6-4-2018

Enhancing Autonomous Vehicles with Commonsense: Smart Mobility in Smart Cities

Priya Persaud
Rutgers - The State University of New Jersey, Newark

Aparna Varde
Montclair State University, vardea@mail.montclair.edu

Stefan Robila
Montclair State University, robilas@mail.montclair.edu

2017 International Conference on Tools with Artificial Intelligence

Enhancing Autonomous Vehicles with Commonsense

Smart Mobility in Smart Cities

Priya Persaud[1], Aparna S. Varde[2] and Stefan A. Robila[2]

1. Center for Law and Justice, Rutgers University, Newark, NJ, pap228@scarletmail.rutgers.edu
2. Department of Computer Science, Montclair State University, Montclair, NJ, (vardea | robilas)@montclair.edu

***Abstract*—Recent advances in AI include a Law firm hiring a robot lawyer and companies developing autonomous vehicles with robot drivers. Findings from our study have gauged the current cognitive capacity of such systems, indicating areas for improvement. We focus on autonomous vehicles, i.e., those that conduct automated driving and need to make autonomous, i.e., independent decisions. We propose an approach enabled with commonsense knowledge (CSK) from worldwide repositories to simulate intuitive humanlike decision-making in autonomous vehicles. We consider the repository WebChild with a multitude of CSK concepts, properties and relations. We investigate this and related domain-specific knowledge bases (domain KBs) to harness them within our proposed approach. Accordingly we build a transportation domain KB incorporating CSK and the needs of autonomous vehicles. This would be useful in guiding automated driving and making the systems get closer to the thresholds of human cognition. This work thereby makes contributions to smart mobility in smart cities. The paper presents our vision with design, implementation, experiments, recommendations and a future roadmap. As a broader impact, it propels more joint work between AI, Law and related areas.**

Keywords – Automated driving; AI and Law; Commonsense knowledge; Domain KBs; Object detection; Smart mobility

I. INTRODUCTION

Machine learning is focused on the idea of creating computers capable of achieving cognitive capacities pursuant to those of humans. With perseverance, computers can be programmed to absorb much data about the real world. Learning is attainable for computers because they behave in accordance with information hardwired into them. Reliance on pre-programmed knowledge is sufficient until computers are faced with novel tasks. In such scenarios, computers must learn to make their own decisions. This is difficult for them due to an important aspect they lack that humans possess: commonsense, which is often intuitive [1]. Commonsense allows humans to logically approach situations in which they have no prior experience. If computers had commonsense knowledge (CSK) and could apply that to specific domains, they would be able to make intuitive and logical decisions analogous to humans [1].

Recent research in autonomous vehicles [2] has shown that even if a robot is equipped with modern tools, it has trouble appropriately reacting to objects in its path. For example, in May 2016, a Tesla Model S crashed into a truck since the Tesla thought the truck was an overpass because of the truck's height. Humans possess commonsense to easily distinguish a truck from an overpass since an overpass is stationary while a truck can move. We hypothesize that CSK-aware robots and autonomous vehicles would be able to distinguish between objects such as a truck and an overpass; and can thus drive better. This paper is an initial step towards bringing this CSK-based reasoning to robots and autonomous vehicles.

With AI advancements such as large-scale commonsense knowledge acquisition and reasoning techniques, CSK-aware machines are no longer a fantasy. Google, Microsoft and Amazon are developing digital assistants like Alexa, Cortana and Echo for context-aware assistance. A Law firm recently hired a robot lawyer ROSS causing much debate [3]. CSK has also been used in smart city applications. A framework with CSK, text mining and ontology has been developed [4] to identify and manage implicit requirements (as opposed to explicit ones given by users) in software development, especially useful for smart city tools. CSK has been used to simulate human judgment in mining social media data for a smarter environment [5]. However, we are not aware of such advances in CSK for smart mobility. This paper heads towards CSK-enabled autonomous vehicles for smarter cities.

Based on the given background and motivation, we define the problem addressed in this work as follows.

- The main goal is to develop an approach embodying commonsense knowledge to enable autonomous vehicles make humanlike decisions.
- The related goal is to use this CSK-enabled approach on a prototype testbed for automated driving and make suitable recommendations accordingly.

II. PROPOSED APPROACH

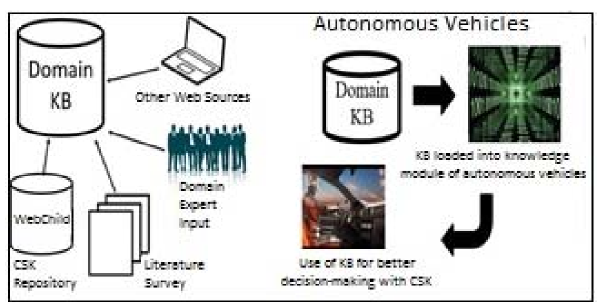

Fig. 1. Overview of CSK-enabled approach for autonomous vehicles

The proposed CSK-enabled approach is depicted in Fig. 1. We capture knowledge from a CSK repository *WebChild* [6], related domain KBs, other Web sources and domain experts. We use this to develop a transportation KB, considering autonomous vehicles. We incorporate this KB in a knowledge module of a simulated autonomous vehicle to test driving. This

2375-0197/17/31.00 ©2017 IEEE
DOI 10.1109/ICTAI.2017.00155

research is particularly focused on automation in the transportation sector of smart cities. It fits into their smart mobility characteristic [7].

The first step of this approach involves investigation of existing scholarship pertaining to CSK and domain-specific knowledge bases (domain KBs). We conduct a thorough study of a worldwide commonsense repository "WebChild" [6] and related domain KBs [8] with reference to smart cities. As a second step, we create a transportation domain KB [9]. Once the transportation sector of WebChild is extracted, it is modified to include images and concepts of obstacles common on roadways while driving. Such images are provided with high resolution over distributed sources and processed using suitable algorithms [10]. In the third step, this domain KB for transportation is programmed into a small robot that serves as a simulated autonomous vehicle by traversing a miniature driving course featuring obstacles. The robot's actions are monitored and recorded. All these steps are elaborated next.

A. *CSK and Domain KB study*

We investigate WebChild with details of its development. It is a huge global Web-based repository with fine-grained commonsense knowledge on everyday concepts. WebChild has an alignment to the lexical database WordNet: this allows distinguishing various senses of concepts, e.g., knowing the different meanings of a *green plant* (in the context of the color "green" vs an energy efficient power plant). As the concepts in WordNet are aligned to nearly 200 different domains, WebChild's alignment to it enables WebChild's slicing for different domains, such as "transportation". Apart from this alignment with WordNet, multimodal content and fine-grained relations make WebChild different from other commonsense KBs (ConceptNet [11] or Cyc [12]). TABLE I summarizes WebChild statistics [6]. It has around 18 million CSK facts with 3 million concepts and activities, connected by thousands of relations. It can be seen from precision values in TABLE I that this gives quite an accurate notion of CSK concepts.

TABLE I: WebChild statistics over 3M concepts and activities

Relation	#Sub-relations	#Assertions	Precision
Properties	19	4.3 M	0.82
Comparatives	6,331	1.1 M	0.90
Locations	1	0.1 M	0.86
Part-whole	3	6.7 M	0.88
Activities	7	6.1 M	0.85

Fig. 2 shows a relevant snapshot of the WebChild browser [6]. It is an example of defining the concept "vehicle" using WebChild data characterization. Commonsense knowledge stored in WebChild has been gathered from Web content using data mining techniques, such as mining Web corpora, movie scripts and visual contents, e.g., images.

In addition to WebChild, we study related domain KBs. In particular, we refer to a smart city KB developed mainly with reference to urban policy [8]. It contains knowledge in the domains: *environment, transport, buildings, vehicles and town planning*. These domains are also found relevant from the perspective of autonomous vehicles. With reference to this KB, a partial snapshot of an entry relevant to "transport" appears in Fig. 3 describing a simple common sense concept "bridge" [8].

It shows properties with values, in commonsense terminology, e.g., "location" is "over road".

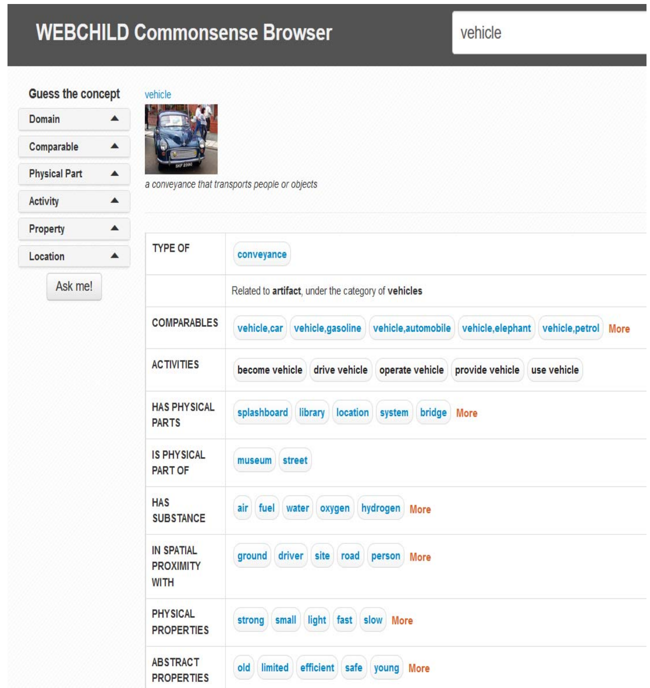

Fig. 2. Vehicle definition by WebChild data characterization

As we can see in the "motion" property here, a bridge is typically "stationary". Incorporating such knowledge in an autonomous vehicle would help it distinguish a bridge from a huge moving truck, potentially avoiding accidents. However, note that the "motion" property includes an exception, namely, a drawbridge. In reality, we know that a drawbridge moves very slowly compared to a truck on the road. However, this needs to be explicitly entered in a KB to be used by autonomous vehicles. Motivated by these and other issues, we develop a transportation KB to suit general purposes and particularly, autonomous vehicles in a smart city context.

Fig. 3. Simple example of CSK concept in domain KB

B. Transportation KB development

The development of a transportation knowledge base follows a three-stage approach of filtering, ranking and augmentation as described next.

Filtering: At the filtering stage, we slice WebChild for the transportation domain. We manually identify 17 domains (approximately 10% of 171 domains in WebChild) relevant to transportation: animals, law, person, geography, buildings, color, town_planning, transport, number, furniture, home, environment, school, architecture, vehicles and railway. We discard seemingly irrelevant domains: biology, sports etc.

Ranking: Here we use assertions relevant to autonomous vehicles. We rank WebChild content for concepts in these 17 domains, so the resulting knowledge pertains to automated driving and vision. It has been shown that properties, part-whole commonsense, spatial commonsense, activities and multimodal content can be useful in vision processing (e.g., overpass vs truck) and their common properties (e.g., overpass is static, while a truck typically moves on the road).

Augmentation: In this stage, we augment and curate the knowledge. We curate the ranked data to remove the noise and to increase the coverage. For this purpose we follow an overall methodology of entering pertinent CSK concepts and their relevant properties in a setup somewhat similar to Fig. 3.

C. Programming a robot for image detection

Based on the current literature and our research, we conduct the programming in two parts. In the first part, Scribbler, a simple, over the shelf robotic platform is used. The Scribbler robot [13] is a popular platform in robotics courses often paired with Fluke 2, an interface / controller card developed by a third party group. The combination results in a low cost platform equipped with a variety of sensors (including light, infrared and imaging), wireless connectivity (Bluetooth) and multi-core computing capability. This robot is easy to use, especially in a Calico Myro programming environment [14]. Apart from classroom use, the robot has been employed in road sign recognition systems and enhanced with the LIDAR capability to improve autonomous navigation [15].

Prior to image collection, we connect the robot to a desktop Calico application via Bluetooth. After initialization, we use simple Python commands to manoeuver the robot towards images of items common in driving: stop sign, construction site, overpass etc. As the robot approaches the images, it is instructed to conduct activities: "stop moving", "take picture", "display picture on laptop screen", and "save picture in Calico folder". Fig. 4 shows examples of such relevant commands.

```
turnLeft(1,10)
pic=takePicture()
show(pic)
savePicture(pic, "construction.jpg")
forward(5)
turnRight(2,10)
pic=takePicture()
show(pic)
savePicture(pic, "stopsign.jpg")
```

Fig. 4. Python commands in Calico for robotic programming

The second part of the programming involves writing code that uses the KB to analyze the images taken by the robot. As the images have already been analyzed with commonsense knowledge, the program should be able to use CSK to choose the appropriate course of action for each obstacle encountered and thus drive accordingly. Given this, the experiments are conducted with the robot on a small scale as a prototype. We discuss these with observations and recommendations.

III. Preliminary Experiments and Discussions

A. Experimental testbed and observations

Experimentation with our prototype system is conducted using a Scribbler robot [13]. The first part of the experimentation solely involves image collection. This part is successful as we have currently produced a program with the capabilities of performing image analysis using CSK. Fig. 5 shows a sample visualization of our robot driving on a small scale. This illustrates the navigation of the robot on a floor and depicts how it deals with obstacles.

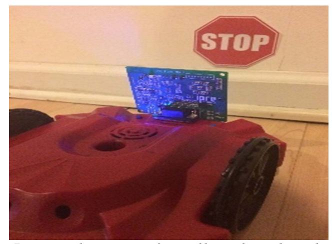

Fig. 5. Visualization of small scale robot driving

We have been able to progress well with these tasks to some extent as good initial steps. Our experiments so far yield the following observations. The prototype for automated driving can navigate well on a small scale, but depicts computational challenges in simultaneously gathering and processing images with reasoning (using the KB) on a large scale. It is found that this occurs due to the rather minute processing capabilities of the Scribbler robot. This leads us to infer that a more powerful system would be needed to test the full functional capabilities of our proposed approach for CSK-enabled autonomous vehicles. Hence our approach, though well outlined in theory with design and implementation of an automated driving prototype using CSK, needs more advanced experiments with comparative studies to be useful in practice.

B. Recommendations from experiments and study

Based on our research, experiments and literature survey, we make the following recommendations that would be useful for further research by us and also by others interested in the area. (Note: our code is in GitHub with access granted upon request). These recommendations would contribute to smart mobility in smart cities.

1. Incorporate more powerful robots with higher computation power. There is a high computational complexity of deep neural model based architectures typically employed in vision tasks. The costs involved in acquiring a powerful robot to cope with these challenges can be very high but experimenting with them can provide useful inputs for automated driving helpful in smart mobility. Acquiring such powerful robots is on our roadmap for large scale experiments (subject to funding).

2. Use newer technologies for image processing (faster and more space efficient). Fast object detectors like YOLO [16],

shown in Fig. 6, offer 100x speedup to state-of-the-art Fast-RCNN models. Space efficient solutions like XNOR-NET [17] allow sophisticated neural models for object detection to run in real-time, even on a smartphone. Yet, accuracy is compromised here to some extent, e.g., *bench* is confused with *overpass*. We claim that augmenting such object detectors with our CSK-enabled approach could potentially enhance accuracy. For example, the "isCollocated" property of WebChild would show high hits between (*overpass*, *road)* but not (*bench, road)*. Such knowledge would thus enable object detectors to distinguish *overpass* from *bench* with reference to context. Running experiments by programming such systems with our CSK-enabled approach to test the accuracy and demonstrating this in automated driving is our future work.

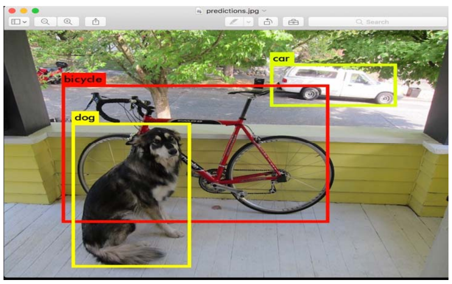

Fig 6. Partial snapshot of YOLO for object detection

C. Broader impacts of this research

As broader impacts, we emphasize that this work is useful in related areas. Humanoid robots with reasoning abilities in applications including automated driving and others (e.g., robot lawyers) demand attention in judicial decision-making. Judicial privileges are safeguarded for humans as they have *innate reasoning* abilities [18]. Research in this paper focuses on enhancing robots with commonsense thus enabling them with *innate reasoning,* hence furthering areas of study for AI and Law. Some issues pose controversies, e.g., lawsuits on autonomous vehicles in case of accidents (who is responsible). Also, employing CSK-enabled robots as lawyers is highly controversial. Such robots if used as prosecution or defendant attorneys would entail no bias and might yield fairer verdicts. Yet, some cases could be so sensitive that is there is need for real humans. Related issues of unemployment for humans are also significant. There could be many debates on autonomous vehicle lawsuits, CSK-enabled robot lawyers etc. Some of them are addressed in our project. Our work on commonsense, robots and human rights has been presented earlier [19] and is being further researched after incorporating aspects from this paper. All of this provides the potential for advanced research, development and debates in AI, Law and other pertinent areas.

IV. Related Work

A. CSK for better smart city applications

A framework called COTIR with common sense knowledge, text mining and ontology has been developed [4] to identify and manage implicit requirements (as opposed to explicit ones given by users) in the requirements specifications phase of software development. The authors claim that this is especially useful in artificial intelligence applications, e.g., smart city tools. Our work in this paper goes a step further in actually using CSK to design and implement a prototype for autonomous vehicles in smart cities.

Concepts pertaining to smart cities and related terms are discussed in [20]. They have categories, e.g., "green city", "sustainable city", "information city" used interchangeably by urban policy makers etc. They address issues of the categories having distinct conceptual perspectives. Their research design entails a Scopus database of articles with criteria on terms in academic literature. They count co-occurring keywords among categories, present evolution of categories over time and build a network structure of author keywords with categories. We find that commonsense knowledge is implicitly used here and hence such work could benefit from CSK repositories and related domain KBs with respect to efficiency, accuracy and the actual deployment of appropriate concepts in urbanization.

Smart cities are attracting attention all over the world today. A huge emphasis exists in Europe which by far has the maximum number of smart cities as described in the literature, e.g., [7]. In Barcelona, buses operate on routes that minimize the number of signals and maximize energy efficiency. In Amsterdam, street and canal lights are automatically dimmed and brightened based on pedestrian usage. Such applications surely embody commonsense. We envisage that they would be further facilitated by explicit commonsense knowledge bases (e.g., WebChild) and pertinent domain KBs.

B. CSK for better object detection

It is noted that object detection has dramatically improved over the last few years, especially with the advent of recent neural models. Faster R-CNNs [21] is an object detection system that works very well on data sets like MS COCO with 100 categories such as *human, chair* etc. This introduces a Region Proposal Network with full image convolution features and uses neural networks with attention mechanisms. It is found to achieve state-of-the-art object detection accuracy.

To further improve object detection, various types of contextual information has been employed. This contextual information can be global scene context or geometric context in the form of 3D surface orientations, relative location, 3D layout or geographic information [22]. Spatial context, global scene context and part-whole knowledge in particular is inspired by CSK. The work in [23] defines relevant co-occurrence of objects and penalizes presence of objects in irrelevant scenes. However, this work assumes that object detection is fairly accurate and only handles missing detected objects, e.g., updating detection of two or zero chairs instead of actually three or one in an image. However, we cannot make the assumption that all the object detection is highly accurate. Unlike our work, they do not rely on context from background knowledge or from text.

While these methods typically use commonsense jointly with object detection, other methods inject CSK into object detection as post processing. Consider an example of a tennis player with a racket where an object detection system confuses the *ball* with a *lemon*. CSK has been used to understand that a *lemon* is not suitable in the context of tennis. Researchers [22] have developed a model to distinguish semantic relatedness from visual relatedness. Such CSK is relevant to our scenarios.

Yet, these models would neither penalize the presence of an *overpass* nor a *truck* in the context of a *car*. They do not use the detailed context-relevant CSK that an *overpass* cannot move or that there is a risk of hitting a truck if it is misclassified as an overpass. We address this in our paper.

C. CSK for better automated vehicles

There is a tremendous amount of work in path planning in the context of AI that intensively makes use of commonsense. However, these systems have typically operated over toy-sized KBs and their primary focus is on reasoning. A system called RoboBrain [24] has been developed at Cornell that learns concepts from the Web, robot trials and computational simulations. It interprets text, pictures and videos and stores this knowledge in a comprehensive manner. This serves to provide a fully integrated KB. The WebChild knowledge base with a broader spectrum of CSK relations can be considered a larger alternative to RoboBrain.

Our work in this paper can benefit from research advances described herein. Yet, we would need decently sized domain KBs to serve as caches and retrieve useful data in specific applications, as opposed to searching vast repositories built by self-learning robots or using generic object detection systems. This would provide an efficient process to get only relevant data for decision-making. In the future, we could expand our work by incorporating knowledge from such systems and by running benchmark studies with CSK-enabled autonomous vehicles using more advanced robots.

V. Conclusions and Future Work

This paper addresses the challenging issue of enhancing autonomous vehicles with commonsense. On the whole, it makes the following contributions so far:

- Investigating the literature in autonomous vehicles to detect anomalies, e.g., *truck* vs *overpass* confusion
- Proposing an approach entailing CSK to enhance intuitive decision-making in automated driving
- Building a transportation KB with filtering, ranking and augmentation, helpful in autonomous vehicles
- Programming a small scale Scribbler robot with CSK for image detection while driving to test the approach
- Making recommendations based on the approach and its execution useful for smart mobility in smart cities
- Encouraging greater participation of professionals in AI, Law and related areas for joint work

Future work entails using powerful robots / advanced object detectors like YOLO to test our CSK-enabled approach. We anticipate that CSK would augment object detection accuracy, improve decision-making in simulated automated driving and thus enhance autonomous vehicles. Hence, this work would make a positive impact on smart mobility in smart cities.

Acknowledgments

Much of this research occurred when Priya Persaud was a BS CS student at Montclair State University (MSU) funded by NSF's LSAMP. We sincerely thank Niket Tandon from Allen Institute for Artificial Intelligence, Seattle, WA, for his valuable inputs and for access to relevant parts of the WebChild code. Early work on this started when Aparna Varde was visiting researcher at Max Planck Institute for Informatics, Saarbruecken, Germany. Stefan Robila's research is supported by a Faculty Scholarship Program at MSU.

References

[1] Davis E., "Commonsense reasoning and commonsense knowledge in artificial intelligence", *Communications of the ACM,* 2015, 58(3):92-103.

[2] Driggs-Campbell K., Shia V., Bajcsy R., "Decisions for autonomous vehicles: Integrating sensors, communication, and control", *ACM HiCoNS*, Berlin, Germany, Apr 2014, pp. 59-60.

[3] Weller C. "Law firms of the future will be filled with robot lawyers", *Business Insider*, Jul 2016.

[4] Emebo O., Varde A., Olawande D., "Common sense knowledge, text mining and ontology for implicit requirements", *DMIN,* Las Vegas, NV, Jul 2016, pp. 146 – 152.

[5] Du X., Emebo O., Varde A., Tandon N., Nag Chowdhury S., Weikum G., "Air quality assessment from social media and structured data", *IEEE ICDM Workshops,* Helsinki, Finland, May 2016, pp. 54-59.

[6] Tandon N., de Melo G., Weikum, G., "Acquiring comparative commonsense knowledge from the Web", *AAAI*, Quebec City, Canada, Jul 2014, pp. 166-172.

[7] Vienna University of Technology (TU Wien), *European smart cities*, Technical Report, Vienna, Austria, 2015.

[8] Varde A., Tandon N., Nag Chowdhury S., Weikum, G., *Common sense knowledge in domain-specific knowledge bases,* Technical Report., Max Planck Institute for Informatics, Saarbruecken, Germany, Aug 2015.

[9] Persaud P., "Domain-specific knowledge bases with common sense knowledge for autonomous vehicles in smart cities", Presented at *NSF GS-LSAMP Undergrad Research Conference,* New Brunswick, NJ, Oct 2016.

[10] Robila S., Ricart S., "Distributed algorithms for unmixing hyperspectral data using non-negative matrix factorization with sparsity constraints", *IEEE IGARSS*, Jul 2013, pp. 2156- 2159.

[11] Speer R., Havasi C., "Representing general relational knowledge in ConceptNet 5", *LREC*, Istanbul, Turkey, May 2012, pp. 3679-3686.

[12] Lenat D., "CYC: A large-scale investment in knowledge infrastructure", *Communications of ACM*, 1995, 38: 32-38.

[13] Kumar D., *Learning computing with robots*, Institute for Personal Robots in Education, 2011.

[14] Blank D., Kay J.., Marshall J., O'Hara K., Russo M., "Calico: A multi-programming-language, multi-context framework designed for computer science education", *ACM SIGCSE*, Raleigh, NC, Mar 2012, pp. 63–68.

[15] Bilgin E., Robila S.A., "Road sign recognition system on Raspberry Pi", *IEEE LISAT*, Farmingdale, NY, 2016, pp. 1-5.

[16] YOLO Real time object detection, https://pjreddie.com/darknet/yolo/

[17] XNOR-NET http://allenai.org/plato/xnornet/

[18] Persaud P., Drake I., *The possibility of including artificial intelligence beings within the justice system*, Technical Report, Dept. of Political Science, Montclair State University, NJ, April 2016.

[19] Persaud P., Varde A., Drake I., "Common sense knowledge, humanoid robots and human rights", *IEEE MIT URTC*. Cambridge, MA, Nov 2016.

[20] de Jong M., Joss S., Schraven D., Zhan C., Weijnen M. "Sustainable-smart-resilient-low carbon eco-knowledge cities; Making sense of a multitude of concepts promoting sustainable urbanization", *Journal of Cleaner Production,* Elsevier, Dec 2015, 109:25-38.

[21] Ren S., He K., Girshick R., Sun J., "Faster R-CNN: towards real time object detection with region proposal networks", *ARXIV e-prints:* 1506.01497, Jan 2015.

[22] Torralba, A., Murphy, K., Freeman, W.T., "Using the Forest to see the Trees: Object Recognition in Context", *Communications of the ACM*, Mar 2010, 53(3):107-114.

[23] Torralba, A., Murphy, K.P., Freeman, W.T., Rubin, M.A., "Context-Based Vision System for Place and Object Recognition", *ICCV*, Nice, France, Oct 2003, 3:273-280.

[24] RoboBrain http://robobrain.me/#/