Deep Learning Object-Recognition in a Design-to-Robotic-Production and -Operation Implementation

Liu Cheng, Alexander; Bier, Henriette; Mostafavi, Sina

DOI
10.1109/ETCM.2017.8247495

Publication date
2017

Document Version
Accepted author manuscript

Published in
Proceedings of the 2nd IEEE Ecuador Technical Chapters Meeting (ETCM 2017)

Deep Learning Object-Recognition in a *Design-to-Robotic-Production and -Operation* Implementation

Alexander Liu Cheng[1,3], Henriette Bier[1,2], Sina Mostafavi[1,2]
[1]Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, The Netherlands
[2]Dessau Institute of Architecture, Anhalt University of Applied Sciences, Dessau, Germany
[3]GRAFT Gesellschaft von Architekten mbH, Berlin, Germany

E-mail: {a.liucheng, h.h.bier, s.mostafavi}@tudelft.nl

Abstract*—This paper presents a new instance in a series of discrete *proof-of-concept* implementations of comprehensively intelligent built-environments based on *Design-to-Robotic-Production and -Operation* (D2RP&O) principles developed at *Delft University of Technology* (TUD). With respect to D2RP, the featured implementation presents a customized design-to-production framework informed by optimization strategies based on point clouds. With respect to D2RO, said implementation builds on a previously developed highly heterogeneous, partially meshed, self-healing, and *Machine Learning* (ML) enabled *Wireless Sensor and Actuator Network* (WSAN). In this instance, a *computer vision* mechanism based on open-source *Deep Learning* (DL) / *Convolutional Neural Networks* (CNNs) for object-recognition is added to the inherited ecosystem. This mechanism is integrated into the system's *Fall-Detection and -Intervention System* in order to enable decentralized detection of three types of events and to instantiate corresponding interventions. The first type pertains to *human-centered* activities / accidents, where cellular- and internet-based intervention notifications are generated in response. The second pertains to *object-centered* events that require the physical intervention of an automated robotic agent. Finally, the third pertains to *object-centered* events that elicit visual / aural notification cues for human feedback. These features, in conjunction with their enabling architectures, are intended as essential components in the on-going development of highly sophisticated alternatives to existing *Ambient Intelligence* (AmI) solutions.

Keywords—Design-to-Robotic-Production & -Operation; Wireless Sensor and Actuator Network; Ambient Intelligence; Computer Vision; Object-Recognition

I. Introduction

Ambient Intelligence (AmI) [1] promotes a vision of the future dwelling space as a digital living room, where automated and intuitive embedded-technologies enhance the inhabitants' experience and comfort. Since its conception, discussions of AmI have centered around ICTs, rendering considerations pertaining to the built-environment as incidental—an overview of the current AmI literature confirms this assertion, e.g., [2, 3]. The implementation of sophisticated ICT systems in static built-environments unnecessarily and inadvisably subjects new technologies and methods to the limitations of outdated modes of building and dwelling. This is tantamount to methodological retrofitting, which hinders the potential of such solutions and the effectiveness of their services. In order to avoid this unintended consequence, the sophistication of built-environments must be commensurate and complementary to that of deployed technologies. In ascertaining so, a more holistic intelligent built-environment emerges, one capable of intuitive, enriching, and effective interactions as well as interconnections between users and their ICTs-integrated built-environments [4]. At present, established research groups and/or projects are developing expressions of such environments (or variations thereof) with promising results—e.g., the *Aware Home Research Initiative* [5]; *The Center for Advanced Studies in Adaptive Systems* [6]; and *PlaceLab* [7], etc.

In this paper, *Design-to-Robotic-Production and -Operation* (D2RP&O) [8, 9] principles inform the development of a discrete and sophisticated *proof-of-concept* intelligent built-environment. Architectural considerations with respect to form, fabrication, and integration of materially heterogeneous physical components are informed by D2RP, while technical and technological considerations pertaining to computational / robotic services deployed in the resulting environment are informed by D2RO. Furthermore, decisions adopted in the physical domain are considered in the computational / robotic and vice versa, resulting in a more deliberate design strategy where neither form nor services are incidental with respect to one another. More specifically, in the present implementation, and with respect to D2RP (Sections II.B and III.A), a real-scale fragment of a conceptual student housing unit is fabricated as a multi-layered hybrid component consisting of concrete and *Expanded Polystyrene* (EPS). Its overall form, distribution of cavities, and densities of porosities are determined by structural optimization, *Interior Environmental Quality* (IEQ) [10] considerations, and the integration of anticipated ICTs. With respect to D2RO (Sections II.A, II.C, III.B), this implementation expands on the *System Architecture* of a previously developed prototype [11] to include object-recognition via *Deep Learning* (DL) / *Convolutional Neural Networks* (CNNs). This vision mechanism is integrated in an inherited *Fall-Detection and -Intervention System* (FADIS) [12] in order to identify three human-centered as well as object-centered events and to instantiate automated interventions accordingly for the promotion of well-being.

II. Concept and Approach

A. *Overview of inherited heterogeneous, scalable, self-healing, and partially meshed System Architecture*

The present implementation inherits a previously developed WSAN, whose *system architecture* consists of four subsystems briefly summarized as follows (for more technical descriptions, see [11]):

1) *Dynamically Clustering Local nodes*

A variety of *Microcontroller Units* (MCUs) and development platforms (e.g., Raspberri Pi 3 (RPi3) and Zero W (RPiZW)) serve as nodes dependent on a local structured environment. More powerful nodes may be clustered dynamically to yield a single node with higher computational power depending on load-requirements. All nodes exchange data in a combination of wired (e.g., Ethernet, USB, Serial) and wireless (WiFi, BLE, and ZigBee) protocols, depending on latency and frequency requirements.

2) *Wearable devices*

A set of three *LightBlue Beans*™ (LBBs) conform the location-dependent wearables while a Fitbit® Charge HR™ activity tracker represents the location-independent wearable. The LBBs detect movement in the upper-body, upper- and lower-extremities and advises the system to listen for *Open Sound Control* (OSC) packets corresponding to accelerometer data sent from a smartphone for *Human Activity Recognition* (HAR). The activity tracker enables a constant feed of physiological data while the user is outside of the structured environment.

3) *Ad hoc Support devices*

In the last five years, smartphones have become convenient and ubiquitous tools for HAR via *Machine Learning* (ML) [13, 14], which in conjunction with their battery life and rechargeability render them a preferred means of accelerometer-data gathering in intelligent built-environment implementations.

4) *Remote / Cloud Services*

Six cloud-based services conform this subsystem: (*I*) external ML mechanism via MATLAB® (in case the local ML mechanism fails); (*II*) data exchange with Fitbit®'s servers via its API [15]; (*III*) cloud data-storage and -plotting via Plotly®'s API [16]; (*IV*) Amazon®'s *Alexa Voice Service* [17]; (*V*) automated SMS notifications, both via Twilio®'s API [18] as well as via a T35 GSM module; and (*VI*) automated email notifications via Gmail©'s API [19].

B. *An Informative Framework for Multi-Layering Materiality Design-to-Robotic-Production* (D2RP)

For the present implementation, an integrated single-occupant student housing unit is conceptualized. This unit is formally defined by optimization strategies based on point-clouds, where each point bears both physical and non-physical information about the envisioned space. Various sets of points provide different types of information—i.e., sets corresponding to spatial definition, structure analysis, heating and cooling, lighting requirements, and the integration of ICT devices (see Fig. 1 and Fig. 2).

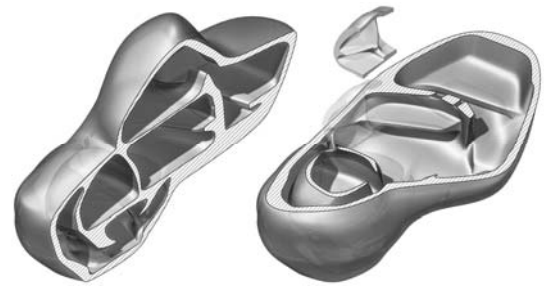

Fig. 1. Schematic design of student housing by M. Moharram, H. Hesham, and M. Elmeligy [20].

The real-scale fragment representative of the conceptual unit is developed in three main steps. First, structural loads and supports locations are determined based on the configuration of the initial form. By using the resulting geometry and the identified locations of supports and loads, a *finite element* model is created and the its corresponding stress lines extracted. These lines are then used to generate the structural elements inside the building component. Second, required lighting is determined based on activities and their corresponding minimum / maximum thresholds of illumination during a typical 24-hour period. This informs the shape and the location of porosities and cavities in the component, enabling the integration of LED-based illumination systems where necessary. Third, heating and cooling requirements—as identified by *Comité Européen de Normalisation* (CEN) Standard EN15251-2007 [21]—inform the orientation of ventilation openings, the integration of intelligent ventilation systems, and the position of required sensors for automated control of said openings to ascertain partial IEQ. Finally, these considerations determined the composition and arrangement of different materials (i.e., concrete and EPS) and identified optimal locations for the integration of ICT devices, which collectively shaped the resulting component.

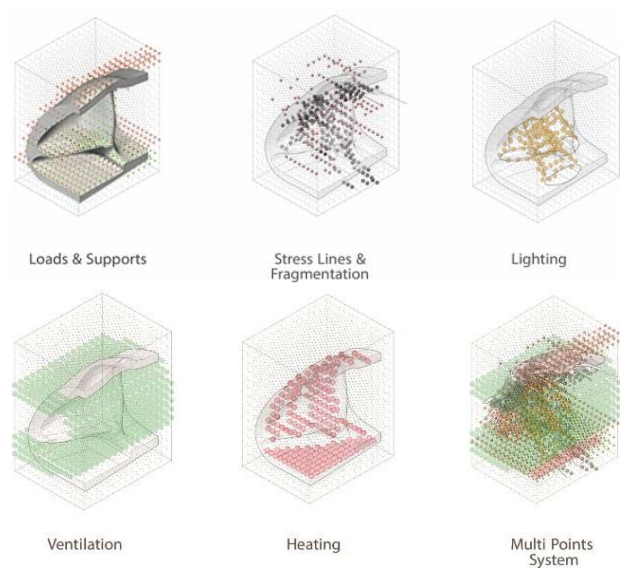

Fig. 2. Sets of point-clouds corresponding to different types of information

C. *Object-Recognition via Deep Learning / Convolutional Neural Networks as a visual component of Design-to-Robotic-Operation* (D2RO)

The newly integrated object-recognition component is intended to enable decentralized detection of three types of events and to instantiate corresponding interventions (see Section IV):

1) Human-centered event: Fall-Detection and -Intervention System, ver. 2.0.

The object-recognition mechanism is integrated with the inherited WSAN's FADIS in order to detect a variety of human- and object-centered events and to yield corresponding reactions / interventions to promote well-being. The existing FADIS adopts a *laser-reflectivity* method [22] in order to detect the presence of collapsed objects and their estimated size. If the shape and size of the detected object corresponds to the dimensions of a person, the system gauges the probability of a collapsed person as high. Consequently, large inanimate objects may cause the system to instantiate false-positives.

The object-recognition component represents an added layer of verification that decreases the probabilities of false-positives, as its enabling DL / ANNs mechanism is trained to detect human faces and shapes. In this particular instance, the object-recognition component needs only one camera, integrated into the ceiling of the scanned environment. However, the system architecture intends for multiple cameras be integrated in the same and in various regions across several nodes in the overall built-environment in order to further increase the probabilities of accurate detections by cross-referencing purported detections. That is to say, we may imagine a scenario where lasers have detected a human-size object and its shape has indeed been identified—via that particular space's integrated ceiling-camera—to correspond to that of a human's. Nevertheless, it may still be a false positive.

In order to reduce the probabilities of this scenario, a number of new features—including the discussed object-recognition component—have been added to FADIS ver. 2.0. First, via a wearable (e.g., LBBs and/or Fitbit activity tracker) and/or smart-device, the presence of the occupant is registered by the WSAN (i.e., the WSAN is programmed to detect the presence and signal intensity of particular MAC addresses within its structured environment). If the occupant is indeed confirmed to be present, and if he/she is detected—via both lasers and a ceiling-camera—to have collapsed in the bathroom, the WSAN will take one final verification step before instantiating appropriate intervention mechanisms (i.e., SMS / email notifications to family-members and/or caretakers). That is, the WSAN requests information from all other nodes controlling the remaining cameras deployed in the overall space in order to detect instances of ambiguities. For example, if the occupant was detected to be in the bathroom as well as in the living-room in the single-occupant unit, then one or more detections may be false-positives. If the occupant has indeed collapsed in the bathroom, then the remaining camera-controlling nodes must not be able to return positive detections.

Hence, and to summarize, if lasers in a given region have detected a collapsed large-object; and if the region's corresponding camera has identified said large-object's shape as that of a human's, and if the wearables / smart-devices associated with the occupant are detected to be within the structured environment; and if no other camera in any other region of the overall unit has detected human-like objects; *then* the built-environment may instantiate aforementioned intervention mechanisms with a high degree of confidence.

2) Object-centered event, robotic intervention:

In this type of events, if FADIS ver. 2.0 detects the unexpected presence of a small object in an otherwise empty region, it engages the object-detection mechanism to attempt to identify it. The identification process is enhanced by cross-referencing the detection via several cameras in whose field of vision the object is found. As may be seen in Fig. 4, an object may be identified by fragments of it—i.e., in said figure, the object-recognition mechanism detects a "cup" by its overall shape and by its handle. This feature is necessary for the present type of events in questions, as its principal purpose is to detect the presence of broken and/or unexpected idle objects on the floor.

In the example of the cup, if from multiple angles (i.e., multiple cameras) a detached handle provides the necessary confidence level for a "cup" to be identified, the WSAN engages a robotic agent inherited from FADIS (i.e., a TurtleBot® [12, 23]) in order to retrieve it. Admittedly, there are considerable limitations with this feature stemming from object-recognition via fragments (see Section IV for a brief discussion of possible solutions).

3) Object-centered event, visual / aural warning:

This type of events is similar to the previous one, except for robotic intervention is replaced with both more passive visual and/or aural interventions as well as palpable vibrations. It may be imagined that an unexpected object, broken or not, may be too large for the TurtleBot® to remove. In such a scenario, visual cues in the form of rapid light-bursts and/or aural warnings in the form of a range of sound-emissions are first instantiated in order to elicit human action. If no response follows, the WSAN sends SMS and email notifications to the occupant, who may consequently instruct the system to ignore the object via an SMS response.

The intervention mechanisms associated with this type of events may provide an additional assistive service to visually and/or aurally impaired individuals who would otherwise have no means of preemptively learning about unexpected collapsed objects.

III. Methodology and Implementation

A. *D2RP: Realization of real-scale fragment*

The design-to-production framework is tested by robotically producing the real-scale fragment / prototype as a multi-layered hybrid component consisting of concrete, EPS, and smart devices. This fragment follows *componentiality* and *hybridity* principles characteristic of D2RP&O. That is

to say, with respect to the former, complex geometries are intelligently divided into components following a structural analysis to identify optimal division-seams that do not compromise physical integrity (see Fig. 3).

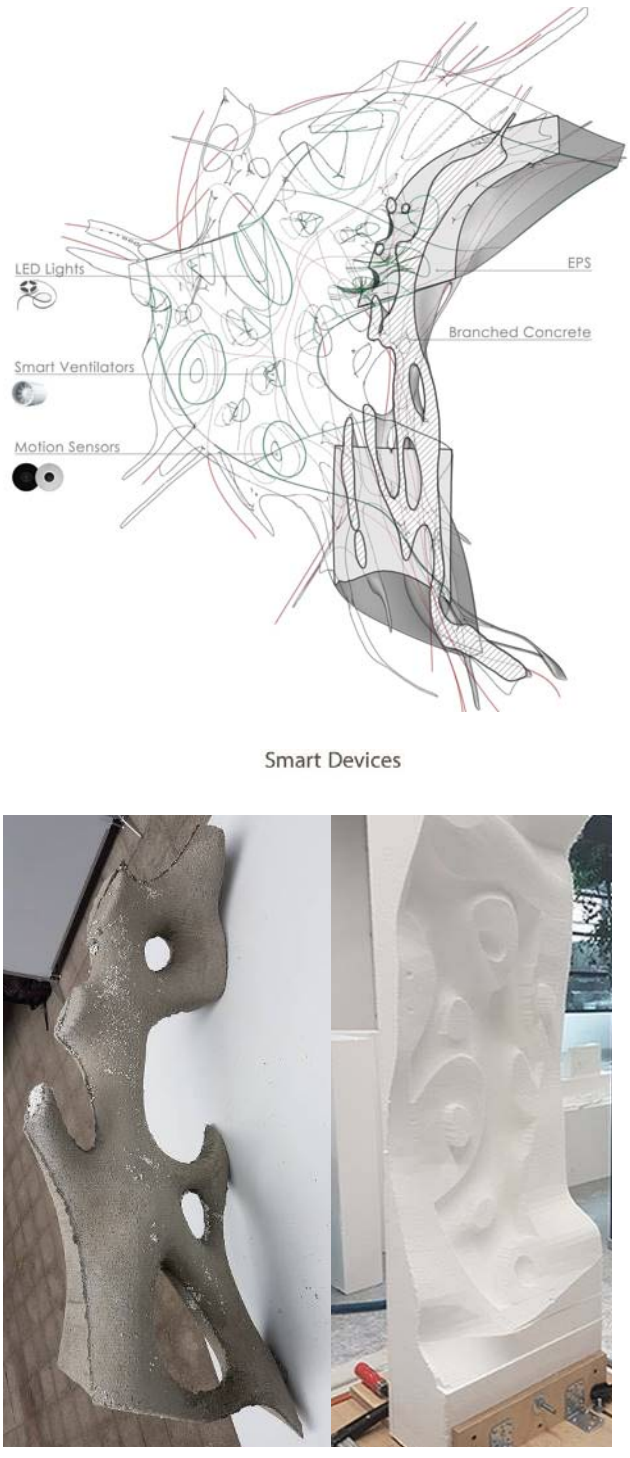

Fig. 3. Top: Real-scale fragment's multi-layered fabrication / integration logic. Bottom: Robotically fabricated concrete (left) and EPS (right) fragments.

With respect to the latter, the composition of each component unit consists of the integration of materially heterogeneous layers, each design in direct response to a purpose or a function. For example, the concrete layer is formed following the stress lines extracted from the *final element* model. Similarly, some of the cavities in the EPS layer are designed with ICT-integration in mind, while others with CEN-identified ventilation requirements.

B. D2RO: Deploying the Object-Recognition Mechanism and Corresponding Intervention Mechanisms

The object-recognition mechanism is implemented with open-source BerryNet® [24], which is built with a classification model (viz., Inception® ver. 3 [25]) as well as a detection model (viz., TinyYOLO® [26]). The classification model uses CNNs, which are at the forefront of ML research [25]. An advantage of BerryNet® is that it is a fully localized DL gateway implementable on a cluster of RPi3s. On an individual RPi3, the *inference* process is slow, requiring a delay between object-recognition sessions. This situation is ameliorated by the dynamic clustering feature of the WSAN (see Fig. 4).

Another benefit-*cum*-limitation is that BerryNet®'s classification and detection models are pretrained, which avoids the need to generate said models locally (see Section IV for an elaboration on limiting consequences as well as the generation of local models if and/or when necessary).

In Section II, it was asserted that the object-recognition mechanism was intended to be deployed across a variety of cameras in the overall built-environment, and that instances of detection were to be cross-referenced to minimize false positives. In order to implement this setup, each RPi3 node in the WSAN was equipped with a low-cost Raspberry Pi Camera® V2.1, then BerryNet® was installed in every node and the *inference* mechanism tested individually. The next step was to enable the nodes to share their detection results, which could be done via WiFi. Nevertheless, in order to reduce energy-consumption for every object-detection cross-referencing instance, ZigBee was preferred. In order to enable ZigBee on BerryNet®'s *detection_server.py* and *classify_server.py* were modified and made compliant with *python-xbee* [27].

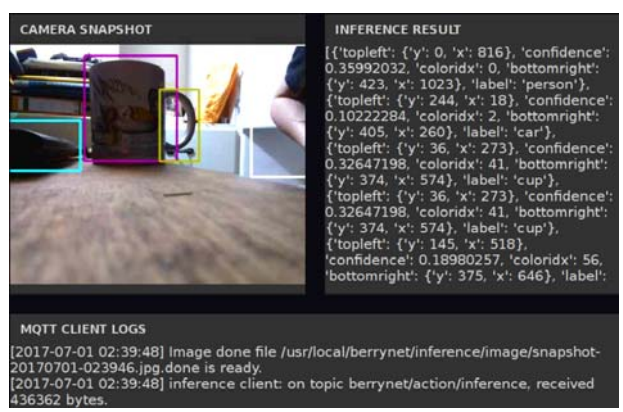

Fig. 4. Multiple-object detection via *BerryNet* [24]: 'person', 'cup'.

IV. Results and Discussion

A. *Human-centered event, limited demonstration:*

The first scenario was verified by having (1) the original FADIS detect a collapsed large-object; (2) a BerryNet®-enabled ceiling-integrated RPi3 detect a 'person'; (3) surrounding BerryNet®-enabled RPi3 nodes (with corresponding cameras) exchange each other's *inference* results via XBee-antennas; (4) corresponding SMS and email notifications sent.

A caveat pertaining to step 3: while a majority of surrounding RPi3 nodes identified the same object—in varying angles—as likely to be a person, not all of them did. Depending on lighting conditions and body-postures, some inference results read 'car', 'lamp', 'sofa'. In these instances, the probability of the object being a person was simply determined by whether the majority of inferring RPi3 nodes returned 'person' or not. One way to improve the probabilities that the majority of nodes identify a same object accurately would be to train the classification models particularly and further, but even this would not ascertain absolute certainty. A better approach would be to keep adding correlation factors via a variety of sensors in order to identify false-positives.

B. *Object-centered event, robotic intervention, limited demonstration:*

The second scenario was verified by having the object-recognition mechanism accurately detect a cup on the floor, which caused the WSAN to relay the XY-coordinates of the object to the TurtleBot® in order for the rover to reach the cup's location (see Fig. 5). *Rviz* [28] was used to enabled the rover to identify the boundaries of its deployment space. In the executed sample runs, the rover was able to arrive at the defined destination while avoiding collision with non-target object on the way. However, at present the TurtleBot® is only able to drag the object away via a rudimentary hook.

Consequently, further development of this system would require the design of a gripper system capable of sensing pressure. Another limitation is that during initial execution, the rover must be placed at its origin position, as defined in the process of generating the environment's map. Over time and inaccuracies of perception, etc., the rover's position and orientation will become uncontrollably inaccurate. This could be ameliorated by the addition of reference touch-sensors or switches in key locations within the environment in order to reset the rover's position and orientation.

C. *Object-centered event, visual / aural warning, limited demonstration:*

The third scenario was verified in tandem with the second. At the detection of the object on the floor, and while the rover was sent to fetch it, a LED and a *buzzer* emitted light and sound, respectively. This was repeated at predetermined intervals, only to stop when a corresponding SMS was sent to the SIM-card installed in the fee-based SMS-sending and -receiving T35 GSM module.

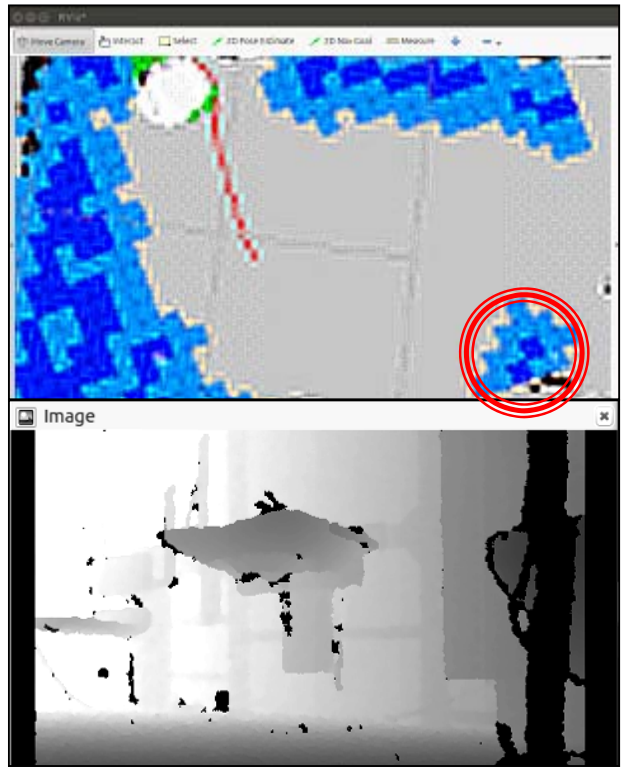

Fig. 5. Robotic intervention based on object-centered events. Top: Rover sent to the location of detected object (circled in red). Bottom: Abstracted robotic vision on the rover.

V. Conclusion

The present discrete *proof-of-concept* implementation adds two important contributions to the on-going development of the Architecture and Systems Architecture of D2RP&O-driven intelligent built-environments.

With respect to Architecture, the developed multi-layered, materially heterogeneous, and structurally optimized fragment herein described illustrates the benefits and feasibility of *componentiality* and *hybridity* in the development of building components. Each layer is informed by a particular consideration (e.g., structural loads, ventilation, illumination requirements, ICT-integration) and therefore justified formally and economically. This approach embeds intelligence from the onset of the design process.

With respect to Systems Architecture, the detailed object-recognition mechanism adds another means for the system to become aware of its built-environment. In this paper, this mechanism has been used to ascertain greater probabilities for accurate object-identification. In the present scope, the capabilities of this mechanism have been limited to recognizing general objects. But it may be extended to detect particular faces and features, which may later be correlated to specific events. In the setup discussed, the deployment scenario was construed as a single-occupant housing unit. But in a scenario with more occupants, the recognition of

each individual may instantiate actuations and transformations in the built-environment specific to each individual's preferences and taste. Furthermore, it may also be extended to conform a *security* mechanism. In this development, ZigBee was used to exchange *inference* data between nodes. But as part of a security mechanism, a WiFi communication layer would render *live-streaming* from any camera to any node (via assigned static IP-addresses) within the WSAN. In conjunction, all security cameras could be trained to identify—via CNNs—one or more particular faces as 'dangerous', etc., and the WSAN would proceed to notify pertinent services accordingly.

ACKNOWLEDGMENT

This paper has profited from the contribution of TUD Robotic Building and DIA researchers, tutors, and students. In particular, the authors acknowledge DIA students Mohammed Moharram, Hossam Hesham, and Mahmoud Elmeligy for the development of the student housing unit as well as its corresponding real-scale fragment.

REFERENCES

[1] E. Zelkha, B. Epstein, S. Birrell, and C. Dodsworth, "From Devices to 'Ambient Intelligence': The Transformation of Consumer Electronics (Conference Keynote)," in *Digital Living Room Conference*, 1998.

[2] H. Lindgren *et al., Ambient intelligence: Software and applications -- 7th International Symposium on Ambient Intelligence (ISAmI 2016).* Switzerland: Springer, 2016.

[3] P. Novais, K. Hallenborg, D. I. Tapia, and J. M. C. Rodríguez, *Ambient Intelligence - Software and Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[4] K. E. Green, *Architectural robotics: Ecosystems of bits, bytes, and biology*. Cambridge, Massachusetts: The MIT Press, 2016.

[5] J. A. Kientz *et al.,* "The Georgia Tech aware home," in *CHI '08: The 26th annual CHI Conference on Human Factors in Computing Systems: April 5-10, 2008 in Florence, Italy*, New York: ACM, 2008, p. 3675.

[6] P. Rashidi and D. J. Cook, "Keeping the Resident in the Loop: Adapting the Smart Home to the User," *IEEE Trans. Syst., Man, Cybern. A*, vol. 39, no. 5, pp. 949–959, 2009.

[7] A. A. Helal, M. Mokhtari, and B. Abdulrazak, *The engineering handbook of smart technology for aging, disability, and independence*. Hoboken, N.J.: Wiley, 2008.

[8] H. H. Bier and S. Mostafavi, "Robotic Building as Physically Built Robotic Environments and Robotically Supported Building Processes," in *Human-computer interaction series, Architecture and interaction: Human computer interaction in space and place*, N. S. Dalton, H. Schnädelbach, M. Wiberg, and T. Varoudis, Eds., Switzerland: Springer International Publishing, 2016, pp. 253–271.

[9] H. H. Bier, "Robotic Building as Integration of Design-to-Robotic-Production & Operation," *Next Generation Building*, no. 3, 2016.

[10] P. M. Bluyssen, *The healthy indoor environment: How to assess occupants' wellbeing in buildings*. London, New York: Routledge/Taylor & Francis Group, 2014.

[11] A. Liu Cheng, H. H. Bier, G. Latorre, B. Kemper, and D. Fischer, "A High-Resolution Intelligence Implementation based on Design-to-Robotic-Production and -Operation strategies," in *Proceedings of the 34th International Symposium on Automation and Robotics in Construction (ISARC 2017)*, 2017.

[12] A. Liu Cheng, C. Georgoulas, and T. Bock, "Fall Detection and Intervention based on Wireless Sensor Network Technologies," *Automation in Construction*, 2016.

[13] J. L. R. Ortiz, *Smartphone-based human activity recognition*. Cham: Springer, 2015.

[14] D. Micucci, M. Mobilio, P. Napoletano, and F. Tisato, "Falls as anomalies?: An experimental evaluation using smartphone accelerometer data," *Journal of Ambient Intelligence and Humanized Computing*, vol. 8, no. 1, pp. 87–99, 2017.

[15] Fitbit®, *Take your place in the future of fitness: Develop Your App With The Fitbit API*. [Online] Available: https://dev.fitbit.com/. Accessed on: Mar. 06 2016.

[16] Plotly®, *API Settings*. [Online] Available: https://plot.ly/settings/api. Accessed on: Feb. 15 2015.

[17] Amazon®, *Alexa Voice Service: Bring Alexa to your connected products with Amazon's intelligent cloud-based voice service.* [Online] Available: https://developer.amazon.com/alexa-voice-service. Accessed on: Apr. 19 2017.

[18] Twilio®, *REST API: Sending SMS or MMS*. [Online] Available: https://www.twilio.com/docs/api/rest/sending-messages. Accessed on: Apr. 01 2017.

[19] Alphabet®, *Google Gmail > API: Flexible, RESTful access to the user's inbox.* Accessed on: Apr. 20 2017.

[20] M. Moharram, H. Elbrrashi, and M. Meligy, *DIA Group 2: Topological Hybridity.* [Online] Available: http://ip.hyperbody.nl/index.php/Msc2G8:Group. Accessed on: Jul. 01 2017.

[21] Comité Européen de Normalisation© (CEN), *Standard EN 15251–2007: Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics.* [Online] Available: http://www.sysecol2.ethz.ch/OptiControl/LiteratureOC/CEN_06_prEN_15251_FinalDraft.pdf. Accessed on: 16/07/07.

[22] Y. Pyo, T. Hasegawa, T. Tsuji, R. Kurazume, and K.'i. Morooka, "Floor sensing system using laser reflectivity for localizing everyday objects and robot," *Sensors (Basel, Switzerland)*, vol. 14, no. 4, pp. 7524–7540, 2014.

[23] C. Georgoulas, A. Raza, J. Güttler, T. Linner, and T. Bock, "Home Environment Interaction via Service Robots and the Leap Motion Controller," in *Proceedings of the 31st International Symposium on Automation and Robotics in Construction (ISARC 2014)*, 2014.

[24] DT42©, Ltd., *BerryNet®: Deep learning gateway on Raspberry Pi.* [Online] Available: https://github.com/DT42/BerryNet. Accessed on: Jun. 21 2017.

[25] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, *Rethinking the Inception Architecture for Computer Vision.* Available: http://arxiv.org/pdf/1512.00567.

[26] J. Redmon and A. Farhadi, *YOLO9000: Better, Faster, Stronger.* Available: http://arxiv.org/pdf/1612.08242.

[27] n.io Innovation©, LLC, *python-xbee: Python tools for working with XBee radios.* [Online] Available: https://github.com/nioinnovation/python-xbee. Accessed on: Jun. 01 2017.

[28] ROS.org®, *ros-visualization: rviz is a 3D visualizer for the Robot Operating System (ROS) framework.* [Online] Available: https://github.com/ros-visualization/rviz. Accessed on: Jul. 16 2015.