessment. ANNs are a set of ML algorithms that are inspired by the neurons of the brain. ANN may be represented as an interconnection of layers of nodes in which the output of one node is an input to another node for subsequent processing layer. Multilayer perceptron (MLP) was the commonest among them [94,42,55,57,59,60]. MLP does not require feature engineering thus necessitate less domain expertise, although a drawback is the fact that they are considered to be black-box having sometimes unpredictable behaviour. MLP achieved very good results for activity recognition and movement classification. Another ANN architecture that was employed is Convolutional Neural Network (CNN) Architectures [43,45,50,54]. They are designed to automatically and adaptively learn spatial hierarchies of features through backpropagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers [96]. CNNs achieved outstanding accuracies in the computer vision field but this did not translate to time-series data structure which is the structure of the data from the sensors. Boukhennoufa et.al [50], encoded the sequential data into images and then employed a popular CNN architecture which is the VGG-16 to then achieve a very high accuracy.

k-Nearest Neighbour (kNN) is another algorithm that was used in three included works [64,51,46]. The kNN classifier is based on distance metric and was widely used in real-time applications as it is free from the underlying assumptions about the distribution of the dataset. Moreover, The setting of different values for 'K' can result in different classification results for the same problem which makes it an additional hyper-parameter to find the most performing model, especially in the activity recognition. Fig. 4 shows an example a wearable sensor based rehabilitation assessment steps.

As per the metric to assess the system, almost all the papers used the

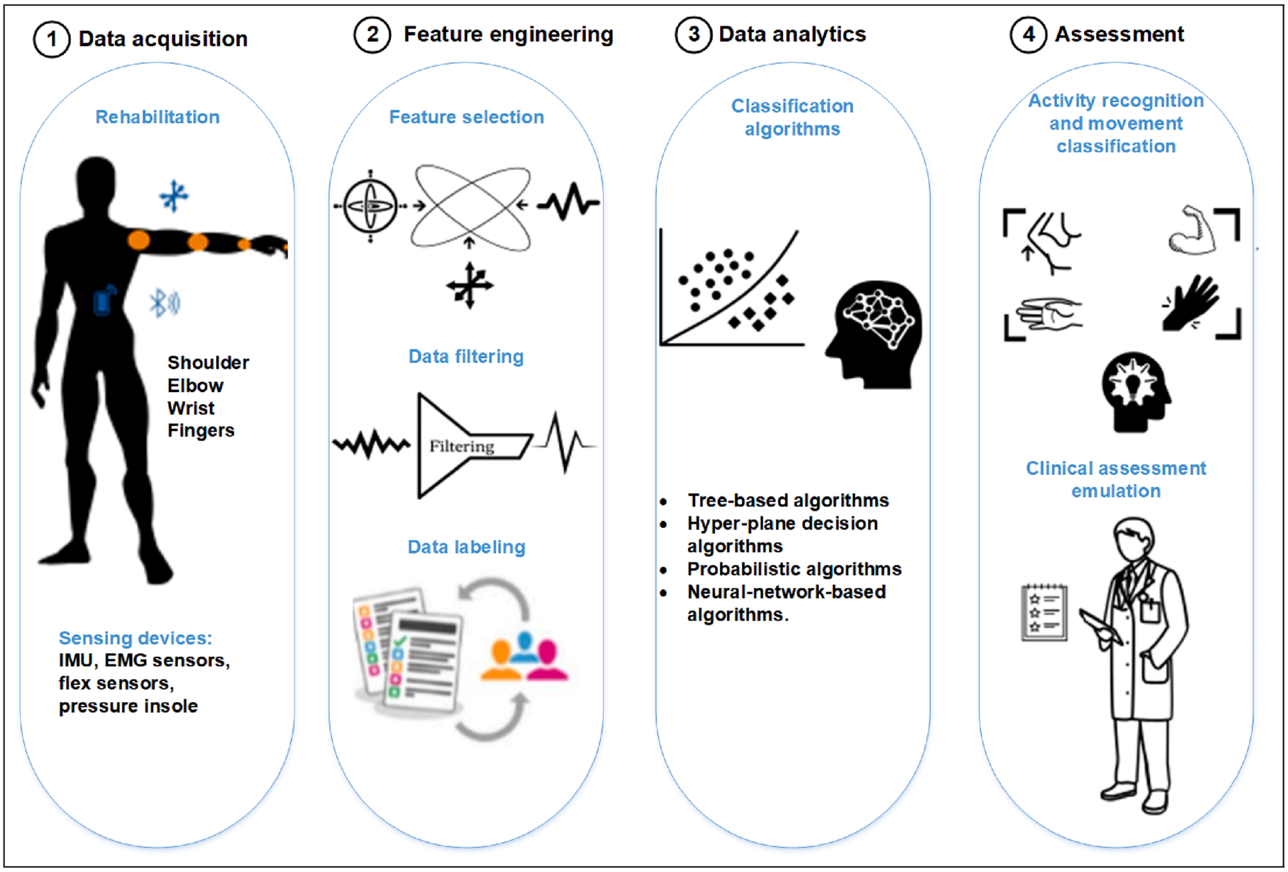

Fig. 4. wearable sensor based rehabilitation assessment steps.

accuracy. It is the proportion of the total number of predictions that were correct. The accuracy was computed using the formula:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) * 100$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.Two works [65,73] used the coefficient of determination, denoted R^2 which is a statistic that will give some information about the goodness of fit of a model in regression models. It was used for the clinical assessment algorithms to compare the predicted score with the score from the clinician.

$$R^2 = 1 - RSS/TSS$$

where *RSS* is the sum of squares of residuals and *TSS* is the total sum of squares

Moreover, another metric used in regression problems and especially in the clinical emulation assessment [53,67,69,71] is the Root Mean Square Error(RMSE) it is defined to be the standard deviation of the residuals (prediction errors):

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(d_i - f_i)^2}$$

Where $d_i$is the predicted score value and f_i is the actual one given by the therapist.

3.7. Quantitative analysis

In this subsection, a quantitative analysis related to the specifics of the included papers is presented, statistics about number of citations, system aim, year and 1st author's country of the different works are given. Fig. 5 shows some statistics charts.

The system aim is almost homogeneously divided between the three categories with eleven of the included works treating activity recognition [41–51], twelve dealing with movement classification [52–61,63,64] and ten with clinical assessment emulation. [65–74]. This demonstrates that the different categories are of equal interests to researchers.

For the publication year a growing interest has been noticed starting from 2020 in the included sample of works. Two studies were published in 2015 [44,48], three in 2016 [56,68,69], four in 2017 [47,61,65,72], six in 2018 [41,42,46,52,59,66], five in 2019 [43,53,55,57,70] and then the maximum eight in 2020 [45,54,58,63,67,71,73,74] and finally five up to now in 2021 (up to August) [49–51,60,64]. Subsequently, publications made in a period of two years and a half (2019, 2020,2021) which consists of a third of the overall time range (seven years and a half) accounted for 55% of the total publications. The growing interest maybe due to the recent Covid-19 pandemic.

As per the academic citations of the included papers, it ranged from no citations at all [50,58,60,64] to 126 citations [56] (up to August 2021). The papers with no citations were tolerated only in the most recent works written in 2021, and that the authors felt it presented an interesting approach worth reviewing. In the same context, statistics of the 1st author's publication are also given. The US, is the country with most publications with thirteen papers [46,47,49,52,54,58, 64–67,69,70,73] followed by china with four papers [41,51,53,68], three papers for South Korea [45,61,74] and Italy [55–57], two papers for the UK [50,71] and Thailand [59,60] and finally a single study for France [42], India [43], Canada [44], Switzerland [48], Germany [63], Singapore [72].

Additionally, as for the targeted limbs from the included papers, upper extremity rehabilitation is the dominant practice with fifteen papers [41,43,53,54,57,58,61,64,65,67–69,71–73]. This is justified by

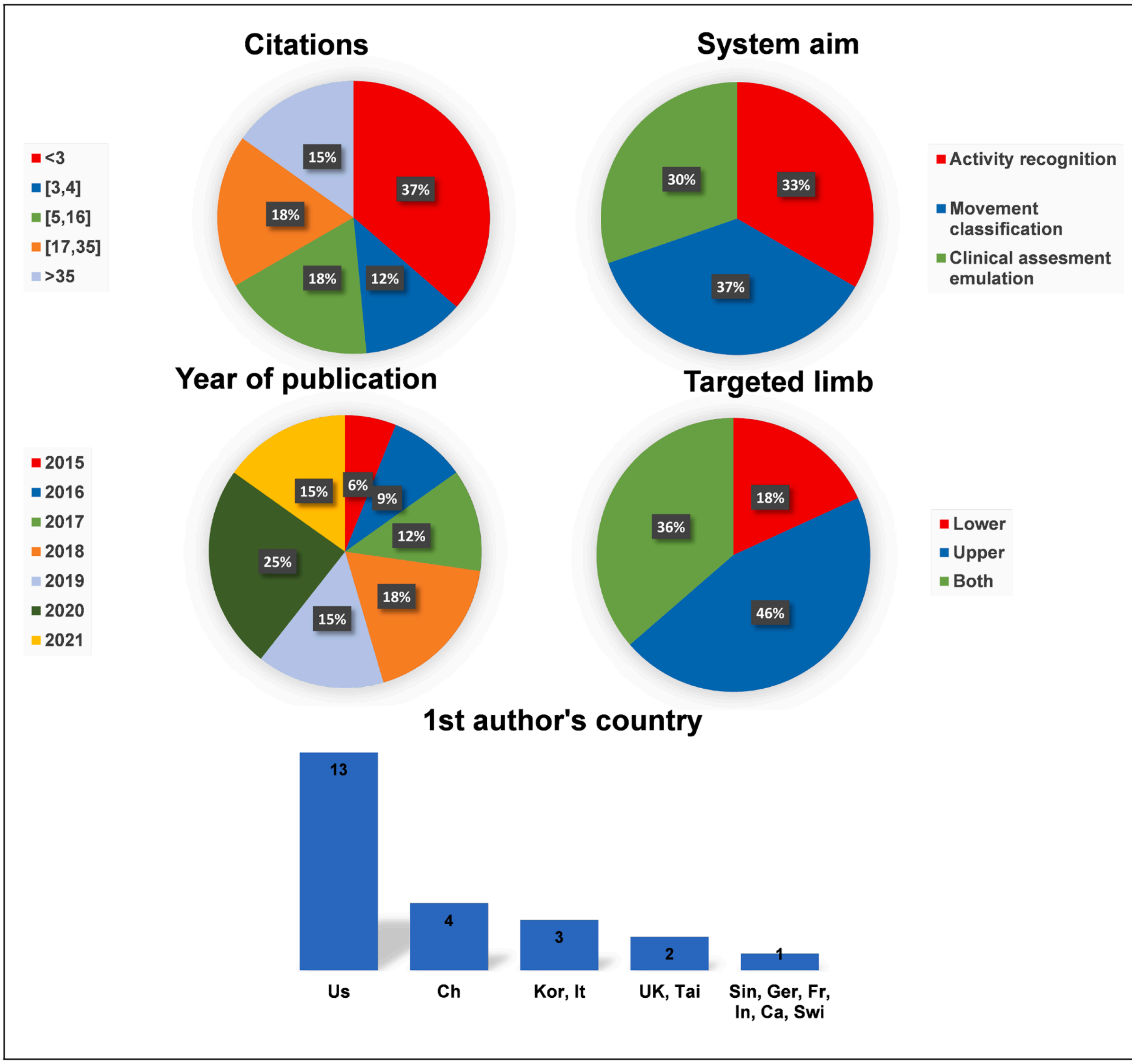

Fig. 5. Statistics on the included papers.

the fact that most of the clinical assessment batteries of test that the researchers tried to emulate are for the upper extremities, as well as the fact that most of the ADLs involve using hands. Both-limb rehabilitation comes second with twelve papers [42,45,46,48–52,63,66,70,74] and finally 6 papers for lower-limbs only [44,47,55,56,59,60].

4. Limitations challenges and potential study directions

In this section we will present in subsection first 4.1 the objective and the limitation for each of the included studies, based on that we present a list of challenges that the researchers in this field most commonly found in subSection 4.1.1 and at this end we will give some tips and some potential study directions to ameliorate the assessment systems in subSection 4.3.

4.1. Limitations

Table 3 presents a summary of each paper and the corresponding limitations and objectives.

4.1.1. Challenges

Based on the limitations presented in Table 3, we assessed different challenges encountered by the researchers in post-stroke tele-rehabilitation.

4.2. Quantity and quality of data

ML based system as it is the case for post-stroke smart tele-rehabilitation, requires rigorous computational models to achieve the desired results and estimate properly the needed parameters. The starting point to construct an efficient model is to have a massive amount of data, besides, the most sophisticated algorithms (e.g. deep learning) require at least 10 times the number of samples as parameters in the network. Indeed, These algorithms thrive in the domains where large amounts of data are easily collected (e.g. computer vision). On the other hand, in the healthcare area and more specifically in post-stroke rehabilitation, the number of patients is limited, and are not always keen to take part into research projects as it is an extra burden they endure. moreover, more than 70 percent of the patients live in low and middle-income countries [98] that do not give enough importance yet to data collection or do not have the necessary means. Subsequently, the available information is still limited to build and train efficient models that would generalise under different conditions and for different cases. Besides, in contrast with other fields where the data is well-structured,

Table 3
Summary of the reviewed papers.

Paper	Study aim	Limitation
[41]	Real-time gesture recognition performance to control a five finger dexterous robot;	• Focused only on user-specific condition, where the training data and the verification data are from the same subject posing a generalisation issue. • Did not test the model on stroke patients. • There is no mention if the final prototype has been used in clinical environments afterwards.
[42]	Monitor the overall body activity and the drinking activity from the liquid level of the mug. Subjects were asked to accomplish while holding the cup some ADL. the resulting data were fused together to increase the performance of the processing algorithm.	• The absence of a study on the acceptability of the smart cup by stroke patients. • No stroke patients were included and no research studies were conducted. • The absence of an assessment system.
[43]	A single sensor was used to collect data from the impaired arm of stroke survivors, the participants execute twenty different arm tasks in two different environment settings: patients at home and patients at labs.	• The absence of a real-time implementation of the system • There is no mention if the final prototype has been used in clinical environments afterwards.
[44]	The study determined signal features that are best suited for activity recognition with various populations (stroke patients, able bodied and elderly participants) independent of the chosen classifier.	• The study did not present any platform. • The classifiers were not customized to the specific HAR application.
[45]	Developed a home-based rehabilitation system that can recognise and record the type and frequency of rehabilitation exercises conducted by the user using a smartwatch.	• The total number of patients who completed the program was relatively small to derive statistically strong evidence. • The actual accuracy of exercise detection at home was not assessed.
[46]	A system that classifies functional and nonfunctional arm movement from accelerometry sensor data.	• Limited activities and ADL tasks that the participants performed..
[47]	Compared HAR performance for persons with stroke while varying the origin of training data, based on either population (AB or SP) or environment setting.	• The healthy cohort did not age match the stroke cohort. • Different sensor placements throughout the study. • Data associated with stroke patients in home setting was small compared to the others.
[48]	Proposed a wearable activity monitoring system based on a fuzzy logic based activity classifier that exploits fused information from the sensors which accounts for behavioral constraints and estimates the body elevation during standing and locomotion.	• Non-uniformity of the number of data samples for the different activities. • Limited number of SP. • There is no mention if the final prototype has been used in clinical environments afterwards.
[49]	Developed a novel prediction model based on ML algorithms and determine the accuracy of detecting different ADLs performed by stroke survivors. The study was conducted in a simulation living room and kitchen. Lastly, additional independent training and testing data were collected to perform external validation to further imitate real-world prediction conditions.	• The sample size is relatively small so the model might not generalise well. • Data was collected in a semi-naturalistic environment instead of participants' homes. • The accuracy might be more improved. • It is a pilot study so it did not yield to an assessment platform yet.
[50]	Encoded time-series data into gray-scale and RGB images and tested different CNN models to profit from computer vision development.	• No SP were included in the collected data. • The presence of confounding movements induced by clinical practitioner patient interactions while performing the exercises.
[51]	A comparative study to investigate the performance of different sensors and different placements for classifying four different ADLs with the purpose to find the optimal placement of a single sensor that achieves best accuracy.	• The study was preliminary. • Only five ADLs were included, they also have similar patterns • There is no mention if the final prototype has been used in clinical environments afterwards.
[52]	Used two sensors to differentiate between goal-directed exercises and ADL as well as detecting the poorly executed exercises following the FMA [76] assessment during in-home rehabilitation exercises.	• The sample size was relatively small and thus the reported results may not be generalised. • Movements that were both goal-directed and non goal-directed in nature were not considered. • There is no mention if the final prototype has been used in clinical environments afterwards.
[53]	The Establishment of a quantitative measurement system of the amount of hand use of 11 motor tasks of ADL using two sensors.	• The proposed technology cannot capture the use of the hands for stabilizing objects (e.g.. holding a cup or stabilizing a piece of steak with a fork) as it focuses on estimating the amount of hand movement. • No stroke patients were included and no research studies were conducted. • There is no mention if the final prototype has been used in clinical environments afterwards.
[54]	Developed an approach that identifies and counts functional primitives that constitute rehabilitation activities in an automated manner.	• The primitives were only recognised, no system has been set up to count them. • No platform was implemented. • Has not been tested in clinical settings.
[55]	SmartPants is used to perform therapy exercises and recognise some ADL lower-limbs taks.	• The number of sensors might be reduced to design a more unobtrusive system. • Did not test the model on stroke patients. • There is no mention if the final prototype has been used in-clinical environments afterwards.
[56]	Validate a general probabilistic modeling approach for the classification of different pathological gaits.	• Metrics for gait assessment were not included.
[57]	Recognise purposeful and non purposeful arms' movements of post-stroke patients when performing ADLs for identifying and promoting the use of the impaired limb during daily life in people affected by stroke. different datasets were investigated to see which gives better results, namely SP, AB, and both.	• Data collected from index and wrist sensors only. • The recruited groups were not age-matched. • Data were collected in a controlled environment.
[58]	Investigated the performance of unimanual, bimanual asymmetric, and bimanual symmetric tasks in participants post-stroke and controls for a	• Accuracy could be further improved. • There is no mention if the final prototype has been used in

(continued on next page)

Table 3 (*continued*)

Paper	Study aim	Limitation
	variety of signal processing and ML tools. The system classifies bi-manual and uni-manual tasks.	clinical environments afterwards. • Accuracy could be further improved. • The sample size was relatively small and thus the reported Results may not be generalised.
[59]	Examined IMU sensor placement configuration with different classification algorithms and differentiate between SP and NDP gaits. It showed that shank placement provided better accuracy.	• Limited sample size. • No clinical application was implemented from this research yet.
[60]	Developed a model that can recognise stroke gait with the help of therapists.	• Limited number of participants which causes the system to not generalise well. • There is no mention if the final prototype has been used in clinical environments afterwards.
[61]	A sensing sub-system placed on a shirt sleeve (smart shirt) collected data that are then processed locally on a smart wireless access point based on Raspberry Pi and then sent to an Android device via a Transmission Control Protocol (TCP) socket by a Wi-Fi master node where the patient is given a personal account that the physicians use to login into and visualise the real-time data. The information is then sent to a data cloud built with MySQL where it is stored and then pushed to a computing cloud that utilises ML algorithms implemented on MATLAB to evaluate the data [76].	• Limited number of activities included. • The data was collected from a single person. • No stroke patients were included and no research studies were conducted. • There is no mention if the final prototype has been used in clinical environments afterwards.
[63]	Proposed three digital biomarkers namely convergence points, physical activity and functional range of motion for the longitudinal performance monitoring and movement evaluation of stroke patients	• Some subtle movement changes require further research to distinguish improved movement ability due to recovery from movement compensation mechanisms. • Further analysis should be done to see how the algorithm would generalise.
[64]	Evaluate the feasibility of using body-worn sensors to track rehabilitation exercises in the inpatient setting and counting exercise repetitions in order to identify stroke severity	• Has only been tested on three basic ADLs. • Gyroscope data did not include all patients. • Stroke patient data were from subjects with at least moderate strength and did not include more severe cases.
[65]	Only two IMU sensors to assess quality of movement Functional Ability Scale scores[78], the results were then correlated with therapists scores giving very high accuracy.	• Further analysis should be done to see how the algorithm would generalise. • No stroke patients were included and no research studies were conducted. • There is no mention if the final prototype has been used in clinical environments afterwards.
[66]	A single sensor to measure upper extremities functional use during ADL and distinguish it from the upper extremities movements that occur while walking.	• Establishing clinical validity requires further research with larger patient populations to determine how well this methodology generalises across stroke survivors.
[67]	Multiple upper-limbs assessment system utilising two different rehabilitation evaluation scoring systems the FAS [97] and FMA associated with different ADL tasks.	• Further analysis on supplementary participants to see how the algorithms would generalise. • Accuracy could be improved further. • Has not been tested in clinical settings.
[68]	Proposes a novel remote quantitative FMA assessment system in home settings. Sensors record the movement information in real time and wirelessly transmit to the computer through ZigBee protocol and finally upload to the web server database through Internet.	• Only seven exercises were included. • The placement of sensors has not been investigated further. • The system is obtrusive. • There is no mention if the final prototype has been used in clinical environments afterwards.
[69]	Evaluated two approaches designed to estimate the quality of post'""stroke upper extremity motion as measured by the FMA subscale for the upper extremity using paretic and non"paretic limb kinematic data.	• Limited number of participants which causes the system to not generalise well. • No clinical application was implemented from this research. • Research was not conducted in a research environment.
[70]	Assessed whether long-term monitoring of seven days or more, in unilaterally impaired stroke patients is useful in determining motor impairment using [81].	• Very small sample size which would not generalise to more data. • The presence of confounding movements induced by clinical practitioner patient interactions while performing the exercises.
[71]	Developed an automated system that can predict the assessment score in an objective manner to do so two new features were proposed.	• There is no mention if the final prototype has been used in clinical environments afterwards. • Included activities are limited
[72]	Employ time–frequency methods to provide a better analytical basis for the derivations.	• Very small data sample, only 78 data segments were collected. • There is no mention if the final prototype has been used in clinical environments afterwards. • No clinical application was implemented from this research. • Research was not conducted in a research environment.
[73]	Proposed an approach to estimate upper-limb impairment in stroke survivors using two wearable inertial sensors, on the wrist and the sternum.	• Small sample size which may not generalise to more data. • Reliance on the performance of large, continuous movements, which can be tiresome for stroke patients.
[74]	Developed an autonomous grading system for stroke patients using NIHSS and MRC scores.	• Very small sample size which may not generalise to more data. • Very limited set of exercises.

healthcare data, in particular and sensor data in general, is heterogeneous, abstruse, noisy and difficult to interpret if not an expert.This makes building a good learning model tricky and requires to address several challenges, such as data-sparsity, missing and dismissed values, sensor miss-calibration issues and noisy segments. In the same context, data bias which is another issue can cause the assessment algorithm to evolve in an unpredictable manner and not generalise to new patients that have different degrees of severity. This was very common among the reviewed papers, where researchers complained about their algorithms not generalising well. [41,49,52,58,60]. Another data-related issue is confidentiality, especially with the growing use of cloud platforms and Internet of Things. Therefore, effort should be spent to secure the data transmission between the platforms to ensure privacy for the users of the assessment systems.

4.2.1. Recruitment related challenges

Dealing with post-stroke participants is a sensitive task and require researchers to follow strict procedures, starting from the recruitment process which requires undertaking tedious ethical approval applications to mounting sensors and collecting data from the participants. In addition, in order to design efficient ML-based assessment systems that generalises to new users a large number of participants should take part with different and variant degrees of severity [43,48,49,69,70], which is not always available and taken into account. Besides, a common issue we found on the included papers when doing cohort studies is not recruiting AB participants that age-match the SP recruited [47,57], this could yield to introduce inequalities that are not caused by stroke disease rather it is by the age difference.

4.2.2. Field complexity and field standards

Understanding illnesses in general and stroke in particular, is a more challenging task than dealing with natural language or image processing. It also requires an advanced expertise since the systems will be deployed to deal with human subjects to assist them in their rehabilitation process or to assess their execution. Moreover, the standards applied in healthcare are highly rigorous, ethical committees have to approve studies involving human subjects, in addition to the privacy restrictions that govern personal patients' data and sensitive information that limit the use of some modern platforms like computing and data clouds. Furthermore, threats introduced by hacking has become a leading cause of breaches in patients' data, and sensing devices are no exception to this since the information is often transmitted wirelessly. All of these reasons resulted in IoT systems locally processing data [41,56,55]. In the same context, some stroke clinical assessments, and some severe cases require particular expertise in dealing with patients to position their limbs, this is usually done with the assistance of experts in the field and is hard to translate to only wearable sensors.

4.2.3. Power consumption and latency issues

The wearable sensors are continuously sensing data, pre-processing and transmitting it to a remote platform for analysis or visualisation. This results in a huge power consumption that may yield to the devices turning off and thus terminating the monitoring process of the patients. In addition to that, absolute dependence on cloud platforms for the analysis of data may result in latency of the processing of information due to the huge amounts of data that these platforms receive at once, this may lead to the loss of the real-time aspect of the system or in worst cases to the complete failure of the system when the internet connection is lost.

4.2.4. Patients' acceptance

Patients' approval should be considered in order to build up platforms that will be used in both clinical and home setting. Sensing devices may turn out to be redundant if the patients or clinicians do not use them. Therefore, the wearable device should be unobtrusive, and easy to operate. It should not influence the ADL of the user. Researchers should also concentrate on the implications of the patients' preferences when designing the systems and more efforts should be spent on making stroke patients more familiar with intelligent sensing devices.

4.3. Potential directions

In this subsection we try to give some potential study directives and tips that might be worth considering to address a few challenges we noticed when reviewing the different papers and therefore design more efficient wearable sensor based post-stroke rehabilitation systems:

- Provide a person-centric approach that considers both what the individual should and can achieve during rehabilitation. Indeed, integrating the quantification and analysis of the present and future conditions of the patient would result in a personalised treatment that takes into account the specificities of the different users.
- Employing additional sensors in conjunction with IMUs to model additional quantities to limbs kinematics depending on the exercise. For batteries of test that involve strength exercises, employing EMG sensors would be an interesting approach to have the muscular activity, for gait related tasks and sit/standing, using insole pressure sensor would add useful information related to which lower limb is more active. For exercises that involve changing body level like standing up or sitting down or going up-stairs using level sensors would be an interesting approach. This will permit to lay out a more holistic and subjective assessment of the movement dysfunction
- Employing non-invasive, unobtrusive wearable sensors and take into more consideration the patient's comfort as now many studies proved the possibility to design effective systems with a very small number of sensors (sometimes a single sensor is sufficient) as is the case in [43,51]. Moreover, making the system simple to use and providing visual tips as using avatars, and giving positive feedback on the execution would attract more users.
- Implementing a more holistic assessment system by combining multiple evaluation categories as a movement classification in conjunction with clinical assessment emulation, this would allow to have different and complementary perspectives and therefore a more effective assessment. From the works reviewed, no paper combined it.
- Taking the assistance of field professionals when designing the systems, as some stroke clinical assessments, and some severe stroke cases require particular expertise in dealing with patients as for example to position their limbs when doing their recovery tasks.
- Making more use of deep learning algorithms as they do not require thorough feature engineering and thus require less signal processing expertise. Moreover, 1D time series deep learning has emerged and they provide better accuracies than conventional ML algorithms and are even much faster like [99].
- As the AI based technologies are going to be an important part of the modern world, it is important to follow universal standards and guidelines that orient the patients well-being in light of the social and ethical issues these AI technologies beget. The recent IEEE 7010–2020 [100] is one good example.

5. Conclusion and study limitations

The COVID-19 global pandemic highlighted the need for designing novel remote-working technologies especially in rehabilitation. In this paper we assessed numerous research works done in the post-stroke telerehabilitation assessment. Based on that, we proposed a new taxonomy on which we can divide the field namely activity recognition, movement classification and clinical assessment emulation. We surveyed the different wearable sensors used for data collection, we found that IMU dominate the other sensors with a slight presence of EMG sensors. Moreover, we scrutinised the sensors' placements, the study designs in addition to the feature engineering and the machine learning used for the assessment. Feature engineering in conventional ML algorithms requires domain expertise which makes a tedious process to implement performant systems. Nevertheless, with the development of computing platforms sophisticated algorithms namely deep learning are taking over which necessitate less domain knowledge. From the reviewed papers, we identified challenges encountered by the researchers in the field that relates to the data-aspect, recruitment to undertake the studies, Field complexity, Power consumption and patients' acceptance. We finally gave some tips to help researchers in the field improve their systems.

A limitation of this work is that We have selected articles based on our inclusion/exclusion criteria which we know is not perfect. We have made our best to select the maximum number of articles, we also decided not to include some papers that satisfied our preliminary criteria because we found the ideas they dealt with are already treated in other works. Moreover, we have excluded some systems that are also used in post-stroke rehabilitation such as exoskeleton because we believe it to be an obtrusive device that exceeds the definition of a simple wearable device. Virtual reality based systems were also not included, as they were treated in many recent reviews.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by the UK Engineering and Physical Sciences Research Council through grants EP/R02572X/1, EP/P017487/1, EP/P015956/1 and EP/V034111/1.

References

[1] E.J. Benjamin, S.S. Virani, C.W. Callaway, A.M. Chamberlain, A.R. Chang, S. Cheng, S.E. Chiuve, M. Cushman, F.N. Delling, R. Deo, et al., Heart disease and stroke statistics-2018 update: a report from the american heart association, Circulation 137 (12) (2018) e67–e492.
[2] P. Langhorne, J. Bernhardt, G. Kwakkel, Stroke rehabilitation, The Lancet 377 (9778) (2011) 1693–1702.
[3] E.H. Lo, M.A. Moskowitz, T.P. Jacobs, Exciting, radical, suicidal: how brain cells die after stroke, Stroke 36 (2) (2005) 189–192.
[4] J. Gomes, A.M. Wachsman, Types of strokes, in: Handbook of clinical nutrition and stroke, Springer, 2013, pp. 15–31.
[5] S.C. Johnston, Transient ischemic attack, New England Journal of Medicine 347 (21) (2002) 1687–1692.
[6] H.B. Van der Worp, J. van Gijn, Acute ischemic stroke, New England Journal of Medicine 357 (6) (2007) 572–579.
[7] S.D. Smith, C.J. Eskey, Hemorrhagic stroke, Radiologic Clinics 49 (1) (2011) 27–45.
[8] A. Rosen, Stroke: who's counting what? Development 38 (2) (2001) 281–289.
[9] J. Gomes, A.M. Wachsman, Types of strokes, in: Handbook of clinical nutrition and stroke, Springer, Totowa, New Jersey, 2013, pp. 15–31.
[10] P. Pound, P. Gompertz, S. Ebrahim, Illness in the context of older age: the case of stroke, Sociology of Health & Illness 20 (4) (1998) 489–506.
[11] D.E. Levy, R.L. Van Uitert, L. Catherine, Delayed postischemic hypoperfusion: a potentially damaging consequence of stroke, Neurology 29 (9 Part 1) (1979) 1245.
[12] N.H.S. Constitution, Rehabilitation after a strok, Proteins (2013) 1–17.
[13] B.H. Dobkin, Strategies for stroke rehabilitation, The Lancet Neurology 3 (9) (2004) 528–536.
[14] S.A. Billinger, R. Arena, J. Bernhardt, J.J. Eng, B.A. Franklin, C.M. Johnson, M. MacKay-Lyons, R.F. Macko, G.E. Mead, E.J. Roth, et al., Physical activity and exercise recommendations for stroke survivors: a statement for healthcare professionals from the american heart association/american stroke association, Stroke 45 (8) (2014) 2532–2553.
[15] R.L. Sacco, R. Adams, G. Albers, M.J. Alberts, O. Benavente, K. Furie, L. B. Goldstein, P. Gorelick, J. Halperin, R. Harbaugh, et al., Guidelines for prevention of stroke in patients with ischemic stroke or transient ischemic attack: a statement for healthcare professionals from the american heart association/ american stroke association council on stroke: co-sponsored by the council on cardiovascular radiology and intervention: the american academy of neurology affirms the value of this guideline, Stroke 37 (2) (2006) 577–617.
[16] J.A. Zivin, D.W. Choi, Stroke therapy, Scientific American 265 (1) (1991) 56–65.
[17] L. Catanese, J. Tarsia, M. Fisher, Acute ischemic stroke therapy overview, Circulation Research 120 (3) (2017) 541–558.
[18] M. Fisher, W. Schaebitz, An overview of acute stroke therapy: past, present, and future, Archives of Internal Medicine 160 (21) (2000) 3196–3206.
[19] J.-F. Nepveu, A. Thiel, A. Tang, J. Fung, J. Lundbye-Jensen, L.A. Boyd, M. Roig, A single bout of high-intensity interval training improves motor skill retention in individuals with stroke, Neurorehabilitation and Neural Repair 31 (8) (2017) 726–735.
[20] R.A. Geiger, J.B. Allen, J. O'Keefe, R.R. Hicks, Balance and mobility following stroke: effects of physical therapy interventions with and without biofeedback/ forceplate training, Physical Therapy 81 (4) (2001) 995–1005.
[21] E. Taub, J.E. Crago, G. Uswatte, Constraint-induced movement therapy: A new approach to treatment in physical rehabilitation, Rehabilitation Psychology 43 (2) (1998) 152.
[22] Z.-S. Hosseini, H. Peyrovi, M. Gohari, The effect of early passive range of motion exercise on motor function of people with stroke: a randomized controlled trial, Journal of Caring Sciences 8 (1) (2019) 39.
[23] J.R. Gladman, C. Sackley, The scope for rehabilitation in severely disabled stroke patients, Disability and Rehabilitation 20 (10) (1998) 391–394.
[24] C.M. Stinear, C.E. Lang, S. Zeiler, W.D. Byblow, Advances and challenges in stroke rehabilitation, The Lancet Neurology 19 (4) (2020) 348–360.
[25] J.A. Young, M. Tolentino, Stroke evaluation and treatment, Topics in Stroke Rehabilitation 16 (6) (2009) 389–410.
[26] A.E. Dickerson, L.J. Molnar, D.W. Eby, G. Adler, M. Bedard, M. Berg-Weger, S. Classen, D. Foley, A. Horowitz, H. Kerschner, et al., Transportation and aging: A research agenda for advancing safe mobility, The Gerontologist 47 (5) (2007) 578–590.
[27] J.H. Cauraugh, S.B. Kim, Chronic stroke motor recovery: duration of active neuromuscular stimulation, Journal of the Neurological Sciences 215 (1–2) (2003) 13–19.
[28] L.H. Schwamm, R.G. Holloway, P. Amarenco, H.J. Audebert, T. Bakas, N. R. Chumbler, R. Handschu, E.C. Jauch, W.A. Knight IV, S.R. Levine, et al., A review of the evidence for the use of telemedicine within stroke systems of care: a scientific statement from the american heart association/american stroke association, Stroke 40 (7) (2009) 2616–2634.
[29] T. Hester, R. Hughes, D.M. Sherrill, B. Knorr, M. Akay, J. Stein, P. Bonato, Using wearable sensors to measure motor abilities following stroke, in: International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06), IEEE, 2006, p. 4.
[30] Y. Zhou, Y. Fang, J. Zeng, K. Li, H. Liu, A multi-channel emg-driven fes solution for stroke rehabilitation, in: International Conference on Intelligent Robotics and Applications, Springer, 2018, pp. 235–243.
[31] K. Rathakrishnan, S.-N. Min, S.J. Park, Evaluation of ecg features for the classification of post-stroke survivors with a diagnostic approach, Applied Sciences 11 (1) (2021) 192.
[32] S. Patel, H. Park, P. Bonato, L. Chan, M. Rodgers, A review of wearable sensors and systems with application in rehabilitation, Journal of Neuroengineering and Rehabilitation 9 (1) (2012) 21.
[33] N. Ahmad, R.A.R. Ghazilla, N.M. Khairi, V. Kasi, Reviews on various inertial measurement unit (imu) sensor applications, International Journal of Signal Processing Systems 1 (2) (2013) 256–262.
[34] R.W. Bohannon, Sit-to-stand test for measuring performance of lower extremity muscles, Perceptual and Motor Skills 80 (1) (1995) 163–166.
[35] D. Podsiadlo, S. Richardson, The timed up & go: a test of basic functional mobility for frail elderly persons, Journal of the American Geriatrics Society 39 (2) (1991) 142–148.
[36] S. Díaz, J.B. Stephenson, M.A. Labrador, Use of wearable sensor technology in gait, balance, and range of motion analysis, Applied Sciences 10 (1) (2020) 234.
[37] K. Shailaja, B. Seetharamulu, M. Jabbar, Machine learning in healthcare: A review, in: Second international conference on electronics, communication and aerospace technology (ICECA), IEEE, 2018, pp. 910–914.
[38] M.A. Alsheikh, S. Lin, D. Niyato, H.-P. Tan, Machine learning in wireless sensor networks: Algorithms, strategies, and applications, IEEE Communications Surveys & Tutorials 16 (4) (2014) 1996–2018.
[39] T. Russell-Rose, J. Chamberlain, F. Shokraneh, A visual approach to query formulation for systematic search, in: Proceedings of the 2019 Conference on Human Information Interaction and Retrieval, 2019, pp. 379–383.
[40] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, Preferred reporting items for systematic reviews and meta-analyses: the prisma statement, Annals of Internal Medicine 151 (4) (2009) 264–269.
[41] G. Yang, J. Deng, G. Pang, H. Zhang, J. Li, B. Deng, Z. Pang, J. Xu, M. Jiang, P. Liljeberg, et al., An iot-enabled stroke rehabilitation system based on smart wearable armband and machine learning, IEEE Journal of Translational Engineering in Health and Medicine 6 (2018) 1–10.
[42] M. Bobin, H. Amroun, M. Boukalle, M. Anastassova, M. Ammi, Smart cup to monitor stroke patients activities during everyday life, in: 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), IEEE, 2018, pp. 189–195.
[43] M. Panwar, D. Biswas, H. Bajaj, M. Jöbges, R. Turk, K. Maharatna, A. Acharyya, Rehab-net: Deep learning framework for arm movement classification using wearable sensors for stroke rehabilitation, IEEE Transactions on Biomedical Engineering 66 (11) (2019) 3026–3037.
[44] N.A. Capela, E.D. Lemaire, N. Baddour, Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients, PloS one 10 (4) (2015), e0124414.
[45] S.H. Chae, Y. Kim, K.-S. Lee, H.-S. Park, Development and clinical evaluation of a web-based upper limb home rehabilitation system using a smartwatch and machine learning model for chronic stroke survivors: Prospective comparative study, JMIR mHealth and uHealth 8 (7) (2020), e17216.
[46] T. Tran, L.-C. Chang, I. Almubark, E.M. Bochniewicz, L. Shu, P.S. Lum, A. Dromerick, Robust classification of functional and nonfunctional arm movement after stroke using a single wrist-worn sensor device, in: 2018 IEEE International Conference on Big Data (Big Data), IEEE, 2018, pp. 5457–5459.
[47] M.K. O'Brien, N. Shawen, C.K. Mummidisetty, S. Kaur, X. Bo, C. Poellabauer, K. Kording, A. Jayaraman, Activity recognition for persons with stroke using mobile phone technology: toward improved performance in a home setting, Journal of Medical Internet Research 19 (5) (2017), e184.
[48] F. Massé, R.R. Gonzenbach, A. Arami, A. Paraschiv-Ionescu, A.R. Luft, K. Aminian, Improving activity recognition using a wearable barometric pressure sensor in mobility-impaired stroke patients, Journal of Neuroengineering and Rehabilitation 12 (1) (2015) 1–15.
[49] P.-W. Chen, N.A. Baune, I. Zwir, J. Wang, V. Swamidass, A.W. Wong, Measuring activities of daily living in stroke patients with motion machine learning

algorithms: A pilot study, International Journal of Environmental Research and Public Health 18 (4) (2021) 1634.
[50] I. Boukhennoufa, X. Zhai, K.D. McDonald-Maier, V. Utti, J. Jackson, Improving the activity recognition using gmaf and transfer learning in post-stroke rehabilitation assessment, in: 2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI), IEEE, 2021, pp. 000391–000398.
[51] L. Meng, A. Zhang, C. Chen, X. Wang, X. Jiang, L. Tao, J. Fan, X. Wu, C. Dai, Y. Zhang, et al., Exploration of human activity recognition using a single sensor for stroke survivors and able-bodied people, Sensors 21 (3) (2021) 799.
[52] S.I. Lee, C.P. Adans-Dester, M. Grimaldi, A.V. Dowling, P.C. Horak, R.M. Black-Schaffer, P. Bonato, J.T. Gwin, Enabling stroke rehabilitation in home and community settings: a wearable sensor-based approach for upper-limb motor training, IEEE Journal of Translational Engineering in Health and Medicine 6 (2018) 1–11.
[53] X. Liu, S. Rajan, N. Ramasarma, P. Bonato, S.I. Lee, The use of a finger-worn accelerometer for monitoring of hand use in ambulatory settings, IEEE Journal of Biomedical and Health Informatics 23 (2) (2019) 599–606, https://doi.org/10.1109/JBHI.2018.2821136.
[54] A. Kaku, A. Parnandi, A. Venkatesan, N. Pandit, H. Schambra, C. Fernandez-Granda, Towards data-driven stroke rehabilitation via wearable sensors and deep learning, arXiv preprint arXiv:2004.08297 (2020).
[55] I. Bisio, C. Garibotto, F. Lavagetto, A. Sciarrone, When ehealth meets iot: A smart wireless system for post-stroke home rehabilitation, IEEE Wireless Communications 26 (6) (2019) 24–29.
[56] A. Mannini, D. Trojaniello, A. Cereatti, A.M. Sabatini, A machine learning framework for gait classification using inertial sensors: Application to elderly, post-stroke and huntington's disease patients, Sensors 16 (1) (2016) 134.
[57] A.H. Butt, C. Zambrana, S. Idelsohn-Zielonka, M. Claramunt-Molet, A. Ugartemendia-Etxarri, E. Rovini, A. Moschetti, C. Molleja, C. Martin, E. O. Salleras, et al., Assessment of purposeful movements for post-stroke patients in activites of daily living with wearable sensor device, in: 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), IEEE, 2019, pp. 1–8.
[58] A. Miller, L. Quinn, S.V. Duff, E. Wade, Comparison of machine learning approaches for classifying upper extremity tasks in individuals post-stroke, in: 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2020, pp. 4330–4336.
[59] W.-C. Hsu, T. Sugiarto, Y.-J. Lin, F.-C. Yang, Z.-Y. Lin, C.-T. Sun, C.-L. Hsu, K.-N. Chou, Multiple-wearable-sensor-based gait classification and analysis in patients with neurological disorders, Sensors 18 (10) (2018) 3397.
[60] F.-C. Wang, S.-F. Chen, C.-H. Lin, C.-J. Shih, A.-C. Lin, W. Yuan, Y.-C. Li, T.-Y. Kuo, Detection and classification of stroke gaits by deep neural networks employing inertial measurement units, Sensors 21 (5) (2021) 1864.
[61] Y. Jiang, Y. Qin, I. Kim, Y. Wang, Towards an iot-based upper limb rehabilitation assessment system, in: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2017, pp. 2414–2417.
[62] G. Rilling, P. Flandrin, P. Goncalves, et al., On empirical mode decomposition and its algorithms, in: IEEE-EURASIP workshop on nonlinear signal and image processing, Vol. 3, NSIP-03, Grado (I), 2003, pp. 8–11.
[63] A. Derungs, C. Schuster-Amft, O. Amft, Wearable motion sensors and digital biomarkers in stroke rehabilitation, Current Directions Biomedical Engineering 6 (3) (2020) 229–232.
[64] N. Balestra, G. Sharma, L.M. Riek, A. Busza, Automatic identification of upper extremity rehabilitation exercise type and dose using body-worn sensors and machine learning: A pilot study, Digital Biomarkers 5 (2) (2021) 158–166.
[65] S. Sapienza, C. Adans-Dester, O. Anne, G. Vergara-Diaz, S. Lee, S. Patel, R. Black-Schaffer, R. Zafonte, P. Bonato, C. Meagher, et al., Using a minimum set of wearable sensors to assess quality of movement in stroke survivors, in: 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), IEEE, 2017, pp. 284–285.
[66] E.M. Bochniewicz, G. Emmer, A. McLeod, J. Barth, A.W. Dromerick, P. Lum, Measuring functional arm movement after stroke using a single wrist-worn sensor and machine learning, Journal of Stroke and Cerebrovascular Diseases 26 (12) (2017) 2880–2887.
[67] C. Adans-Dester, N. Hankov, A. O'Brien, G. Vergara-Diaz, R. Black-Schaffer, R. Zafonte, J. Dy, S.I. Lee, P. Bonato, Enabling precision rehabilitation interventions using wearable sensors and machine learning to track motor recovery, NPJ Digital Medicine 3 (1) (2020) 1–10.
[68] L. Yu, D. Xiong, L. Guo, J. Wang, A remote quantitative fugl-meyer assessment framework for stroke patients based on wearable sensor networks, Computer Methods and Programs in Biomedicine 128 (2016) 100–110.
[69] S. Chaeibakhsh, E. Phillips, A. Buchanan, E. Wade, Upper extremity post-stroke motion quality estimation with decision trees and bagging forests, in: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2016, pp. 4585–4588.
[70] A. Lucas, J. Hermiz, J. Labuzetta, Y. Arabadzhi, N. Karanjia, V. Gilja, Use of accelerometry for long term monitoring of stroke patients, IEEE Journal of Translational Engineering in Health and Medicine 7 (2019) 1–10.
[71] X. Chen, Y. Guan, J.-Q. Shi, X.-L. Du, J. Eyre, Automated stroke rehabilitation assessment using wearable accelerometers in free-living environments, arXiv preprint arXiv:2009.08798 (2020).
[72] T.K. Lee, K.-H. Leo, S. Sanei, E. Chew, L. Zhao, Triaxial rehabilitative data analysis incorporating matching pursuit, in: 25th European Signal Processing Conference (EUSIPCO), IEEE, 2017, pp. 434–438.
[73] B. Oubre, J.-F. Daneault, H.-T. Jung, K. Whritenour, J.G.V. Miranda, J. Park, T. Ryu, Y. Kim, S.I. Lee, Estimating upper-limb impairment level in stroke survivors using wearable inertial sensors and a minimally-burdensome motor task, IEEE Transactions on Neural Systems and Rehabilitation Engineering 28 (3) (2020) 601–611.
[74] E. Park, K. Lee, T. Han, H.S. Nam, Automatic grading of stroke symptoms for rapid assessment using optimized machine learning and 4-limb kinematics: clinical validation study, Journal of Medical Internet Research 22 (9) (2020), e20641.
[75] S. Katz, Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living, Journal of the American Geriatrics Society 31 (12) (1983) 721–727.
[76] K.J. Sullivan, J.K. Tilson, S.Y. Cen, D.K. Rose, J. Hershberg, A. Correa, J. Gallichio, M. McLeod, C. Moore, S.S. Wu, et al., Fugl-meyer assessment of sensorimotor function after stroke: standardized training procedure for clinical practice and clinical trials, Stroke 42 (2) (2011) 427–432.
[77] Y.-W. Hsieh, I.-P. Hsueh, Y.-T. Chou, C.-F. Sheu, C.-L. Hsieh, G. Kwakkel, Development and validation of a short form of the fugl-meyer motor scale in patients with stroke, Stroke 38 (11) (2007) 3052–3054.
[78] S.L. Wolf, J.P. McJunkin, M.L. Swanson, P.S. Weiss, Pilot normative database for the wolf motor function test, Archives of Physical Medicine and Rehabilitation 87 (3) (2006) 443–445.
[79] S.L. Wolf, P.A. Catlin, M. Ellis, A.L. Archer, B. Morgan, A. Piacentino, Assessing wolf motor function test as outcome measure for research in patients after stroke, Stroke 32 (7) (2001) 1635–1639.
[80] M. McDonnell, Action research arm test, Aust J Physiother 54 (3) (2008) 220.
[81] T. Da Roza, T. Mascarenhas, M. Araujo, V. Trindade, R.N. Jorge, Oxford grading scale vs manometer for assessment of pelvic floor strength in nulliparous sports students, Physiotherapy 99 (3) (2013) 207–211.
[82] S.R. Barreca, P.W. Stratford, L.M. Masters, C.L. Lambert, J. Griffiths, C. McBay, Validation of three shortened versions of the chedoke arm and hand activity inventory, Physiotherapy Canada 58 (2) (2006) 148–156.
[83] S. Majumder, T. Mondal, M.J. Deen, Wearable sensors for remote health monitoring, Sensors 17 (1) (2017) 130.
[84] N. Ahmad, R.A.R. Ghazilla, N.M. Khairi, V. Kasi, Reviews on various inertial measurement unit (imu) sensor applications, International Journal of Signal Processing Systems 1 (2) (2013) 256–262.
[85] K. Tanja-Dijkstra, M.E. Pieterse, The psychological effects of the physical healthcare environment on healthcare personnel, Cochrane Database of Systematic Reviews (2011) 1.
[86] D.W. Nicoll, Users as currency: Technology and marketing trials as naturalistic environments, The Information Society 16 (4) (2000) 303–310.
[87] A. Zheng, A. Casari, Feature engineering for machine learning: principles and techniques for data scientists, O'Reilly Media Inc, 2018.
[88] R. Miotto, F. Wang, S. Wang, X. Jiang, J.T. Dudley, Deep learning for healthcare: review, opportunities and challenges, Briefings in Bioinformatics 19 (6) (2018) 1236–1246.
[89] J. Himberg, K. Korpiaho, H. Mannila, J. Tikanmaki, H.T. Toivonen, Time series segmentation for context recognition in mobile devices, in: Proceedings 2001 IEEE International Conference on Data Mining, IEEE, 2001, pp. 203–210.
[90] T.-C. Fu, A review on time series data mining, Engineering Applications of Artificial Intelligence 24 (1) (2011) 164–181.
[91] J. Heaton, An empirical analysis of feature engineering for predictive modeling, in, in: SoutheastCon 2016, IEEE, 2016, pp. 1–6.
[92] Z. Wang, T. Oates, Imaging time-series to improve classification and imputation, arXiv preprint arXiv:1506.00327 (2015).
[93] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, Y. Wang, Artificial intelligence in healthcare: past, present and future, Stroke and Vascular Neurology 2 (4) (2017).
[94] T.J. Quinn, P. Langhorne, D.J. Stott, Barthel index for stroke trials: development, properties, and application, Stroke 42 (4) (2011) 1146–1151.
[95] P. Khera, N. Kumar, Role of machine learning in gait analysis: a review, Journal of Medical Engineering & Technology 44 (8) (2020) 441–467.
[96] R. Yamashita, M. Nishio, R.K.G. Do, K. Togashi, Convolutional neural networks: an overview and application in radiology, Insights into Imaging 9 (4) (2018) 611–629.
[97] H.J. Michielsen, J. De Vries, G.L. Van Heck, Psychometric qualities of a brief self-rated fatigue measure: The fatigue assessment scale, Journal of Psychosomatic Research 54 (4) (2003) 345–352.
[98] W. Johnson, O. Onuma, M. Owolabi, S. Sachdev, Stroke: a global response is needed, Bulletin of the World Health Organization 94 (9) (2016) 634.
[99] H.I. Fawaz, B. Lucas, G. Forestier, C. Pelletier, D.F. Schmidt, J. Weber, G.I. Webb, L. Idoumghar, P.-A. Muller, F. Petitjean, Inceptiontime: Finding alexnet for time series classification, Data Mining and Knowledge Discovery 34 (6) (2020) 1936–1962.
[100] D. Schiff, A. Ayesh, L. Musikanski, J.C. Havens, Ieee 7010: A new standard for assessing the well-being implications of artificial intelligence, in: IEEE international conference on systems, man, and cybernetics (SMC), IEEE, 2020, pp. 2746–2753.