Post-print of: Silva de Oliveira C., Sanin C., Szczerbicki E. (2020) Image Representation for Cognitive Systems Using SOEKS and DDNA: A Case Study for PPE Compliance. In: Nguyen N., Jearanaitanakij K., Selamat A., Trawiński B., Chittayasothorn S. (eds) Intelligent Information and Database Systems. ACIIDS 2020. Lecture Notes in Computer Science, vol 12033. Springer, Cham
DOI: 10.1007/978-3-030-41964-6_19

Image Representation for Cognitive Systems using SOEKS and DDNA: A case study for PPE compliance

Caterine Silva de Oliveira[1], Cesar Sanin[1] and Edward Szczerbicki[2]

[1] The University of Newcastle, Newcastle NSW, Australia
caterine.silvadeoliveira@uon.edu.au,
cesar.maldonadosanin@newcastle.edu.au
[2] Gdansk University of Technology, Gdansk, Poland
edward.szczerbicki@newcastle.edu.au

Abstract. Cognitive Vision Systems have gained significant interest from academia and industry during the past few decade, and one of the main reasons behind this is the potential of such technologies to revolutionize human life as they intend to work under complex visual scenes, adapting to a comprehensive range of unforeseen changes, and exhibiting prospective behavior. The combination of these properties aims to mimic the human capabilities and create more intelligent and efficient environments. Nevertheless, preserving the environment such as humans do still remains a challenge in cognitive systems applications due to the complexity of such process. Experts believe the starting point towards real cognitive vision systems is to establish a representation which could integrate image/video modularization and virtualization, together with information from other sources (wearable sensors, machine signals, context, etc.) and capture its knowledge. In this paper we show through a case study how Decisional DNA (DDNA), a multi-domain knowledge structure that has the Set of Experience Knowledge Structure (SOEKS) as its basis can be utilized as a comprehensive embedded knowledge representation in a Cognitive Vision System for Hazard Control (CVP-HC). The proposed application aims to ensure that workers remain safe and compliant with Health and Safety policy for use of Personal Protective Equipment (PPE) and serves as a showcase to demonstrate the representation of visual and non-visual content together as an experiential knowledge in one single structure.

Keywords: Cognitive Vision Systems, Knowledge Representation, SOEKS, DDNA, PPE compliance, Hazard Control

1 Introduction

Cognitive Vision Systems have gained considerable interest from academia and industry during the past few decade, and one of the main reasons behind this is the potential of such technologies to revolutionize human life as they intend to work under complex visual scenes, adapting to a comprehensive range of unforeseen changes, and exhibiting prospective behavior [1]. The combination of these properties aims to mimic the human capabilities and create more intelligent and efficient environments [2].

Nonetheless, preserving the environment such as humans do still remains a challenge in cognitive systems applications due to the complexity of such process. It involves understanding the context and gathering visual and other sensorial information available and translating it into knowledge to be useful. Moreover, past experiences also plays an important role when it comes to perception [3] and must also be considered as an important element in this process. Smart cognitive systems that have been proposed so far oversight the potential of using these experiences to enrich the application with smartness while, at the same time, creating decisional fingerprints. This would allow the system knowledge growth through daily operation autonomously, just like human experience do in real life [4].

Experts believe the starting point towards real cognitive vision systems is to establish a representation which could integrate image/video modularization and virtualization, together with information from other sources (wearable sensors, machine signals, context, etc.) and capture its knowledge. In this context, Decisional DNA (DDNA), a multi-domain knowledge structure based on experience, has been extended to the visual domain to be used as a comprehensive embedded knowledge representation for Cognitive Systems [5]. DDNA has the Set of Experience Knowledge Structure (SOEKS) [6] as its basis and allow the creation of a multi-modal space composed of information from different sources, such as contextual, visual, auditory etc., in a form of a structure and explicit experiential knowledge [7].

The applicability of such representation have been tested over a Cognitive Vision Platform for Hazard Control (CVP-HC). The CVP-HC is scalable yet adaptable platform capable of working in a variety of video analysis scenarios whilst meeting specific safety requirements of industries [8]. This platform aims to assist the safety management process in industrial environments, and the special case of PPE compliance is presented in this paper.

This paper is organized as follow: In Section 2, some fundamental concepts are presented, including the evolution of systems towards augmented cognitive technologies and the challenge of representation and management of knowledge in these systems. The proposed representation based on SOEKS and DDNA is also explained. In Section 3 a case study for the case of PPE compliance is presented, including its applicability, design and experimental results achieved so far. Finally, in Section 5 conclusions and future work is presented.

2 Fundamental Concepts

In order to offer a more complete view, we briefly introduce concepts that have driven the proposed research as well as the technologies involved.

2.1 From Computer Vision to Cognitive Vision Systems

The use of computer vision techniques can support automatic detection and tracking of objects and people with reasonable accuracy [9-13]. Visual sensing facilities, such as video cameras can gather a large amount of data, such as video sequences or digitized visual information that, with support of machine learning technologies and powerful machines, can operate in real time [14]. For those reasons, computer vision systems have been a research focus for a long time in surveillance systems, human detection, and tracking.

However, computer vision systems have their own inherent limits, especially those whose task is to work in unidentified environments and deal with unknown scenarios and specifications. Besides the significant improvements in computer vision technologies, they are still challenged by issues such as occlusion or position accuracy; and background changes result in the necessity of adapting the algorithms for different conditions, clients and situations. To date, the creation of a general-purpose vision system with the robustness and resilience comparable to human vision still remains a challenge [13].

In this context, methods incorporating prior knowledge and context information have gained interest. The understanding about scene composition in an image (which set of objects are present) can improve recognition performance about the scene where they are inserted [15]. For instance, the presence of multiple cutlery items in an image can aid the recognition of a kitchen image. This relationship is held both ways, as contextual knowledge can also offer insights about the function of an object in a scene, reducing the impacts of sensor noise or occlusions [16]. These technologies are known as knowledge-based systems. For instance, an automatic semantic and flexible annotation service able to work in a variety of video analysis with little modification to the code using Set of Experience Knowledge Structure (SOEKS) was proposed in work by Zambrano et al. [17]. This system is a pathway towards cognitive vision and it is composed, basically, by the combinations of detection algorithms and an experience based approximation.

The design of a general-purpose vision system with the robustness and resilience of the human vision is still a challenge. One of the latest trends in computer vision research to mimic the human-like capabilities is the joining of cognition and computer vision into cognitive computer vision. Cognitive Systems have been defined as “a system that can modify its behavior on the basis of experience” [18]. Although, most experts tend to agree that such systems only exists in theory, that is, systems that can independently process, reason and create in the same capacity as the human brain has not yet been implemented successfully [19].

In this scenario, the concept of Augmented Intelligence, also known as Cognitive Augmentation or Intelligence Amplification (IA) comes into play [20]. For any specific application humans being and machines have both their own strengths and weaknesses. Machines are very efficient in numerical computation, information retrieval, statistical reasoning, with almost unlimited storage. Machines can capture many categories of information from the environment through various sensors, such as range sensors, visual sensors, vibration sensors, acoustic sensors, and location sensors [21]. On the other hand, humans have their own cognitive capabilities which includes consciousness, problem-solving, learning, planning, reasoning, creativity, and perception. These cognitive functions allows humans to learn from last experiences and use this experiential knowledge to adapt to new situations and to handle abstract ideas to change their environment. Therefore, the combination of both human experiential knowledge and information collected by a system can be used to enhance smartness of systems and for improved decision making [22]. Fig. 1 shows the steps towards Augmented Cognitive Vision and a synthesis of components involved in each stage.

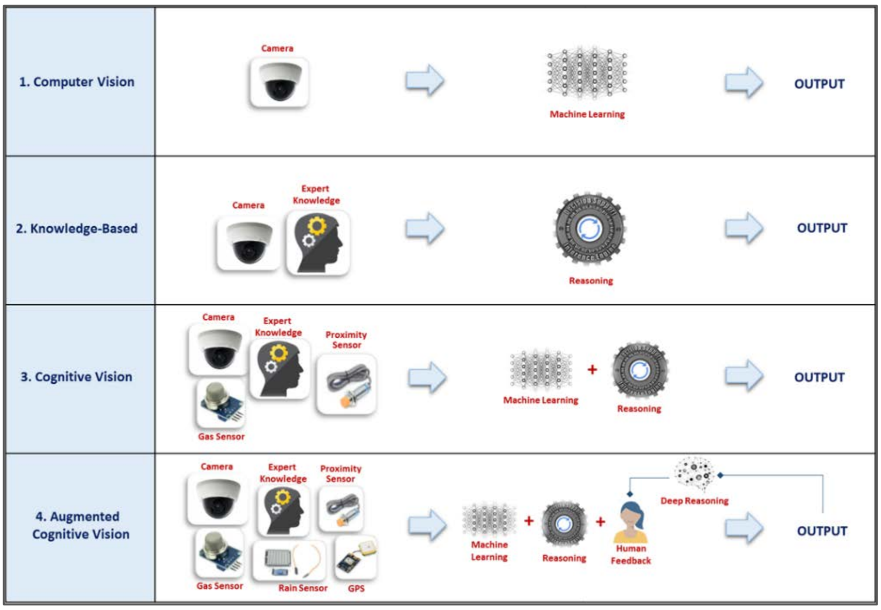


Fig. 1. Steps towards Augmented Cognitive Vision.

2.2 Knowledge Representation for Cognitive Systems

The implementation of cognitive vision systems require the design of functionalities for knowledge engineering (acquisition and formalism), recognition and categorization, reasoning about events for decision making, and goal specification, all of which are concerned with the semantics of the relationship between the visual agents and their environments i.e. context [23]. These functionalities direct cognitive vision systems towards purposeful behavior, adaptability, anticipation, such as human beings.

In this context, knowledge and leaning are central to cognitive vision. To be readily articulated, codified, accessed and shared, knowledge must be represented in an explicit and structured way [24]. In addition, the choice of a suitable representation greatly facilitates obtaining methods that efficiently learn the relevant information available. Therefore, an appropriate knowledge representation is crucial for the success in designing of cognitive systems.

Nevertheless, most approaches that have been proposed on past years, even though they present some principles for intelligent cognitive vision, they fail in providing a unique standard that could integrate image/video modularization, its virtualization, and capture its knowledge [6]. To address these issues an experience-based technology that allows a standardization of image/video and the entities within together with any other information as a multi-source knowledge representation (required for the further development of cognitive vision) without limiting their operations to a specific domain and/or following a vendor's specification has been proposed [25]. This representation supports mechanisms for storing and reusing experience gained during cognitive vision decision-making processes through a unique, dynamic, and single structure called Decisional DNA (DDNA) [5]. DDNA makes use of Set of Experience (SOE) in an extended version for the use of storing formal decision events related to image and video. DDNA and SOE provide a knowledge structure that has been proven to be multi-domain independent [7].

Set of Experience Knowledge Structure (SOEKS) and Decisional DNA (DDNA). The Set of Experience Knowledge Structure (SOEKS) is a knowledge representation structure created to acquire and store formal decision events in a structured and explicit way. It is composed by four key elements: variables, functions, constraints, and rules. Variables are commonly used to represent knowledge in an attribute-value form, following the traditional approach for knowledge representation. Functions, Constraints, and Rules of SOEKS are ways of relating variables. Functions define relationships between a set of input variables and a dependent variable; thus, SOEKS uses functions as a way to create links among variables and to build multi-objective goals. Constraints are functions that act as a way to limit possibilities, limit the set of possible solutions and control the performance of the system in relation to its goals. Lastly, rules are relationships that operate in the universe of variables and express the condition-consequence connection as "if-then-else" and are used to represent inferences and associate actions with the conditions under which they should be implemented [6]. Rules are also ways of inputting expert knowledge into the system. The Decisional DNA consists is a structure capable of capturing decisional fingerprints of an individual or organization and has the SOEKS as its basis. Multiple Sets of Experience can be collected, classified, organized and then grouped into decisional chromosomes, which accumulate decisional strategies for a specific area of an organization. The set of chromosomes comprise, finally, what is called the Decisional DNA (DDNA) of the organization [5].

3 Case Study: PPE Safety Compliance

Hazards are present in all workplaces and can result in serious injuries, short and long-term illnesses, or death [26]. Reports HSE UK report has shown that over 80% of reported workplace injuries are sustained due to a person not wearing correct protective clothing [27]. In this context, the verification of PPE compliance becomes essential in the management of safety to ensure the occupational health of workers. Technologies to support its practical and automated implementation have emerged as a need, but the current technologies available still face considerable limitations [9, 13, 15, 28].

The combination of vision and sensor data together with the resulting necessity for explicit and formal representations builds a central element of an autonomous system for detection and tracking of laborers in workplaces environments. To be able to perform in a variety of plants and scenes, making sure employees remain safe and compliant with Health & Safety policy without the necessity of recoding the application for each specific case scenario, the system must be adaptable and perceive the environment as automatically as possible and change its behavior accordingly. However, computer vision systems have their own inherent limits, especially those whose task is to work in unidentified environments and deal with unknown scenarios and specifications [29].

The gaps of current systems may be filled by connecting the probabilistic area of detection of events with the logical area of formal reasoning in a Cognitive Vision Platform for Hazard Control (CVP-HC) [29]. This platform verifies the PPE compliance in variety of video analysis scenarios whilst meeting specific safety requirements of industries [25]. The proposed system is based on the Set of Experience Knowledge Structure (SOEKS or SOE in short) and Decisional DNA (DDNA).

3.1 Applications

Automated verification of PPE compliance can be useful in a variety of industries (e.g. Oil & Gas, Manufacturing & Production, Construction, Engineering, Pharmaceuticals, etc.) and applied in a range use case scenarios to ensure employees remain safe [30]. Below we exemplify two main applications that the proposed solution can address.

Access Control. With cameras positioned above an entrance/exit of a site or facility, the system is able to visually verify that laborers are wearing the protective equipment according to the safety requirements of that industry/area before allowing entry. In case of any equipment being missed at the point of entry, then the system will not permit a gate to open and will advise which items must be worn in order to enable access. Once all the mandatory equipment are detected the access is granted. The visual information from the cameras can be combined with other sensor data to give extra information about crucial required equipment (e.g. oxygen mask when oxygen level is critically low).

Continuous Monitoring. Another solution can address the continuous monitoring of works by the use of cameras and other sensor data covering the site or facility to ensure that employees remain wearing the required PPE in a given context. If laborers remove a required equipment then the system will recognize this in real-time and carry out an action based on a set of given preferences or recommendations. For instance, an alert can be sent directly to the employee or manager for correction on site; the event can be logged for future reports and analysis, etc. If sensors detect any abnormality, which changes the status of the required equipment, workers can also be advised of that for a quick action.

3.2 Representation of variables, constrains, functions and rules

For the case study in analysis, a set of variables, functions, constrains and rules are represented as a Set of Experience Knowledge Structure (SOEKS). SOEKS allows the representation, use, storing and retrieval of visual and non-visual knowledge content together in one single standardized structure [25].

Variables. The variables in our system are composed by each image/frame being analyzed, body parts of workers, and annotations of each Personal Protective Equipment (PPE). In addition, we include, as part of the set of variables, the calculation of area of intercept A_I between the bounding boxes containing a body part and a corresponding PPE, as well as the area of each PPE in the scene, which is defined respectively by:

$$A_I = \begin{array}{c} \boldsymbol{max}(0, \boldsymbol{min}(ppe_{xmax}, bp_{xmax}) - \boldsymbol{ma\,x}(ppe_{xmin}, bp_{xmin})) * \\ \boldsymbol{max}\big(0, \boldsymbol{min}\big(ppe_{ymax}, bp_{ymax}\big) - \boldsymbol{ma\,x}\big(ppe_{ymin}, bp_{xmin}\big)\big) \end{array} \quad (1)$$

$$A_{ppe} = \Big((ppe_{xmax} - ppe_{xmin}) * \big(ppe_{ymax} - ppe_{ymin}\big)\Big) \quad (2)$$

Finally, the last two variables considered are: the dependent variable resulting from the creation of the overlap function $O_{I,ppe}$ (eq. 3), and the safety status of scene, to be defined by the set of rules. Both variables will explained in the following subsections.

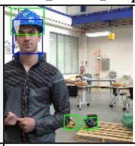
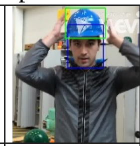
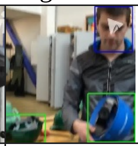
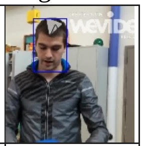
Functions. As defined before, function establishes relationships among input and dependent variables as a way to find more elements of decision-making that reduce the possibility of duality, while facilitating knowledge elicitation [cite Cesar' Thesis]. In our application, for each body part of a person detected there may be a range of compatible surrounding PPEs that can be associated with it, including ones belonging to other people in the scene. For instance, let's imagine a scene where two people are being detected, one is wearing a respirator and another one is not (the second's person respirator is placed next to them, on the floor). In this case we have four interceptions being computed and inputted into the system, producing different states in relation to the safety status of the scene. In this case, it is necessary to reduce the possibilities of duality in finding an optimal unique set of variables that identifies a unique state while reducing ambiguity [cite Cesar' Thesis]. Therefore, we calculate the overlap between the areas of intercept A_I and PPEs A_{ppe} as a function, which objective is to maximize the area of overlap, associating the PPE to the closest conforming body part.

The maximum overlap $O_{I,ppe}$ between intercept and corresponding PPE goes from 0 (disjoint) to 1 (complete overlap) is calculated as following:

$$O_{I,\mathrm{ppe}} = \left\{[max]\,\frac{A_I}{A_{ppe}}\right\} \quad (3)$$

Table 1 shows values of maximum $O_{I,\mathrm{helmet}}$ for a sequence of frames and the status of *wearing/not wearing* associated with them.

Table 1. Examples of [max] $O_{I,\text{helmet}}$ and respective wearing/not wearing status.

Frame					
$[\boldsymbol{max}]\boldsymbol{O}_{\boldsymbol{I},\textbf{helmet}}$	0.49	0.44	0.40	0.00	0.00
Wearing helmet?	YES	YES	YES	NO	NO

Constraints. In our analysis, we only consider the XY plane, i.e. no depth information is taking into consideration. When not taking the Z plane, protective equipment on the background may be wrongly associated with the body parts even being meters distant on the depth plane and vice versa. To minimize the set of possible misleading associations of body parts and PPEs that are distant from each other on the Z plane, we create a set of constraints. These constraints restrict the possible size of the PPE that can be associated with each body part being detected.

Rules. To ensure flexibility and as well as to attend each specific requirements of different industries and scenarios, a set of rules is created. These rules are also a way of allowing expert knowledge to be included in the system reasoning as they can be easily changed and adjusted to attend specific requisites and situations. For this analysis in specific, the following set of rules are considered:

Rule 1:
IF $O_{I,respirator}$ > threshold
  THEN safety_status = SAFE
ELSE safety_status = UNSAFE

Rule 2:
IF $O_{I,helmet}$ > threshold
  THEN safety_status = SAFE
ELSE safety_status = UNSAFE

Rule 3:
IF $O_{I,respirator}$ > threshold & $O_{I,helmet}$ > threshold
  THEN safety_status = SAFE
ELSE safety_status = UNSAFE

Rule 4:
IF $O_{I,hivis}$ > threshold & $O_{I,boot}$ > threshold
  THEN safety_status = SAFE
ELSE safety_status = UNSAFE

Rule 5:
IF $O_{I,respirator}$ > threshold & $O_{I,helmet}$ > threshold & $O_{I,googles}$ > threshold
  THEN safety_status = SAFE
ELSE safety_status = UNSAFE

Rule 6:
IF $O_{I,harness}$ > threshold & $O_{I,helmet}$ > threshold & $O_{I,glove}$ > threshold
  THEN safety_status = SAFE
ELSE safety_status = UNSAFE

The *threshold* that defines wearing/not wearing is set to 0.4 for all overlaps in this analysis but can be modified to better suit each application's requirement.
A summary of all variables, functions, constraints and rules considered in this analysis is presented in Table 2.

Table 2. Set of variables, functions, constraints and rules considered in analysis.

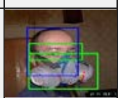
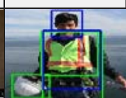
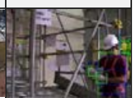
Elements	Term
Variable	Image
	Body Parts: head, forearm, legs, torso etc.
	PPEs: boot, earmuff, respirator, etc.
	Area of intercept A_I between body part and PPE
	Area of PPE A_{ppe}
	$O_{I,ppe}$
	safety_satatus of the scene
Function	Maximum overlap $O_{I,ppe}$
Constraint	Size of PPEs relative to body part
Rule	Set of Rules (1, 2, 3, 4, 5 and 6)

3.3 Experimental Results

The system has been tested over collection of frames (representing different industrial settings) of successful detections of body parts and PPEs. Only successful detections of PPEs are considered, as the goal at this stage is to evaluate the reasoning only. These images have been tested for two different set of rules, totalizing 150 observations.

Table 3 shows examples of the outputs representing the safety status of the frame in analysis for the given rule. Body parts are represented on blue rectangles and PPEs as green rectangles on the input frames.

Table 3. Output of system for each given set of rules.

	Rule 1	**Rule 2**	**Rule 3**	**Rule 4**	**Rule 5**	**Rule 6**
Frame						
Required equipment	Respirator	Helmet	Respirator and Helmet	High Visibility Clothes and Boots	Respirator, Helmet and Googles	Harness, Helmet and Gloves
Output	SAFE	UNSAFE	UNSAFE	SAFE	UNSAFE	SAFE

The outputs were manually verified to check the suitability of such approach. It has been measured the number of True Positive (TP), which is the number of frames tagged correctly as UNSAFE; True Negative (TN), the number of frames marked appropriately as SAFE; False Positive (FP), which is amount of frames that should have been identified as SAFE by the system but wrongly outputted the status as UNSAFE; and finally False Negative (FN), the number of frames the system tagged as UNSAFE mistakenly. The sensitivity and specificity rates also known by True Positive Rate (TPr) and True Negative Rate (TNr) respectively, have also been calculated [31]. Table 4 shows the results for evaluation of performance.

Table 4. Evaluation of performance.

Parameter	TP	TN	FP	FN	TP_r	TN_r	Accuracy
Value	97	50	2	1	98.98%	96.15%	98.00%

Given a set of successful detections, the methodology works effectively in recognising the safety status of the scene. For real time applications, the wrong status of safety may happen due to wrong status of each variable inputted into the system reasoning (e.g. wrong detections of body parts and PPEs) or mistakes in the interpretation of these variables during the reasoning process. One of the advantages of explicit representation of knowledge is the possibility to evaluate the causes of unreasonable outputs by checking the status of each variable involved. This way, if the issues are found to be related to the status of the variables, calibration the classifiers can be done as well as adjust on the data gathering process that could lead to such mislead. In addition, if the status of variables are found to be accurate, correction to reasoning can be made by adding a new set of constrains, functions or rules that adjust the output to the correct value for future observations.

4 Conclusions

In this paper we have shown through a case study how Decisional DNA (DDNA), a multi-domain knowledge structure that has the Set of Experience Knowledge Structure (SOEKS) as its basis can be utilized as a comprehensive embedded knowledge representation in a Cognitive Vision System for Hazard Control (CVP-HC). The proposed application aims to ensure that workers remain safe and compliant with Health and Safety policy for use of Personal Protective Equipment (PPE) and serves as a showcase to demonstrate the representation of visual and non-visual content together as an experiential knowledge in one single structure. At this point the implementation is working in offline mode, i.e. the application has been tested over images coming from a database. For next steps, more complex scenarios will be explored for the creation of more complexes set of rules and analyses of the results presented for online operation of the system in which the input images and context variables are gathered from video cameras and sensors in real time.

References

1. Sanin, C., Haoxi, Z., Shafiq, I., Waris, M. M., de Oliveira, C. S., & Szczerbicki, E. (2018). Experience based knowledge representation for Internet of Things and Cyber Physical Systems with case studies. Future Generation Computer Systems.
2. Vernon, D.: The space of cognitive vision. In Cognitive Vision Systems, 7-24, Springer, Berlin, Heidelberg (2006).
3. Gregory, R. L. (1973). Eye and brain: The psychology of seeing. McGraw-Hill.
4. De Oliveira, C. S., Sanin C., Szczerbicki E. Visual content representation and retrieval for Cognitive Cyber Physical Systems. In 23rd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (2019).
5. Sanin, C., Toro, C., Haoxi, Z., Sanchez, E., Szczerbicki, E., Carrasco, E.,. & Man-cilla-Amaya, L. (2012). Decisional DNA: A multi-technology shareable knowledge structure for decisional experience. Neurocomputing, 88, 42-53.
6. C. Sanin, E. Szczerbicki, Experience-based Knowledge Representation SOEKS. Cybernet Sys. 40(2) (2009) 99-122.
7. Sanin, C., Szczerbicki, E.: Decisional DNA and the smart knowledge management system: A process of transforming information into knowledge. In Techniques and tool for the design and implementation of enterprise information systems, ed. A. Gunasekaran, 149–175. New York: IGI Global (2008).
8. de Oliveira, C. S., Sanin, C., & Szczerbicki, E. (2018, September). Cognition and decisional experience to support safety management in workplaces. In International Conference on Information Systems Architecture and Technology (pp. 266-275). Springer, Cham.
9. Han, S., & Lee, S. (2013). A vision-based motion capture and recognition framework for behavior-based safety management. Automation in Construction, 35, 131-141.
10. Ciresan, D. C., Meier, U., Masci, J., Maria Gambardella, L., & Schmidhuber, J. (2011, July). Flexible, high performance convolutional neural networks for image classification. In IJCAI Proceedings-International Joint Conference on Artificial Intelligence (Vol. 22, No. 1, p. 1237).
11. Little, S., Jargalsaikhan, I., Clawson, K., Nieto, M., Li, H., Direkoglu, C., ... & Liu, J. (2013, April). An information retrieval approach to identifying infrequent events in surveillance video. In Proceedings of the 3rd ACM conference on International conference on multimedia retrieval (pp. 223-230). ACM.
12. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).
13. Mosberger, R., Andreasson, H., & Lilienthal, A. J. (2013, November). Multi-human tracking using high-visibility clothing for industrial safety. In Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on (pp. 638-644). IEEE.
14. Chen, L., Hoey, J., Nugent, C. D., Cook, D. J., & Yu, Z. (2012). Sensor-based activity recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42(6), 790-808.
15. Zambrano, A., Toro, C., Nieto, M., Sotaquirá, R., Sanín, C., Szczerbicki, E.: Video semantic analysis framework based on run-time production rules – towards cognitive vision. J. Univ. Comput. Sci. 21(6), 856–870 (2015).
16. S. Aditya, Y. Yang, C. Baral, Y. Aloimonos and C. Fermuller, "Image Understanding using vision and reasoning through Scene Description Graph," Computer Vision and Image Understanding, pp. 1-13, 2017.

17. de Oliveira, C. S., Sanin, C., & Szczerbicki, E. (2018). Flexible Knowledge–Vision–Integration Platform for Personal Protective Equipment Detection and Classification Using Hierarchical Convolutional Neural Networks and Active Leaning. Cybernetics and Systems, 49(5-6), 355-367.
18. Hollnagel, E., & Woods, D. D. (2005). Joint cognitive systems: Foundations of cognitive systems engineering. CRC Press.
19. Cole, G. S. (1990). Tort liability for artificial intelligence and expert systems. Computer/LJ, 10, 127.
20. Ashby, W. R. (1961). An introduction to cybernetics. Chapman & Hall Ltd.
21. Yu, Y., Pan, G., Gong, Y., Xu, K., Zheng, N., Hua, W., ... & Wu, Z. (2016). Intelligence-augmented rat cyborgs in maze solving. PloS one, 11(2), e0147754.
22. Pathak, N. (2017). The Future of AI. In Artificial Intelligence for. NET: Speech, Language, and Search (pp. 247-259). Apress, Berkeley, CA.
23. Vernon, D.: The space of cognitive vision. In Cognitive Vision Systems, 7-24, Springer, Berlin, Hei-delberg (2006).
24. Brézillon, P., & Pomerol, J. C.: Contextual knowledge and proceduralized context. In Proceedings of the AAAI-99 Workshop on Modeling Context in AI Applications, Orlando, Florida, USA, July. AAAI Technical Report (1999).
25. de Oliveira, C. S., Sanin, C., & Szczerbicki, E. (2019, April). Towards Knowledge Formalization and Sharing in a Cognitive Vision Platform for Hazard Control (CVP-HC). In Asian Conference on Intelligent Information and Database Systems (pp. 53-61). Springer, Cham.
26. Safe Work Australia: Australian Work Health and Safety Strategy 2012–2022. Creative Commons (2012).
27. Health and Safety Executive (HSE) (2018). Measuring the Effectiveness of HSES Field Activities. HSE Occasional Paper Series, OPll, Health and Safety Commission. London: HSE.
28. DeJoy, D. M. (2005). Behavior change versus culture change: Divergent approaches to managing workplace safety. Safety Science, 43(2), 105-129.
29. de Oliveira, C. S., Sanin, C., & Szczerbicki, E. (2018). Contextual Knowledge to Enhance Workplace Hazard Recognition and Interpretation in a Cognitive Vision Platform. Procedia Computer Science, 126, 1837-1846.
30. Au, K. W., Davalos, P., Venkatesha, S., Khurana, H., Bedros, S. J., Mohideen, M. I., ... & Cabuz, C. (2017). U.S. Patent No. 9,695,981. Washington, DC: U.S. Patent and Trademark Office.
31. Cortes, C., & Mohri, M. (2004). AUC optimization vs. error rate minimization. In Advances in neural information processing systems (pp. 313-320).