Automation in Construction 142 (2022) 104497

Contents lists available at ScienceDirect

Automation in Construction

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

Review

Human motion prediction for intelligent construction: A review

Xiaolu Xia [a], Tianyu Zhou [b], Jing Du [b], Nan Li [a,c,*]

[a] Department of Construction Management, Tsinghua University, China
[b] Department of Civil and Costal Engineering, University of Florida, United States of America
[c] Hang Lung Center for Real Estate, Tsinghua University, China

ARTICLE INFO

Keywords:
Human motion prediction
Human motion and intent
Kinematics
Ergonomics
Cognition
Review

ABSTRACT

Intelligent construction is an important construction trend. With the growing number of intelligent autonomous systems implemented in the construction area, understanding and predicting human motion becomes increasingly important. Based on such predictions, the autonomous systems can optimize their actions to improve the efficiency of human-robot interactions, and supervisors can make informed decisions about when and where to intervene in human motion to avoid collisions. This paper presents a comprehensive review of existing literature on human motion prediction (HMP). Relevant studies from a wide range of fields are reviewed, analyzed and synthesized, in terms of prediction indicators, methods and applications, based on a three-level taxonomy. The taxonomy is structured based on the levels of human information required by different prediction methods, and reflects different understandings of the underlying causality and mediators of human motions and intent. The paper also discusses the evolutions of the theoretical understanding and methodological development of HMP, its application scenarios in and beyond the construction domain, and possible directions for future research. This review is expected to increase the visibility of this rapidly expanding research area, and inspire future studies and advancements for human-robot interactions in construction.

1. Introduction

Human motion prediction (HMP) refers to predicting how a person moves or acts in the near future by conditioning on a series of historical movements or certain leading indicators [1]. Effective HMP is the basis and the enabler of scientific advancements in a variety of disciplines, including ergonomics, medical science, human factors engineering, transportation engineering, and especially human-robot interactions in the construction industry [2]. In any human-engineering coupled system, understanding and predicting human motion are the building blocks for a safer and better workplace [3]. In the construction literature, there is a growing interest in adopting HMP for various intelligent system designs and applications [4]. For example, the in-situ worker motion predictions can provide a warning system for reducing falling risks based on the relationship between the environment and the worker motion features [5]. Early injury prevention for critical events can also be realized via human motion simulations and predictions [6]. The human motion trajectory prediction also helps solve the problems of coexistence and interaction of intelligent systems and human agents in collaborative construction tasks, such as the human-robot collaborative operations [7].

However, due to the complexity of human behaviors, HMP is still facing critical challenges. The most prominent source of challenges in HMP roots in the diversity of human motion modalities and contexts [3]. In a broader context, human motions can refer to articulated full-body motions, gestures and facial expressions, and spatial navigations with mobility devices or vehicles [3]. Without losing its generality, this paper focuses on the motions and navigations of an individual in indoor or outdoor environments, as well as the sophisticated body motions in workplaces. Understanding the status quo, challenges, and trends of these two HMP themes will contribute to a deeper understanding of how the industrial systems run with humans in the loop, in built environments, and in the new context of human-robot collaboration and automation [8].

In the past two decades, HMP methods have been substantially enriched and enhanced. As for the sources of data, literature has explored motion captures, imagery data, physiological signals, and neurofunctional data [9–11]. The modeling methods also spanned from the straightforward physics-based approaches to the most recent deep learning models [12–14]. Along with the methodological innovations,

* Corresponding author.
E-mail address: nanli@tsinghua.edu.cn (N. Li).

https://doi.org/10.1016/j.autcon.2022.104497
Received 18 October 2021; Received in revised form 17 July 2022; Accepted 23 July 2022
Available online 4 August 2022

our basic understanding of the causality and mediators of human motions and intent have greatly advanced. More evidence has been driven by discoveries on the role of human-environment relationships in shaping motion intents [3]. Human ergonomic features have been quantified to expand prediction robustness [3,15,16]. Certain cognitive processes are measured to forecast motion intents long before actions can be observed [17,18]. Most recently, psychological frameworks are being tested to explain the decision-intent-motion-feedback loop from a higher level [19]. With all said, the HMP community is growing exponentially and benefiting science and engineering discoveries of more disciplines. Previous review papers related to this topic [15,17,18,20] focused only on specific and implicit aspects of HMP predictors (such as kinematic features of human motions), instead of providing a comprehensive discussion on the variety of multimodal data, which is herein referred to as human information. Human information, by providing a deeper insight into the human physiological functions, metabolism, psychological status and cognitive state of a person, can contribute to more effective HMP. The most recent advancements in the HMP literature, including the paradigm evolution, algorithmic changes and new assumptions, that pertain to the deeper human information need a thorough examination.

This paper aims to provide a comprehensive review of the HMP literature, including the roots of different paradigms, the state of the art, and future trends, particularly in the context of navigation and body motions in workplaces. This will help construction scholars gain a high-level insight into the relevance of different HMP methods and assumptions in their research areas, such as productivity, safety, emergency management, and automation. Our review finds that HMP methods and models present a multilevel, multidimensional knowledge structure. The existing HMP literature is based on various domains and has attracted scholars from a wide spectrum of disciplines such as self-driving vehicles [21–24], service robots [25–29], advanced surveillance systems [30,31], computer vision [32–35], and so on. The human information that has been used in HMP covers multi-dimensional information of a person, including not only the direct visual or imagery information but also the kinematics and biomechanics of human agents, as well as the complex physiological and biochemical information that can be detected by various sensors and auxiliary systems [36]. Based on a comprehensive literature review, we proposed to categorize HMP methods into three levels depending on the classes of human information used in the prediction. The first level of human information includes the particle-based features, in which a person is modeled as a point and the algorithms used to predict human-space interactions or human-human interactions are based on physics laws. The methods used are based on physics laws, physics-based constraints and/or physics-inspired objective functions. The second level is body information, in which the features include the biomechanics and complex physiological and biochemical information of a person. The highest level is the decision context information, in which how humans make motion decisions are considered.

Based on the above categorization framework, this paper reviews the state of the art, discusses typical properties, advantages, and drawbacks of each type of HMP method, and outlines the challenges for future research. Specifically, we aim to answer the following four questions in this review: 1) What are the methods and data for HMP based on human information? 2) What are the differences and similarities in the methods used to predict human motions in different domains? 3) What are the pros and cons of each method used in HMP using different levels of human information? and 4) What are the trends and limitations in the HMP literature? The contribution of this review is twofold: 1) it reviews the indicators, methods, and cutting-edge research in various fields of HMP, and presents a comprehensive and comparative overview of the latest accomplishments; and 2) it analyzes the evolution of HMP paradigms based on different levels of human information, reveals the rationales behind them, and puts forward recommendations for future research especially in construction.

2. Methodology

A wide range of materials of various types, including journal articles and conference proceedings were reviewed to synthesize the latest accomplishments in the area of HMP. Web of Science Core Collection was used to search for these materials to include high-impact publications. To find the published works related to HMP, different combinations of keywords were used. TS, which is the "topic" field in Web of Science databases, includes papers' keywords, title, and abstract. Considering that the predicting indicators, the applicable scenarios, used methods and the associated keywords of the HMP are different, each level of information that HMP is based on was searched separately. For particle-based prediction, "TS = (human OR pedestrian OR people OR participant OR crowds) AND trajector* AND (predict* OR estimat* OR understand* OR forecast*)" was searched. For body information-based prediction, "TS = (((human OR pedestrian) NEAR/1 (intent* OR action OR trajector* OR motion OR path OR movement) NEAR/1 predict*) AND (body OR kinematic OR posture OR gait OR head OR eye OR gaze OR hand OR trunk OR hip OR knee OR ankle OR EEG OR EMG))" was searched. For decision context knowledge-based prediction, "TS= (((human OR pedestrian) NEAR/1 (intent* OR action OR trajector* OR motion OR path OR movement) NEAR/1 predict*) AND ("human decision-making" OR "human decision making" OR "psychology"))" was searched. Moreover, based on the above search results, the backward and forward snowballing strategy [37] was used to search for additional relevant publications not indexed in Web of Science Core Collection. In addition, to filter irrelevant publications in the search results, the title and abstract, as well as the full text when necessary, of every publication in the search results were manually screened. To be considered relevant, a publication must provide information or evidence to meet the aforementioned research scope of using human information to predict human motion intention in an indoor or outdoor environment or workplace. Publications that did not meet the above criteria were considered irrelevant and hence excluded from the review. The systematic protocol for paper collection and screening is shown in the Fig. 1.

The remainder of this literature review is organized as follows. Section 3 reviews particle-based HMP methods and applications. Section 4 reviews methods and applications of HMP based on body information and decision-context knowledge-based information including ergonomics, physiological, and cognitive information. Section 5 discusses the findings, and Section 6 concludes the paper.

3. Single-Point Kinematics-based Prediction (SPK)

The earliest, and yet still the most prevailing paradigm in HMP is the single-point kinematics-based modeling approach or SPK, which predicts human motion using particle-based human information, i.e., human is modeled as a particle and the human information is based on physics laws. SPK models a human as a single target point of interest, without looking into the internal ergonomic mechanisms about how a motion begins or proceeds. Specifically, this type of approach focuses on trending out the trajectories of the human body (usually modeled as a single point or a "particle" in physics modeling) based on the historical motion trajectory and momentum data. With that said, SPK heavily relies on statistical modeling and recent machine learning (ML) methods for spatiotemporal analysis and prediction. Moving from the early assumption of particle models [38], recent advances in this area also start to predict human motion based on human-space interactions, human-human interactions, and so on. Attributed to the richness of spatiotemporal modeling approaches and recent ML advancements, a large variety of SPK models have been developed in the target tracking and automatic control communities to predict the trajectory of humans to facilitate autonomous driving, robotics, and abnormal crowded behavior detection.

Although SPK represents one of the oldest HMP paradigms, it remains a challenging task because the precise prediction requires strong

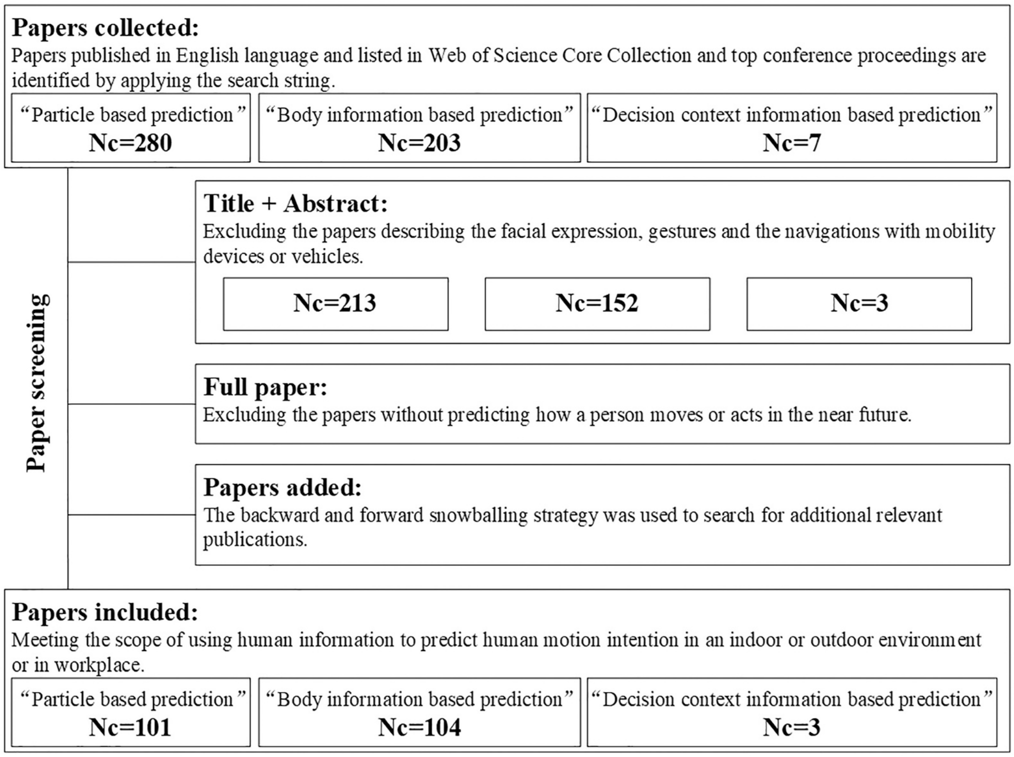

Fig. 1. The systematic protocol for paper collection and screening.

reasoning about human agents' past movements, social interactions among varying numbers and kinds of agents, constraints from scene contexts, and the stochasticity of human behavior [39]. Scholars have often borrowed methods, such as clustering [40] and the social force modeling [41], from the ML and behavioral literature. Recent research has become more successful by leveraging the emerging deep learning methods including the convolutional neural networks (CNN) [42], recurrent neural network (RNN) [43], long short-term memory (LSTM) [38], gated recurrent units (GRU) [44], and generative adversarial network (GAN) [45]. These new deep learning methods seem to be promising in addressing the data scarcity issues usually faced by SPK. We use the complexity and comprehensiveness of the specific spatio-temporal modeling approaches to review and categorize the first level of human information. We found that the following categorization also lines up with the evolving timeline of this research area: (1) trajectory categorization, (2) social force methods, (3) time-series deep learning, and (4) generative-based methods.

In addition, two popular human-trajectory datasets are often used in the literature to evaluate the HMP methods, namely ETH [34] and UCY [46]. The ETH dataset contains two scenes, including ETH and HOTEL. The UCY dataset contains three scenes, including ZARA1, ZARA2, and UCY. The following two metrics are commonly used to quantify the accuracy of prediction results: average displacement error (ADE) [47], which measures the mean square error (MSE) over all estimated points of a trajectory and the true points, and final displacement error (FDE) [48], which measures the distance between the predicted final destination and the true final destination at end of the prediction period. These two metrics can be calculated as follows:

$$ADE = \frac{\sum_{i=1}^{N}\sum_{t=0}^{T}\left\|\widetilde{Y}_t^i - Y_t^i\right\|_2}{N^*T_{pred}} \tag{1}$$

$$FDE = \frac{\sum_{i=1}^{N}\left\|\widetilde{Y}_t^i - Y_t^i\right\|_2}{N} \tag{2}$$

where $\widetilde{Y}_t^i$ and Y_t^i are the predicted and the true locations of target i at time t, respectively, N is the number of targets, and T is the predicted trajectory length.

The performance of main HMP methods reviewed in this section, measured by the ADE and FDE metrics, are summarized in Table 1.

Table 1
Quantitative results of different methods across ETH and UCY datasets.

Dataset method	ETH	HOTEL	UCY	ZARA1	ZARA2	AVG
Linear [47]	1.33/ 2.94	0.39/ 0.72	0.82/ 1.59	0.62/ 1.21	0.77/ 1.48	0.79/ 1.59
LSTM [91]	1.09/ 2.41	0.86/ 1.91	0.61/ 1.31	0.41/ 0.88	0.52/ 1.11	0.70/ 1.52
Social-LSTM [47]	1.09/ 2.35	0.79/ 1.76	0.67/ 1.40	0.47/ 1.00	0.56/ 1.17	0.72/ 1.54
Social-GAN-P [91]	0.87/ 1.62	0.67/ 1.37	0.76/ 1.52	0.35/ 0.68	0.42/ 0.84	0.61/ 1.21
SoPhie [109]	0.70/ 1.43	0.76/ 1.67	0.54/ 1.24	0.30/ 0.63	0.38/ 0.78	0.54/ 1.15
SR-LSTM-2 [89]	0.63/ 1.25	0.37/ 0.74	0.51/ 1.10	0.41/ 0.90	0.32/ 0.70	0.45/ 0.94
CGNS [107]	0.62/ 1.40	0.70/ 0.93	0.48/ 1.22	0.32/ 0.59	0.35/ 0.71	0.49/ 0.97
PIF [88]	0.73/ 1.65	0.30/ 0.59	0.60/ 1.27	0.38/ 0.81	0.31/ 0.68	0.46/ 1.00
STSGN [116]	0.75/ 1.63	0.63/ 1.01	0.48/ 1.08	0.30/ 0.65	0.26/ 0.57	0.48/ 0.99
GAT [114]	0.68/ 1.27	0.68/ 1.40	0.57/ 1.27	0.29/ 0.60	0.37/ 0.75	0.52/ 1.07
Social-BiGAT [114]	0.69/ 1.27	0.49/ 1.01	0.55/ 1.32	0.30/ 0.62	0.36/ 0.75	0.48/ 1.00
Social-STGCNN [115]	0.64/ 1.11	0.49/ 0.85	0.44/ 0.79	0.34/ 0.53	0.30/ 0.48	0.44/ 0.75

Note: the results are based on the task of predicting 12 future timesteps, given the 8 previous ones. Error metrics reported are ADE/FDE in meters.

3.1. Trajectory categorization

Unsupervised ML was tested as an early effort because of its easiness of adoption and the relatively lighter requirements on data. A representative method is the clustering analysis. Clustering is to group data into clusters, making the data in one group more similar than those of others [49]. Clustering can be used in HMP because it can help predict a particular human motion based on the patterns learned from a similar dataset, identify similarity among trajectory datasets and thus reduce the dimensionality of prediction, and/or help distinguish undesired behaviors (such as outliers) [50]. Literature shows that human intentions and trajectories can be predicted based on the trajectory clustering method [51].

The previous review summarized the clustering methods into six categories and introduced them separately [50]: spatial based clustering [52], time depend clustering [53], partition and group-based clustering [54], uncertain trajectory clustering [55], semantic trajectory clustering [56], and road network-based moving object clustering [57]. Spatial information and time information are the basic features of a moving object's trajectory. Clustering from the spatial dimension is intuitive for discovering the activities of moving objects. The time information is crucial for analyzing moving objects' spatial information which is changing over time. For partition and group-based clustering, subpattern of trajectories not only can be partitioned with low input, output, and time costs, but also can obtain as many features as the original trajectories. The uncertain trajectory clustering means that objects move continuously while their locations can only be updated at discrete times, leaving the location of a moving object between two updates uncertain [58]. The semantic trajectory clustering integrates background geographic information and moving objects' characteristics into trajectories. Existing approaches for trajectory data mining and knowledge discovery have focused on the geometrical properties of trajectories, without considering the background geographic information. For many application domains, useful information may only be extracted from trajectory data if their semantics and the background geographic information are considered [59]. In road network-based moving object clustering, the moving objects' trajectory data can be divided into two classes, including road network constraint data [60] and unconstraint data in the free space [61]. However, most of the methods and literature do not consider internal ergonomic mechanisms and regard all moving objects as points. The spatial location, time stamp, and environment around the moving object are the most important information of clustering methods.

Compared with other human trajectory prediction methods, clustering-based methods are more robust and can be used for other moving objects (such as vehicles [62] and ships [63]). It is also an efficient way to analyze and find the massive, hidden, unknown, and interesting knowledge in scale datasets [50]. However, the neglected human information can make the prediction of human trajectory inaccurate. Some trajectories are physically possible but socially unacceptable. Pedestrians are governed by social norms like yielding right-of-way or respecting personal space. Therefore, trajectory prediction that ignores human factors may not be applicable in many realistic daily-life scenarios.

3.2. Social force models

Clustering is efficient when human motion data is clear and follows simple rules, analogous to the rules seen in classic kinetics. However, literature soon began to recognize that human motions in real-world settings are far more complex. As a result, social force-based methods have been proposed. Unlike clustering-based methods which treat humans as independent moving objects, the social force-based methods model human interactions as well. Helbing and Molnar [41] simulated pedestrian behavior with a combination of attractive forces guiding the pedestrians toward their goal and repulsive forces encouraging collision avoidance as the social force model. A large number of features reported in the literature have been "handcrafted" with an eye for overcoming specific issues like occlusions and variations in scale and illumination. The design of handcrafted features often involves finding the right trade-off between accuracy and computational efficiency [64]. Social force methods for predicting or classifying trajectories use handcrafted features and cost functions to model the interactions and constraints [65].

Most of the social force-based models attempt to learn the parameters of the force functions from real-world crowd datasets. When sailing in a crowd, humans have the innate ability to read the behavior of others, and everyone's actions depend on the people around them. The social force has been shown to achieve competitive results on pedestrian datasets [34,46], surveillance systems [11], autonomous driving [41], robotics [27,28], intelligent human tracking systems [27,28,30], abnormal crowded behavior detection [66], improving data association with people in the crowd [67], and other trajectory prediction applications [32,68]. Other methods, such as continuum dynamics [69], discrete choice framework [70], topics models [71], and Gaussian processes [25,30], have also been combined with the social force method to model human-human interactions with strong priors. These works targeted smooth motion paths and did not handle the problems associated with discretization.

Another line of work uses well-engineered features and attributes to improve prediction. Feature engineering is the process of transforming raw data into features that can better represent the potential problems of the predictive model, thereby improving the accuracy of the model for invisible data [72]. Alahi, et al. [73] presented a social affinity feature by learning human relative positions from their trajectories in the crowd. Hosseini and Maghrebi [74] utilized a social force model-based simulation engine to analyze the human behaviors, risks of fire emergency occurrence, and emergency evacuation in complex construction sites. Yi, et al. [75] proposed the use of human attributes to improve prediction in dense crowds. They also used an agent-based model to improve the accuracy in systems with a large amount of interconnected human agents. Rodriguez, et al. [76] analyzed videos of high-density crowds to track and count people. With the aid of topic models, Wang, et al. [77] were able to learn motion patterns in crowd behavior without tracking objects. These approaches were extended to incorporate spatiotemporal dependencies [78]. A mixture model of dynamic pedestrian agents that considers the temporal ordering of the observations was also presented [79]. Most of these models provide handcrafted features based on relative distances and rules for specific scenes.

While existing social force-based methods have made notable progress in addressing human trajectory prediction challenges, they suffer from the following limitation. Most of these methods provide handcrafted features based on relative distances and rules for specific scenes [47]. This results in favoring models that capture simple interactions, such as repulsion and attractions, and might fail to generalize for more complex and dynamic crowded settings. In contrast, over the past few years, time-series deep learning methods have been used to outperform the above traditional ones.

3.3. Time series deep learning

Because SPK is time series prediction in nature, scholars start to examine the recent deep learning methods for time series data analysis as a potential solution for further data mining and performance improvement, i.e., keeping improving prediction performance based on the same available datasets. RNN [43] is one of the most widely used deep learning methods to tackle the trajectory prediction problem. Combined with the LSTM [38] network, the RNN model retains long-term dependencies and avoids the vanishing and exploding gradient problems. Highlighted by the capability to perform sequence-to-sequence modeling, LSTMs have been successful in learning temporal sequences like future pedestrian trajectories conditioned on the history trajectories [47,80]. The applications of these methods were extended to robotics [81] and intelligent human tracking systems [73]. RNN models have also proven to be effective for tasks with densely connected data such as semantic segmentation [82] and scene labeling [83].

Inspired by the success of LSTM for different sequence prediction tasks such as handwriting [84] and speech generation [85], Alahi, et al. [47] extended LSTM for human trajectory prediction as well. Social-LSTM [47] is one of the earliest LSTM-based models focusing on pedestrian trajectory prediction. One limitation is that this model only analyzes nearby agents within a predefined distance range. Later work such as context-aware pool [86], LSTM-MDL [87], peek into the future (PIF) [88], Scene-LSTM [48], states refinement LSTM (SR-LSTM) [89], and CAR-Net based method [90] extended Social-LSTM with visual features and new pooling mechanisms to improve the prediction range and precision.

Some LSTM-based methods were proposed to capture the dynamic interactions of pedestrians, where the latent motions represented with the hidden states of LSTMs were shared by various mechanisms including "pooling" [47,91] and "attention" [80,92]. The "pooling" proposes to use social pooling on occupancy maps to collect the latent motion dynamics of pedestrians involved in a local neighborhood or the whole scene. Unlike the limitations of the local neighborhood or whole scene hypothesis, the "attention" mechanism helps to encode the relative influence and potential spatial interactions between pedestrians, because neighboring pedestrians have different importance for trajectory prediction. Compared with the "pooling" scheme, by assigning the different and adaptive importance to the pedestrians, attention-based models can achieve a better understanding of the crowd behaviors based on spatial interactions [93].

In addition, several attempts have been made to improve the performance of the time-series deep learning method. Behavior-CNN [94] used Gaussian processes to avoid the problems associated with discretization and could generate a smooth moving path [81]. Rodriguez, et al. [76] analyzed videos of high-density crowds to track and count people. Alahi, et al. [73] learned the pedestrian trajectories with relative positions as social affinity. Xu, et al. [95] assigned different weights to nearby pedestrians based on spatial affinity. Cai, et al. [96] used the context-based LSTM to integrate both individual movement and workplace contextual information to predict a sequence of target positions from a sequence of observations. They predicted workers' trajectories on unstructured and dynamic construction sites to ensure workplace safety. Due to the unscripted nature of construction sites, which places workers and equipment in close proximity, near-miss incidents or life-threatening contact crashes are possible. Rashid and Behzadan [97] combined polynomial regression (PR) and the hidden Markov model (HMM) to build a preemptive proximity-based safety framework to solve this problem. Zhu, et al. [98] proposed novel Kalman filters that could predict positions of the equipment and workers based on estimates from multiple video cameras to address the risks of injuries and deaths due to the workers being struck by mobile equipment on sites. Other methods use multiple networks to capture complex human-human and human-scene interactions [47,99,100]. These works showed that RNN models are capable of learning the dependencies between spatially correlated data such as image pixels.

However, time-series deep learning methods lack high-level and spatiotemporal structure [101]. They only focus on modeling interactions among people in close proximity to each other to avoid immediate collisions. They do not model human-human interactions in crowded scenes and do not anticipate interactions that could occur in the more distant future [47,91]. Hence, most of these methods only take advantage of the local interactions, but they do not have the capacity to model interactions between all people in a scene in a computationally efficient fashion, which leads to insufficient accuracy in long-term predictions.

3.4. Generative deep learning

The most recent advancements in deep learning have further provided promising solutions for SPK to address the data scarcity issue. Many data-driven approaches learn to predict deterministic future trajectories of humans by minimizing reconstruction loss [47]. However, human behavior is stochastic. The learning-based methods are considered to be more capable of handling the stochastic nature of human behavior as compared to feature-matching methods [102].

This aspect has been studied in the context of route choices to take at intersections [103,104]. Pedestrians have multiple navigation styles in crowded scenes such as mild or aggressive styles [105]. Therefore, the forecasting task entails outputting different possible outcomes. Generative models such as variational autoencoders [106] were trained by maximizing the lower bound of training data likelihood. An alternative approach, GANs [45], has been proposed where the training procedure is a minimax game between a generative model and a discriminative model. GANs overcome the difficulties in approximating intractable probabilistic computation and behavioral inference, and have been employed to approximate socially accepted motion trajectories in crowds [91,107–109], and GANs have shown promising results in tasks such as the generation of multiple socially acceptable trajectories given an observed past [91].

As aforementioned, Social-LSTM does not consider the interaction between agents and remote agents. Gupta, et al. [91] solved this problem by developing a novel pooling mechanism termed as Social-GAN to aggregate information across people involved in a global scene. Although these max-pooling methods handle multiple agents well, the permutation invariant functions such as max may discard important information as they might lose the uniqueness of their inputs [109].

In contrast, the social attention model [92] and scene semantic segmentation model [109] addressed the heterogeneity of social interaction differently among different agents by attention mechanisms [110] and spatial-temporal graphs [111]. Attention mechanisms encode which other agents are the most important to focus on when predicting the trajectory of a given agent [39]. The variation of construction machine poses can cause interactive on-site safety issues such as struck-by hazards. Luo, et al. [112] used GRU to recognize machine activities considering working patterns and interaction characteristics to predict future machine poses. Sadeghian, et al. [109] combined deep neural network features from the scene semantic segmentation model and GAN using attention to model human trajectory. Kim, et al. [113] developed a trajectory prediction model for mobile construction resources capable of predicting more than five seconds based on Social-GAN. For construction, this network can automate proximity monitoring and struck-by hazard detection to avoid construction accidents. GANs have been widely used in the study of human trajectory prediction and have become the most popular tool recently. However, both Social-GAN [91] and the scene semantic segmentation model [109] fell short of learning the truly multimodal distribution of human behavior and instead learned a single mode of behavior with high variance [114]. Unlike most of the previous work about human trajectories prediction [47,48,91,103,109], which oversimplified a person as a point in space, Liang, et al. [88] encoded a person through high semantic features about visual appearance, body movement, and interaction with the surrounding, motivated by the fact that humans derive such predictions by relying on one similar visual cues. Kosaraju, et al. [114] constructed a generative model term as Social-BiGAT that can learn these essential multimodal trajectory distributions, directly capture the interactions between pedestrians, and predict future paths based on the graph representation. Mohamed, et al. [115] developed Social-STGCNN which uses spatio-temporal graph to model the scenes based on Social-BiGAT [114], thereby benefiting more for graph representation [114]. These works show that GANs can not only predict possible interactions in the more distant future, but also model human-human interactions in crowded scenes.

3.5. Summary of the SKP methods

In summary, scholars initially used the clustering approach [40] for SKP-based HMP, but one major limitation of this approach is that the neglected human information can make the prediction of human trajectory inaccurate. After that, scholars introduced the social force modeling [41] which, however, still bore limitations related to relative distances and rules for specific scenes, preventing it from being generalized for more complex and dynamic crowded settings. With the development of data sensing technologies, deep learning methods, particularly RNN [43] and LSTM [38], began to be used in HMP. However, these methods lack high-level and spatiotemporal structure [101]. They are not able to model human-human interactions in crowded scenes or anticipate interactions that could occur in the more distant future [47,91]. More recently, GANs, which can overcome the difficulties in approximating intractable probabilistic computation and behavioral inference, have been employed in HMP [91,107–109]. GANs-based methods have shown promising results, as they can not only predict possible interactions in the more distant future, but also model human-human interactions in crowded scenes [91]. A summary of all SPK methods is illustrated in Fig. 2.

4. Ergonomics, Physiology and Cognition-based Prediction (EPC)

Despite the advancements of SPK methods, a pressing challenge still presents due to the lack of state information on the decision-making points, such as when/why a person will move forward. The lack of human information makes it challenging for SPK to predict human intentions. Higher granularity and causality of motions are neglected. Hence, how and why humans have their motions are not clear. The performance of HMP is insufficient for the requirements of intelligent system applications in construction, such as collision avoidance in human-robot collaboration in construction workplaces. To solve this problem, various human information needs to be considered in HMP. This level of body information includes ergonomics and physiological information. First, ergonomic methods can capture the initiations and key kinematic points of motions based on delicate body-oriented features, hence providing information about when humans would move. Second, physiological methods can explore the leading physiological signals for predicting the near future state change of motion intents, which can help model how a motion starts. Lastly, cognitive methods can move a step further to incorporate high-level cognitive and decision-making data into the prediction of an early intent for a motion. To summarize, it is necessary to consider the deeper driving factors of human motion. EPC-based HMP methods vary significantly in terms of the inherent analytical method. Depending on the specific prediction indicators, they may involve classic ergonomics models, statistical inferences, and ML, as reviewed in detail in the remainder of this section.

4.1. Ergonomics-based modeling

The ergonomic approach captures the initiations and key kinematic features of human motions based on delicate body-oriented data. Motion capture technologies are often needed to examine the finite sets of movements of key body components, such as joints, arms, low back, and feet. The information of motion capture provides rich information for modeling human ergonomics and predicting human motion.

4.1.1. Classic ergonomic models

The classic ergonomic models aim to predict human motion by digital human modeling and simulation [117], which contain a biomechanical representation of the human body along with the computational algorithms. For construction, the biomechanical simulation and prediction of human motion can solve the serious professional injuries problem that workers are suffering from [118]. There are two methods used for motion prediction, namely data-based models and physics-based methods [119]. The first method uses the anthropometric data and motion data collected from motion capture experiments to predict human motion. This method involves functional regression and data-driven model. The second method uses mathematical models (e.g., dynamic-based models or optimization-based models) to predict human

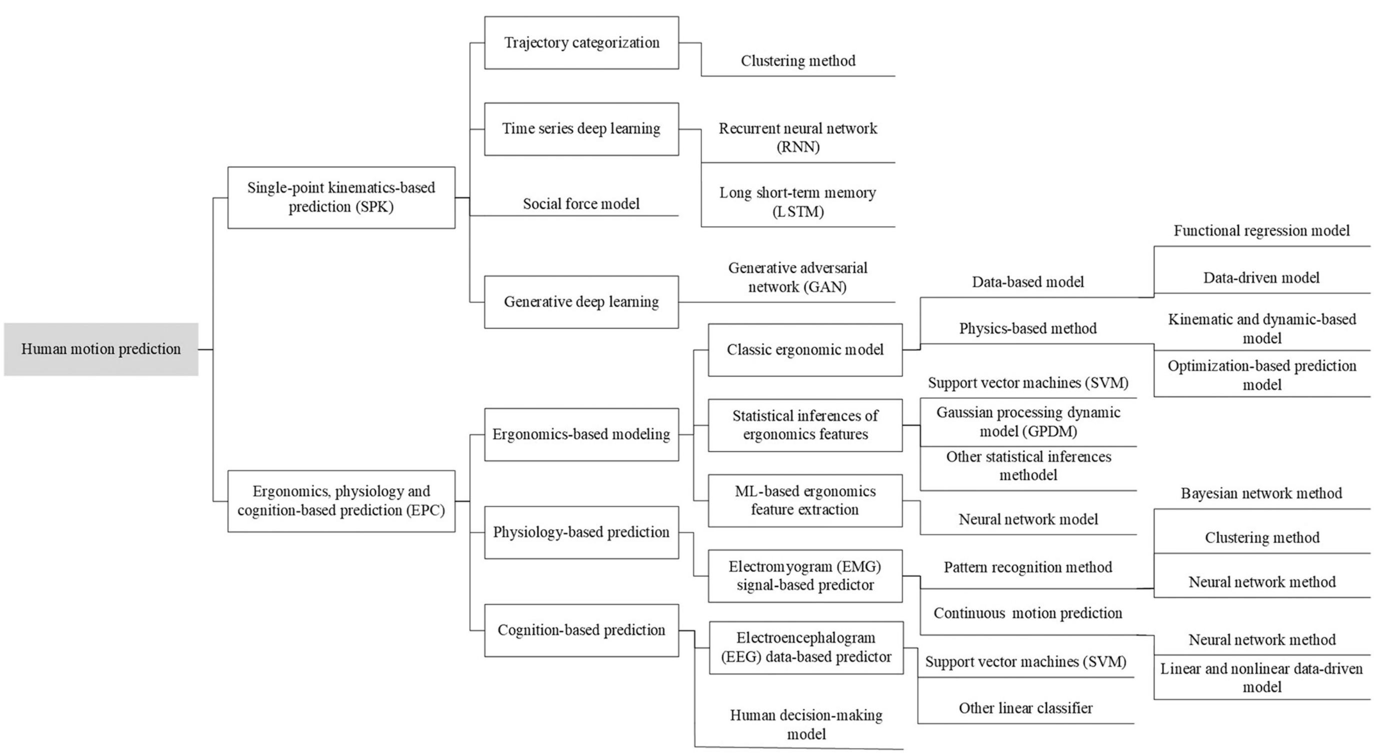

Fig. 2. The summary of HMP methods.

motion based on the biomechanics and kinematics data [117].

First, one of the widely used data-based models is functional regression [120], which uses a linear combination of the parameter functions. For example, the right-arm reaching motion was modeled, revealing the influence of height, age and personality on the choice of stretching exercise [121]. The method of non-parametric regression modeling was used to predict human motion [122] using motion capture data. Faraway and Reed [123] proposed a statistical method for digital human motion modeling. Mavrikios, et al. [124] used an additive model based on anthropometric parameters and motion coordinates to predict human motion. The lifting motion and the force-exertion postures were also predicted using the regression model [125]. One limitation with the above statistical models is that their predictive power is limited when extrapolating to novel, untested conditions. In addition, these models require high-quality data, which are typically collected either from thousands of experiments with human participants. Second, the data-driven models utilize a database to predict human motion or combine existing motions in order to generate new ones [119]. For instance, Park, et al. [126] introduced a novel simulation approach termed memory-based motion simulation (MBMS), which included a "motion modification" (MoM) algorithm to predict both seated upper body reaching and whole-body load-transfer motions. Woojin, et al. [127] developed a memory-based HMP model, which used real human movement samples recorded in motion capture experiments as templates to simulate new movements. This method could simulate the variability of human motion, but the representation of the human body was relatively simple, which limited the accuracy of posture planning.

As for physics-based methods, the direct inverse kinematics (IK) approach to posture prediction has received substantial attention. For example, Jung, et al. [128] developed an analytic reach prediction algorithm by employing the IK method. Similarly, a geometric IK algorithm to predict arm reach postures was proposed based on the criterion of minimization of the norm of joint angular velocities [129]. Lin, et al. [130] reported a dynamic simulation model based on biomechanical features for analyzing lifting activities, in which the IK method was employed. Hauberg and Pedersen [131] proposed a probabilistic interpretation of IK and extended it to sequence data to estimate the articulated human motion in visual data. There are an infinite number of possibilities to determine a posture due to excessive degrees of freedom (DOF) possessed by the human body, which, often referred to as the kinematic redundancy problem, still needs to be addressed [117].

A variety of optimization-based approaches have been proposed for addressing kinematic redundancy and computational complexity issues in HMP, where certain cost functions or performance criteria were hypothesized to represent presumed optimal strategies. First of all, optimization-based IK approach is one of the optimization-based approaches that can resolve the kinematic redundancy efficiently. For instance, the optimization-based differential IK approach was used for modeling three-dimensional (3D) human seated reaching motions [132], and predicting sitting and stretching motions [133] and gait motions [134]. Zou, et al. [135] proposed a two-layer nonlinear inverse optimization method to simulate and optimize a 52 degree-of-freedom human model. Xiang, et al. [136] proposed to use hybrid predictive dynamics (HPD) with the IK method to simulate, predict and track human motion. Second, in addition to optimization-based IK approach, the direct optimization-based method has also been considered, assuming that humans choose a posture to minimize certain motion objective functions. Multi-objective optimization has been widely used to predict the posture [137], upper body motion of humans in the car [138], human lifting motion [139] and human walking [140]. By minimizing the power cost and discomfort function at the same time, a multi-objective optimization method was proposed in [9] to predict human posture. A direct optimization approach was also used to predict human body joints' profiles [141]. To reduce the amount of required data for model input, Farahani, et al. [142] developed an anatomically detailed 3D human squat jumping model based on a zeroth-order optimization algorithm independent of the gradient information. Predictive dynamics was a novel optimization method for predicting human motion [143]. The planning and prediction of human arm motion is of particular interest to direct optimization-based methods, which can be further divided into deterministic and stochastic. For deterministic methods, the cost is typically expressed as the integral of certain deterministic functions over the movement time. For example, to predict the trajectory of the human hand, the minimum hand jerk criterion was proposed [144]. Through the optimization of the cost function, the robust optimization-based deterministic model was used to simulate and predict human arm motion [145]. For stochastic methods, random disturbances are included and the expected value of the cost function can be minimized. Mainprice, et al. [29] used an inverse reinforcement learning algorithm to learn the cost function from hand trajectories, and used the cost function and the motion planning stochastic trajectory optimizer (STOMP) to iteratively re-plan the trajectories to predict human hand motions. In another study, Svinin, et al. [146] tested four different dynamic models for dynamic prediction and evaluation of the human hand motions.

By considering the kinematic redundancy and computational complexity, the above classic ergonomic models have notably improved the performance of HMP. However, they also bear a few limitations. First, some of these models use optimization methods that are slow to converge on a solution [147]. Hence, the computational tractability of these models and their speed toward the real-time simulation of complex human motions still need to be strengthened. Second, most of these models were tested with limited motion databases [147], therefore, it has remained challenging for them to fully and accurately capture the characteristics of complex and large-scale human motions. To capture the realistic complexity of human motions, the level of physical realism in mathematical representations of the human motion should be improved [117,148]. It has been found that the psychophysical and biomechanical models could provide a potential tool for developing human posture and motion prediction models to improve the performance of dynamic prediction for complex human motions [148]. The accuracy of HMP can still be improved with a further consideration of ergonomic features.

4.1.2. Statistical inferences of ergonomics features

According to a series of behavioral experiments [149], people mainly judge the intentions of others by observing the gaze of others' eyes and the motion of others' heads as behavioral cues. By adding these behavioral cues, the intention prediction could be made more accurate. This has motivated a number of studies that aim to use behavioral cues to predict human motion intentions, by building statistical models that classify different types of data based on certain statistical features.

For HMP based on motion capture data and image motion contour data, the support vector machines (SVM) has been used effectively for classifying between two or more classes or estimating a continuous variable using regression, and it can be easily extended to multiclass classification through a single optimization. In addition, it has been used widely for walking direction estimation, walking intention and motion prediction in computer vision because of its efficiency and simplicity. For motion capture data-enabled HMP, Kadu and Kuo [150] used an SVM classifier based on a position histogram to classify human motion. For image motion contour data, Gandhi and Trivedi [151] proposed a probabilistic prediction model of pedestrian direction, which uses an SVM-based scheme to estimate the pedestrian direction, and uses HMM to model the conversion between directions and integrate the direction probabilities at different times. HMP based on motion contour histogram of oriented gradient (MCHOG) and SVM algorithm were also used to predict pedestrians' intention of entering the traffic lane [152], the initial gait of walking [153], status change between walking, standing, stopping and bending [154,155]. In addition, the standardized SVM [156] and HMM [157] were also used to determine whether pedestrians are willing to enter the road based on their postures and the position and displacement of the joints. The SVM classification is also employed to recognize construction workers' awkward motions to protect workers from cumulative trauma disorders. The method with the constructing tensor solves the problem that the recognition typically requires huge computational resources and complicated processes, and the constructing tensor has the potential in efficiently integrating multiple heterogeneous data sources [158].

The Gaussian processing dynamic model (GPDM) has also been used for imagery information-based HMP, allowing for nonlinear mapping from the latent space to the observation space, as well as a smooth prediction of latent points. For instance, Keller, et al. [159] classified human motion intention and realized short-term path prediction based on image motion features. Position cues obtained from detector information and motion features obtained from dense optical flow information were used in the Gaussian function through principal component analysis (PCA). Zhou, et al. [160] used the Kalman filter (KF) algorithm based on the Gaussian mixture model for pose correction, and then used Gaussian mixture regression (GMR) to predict the human motion pose. To leverage the higher-level information, Keller and Gavrila [161] proposed a prediction system that integrated the GPDM, the probabilistic hierarchical matching algorithm, the KF method, and its extension interacting with multiple model KF. The system was used to predict the path of pedestrians and to classify the moving action of the pedestrians. GPDM was also used to realize the prediction for walking and stopping trajectories [162], and to be integrated into a prediction framework to achieve a prediction for stopping, walking, and starting trajectories [163]. Using the balanced Gaussian process dynamics (B-GPDM) and the original Bayesian classifier, the position, posture, and intentions of the pedestrians could be predicted through the information obtained from a stereo camera on the vehicle [164]. Using the B-GPDM and the combination of different KF, the position, posture, and intentions of the pedestrians also could be predicted [165].

There are also other relatively lesser used statistical inferences methods for HMP. Based on different kinds of data including motion capture data with or without labeling, electromyography data, and accelerometer data, an unsupervised data segmentation method was used to automatically segment human motion data into different actions [166]. In addition, Matsubara, et al. [167] proposed a generative model including low-dimensional state dynamics and a two-factor (state-related observation basis and style parameters) observation model to predict the action sequence in real-time. As for image-based motion contour data, based on speed, time, and trajectory data extracted from stationary video-based marker and head-detection data, Goldhammer, et al. [168] used a piecewise linear model and S-type model to reduce the initial prediction time of gait and to improve the prediction performance. Dynamic Bayesian networks (DBN) [169] were used to predict human motion intention using the posture of the head and to determine whether pedestrians would enter a shopping mall. With the context information considered, Kooij, et al. [170] used DBN to capture pedestrians' situational awareness, situational criticality, and spatial layout of the environment as a potential state transition linear dynamic system, which was used to predict pedestrians' status changes between stopping and walking. Another method seen in the literature is the conditional random field modeling. Using the data obtained from head tracking, the lateral position of path prediction was performed using different models [171].

In summary, the human motion intention prediction using features extracted from the image is mainly based on statistical inference methods. However, the online prediction is difficult when using these methods due to computational cost issues, and the prediction ignoring the randomness of behavior may result in remarkable deviation from the reality. To address such limitations, more advanced analytical methods, particularly from the ML area, started to gain popularity.

4.1.3. ML-based ergonomics feature extraction

With more features considered in the HMP, the ML methods have been used in the ergonomics-based prediction. These ML-based ergonomics methods are groups based on different types of indicators and reviewed in detail as follows.

Based on inertial motion sensor data and with the recent advancements in ML, ANN [172] and RNN [173,174] were used for characterizing human postures, recognizing skeleton-based actions, and detecting the actions of skeletons in real-time. The dropout autoencoder LSTM (DAELSTM) was used to improve the prediction of long-term action sequences [175] and a global context-aware attention LSTM (GCA-LSTM) neural network [176] was proposed to realize online motion recognition from skeleton data and predict human motion in both online and offline states considering the context information. The radial based network (RBN) has also been used for similar purposes, as it could eliminate the ambiguity of the objective function, and improve the accuracy of prediction compared [177]. In addition, raw data was preprocessed to achieve accurate predictions with the ML methods. Fragkiadaki, et al. [178] used an encoding recursive decoder (ERD) model to recognize and predict human poses with motion capture data. The ERD model is a RNN that combines a non-linear encoder and decoder network. Li, et al. [179] proposed a new convolutional hierarchical autoencoder (CHA) framework. The CHA framework is superior to previous approaches in terms of computational complexity, storage capacity, and long-term and short-term predictions, although it demonstrates an unbalanced preference of actions.

As for imagery data, Goldhammer, et al. [180] used a self-learning approach based on ANN to classify and predict pedestrians' walking, starting, stopping and bending. Based on the head position and velocity information obtained by head tracking, polynomial least-squares approximation and multilayer perceptron neural networks were used to predict the continuous position over a period of time [181]. As more models have gradually begun to consider context information, a new fuzzy finite automata prediction method was proposed to improve the prediction of pedestrian intentions [182]. Later, Kwak, et al. [183] used the dynamic fuzzy automata (DFA) method to predict the intention of pedestrians in a continuous sequence, based on features of the distance between curbs and pedestrians, the pedestrians' speed and head direction.

In summary, the ML methods have the capacity of handling the stochastic nature of human motions. They also help achieve higher accuracy by solving the challenges associated with long-term data sequences and online prediction. These methods have provided promising and evolving solutions in ergonomics-based HMP.

4.2. Physiology-based prediction

The physiology-based prediction methods predict human motions by interpreting humans' biological signals. As human motions are controlled by muscles, the electromyogram (EMG) signal, which contains information about joint angle, joint torque and muscle force, is one of the widely used biological signals for human motion intention prediction. There are two types of EMG signals including surface EMG (sEMG) acquired with electrodes that are placed on the surface of the skin just above the target muscle, and intramuscular EMG (imEMG) detected with needles or wires that are inserted into muscles. There is no significant difference between the two types of EMG signals in terms of measuring motion classification accuracy [184], hence we make no distinction between these two types of signals hereafter.

Existing EMG-based HMP methods are based on either pattern recognition or continuous motion prediction. For pattern recognition, there are several commonly used methods including Bayesian network methods, clustering methods, and neural network methods. Firstly, LDA, which is a special Bayesian network method, is the most commonly used method for HMP with EMG. Multiple types of upper extremity motions were recognized based on the LDA classifier [185]. The state-space model [186] was also used to estimate arm force and motion using EMG signals for the control of exoskeletons. Secondly, the K-mean-based clustering method was used to explore the relationship between the activities of peris shoulder EMG and different handshake postures and arm directions for the applications of prosthesis [187]. Fukuda, et al. [188] used the electromyographic EMG signals to recognize eight different hand motions using log-linearized Gaussian mixture network (LLGMN). To explore the best filtering and classification methods for hand motion prediction based on sEMG, Zhou, et al. [189] used three different types of classifiers, and found that the Gaussian mixture model had the best performance. Thirdly, as for the neural network methods, the ANN was used to recognize six kinds of wrist motions using the myoelectric lower arm signal for the application in a non-invasive EMG computer interface [190]. ANN was also used for the prediction of hand tasks [191], the prediction of upper-limb [192], and robotic prosthesis myoelectric control based on motion prediction [193]. The decision theory and context information were used for EMG-based human-robot interfaces with Bayesian and neural networks to reduce the difficulty of classification and improve the accuracy [194]. In order to solve the problem of a limited number of pattern recognition systems of EMG signals, an adaptive neuro-fuzzy inference system was developed for the sEMG-based identification of hand motion commands [195]. Batzianoulis, et al. [196] used the dynamic EMG signal in the pre-motion of grasping to decode the grasping posture for robot control in human-computer interaction. Duan, et al. [197] used a wavelet neural network combined with the discrete wavelet transform to predict hand motion commands with a small number of signals. The methods used for EMG-based pattern recognition have increasingly focused on how to use the least signal channels to achieve real-time prediction and higher accuracy. However, EMG based on pattern recognition only realizes the discrete motion prediction, which is different from the natural and smooth motion of human beings. To solve this problem, the continuous HMP based-EMG signal has been explored.

The commonly used method for continuous HMP is ANN, which can be utilized to approximate the nonlinear relationship between s-EMG and continuous motions. Considering muscle contractions and the contact between the body and the objects, Kwon and Kim [198] used sEMG and joint angular velocity to achieve real-time prediction of upper limb motion based on the method of ANN. ANN was also used to predict the human shoulder joint angle from zero to ninety degrees in a virtual reality rehabilitation system [199], identify forearm flexion and extension motions in real-time [200], and predict simultaneous and continuous shoulder and elbow motion [201]. The back propagation neural network (BPNN) calculates the derivative flows backward through the network, which improves the estimation accuracy and realizes real-time prediction. This nature of BPNN is used to minimize the error of the network using the derivatives of the error function. Zhang, et al. [202] achieved sEMG-based continuous prediction on joint angles of human legs by using BPNN. The angle of the shoulder and elbow joints [203], the joint motion [204], and the continuous joint angle prediction in a short time [205] were also estimated using the BPNN. More different neural networks have been used for continuous motion prediction for different purposes. Loconsole, et al. [206] adopted the EMG from five muscles in the shoulder and elbow to predict joint torque online for the purpose of online control of exoskeleton by using the time-delay neural network. Akhtar, et al. [207] explored the extent to which the EMG in different positions including shoulder, upper arm and forearm can be used as an indicator to predict the angle of distal arm joint for the normal and disabled people using the methods of locally weighted projection regression (LWPR) and time-delay adaptive neural network (TDANN). The other methods include linear and nonlinear data-driven models. Firstly, there were other state-space models, including state-space vector model [10] and state-space EMG model [208,209], that were used to continuously predict human joint angles. Then online projection regression [210] and temporally smoothed multilayer perceptron (MLP) regression [211] were used to predict the joint angles and continuous lower limbs' motion respectively. In continuous HMP based-EMG, the methods study the basis of continuous prediction, and aim for high-accuracy, real-time and smooth HMP in human-computer interaction.

4.3. Cognition-based prediction

Unlike the rich body of literature on HMP based on the ergonomic and physiological modeling, there have been fewer studies on HMP relying on the neurofunctional data and the corresponding cognitive theory. EEG can provide more information about how humans think and whether they decide to initiate a motion. The intention to move is associated with at least two cortical activities including movement-related cortical potentials (MRCP) and event-related desynchronization over sensorimotor and supplementary motor cortices [212]. Therefore, it is possible to use pre-motion EEG to classify the motion pattern and realize the online prediction of human motion intention. In the meanwhile, the cognitive theory can explain human behavior by understanding human processes of thoughts and decisions. It provides a deeper understanding of how to predict human motions long before they happen.

Efforts have been made to leverage electroencephalogram (EEG) data for human movement decision prediction. The combined methods of independent component analysis (ICA), power spectral density estimation (PSD) and SVM [213] were used to predict human motion from single-trial EEG. Shafiul Hasan, et al. [212] also used an SVM classifier in a ten-fold cross-validation scheme to predict the intention of walking using EEG data. SVM, Mahalanobis linear distance (MLD) classifier, GA-based MLD classifier and decision tree classifier were all used for HMP, though there was no significant difference among these three methods [214]. In addition, a new PCA [192] and neural network were also proposed to predict human intention regarding hand motions using the combined data of the EEG and EMG. For other linear classifiers, Sburlea, et al. [215] used sparse linear discriminant analysis (SLDA) to detect the intention to walk of healthy people and stroke patients, respectively. It allows HMP to be made with a small amount of EEG data. In addition to the pursuit of high accuracy of offline classification with fewer data and time, subsequent research has been focusing on how to achieve online predictions, which is demanding in terms of computing power and speed. For instance, Bai, et al. [216] explored how to predict human voluntary motion in real-time before it occurs using EEG signals. They used the MLD classifier as a modeling method to predict motion intention in advance before the actual motion was monitored.

Most recently, several studies have attempted to predict human motion intention based on cognitive frameworks. Turnwald, et al. [217] applied the game theory to the decision-making process of human navigation to predict actual human trajectory. By generating a framework to identify potential factors in the decision-making process with an unsupervised method, Zou, et al. [100] introduced a model to deconstruct the decision-making process to infer the future trajectories and intention of pedestrians. In addition, Yang and Howard [19] integrated the final comfort state theory in the optimization-based motion prediction formulation to predict hand posture. These methods solved the problem that the methods without cognition theory have no decision-making or task planning capabilities, thus severely limiting the potential power of HMP. In summary, the cognition theory focuses on the motion control, planning, and decision-making aspects behind human motion and it strives to answer the "why" of human motion. The adoption of cognition theories has shown a substantial potential to improve the accuracy of predicting human motion.

4.4. Summary of the EPC methods

In summary, developed upon different theories and approaches, the EPC methods can be divided into three groups including ergonomics-based modeling, physiology-based prediction and cognition-based prediction. First of all, early EPC methods mainly relied on ergonomic models, which only considered the behavioral and ergonomic characteristics of humans [117]. Later on, motivated by the increased accessibility to the anthropometric data and motion data, research began to explore the data-based models, while at the same time the physics-based methods were also introduced based on the biomechanics and kinematics data. Although these data-based methods have improved the prediction of human posture and movement, they typically involve high computational complexity and lack the level of physical realism in mathematical representations of human motion [148]. With behavioral cues being added, statistical models that classify different types of data based on certain statistical features are used. SVM, GPDM and other statistical inferences methods show promising performance in predicting human motion, although the online prediction issue has remained challenging because of the high computational cost. As ML methods have the capacity of handling the stochastic nature of human motions, recent studies have begun to explore ML methods, such as ANN [172], RNN [173,174] and RBN [177] to improve the HMP accuracy. Second, for physiology-based prediction, EMG is a widely used indicator that can provide useful information for predicting human motion, and various pattern recognition methods, such as Bayesian network methods [185], clustering methods [187] and neural network methods [191] have been applied for interpreting the EMG data to predict human motion. In addition, the continuous motion prediction using neural network models [199], and linear and nonlinear data-driven models [208,209] have also been explored. These physiology-based prediction methods can contribute to high-accuracy, real-time and smooth HMP, though they have not been widely integrated with ergonomic models. Lastly, in order to better understand people's motion intentions so as to predict human motion more accurately, researchers have recently begun to explore cognition-based prediction methods, which use EEG data-based predictors (using SVM [213] and other linear classifiers [216]) and decision-making model for human motion decision prediction [19]. With these methods, HMP has the potential to consider the human decision-making or task planning. Limited by the scarcity of relevant literatures, the continuous and real-time prediction still remains to be explored. A summary of all EPC methods is illustrated in Fig. 2.

5. Discussions

5.1. The evolution of HMP rationale

This review finds that the rationale and main assumption behind the HMP methods has been evolving rapidly in the past three decades, representing a deepening understanding of the subject matter. Initially, the HMP mechanism is based on the straightforward physical characteristics of the human motions, modeling humans as a point or particle. Only the features of trajectory kinetics such as the past path, the speed, and the direction traveled by the agent in the past period are considered. The major assumption is that the trajectory of an agent in the past represents the momentum of motions of the agent, and the location and motion characteristics of the agent in the future can be inferred according to past trajectory. With this basic assumption, the HMP is not different from the trajectory prediction of any other moving objects, that they are all treated as moving particles controlled by basic physics. It should be noted that this is an oversimplified assumption as when there are interactions among humans, between humans and the environment, and/or when the stochasticity of human motions cannot be ignored, the prediction can fail. As a result, the interactions among human agents and aspects of interaction with the surrounding environment are added to the formula or as constraints to enable a more accurate prediction result. Although the influence of other agents in the same environment is considered in these improved methods, these early efforts of modeling interactions are fairly basic. To further push the methodological frontier, the complexity and stochasticity of human interactions begin to be captured in the newer studies. Besides, newer methods to model the interaction between humans and the environment are also incorporated gradually, including the influence of environmental constraints. Examples of new variables include the distance and the interaction rules between an individual and the environment in the process of moving. As for the methodological innovations, the KF, multiple linear dynamical models, nonparametric stochastic models, and trajectory matching models started to appear in the literature.

With the gradual deepening of the understanding of the subject matter, scholars in this area begin to investigate the behavioral and ergonomic characteristics of humans that ultimately lead to the initiation and triggering of movement, and to predict human motions based on these causal relationships, or the so-called motion intention prediction. The intention prediction tends to be more accurate by adding behavioral cues, for these cues are necessary for humans to judge the intentions of others and make movement decisions. The richer information about the ergonomic and behavioral features, or the body information, includes the biomechanics data of the humans, the intuitive imagery information of the different phases of a motion, and the complex physiological and biochemical features that provide more information about motion intentions. The body information adds more clues about key nodes of the human body to human motion intention prediction. The whole motion is disassembled and the key nodes of the human body provide a basis for judging whether humans have the intention to move. In addition, any human motion is achieved by the control of musculoskeletal processes, and all the peripheral nervous systems are controlled by the brain. The biochemical information such as the EMG can provide information about how humans move, think, and whether they will move. As such, researchers have begun to think about how any human motion occurs and why. In particular, EEG reveals the neural mechanism of human brain cognitive processing and provides us with a deeper cognitive understanding of human motion. Meanwhile, to handle the increasing computing needs, ML methods have been gradually used in the HMP literature. Most recently, the HMP literature starts to explore the cognitive basis of motion planning, based on EEG which reveals when humans intend to move, and existing decision-making theories in which how humans make a decision is considered. Although representing a relatively new direction of research, several frameworks of human decision-making combining the trajectory information and human behavioral cues are introduced in recent literature for a better HMP over a long period of prediction window.

To sum up, our understanding of HMP has shifted from relying on the basic trajectory-based methods to be on deeper physiological signaling mechanisms. In addition, the evolution of information fusion by integrating multiple heterogeneous data sources has also been observed. With the increasing understanding of the underlying mechanisms of human motion, the intentional and decisional variables are captured. Hence, we can also observe that in the construction industry, more and more sensors are being applied to predict human motion, examples of which include inertial measurement units (IMUs), EEG, and heart rate sensors, among others [158]. The increased number of variables in the HMP methods also extends the required prediction time for bigger and more complex models. Based on the specific HMP needs, choosing a subset of relevant information may be more conducive to enabling a more efficient HMP for context-specific applications.

5.2. Evolution of HMP methods

Another clear trend this review reveals is the evolution of prediction and modeling methods that are becoming more apt to cope with the scarcity and quality issues of the training data in HMP. Like many other predictions, the performance of HMP builds on a significant amount of quality data for feature exaction, modeling, and predicting. However, human motions usually represent substantially dynamic and uncertain processes, showing apparent differences among different people, or the changing behavior of the same person at different time points or in different contexts and/or environments. These differences can be driven by the variations in human cognitive and physiological status, a perceived difference in task requirements, contexts and environments, and so on. Early unsupervised learning methods, such as clustering analysis [40], are less concerned about the driven factors of human motion, assuming that a "top-down" similarity clustering would help reveal similar patterns and therefore reduce the dimensionality in prediction. The requirements of these early methods on the quantity and quality of data are both minimum but are more biased toward an analysis of only the outcomes of human motions (such as labeling a motion), instead of the continuous spatiotemporal analysis or causal analysis of human motions. Later methods, such as social force molding [41] and activity forecasting methods [103], began to look into the driving mechanism in the initiation and development of motions (such as interpersonal avoidance behaviors in crowded situations). These methods are more effective in capturing the nature and spatiotemporal features of human motion, however, they are also more demanding on the amount and quality of the required input data, which sometimes could be highly challenging. As a result, scholars start to turn to the latest advancements in the deep learning area for possible solutions. The resulting newer methods are either capable of retaining long-term dependencies and avoiding the vanishing and exploding gradient problems (e.g., RNNs [43]), or can generate simulated data based on the learned patterns (e.g., GANs [45]). Besides RNNs, it has been recently reported that CNN-based models are well-suited to perform sequence-to-sequence tasks [218,219]. Unlike the LSTM-based models which attend sequentially to each frame, the CNN-based models support increased parallelism and effective temporal representation [220]. While the predictions with RNNs are inherently sequential and the later time-steps cannot be computed until the earlier ones are completed, CNN takes advantage of its parallelism and can perform training and inference in a more time-efficient manner [220].

By preserving or enriching the most relevant information for feature extraction and prediction, these deep learning methods show a great promise in HMP with limited or flawed training data. The HMP performance, in terms of precision and accuracy, has greatly improved fueled by the new deep learning methods. Meanwhile, various data are used to predict HMP using these deep learning methods. More importantly, because of the improved robustness to low-quality training data, scholars are able to expand their research focus from simpler human motions (such as upper extremity motions [185]), to more complex motions (such as human voluntary motion [216]), and in a more realistic setting such as that with interpersonal impacts and dynamic environments [47,91]. These deep learning methods have also been increasingly applied in the practice of HMP. From this methodological roadmap, it is not difficult to find that HMP remains a prosperous research area because of the additional benefits and intellectual merits (from the data science perspective) of employing new data analytics and modeling methods in promoting HMP. This is a relatively independent research roadmap in parallel with the paradigm changes but is also supporting the development of new prediction paradigms.

5.3. Application domains using HMP

Another finding is that the information used for HMP differs across various application domains. In the fields of autonomous driving, robotics, and abnormal crowded behavior detection, the historical trajectory information is easy to obtain and is therefore widely used to predict human motion. The prediction of human motion intention based on body information is also extensively utilized in many other fields. In exoskeletons, robotics, and healthcare applications, the EMG has been gradually used to pursue the accuracy of continuous motion prediction. The development and use of brain-computer interfaces have also gained popularity for classified offline EEG data to achieve online real-time prediction before the next action. The motion capture data, the body information such as body orientation, and the historical trajectory data have been used in traffic research for predicting pedestrian trajectory and pedestrian intention. In the human-computer interaction, the motion capture data and the historical trajectory data are all used to realize HMP for the simple interaction of humans and computers. With the pursuit of real-time prediction in different fields, the speed of computation is critical. However, the tradeoff between having multiple indicators for improving the accuracy and having fewer indicators for reduced computational complexity is still a debatable issue in the literature.

As for the areas of construction, built environments, and engineering operations, the HMP has also seen various valuable applications. For example, abnormal crowd behavior detection using workers' trajectories can be used by project managers to monitor workers' unsafe behaviors and intervene in them in time. HMP can also be applied for autonomous cars and robotics modeling in construction, as the trajectories and behavioral cues can also be used for intelligent systems to predict human motions and to avoid human-robot collisions during construction activities. The biochemical information such as the EMG and the EEG can be collected for HMP to avoid the undesired operations and prevent occupational injuries at construction job sites.

5.4. Gaps and directions for future research

Findings in this review indicate that several gaps have remained to be addressed in the future. Firstly, from the comprehensive review of the HMP literature, the challenge is the variety and diversity of the existing data and methods. For SPK, researchers predict human trajectories based on human-human interactions and human-environment interactions. Due to human behavior stochasticity and environmental complexity, a variety of methods have been developed for different scenarios. In the study of EPC, the data and models need to focus on different parts of the human body, leading to major differences in prediction methods and results. Scholars in different fields need to spend significant amounts of time searching literature related to their own research. Given the complexity of the HMP problem, there is no single methodology that can fit all research needs. As such, researchers may be benefited from a descriptive framework that can properly categorize different HMP methodologies existing in the literature, associate them with diverse application scenarios, and provide clear guidance to help researchers select the appropriate HMP methodology to meet specific research needs.

Secondly, the understanding of the dynamics of work context in existing research can be further improved. Most existing approaches have focused on predicting human motion based on well-defined tasks for a relatively short period of time, which are relatively static settings without fully considering the changes in work context and uncertainty in human motion. It should be noted that in reality, human motion is dynamic and not entirely dictated by well-defined restrictions, meaning that the changes in human motion patterns (such as how hand motions and gaze focus are coupled, see [221,222]) are inevitable, especially in complex and dynamic settings.

Thirdly, although some existing approaches have considered different human motions, the stochasticity uncertainty, i.e., differences in behavior among different individuals, or the changing motion patterns of the same individual at different points of time or in different contexts and/or environments, still needs further consideration. In addition, driven by the complexity and the variability of human motions, there is usually a tradeoff between the length of the prediction window and the accuracy of prediction results. How to build a model to fully extract the features of the stochasticity uncertainty in human motions, and to analyze different human behaviors has remained an open problem. A possible solution could be explored to address the stochastic uncertainty. Instead of assuming that a person remains consistent in terms of motion patterns, data collected from the same person should be further segmented and categorized into distinct pattern groups based on the unique motion characteristics. Secondary metrics, such as EEG and EMG, can be used to identify motion pattern changes in a continuous activity, facilitating an automated segmentation and categorization of human motion data.

Following the existing literature, a possible advancement for HMP may be integrated prediction that employs and combines principles from multiple methods and/or disciplines of knowledge. For example, the social force method was often used alone in earlier efforts [41]. Later studies have attempted to combine it with other methods, such as social LSTM [47] and social GAN [91], to obtain more efficient and accurate results. Similarly, particle-based methods have been coupled with other methods, such as Gaussian processes, to improve the performance of prediction. We believe that there is potential room for further exploration in multi-method combinations, such as combining RNN and GAN. For instance, for the EPC method, combining the SVM, DBN, and the LDA classifier may improve the prediction performance. In addition, considering the complexity and the variability of human intention, the multi-method combination may better adapt to the variety and diversity of data. Recent studies have attempted to combine insights from social sciences with methods rooted in computer vision and pattern recognition to analyze the motion of groups and crowds [223]. The particle-based prediction can also be advanced by incorporating knowledge models (e.g., social force) and secondary leading indicators (e.g., EEG or EMG). Based on the growing evidence in the literature, we expect that multidisciplinary approaches will have a great influence on HMP in complex, dynamic and crowded environments [47].

In addition, with the latest development of body-carried sensors and data-sensing technologies, there is potential to collect new types of human data for training HMP models. For example, high precision eye-tracking can be achieved with head-mounted systems. These systems mount an eye tracker on a helmet or glasses-like structure near the eyes, such as Applied Science Laboratories (ASL), SensoMotoric Instruments (SMI) Tobii Pro [224], Ergoneers, and SR Research [225]. As for capturing the human hand motion, marker-based motion capture and IMU-based sensors are increasingly used [226,227]. These data may motivate a new direction of HMP methods that leverages the emerging approaches in the deep learning domain. An example of such deep learning approaches is the automated machine learning (AutoML). By enabling the automated construction of an ML pipeline on a limited computational budget [228], AutoML provides methods and processes to make ML available for non-ML experts, and improves the efficiency of ML. Existing AutoML applications in object detection [229], GAN [230] and video tasks [231] may inspire the future advances of HMP in the construction field.

For the construction industry, a few datasets for human motion recognition research are available. For example, Maurice, et al. [232] created a dataset of human motions in industry-like activities. Guerra-Filho and Biswas [233] built a dataset including human motion and cognition. However, the majority of these datasets are only for generic motion recognition purposes and are not adapted to specific settings or needs in construction. The mature and sufficiently large datasets are lacking in the construction field to develop new methods for human motion prediction, and its importance has been gradually realized by researchers. One of the purposes of this paper is to urge the construction research community and industry to collect more high-quality data for construction-specific human motion prediction modeling. One of the proposed efforts is to utilize remote and body-carried sensors for data collection to build potential datasets for the construction industry. It is worth noting that construction operations are characterized by the open and evolving work environment, dynamic and changing workflows, and hard-to-define human-robot cooperation requirements [234]. Most existing datasets can hardly represent realistic and complex construction tasks, and there is a pressing need for diverse datasets that capture and integrate the representative human information in construction. This can be facilitated by the use of advanced data collection technologies. For example, for eye-tracking data, high-precision eye-tracking can be achieved with head-mounted systems. Considering that construction workers typically wear protective goggles, the eye trackers could be mounted on them and a satisfactory accuracy could be achieved [235]. For human hand motion data, it can be collected by marker-based motion capture systems and IMU. Advanced camera systems, such as the OptiTrack camera system (Nature Point, OR, USA) and high-end Vicon MX-f20 camera system (Vicon, Oxford, UK), have also been introduced and proven to be effective for human hand motion capture [227]. As for physiological data, a large number of artifacts are expected if physiological signals are collected when workers are at work. To address this issue, recent research has integrated EEG into the safety helmets worn by construction workers to assess their attention level [236], and the results demonstrated the feasibility of combining physiological signal-gathering devices with safety gears to help build motion recognition datasets. In addition, an affordable open-source hardware and software platform, such as the one proposed in [237], could be utilized to enable multiple groups to collaborate in developing the datasets. Last but not the least, the data-sharing culture should be further cultivated to encourage the creation and sharing of human motion recognition datasets in construction.

6. Conclusions

This paper presented a thorough review of the human motion intention prediction. The literature about the motion and path planning of an individual in an indoor or outdoor environment, and articulated full-body motions in workplaces across multiple domains was surveyed. A taxonomy of motion prediction techniques was proposed based on the theoretical understanding of human motion. This taxonomy has three important levels of information, namely physical-based kinematics information, body information, and decision context information, which differ in the degree of human cognition considered in prediction and the assumptions made about human behaviors. Based on the above taxonomy, the indicators, methods, and application domains of HMP using each level of human information were synthesized. Finally, we summarized and discussed the historical evolution of the theoretical understanding of human motion and the methodological solutions to HMP, and outlined a few potential directions for future research. This paper provides enlightenment for better fulfilling the need of HMP in various domains and in the construction domain in particular. It will hopefully increase the visibility of this rapidly expanding research area and stimulate further research for continuous theoretical, methodological and technical advancements.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

References

[1] J. Martinez, M.J. Black, J. Romero, On human motion prediction using recurrent neural networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2891–2900, https://doi.org/10.1109/CVPR.2017.497.

[2] K.M. Rashid, A.H. Behzadan, Enhancing motion trajectory prediction for site safety by incorporating attitude toward risk, Comput. Civil Eng. 2017 (2017) 425–433, https://doi.org/10.1061/9780784480847.053.

[3] A. Rudenko, L. Palmieri, M. Herman, K.M. Kitani, D.M. Gavrila, K.O. Arras, Human motion trajectory prediction: a survey, Int. J. Robot. Res. 39 (8) (2019) 895–935, https://doi.org/10.1177/ToBeAssigned.

[4] K.M. Rashid, S. Datta, A.H. Behzadan, Coupling risk attitude and motion data mining in a preemtive construction safety framework, winter simulation conference (WSC), Las Vegas, NV (2017) 2413–2424, https://doi.org/10.1109/WSC.2017.8247971.

[5] R.S. Novin, E. Taylo, T. Hermans, A. Merryweather, Development of a novel computational model for evaluating fall risk in patient room design, Fall Risk Eval. Healthc. Facilit. 14 (2) (2020) 350–367, https://doi.org/10.1177/1937586720959766.

[6] M. Bataineh, T. Marler, K. Abdel-Malek, J. Arora, Neural network for dynamic human motion prediction, Expert Syst. Appl. 48 (2016) 26–34, https://doi.org/10.1016/j.eswa.2015.11.020.
[7] G. Ferrer, A. Sanfeliu, Bayesian human motion intentionality prediction in urban environments, Pattern Recogn. Lett. 44 (2014) 134–140, https://doi.org/10.1016/j.patrec.2013.08.013.
[8] L. Lentz, J.R. Randall, D.P. Gross, A. Senthilselvan, D. Voaklander, The relationship between physical fitness and occupational injury in emergency responders: a systematic review, Am. J. Ind. Med. 62 (1) (2019) 3–13, https://doi.org/10.1002/ajim.22929.
[9] H.-J. Kwon, H.-J. Chung, Y. Xiang, Human gait prediction with a high Dof upper body: a multi-objective optimization of discomfort and energy cost, Int. J. Human. Robot. 14 (01) (2017) 1650025, https://doi.org/10.1142/s0219843616500250.
[10] Q.C. Ding, A.B. Xiong, X.G. Zhao, J.D. Han, A novel emg-driven state space model for the estimation of continuous joint. International Conference on Systems 2011, IEEE, 2011, pp. 2891–2897. doi:10.1109/ICSMC.2011.6084104.
[11] M. Yasuno, N. Yasuda, M. Aoki, Pedestrian detection and tracking in far infrared images. 2004 Conference on Computer Vision and Pattern Recognition Workshop, IEEE, 2004, p. 125. doi:10.1109/CVPR.2004.417.
[12] D. Tran, W. Sheng, L. Liu, M. Liu, A. Hidden, Markov model based driver intention prediction system. 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), IEEE, 2015, pp. 115–120. doi:10.1109/CYBER.2015.7287920.
[13] K. Saleh, M. Hossny, S. Nahavandi, Intent prediction of vulnerable road users from motion trajectories using stacked Lstm network, in: 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), IEEE, 2017, pp. 327–332, https://doi.org/10.1109/ITSC.2017.8317941.
[14] S. Hellbach, J.P. Eggert, E. Körner, H.-M. Gross, Time series analysis for long term prediction of human movement trajectories, in: International Conference on Neural Information Processing, Springer, 2008, pp. 567–574, https://doi.org/10.1007/978-3-642-03040-6_69.
[15] B. Morris, M. Trivedi, A survey of vision-based trajectory learning and analysis for surveillance, IEEE Trans. Circuits Syst. Video Technol. 18 (8) (2008) 1114–1127, https://doi.org/10.1109/TCSVT.2008.927109.
[16] S. Lefèvre, D. Vasquez, C. Laugier, A survey on motion prediction and risk assessment for intelligent vehicles, Robomech J. 1 (1) (2014) 1–14, https://doi.org/10.1186/s40648-014-0001-z.
[17] G. Cheron, M. Duvinage, C. De Saedeleer, T. Castermans, A. Bengoetxea, M. Petieau, K. Seetharaman, T. Hoellinger, B. Dan, T. Dutoit, F. Sylos Labini, F. Lacquaniti, Y. Ivanenko, From spinal central pattern generators to cortical network: integrated Bci for walking rehabilitation, Neural Plast. 2012 (2012), 375148, https://doi.org/10.1155/2012/375148.
[18] A. Shakeel, M.S. Navid, M.N. Anwar, S. Mazhar, M. Jochumsen, I.K. Niazi, A review of techniques for detection of movement intention using movement-related cortical potentials, Comput. Math. Methods Med. 2015 (2015), 346217, https://doi.org/10.1155/2015/346217.
[19] J. Yang, B. Howard, Prediction of initial and final postures for motion planning in human manual manipulation tasks based on cognitive decision making, J. Comput. Inf. Sci. Eng. 20 (1) (2020), 011007, https://doi.org/10.1115/1.4044586.
[20] L. Bi, A.G. Feleke, C. Guan, A review on Emg-based motor intention prediction of continuous human upper limb motion for human-robot collaboration, Biomed. Signal Process. Control 51 (2019) 113–127, https://doi.org/10.1016/j.bspc.2019.02.011.
[21] J. Walker, A. Gupta, M. Hebert, Patch to the future: unsupervised visual prediction, in: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2014, pp. 3302–3309, https://doi.org/10.1109/CVPR.2014.416.
[22] M.W. Turek, A. Hoogs, R. Collins, Unsupervised learning of functional categories in video scenes, in: European Conference on Computer Vision 6312, Springer, 2010, pp. 664–677, https://doi.org/10.1007/978-3-642-15552-9_48.
[23] G. Zen, E. Ricci, Earth mover's prototypes: a convex learning approach for discovering activity patterns in dynamic scenes. Conference on Computer Vision and Pattern Recognition 2011, IEEE, 2011, pp. 3225–3232. doi:10.1109/CVPR.2011.5995578.
[24] D. Kuettel, M.D. Breitenstein, L. Van Gool, V. Ferrari, What's going on? discovering spatio-temporal dependencies in dynamic scenes. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2010, pp. 1951–1958. doi:10.1109/CVPR.2010.5539869.
[25] M.K.C. Tay, C. Laugier, Modelling smooth paths using Gaussian processes, in: Field and Service Robotics 42, Springer, 2008, pp. 381–390, https://doi.org/10.1007/978-3-540-75404-6_36.
[26] R.P.D. Vivacqua, M. Bertozzi, P. Cerri, F.N. Martins, R.F. Vassallo, Self-localization based on visual lane marking maps: an accurate low-cost approach for autonomous driving, IEEE Trans. Intell. Transp. Syst. 19 (2) (2017) 582–597, https://doi.org/10.1109/TITS.2017.2752461.
[27] M. Luber, J.A. Stork, G.D. Tipaldi, K.O. Arras, People tracking with human motion predictions from social forces, in: 2010 IEEE International Conference on Robotics and Automation, IEEE, 2010, pp. 464–469, https://doi.org/10.1109/ROBOT.2010.5509779.
[28] A. Bera, S. Kim, T. Randhavane, S. Pratapa, D. Manocha, Glmp-realtime pedestrian path prediction using global and local movement patterns, in: 2016 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2016, pp. 5528–5535, https://doi.org/10.1109/ICRA.2016.7487768.
[29] J. Mainprice, R. Hayne, D. Berenson, Goal set inverse optimal control and iterative replanning for predicting human reaching motions in shared workspaces, IEEE Trans. Robot. 32 (4) (2016) 897–908, https://doi.org/10.1109/TRO.2016.2581216.
[30] P. Trautman, A. Krause, Unfreezing the robot: navigation in dense, interacting crowds. 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2010, pp. 797–803. doi:10.1109/IROS.2010.5654369.
[31] J.J. Leonard, H.F. Durrant-Whyte, Application of multi-target tracking to sonar-based mobile robot navigation. 29th IEEE Conference on Decision and Control, IEEE, 1990, pp. 3118–3123. doi:10.1109/CDC.1990.203365.
[32] L. Leal-Taixé, M. Fenzi, A. Kuznetsova, B. Rosenhahn, S. Savarese, Learning an image-based motion context for multiple people tracking, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 3542–3549, https://doi.org/10.1109/CVPR.2014.453.
[33] F. Solera, S. Calderara, R. Cucchiara, Learning to divide and conquer for online multi-target tracking, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 4373–4381, https://doi.org/10.1109/ICCV.2015.497.
[34] S. Pellegrini, A. Ess, K. Schindler, L. Van Gool, You'll never walk alone: modeling social behavior for multi-target tracking, in: 2009 IEEE 12th International Conference on Computer Vision, IEEE, 2009, pp. 261–268, https://doi.org/10.1109/ICCV.2009.5459260.
[35] D. Xie, S. Todorovic, S.-C. Zhu, Inferring "dark matter" and "dark energy" from videos, in: Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 2224–2231, https://doi.org/10.1109/ICCV.2013.277.
[36] R.W. Proctor, K.-P.L. Vu, Human Information processing: an overview for human–computer interaction, in: The Human-Computer Interaction Handbook, 2002, pp. 67–83. ISBN: 1420088815.
[37] S. Jalali, C. Wohlin, Systematic literature studies: database searches vs. backward snowballing, in: The ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, 2012, pp. 29–38, https://doi.org/10.1145/2372251.
[38] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.
[39] T. Zhao, Y. Xu, M. Monfort, W. Choi, C. Baker, Y. Zhao, Y. Wang, Y.N. Wu, Multi-agent tensor fusion for contextual trajectory prediction, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 12126–12134, https://doi.org/10.1109/CVPR.2019.01240.
[40] F. Giannotti, M. Nanni, F. Pinelli, D. Pedreschi, Trajectory pattern mining, in: Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2007, pp. 330–339, https://doi.org/10.1145/1281192.1281230.
[41] D. Helbing, P. Molnar, Social force model for pedestrian dynamics, Phys. Rev. E 51 (5) (1995) 4282, https://doi.org/10.1103/PhysRevE.51.4282.
[42] F. Visin, K. Kastner, K. Cho, M. Matteucci, A. Courville, Y. Bengio, Renet: a recurrent neural network based alternative to convolutional networks, Comput. Sci. 25 (7) (2015) 2983–2996, https://doi.org/10.48550/arXiv.1505.00393.
[43] T. Mikolov, S. Kombrink, L. Burget, J. Černocký, S. Khudanpur, Extensions of recurrent neural network language model, in: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2011, pp. 5528–5531, https://doi.org/10.1109/ICASSP.2011.5947611.
[44] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, ArXiv Preprint (2014), https://doi.org/10.48550/arXiv.1412.3555. ArXiv:1412.3555.
[45] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Adv. Neural Inf. Proces. Syst. 27 (2014) 2672–2680, https://doi.org/10.5555/2969033.2969125.
[46] A. Lerner, Y. Chrysanthou, D. Lischinski, Crowds by example, computer graphics forum, Wiley Online Libr. 26 (3) (2007) 655–664, https://doi.org/10.1111/j.1467-8659.2007.01089.x.
[47] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, S. Savarese, Social Lstm: human trajectory prediction in crowded spaces, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 961–971, https://doi.org/10.1109/CVPR.2016.110.
[48] H. Manh, G. Alaghband, Scene-Lstm: a model for human trajectory prediction, ArXiv Preprint (2018), https://doi.org/10.48550/arXiv.1808.04018. ArXiv:1808.04018.
[49] T.W. Liao, Clustering of time series data—a survey, Pattern Recogn. 38 (11) (2005) 1857–1874, https://doi.org/10.1016/j.patcog.2005.01.025.
[50] G. Yuan, P. Sun, J. Zhao, D. Li, C. Wang, A review of moving object trajectory clustering algorithms, Artif. Intell. Rev. 47 (1) (2017) 123–144, https://doi.org/10.1007/s10462-016-9477-7.
[51] D. Makris, T. Ellis, Learning semantic scene models from observing activity in visual surveillance, in: IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 35(3), 2005, pp. 397–408, https://doi.org/10.1109/TSMCB.2005.846652.
[52] S. Tsumoto, S. Hirano, Behavior grouping based on trajectory mining, in: Social Computing and Behavioral Modeling, Springer, 2009, pp. 1–8, https://doi.org/10.1007/978-1-4419-0056-2_28.
[53] D. Birant, A. Kut, St-Dbscan: an algorithm for clustering spatial–temporal data, Data Knowl. Eng. 60 (1) (2007) 208–221, https://doi.org/10.1016/j.datak.2006.01.013.
[54] G. Yuan, S. Xia, L. Zhang, Y. Zhou, C. Ji, An efficient trajectory-clustering algorithm based on an index tree, Trans. Inst. Meas. Control. 34 (7) (2012) 850–861, https://doi.org/10.1177/0142331211423284.

[55] N. Pelekis, I. Kopanakis, E. Kotsifakos, E. Frentzos, Y. Theodoridis, Clustering trajectories of moving objects in an uncertain world. 2009 Ninth IEEE International Conference on Data Mining, IEEE, 2009, pp. 417–427. doi:10.1109/ICDM.2009.57.
[56] A.T. Palma, V. Bogorny, B. Kuijpers, L.O. Alvares, A clustering-based approach for discovering interesting places in trajectories, in: Proceedings of the 2008 ACM Symposium on Applied Computing, 2008, pp. 863–868, https://doi.org/10.1145/1363686.1363886.
[57] B. Han, L. Liu, E. Omiecinski, Neat: road network aware trajectory clustering. 2012 IEEE 32nd International Conference on Distributed Computing Systems, IEEE, 2012, pp. 142–151. doi:10.1109/ICDCS.2012.31.
[58] Y. Zheng, Trajectory data mining: an overview, ACM Trans. Intellig. Syst. Technol. (TIST) 6 (3) (2015) 1–41, https://doi.org/10.1145/2743025.
[59] C. Parent, S. Spaccapietra, C. Renso, G. Andrienko, N. Andrienko, V. Bogorny, M. L. Damiani, A. Gkoulalas-Divanis, J. Macedo, N. Pelekis, Semantic trajectories modeling and analysis, ACM Comput. Surv. (CSUR) 45 (4) (2013) 1–32, https://doi.org/10.1145/2501654.2501656.
[60] J.-D. Chen, X.-F. Meng, C.-F. Lai, Clustering objects in a road network, J. Software 18 (2) (2007) 332–344, https://doi.org/10.1360/jos180332.
[61] K. Buchin, M. Buchin, J. Gudmundsson, Constrained free space diagrams: a tool for trajectory analysis, Int. J. Geogr. Inf. Sci. 24 (7) (2010) 1101–1125, https://doi.org/10.1080/13658810903569598.
[62] J. Wiest, M. Höffken, U. Kreßel, K. Dietmayer, Probabilistic trajectory prediction with gaussian mixture models. 2012 IEEE Intelligent Vehicles Symposium, IEEE, 2012, pp. 141–146. doi:10.1109/IVS.2012.6232277.
[63] S. Hexeberg, A.L. Flåten, E.F. Brekke, Ais-based vessel trajectory Pprediction. 2017 20th International Conference on Information Fusion (Fusion), IEEE, 2017, pp. 1–8. doi:10.23919/ICIF.2017.8009762.
[64] L. Nanni, S. Ghidoni, S. Brahnam, Handcrafted Vs. Non-Handcrafted Features for Computer Vision Classification, Pattern Recognition 71 (2017) 158–172, https://doi.org/10.1016/j.patcog.2017.05.025.
[65] J. Xu, S. Denman, C. Fookes, S. Sridharan, Unusual scene detection using distributed behaviour model and sparse representation. 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance, IEEE, 2012, pp. 48–53. doi:10.1109/AVSS.2012.80.
[66] R. Mehran, A. Oyama, M. Shah, Abnormal crowd behavior detection using social force model. 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2009, pp. 935–942. doi:10.1109/CVPR.2009.5206641.
[67] S. Pellegrini, A. Ess, L. Van Gool, Improving data association by joint modeling of pedestrian trajectories and groupings, in: European Conference on Computer Vision 6311, Springer, 2010, pp. 452–465, https://doi.org/10.1007/978-3-642-15549-9_33.
[68] L. Leal-Taixé, G. Pons-Moll, B. Rosenhahn, Everybody needs somebody: modeling social and grouping behavior on a linear programming multiple people tracker. 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), IEEE, 2011, pp. 120–127. doi:10.1109/ICCVW.2011.6130233.
[69] A. Treuille, S. Cooper, Z. Popović, Continuum crowds, ACM Trans. Graphics (TOG) 25 (3) (2006) 1160–1168, https://doi.org/10.1145/1141911.1142008.
[70] G. Antonini, M. Bierlaire, M. Weber, Discrete choice models of pedestrian walking behavior, Transp. Res. B Methodol. 40 (8) (2006) 667–687, https://doi.org/10.1016/j.trb.2005.09.006.
[71] X. Wang, X. Ma, W.E.L. Grimson, Unsupervised activity perception in crowded and complicated scenes using hierarchical Bayesian models, IEEE Trans. Pattern Anal. Mach. Intell. 31 (3) (2008) 539–555, https://doi.org/10.1109/TPAMI.2008.87.
[72] P. Flach, Machine Learning: The Art and Science of Algorithms that Make Sense of Data, Cambridge University Press, 2012. ISBN: 1107096391.
[73] A. Alahi, V. Ramanathan, L. Fei-Fei, Socially-aware large-scale crowd forecasting, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2203–2210, https://doi.org/10.1109/CVPR.2014.283.
[74] O. Hosseini, M. Maghrebi, Risk of fire emergency evacuation in complex construction sites: integration of 4d-Bim, social force modeling, and fire quantitative risk assessment, Adv. Eng. Inform. 50 (2021), 101378, https://doi.org/10.1016/j.aei.2021.101378.
[75] S. Yi, H. Li, X. Wang, Understanding pedestrian behaviors from stationary crowd groups, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3488–3496, https://doi.org/10.1109/CVPR.2015.7298971.
[76] M. Rodriguez, J. Sivic, I. Laptev, J.-Y. Audibert, Data-driven crowd analysis in videos. 2011 International Conference on Computer Vision, IEEE, 2011, pp. 1235–1242. doi:10.1109/ICCV.2011.6126374.
[77] X. Wang, K.T. Ma, G.-W. Ng, W.E.L. Grimson, Trajectory analysis and semantic region modeling using nonparametric hierarchical Bayesian models, Int. J. Comput. Vis. 95 (3) (2011) 287–312, https://doi.org/10.1007/s11263-011-0459-6.
[78] T. Hospedales, S. Gong, T. Xiang, A Markov clustering topic model for mining behaviour in video. 2009 IEEE 12th International Conference on Computer Vision, IEEE, 2009, pp. 1165–1172. doi:10.1109/ICCV.2009.5459342.
[79] B. Zhou, X. Tang, X. Wang, Learning collective crowd behaviors with dynamic pedestrian-agents, Int. J. Comput. Vis. 111 (1) (2015) 50–68, https://doi.org/10.1007/s11263-014-0735-3.
[80] T. Fernando, S. Denman, S. Sridharan, C. Fookes, Soft+ hardwired attention: an Lstm framework for human trajectory prediction and abnormal event detection, Neural Netw. 108 (2018) 466–478, https://doi.org/10.1016/j.neunet.2018.09.002.
[81] P. Trautman, J. Ma, R.M. Murray, A. Krause, Robot navigation in dense human crowds: statistical models and experimental studies of human–robot cooperation, Int. J. Robot. Res. 34 (3) (2015) 335–356, https://doi.org/10.1177/0278364914557874.
[82] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang, P. H. Torr, Conditional random fields as recurrent neural networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1529–1537, https://doi.org/10.1109/ICCV.2015.179.
[83] P. Pinheiro, R. Collobert, Recurrent convolutional neural networks for scene labeling, in: International Conference on Machine Learning, PMLR 32(1), 2014, pp. 82–90, https://doi.org/10.5555/3044805.3044816.
[84] A. Graves, Generating sequences with recurrent neural networks, ArXiv Preprint (2013), https://doi.org/10.48550/arXiv.1308.0850. ArXiv:1308.0850.
[85] A. Graves, N. Jaitly, Towards end-to-end speech recognition with recurrent neural networks, Int. Conf. Machine Learn. PMLR 32 (2) (2014) 1764–1772, https://doi.org/10.5555/3044805.3045089.
[86] F. Bartoli, G. Lisanti, L. Ballan, A. Del Bimbo, Context-aware trajectory prediction. 2018 24th International Conference on Pattern Recognition (ICPR), IEEE, 2018, pp. 1941–1946. doi:10.1109/ICPR.2018.8545447.
[87] R. Hug, S. Becker, W. Hübner, M. Arens, Particle-based pedestrian path prediction using lstm-mdl models. 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE, 2018, pp. 2684–2691. doi:10.1109/ITSC.2018.8569478.
[88] J. Liang, L. Jiang, J.C. Niebles, A.G. Hauptmann, L. Fei-Fei, Peeking into the future: predicting future person activities and locations in videos, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 5725–5734, https://doi.org/10.1109/CVPRW.2019.00358.
[89] P. Zhang, W. Ouyang, P. Zhang, J. Xue, N. Zheng, Sr-Lstm: state refinement for Lstm towards pedestrian trajectory prediction, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 12085–12094, https://doi.org/10.1109/CVPR.2019.01236.
[90] N. Jaipuria, G. Habibi, J.P. How, A transferable pedestrian motion prediction model for intersections with different geometries, ArXiv Preprint (2018), https://doi.org/10.48550/arXiv.1806.09444. ArXiv:1806.09444.
[91] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, A. Alahi, Social Gan: socially acceptable trajectories with generative adversarial networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 2255–2264, https://doi.org/10.1109/CVPR.2018.00240.
[92] A. Vemula, K. Muelling, J. Oh, Social attention: modeling attention in human crowds, in: 2018 IEEE international Conference on Robotics and Automation (ICRA), IEEE, 2018, pp. 1–7, https://doi.org/10.1109/ICRA.2018.8460504.
[93] Y. Huang, H. Bi, Z. Li, T. Mao, Z. Wang, Stgat: modeling spatial-temporal interactions for human trajectory prediction, in: Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 6272–6281, https://doi.org/10.1109/ICCV.2019.00637.
[94] S. Yi, H. Li, X. Wang, Pedestrian behavior understanding and prediction with deep neural networks, in: European Conference on Computer Vision 9905, Springer, 2016, pp. 263–279, https://doi.org/10.1007/978-3-319-46448-0_16.
[95] Y. Xu, Z. Piao, S. Gao, Encoding crowd interaction with deep neural network for pedestrian trajectory prediction, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 5275–5284, https://doi.org/10.1109/CVPR.2018.00553.
[96] J. Cai, Y. Zhang, L. Yang, H. Cai, S. Li, A context-augmented deep learning approach for worker trajectory prediction on unstructured and dynamic construction sites, Adv. Eng. Inform. 46 (2020), 101173, https://doi.org/10.1016/j.aei.2020.101173.
[97] K.M. Rashid, A.H. Behzadan, Risk behavior-based trajectory prediction for construction site safety monitoring, J. Constr. Eng. Manag. 144 (2) (2018) 04017106, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001420.
[98] Z. Zhu, M.-W. Park, C. Koch, M. Soltani, A. Hammad, K. Davari, Predicting movements of onsite workers and mobile equipment for enhancing construction site safety, Autom. Constr. 68 (2016) 95–101, https://doi.org/10.1016/j.autcon.2016.04.009.
[99] T. Fernando, S. Denman, A. McFadyen, S. Sridharan, C. Fookes, Tree memory networks for modelling long-term temporal dependencies, Neurocomputing. 304 (2018) 64–81, https://doi.org/10.1016/j.neucom.2018.03.040.
[100] H. Zou, H. Su, S. Song, J. Zhu, Understanding human behaviors in crowds by imitating the decision-making process, ArXiv Preprint (2018), https://doi.org/10.48550/arXiv.1801.08391. ArXiv:1801.08391.
[101] J. Liu, A. Shahroudy, D. Xu, G. Wang, Spatio-temporal Lstm with trust gates for 3d human action recognition, in: European Conference on Computer Vision 9907, Springer, 2016, pp. 816–833, https://doi.org/10.1007/978-3-319-46487-9_50.
[102] Y.F. Chen, M. Everett, M. Liu, J.P. How, Socially aware motion planning with deep reinforcement learning, in: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2017, pp. 1343–1350, https://doi.org/10.1109/IROS.2017.8202312.
[103] K.M. Kitani, B.D. Ziebart, J.A. Bagnell, M. Hebert, Activity forecasting, in: European Conference on Computer Vision 7575, Springer, 2012, pp. 201–214, https://doi.org/10.1007/978-3-642-33765-9_15.
[104] Z. Yu, D.S. Moirangthem, M. Lee, Continuous timescale long-short term memory neural network for human intent understanding, Front. Neurorobot. 11 (2017) 42, https://doi.org/10.3389/fnbot.2017.00042.
[105] A. Robicquet, A. Sadeghian, A. Alahi, S. Savarese, Learning social etiquette: human trajectory understanding in crowded scenes, in: European Conference on Computer Vision 9912, Springer, 2016, pp. 549–565, https://doi.org/10.1007/978-3-319-46484-8_33.

[106] D.P. Kingma, M. Welling, Auto-encoding variational bayes, ArXiv Preprint (2013), https://doi.org/10.48550/arXiv.1312.6114. ArXiv:1312.6114.
[107] J. Li, H. Ma, M. Tomizuka, Conditional generative neural system for probabilistic trajectory Prediction, ArXiv Preprint (2019), https://doi.org/10.48550/arXiv.1905.01631. ArXiv:1905.01631.
[108] J. Amirian, J.-B. Hayet, J. Pettré, Social ways: learning multi-modal distributions of pedestrian trajectories with Gans, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019, p. 0, https://doi.org/10.1109/CVPRW.2019.00359.
[109] A. Sadeghian, V. Kosaraju, A. Sadeghian, N. Hirose, H. Rezatofighi, S. Savarese, Sophie: an attentive Gan for predicting paths compliant to social and physical constraints, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 1349–1358, https://doi.org/10.1109/CVPR.2019.00144.
[110] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, Adv. Neural Inf. Proces. Syst. 30 (2017) 5998–6008, https://doi.org/10.5555/3295222.3295349.
[111] A. Jain, A.R. Zamir, S. Savarese, A. Saxena, Structural-Rnn: deep learning on spatio-temporal graphs, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 5308–5317, https://doi.org/10.1109/CVPR.2016.573.
[112] H. Luo, M. Wang, P.K.-Y. Wong, J. Tang, J.C. Cheng, Construction machine pose prediction considering historical motions and activity attributes using gated recurrent unit (Gru), Autom. Constr. 121 (2021), 103444, https://doi.org/10.1016/j.autcon.2020.103444.
[113] D. Kim, M. Liu, S. Lee, V.R. Kamat, Trajectory prediction of mobile construction resources toward pro-active struck-by hazard detection, in: Proceedings of the International Symposium on Automation and Robotics in Construction (IAARC), 2019, https://doi.org/10.22260/ISARC2019/0131.
[114] V. Kosaraju, A. Sadeghian, R. Martín-Martín, I. Reid, H. Rezatofighi, S. Savarese, Social-Bigat: multimodal trajectory forecasting using bicycle-Gan and graph attention networks, Adv. Neural Inf. Proces. Syst. 32 (2019) 137–146, https://doi.org/10.5555/3454287.3454300.
[115] A. Mohamed, K. Qian, M. Elhoseiny, C. Claudel, Social-Stgcnn: a social spatio-temporal graph convolutional neural network for human trajectory prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 14424–14432, https://doi.org/10.1109/CVPR42600.2020.01443.
[116] L. Zhang, Q. She, P. Guo, Stochastic trajectory prediction with social graph network, ArXiv Preprint (2019), https://doi.org/10.48550/arXiv.1907.10233. ArXiv:1907.10233.
[117] X. Zhang, D.B. Chaffin, Digital human modeling for computer-aided ergonomics, in: Handbook of Occupational Ergonomics, 2005, pp. 1–20, https://doi.org/10.1201/9781420003642-14.
[118] S.H. Hsiang, M.M. Ayoub, Development of methodology in biomechanical simulation of manual lifting, Int. J. Ind. Ergon. 13 (4) (1994) 271–288, https://doi.org/10.1016/0169-8141(94)90085-X.
[119] H.O. Demirel, S. Ahmed, V.G. Duffy, Digital human modeling: a review and reappraisal of origins, present, and expected future methods for representing humans computationally, Int. J. Hum. Comp. Interact. 38 (10) (2021) 1–41, https://doi.org/10.1080/10447318.2021.1976507.
[120] A. Wolf, J. Miehling, S. Wartzack, Challenges in interaction modelling with digital human models - a systematic literature review of interaction modelling approaches, Ergonomics 63 (11) (2020) 1442–1458, https://doi.org/10.1080/00140139.2020.1786606.
[121] D.B. Chaffin, J.J. Faraway, X. Zhang, C. Woolley, Stature, age, and gender effects on reach motion postures, Hum. Factors 42 (3) (2000) 408–420, https://doi.org/10.1518/001872000779698222.
[122] J. Faraway, Human animation using nonparametric regression, J. Comput. Graph. Stat. 13 (3) (2004) 537–553, https://doi.org/10.1198/106186004x2507.
[123] J. Faraway, M.P. Reed, Statistics for digital human motion modeling in ergonomics, Technometrics 49 (3), 49(3) (2007) 277–290, https://doi.org/10.1198/004017007000000281.
[124] D. Mavrikios, V. Karabatsou, M. Pappas, G. Chryssolouris, An efficient approach to human motion modeling for the verification of human-centric product design and manufacturing in virtual environments, Robot. Comput. Integr. Manuf. 23 (5) (2007) 533–543, https://doi.org/10.1016/j.rcim.2006.05.009.
[125] D.W. Wagner, M.P. Reed, D.B. Chaffin, The development of a model to predict the effects of worker and task factors on foot placements in manual material handling tasks, Ergonomics 53 (11) (2010) 1368–1384, https://doi.org/10.1080/00140139.2010.523482.
[126] W. Park, D.B. Chaffin, B.J. Martin, Toward memory-based human motion simulation: development and validation of a motion modification algorithm, IEEE Trans. Syst. Man Cybern. Syst. Hum. 34 (3) (2004) 376–386, https://doi.org/10.1109/tsmca.2003.822965.
[127] P. Woojin, D.B. Chaffin, B.J. Martin, Y. Jonghwa, Memory-based human motion simulation for computer-aided ergonomic design, IEEE Trans. Syst. Man Cybern. Syst. Hum. 38 (3) (2008) 513–527, https://doi.org/10.1109/tsmca.2008.918588.
[128] E.S. Jung, D. Kee, M.K. Chung, Upper body reach posture prediction for ergonomic evaluation models, Int. J. Ind. Ergon. 16 (2) (1995) 95–107, https://doi.org/10.1016/0169-8141(94)00088-K.
[129] X. Wang, J.P. Verriest, A geometric algorithm to predict the arm reach posture for computer-aided ergonomic evaluation, J. Vis. Comput. Animat. 9 (1) (1998) 33–47, https://doi.org/10.1002/(SICI)1099-1778(199801/03)9:1%3C33::AID-VIS174%3E3.0.CO;2-Q.
[130] C.J. Lin, M.M. Ayoub, T.M. Bernard, Computer motion simulation for sagittal plane lifting activities, Int. J. Ind. Ergon. 24 (2) (1999) 141–155, https://doi.org/10.1016/S0169-8141(98)00010-9.
[131] S. Hauberg, K.S. Pedersen, Predicting articulated human motion from spatial processes, Int. J. Comput. Vis. 94 (3) (2011) 317–334, https://doi.org/10.1007/s11263-011-0433-3.
[132] X. Zhang, A.D. Kuo, D.B. Chaffin, Optimization-based differential kinematic modeling exhibits a velocity-control strategy for dynamic posture determination in seated reaching movements, J. Biomech. 31 (11) (1998) 1035–1042, https://doi.org/10.1016/S0021-9290(98)00117-1.
[133] X. Zhang, D. Chaffin, A three-dimensional dynamic posture prediction model for simulating in-vehicle seated reaching movements: development and validation, Ergonomics 43 (9) (2000) 1314–1330, https://doi.org/10.1080/001401300421761.
[134] H.J. Kim, Q. Wang, S. Rahmatalla, C.C. Swan, J.S. Arora, K. Abdel-Malek, J. G. Assouline, Dynamic motion planning of 3d human locomotion using gradient-based optimization, J. Biomech. Eng. 130 (3) (2008), 031002, https://doi.org/10.1115/1.2898730.
[135] Q. Zou, Q. Zhang, J. Yang, A. Cloutier, E. Pena-Pitarch, Nonlinear inverse optimization approach for determining the weights of objective function in standing reach tasks, Comput. Ind. Eng. 63 (4) (2012) 791–801, https://doi.org/10.1016/j.cie.2012.05.001.
[136] Y. Xiang, J.S. Arora, K. Abdel-Malek, Hybrid predictive dynamics: a new approach to simulate human motion, Multibody Syst. Dyn. 28 (3) (2012) 199–224, https://doi.org/10.1007/s11044-012-9306-y.
[137] R.T. Marler, J.S. Arora, J. Yang, H.-J. Kim, K. Abdel-Malek, Use of multi-objective optimization for digital human posture prediction, Eng. Optim. 41 (10) (2009) 925–943, https://doi.org/10.1080/03052150902853013.
[138] J. Yang, T. Marler, S. Rahmatalla, Multi-objective optimization-based method for kinematic posture prediction: development and validation, Robotica 29 (2) (2010) 245–253, https://doi.org/10.1017/s026357471000010x.
[139] Y. Xiang, J.S. Arora, S. Rahmatalla, T. Marler, R. Bhatt, K. Abdel-Malek, Human lifting simulation using a multi-objective optimization approach, Multibody Syst. Dyn. 23 (4) (2010) 431–451, https://doi.org/10.1007/s11044-009-9186-y.
[140] Y. Xiang, J.S. Arora, S. Rahmatalla, K. Abdel-Malek, Optimization-based dynamic human walking prediction: one step formulation, Int. J. Numer. Methods Eng. 79 (6) (2009) 667–695, https://doi.org/10.1002/nme.2575.
[141] K. Abdel-Malek, Z. Mi, J. Yang, K. Nebel, Optimization-based trajectory planning of the human upper body, Robotica 24 (6) (2006) 683–696, https://doi.org/10.1017/s0263574706002852.
[142] S.D. Farahani, M.S. Andersen, M. de Zee, J. Rasmussen, Optimization-based dynamic prediction of kinematic and kinetic patterns for a human vertical jump from a squatting position, Multibody Syst. Dyn. 36 (1) (2015) 37–65, https://doi.org/10.1007/s11044-015-9468-5.
[143] Y. Xiang, H.-J. Chung, J.H. Kim, R. Bhatt, S. Rahmatalla, J. Yang, T. Marler, J. S. Arora, K. Abdel-Malek, Predictive dynamics: an optimization-based novel approach for human motion simulation, Struct. Multidiscip. Optim. 41 (3) (2009) 465–479, https://doi.org/10.1007/s00158-009-0423-z.
[144] M. Svinin, I. Goncharenko, Z.W. Luo, S. Hosoe, Reaching movements in dynamic environments: how do we move flexible objects?. Proceedings of the 2005 IEEE International Conference on Robotics and Automation vol.22 IEEE, 2005, pp. 724–739.
[145] A. Biess, M. Nagurka, T. Flash, Simulating discrete and rhythmic multi-joint human arm movements by optimization of nonlinear performance indices, Biol. Cybern. 95 (1) (2006) 31–53, https://doi.org/10.1007/s00422-006-0067-7.
[146] M. Svinin, I. Goncharenko, V. Kryssanov, E. Magid, Motion planning strategies in human control of non-rigid objects with internal degrees of freedom, Hum. Mov. Sci. 63 (2019) 209–230, https://doi.org/10.1016/j.humov.2018.12.004.
[147] D.B. Chaffin, Improving digital human modelling for proactive ergonomics in design, Ergonomics 48 (5) (2005) 478–491, https://doi.org/10.1080/00140130400029191.
[148] D.B. Chaffin, Human motion simulation for vehicle and workplace design, Hum. Factors Ergonom. Manuf. 17 (5) (2007) 475–484, https://doi.org/10.1002/hfm.20087.
[149] C.M. Huang, S. Andrist, A. Sauppe, B. Mutlu, Using gaze patterns to predict task intent in collaboration, Front. Psychol. 6 (2015) 1049, https://doi.org/10.3389/fpsyg.2015.01049.
[150] H. Kadu, C.C.J. Kuo, Automatic human mocap data classification, IEEE Trans. Multimedia 16 (8) (2014) 2191–2202, https://doi.org/10.1109/TMM.2014.2360793.
[151] T. Gandhi, M.M. Trivedi, Image based estimation of pedestrian orientation for improving path prediction, in: 2008 IEEE Intelligent Vehicles Symposium, 2008, pp. 506–511, https://doi.org/10.1109/IVS.2008.4621257.
[152] S. Kohler, M. Goldhammer, S. Bauer, K. Doll, U. Runsmann, K. Dietmayer, Early detection of the pedestrian's intention to cross the street. 2012 15th International IEEE Conference on Intelligent Transportation Systems, IEEE, 2012, pp. 1759–1764.
[153] S. Kohler, B. Schreiner, S. Ronalter, K. Doll, U. Brunsmann, K. Zindler, Autonomous evasive maneuvers triggered by infrastructure-based detection of pedestrian intentions, in: 2013 IEEE Intelligent Vehicles Symposium (IV), 2013, pp. 519–526, https://doi.org/10.1109/IVS.2013.6629520.
[154] S. Koehler, M. Goldhammer, S. Bauer, S. Zecha, K. Doll, U. Brunsmann, K. Dietmayer, Stationary detection of the pedestrian? S intention at intersections, IEEE Intell. Transp. Syst. Mag. 5 (4) (2013) 87–99, https://doi.org/10.1109/MITS.2013.2276939.

[155] S. Kohler, M. Goldhammer, K. Zindler, K. Doll, K. Dietmeyer, Stereo-vision-based pedestrian's intention detection in a moving vehicle, in: 2015 IEEE 18th International Conference on Intelligent Transportation Systems, 2015, pp. 2317–2322, https://doi.org/10.1109/ITSC.2015.374.
[156] Z. Fang, D. Vazquez, A.M. Lopez, On-board detection of pedestrian intentions, Sensors 17 (10) (2017) 2193, https://doi.org/10.3390/s17102193.
[157] R. Quintero, I. Parra, J. Lorenzo, D. Fernandez-Llorca, M.A. Sotelo, Pedestrian intention recognition by means of a Hidden Markov model, in: 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 2017, pp. 1–7, https://doi.org/10.1109/ITSC.2017.8317766.
[158] J. Chen, J. Qiu, C. Ahn, Construction worker's awkward posture recognition through supervised motion tensor decomposition, Autom. Constr. 77 (2017) 67–81, https://doi.org/10.1016/j.autcon.2017.01.020.
[159] C.G. Keller, C. Hermes, D.M. Gavrila, Will the pedestrian cross probabilistic path prediction based on learned motion features, in: Joint Pattern Recognition Symposium 6835, Springer, 2011, pp. 386–395, https://doi.org/10.1007/978-3-642-23123-0_39.
[160] Y. Zhou, M. Chen, G. Du, P. Zhang, X. Liu, Intelligent grasping with natural human-robot interaction, Industr. Robot 45 (1) (2018) 44–53, https://doi.org/10.1108/ir-05-2017-0089.
[161] C.G. Keller, D.M. Gavrila, Will the pedestrian cross? A study on pedestrian path prediction, IEEE Trans. Intell. Transp. Syst. 15 (2) (2014) 494–506, https://doi.org/10.1109/TITS.2013.2280766.
[162] R. Quintero, J. Almeida, D.F. Llorca, M.A. Sotelo, Pedestrian path prediction using body language traits, in: 2014 IEEE Intelligent Vehicles Symposium Proceedings, 2014, pp. 317–323, https://doi.org/10.1109/ivs.2014.6856498.
[163] R. Quintero, I. Parra, D.F. Llorca, M.A. Sotelo, Pedestrian path prediction based on body language and action classification, in: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2014, pp. 679–684, https://doi.org/10.1109/itsc.2014.6957768.
[164] R. Quintero, I. Parra, D.F. Llorca, M.A. Sotelo, Pedestrian intention and pose prediction through dynamical models and behaviour classification, in: 2015 IEEE 18th International Conference on Intelligent Transportation Systems, 2015, pp. 83–88, https://doi.org/10.1109/itsc.2015.22.
[165] R.Q. Mínguez, I.P. Alonso, D. Fernández-Llorca, M.Á. Sotelo, Pedestrian path, pose, and intention prediction through Gaussian process dynamical models and pedestrian activity recognition, IEEE Trans. Intell. Transp. Syst. 20 (5) (2019) 1803–1814, https://doi.org/10.1109/tits.2018.2836305.
[166] B. Krüger, A. Vögele, T. Willig, A. Yao, R. Klein, A. Weber, Efficient unsupervised temporal segmentation of motion data, IEEE Trans. Multimedia 19 (4) (2016) 797–812, https://doi.org/10.1109/TMM.2016.2635030.
[167] T. Matsubara, S.-H. Hyon, J. Morimoto, Real-time stylistic prediction for whole-body human motions, Neural Netw. 25 (1) (2012) 191–199, https://doi.org/10.1016/j.neunet.2011.08.008.
[168] M. Goldhammer, M. Gerhard, S. Zernetsch, K. Doll, U. Brunsmann, Early prediction of a pedestrian's trajectory at intersections, in: Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems, 2013, pp. 237–242, https://doi.org/10.1109/ITSC.2013.6728239.
[169] Y. Huang, J. Cui, F. Davoine, H. Zhao, H. Zha, Head pose based intention prediction using discrete dynamic Bayesian network, in: 2013 Seventh International Conference on Distributed Smart Cameras (ICDSC), 2013, pp. 1–6, https://doi.org/10.1109/ICDSC.2013.6778228.
[170] J. Kooij, P. Francisco, N. Schneider, F. Flohr, D.M. Gavrila, Context-based pedestrian path prediction, in: European Conference on Computer Vision 8694, Springer, 2014, pp. 618–633, https://doi.org/10.1007/978-3-319-10599-4_40.
[171] A.T. Schulz, R. Stiefelhagen, A controlled interactive multiple model filter for combined pedestrian intention recognition and path prediction. 2015 IEEE 18th International Conference on Intelligent Transportation Systems, Gran Canaria, Spain, IEEE, 2015, pp. 173–178.
[172] B. Zhang, I. Horváth, J.F.M. Molenbroek, C. Snijders, Using artificial neural networks for human body posture prediction, Int. J. Ind. Ergon. 40 (4) (2010) 414–424, https://doi.org/10.1016/j.ergon.2010.02.003.
[173] Y. Du, W. Wang, L. Wang, Hierarchical recurrent neural network for skeleton based action recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Boston, USA, 2015, pp. 1110–1118.
[174] Y. Li, C. Lan, J. Xing, W. Zeng, C. Yuan, J. Liu, Online human action detection using joint classification-regression recurrent neural network, in: European Conference on Computer Vision 9911, Springer, Cham, 2016, https://doi.org/10.1007/978-3-319-46478-7_13.
[175] P. Ghosh, J. Song, E. Aksan, O. Hilliges, Learning human motion models for long-term predictions, in: 2017 International Conference on 3D Vision, 2017, pp. 458–466, https://doi.org/10.1109/3DV.2017.00059.
[176] J. Liu, G. Wang, L.-Y. Duan, P. Hu, A.C. Kot, Skeleton-based human action recognition with global context-aware attention Lstm networks, IEEE Trans. Image Process. 27 (4) (2017) 1586–1599, https://doi.org/10.1109/TIP.2017.2785279.
[177] Y. Kim, E. Soo Choi, J. Seo, W.S. Choi, J. Lee, K. Lee, A novel approach to predicting human ingress motion using an artificial neural network, J. Biomech. 84 (2019) 27–35, https://doi.org/10.1016/j.jbiomech.2018.12.009.
[178] K. Fragkiadaki, S. Levine, P. Felsen, J. Malik, Recurrent network models for human dynamics. 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, IEEE, 2015, pp. 4346–4354.
[179] Y. Li, Z. Wang, X. Yang, M. Wang, S.I. Poiana, E. Chaudhry, J. Zhang, Efficient convolutional hierarchical autoencoder for human motion prediction, Vis. Comput. 35 (6–8) (2019) 1143–1156, https://doi.org/10.1007/s00371-019-01692-9.
[180] M. Goldhammer, K. Doll, A. Gensler, B. Sick, Pedestrian's trajectory forecast in public traffic with artificial neural networks. 2014 22nd International Conference on Pattern Recognition, IEEE, Stockholm, Sweden, 2014, pp. 4110–4115. doi:10.1109/ICPR.2014.704.
[181] M. Goldhammer, S. Köhler, K. Doll, B. Sick, Camera based pedestrian path prediction by means of polynomial least-squares approximation and multilayer perceptron neural networks, SAI intelligent systems conference 2015, IEEE (2015) 390–399, https://doi.org/10.1109/IntelliSys.2015.7361171.
[182] J.-Y. Kwak, E.-J. Lee, B. Ko, M. Jeong, Pedestrian's intention prediction based on fuzzy finite automata and spatial-temporal features, Electron. Imaging 2016 (3) (2016) 1–6, https://doi.org/10.2352/ISSN.2470-1173.2016.3.VSTIA-512.
[183] J.-Y. Kwak, B.C. Ko, J.-Y. Nam, Pedestrian intention prediction based on dynamic fuzzy automata for vehicle driving at nighttime, Infrared Phys. Technol. 81 (2017) 41–51, https://doi.org/10.1016/j.infrared.2016.12.014.
[184] M. Simao, N. Mendes, O. Gibaru, P. Neto, A review on electromyography decoding and pattern recognition for human-machine interaction, IEEE Access 7 (2019) 39564–39582, https://doi.org/10.1109/access.2019.2906584.
[185] K. Englehart, B. Hudgins, A. Robust, Real-time control scheme for multifunction myoelectric control, IEEE Trans. Biomed. Eng. 50 (7) (2003) 848–854, https://doi.org/10.1109/TBME.2003.813539.
[186] P.K. Artemiadis, K.J. Kyriakopoulos, Estimating arm motion and force using Emg signals: on the control of exoskeletons, in: 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008, pp. 279–284, https://doi.org/10.1109/IROS.2008.4650949.
[187] J. González, Y. Horiuchi, W. Yu, Classification of upper limb motions from around-shoulder muscle activities hand biofeedback, Open Med Inform J 4 (2010) 74–80, https://doi.org/10.2174/1874431101004020074.
[188] O. Fukuda, T. Tsuji, M. Kaneko, A. Otsuka, A human-assisting manipulator teleoperated by Emg signals and arm motions, IEEE Trans. Robot. Autom. 19 (2) (2003) 210–222, https://doi.org/10.1109/TRA.2003.808873.
[189] S. Zhou, K. Yin, F. Fei, K. Zhang, Surface electromyography–based hand movement recognition using the Gaussian mixture model, multilayer perceptron, and Adaboost method, Int. J. Distributed Sensor Netw. 15 (4) (2019) 1–13, https://doi.org/10.1177/1550147719846060.
[190] C. Choi, S. Micera, J. Carpaneto, J. Kim, Development and quantitative performance evaluation of a noninvasive Emg computer interface, IEEE Trans. Biomed. Eng. 56 (1) (2009) 188–191, https://doi.org/10.1109/TBME.2008.2005950.
[191] N. Jiang, S. Muceli, B. Graimann, D. Farina, Effect of arm position on the prediction of kinematics from Emg in amputees, Med. Biol. Eng. Comput. 51 (1–2) (2013) 143–151, https://doi.org/10.1007/s11517-012-0979-4.
[192] K. Kiguchi, Y. Hayashi, Task estimation of upper-limb using Eeg and Emg signals, in: 2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Besançon, France, 2014, pp. 548–553, https://doi.org/10.1109/AIM.2014.6878135.
[193] C. Calderon-Cordova, C. Ramirez, V. Barros, P.A. Quezada-Sarmiento, L. Barba-Guaman, Emg signal patterns recognition based on feedforward artificial neural network applied to robotic prosthesis myoelectric control, in: 2016 Future Technologies Conference (FTC), 2016, pp. 868–875, https://doi.org/10.1109/FTC.2016.7821705.
[194] B. Nan, M. Okamoto, T. Tsuji, A hybrid motion classification approach for Emg-based human–robot interfaces using Bayesian and neural networks, IEEE Trans. Robot. 25 (3) (2009) 502–511, https://doi.org/10.1109/tro.2009.2019782.
[195] M. Khezri, M. Jahed, A neuro–fuzzy inference system for semg-based identification of hand motion commands, IEEE Trans. Ind. Electron. 58 (5) (2011) 1952–1960, https://doi.org/10.1109/TIE.2010.2053334.
[196] I. Batzianoulis, S. El-Khoury, E. Pirondini, M. Coscia, S. Micera, A. Billard, Emg-based decoding of grasp gestures in reaching-to-grasping motions, Robot. Auton. Syst. 91 (2017) 59–70, https://doi.org/10.1016/j.robot.2016.12.014.
[197] F. Duan, L. Dai, W. Chang, Z. Chen, C. Zhu, W. Li, Semg-based identification of hand motion commands using wavelet neural network combined with discrete wavelet transform, IEEE Trans. Ind. Electron. 63 (3) (2016) 1923–1934, https://doi.org/10.1109/TIE.2015.2497212.
[198] S. Kwon, J. Kim, Real-time upper limb motion estimation from surface electromyography and joint angular velocities using an artificial neural network for human-machine cooperation, IEEE Trans. Inf. Technol. Biomed. 15 (4) (2011) 522–530, https://doi.org/10.1109/TITB.2011.2151869.
[199] Y.M. Aung, A. Al-Jumaily, Semg based Ann for shoulder angle prediction, Proc. Eng. 41 (2012) 1009–1015, https://doi.org/10.1016/j.proeng.2012.07.276.
[200] R. Raj, R. Rejith, K.S. Sivanandan, Real time identification of human forearm kinematics from surface Emg signal using artificial neural network models, Proc. Technol. 25 (2016) 44–51, https://doi.org/10.1016/j.protcy.2016.08.079.
[201] Q. Zhang, R. Liu, W. Chen, C. Xiong, Simultaneous and continuous estimation of shoulder and elbow kinematics from surface Emg signals, Front. Neurosci. 11 (2017) 280, https://doi.org/10.3389/fnins.2017.00280.
[202] F. Zhang, P. Li, Z.-G. Hou, Z. Lu, Y. Chen, Q. Li, M. Tan, Semg-based continuous estimation of joint angles of human legs by using Bp neural network, Neurocomputing 78 (1) (2012) 139–148, https://doi.org/10.1016/j.neucom.2011.05.033.
[203] Y.M. Aung, A. Al-Jumaily, Estimation of upper limb joint angle using surface Emg signal, Int. J. Adv. Robot. Syst. 10 (10) (2013) 1–8, https://doi.org/10.5772/56717.
[204] Q. Zhang, R. Hosoda, G. Venture, Human joint motion estimation for electromyography (Emg)-based dynamic motion control. 35th Annual International Conference of the IEEE EMBS, Osaka, Japan, IEEE, 2013, pp. 21–24.

[205] L. Tong, F. Zhang, Z.-G. Hou, W. Wang, L. Peng, Bp–Ar-based human joint angle estimation using multi-channel Semg, Int. J. Robot. Autom. 30 (3) (2015) 227–237, https://doi.org/10.2316/Journal.206.2015.3.206-4080.

[206] C. Loconsole, S. Dettori, A. Frisoli, C.A. Avizzano, M. Bergamasco, An Emg-based approach for on-line predicted torque control in robotic-assisted rehabilitation. IEEE Haptics Symposium 2014, Houston, TX, USA, IEEE, 2014, pp. 181–186.

[207] A. Akhtar, N. Aghasadeghi, L. Hargrove, T. Bretl, Estimation of distal arm joint angles from Emg and shoulder orientation for transhumeral prostheses, J. Electromyogr. Kinesiol. 35 (2017) 86–94, https://doi.org/10.1016/j.jelekin.2017.06.001.

[208] Q. Ding, J. Han, X. Zhao, Continuous estimation of human multi-joint angles from Semg using a state-space model, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (9) (2017) 1518–1528, https://doi.org/10.1109/TNSRE.2016.2639527.

[209] J. Han, Q. Ding, A. Xiong, X. Zhao, A state-space Emg model for the estimation of continuous joint movements, IEEE Trans. Ind. Electron. 62 (7) (2015) 4267–4275, https://doi.org/10.1109/TIE.2014.2387337.

[210] Y. Chen, X. Zhao, J. Han, Hierarchical projection regression for online estimation of elbow joint angle using Emg signals, Neural Comput. & Applic. 23 (3–4) (2012) 1129–1138, https://doi.org/10.1007/s00521-012-1045-8.

[211] Z. Li, D. Zhang, X. Zhao, F. Wang, B. Zhang, D. Ye, J. Han, A temporally smoothed Mlp regression scheme for continuous knee/ankle angles estimation by using multi-channel Semg, IEEE Access 8 (2020) 47433–47444, https://doi.org/10.1109/ACCESS.2020.2979008.

[212] S.M. Shafiul Hasan, M.R. Siddiquee, R. Atri, R. Ramon, J.S. Marquez, O. Bai, Prediction of gait intention from pre-movement Eeg signals: a feasibility study, J. Neuro Eng. Rehabil. 17 (1) (2020) 50, https://doi.org/10.1186/s12984-020-00675-5.

[213] O. Bai, P. Lin, S. Vorbach, J. Li, S. Furlani, M. Hallett, Exploration of computational methods for classification of movement intention during human voluntary movement from single trial Eeg, Clin. Neurophysiol. 118 (12) (2007) 2637–2655, https://doi.org/10.1016/j.clinph.2007.08.025.

[214] D. Huang, P. Lin, D.Y. Fei, X. Chen, O. Bai, Decoding human motor activity from Eeg single trials for a discrete two-dimensional cursor control, J. Neural Eng. 6 (4) (2009) 46–57, https://doi.org/10.1088/1741-2560/6/4/046005.

[215] A.I. Sburlea, L. Montesano, R. Cano de la Cuerda, I.M. Alguacil Diego, J. C. Miangolarra-Page, J. Minguez, Detecting intention to walk in stroke patients from pre-movement Eeg correlates, J. Neuro Eng. Rehabil. 12 (1) (2015) 1–12, https://doi.org/10.1186/s12984-015-0087-4.

[216] O. Bai, V. Rathi, P. Lin, D. Huang, H. Battapady, D.-Y. Fei, L. Schneider, E. Houdayer, X. Chen, MarkHallett, Prediction of human voluntary movement before it occurs, Clin. Neurophysiol. 122 (2) (2011) 364–372, https://doi.org/10.1016/j.clinph.2010.07.010.

[217] A. Turnwald, D. Althoff, D. Wollherr, M. Buss, Understanding human avoidance behavior: interaction-aware decision making based on game theory, Int. J. Soc. Robot. 8 (2) (2016) 331–351, https://doi.org/10.1007/s12369-016-0342-2.

[218] J. Gehring, M. Auli, D. Grangier, D. Yarats, Y.N. Dauphin, Convolutional sequence to sequence learning, ArXiv Preprint 70 (2017) 1243–1252, https://doi.org/10.48550/arXiv.1705.03122.

[219] S. Bai, J.Z. Kolter, V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, ArXiv Preprint 1 (2018) 1–14, https://doi.org/10.48550/arXiv.1803.01271.

[220] N. Nikhil, B. Tran Morris, Convolutional neural network for trajectory prediction, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 1–11, https://doi.org/10.1007/978-3-030-11015-4_16.

[221] T. Zhou, Y. Wang, Q. Zhu, J. Du, Human hand motion prediction based on feature grouping and deep learning: pipe skid maintenance example, Autom. Constr. 138 (2022), 104232, https://doi.org/10.1016/j.autcon.2022.104232.

[222] T. Zhou, Q. Zhu, Y. Shi, J. Du, Construction robot teleoperation safeguard based on real-time human hand motion prediction, J. Constr. Eng. Manag. 148 (7) (2022) 04022040, https://doi.org/10.1061/(ASCE)CO.1943-7862.0002289.

[223] V. Murino, M. Cristani, S. Shah, S. Savarese, Group and crowd behavior for computer vision, Academic Press, 2017. ISBN: 978-0-12-809276-7.

[224] D.C. Niehorster, R.S. Hessels, J.S. Benjamins, Glassesviewer: open-source software for viewing and analyzing data from the Tobii pro glasses 2 eye tracker, Behav. Res. Methods 52 (2020) 1244–1253, https://doi.org/10.3758/s13428-019-01314-1.

[225] M. Cognolato, M. Atzori, H. Müller, Head-mounted eye gaze tracking devices: an overview of modern devices and recent advances, J. Rehabil. Assist. Technol. Eng. 5 (2018) 1–13, https://doi.org/10.1177/2055668318773991.

[226] M.M. Morrow, B. Lowndes, E. Fortune, K.R. Kaufman, M.S. Hallbeck, Validation of inertial measurement units for upper body kinematics, J. Appl. Biomech. 33 (3) (2017) 227–232, https://doi.org/10.1123/jab.2016-0120.

[227] D. Thewlis, C. Bishop, N. Daniell, G. Paul, Next-generation low-cost motion capture systems can provide comparable spatial accuracy to high-end systems, J. Appl. Biomech. 29 (1) (2013) 112–117, https://doi.org/10.1123/jab.29.1.112.

[228] X. He, K. Zhao, X. Chu, Automl: a survey of the state-of-the-art, Knowl.-Based Syst. 212 (5) (2021), 106622, https://doi.org/10.1016/j.knosys.2020.106622.

[229] G. Ghiasi, T.-Y. Lin, Q.V. Le, Nas-Fpn: learning scalable feature pyramid architecture for object detection, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 7036–7045, https://doi.org/10.1109/CVPR.2019.00720.

[230] X. Gong, S. Chang, Y. Jiang, Z. Wang, Autogan: neural architecture search for generative adversarial networks, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 3224–3234, https://doi.org/10.48550/arXiv.1908.03835.

[231] W. Peng, X. Hong, G. Zhao, Video action recognition via neural architecture searching. 2019 IEEE International Conference on Image Processing (ICIP), IEEE, 2019, pp. 11–15. doi:10.1109/ICIP.2019.8802919.

[232] P. Maurice, A. Malaisé, C. Amiot, N. Paris, G.-J. Richard, O. Rochel, S. Ivaldi, Human movement and ergonomics: an industry-oriented dataset for collaborative robotics, Int. J. Robot. Res. 38 (14) (2019) 1529–1537, https://doi.org/10.1177/0278364919882089.

[233] G. Guerra-Filho, A. Biswas, The human motion database: a cognitive and parametric sampling of human motion, Image Vis. Comput. 30 (3) (2012) 251–261, https://doi.org/10.1109/FG.2011.5771381.

[234] A. Dubois, L.-E. Gadde, The construction industry as a loosely coupled system: implications for productivity and innovation, Constr. Manag. Econ. 20 (7) (2002) 621–631, https://doi.org/10.1080/01446190210163543.

[235] H. Wang, J. Pi, T. Qin, S. Shen, B.E. Shi, Slam-based localization of 3d gaze using a mobile eye tracker, in: Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications, 2018, pp. 1–5, https://doi.org/10.1145/3204493.3204584.

[236] D. Wang, J. Chen, D. Zhao, F. Dai, C. Zheng, X. Wu, Monitoring workers' attention and vigilance in construction activities through a wireless and wearable electroencephalography system, Autom. Constr. 82 (2017) 122–137, https://doi.org/10.1016/j.autcon.2017.02.001.

[237] C. Llamas, M.A. González, C. Hernández, J. Vegas, Open source platform for collaborative construction of wearable sensor datasets for human motion analysis and an application for gait analysis, J. Biomed. Inform. 63 (2016) 249–258, https://doi.org/10.1016/j.jbi.2016.08.025.