International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.2 (2025) pp. 1943-1964

http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Human Activity-Based Machine Learning and Deep Learning Techniques

Mohammed F Ibrahim Alsarraj[1,2*], Aliza Binti Abdul Latif[3], Rohaini Binti Ramli[4]

[1]College of Graduate Studies, Universiti Tenaga Nasional, Jalan Ikram-UNITEN, 43000 Kajang, Selangor, Malaysia.
[2]Technical College of Management, Northern Technical University, Ninevah 41001, Iraq.
* **Corresponding Author Email:** PT21274@student.uniten.edu.my,mohammed_alsarraj@ntu.edu.iq - **ORCID:**0000-0001-7886-4464

[3]College of Computing and Informatics, Universiti Tenaga Nasional, Jalan Ikram-UNITEN, 43000 Kajang, Selangor, Malaysia.
Email: aliza@uniten.edu.my - **ORCID:** 0000-0002-8293-6130

[4]College of Computing and Informatics, Universiti Tenaga Nasional, Jalan Ikram-UNITEN, 43000 Kajang, Selangor, Malaysia.
Email: Rohaini@uniten.edu.my - **ORCID:** 0000-0002-4883-8967

Article Info:

DOI: 10.22399/ijcesen.1368
Received : 28 December 2024
Accepted : 20 March 2025

Keywords :

Human Activity Recognition,
Machine learning,
Deep Learning,
Sensors,
Healthcare,
Activity Monitoring.

Abstract:

Human activity recognition (HAR) has been hot research issues in recent years. The studies have differences in data types, data processing, feature description, etc. HAR constitutes a fundamental component of intelligent health monitoring systems, wherein the underlying intelligence of the services is derived from and enhanced by sensor data. Researchers have proposed multiple HAR systems designed to convert smartphone readings into other forms of physical activity. This review synthesizes the current methodologies for smartphone-based Human Activity Recognition (HAR) with focusing on healthcare application. For this purpose, we systematically searched for peer-reviewed articles regarding the utilization of cell phones for Human Activity Recognition (HAR). We collect information regarding smartphone body placement, sensors, types of physical activities examined, as well as the data transformation methodologies and classification frameworks employed for activity recognition. Thus, we selected these articles and delineated the diverse methodologies employed for data gathering, preprocessing, extraction of features, and activity classification, highlighting the predominant practices and their alternatives. We determine that cell phones are very appropriate for HAR research within the health sciences. Future studies should prioritize enhancing the quality of data gathered, addressing data gaps, incorporating a more diverse array of participants and activities, relaxing phone placement requirements, providing comprehensive documentation for study participants, and sharing the source code of the employed methods and algorithms to achieve population-level impact.

1. Introduction

Human activity recognition is one of the hot spots in the research of metaverse applications and digital sports training, and different human movements can be recorded by some information carriers, such as cameras, sensors, radars, Wi-Fi signals, etc. [1]. With the fast development of micro-electromechanical systems, the integration of sensor modules has become increasingly high[2] .Highly integrated intelligent sensor modules have promoted the continuous progress and improvement of human activity recognition applications in ubiquitous environments. In particular, portable wearable sensor devices represented by smartphones are becoming increasingly intelligent and popular. Sensor-based activity recognition has become a better way in many intelligent scenarios [3,4]. Recently, this topic has gained attention in the field of machine learning studies community; as of the present time, several studies have been presented on HAR techniques utilizing smartphones. This is a significant escalation from merely a few pieces published some years prior. As smartphone data collecting becomes more accessible, the interpretation of the gathered data is increasingly recognized as the primary impediment in health research. To address the analytical issues of HAR, researchers have developed diverse algorithms that significantly vary in the types of data utilized, the methods of data manipulation, and the statistical techniques employed for inference and/or classification. [5,6]. Published research utilize traditional approaches and introduce new methods for the collection, processing, and categorizing activities of daily living. Authors often study data filtering along with feature selection methodologies, comparing the accuracy of diverse machine learning classifiers on either pre-existing datasets or datasets they have assembled specifically for the study. The outcomes are generally encapsulated through categorization accuracy across many activity categories, including ambulation, locomotion, and exercise. [3-7]. To successfully implement advancements in HAR into public health and medical research, it is necessary to understand the traditional techniques and identify their possible limitations. Methods must consider physiological (such as weight, height, age) in addition to habitual (such as posture, walks) variations among smartphone users, along with gaps in architectural design (such as structures and green spaces) which influence the social and physical context for human activities. Furthermore, the data collecting and statistical methodologies commonly employed in Human Activity Recognition (HAR). This work systematically reviews the developing literature on the utilization of smartphones for human activity recognition in health research within free-living environments. Recognizing that the primary problem in this domain is transitioning from data gathering to data analysis, we concentrate our examination on the methodologies employed for data collecting, data preprocessing, extraction of features, and activity classification. We elucidate the complexity and multiple dimensions of Human Activity Recognition (HAR) using cellphones, the categories of data gathered, and the techniques employed to convert digital metrics into human actions. We examine the reproducibility and generalizability of methodologies, namely the characteristics that are crucial and relevant to extensive and varied groups of study participants. Finally, we identify obstacles that must be addressed to enhance the broader application of smartphone-based HAR for public health research. In this work, we investigated sensor-based HAR models with an overview of the new type of machine learning, deep learning and hybrid models that overcome the HAR shortcomings. Readout training is a simple supervised learning linear regression problem, which has simplicity and high learning efficiency.

2. Method

2.1 Review Planning

This systematic review has been carried out by searching for published literature on PubMed, Scopus, IEEE, Web of Science, ACM, and Springer databases. The databases were screened for titles, abstracts, and keywords containing phrases ("HAR" OR human activity recognition") AND "activity" AND ("recognition" OR "estimation" OR "classification") AND ("smartphone" OR "cell phone" OR "mobile phone" OR "wearable device") AND ("deep learning" OR "machine learning") The search had been limited to discovering all journal papers published in English. Subsequent to the elimination of duplicates, we reviewed the abstracts and titles of the remained publications. Studies which were not investigating HAR methods have been excluded from additional assessment. We then excluded studies that utilized supplemental equipment, such as wearable and ambient devices, as well as those requiring the transport of several smartphones. Only research utilizing commercially accessible consumer-grade smartphones, whether personal or borrowed, were thoroughly reviewed. We excluded studies utilizing smartphone microphones or video cameras for activity classification, as they may capture data regarding an individual's environment, including details about unconsented individuals, thereby impeding the wide use of this method due to privacy issues. To focus on research that mimicked free-living environments, we neglected those utilizing equipment attached or glued to the human body in a fixed way.

2.2 Review Conducting:

The study selection procedure adhered to the PRISMA principles, as illustrated in Figure 1. A collection of 461 articles were first identified through searches across seven databases: (Scopus 112, IEEE 66, ResearchGate 97, Springer 29, Web of Science 72, ACM 23, and PubMed 62) initially identified 461 records. After the removal of 152 duplicates, 309 records were screened based on their titles and abstracts. After screening titles and abstracts, 68 articles were assessed for eligibility based on the inclusion criteria, of which 36 records were screened based on full-text reading. During the entire article review stage of the secondary search, more relevant research was identified by examining the references for the selected articles., and new 8 studies met the inclusion criteria and were incorporated into the systematic review, resulting in 71 studies being involved in the final synthesis. All of the included studies had their critical data extracted using a standardized form. Items such as study ID, published year, authors name, research method, and associated limitations were retrieved from the database. The proportion of articles sourced from each database is displayed in Figure 2. Accordingly, several HAR techniques, datasets, and results were reviewed and compared to identify noticeable trends and conclusions.

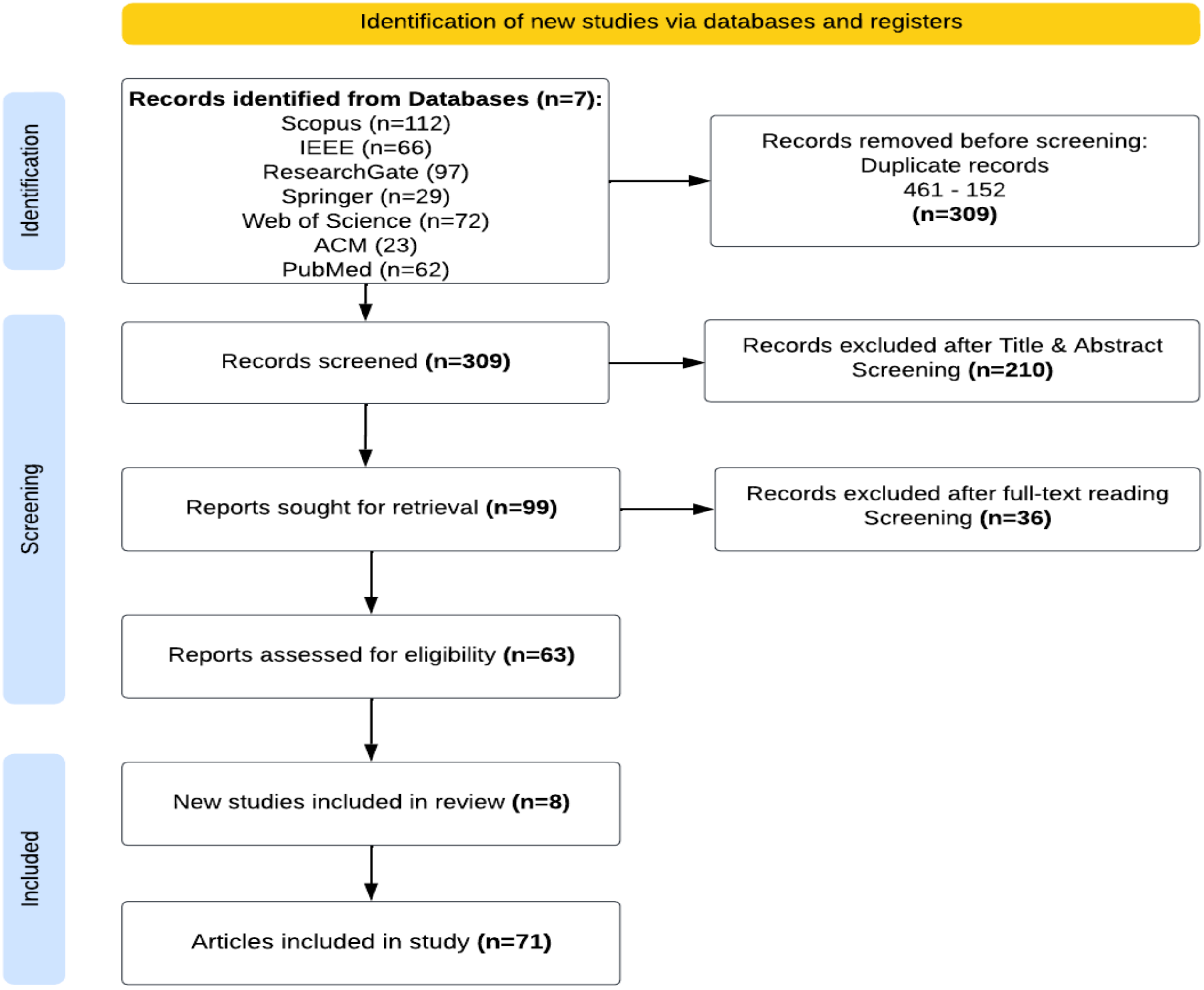

Figure 1. *PRISMA Flowchart*

This process enabled the determination of various HAR algorithms, which include RNN, SVM, CNN, DL, and ML, and made possible a comparison of their performance measurement in detail regarding HAR outcome prediction.

2.3 Review Reporting

A selection of primary studies of the literature were reviewed using the quality evaluation questions during the reporting phase. Each study's validity was determined using the quality evaluation criteria mentioned in Phase 1 above. A score of 0 meant poor quality or lack methodological detail, while a score of 6 represented good quality and solid technique. The study evaluation of quality varied from 0 to 6.

2.4 Scope Validation: Ensuring the Accuracy of the Selected Articles

Figure 3 illustrates how the bibliometric analysis identified a total of 44 keywords from the collected papers. To verify our investigation's scope, we evaluated these terms and classified them based on their co-occurrences. Then, we defined a threshold indicating all co-occurrences for every key word throughout all publications. Therefore, we found 10 terms from a total of 44 that satisfied the criteria. All of these ten keywords appeared a minimum of 3. Figure 4 shows the connections between these ten terms. The size of every circle represents how frequently a particular relevant keyword appeared. The larger the circle size, more frequently a keyword appears. Hence, the term "Human Activity Recognition" shows the largest circle size within the diagram, indicating that it appears the most frequently in the gathered articles. The second factor involves the color, that indicates how frequently a specific keyword appears each year. The last aspect is the total link strength, which indicates the total connection of a keyword to other keywords. The more frequently two terms appear in the identical article, the thicker the line between them. As an example, the phrases "deep learning"

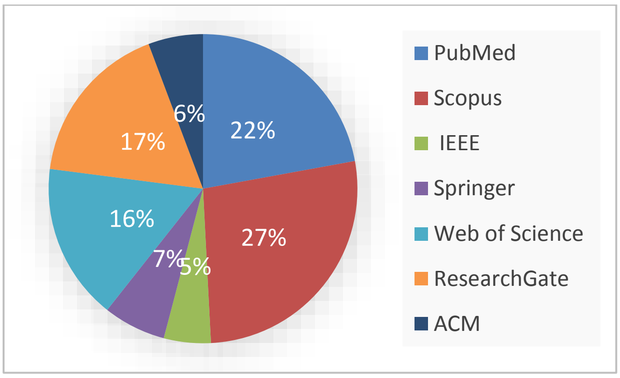

Figure 2. *The proportion of research articles collected from all databases*

and "Human Activity Recognition" are connected by a thicker line rather than the line that connected the phrases "machine learning" and "Human Activity Recognition," indicating that "deep learning" and "Human Activity Recognition" was seen together in the collected articles more than "machine learning" and "Human Activity Recognition." This indicates that DL has been used less than other ML techniques in Human Activity Recognition.Table 1 lists the ten keywords' occurrences together with their overall link strength. The most often appearing keyword is Human Activity Recognition, which has specifically shown up 32 times in the gathered papers and 40 times alongside other keywords. Deep learning happened 23 times with 36 links to other keywords; keyword machine learning happened 18 times with 28 links to other keywords. Finally, these highest ratings for the specified keywords statistically proved the validity in our search query utilized to compile scholarly works. Moreover, it shows that our study focus is on three primary keywords: Machine learning, deep learning and Human Activity Recognition as among them their circle sizes are the highest and their link is strongest.

3. Literature review

3.1 Human Activity Recognition Methods

According to the different methods of human activity data collection, human activity recognition mainly uses contact sensors, non-contact vision and wireless signals [8]. Contact sensors are a type of sensor that can detect the physical characteristics or movement of an object by contacting its surface. Due to direct contact with the target object, it can provide more direct and accurate measurement results [9]. Moreover, because contact sensors can directly contact the target object and measure its specific

Table 1. *Keywords occurrence*

Keyword	Occurrences	Total Link Strength
Human activity recognition	32	40
Deep learning	23	36
Machine learning	18	28
Sensors	10	18
CNN	9	17
Wearable sensors	8	12
Action recognition	8	11
Activity recognition	10	10
Classification	7	4
healthcare	6	4

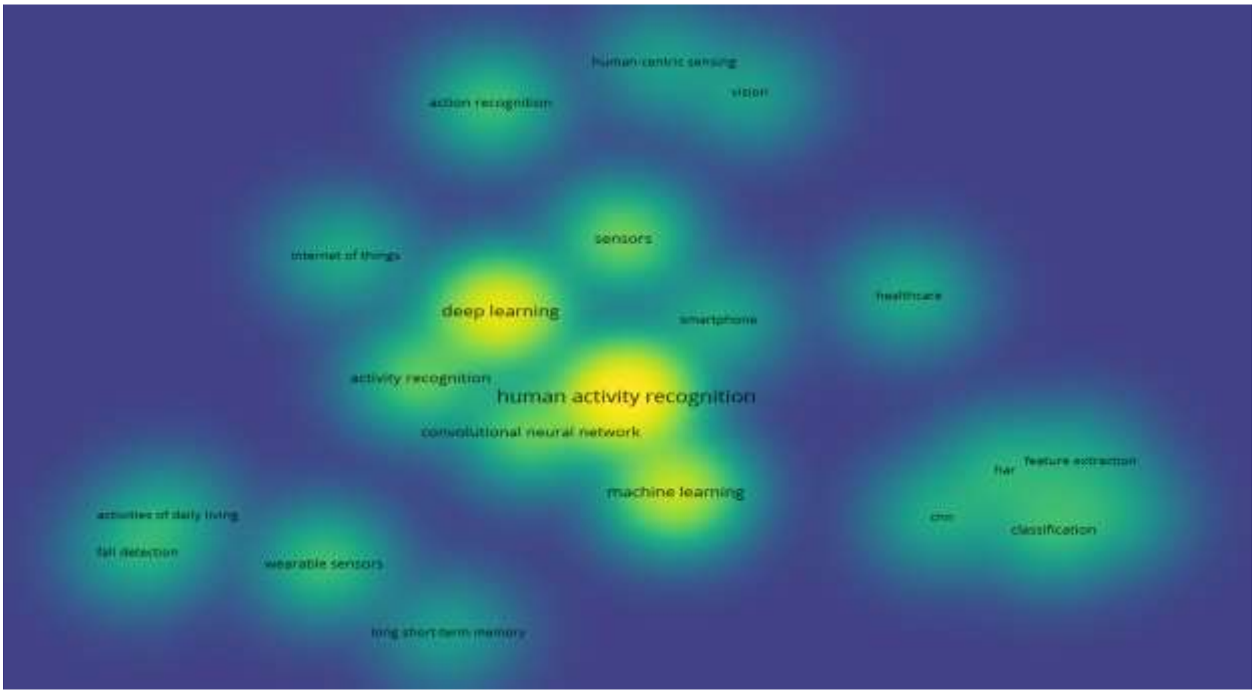

Figure 3. *Density bibliometric analysis Flowchart.*

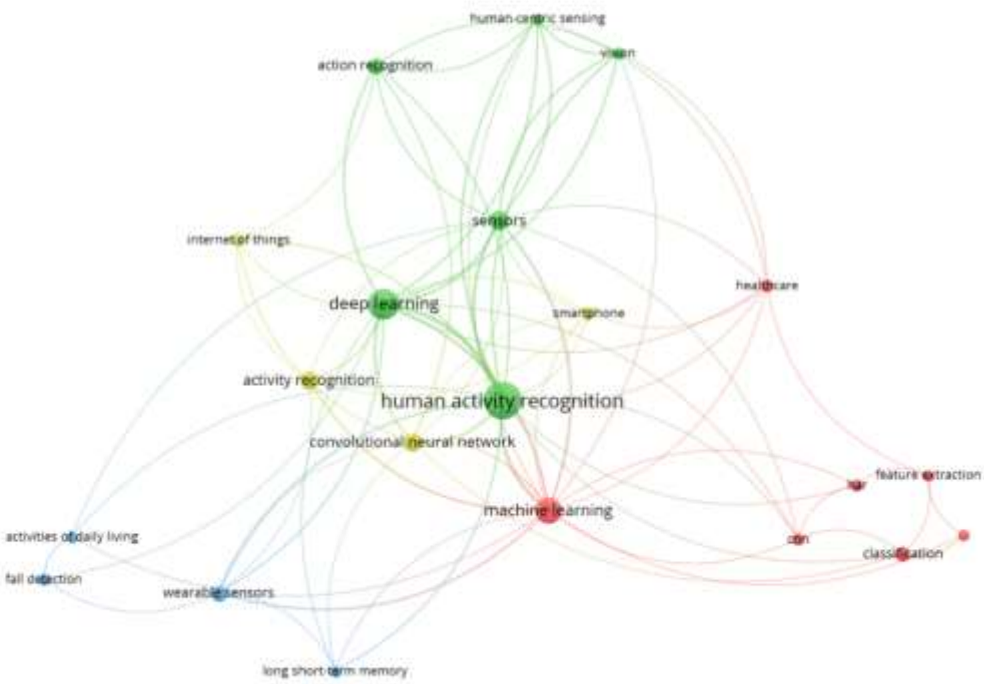

Figure 4. *Diagram of scope validation.*

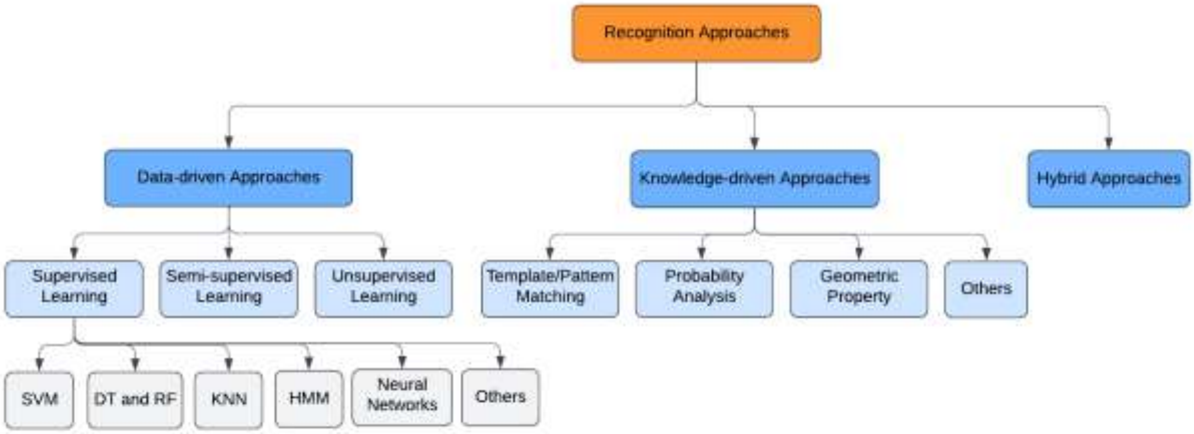

Figure 5. *Methods of recognition in HAR based on mobile devices.*

parameters, they usually have high measurement accuracy, which makes them effective in applications that require high-precision detection [10]. In addition, contact sensors directly contact the target object and measure parameters in real time, providing almost instant feedback and data output. However, the contact sensor needs to be in contact with the target object during the measurement process, which may have a certain impact on the target object. At the same time, since the contact sensor needs to be in direct contact with the target object, it is necessary to consider the wear and damage of the sensor during long-term use. Factors such as the surface texture, hardness and shape of the object may also affect the sensor [11]. With the pre-processed sensor data, the recognition approaches will be adopted to map the sensor data into one type of activity. Recognition methodologies are typically categorized into three types: data-driven approaches that employ training data to derive clusters or classifiers for activity recognition; knowledge-driven approaches that leverage domain knowledge to analyze sensor data with activity recognition; and hybrid approaches that integrate both data-driven and knowledge-driven methods, as illustrated in Figure 5.

Machine Learning Classifier

Traditional machine learning methods are the most commonly used methods in sensor-based HAR research. Compared with the emerging deep learning classification models, these algorithms are usually less computationally intensive and simple to implement and are more suitable for ubiquitous computing environments [12,13].

Decision tree (DT): Decision tree is a very common classification algorithm, which has the advantages of simple concepts and easy-to-understand classification rules. Intuitively, DT is a set of multi-branch trees with classification conditions. The computational model has low complexity and high efficiency. However, the DT model is prone to overfitting and is more suitable for classification problems with discrete features [12]. The currently more commonly used CART (Classification and Regression Tree) model optimizes this situation. CART uses a Gini index estimation function based on the minimum distance, which is very flexible and allows for partial misclassification of samples [14].

Support vector machine (SVM): Vaonik et al. [14] of Bell Labs formally published a paper proposing the superior performance of the SVM algorithm in classification tasks [15]. The improved soft-margin nonlinear SVM algorithm can be regarded as a landmark achievement in SVM-related research on handwritten character recognition [16]. Initially, SVM was mainly aimed at binary classification problems. The algorithm idea is to find the optimal hyperplane that can distinguish the two categories and has the largest margin between the two categories in the sample space of the data [17-19]. Support Vector Machine (SVM) is a highly effective classifier that has been utilized or enhanced in numerous analogous investigations.

Naïve Bayes classifier: The naïve Bayes model's way of recognizing things varies from most ways of classifying things. Its main goal is to find the form of the combined distribution P(X, Y). The formula of P(X, Y) is as follows [20]:

$$P(X,Y) = P(Y \mid X) * P(X) \qquad (1)$$

what X stands for and what Y stands for is the action label. The naïve Bayes algorithm works well with small amounts of data and is simple to set up [17]. It does, however, assume that different traits don't depend on each other, so the performance of the naïve Bayes classifier on data with high coupling characteristics is not very good.

K-nearest neighbor: The K-nearest neighbor technique is a basic method for regression and classification. In its most basic form, it uses majority vote as a lazy way to learn. The core is to use a distance metric to count the labels with the largest number of nodes closest to the target point and assign them to the target point [21]. Table 2 shows research work using ML classifiers.

Deep Learning (DL) Models

In the last few years, traditional pattern recognition techniques have advanced remarkably. But these techniques can rely mostly on heuristic hand-crafted extraction of features, which could limit their generalizing capability. Recent developments in DL have made autonomous high-level feature extraction feasible, hence enabling interesting performance in many domains. With little to no data and no feature engineering, DL techniques including RNN and one-dimensional CNNs have proved to produce state-of-the-art results on difficult activity detection challenges [37,38]. DL is one of the most current and novel ways since it allows for extracting features from input. CNN is the common type of DL algorithm which is highly effective in many different applications focused on learning complex behaviors and actions. CNN has the property of not being affected by scale invariance [39-41]. Furthermore, HAR faces challenges related to the ground truth annotation, such as identifying a specific activity of several users. Other issues include sensors containing fault values, sensing unit heterogeneity, and activity models being implemented in the form of one kind of domain to another. Avilés-Cruz et al., (2019) [42], suggests a book CNN method is used to track the movements of a single smartphone user. It took three CNNs working at the same time to pull out local features, which were then combined during the classification step. Walking, going up and down stairs, reclining, standing, or lying are the six actions that have been successfully grouped. A test of the suggested CNN's performance shows that it can recognize things 100% of the time on average. Gjoreski et al., (2020)[33], proposed an approach based in DL techniques for the purpose of HAR. These activities, involving walking, running, biking, driving a car,taking a bus, and using a subway, were determined by the analysis of smartphone sensor data. An additional model, which relied on the Hidden Markov Model (HMM), was introduced. Additionally, a baseline DL model was depicted. The results show that the use of an end-to-end DL structure, the multi-ResNet obtained an accuracy of about 89.4%. After hyper-parameter optimization, the models significantly improved. By combining the ML and DL models a significantly enhanced accuracy (from 89.4% to 93.3%). In addition, the accuracy was increased by 4% (from 93.8% to 97.8%) when utilized the HMM smoothing algorithm. Huang et al., (2021) [43], Propose a shallow CNN that analyzes cross-channel communication in a HAR scenario, in which all channels in the same layer interact comprehensively to capture more distinguishing features of sensor input. Extensive experiments on multi benchmark HAR datasets, including UCI-HAR, OPPORTUNITY, and UniMib-SHAR, demonstrate

Table 2. *Research Work Using ML Classifiers*

Authors	Classifier Model	dataset	Accuracy	Limitations\Gaps
Wang et al. [22]	SVM	SC	96.49%	Accuracy varies with body posture and external noise, Furthermore, the number of recognitions is limited due to difficulty in estimating the directionality of gestures.
Watanabe et al. [23]	RF	SC	68.9%-87.1%	Lower recognition Accuracy and model still face challenges with complex or subtle gestures, and with limited Gestures, the study tested only six gestures with four participants.
Xu et al. [24]	SVM	SC	94.80%	While the system can recognize several inattentive driving events, it may not cover all possible scenarios, the system's accuracy might vary depending on the driving environment and external noise levels, as well as sample data Size and the data collection period might not capture long-term variations in driving behavior.
Vu et al. [25]	RF	SC	92%-95%	The dataset may not cover a wide range of modernization scenarios and there is a need for better evaluation metrics to assess the effectiveness of modernization strategies comprehensively.
Lu et al. [26]	SVM, SVDD	SC	90.21%, 93.1%	The dataset is small sized and limited to one dataset.
Voigt et al. [27]	RF	SC	60.6%-94.6%	Limited number of participants, which may affect generalizability, as well as the performance influenced by lighting, background clutter, and occlusions which not take in consideration.
Lu et al. [28]	SVM	SC	94.6%, 98.4%, 96.3%	Evaluations were conducted only in stationary conditions, not under mobile conditions. Further investigations are needed to understand performance in more dynamic scenarios. Furthermore, the study did not include user elicitation studies to design gestures in particular scenarios.
Wang et al. [29]	SVC	SC	77.89%-84.38%	The evaluations are likely conducted under controlled conditions, which may not fully represent diverse real-life usage scenarios, and the dataset used for training and evaluation may not encompass a wide range of user behaviors and eye movement patterns, limiting the model's generalizability.
Shiet al. [30]	SVM	SC	Above 95%	The current context awareness module only includes dynamic, semi-static, and static contexts, which may not cover all real-world body postures such as lying down, walking up and down stairs, and jogging. Additionally, the study's data collection primarily involved participants only aged 19-30.
Caoet al. [31]	RF	SC	Precision: 95%, Recall: 94.84%	The study evaluated a small set of tongue-jaw movements with twenty participants. Expanding the dataset to include a broader range of movements and more participants would enhance the system's robustness, and variations in ear canal shapes among different users can impact the consistency of acoustic reflections and recognition accuracy.
Chenet al. [32]	RF	SC	Fl: 96.49%	Weak accelerometer readings along the three axes can lead to difficulty in detecting respiration patterns. Furthermore, the clinical study does not control wrist movement during sleep, which can lead to inconsistencies in the data and potential misclassifications due to sudden changes in acceleration data.
Gjoreski et al., [33]	HMM	UCI-HAR	93.3%	The model accuracy needs to improve.
Chang et al. [34]	KNN, HMM	SC	Above 80%	Low accuracy and validation limited to one dataset.
Hamatani et al. [35]	CRF	SC	Above 82%	User Interaction: Reliance on users consistently wearing and interacting with the smartwatch may impact data reliability. The generalizability: Results may not be broadly applicable due to the study's specific sample and conditions.
Jiokeng et al. [36]	LR	SC	92.57%	While the study involved over 100 volunteers, expanding the sample size and including a more diverse participant pool could help validate the generalizability of the system across different demographics.

that the proposed approach enables shallower CNNs to gather more useful information than baseline deep networks together with competing methods. The pace of inference is measured by deploying HAR systems in an embedded system. The results show that the model achieves 92.4% in average accuracy. However, stacked autoencoders are composed of many encoders placed on top of one another. Therefore, in this application scenario, utilizing stacked autoencoders is optimal [44]. The stacked autoencoder is employed to pretrain every single layer in an unsupervised manner in order to achieve improved weight initialization when performing subsequent supervised training. This improved weight initialization typically mitigates the possibility of disappearing gradients for feed forward neural networks [45-47]. An advantage of utilizing these strategies lies in their capacity to automatically develop features. In addition, the stacked autoencoders method demonstrates superior classification outcomes compared to the current leading methods, assuming that the parameters associated with the deep neural network are appropriately adjusted [44,48]. Prabono, Yahya and Lee, (2021) [49], presents a new approach for domain adaptation within the framework of Human Activity Recognition (HAR). The suggested technique employs a two-phase autoencoder framework, achieving a model performance of 67.9%. Moreover, Garcia et al. (2021) [69] presented a highly effective multi-class methodology that consists of a group of autoencoders, each autoencoder being distinctly associated with a specific class. This procedure exclusively necessitates the integration of new autoencoders, obviating the requirement for retraining the complete model. The findings indicate that the model attains an average accuracy of 71%. Challa, Kumar and Semwal, (2022) [50], With the help of ubiquitous sensor data, Challa suggested a robust classification framework for HAR. Additionally, she proposed a hybrid deep learning model that integrated CNN and BiLSTM. The model is able to learn both short-term and long-term relationships at sequential data thanks to the model's capabilities. For the purpose of determining whether or not the model is effective, the benchmark datasets WISDM, UCI-HAR, along with PAMAP2 are utilized. There was a 96.05% accuracy on the WISDM dataset, a 96.37% accuracy on the UCI-HAR dataset, and a 94.29% accuracy on the PAMAP2 dataset. Table 3 presents a comparison of those works that are relevant to HAR.

Table 3. *Research Work Using DL Classifiers*

Authors	Classifier Model	dataset	Accuracy	Limitations\Gaps
Lu et al. [51]	RNN	SC	Above 90%	Small Size of Data and limited training data can impact the robustness and generalization of the authentication framework, and the dataset might not capture the full range of user behaviours, which could affect the system's ability to recognize genuine and impostor inputs accurately.
Lu et al. [52]	RNN	SC	92.7%, 91.4%	More complex or less predictable password patterns can reduce the accuracy of the system.
Hassan et al [53]	DBN	HAPT	95.85%	A limited number of participants might restrict the generalizability of the results. Furthermore, the dataset may not cover a wide range of activities, impacting the model's ability to generalize, and using single dataset.
Du et al. [54]	CNN	SC, HHAR	79.80%	The system may struggle with real-time processing and recognition due to the computational demands of DL models. Furthermore, the dataset may not encompass a wide variety of writing styles, surfaces, and tools, limiting the generalizability of the results.
Gong et al. [55]	CNN	SC	Above 65%	The few-shot learning approach may still struggle with extremely limited data, impacting its effectiveness.
Becker et al. [56]	CNN	SC	97.20%	Validation limited to one dataset.
Brunner et al. [57]	CNN	SC	Fl: 97.4%	The study may have a limited number of participants, which can affect the generalizability of the results, and limited to single dataset with small size.
Liu et al. [58]	RNN	SC	92.7%, 91.4%	Compared to camera-based and other advanced tracking methods, ArmTroi's accuracy is lower, highlighting an area for improvement, and it needs pre-knowledge of user-specific metrics (e.g., torso length, shoulder breadth, upper-arm length) to generate point clouds.
Hou et al. [59]	LSTM	SC	Above 99%	The performance of the system can be affected by practical issues, such as how the smartwatch is worn on the user's wrist. Ensuring the system performs well under various practical conditions remains a challenge, as well as collecting a large amount of training data from individuals is time-consuming and not user-friendly. Training a user-specific model requires significant data, but a more practical solution would be to fine-tune a generic model for each user through user adaptation.
Giallanza et al. [60]	CNN, RNN	SC	27%-41.8%	Both the real-time analysis and the primary data collection included few users, and limited Training Data.
Liu et al. [61]	CNN	SC	Above 80%	Handling both uppercase and lowercase letters in the same word requires additional steps to ensure accurate recognition, such as coordinate transformation and scaling.

Yin et al. [62]	LSTM	SC	64.96%-94.86%	The word selection method based on the multi-class bi-gram language model may not be sufficient to handle all possible letter combinations and word formations, affecting recognition accuracy in practical use.
Gao et al. [63]	DNN	SC	WER: 8.33%	The study used a limited number of training samples, which may result in skewed phoneme-level samples. Expanding the dataset to include more diverse examples could improve the system's accuracy and robustness.
Liu et al. [64]	DNN	RealWorld	86.11%	While the global attention module is designed to be lightweight, the added complexity of integrating attention mechanisms can still introduce some computational overhead, especially in more resource-constrained environments.
Wang et al. [65]	CNN, LSTM	SC	98.40%	The system assumes that the user's torso and device are relatively static. In practice, dynamic conditions (e.g., walking with a mobile phone in the pocket) can lead to inconsistent hand movements and affect performance.
Park et al. [66]	CNN	SC	91%	With nearly 5,000 commonly used words in most sign languages, further research is required to validate SUGO's scalability. Although 50 words with diverse motion characteristics were chosen, expanding and testing the system with a larger vocabulary is essential for real-world deployment.
Zhai et al. [67]	Res Deep CNN	AWSS	78.20%	The study uses two publicly available datasets, which may not fully represent the diversity of sleep patterns across different populations and environments. Expanding the dataset to include more diverse samples could improve the model's generalizability.
Chen et al. [68]	GRU	SC	90%, 91%	Extensive data collection is necessary to train the model effectively. Gathering a large, diverse dataset can be time-consuming and resource-intensive, but it is crucial for improving the system's accuracy and generalizability. While five training-free gestures achieve 100% accuracy, the study needs to expand the range of gestures and validate their accuracy in different real-world scenarios.
Ouyang et al. [69]	DNN, CNN	SC	Above 80%	Although raw user data is not exposed during the learning process, model updates transmitted in ClusterFL may still reveal certain information about user activities. Future work should integrate privacy-preserving techniques and investigate the trade-off between privacy and utility.
Khaertdinov et al. [70]	CNN, Transformer	MobiAct, HAR	Fl: 81.13%, 91.14%	The study uses three widely used public datasets. To fully validate the robustness of the method, it is essential to test it across a more diverse range of datasets, representing different environments and user behaviors.
Zhang et al. [71]	CNN, ED	SC	91.2% in Accuracy; 7.1% in WER:	The current system achieves high accuracy with specific datasets. Expanding the dataset to include more diverse user scenarios and conditions is necessary to validate its generalizability and robustness.
Zhang et al. [72]	CNN, ED	SC	CER: 9.3%, 3.8%	While WriteAS demonstrates promising accuracy for recognizing continuous handwriting, it may still face challenges in adapting to various individual writing styles, especially for new users.
Lu et al. [73]	CNN	SHAR. HAR	Above 60%	Collecting sufficient labelled data for building human activity recognition (HAR) models is expensive and time-consuming. Training on existing data can lead to models being biased towards the training data distribution, which may not generalize well to test data with different distributions.
Xu et al. [74]	CNN	SC	55.3%-87.2%	The user study involved a relatively small number of participants (N=20), which may not fully capture the diversity of user behaviours and preferences. Expanding the study to include more participants and diverse scenarios is necessary for broader validation.
Xie et al. [75]	CNN	SC	93.44%	The system relies on detecting hand movements and chest fluctuations, which can be influenced by other activities or movements, potentially leading to false positives or negatives.

Ling et al. [76]	CNN	SC	99%	Although the system demonstrates high accuracy for 12 gestures, expanding the dataset to include a wider range of gestures and more complex movements is necessary to ensure broader applicability and robustness.
Mollyn et al. [77]	CNN, VGG	SC	92.20%	While the system achieves a recognition accuracy of 92.2% across 26 daily activities, there is still room for improvement, especially in more complex or noisy environments
Wang et al. [78]	ResNet 18	SC	96%	The study needs to include a wider range of gestures and more complex hand movements to ensure the system's robustness and generalizability across different scenarios
Tanigaki et al. [79]	AIP-Net	RealWorld	Added by 20%	The system's ability to generalize across different datasets and activities is essential. Further validation is needed to assess how well AIP-Net performs on a wide range of HAR tasks beyond those included in the study.
Raza et al. [80]	Transformer	WISDM	98.89%	• The dataset used in the study was constructed with only five human participants in a simulated home care scenario. This limited and artificial setting may not fully represent real-world conditions. More realistic home healthcare settings and larger, diverse participant pools are necessary to validate the findings. • The current dataset was created under controlled conditions. Expanding data collection to include real-world environments and diverse user groups will be crucial for validating the proposed HAR classification method.
EK et al. [81]	Transformer	HAR, SHL, MotionSense, RealWorld, HHAR	92.6%-98.72%	• Although MobileHART shows robustness in handling domain shifts, further research is needed to understand its limitations in completely unseen situations and diverse real-world conditions. • The datasets used in the study may not cover all possible activities and scenarios. Expanding the dataset to include a wider variety of activities, environments, and user behaviours is crucial for ensuring the model's generalizability.
Miao et al. [82]	DNN	RealWorld	over 75%	The framework shown promising results in controlled experiments, its ability to generalize across different real-world scenarios needs further validation.
You et al. [83]	SAN	SC, SHL, SHO	over 83%	The study may need to include a wider range of activities and scenarios to validate the model's robustness and generalizability.
Yoshimura et al. [84]	LOSNet, ED	SC	over 50%	While the Lightweight Orderedwork Segmentation Network (LOSNet) shows promising results in recognizing repetitive works, its performance may vary across different industrial settings and activities.
Li et al. [85]	RiskNet	SC	80.10%	The study recruited 988 healthy individuals and 417 patients with various conditions (PD, TBI, stroke). Expanding the dataset to include a wider range of activities, environments, and user behaviours would help improve the system's robustness and generalizability, and lower accuracy achieved need to improving.
Augustinov et al. [86]	DNN, Transformer	CogAge	73.36%	Complex Activity Variability: The study focuses on seven ADLs, but the results may vary with a broader range of complex activities. Further validation is required to ensure consistent performance across diverse activities.
Sharma et al. [87]	CNN, Transformer	harAGE	Recall: 75.9%	The study may need to include a wider range of activities and scenarios to validate the model's robustness and generalizability.
Zhang et al. [88]	CNN, Transformer	SHL2018, HEAR, HAPT, MobiAct MotionSense, ,	Fl-score: 78.55% and 95.66%	Although the study used eight benchmark datasets, expanding the dataset to include a wider variety of activities, environments, and user behaviours is essential for validating the model's robustness and generalizability.
Lee et al. [89]	CNN	SC	97.44%	• The system's reliance on sound characteristics means it is susceptible to external noise, which can affect the accuracy of transportation mode recognition. • The study used seven different Android phone models for data collection. However, the performance across a

				broader range of smartphone models with varying hardware specifications needs to be assessed to ensure robustness and generalizability.
Ding et al. [90]	TSNN	SC	75.3%, 86.4%, 79%	The study may need to include a wider range of handwriting samples, including different languages, scripts, and writing styles, to validate the model's robustness and generalizability.
Mishra et al. [91]	DNN	TMD, SHL, SC	59.41 % and 94.21%	The study may need to include a more diverse range of locomotion modes, user behaviours, and environmental conditions to validate the model's robustness and generalizability.
Xiao et al. [92]	CNN, ED	SC	83.8% and92.2%	Single dataset for validation and the accuracy needs to improve.
Avilés-Cruz et al., [93]	CNN	UCI-HAR and WISDM	Average recognition of 100%.	The validation is limited to one class (walk).
Huang et al., [43]	shallow CNN	UCI-HAR, OPPORTUNITY, UniMib-SHAR UCI-HAR WISDM	92.4%.	Accuracy needs to be improved, Model used based shallow CNN while it needs to design a new model.

Hybrid Deep Learning Models

A hybrid deep learning model involves the integration of several methods for modeling with DL to create a more proficient model for accurate streamflow forecasting, moving away from using solo DL models. Hybrid DL methods have a primary benefit over solo models in that they can amalgamate the advantages of many methods of modeling for streamflow forecasting. Hybrid algorithms can expedite convergence and boost performance by amalgamating diverse methods. In addition, hybrid algorithms are often more efficient than pure optimization algorithms in tackling engineering challenges [94-96]. A hybrid model combines the characteristics of LSTM networks in sequence learning with the spatial feature detection capabilities of CNN. Before extracting important spatial features from input data, like video frames and images, the CNN processes this configuration. It is then followed by LSTM, which is capable of identifying temporal patterns across time [97,98], [99]. For tasks requiring awareness of both time and space, such as identifying actions within videos or comprehending the sequence of events within a series of images, this combination is especially effective. In order to train this model, data is typically fed through CNN, which provides the LSTM with detailed feature maps that are utilized as input. Subsequently, the combined system is fine-tuned to improve its ability to predict or classify complex, sequence-dependent data [43], [100]. Mukherjee et al., (2020) [101], shows a new group of three classification models, CNN-Net, Encoded-Net, and CNN-LSTM. The name for this group is Ensem-ConvNet. All of the above classification models are built on top of a simple 1D CNN. But these models are not all the same when it comes to the number from dense layers, the size of the kernel, and other important design differences. Each model looks at time series data within the form of a 2D grid. At any given time, a window of data is looked at to find useful information and make guesses about what people are doing. Three standard datasets are used to test their suggested model: MobiAct, WISDM activity forecasting, and UniMiB SHAR. They compared their EnsemConvNet model to a number of well-known deep learning models, such as Multi Headed CNN, which is a mix of CNN and LSTM models. When compared with the other models talked about in this study, the Ensem-ConvNet system is more accurate, reaching up to 97.7%. Ihianle et al., (2020) [45], proposed a methodology for detecting the existence of Human Activity Recognition via Deep Learning techniques. The suggested algorithms are based on CNN and Bi-LSTM. Bi-LSTM exhibited its superiority by traversing both backwards and forwards through a specified sequence to augment the extracted characteristics. The final model was proposed based on multi-channel convolutional Bi-LSTM (MCBLSTM). Two principal datasets, MHEALTH along with WISDM, were utilized to assess the efficacy of the proposed methodologies. Utilizing each data set, the accuracy achieved by MCBLSTM exhibited the maximum performance at 97.7%. In regard to the research work using these classifiers, they can be found in Table 4.

4. Application of human activity recognition

The data provided by contact sensors are also becoming increasingly popular due to their widespread availability, ease of installation and non-intrusive nature. Many smart products, including watches and smartphones, integrate inertial sensors such as accelerometers and gyroscopes [111,112] that can continuously record various human activity data to identify human movements and detect human

Table 4. *Research Work Using Hybrid Classifiers*

Authors	Classifier Model	dataset	Accuracy	Limitations\Gaps
Shen et al. [102]	HMM, CRF, RF, DT, SVM	SC	Over 90%	Accuracy needs to improve, use one dataset.
Ahuja et al. [103]	SVM, CNN	SC	Between 81.4% and 100%	Accurate eye tracking may require precise calibration, which can be challenging with off-the-shelf virtual reality headsets, and the system may experience latency issues, affecting real-time interaction quality.
Yang et al. [104]	DenseNet, SVM, C4.5 DT	SC	over 93%	The ability of the model to generalize to different geographical areas or conditions not covered in the training data might be limited. Furthermore, the model was tested on specific benchmark datasets, which may not fully represent the variety of real-world scenarios, limiting the applicability of the results to other contexts.
Luo et al. [105]	LSTM, KNN, DT, SVM, AT-	SC	Over 90%	Distinguishing between different brushing actions remains challenging due to the subtle differences in the movements, which may impact the system's precision, as well as the dataset may not include a diverse range of users, potentially limiting the generalizability of the system across different demographics.
Yin et al. [106]	RF, CNN	SC	91.6%, 94.3%	During the writing process, if a user is walking, the sensor data gets mixed between body and arm movements. The system needs to filter out body movements to accurately recognize arm gestures.
Zhang et al. [107]	MT-KNN, OCSVM	SC	Achieve EER about 4.9%	TouchID focuses on graphic pattern-based touch gestures, which have layout constraints. In some scenarios, touch gestures may not follow these constraints, affecting recognition accuracy. Additionally, the current system primarily supports single-touch gestures. The system's ability to handle multi-touch gestures.
Song et al. [108]	ResNet-50, BiLSTM, KNN	TFST	EER: below 2%	While the Feature Regularization Net (FRN) approach helps mitigate behavioural variability, it may still face challenges in handling extreme variations in user behaviour over time.
Bhattacharya et al. [109]	RF, NB, Deep Conv-LSTM, CNN, and Discriminate	SC	Fl: 89.7%-94.3%; 30%-55.8%	The semi-naturalistic dataset focused on one activity at a time and did not account for overlapping activities. In contrast, real-world settings often involve multitasking with multiple activities occurring simultaneously, which the model struggled to recognize accurately.
Challa, Kumar and Semwal, [110]	Hybrid DL model CNN-BiLSTM.	UCI-HAR and WISDM	96.05%, 96.37%, and 94.29%, respectively.	Complex model and high compositional requirements, and despite automatic feature extraction capabilities, handling the complexities of time-series data remains challenging. The model's ability to effectively extract relevant features from noisy and raw data needs continual improvement.
Ihianle et al., [45]	Hybrid CNN and Bi-LSTM	MHEALTH and WISDM	97.7%.	Complex model and missing automatically tune and adjust parameters.
Mukherjee et al., [101]	CNN-Net, Encoded-Net, CNN-LSTM.	WISDM, UniMiB SHAR and MobiAct	higher accuracy reach up to 97.7%.	Complex model and missing processing of time-series data before being fitted into a model.
Prabono, [30]	domain adaptation	UCI-HAR, OPPORTUNITY, PAMAP2	67.9%.	Complex model and low accuracy.

health.With the development of wearable human body sensing technology, human activity recognition can be used in the clothing industry. In recent years, textile-based sensors have been used for activity recognition. Using the latest electronic textiletechnology, sensors can be integrated into clothing, allowing users to wear them comfortably and enjoy long-term human movement recording. The recent development of electronic textiles has made it possible to integrate sensors into clothing [113-115]. This has significant advantages, such as the ability to capture natural behavior and ensure wearer comfort through unobtrusive sensing, and the sensors can be attached anywhere on the garment. Human activity recognition based on contact sensors can also be used to monitor athletes' posture, movements and strength to provide personalized fitness guidance and feedback. For example, by measuring the data of the pressure sensor, the correctness of the posture in weight training can be identified and the user can be reminded to adjust the posture [115-118]. In addition, human activity recognition based on contact sensors can also be used to monitor the range of motion, posture and activity level of rehabilitation patients to help evaluate rehabilitation progress and guide treatment plans. For example, by measuring the data of joint sensors, the range of motion of the joints can be monitored to help monitor and adjust rehabilitation training [119] Overall, the application of human activity recognition based on contact sensors has broad prospects and will play an important role in health, rehabilitation, smart products and other fields. As technology and algorithms continue to advance, we can look forward to the development of more accurate, intelligent and personalized human activity recognition applications. Human action recognition has attracted great attention in the field of computer vision due to its applications in the real world. Vision-based systems have a solid theoretical foundation and perform well in identifying human activities. Action recognition in computer vision is generally divided into action recognition in videos and action recognition in still images. Action recognition in still images aims to recognize human activities in static images without any temporal information. Since human activities such as running and smoking can be recognized through a single input image without additional motion cues, action recognition based on non-contact sensors has received great attention. Previously, due to the emergence of the COVID-19 pandemic, which caused people to avoid close contact with devices, the demand for accurate and efficient vision-based communication increased significantly [120]. Unlike traditional wireless mice and keyboards, vision-based interaction, in addition to providing effective remote control, can also control electronic devices without touching any part of the electronic device. With the popularity of electronic health in various smart home application fields in recent years, human motion recognition technology is increasingly used in rehabilitation systems, chronic disease management and personal health monitoring of the elderly. For example, by monitoring the daily behaviour of an elderly person, the assistant service can track how completely and consistently his daily behaviour is performed, and on this basis determine whether and when intervention or assistance is needed [122], for example, whether the person Falls in the bathroom, etc. Research can also be carried out around the activity sensing technology of unstructured intelligent space home care robots, and the integrated development and application verification of the monitoring robot prototype system can be carried out to provide a basis for the discovery of the elderly's activity intentions, medical diagnosis, behavioural intervention and active robot services. , laying a theoretical and technical foundation for the research and development of robot-based intelligent elderly care products. In addition, the development of autonomous driving also brings requirements for intelligence, safety and stability. Therefore, human action recognition technology also has important applications in this regard. In dynamic, complex and uncertain environments, it is necessary to establish an effective form of interactive cognition between pedestrians and vehicles. Vehicles need to make appropriate decisions before Detect pedestrians through video sensors, identify their body movements, and understand the meaning of their movements [123]. In the future, with the continuous development of deep learning and computer vision technology, human motion recognition will be applied in more fields, such as medicine, agriculture, textile industry, manufacturing [124], etc.

4.1 Applications and Challenges of HAR in Healthcare

HAR focuses on two main types of methods: those that use video and those that use sensors. Video-based methods give us a lot of features that let us do fine-grained investigation in HAR. However, it has a very complicated background of images because the data collection process needs a very strict environment with well-positioned cameras and people, which costs a lot in terms of price, computer power, and energy use. So, video-based methods are still not very useful in epidemiological studies that need a reliable, accurate, and inexpensive way to measure daily physical exercise. Sensor-based methods are commonly used in scientific studies of physical exercise because they are more flexible in changing environments, are more accurate at recognizing people, and use less power. Also, accelerometers are the most commonly used sensor for research because they are easy to acquire and come with most smart gear. Also, accelerometers are thought to be a pretty good tool for detecting a wide range of activities, since most of them involve simple body movements [120].

4.2 Developing an Intelligent Healthcare System based HAR

Activity recognition for the SHCS enhances healthcare for patient treatment and care, alleviates the workload of healthcare professionals, reduces expenses, and improves the quality of life for the aged. Medical professionals assert that automatic activity detection is an optimal method for identifying and uncovering new medical disorders to track daily activities [120,125,126]. HAR comprises six primary components: data gathering, preprocessing, extraction of features, feature selection, learning, and recognition. Figure 6 illustrates that preprocessing is an essential phase in data processing, encompassing discrimination, windowing, and filtering. Initially, the signals have been discretized, and their temporal and frequency properties are analyzed, which are extensively.

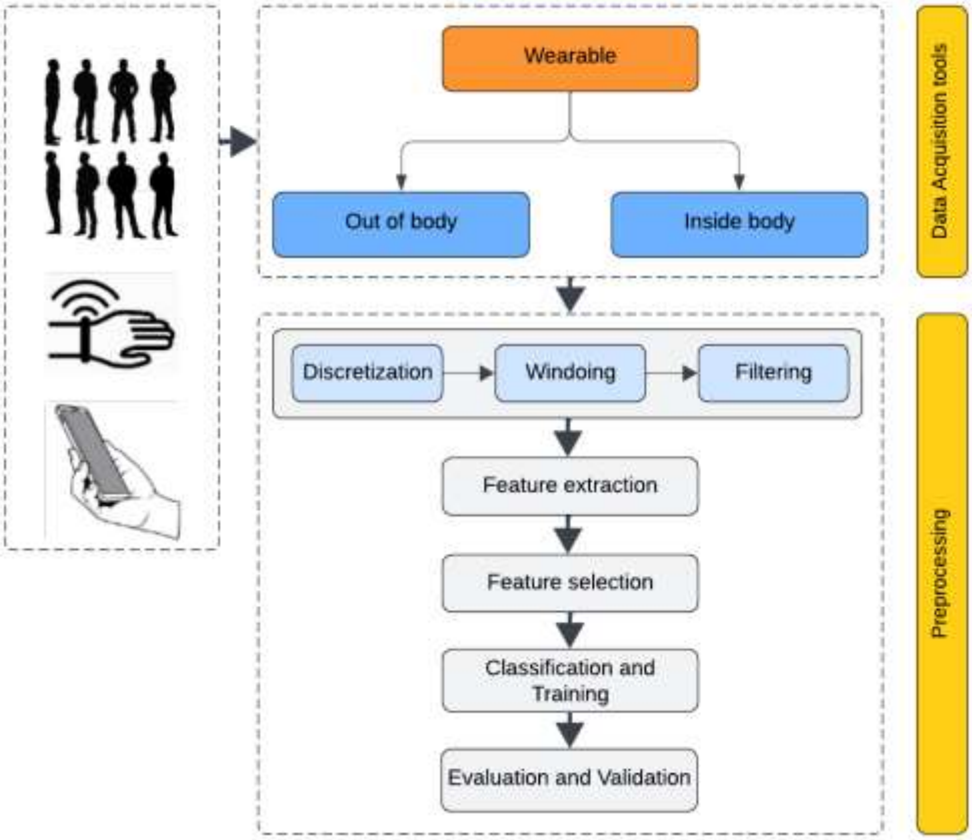

Figure 6. *Senior based HAR for Senior Healthcare Services*

utilized for feature calculation. The median, or the mean and variance are attributes of the temporal domain. Windowing techniques are employed to partition the sensor signals into segments [6,15,121,127]The most effective window-based strategies are based on activities, time-based, sensor-based, latest sensor-based, and sensor-dependent approaches. In activity-based frames, data are partitioned at the moment of activity transition. Activity streams of data are segmented into fixed time-based periods. In the sensing-based window series $(s_1,\ s_2,\dots,s_N)$, the sequence is divided into windows containing an equal amount of sensor inputs $w_1, w_2,\dots,w_M$, with the w_1 window denoted as $[(s_i\ -\ \Delta s), s_i]$. The outcomes of the window's length differ from one window towards another. In temporal dependence, 2 sensor events transmitted independently may belong to the same interval [122]. The filtering method assists in substituting missing values and eliminating outlier values. The e HAR component encompasses feature extraction and selection, followed by learning and recognition. Data mining represents the practice of analyzing data to uncover concealed knowledge. Feature extraction from raw data is executed with split and classification methods for each window, respectively. Feature extraction is performed both linearly and in a non- to diminish dimensionality, such as LDA and PCA techniques. Value-based criteria have been chosen to enhance the precision of activity identification. Feature selection methods encompass filtering (canonical analysis of correlation (CCA)), wrapper techniques (such as support vector machines (SVM) and neural networks (NN)) and embedding methods [123,128]. The gathered data must be transmitted to the HAR element for analysis utilizing technologies including Wi-Fi. Identified activity can be disseminated through [101], [123] upon the new hardware architecture. All components depicted in this design (Figure 6) are explained in the subsequent sections. However, as proved in real healthcare situations, several significant challenges emerge about the lack of available labeled data necessary for constructing a classification model in relation to the overall velocity and volume of sensor-generated data. Moreover, the discriminatory capacity of features is frequently challenging to ascertain across various classes, as the range of motion patterns in specific patient groups, such as those with obesity or even geriatric patients, is constrained and preserved throughout time. Another concern is the typical class imbalance present in data recorded from these sensor data streams. Samples reflecting various consistent postures, including sleeping, sitting, active, and inactive, are perceptually more abundant compared to others like jogging or climbing stairs. Consequently, these issues need the creation and implementation of hybrid data-driven methodologies, wherein semi-supervised models serve as the foundation of data processing processes, typically including contemporary Big Data technology [124,128,129].

5. Results and Discussions

Tables 2, 3, and 4 describe the characteristics of popular classifiers, specifically the common classifiers KNN, SVM, HMM, and RF, alongside contemporary neural networks, from various perspectives. In terms of input, classical classifiers often utilized feature vectors, whereas neural networks commonly employed preprocessed sensor data. Traditional classifiers typically utilized handmade features, whereas neural networks generally extracted features autonomously. Occasionally, traditional classifiers may use automated features derived from other neural networks. Traditional classifiers such as SVM, RF, KNN, and HMM mostly concentrated on obtaining efficient features, whereas neural networks emphasized the building of effective models and architectures. Regarding training size, the data utilized in SVM, RF, and KNN is often minimal, while the learning size for HMM is moderate, but neural networks generally require substantial training data. Neural networks frequently necessitate an appropriate number of samples for training the models. Typically, sensor data acquired for classical classifiers or artificial neural networks was transmitted to a server via Bluetooth, WiFi, or mobile data networks for model training, as the computationally intensive model training was generally conducted on a server rather than on mobile devices. The trained classifier/model can be installed on either a mobile device or a server. Considering model size along with computational overhead, classical classifiers such as SVM, RF, and KNN can be implemented on mobile devices for Human Activity Recognition, whereas neural network models are often deployed on servers. Consequently, while implementing neural networks for Human Activity Recognition (HAR), sensor data was typically relayed to a server to be processed and categorization. Regarding recognition performance, given effective features or well-crafted models, both the conventional classifiers and deep artificial neural networks can show a decent performance in activity recognition. Deep neural networks typically outperformed others, nevertheless, when there were enough training data. Regarding scalability, the conventional classifiers had low scalability since they rely on extracted characteristics, which could differ between one dataset to another. Using the same classifier in another scenario (or dataset), it needed to extract new features again. Differently, neural networks have higher scalability, because they can automatically extract features from sensor data, thus can be more easily applied to another activity recognition challenge (datasets). Table 1 shows that the DL based methods performed very differently on various datasets. With regard to computation overhead, SVM, KNN and RF generally have a low overhead; HMM has a medium cost, while neural networks typically have a high overhead. Regarding implementation complexity, the conventional classifiers—especially the SVM, RF, and KNN—which were usually used for lightweight devices are simpler. Regarding neural networks, they are difficult to apply on mobile vices and frequently run on a server. If an artificial neural network was supposed to work with mobile devices, the algorithms for compression and improvement are usually applied to reduce the complexity the network. Regarding frequency of use, the SVM and also artificial neural networks were most often used; the RF was likewise a little widely used; KNN and HMM had been less often utilized. These features of popular classifiers are supposed to be taken into account while choosing or building classifiers for activity recognition. In Figure 7, we provide the statistics of recognition approaches, including ML, DL and hybrid approaches. It can be found that a lot of research work preferred to adopt data-driven approaches, especially supervised learning-based approaches, which often have the common workflow. Neural networks with strong feature extraction capabilities were most commonly employed in supervised learning.

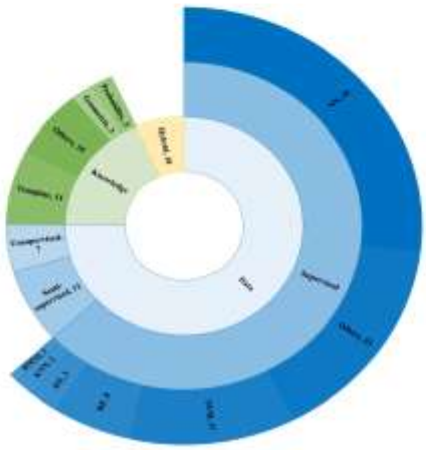

Figure 7. *Distribution of HAR approaches*

Another well-liked classifier with a lower processing overhead was the SVM classifier. Additionally, combining several classifiers was also commonly used for different recognition tasks and performance comparisons. Different from the most popular supervised learning-based methods, the semi-supervised learning based and unsupervised learning based methods occurred sporadically and were not often adopted. The hybrid and knowledge-driven approaches were less widely used and were typically created for particular needs. We also offer the statistics of current approaches for each year in order to better examine the research developments in recognition approaches. As shown in Figure 8, the data-driven approaches become more and more popular. The knowledge-driven approaches is less used than other approaches in period between 2018 to 2021 while it rised in year 2021 then it become close to hybrid approaches from 2022 to date. When considering the popularity of supervised learning based methods in data-driven approaches, we also provide the statistics of supervised learning based approaches in each year.

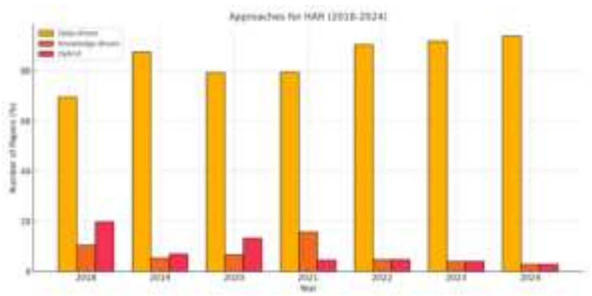

Figure 8. *HAR approaches through years*

As shown in Figure 9, the DL classifiers are more used than other techniques from 2018 to 2024 and it has been rise every year. Hybrid approaches is the second commonly used method especially in period between 2018 to 2020. Knowledge-driven approaches were less used than other approaches. However, it become preferred or close to hybrid models from 2020. Hence, it clear that the DL method is more likely to used in most approaches this due its higher performance compared to other techniques.

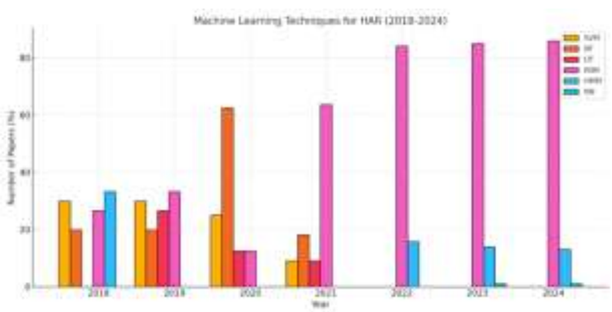

Figure 9. *HAR approaches-based ML/DL techniques*

In regard to other approaches using different or multiple classifiers, they were adopted as needed, and received good attention during these years. Designing approaches to mobile device-based HAR can be guided by the following uniqueness, qualities, similarities and research developments of recognition methodologies. Open issues are: First of all, particularly with deep learning-based approaches, it is usually required to give adequate training data if one is training a classifier for HAR. Still, gathering and labeling sensor data on human activities calls for a lot of work. For HAR, how to lower the cost for data annotation is fairly important and should be investigated more. Second, the current HAR systems cannot identify new-class events; they only acknowledge actions in designated classes. In a real-life situation, nevertheless, there are many different kinds of activities, therefore appreciating new-class events is significant and relevant. It is a difficult work and has not received enough research. Thirdly, a HAR solution should operate under several conditions considering the variations of people, surroundings, and tools. Still, the current methods were sometimes judged against a small number of possibilities. One should pay greater attention to the generalization in HAR methods. Fourthly, considering the restricted resources of mobile devices, many current methods—especially deep learning based approaches—process data offline and can rarely operate on mobile device. Encouragement of the mobile device as well as a server running HAR method helps to ensure a timely HAR feedback by means of data transfer. Furthermore envisaged is the invention of lightweight variants, which can fit mobile devices for HAR.

6. Conclusions

This paper takes the human activity recognition process as the main line and sorts of various key technologies and related work. Although the current research on human activity recognition technology based on sensors can achieve good results, it still faces many new problems brought by more complex scenarios in actual scenarios. For example, the system is not robust enough to the offset of sensor equipment, the amount of data for training deep learning classification models is too small, and the contradiction between balancing equipment energy consumption and service quality.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of 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

- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.

- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] H. Kaur, V. Rani, and M. Kumar, (2024). Human activity recognition: A comprehensive review," *Expert Syst,* 41(11);e13680, doi: 10.1111/EXSY.13680.

[2] S. A. Tadigadapa and N. Najafi, (2003) Developments in microelectromechanical systems (MEMS): A manufacturing perspective, *J Manuf Sci Eng*, 125(4);816–823, doi: 10.1115/1.1617286.

[3] V. Dentamaro, V. Gattulli, D. Impedovo, and F. Manca, (2024) Human activity recognition with smartphone-integrated sensors: A survey, doi: 10.1016/j.eswa.2024.123143.

[4] A. Ravuri *et al.*, (2024) A Systematic Literature Review on Human Activity Recognition,

[5] A. Búzás et al., (2024) Hierarchical organization of human physical activity, *Sci Rep*, 14(1), doi: 10.1038/s41598-024-56185-0.

[6] F. Serpush, M. B. Menhaj, B. Masoumi, and B. Karasfi, (2022) Wearable Sensor-Based Human Activity Recognition in the Smart Healthcare System, *Comput Intell Neurosci*, 2022;1391906, doi: 10.1155/2022/1391906.

[7] Z. Zhou, H. Yu, and H. Shi, (2020) Human Activity Recognition Based on Improved Bayesian Convolution Network to Analyze Health Care Data Using Wearable IoT Device, *IEEE Access,* 8;86411–86418, doi: 10.1109/ACCESS.2020.2992584.

[8] S. Zhang et al., (2022) Deep Learning in Human Activity Recognition with Wearable Sensors: A Review on Advances, *Sensors* 22(4);1476, doi: 10.3390/S22041476.

[9] T. Özyer, D. S. Ak, and R. Alhajj, (2021) Human action recognition approaches with video datasets—A survey, *Knowl Based Syst,* 222, doi: 10.1016/j.knosys.2021.106995.

[10] M. Munoz-Organero and A. Lotfi, (2016) Human movement recognition based on the stochastic characterisation of acceleration data, *Sensors (Switzerland),* 16(9), doi: 10.3390/s16091464.

[11] N. Tasnim and J. H. Baek, "Dynamic Edge Convolutional Neural Network for Skeleton-Based Human Action Recognition," Sensors, vol. 23, no. 2, Jan. 2023, doi: 10.3390/s23020778.

[12] S. Agac and O. D. Incel, (2024) Resource-efficient, sensor-based human activity recognition with lightweight deep models boosted with attention, *Computers and Electrical Engineering*, 117;109274, doi: 10.1016/J.COMPELECENG.2024.109274.

[13] K. Kaur and J. S. Batth, (2024). Human Activity Recognition Using Machine Learning, *2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0*, pp. 1–5, doi: 10.1109/OTCON60325.2024.10688067.

[14] T. Daniya, M. Geetha, and K. S. Kumar, (2020) Classification and regression trees with gini index, *Advances in Mathematics: Scientific Journal*, 9(10);8237–8247, doi: 10.37418/amsj.9.10.53.

[15] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, (2024) An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review, *Information (Switzerland)*, 15(4), doi: 10.3390/info15040235.

[16] J. Liu, L. W. Huang, Y. H. Shao, W. J. Chen, and C. N. Li, (2024) A nonlinear kernel SVM classifier via L0/1 soft-margin loss with classification performance, *J Comput Appl Math,* 437;115471, doi: 10.1016/J.CAM.2023.115471.

[17] A. Mammone, M. Turchi, and N. Cristianini, "Support vector machines", doi: 10.1002/wics.049.

[18] A. Rehman, F. Noor, J. I. Janjua, A. Ihsan, A. Q. Saeed, and T. Abbas, (2024) Classification of Lung Diseases Using Machine Learning Technique, *2024 International Conference on Decision Aid Sciences and Applications (DASA),* pp. 1–7, doi: 10.1109/DASA63652.2024.10836302.

[19] S. Akhtar, S. Aftab, S. Kousar, A. Rehman, M. Ahmad, and A. Q. Saeed, (2024). A Severity Grading Framework for Diabetic Retinopathy Detection using Transfer Learning, *2024 International Conference on Decision Aid Sciences and Applications (DASA),* pp. 1–5, doi: 10.1109/DASA63652.2024.10836441.

[20] T. M. Mitchell, (2024) Machine Learning: Tom M Mitchell: 9781259096952: Amazon.com: Books.. [Online]. Available: https://www.amazon.com/Machine-Learning-Tom-M-Mitchell/dp/1259096955

[21] R. K. Halder, M. N. Uddin, M. A. Uddin, S. Aryal, and A. Khraisat, (2024) Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications, *J Big Data*, 11(1), doi: 10.1186/s40537-024-00973-y.

[22] L. Wang, (2018) Unlock with Your Heart : Heartbeat-based Authentication on Commercial Mobile Phones,.

[23] H. Watanabe and T. Terada, (2018) Improving ultrasound-based gesture recognition using a partially shielded single microphone, *Proceedings - International Symposium on Wearable Computers, ISWC,* pp. 9–16,, doi: 10.1145/3267242.3267274/SUPPL_FILE/P9-WATANABE.MP4.

[24] X. Xu, J. Yu, Y. Chen, Y. Zhu, S. Qian, and M. Li, (2018). Leveraging Audio Signals for Early Recognition of Inattentive Driving with Smartphones, *IEEE Trans Mob Comput,* 17(7);1553–1567, doi: 10.1109/TMC.2017.2772253.

[25] T. H. VU, A. MISRA, Q. ROY, K. T. W. CHOO, and Y. LEE, (2018). Smartwatch-based early gesture detection & trajectory tracking for interactive gesture-driven applications, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 2(1);39: doi: 10.1145/3191771.

[26] L. Lu et al., (2018) LipPass: Lip Reading-based User Authentication on Smartphones Leveraging Acoustic Signals, *Proceedings - IEEE INFOCOM*, 2018;1466–1474, doi: 10.1109/INFOCOM.2018.8486283.

[27] P. Voigt, M. Budde, E. Pescara, M. Fujimoto, K. Yasumoto, and M. Beigl, (2018) Feasibility of human activity recognition using wearable depth cameras, *Proceedings - International Symposium on Wearable Computers, ISWC*, pp. 92–95, doi: 10.1145/3267242.3267276/SUPPL_FILE/P92-VOIGT.JPG.

[28] Y. Lu, B. Huang, C. Yu, G. Liu, and Y. Shi, (2020) Designing and evaluating hand-to-hand gestures with dual commodity wrist-worn devices, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 4(1), doi: 10.1145/3380984/SUPPL_FILE/LU.ZIP.

[29] Y. Wang, W. Cai, T. Gu, and W. Shao, (2020) Your eyes reveal your secrets: An eye movement based password inference on smartphone, *IEEE Trans Mob Comput*, 19(11);2714–2730, doi: 10.1109/TMC.2019.2934690.

[30] D. Shi et al., (2021) Fine-Grained and Context-Aware Behavioral Biometrics for Pattern Lock on Smartphones, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 5(1), doi: 10.1145/3448080.

[31] Y. Cao, H. Chen, F. Li, and Y. Wang, (2021) CanalScan: Tongue-jaw movement recognition via ear canal deformation sensing, *Proceedings - IEEE INFOCOM*, 2021, doi: 10.1109/INFOCOM42981.2021.9488852.

[32] X. Chen, J. Fernandez-Mendoza, Y. Xiao, Y. Tang, and G. Cao, (2021) ApneaDetector: Detecting Sleep Apnea with Smartwatches, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol*, 5, doi: 10.1145/3463514.

[33] M. Gjoreski et al., (2020). Classical and deep learning methods for recognizing human activities and modes of transportation with smartphone sensors, *Information Fusion*, 62;47–62, doi: 10.1016/j.inffus.2020.04.004.

[34] ChangLiqiong et al., (2018). SleepGuard, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 2(3);1–34, doi: 10.1145/3264908.

[35] T. Hamatani, M. Elhamshary, A. Uchiyama, and T. Higashino, (2018) FluidMeter," *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 2(3), doi: 10.1145/3264923.

[36] K. Jiokeng, G. Jakllari, and A. L. Beylot, I Want to Know Your Hand: Authentication on Commodity Mobile Phones Based on Your Hand s Vibrations, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 6(2), doi: 10.1145/3534575.

[37] I. H. Sarker, (2021) Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions, doi: 10.1007/s42979-021-00815-1.

[38] J. Wu, Y. Feng, and P. Sun, (2018)Sensor fusion for recognition of activities of daily living, *Sensors (Switzerland)*, 8(11), doi: 10.3390/s18114029.

[39] M. M. Islam, S. Nooruddin, F. Karray, and G. Muhammad, (2022) Human activity recognition using tools of convolutional neural networks: A state of the art review, data sets, challenges, and future prospects, *Comput Biol Med*, 149;106060, doi: 10.1016/J.COMPBIOMED.2022.106060.

[41] Z. He, Y. Sun, and Z. Zhang, (2024) Human Activity Recognition Based on Deep Learning Regardless of Sensor Orientation, *Applied Sciences* 14(9);3637, doi: 10.3390/APP14093637.

[42] C. Avilés-Cruz, A. Ferreyra-Ramírez, A. Zúñiga-López, and J. Villegas-Cortéz, (2019) Coarse-fine convolutional deep-learning strategy for human activity recognition, *Sensors (Switzerland)*, 19(7), doi: 10.3390/s19071556.

[43] W. Huang, L. Zhang, W. Gao, F. Min, and J. He, (2021) Shallow Convolutional Neural Networks for Human Activity Recognition Using Wearable Sensors, *IEEE Trans Instrum Meas*, 70, doi: 10.1109/TIM.2021.3091990.

[44] J. Jiao, M. Zhao, J. Lin, and K. Liang, (2020) A comprehensive review on convolutional neural network in machine fault diagnosis, *Neurocomputing*, 417;36–63, doi: 10.1016/j.neucom.2020.07.088.

[45] I. K. Ihianle, A. O. Nwajana, S. H. Ebenuwa, R. I. Otuka, K. Owa, and M. O. Orisatoki, (2020) A deep learning approach for human activities recognition from multimodal sensing devices, *IEEE Access*, 8;179028–179038, doi: 10.1109/ACCESS.2020.3027979.

[46] L. A. Jasim et al., (2024) Network Fortification: Leveraging Support Vector Machine for Enhanced Security in Wireless Body Area Networks, *International Journal of Safety and Security Engineering*, 14(3);923–931, doi: 10.18280/ijsse.140323.

[47] S. Akhtar, S. Aftab, S. Kousar, A. Rehman, M. Ahmad, and A. Q. Saeed, (2024) Performance Analysis of Training Optimizers on Diabetic Retinopathy Detection Using SP-Net CNN, *2024 International Conference on Decision Aid Sciences and Applications (DASA)*, pp. 1–5, doi: 10.1109/DASA63652.2024.10836543.

[48] A. Q. Saeed, S. N. H. S. Abdullah, J. Che-Hamzah, and A. T. A. Ghani, (2024) FINE VESSEL SEGMENTATION WITH REFINEMENT GATE IN DEEP LEARNING ARCHITECTURES, *Malaysian Journal of Computer Science*, 37(3);205–224, doi: 10.22452/MJCS.VOL37NO3.2.

[49] S. Mekruksavanich and A. Jitpattanakul, (2021) Deep convolutional neural network with rnns for complex activity recognition using wrist-worn wearable sensor data, *Electronics (Switzerland)*, 10(14), doi: 10.3390/electronics10141685.

[50] S. Gupta, (2021). Deep learning based human activity recognition (HAR) using wearable sensor data, *International Journal of Information Management Data Insights*, 1(2), doi: 10.1016/j.jjimei.2021.100046.

[51] C. X. Lu et al., (2018) DeepAuth: In-situ authentication for smartwatches via deeply learned behavioural biometrics, *Proceedings - International Symposium on Wearable Computers, ISWC*, pp. 204–207, doi: 10.1145/3267242.3267252.
[52] C. X. Lu et al., (2019) Snoopy: Sniffing Your Smartwatch Passwords via Deep Sequence Learning," *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 1(4);1–29, doi: 10.1145/3161196.
[53] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren, (2018) A robust human activity recognition system using smartphone sensors and deep learning, *Future Generation Computer Systems*, 81;307–313, doi: 10.1016/J.FUTURE.2017.11.029.
[54] H. Du, P. Li, H. Zhou, W. Gong, G. Luo, and P. Yang, (2018) WordRecorder: Accurate Acoustic-based Handwriting Recognition Using Deep Learning, *Proceedings - IEEE INFOCOM*, 2018;1448–1456, doi: 10.1109/INFOCOM.2018.8486285.
[55] T. Gong, Y. Kim, J. Shin, and S.-J. Lee, (2019) MetaSense: Few-Shot Adaptation to Untrained Conditions in Deep Mobile Sensing,, doi: 10.1145/3356250.3360020.
[56] V. Becker, L. Fessler, and G. Sörös, GestEar: Combining Audio and Motion Sensing for Gesture Recognition on Smartwatches, doi: 10.1145/3341163.3347735.
[57] G. Brunner, D. Melnyk, B. Sigfússon, and R. Wattenhofer, (2019) Swimming style recognition and lap counting using a smartwatch and deep learning, *Proceedings - International Symposium on Wearable Computers, ISWC*, pp. 23–31, doi: 10.1145/3341163.3347719.
[58] Y. Liu, Z. Li, Z. Liu, and K. Wu, (2019). Real-time arm skeleton tracking and gesture inference tolerant to missing wearable sensors, *MobiSys 2019 - Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 287–299, doi: 10.1145/3307334.3326109.
[59] J. Hou et al., SignSpeaker: A Real-time, High-Precision SmartWatch-based Sign Language Translator, 19, doi: 10.1145/3300061.3300117.
[60] T. Giallanza et al., (2019) Keyboard Snooping from Mobile Phone Arrays with Mixed Convolutional and Recurrent Neural Networks, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 3(2);1–22, doi: 10.1145/3328916.
[61] Y. Liu, K. Huang, X. Song, B. Yang, and W. Gao, (2020). MagHacker: Eavesdropping on Stylus Pen Writing via Magnetic Sensing from Commodity Mobile Devices, 13, doi: 10.1145/3386901.3389030.
[62] H. Yin, A. Zhou, G. Su, B. Chen, L. Liu, and H. Ma, (2020). Learning to Recognize Handwriting Input with Acoustic Features, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 4(2); doi: 10.1145/3397334.
[63] Y. Gao, Y. Jin, J. Li, S. Choi, and Z. Jin, (2020) EchoWhisper: Exploring an Acoustic-based Silent Speech Interface for Smartphone Users, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 4(3), doi: 10.1145/3411830.
[64] S. Liu et al., (2020) GlobalFusion: A global attentional deep learning framework for multisensor information fusion, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 4(1), doi: 10.1145/3380999.
[65] Y. Wang, J. Shen, and Y. Zheng, "Push the Limit of Acoustic Gesture Recognition," *Proceedings - IEEE INFOCOM*, 2020;566–575, doi: 10.1109/INFOCOM41043.2020.9155402.
[66] H. Park, S. Korea, Y. Lee, and J. Ko, (2021) Enabling Real-time Sign Language Translation on Mobile Platforms with On-board Depth Cameras,, doi: 10.1145/3463498.
[67] B. Zhai, Y. Guan, M. Catt, and T. Ploetz, (2021) Ubi-SleepNet: Advanced Multimodal Fusion Techniques for Three-stage Sleep Classification Using Ubiquitous Sensing, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 5(4), doi: 10.1145/3494961.
[68] W. Chen, L. Chen, M. Ma, F. S. Parizi, S. Patel, and J. Stankovic, (2021) ViFin: Harness Passive Vibration to Continuous Micro Finger Writing with a Commodity Smartwatch, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 5(1), doi: 10.1145/3448119.
[69] X. Ouyang, Z. Xie, J. Zhou, J. Huang, and G. Xing, ClusterFL: A Similarity-Aware Federated Learning System for Human Activity Recognition Human activity recognition, Federated learning, Clustering, Multi-task learning, *Communication optimization ACM Reference Format*, p. 13, doi: 10.1145/3458864.3467681.
[70] B. Khaertdinov, E. Ghaleb, and S. Asteriadis, (2021) Contrastive self-supervised learning for sensor-based human activity recognition, *2021 IEEE International Joint Conference on Biometrics, IJCB*, doi: 10.1109/IJCB52358.2021.9484410.
[71] Q. Zhang, D. Wang, R. Zhao, and Y. Yu, (2021) SoundLip: Enabling Word and Sentence-level Lip Interaction for Smart Devices, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 5(1), doi: 10.1145/3448087.
[72] Q. Zhang, D. Wang, R. Zhao, Y. Yu, and J. Z. Jing, (2021) Write, Attend and Spell: Streaming End-to-end Free-style Handwriting Recognition Using Smartwatches, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 5(3), doi: 10.1145/3478100.
[73] W. Lu, J. Wang, Y. Chen, S. J. Pan, C. Hu, and X. Qin, (2022). Semantic-Discriminative Mixup for Generalizable Sensor-based Cross-domain Activity Recognition, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 6(2); doi: 10.1145/3534589.
[74] X. Xu et al., (2022). Enabling hand gesture customization on wrist-worn devices, *Conference on Human Factors in Computing Systems - Proceedings*, p. 19, doi: 10.1145/3491102.3501904.
[75] Y. Xie, F. Li, Y. Wu, S. Yang, and Y. Wang, (2022) HearSmoking: Smoking Detection in Driving Environment via Acoustic Sensing on Smartphones,

IEEE Trans Mob Comput, 21(8);2847–2860, doi: 10.1109/TMC.2020.3048785.

[76] K. Ling, H. Dai, Y. Liu, and A. X. Liu, (2018) Ultragesture: Fine-grained gesture sensing and recognition, 2018 *15th Annual IEEE International Conference on Sensing, Communication, and Networking, SECON*, pp. 1–9, doi: 10.1109/SAHCN.2018.8397099.

[77] MollynVimal, AhujaKaran, VermaDhruv, HarrisonChris, and GoelMayank, (2022) SAMoSA, *Proc ACM Interact Mob Wearable Ubiquitous Technol,* 6(3);331–332, doi: 10.1145/3550284.

[78] P. Wang, R. Jiang, and C. Liu, (2022) Amaging: Acoustic Hand Imaging for Self-adaptive Gesture Recognition, *Proceedings - IEEE INFOCOM,* 2022; 80–89,, doi: 10.1109/INFOCOM48880.2022.9796906.

[79] K. Tanigaki, T. C. Teoh, N. Yoshimura, T. Maekawa, and T. Hara, (2022) Predicting Performance Improvement of Human Activity Recognition Model by Additional Data Collection, *Proc ACM Interact Mob Wearable Ubiquitous Technol,* 6(3);, doi: 10.1145/3550319.

[80] A. Raza et al., (2021). Lightweight Transformer in Federated Setting for Human Activity Recognition, [Online]. Available: https://arxiv.org/abs/2110.00244v3

[81] S. EK, F. Portet, and P. Lalanda, (2022). Lightweight Transformers for Human Activity Recognition on Mobile Devices, *Pers Ubiquitous Comput,* 27(6);2267–2280, doi: 10.1007/s00779-023-01776-3.

[82] S. Miao, L. Chen, R. Hu, and Y. Luo, (2022) Towards a Dynamic Inter-Sensor Correlations Learning Framework for Multi-Sensor-Based Wearable Human Activity Recognition, *Proc ACM Interact Mob Wearable Ubiquitous Technol,* 6(3), doi: 10.1145/3550331.

[83] X. You, L. Zhang, H. Yu, M. Yuan, and X. Y. Li, (2021) KATN: Key activity detection via inexact supervised learning," *Proc ACM Interact Mob Wearable Ubiquitous Technol,* 5(4), doi: 10.1145/3494957.

[84] N. Yoshimura, T. Maekawa, T. Hara, and A. Wada, (2022) Acceleration-based Activity Recognition of Repetitive Works with Lightweight Ordered-work Segmentation Network, *Proc ACM Interact Mob Wearable Ubiquitous Technol,* 6(2), doi: 10.1145/3534572.

[85] H. Li et al., "NeuralGait: Assessing Brain Health Using Your Smartphone," Proc ACM Interact Mob Wearable Ubiquitous Technol, vol. 6, no. 4, Jan. 2023, doi: 10.1145/3569476.

[86] G. Augustinov et al., (2022). Transformer-Based Recognition of Activities of Daily Living from Wearable *Sensor Data,* pp. 1–8, doi: 10.1145/3558884.3558895.

[87] G. Sharma, A. Dhall, and R. Subramanian, (2022). A Transformer Based Approach for Activity Detection, *MM 2022 - Proceedings of the 30th ACM International Conference on Multimedia*, pp. 7155–7159, doi: 10.1145/3503161.3551598.

[88] Y. Zhang, L. Wang, H. Chen, A. Tian, S. Zhou, and Y. Guo, (2022). IF-ConvTransformer: A Framework for Human Activity Recognition Using IMU Fusion and ConvTransformer, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 6(2); doi: 10.1145/3534584.

[89] S. Lee, J. Lee, and K. Lee, (2023). DeepVehicleSense: An Energy-Efficient Transportation Mode Recognition Leveraging Staged Deep Learning Over Sound Samples, *IEEE Trans Mob Comput*, 22(6);3270–3286, doi: 10.1109/TMC.2022.3141392.

[90] D. Ding, L. Yang, Y. C. Chen, and G. Xue, (2023). Handwriting Recognition System Leveraging Vibration Signal on Smartphones, *IEEE Trans Mob Comput*, 22(7);3940–3951, doi: 10.1109/TMC.2022.3148172.

[91] R. Mishra, A. Gupta, and H. P. Gupta, (2023) Locomotion Mode Recognition Using Sensory Data With Noisy Labels: A Deep Learning Approach, *IEEE Trans Mob Comput*, 22(6);3460–3471, doi: 10.1109/TMC.2021.3135878.

[92] Z. Xiao, T. Chen, Y. Liu, J. Li, and Z. Li, (2023) Keystroke Recognition With the Tapping Sound Recorded by Mobile Phone Microphones, *IEEE Trans Mob Comput*, 22(6);3407–3424, doi: 10.1109/TMC.2021.3137229.

[93] C. Avilés-Cruz, A. Ferreyra-Ramírez, A. Zúñiga-López, and J. Villegas-Cortéz, (2019). Coarse-Fine Convolutional Deep-Learning Strategy for Human Activity Recognition, *Sensors* 19(7);1556, doi: 10.3390/S19071556.

[94] A. Sarkar, S. K. S. Hossain, and R. Sarkar, (2023). Human activity recognition from sensor data using spatial attention-aided CNN with genetic algorithm, *Neural Comput Appl*, 35(7);5165–5191, doi: 10.1007/s00521-022-07911-0.

[95] J. I. Janjua, S. Kousar, A. Khan, A. Ihsan, T. Abbas, and A. Q. Saeed, (2024). Enhancing Scalability in Reinforcement Learning for Open Spaces, *2024 International Conference on Decision Aid Sciences and Applications (DASA)*, pp. 1–8, doi: 10.1109/DASA63652.2024.10836237.

[96] O. Saidani, M. Alsafyani, R. Alroobaea, N. Alturki, R. Jahangir, and L. Jamel, (2023). An Efficient Human Activity Recognition Using Hybrid Features and Transformer Model, *IEEE Access*, 11;101373–101386, doi: 10.1109/ACCESS.2023.3314492.

[97] M. K, A. Ramesh, R. G, S. Prem, R. A A, and D. M. P. Gopinath, (2021). 1D Convolution approach to human activity recognition using sensor data and comparison with machine learning algorithms, *International Journal of Cognitive Computing in Engineering*, 2;130–143, doi: 10.1016/j.ijcce.2021.09.001.

[98] K. Xia, J. Huang, and H. Wang, (2020). LSTM-CNN Architecture for Human Activity Recognition, *IEEE Access*, 8;56855–56866, doi: 10.1109/ACCESS.2020.2982225.

[99] E. H. Bahadur, A. Kadar Muhammad Masum, A. Barua, M. G. Rabiul Alam, M. A. U. Zaman Chowdhury, and M. R. Alam, (2019). LSTM based approach for diabetic symptomatic activity

recognition using smartphone sensors, *22nd International Conference on Computer and Information Technology, ICCIT 2019, Institute of Electrical and Electronics Engineers Inc*.. doi: 10.1109/ICCIT48885.2019.9038185.

[100]M. Mahfuri, S. Ghwanmeh, A. Q. Saeed, and M. Khishe, (2024). Artificial Intelligence in Account Management: Innovation, Challenges, and Strategic Outlook, *2024 International Conference on Decision Aid Sciences and Applications (DASA)*, pp. 1–7, doi: 10.1109/DASA63652.2024.10836501.

[101]D. Mukherjee, R. Mondal, P. K. Singh, R. Sarkar, and D. Bhattacharjee, (2020). EnsemConvNet: a deep learning approach for human activity recognition using smartphone sensors for healthcare applications, *Multimed Tools Appl*, 79(41–42);31663–31690, doi: 10.1007/s11042-020-09537-7.

[102]C. Shen, B. J. Ho, and M. Srivastava, (2018). MiLift: Efficient Smartwatch-Based Workout Tracking Using Automatic Segmentation, *IEEE Trans Mob Comput*, 17(7);1609–1622, doi: 10.1109/TMC.2017.2775641.

[103]K. Ahuja, R. Islam, V. Parashar, K. Dey, C. Harrison, and M. Goel, (2018). EyeSpyVR, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 2(2);1–10, doi: 10.1145/3214260.

[104]Z. Yang, C. Yu, F. Zheng, and Y. Shi, (2019). *ProxiTalk*, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 3(3);1–25, doi: 10.1145/3351276.

[105]C. Luo et al., (2019). Brush like a Dentist: Accurate Monitoring of Toothbrushing via Wrist-Worn Gesture Sensing, *Proceedings - IEEE INFOCOM*, 2019;1234–1242, doi: 10.1109/INFOCOM.2019.8737513.

[106]Y. Yin, L. Xie, T. Gu, Y. Lu, and S. Lu, (2019) AirContour: Building contour-based model for in-Air writing gesture recognition," *ACM Trans Sens Netw*, 15(4); doi: 10.1145/3343855.

[107]X. Zhang, Y. Yin, L. Xie, H. Zhang, Z. Ge, and S. Lu, (2020) Touchid: User authentication on mobile devices via inertial-Touch gesture analysis, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 4(4), doi: 10.1145/3432192.

[108]Y. Song and Z. Cai, (2022) Integrating Handcrafted Features with Deep Representations for Smartphone Authentication, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 6(1), doi: 10.1145/3517332.

[109]S. Bhattacharya, R. Adaimi, and E. Thomaz, (2022) Leveraging Sound and Wrist Motion to Detect Activities of Daily Living with Commodity Smartwatches, *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 6(2), doi: 10.1145/3534582.

[110]S. K. Challa, A. Kumar, and V. B. Semwal, (2022) A multibranch CNN-BiLSTM model for human activity recognition using wearable sensor data," *Visual Computer*, 38(12);4095–4109, doi: 10.1007/s00371-021-02283-3.

[111]S. Schober, E. Schimbäck, K. Pendl, K. Pichler, V. Sturm, and F. Runte, (2024). Human activity recognition system using wearable accelerometers for classification of leg movements: a first, detailed approach, *Journal of Sensors and Sensor Systems*, 13(2);187–209, doi: 10.5194/jsss-13-187-2024.

[112]A. G. Prabono, B. N. Yahya, and S. L. Lee, (2021) Atypical Sample Regularizer Autoencoder for Cross-Domain Human Activity Recognition, *Information Systems Frontiers*, 23(1);71–80, doi: 10.1007/s10796-020-09992-5.

[113]Y. Li, G. Yang, Z. Su, S. Li, and Y. Wang, (2023) Human activity recognition based on multienvironment sensor data, *Information Fusion*, 91;47–63, doi: 10.1016/j.inffus.2022.10.015.

[114]K. D. Garcia et al., (2021) An ensemble of autonomous auto-encoders for human activity recognition, *Neurocomputing*, 439;271–280, doi: 10.1016/j.neucom.2020.01.125.

[116]E. Essa and I. R. Abdelmaksoud, (2023) Temporal-channel convolution with self-attention network for human activity recognition using wearable sensors, *Knowl Based Syst*, 278;110867, doi: 10.1016/J.KNOSYS.2023.110867.

[117]M. M. Jankowska, J. Schipperijn, J. Kerr, M. M. Jankowska, J. Schipperijn, and J. Kerr, (2015). A Framework for Using GPS Data in Physical Activity and Sedentary Behavior Studies. [Online]. Available: www.acsm-essr.org

[118]N. A. Al-Dmour, S. Kousar, A. Khan, A. Ihsan, T. Abbas, and A. Q. Saeed, (2024). Enhancing Email Spam Detection Using Advanced AI Techniques, *2024 International Conference on Decision Aid Sciences and Applications (DASA)*, pp. 1–6, doi: 10.1109/DASA63652.2024.10836525.

[119]T. Shen, I. Di Giulio, and M. Howard, (2023) A Probabilistic Model of Human Activity Recognition with Loose Clothing *Sensors*, 23(10), doi: 10.3390/s23104669.

[120]J. Park et al., (2023). Development of a Gait Feature-Based Model for Classifying Cognitive Disorders Using a Single Wearable Inertial Sensor, *Neurology*, 101(1);E12–E19, doi: 10.1212/WNL.0000000000207372.

[122]H. A. Dehkordi, A. S. Nezhad, H. Kashiani, S. B. Shokouhi, and A. Ayatollahi, (2022) Multi-expert human action recognition with hierarchical super-class learning, *Knowl Based Syst*, 250, doi: 10.1016/j.knosys.2022.109091.

[123]Y. Wang and F. Feng, (2021). Reliability enhancement algorithm of human motion recognition based on knowledge graph, *International Journal of Distributed Systems and Technologies*, 12(1);1–15, doi: 10.4018/IJDST.2021010101.

[124]L. Yao et al., (2018). WITS: an IoT-endowed computational framework for activity recognition in personalized smart homes, *Computing*, 100(4);369–385, doi: 10.1007/s00607-018-0603-z.

[125]A. Zahin, L. T. Tan, and R. Q. Hu, (2019) Sensor-based human activity recognition for smart healthcare: A semi-supervised machine learning, *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, 287;450–472, doi: 10.1007/978-3-030-22971-9_39.

[126]L. Bibbò and M. M. B. R. Vellasco, (2023). Human Activity Recognition (HAR) in Healthcare, *Multidisciplinary Digital Publishing Institute* doi: 10.3390/app132413009.

[127]G. Ogbuabor and R. La, (2018). Human activity recognition for healthcare using smartphones, *ACM International Conference Proceeding Series*, pp. 41–46, doi: 10.1145/3195106.3195157.

[128]L. Chen et al., (2020). Survey of Pedestrian Action Recognition Techniques for Autonomous Driving.

[129]L. C. Günther, S. Kärcher, and T. Bauernhansl, (2019). Activity recognition in manual manufacturing: Detecting screwing processes from sensor data, *Procedia CIRP*, 1177–1182. doi: 10.1016/j.procir.2019.03.288.