A semantic model for computer vision-based hazard identification in construction

Weili Fang[1], Hanbin Luo[1*] and Peter E.D. Love[2]
[1] Dept. of Construction Management, School of Civil Engineering and Mechanics, Huazhong University of Science and Technology, 430074, Hubei, China
[2] Dept. of Civil Engineering, Curtin University, Perth, Western Australia, Australia
* email: luohbcem@hust.edu.cn

Abstract

Potential hazards for people working on construction sites include falls from heights (FFH), the collapse of trenches and scaffolding, electric shocks and failure to use proper personal protective equipment. Such hazards are major contributors to accidents and fatalities. To assist with the mitigation of accidents and fatalities, computer vision has been used to automatically detect safety hazards. Yet as the definition of safety violation becomes more complicated, interdependent and stringent, there is a likelihood that prevailing computer vision approaches will be unable to detect emerging hazards, which would render them to be obsolete unless they can accommodate change and the nuances of construction. To address this problem, this paper combines computer vision algorithms with ontology to develop a novel and robust semantic approach that is able to automatically and accurately recognize hazards. Our proposed semantic computer vision approach consists of (1) an ontological model for safety hazards; (2) detection of object entities and their attributes; (3) extraction of spatial relationships from images, and (4) reasoning data for hazard identification with a graph database. This paper uses the detection of FFH hazards as an example to illustrate the proposed approach. The research demonstrates that the proposed approach can successfully detect FFH hazards in a variety of contexts from images.

Keywords: Computer vision, hazard identification, ontology, safety

1. Introduction

According to the Occupation Safety and Health Administration (OSHA), for example, the construction industry is responsible for more than 20% of worker fatalities in the United States [1]. In the United Kingdom, for example, a similar scenario occurs where construction accounts for the greatest number of fatalities across all sectors [2]. The industry is having to work with less experienced people with equipment in high demand, which generates fresh hazards in a business that is inherently risky.

To address this problem, hazard analysis is typically undertaken prior to construction, which may be performed using manual methods and/or three-dimensional (3D) models [3-4]. Yet hazards can change once construction commences and their identification then needs to be undertaken manually, which can be a labour-intensive and time-consuming process. To overcome the drawbacks of manually identifying hazards, computer vision-based approaches have been developed (e.g., [5-10]). For example, Fang et al. (2018a) [8] proposed a combination of Faster R-CNN and a classification network to identify workers when they were not wearing their safety harness while working at heights. Similarly, Fang et al. (2019) [10] developed a computer vision approach with Mask R-CNN to identify worker traversing on structural supports. Usually, one computer vision algorithm is used to identify a single safety hazard in a scene. For example, computer vision has been widely used to determine if an individual is wearing their safety helmet. Realistically, however, a safety hazard involves a combination of conditions. Therefore, there is a need to move away from using a single computer vision algorithm to a position where a number of safety hazards can be simultaneously detected.

To address this problem, a novel approach that combines computer vision and ontology to automatically detect safety hazards is proposed. The goal of our research is to determine whether hazards with complex rules can be identified using our proposed semantic vision-based approach. Ontology is used to enable computer applications to easily represent and reason safety knowledge, and with computer vision being used to automatically detect objects and extract attributes from images (i.e., classes and distance). The paper commences by providing a review of computer vision-based object detection approaches and applications of ontology-based risk management that have been developed in construction.

2. Related Work

2.1 Computer Vision-based Object Detection

Computer vision has been utilized to undertake a number of tasks in construction such as productivity analysis [11], progress monitoring [12], and the recognition of unsafe behaviour [8-10]. Vision-based identification of different objects is an innate feature of vision-based applications. Vision-based object detection within the domain of construction has focused on utilizing the following approaches: (1) handcrafted features; and (2) deep learning.

Handcrafted feature-based methods tend to employ a two-stage procedure. Firstly, features from images and videos are extracted by descriptors such as Histogram of Oriented Gradients (HOG) [13] or Histogram of Optical Flow (HOF) [14]. Secondly, these extracted features are put into a classifier such as Support Vector Machine (SVM). For example, Memarzadeh (2012) [15] combined a HOG and color features with a new multiple binary SVM classifier to automatically detect and distinguish between a person and equipment using videos. Despite the success of hand-crafted feature-based approaches they are manually created and therefore there is a trade-off between detection accuracy and computational efficiency (i.e., speed) arises.

Deep learning has provided the ability to automatically extract and learn features in an end-to-end manner from images with higher levels of accuracy [16]. Several studies have demonstrated the potential of CNN's for object detection and action recognition on construction sites [9, 17-18]. For example, Fang et al. (2018b) [9] developed an improved Faster R-CNN to identify objects from images and have achieved an accuracy with 91% and 95% when detecting individuals and excavators, respectively.

A review of computer vision-based studies in construction reveals that they have achieved acceptable levels of accuracy (i.e., precision, recall) on object detection and attributes (e.g. distance). For example, Kim et al. (2017) [19] applied a transformation matrix to determine the distance between objects from a single image. Drawing on the work of Kim et al. (2017) [19] and the research of Fang et al. (2018b) [9] an improved Faster R-CNN is selected as an approach to detect objects from two-dimensional (2D) images, and a transformation matrix [19] is used to compute the distance of objects from a single image.

2.2. Ontology-based Risk Knowledge Management

Ontology is a formal conceptualization of knowledge, which is a simplified view of a domain that describes objects, concepts, and relationships between them. Semantic Web technology, for example, can allow various sources of information to be made available in a format that can be searched and retrieved from the Internet [20]. Thus, the combination of semantic web technology with ontology can enable lots of advantages to be realized, such as, knowledge extending or changing. [20-22]. Ontologies-based approaches have been extensively applied to numerous aspects of construction, such as energy management [23-24], and risk management [25]. For example, Zhou *et al*. (2016) [25] reviewed ontology-based research to identify gaps in knowledge within the extant literature.

The aforementioned studies demonstrate the potential of ontology technology in supporting risk management, especially as it can be used to raise the level of safety awareness. By organizing knowledge as a logical semantic expression, it can be shared using ontology technologies and therefore enable semantic interoperability. As a result, the structured and unified knowledge in the ontology can be understood and readily operated by different parties and computer applications and thus enable the re-use of knowledge to be promoted. However, to the best of our knowledge, there has been no research that has combined computer vision and ontology to identify hazards on construction sites.

3. Research Approach

Our proposed semantic computer vision-based approach comprises four procedures: (1) ontological model for construction hazards; (2) computer vision-based entity and attributes detection; (3) extraction of spatial-relationship from images; and (4) reasoning data for hazard identification with a customized graph database. Figure 1 presents the workflow of our proposed hybrid semantic computer vision-based approach, which integrates computer vision and ontology. The procedure to implement our approach is described in the following steps that are described in further detail in the sections hereinafter.

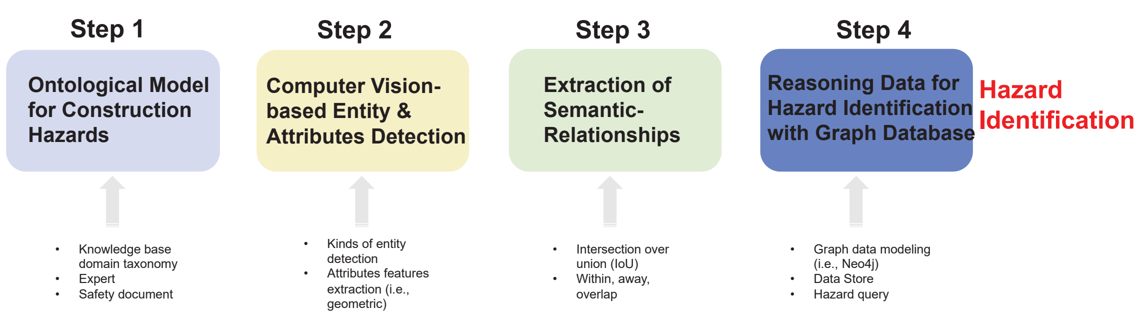

Fig. 1. Workflow of proposed hybrid semantic computer vision approach

Step 1: Ontological Model for Hazards

The initial process for implementing our semantic computer vision-based hazard identification model was to develop a taxonomy of hazards. The Chinese code for 'Quality and Safety Inspection Guide of Urban Rail Transit Engineering', for example, was selected as a point of reference to examine hazards for a metro-rail project that is being constructed in Wuhan, China.

Based on the taxonomies, the 'HowNet' structure was further extended, and then applied to the ontological model used for identifying hazards. HowNet is "an on-line common-sense knowledge base, unveiling inter-conceptual relations and inter-attribute relations of concepts as connoting in Chinese and English bilingual lexicons" [26-27]. Within the context of construction, a hazard can be defined by its: given time and space, and the entities (with specific attributes) undergo certain activities [28-29]. Thus, a hazard event consists of semantic information that specifies its: (1) entity; (2) activity; (3) location; (3) time; (4) attribute.

Step 2: Computer Vision-based Entity and Attributes Detection

The goal of our research is to develop a computer vision approach that can be used to early warn construction hazards. With this in mind, we use computer vision to identify contextual information from construction sites, including hazard entities and its attributes. Thus, the following types of information (e.g., classes of entities and their attributes) are extracted to identify hazards by using reasoning modelling.

- Entity Detection

Entities can be divided into four types of objects: (1) people; (2) equipment; (3) materials; and (4) environment. In this research, two types of detection approaches are used: (1) object detection; and (2) scene recognition. Here, an object detection approach is used to detect people, equipment (i.e., excavator), and materials (e.g., structural support). The scene recognition approach, one of the hallmark tasks of computer vision, allowing defining a context for object recognition, is used to detect working at a height and foundation pit. In this research, built on our previous study [9], Faster R-CNN is applied to detect an individual, excavator, and structural support. In addition, a Unified Perceptual Parsing approach (UPP) is used to effectively segment concepts from images that can be used to recognize scenes [30].

- Attributes Extraction

As our research focuses on identifying hazards based on distance and spatial features, we only need to extract two types of attributes: (1) the coordinates of each object in the image; and (2) their distance between objects. We, therefore, utilized the transformation matrix [19] within our hybrid semantic computer vision model to compute distances between objects.

Step 3: Extraction of Spatial-Relationships from Images

After identifying an object's types and their attributes, three spatial relationships between the objects can be computed: (1) within; (2) overlap; and (3) away. An example of a spatial relationship is presented in Figure 2. In this research, the choice of terminology and semantics for the spatial relation is based on the distance between objects (between two geometries a and b) according to the safety rules that were established from the prevailing Chinese codes.

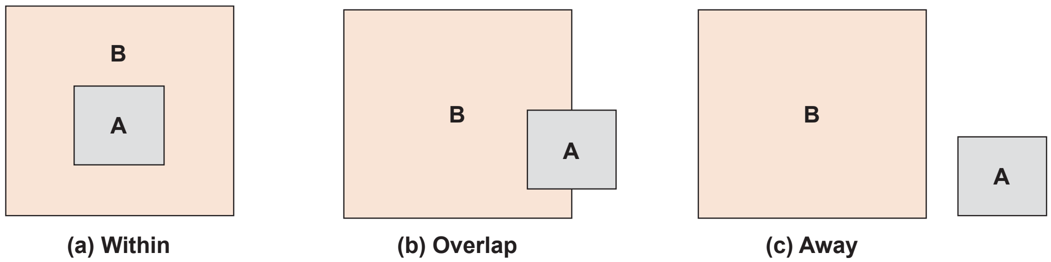

Fig. 2 Examples of spatial relationship

The spatial relationship of object A and object B is defined as the IoU of the bounding box A and bounding box B, as shown in Eq. [1]:

$$IoU(A,B)=\frac{area(A)\cap area(B)}{\min\left\{area(A),area(B)\right\}}=\begin{cases}1 & within\\ [0,1] & overlap\\ 0 & away\end{cases}$$

Eq. [1]

Step 4: Reasoning Data for Hazard Identification Using Graph Database

We use a graph database, Neo4j graph database, to present the safety knowledge for reasoning in a highly accessible way [31-32]. To automatically identify hazards, we perform the following tasks: (1) data modelling; and (2) automated reasoning and query.

- Data Modelling

The procedure to extract objects classes and their spatial relationships have been described above. The outputs from these procedures are saved as a ".CSV" file and loaded in the Neo4j database. The Neo4j database automatically processes the data and then provides an output.

- Automated Reasoning and Query

On completion of the modelling process in the graph database, the final step is to identify the hazards by querying the unsafe behaviour rules that had been defined in the model. The as-built graph database is constructed based on the objects and their spatial relationship, the unsafe rules are derived from the safety codes, which were re-defined as the queries.

4. Case Study

To demonstrate and test the validity of our developed semantic model, we can focus on identifying the unsafe condition that may lead to FFH (Table 2). We have selected an urban subway system under construction in Wuhan China to evaluate the effectiveness of detection for the developed semantic approach.

Table 1. Checklist of unsafe behaviour related to FFH.

Number	Unsafe behaviour description
1	There should be no more than two people in a man basket
2	Workers should not walk on the support of excavation if there has no guardrail
3	Edges of excavations (over 2m deep) should be protected with a guardrail
4	People should not stand on machinery when hoisting
5	People should wear a safety harness when working above a certain height
6	There should not use car hopper to pick up people

4.1 Development *of Ontology for FFH*

A taxonomy of hazards related to FFH was developed based on the checklist in Table 2. The core concepts identified are classified and serve as an extension to the taxonomy. For example, a hazard is "There should be no more than two people in a construction man basket". Thus, the hazard entity is 'people' and 'construction man basket'. Activity is 'stand'. Attribute is 'number' and coordinate'. The relationship is 'overlapped/within'.

4.2 Hazard Identification Results

We initially used computer vision to detect objects and their attributes. Individuals, structural support, foundation pit are identified (Figure 3(b) and Figure 3(c)). The spatial relationships of objects are recognized using IoU and distance between objects (Figure 3(c)). As previously mentioned, the results are stored in the Neo4j database (Figure 3(d)) in order to identify unsafe conditions using rule "MATCH (x:laborer)-[r:touch]-(y:structure) RETURN x,r,y" (Figure 3(e)).

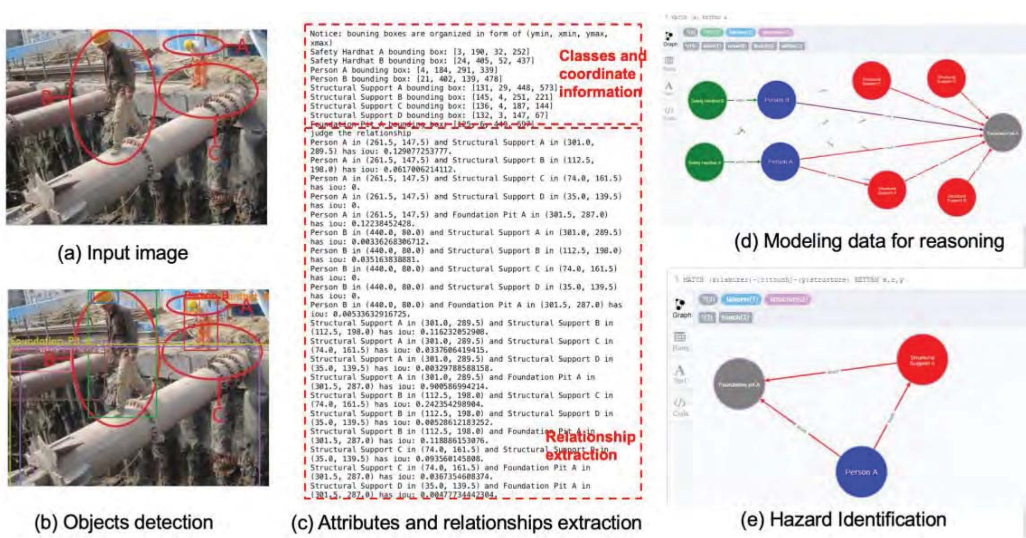

Fig. 3. Semantic computer vision detection results

5. Conclusion, limitation, and future work

We have introduced a novel semantic model to combine computer vision and ontology to automatically identify hazards from images. We utilized the following tools to develop our model: (1) computer vision algorithms which are used to detect objects and attributes; and (2) ontological reasoning to identify unsafe conditions based on the identified distance and spatial information. To validate our approach, a database of FFH from several construction sites is used as a case in our research. It was revealed that our semantic model can accurately recognize safety hazards from images with complex rules. We also suggest that our proposed semantic model can be used by site management to automatically identify potential hazards and therefore put in place strategies to mitigate injuries and accidents.

Despite its success of our approach, it should be acknowledged that several limitations exist. Firstly, our research relied on distance and coordinate information to extract spatial relationship for reasoning hazards. In fact, many hazards comprise safety rules with features. Secondly, our research extracts the coordinates and the distance between objects from 2D images and then obtains spatial-relationship in accordance with the information obtained (i.e., coordinate, distance). Mistakes can be made when using the transformation matrix to compute the distance of objects from single images. Finally, we have also assumed that a variety of objects can be accurately detected by Faster R-CNN. However, if an object is occluded or lack of an available database for training object detection model, the error rate of object detection may be high.

Our future research will focus on: (1) combining temporal information and spatial information to identify construction hazards from videos streaming; (2) using stereo camera to collect data, and compute 3D depth information from stereo videos; (3) combing other information techniques and computer vision to extract more features, such as, size of foundation pit, and colour of safety hardhat, for identifying more types of hazards.

Acknowledgements

The authors would like to acknowledge the financial support provided by the National Natural Science Foundation of China (Grant No.71732001, No.51678265, No.51878311) and the China Scholarship Council (CSC).

References

[1] OSHA, Commonly Used Statistics, https://www.osha.gov/oshstats/commonstats. html2014 (accessed January 17, 2015).

[2] Edwards, H. (2017). "Fatal injuries arising from accidents at work in Great Britain 2017."

[3] Hadikusumo, B.H.W., and Rowlinson, S., (2002) "Integration of virtually real construction model and design-for-safety-process database". Automation in Construction. 11(5), 501-509. https://doi.org/10.1016/S0926-5805(01)00061-9.

[4] Martinez-Aires, M., Lopez-Alonso, M., Martinez-Rojas, M., (2018) "Building information modeling and safety management: A systematic review". Safety Science, 101, 11-18. https://doi.org/10.1016/j.ssci.2017.08.015.

[5] Han, S., and Lee, S. (2013). "A vision-based motion capture and recognition framework for behavior-based safety management". Automation in Construction, 35, 131-141. https://doi.org/10.1016/j.autcon.2013.05.001.

[6] Han, S., Lee, S., and Peña-Mora, F. (2014). "Comparative Study of Motion Features for Similarity-Based Modeling and Classification of Unsafe Actions in Construction". Journal of Computing in Civil Engineering, 28, 10.1061/(ASCE)CP.1943-5487.0000339, A4014005.

[7] Ding, L., Fang, W., Luo, H., Love, P.E.D., Zhong, B., and Ouyang, X. (2018). "A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory". Automation in Construction, 86, 118-124. https://doi.org/10.1016/j.autcon.2017.11.002.

[8] Fang, W., Ding, L., Luo, H., Love, P. E. D., and Zhong, B. (2018a). "Falls from Heights: A Computer Vision-based Approach for Safety Harness Detection". Automation in Construction. 91, 53-61. https://doi.org/10.1016/j.autcon.2018.02.018.

[9] Fang, W., Ding, L., Zhong, B.T., P.E.D., and Luo, H. B., (2018b). "Automated Detection of Workers and Heavy Equipment on Construction Sites: A Convolutional Neural Network Approach". Advanced Engineering Informatics, 37, 139-149. https://doi.org/10.1016/j.aei.2018.05.003.

[10] Fang, W., Zhong, B., Zhao, N., Love, P.E.D., Luo, H., Xue, J., Xu, S. (2019). A deep learning-based approach for mitigating falls from a height with computer vision: Convolutional neural network. Advanced Engineering Informatics. 39: 170-177. https://doi.org/10.1016/j.aei.2018.12.005.

[11] Gong, J., and Caldas, C. H. (2009). "Computer vision-based video interpretation model for automated productivity analysis of construction operations". Journal of Computing in Civil Engineering, 24, 252-263. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000027.

[12] Han, K. K., and Golparvar-Fard, M. (2015). "Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs". Automation in Construction, 53, 44-57. https://doi.org/10.1016/j.autcon.2015.02.007.

[13] Dalal, N., and Triggs, B. (2005). "Histograms of Oriented Gradients for Human Detection". IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005. (pp. 886-893). https://doi.org/10.1109/cvpr.2005.177.

[14] Wang, H., and Schmid, C. (2014). "Action Recognition with Improved Trajectories". In IEEE International Conference on Computer Vision (pp. 3551-3558). https://doi.org/10.1109/iccv.2013.441.

[15] Memarzadeh, M., Heydarian, A., Golparvarfard, M., and Niebles, J. C. (2012). "Real-Time and Automated Recognition and 2D Tracking of Construction Workers and Equipment from Site Video Streams". International Conference on Computing in Civil Engineering (2012) (pp. 429-436), June 17-20, Florida, United States. https://doi.org/10.1061/9780784412343.0054.

[16] LeCun, Y., Bengio, Y., Hinton, G., "Deep learning", Nature, 521, 436-444.

[17] Luo, X.C., Li, H., Cao, D.P., Yu, Y., Yang, X.C., Huang, T., (2018a) "Towards efficient and objective work sampling: Recognizing workers' activities in site surveillance videos with two-stream convolutional network". Automation in Construction, 94, 360-370. https://doi.org/10.1016/j.autcon.2018.07.011

[18] Luo, X.C., Li, H., Cao, D.P., Dai, f., Seo, JO., Lee, SH., (2018b)"Recognizing Diverse Construction Activities in Site Images via Relevance Networks of Construction-Related Objects Detected by Convolutional Neural Networks", 32(3): 04018012, 10.1061/(ASCE)CP.1943-5487.0000756.

[19] Kim, K., Kim, K., Kim, K., (2017). "Image-based construction hazard avoidance system using augmented reality in a wearable device". Automation in construction, 83, 390-403. https://doi.org/10.1016/j.autcon.2017.06.014.

[20] Elghamrawy, T., Boukamp, F., and Kim, H. S. (2009). "Ontology-Based, Semi-Automatic Framework for Storing and Retrieving On-Site Construction Problem Information — An RFID-Based Case Study". In Construction Research Congress (pp. 457-466). https://doi.org/10.1061/41020(339)47.

[21] Anumba, C. J., Issa, R. R. A., Pan, J., and Mutis, I. (2008). "Ontology-based information and knowledge management in construction". Construction Innovation, 8, 218-239. https://doi.org/10.5220/0003053903310334.

[22] Ding, L. Y., Zhong, B. T., Wu, S., and Luo, H. B. (2016). "Construction risk knowledge management in BIM using ontology and semantic web technology". Safety Science, 87, 202-213. https://doi.org/10.1016/j.ssci.2016.04.008.

[23] Corry, E., Pauwels, P., Hu, S.S., Keane, M., O'Donnell, J., (2015) "A performance assessment ontology for the environmental and energy management of buildings", Automation in Construction, 57, 249-259. http://dx.doi.org/10.1016/j.autcon.2015.05.002

[24] Hippolyte J.L., Rezgui, Y., Li, H., Jayan, B., Howell, S., (2018) "Ontology-driven development of web services to support district energy applications". Automation in Construction, 86, 210-225. http://dx.doi.org/10.1016/j.autcon.2017.10.004.

[25] Zhou, Z.P, Goh, Y.M., Shen, L.J. (2016). "Overview and Analysis of Ontology Studies Supporting Development of the Construction Industry". J. Computing in Civil Engineering, 30(6), 10.1061/(ASCE)CP.1943-5487.0000594, 0406026.

[26] Dong, Z. and Dong, Q. (2003). HowNet – a hybrid language and knowledge resource. Proceedings of the International Conference on Natural Processing Language and Knowledge Engineering, 26-29th October, IEEE, Beijing, China, 10.1109/NLPKE.2003.1276017

[27] Dong, Z., Dong, Q., (2006) "HowNet and the computation of meaning", World Scientific Publishing Co., Inc., https://doi.org/10.1142/9789812774675.

[28] Goh, Y. M., and Chua, D. K. H. (2010). "Case-Based Reasoning Approach to Construction Safety Hazard Identification: Adaptation and Utilization." Journal of Construction Engineering and Management, 136(2): 170-178. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000116.

[29] Guo, B. H. W., and Goh, Y. M. (2017). "Ontology for the design of active fall protection systems." Automation in Construction, 82(Supplement C), 138-153. https://doi.org/10.1016/j.autcon.2017.02.009.

[30] Xiao, T., Liu, Y.C., Zhou, B.L., Jiang, Y.N., Sun, J., (2018) "Unified Perceptual Parsing for Scene Understanding" Lecture Notes in Computer Science, 432-448. https://doi.org/10.1007/978-3-030-01228-1_26.

[31] Johnpaul, C.I., Mathew, T. (2017). "A cypher query based NoSQL data mining on protein databases using Neo4j graph database" 4th International Conference on Advanced Computing and Communication Systems. Coimbatore, India. https://doi.org/10.1109/ICACCS.2017.8014558.

[32] Guia, J., Soares, v. G., Bernardina, S., (2017) "Graph Database: Neo4j Analysis", Proceedings of the 19th International Conference on Enterprise Information System, 351-356, Porta, Portugal. https://doi.org/10.5220/0006356003510356.