Chapter 16
Digital Twin Models Supporting Cognitive Buildings for Ambient Assisted Living

Alessandra Corneli, Leonardo Binni, Berardo Naticchia, and Massimo Vaccarini

Abstract The rapid and global aging of population is outlining the need for environments that can provide support for these individuals during their daily activities. The challenge of an aging society is being addressed through the incorporation of new technologies into the home environment, which is nothing less than Ambient Assisted Living (AAL). To date, some of the AAL solutions exploit AI models to recognize the elderly's behaviors through data collected by sensors. In recent times, Digital Twins (DTs) at building level have begun to appear on the construction domain. These are still under development but through the integration of users into assessments, they improve efficiency, prevention, and prediction of likely events through real-time AI computing. The integration of DT and AAL defines cognitive buildings which aim to learn at scale, reason with a purpose, and co-operate with users in a natural way. This research aims to develop DT models to achieve scenario awareness to provide support to elderly people living alone and suffering from cognitive disorders. The proposed multi-agent architecture is based on a five-layer system that autonomously develops high-level knowledge to detect anomalies in the home environment scenarios and therefore support the user. Bayesian networks (BNs) are exploited to perform high-level deductive reasoning on low-level multi-modal information, thus recognizing senseless or dangerous behaviors, environmental disruptions, changes in behavioral patterns, and serious medical events. Bi-directional user-system interaction provides user support by leveraging Speech-To-Text and Text-To-Speech AI agents. Three main functions were tested: real-time data integration, anomaly detection, and two-way interaction.

A. Corneli (✉) · L. Binni · B. Naticchia · M. Vaccarini
Polytechnic University of Marche, Ancona, Italy
e-mail: a.corneli@univpm.it

L. Binni
e-mail: l.binni@pm.univpm.it

B. Naticchia
e-mail: b.naticchia@univpm.it

M. Vaccarini
e-mail: m.vaccarini@staff.univpm.it

16.1 Introduction

Projections about the increase of people in older age from the United Nations show that the number of people aged 65 or older will double by 2050 (United Nations 2020a, b), and this will cause significant effects on the share of the population suffering from geriatric diseases such as cognitive impairments, and consequently on the share of the population needing care due to loss of independence (Berryhill et al. 2012).

Advanced care planning, supported decision-making, and availability of assistive devices can enhance autonomy regardless of an elderly person's level of capacity (World Health Organization 2017).

Nevertheless, solutions are still far from predictive environments as intended by the Ambient Assisted Living (AAL) policies (World Health Organization, Regional Office for Europe 2017). This gap could be narrowed by the Digital Twin (DT) technology. Accordingly, this paper provides prototypes of DT models to support elderly people within the home environment in detecting anomalies in daily scenarios.

16.2 Literature Review

An increasing number of DT models are currently emerging within the built environment. The development of such innovating systems aims to achieve a variety of objectives depending on domains (Liu et al. 2021; Opoku et al. 2021; Sharma et al. 2020). Among the different assets that are currently being mirrored by DTs, both prevention and prediction of probable events during the whole life cycle of a building can be performed by means of a building-level DT, enhancing the building's efficiency as well. Building-level DTs consider both the environment and its user processing real-time information to offer appropriate services.

Hence, they could be also referred to as cognitive buildings (Yitmen et al. 2021).

AAL is a field that combines information and communication technologies, sociological sciences, and medical research and its purposes can be summarized as the development of products and services for countering the effects of a growing elderly population (Li et al. 2015; Dobre et al. 2017). Cognitive environments in AAL domain should be able to learn at scale, reason with purpose and co-operate with users in a natural way. Accordingly, some cognitive human-centered environments have recently begun to appear (De Paola et al. 2017; Rafferty et al. 2017; Patel and Shah 2020; Calderita et al. 2020) due to the significant influence that the home environment has on AAL's objectives.

Since pursuing and completing Activities of Daily Living (ADLs) allows autonomous well-being in older ages, this kind of system usually aims at encouraging, supporting, and easing the users in their ADLs. A variety of sensors and devices are exploited to collect data that can be processed by AI algorithms to make analysis on either user or environmental conditions. However, visual sensors (e.g., cameras) are not fully exploited yet.

16.3 Methodology

16.3.1 Mirroring Real Environment

A consistent virtual representation of the context must include both the building elements and the user. Accordingly, a real-time representation of the user is combined with BIM information within a game engine (GE), i.e., unity. The user virtual counterpart is synthesized through its posture, which is typically referred to as Skeleton.

A LiDAR camera is used as visual sensor (Intel RealSense L515). Nuitrack AI is used as the Skeleton tracking algorithm to process camera's raw 3D data and yield the user Skeleton. Besides, dynamic features such as appliance states, environmental temperature, and so on can also be associated with BIM elements into the GE to retrieve further information from sensor readings.

Though, further information is required to define a complete semantic of a user suffering from cognitive disorders. The activities that he or she performs can be detected through an Activity Recognition (AR) model. The AR task can be performed following either data-driven or knowledge-driven approaches (Rafferty et al. 2017). A data-driven approach is followed in this work since it enables the modeling of uncertainty and exploits increasingly available activity datasets. Specifically, the model developed in Liu et al. (2020) is integrated as the AR agent of the system. The MS-G3D model is based on Spatial–Temporal Graph Convolutional Networks (ST-GCNs), firstly proposed in Yan et al. (2018) for the AR task, and is therefore fed with the 3D coordinates of the Skeleton joints instead of RGB images as required by the previous Convolutional Neural Network (CNN). This results in a lightweight model that outperforms existing methods for AR.

16.3.2 Knowledge Contextualization

BIM information combined with its dynamic features and the user-related information define a low-level knowledge on the real asset of its virtual counterpart. A contextualization of such information allows interpreting the real world and thus the twinning of scenarios that includes the environment and the user with its behaviors,

habits, intentions, activities, and situations. Emergent scenarios that the system aims to detect are those including anomalies. Accordingly, knowledge contextualization can be achieved through an agent acting as the reasoner of the system. A rule-based reasoner has been proposed in De Paola et al. (2017). Their module consists in if, else conditional rules and takes basic decisions such as turning the heating/cooling system on/off depending on the user satisfaction. Since complex rules cannot cope with scalability and are not easily reusable, above all when considering an elderly person suffering from cognitive disorders, a probabilistic approach is followed by means of Bayesian Networks (BNs).

Such a probabilistic model is based on conditional probabilities that an event may occur depending on evidence or other variables. BNs are probabilistic graphical models that represent a set of variables and their conditional dependencies via a directed acyclic graph. Expert knowledge could be elicited in Conditional Probability Tables (CPTs) of the nodes that represent the events. BNs are ideal for taking an event that occurred and predicting the likelihood that any one of several possible known causes was the contributing factor (De Grassi et al. 2009). In this work, BNs are used to infer different types of scenario anomalies:

- Wasteful and senseless situations (e.g., window open while heating system is turned on)
- Unusual behaviors (e.g., skipping meals)
- Dangerous situations (e.g., something dropped on the ground)
- Emergencies (e.g., falls).

Once the real-world scenario has been recognized by the system, it should offer the appropriate supportive services. A dialog system is implemented through a flow-based programming tool (i.e., Node-RED) to enable a bidirectional interaction between the user and the system itself. This platform can bridge the gap between the reasoner and the services to deliver.

16.4 System Architecture

The architecture of the proposed system has been outlined following the guidelines stated in Lu et al. (2020), which define the structure of a building-level DT. Thus, our model consists of five layers, namely data acquisition layer, transmission layer, digital modelling layer, data/model integration layer, and service layer. The proposed architecture defines a system able to autonomously perform high-level reasoning to detect anomalies in daily scenarios and consequently offer support to the user. Figure 16.1 shows the architecture of the system.
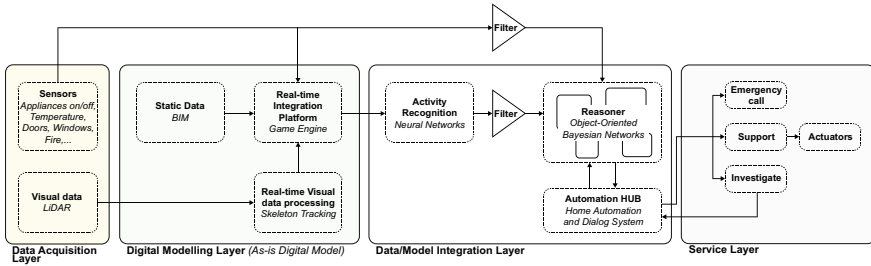
Fig. 16.1 System architecture. Arrows indicate data flow (transmission layer)

16.4.1 3D Real-Time Representation

Data acquisition and digital modelling layers hold the computation that manages the 3D real-time representation of the context. The virtual scenario built upon BIM and Skeleton data requires some adjustments. Two different filtering algorithms are developed for the following reasons:

- Filtering non-confident Skeleton data. Some Skeleton joints may have low confidence values due to obstructed camera field of view and can thus be less reliable leading to distortions of the user's avatar within the GE. Consequently, a threshold is introduced to discard data regarding Skeleton joints with a confidence value below 10%.
- Enhancing Skeleton stabilization. Once non-confident data are discarded, the avatar should move following natural movements. The avatar is stabilized through an autoregressive filter that acts on joint's position, orientation, and avatar's height.

$$X_{(t+1)} = (1 - a) \cdot X_{(t)} + a \cdot X_{(t+1)}^{\text{raw}} \qquad (16.1)$$

where $X_{(t+1)}$ is the processed data at the time $(t + 1)$, $X_{(t)}$ is the processed data at the time (t), a is a corrective factor with a value that ranges between 0 and 1, and $X_{(t+1)}^{\text{raw}}$ is the raw value of the data at the time $(t + 1)$.

These filters allow avoiding the ambiguities that the avatar had. Furthermore, cleaner movements mean more consistent output data that will consequently feed the AR model. To this end, the avatar joints have been remapped as the Kinect v2 Skeleton, used to build the NTU RGB + D dataset (Shahroudy et al. 2016) on which the MS-G3D model is pre-trained.

16.4.2   Scenario Awareness

The data/model integration layer is responsible for analyzing and processing the data to achieve scenario awareness. It consists of three intelligent agents that can, respectively, detect the activities performed by the user, reason on the current scenario detecting the anomalous ones, and act accordingly to support the individual if necessary.

The agent acting as the reasoner of the system is formalized through an Object-Oriented Bayesian Network (OOBN), composed in turn of four sub-modules (Fig. 16.2).

The approach followed to formalize the OOBN firstly considers general symptoms that could lead to anomalous scenarios (confusion, depression, loss of memory, and
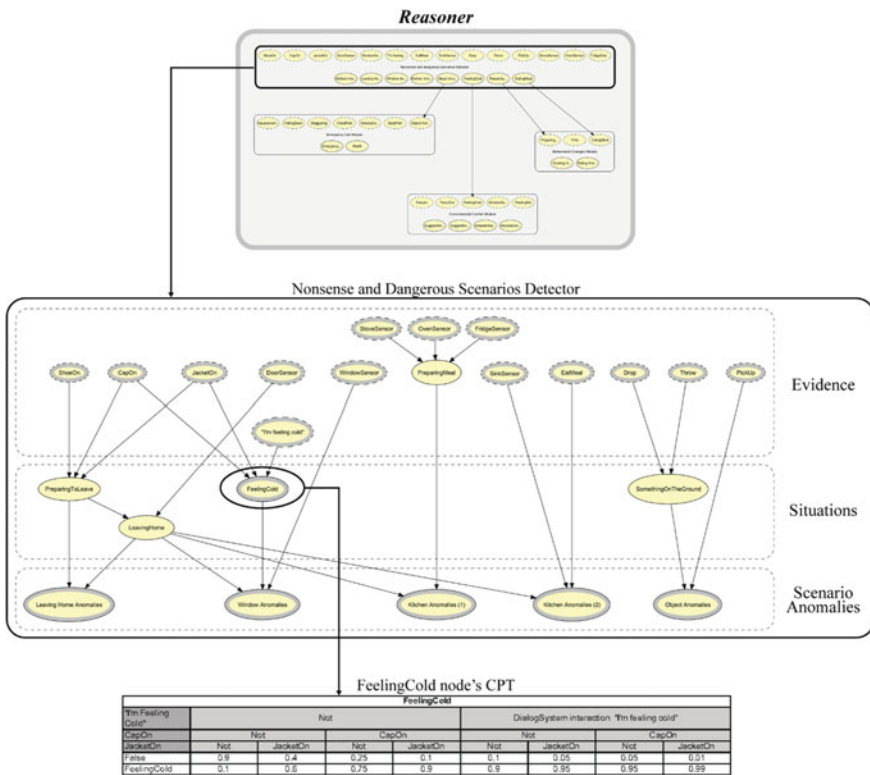
Fig. 16.2 The reasoner consists of an Object-Oriented Bayesian Network, decomposed in turn into four underlying OOBN modules. Zooming in shows the Nonsense and Dangerous Scenarios Detector module and the CPT of the "FeelingCold" situation node. Dashed gray-edged nodes represent input nodes. Solid gray-edged nodes are output nodes

so forth) (Berryhill et al. 2012; Dillon et al. 2013; scie.org, https://www.scie.org.uk/dementia). Then, a semantic regarding probable events, situations, scenarios, and anomalies in a AAL environment is built: Evidence is captured by sensors (turning on/off appliances, indoor/outdoor temperature, open/closed window, and so forth) as are the results of the AR model and the user-system vocal interactions; situations are combinations of evidence and represent feelings, behaviors, events, or intentions (feeling hot/cold, getting dressed, something on the ground, leaving home, and so forth).

By associating and combining available evidence and recognizable situations, probable scenarios are theorized as the anomalies that may occur. Anomalous scenarios therefore include time disorientation, difficulties arranging, indifference to the environment, getting easily overwhelmed, mishandling appliances, and changes in eating patterns.

The Automation HUB, based on Node-RED, can integrate applications to offer appropriate support to the user. In this work, a dialog system is implemented. Speech-to-Text (STT) and Text-to-Speech (TTS) services that rely on Machine Learning models are integrated to define a dialog system whereby bidirectional vocal interactions between the user and the cognitive layer of the building can be performed.

Is essential not to have Hot Phrases (HP) since the user may forget them due to cognitive impairments. HP are phrases typically used to trigger common dialog system such as Alexa and Google Assistant.

16.5 System Implementation

Combining BIM data and the Skeleton allows achieving a reliable real-time virtual representation of the physical asset which is shown in Fig. 16.3. BIM data from a home environment are converted to Industry Foundation Classes (IFC) format using Autodesk Revit. Importing IFC files into the unity game engine recognizes all BIM objects as Prefabs. Prefabs preserve all information related to BIM objects. Working with physics engines, Unity allows additional properties to be assigned to Prefabs achieving greater realism. Indeed, the mesh collider attribute is applied to all tangible components to avoid inconsistencies. Additionally, dynamic features can be added to Prefabs to extract real-time sensor readings about BIM objects. Besides, the LiDAR camera is placed at a height of 1 m and leveled horizontally. The tests show that the distance between the user and the camera should be unobstructed and not exceed 5 m to obtain consistent results.

To evaluate the effectiveness of the Object-Oriented Bayesian Networks developed in this work, the node's CPTs are filled eliciting the knowledge of the authors. Figure 16.2 shows the CPT relative to the "FeelingCold" situation node. Then, possible combinations of evidence are set up by manually activating input nodes, and the expected consequences achieve high percentage values meaning that predictable anomalies within the scenario are fully recognized. Figure 16.4 shows an example of anomaly detection within the Nonsense and Dangerous Scenario module. Four input
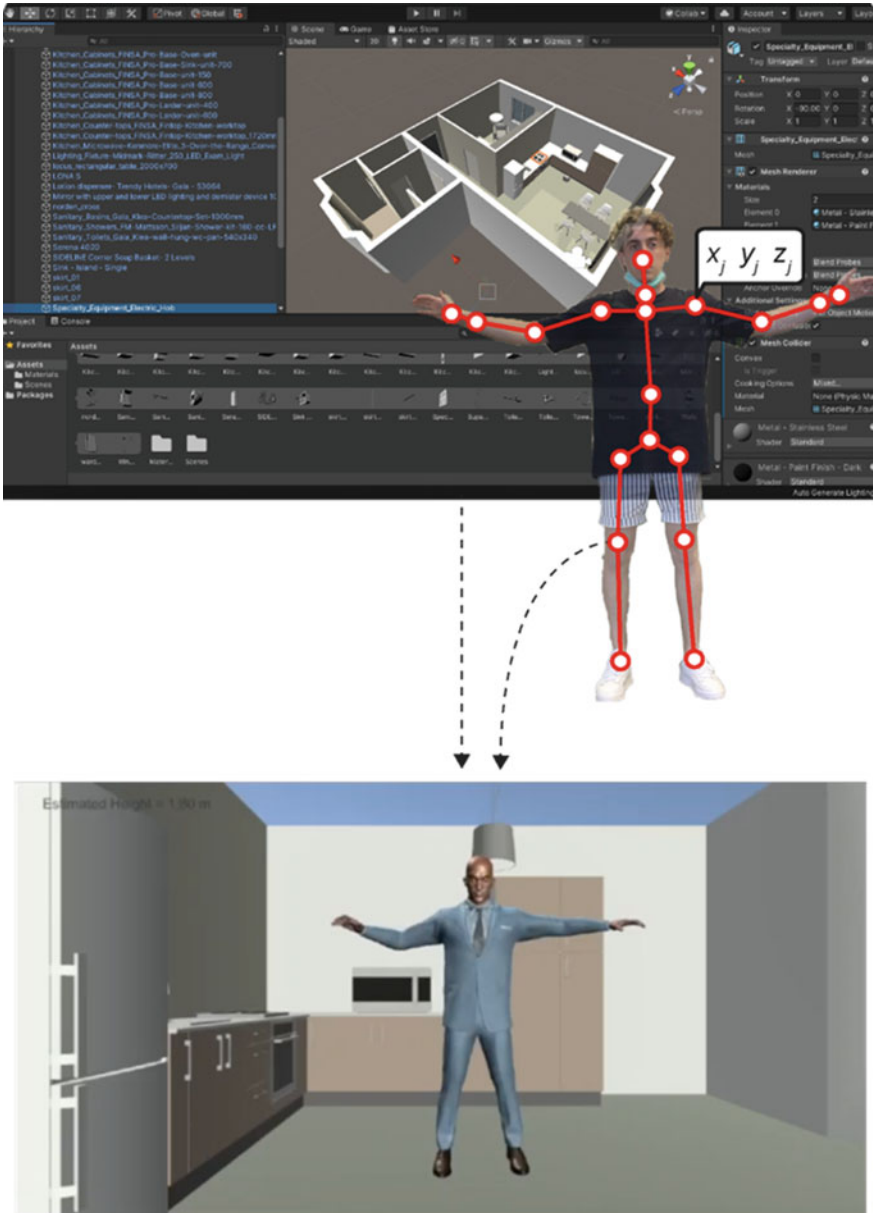
Fig. 16.3 Real-time 3D representation of the context

nodes are set up to represent a scenario where the user is barefoot, not wearing a jacket and hat, and is opening the door. Specifically, the ShoeOn, CapOn, and JacketOn input nodes have been set up to false (activities recognizable through the AR model), while the DoorSensor input node has been set up to open. The user is not preparing to leave (98.90% false) but is actually leaving home (90.09% true). The "Leaving Home Anomalies" output node detects a likelihood of 87.34% that the user is leaving undressed.

Figure 16.5 shows the dialog system prototype built upon the STT and TTS processes. STT module starts by recording the user's speech without requiring HP. Then, the record is managed by the IBM's Watson STT service that returns a transcription of the speech. Finally, the transcription is shown in the Node-RED's debug tab. By contrast, the TTS process is automatically triggered by the system depending on the output of the reasoner. Tailored messages can be played depending on the needs of the user. These written messages are converted through the IBM's Watson TTS service. Finally, the converted speech is played by the speakers.
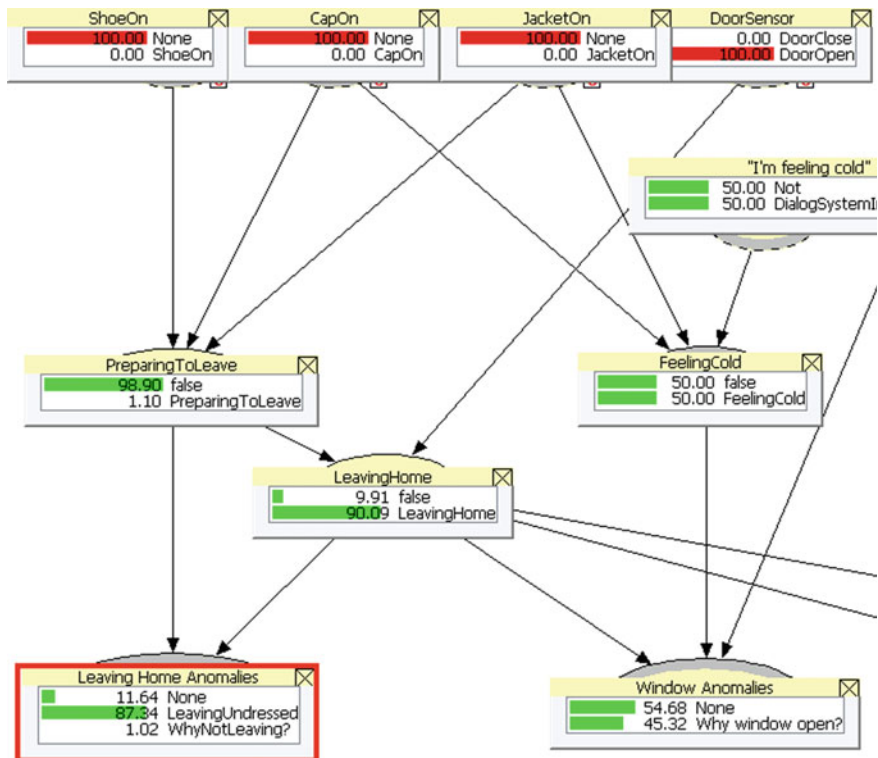

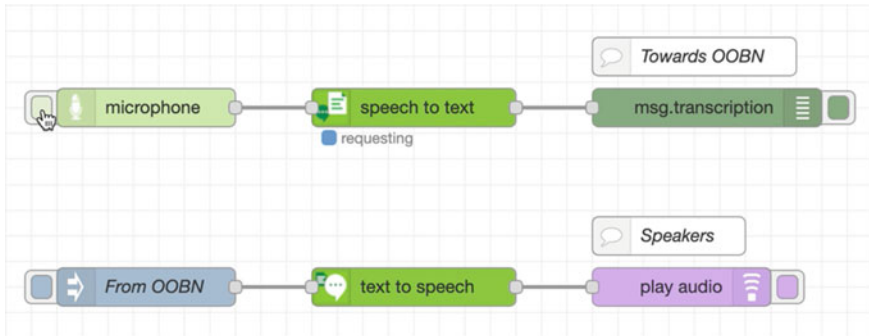**Fig. 16.4**   Detected anomaly within the Nonsense and Dangerous Scenario module

Fig. 16.5 SST and TTS processes within the Node-RED editor

16.6   Conclusion

The increase in the number of elderly people and consequently the increased occur-rence we will see in the future of geriatric cognitive disorders requires new systems for developing AAL solutions. Therefore, this research work aims to propose the development of Cognitive Building through the exploitation of the DT paradigm.

The grounded multi-agent system architecture defines a model able to autonomously perform real-time high-level reasoning, that allows the detection of anomalies in daily scenarios, and consequently offers support to the user. The knowl-edge development applied here is a major strength: the raw data that is captured by multi-modal sensors (visual and non-visual) and subsequently reported in 3D in real time, but also the reasoning applied at a high level when anomalies are detected. AR is performed using a neural network model that leverages 3D data derived from the user's pre-processed real-time 3D representation. On the other hand, the OOBN can recognize wasteful, meaningless, or dangerous behaviors, environmental distress, changes in behavioral patterns, and serious medical situations or events, and then trigger specific services. Two-way voice interaction with the individual is performed by the dialogue system, implemented in the Automation HUB agent, based on Node-RED. A number of improvements can be addressed as future work of this study: implementing the MS-G3D model; learning the OOBN modules through data captured from a real-world AAL environment; and fully testing the entire pipeline in an end-to-end manner.

References

Berryhill ME, Peterson D, Jones K, Tanoue R (2012) Cognitive disorders. In: Encyclopedia of human behavior, 2 edn, vol 1. Academic Press, pp 536–542

Calderita L, Vega A, Barroso-Ramírez S, Bustos P, Núñez P (2020) Designing a cyber-physical system for ambient assisted living: a use-case analysis for social robot navigation in caregiving centers. Sensors (Switzerland) 20(14):1–25

De Grassi M, Naticchia B, Giretti A, Carbonari A (2009) Reti Bayesiane con applicazioni all'edilizia e alla gestione del territorio. Franco Angeli, Milano

De Paola A, Ferraro P, Gaglio S, Re G, Morana M, Ortolani M, Peri D (2017) An ambient intelligence system for assisted living. In: AEIT international annual conference, Cagliari, Italy, pp 1–6

Dillon C, Serrano CM, Castro D, Leguizamon PP, Heisecke SL, Taragano FE (2013) Behavioral symptoms related to cognitive impairment. Neuropsychiatr Dis Treat 9:1443–1455

Dobre C, Mavromoustakis C, Garcia N, Goleva R, Mastorakis G (2017) Ambient assisted living and enhanced living environments, 1st edn. Butterworth-Heinemann

Li R, Lu B, McDonald-Maier K (2015) Cognitive assisted living ambient system: a survey. Digit Commun Netw 1:229–252

Liu Z, Zhang H, Chen Z, Wang Z, Ouyang W (2020) Disentangling and unifying graph convolutions for skeleton-based action recognition. arXiv:2003.14111v2

Liu Y, Chen K, Ma L, Tang S, Tan T (2021) Transforming data into decision making: a spotlight review of construction digital twin. In: Proceedings of the international conference on construction and real estate management 2021, pp 289–296

Lu Q, Parlikad A, Woodall P, Don Ranasinghe G, Xie X, Schooling J (2020) Developing a digital twin at building and city levels: case study of West Cambridge campus. J Manage Eng 36(3)

Opoku D, Perera S, Osei-Kyei R, Rashidi M (2021) Digital twin application in the construction industry: a literature review. J Build Eng 40:102726

Patel A, Shah J (2020) Real-time human behaviour monitoring using hybrid ambient assisted living framework. J Reliab Intell Environ 6(2):95–106

Rafferty J, Nugent C, Liu J, Chen L (2017) From activity recognition to intention recognition for assisted living within smart homes. IEEE Trans Hum-Mach Syst 47(3):368–379

scie.org. https://www.scie.org.uk/dementia

Shahroudy A, Liu J, Ng T-T, Wang G (2016) NTU RGB+D: a large scale dataset for 3D human activity analysis

Sharma A, Kosasih E, Zhang J, Brintrup A, Calinescu A (2020) Digital twins: state of the art theory and practice, challenges, and open research questions. Tratto da https://arxiv.org/abs/2011.02833

United Nations (2020a) World population ageing 2019. United Nations, Department of Economic and Social Affairs, Population Division

United Nations (2020b) World population ageing 2020 highlights: living arrangements of older persons. United Nations, Department of Economic and Social Affairs, Population Division

World Health Organization (2017) Global strategy and action plan on ageing and health. WHO, Geneva

World Health Organization, Regional Office for Europe (2017) Age-friendly environments in Europe. A handbook of domains for policy action. Denmark

Yan S, Xiong Y, Lin D (2018) Spatial temporal graph convolutional networks for skeleton-based action recognition. In: AAAI conference on artificial intelligence

Yitmen I, Alizadehsalehi S, Akiner I, Akiner ME (2021) An adapted model of cognitive digital twins for building lifecycle management. Appl Sci 11(9)