Visual Relationship Detection using Knowledge Graphs for Neural-Symbolic AI

Dave Herron

City, University of London, London, United Kingdom

Abstract

Momentum is surging behind the consensus that neural-symbolic AI is the right road for AI to take today. We propose to travel this road using Semantic Web technologies to represent the symbolic AI tradition. Our objective is to investigate and compare the efficacy of a variety of strategies for combining the capabilities of deep neural networks for statistical learning from data with those of OWL ontologies and knowledge graphs for symbolic knowledge representation and reasoning. Our application area is visual relationship detection within images. Deep learning is data hungry and struggles to generalise to examples outside the training distribution. We seek to show that combining Semantic Web domain knowledge and reasoning with deep learning can deliver superior performance, can substitute for plentiful training data, and can deliver robust generalisation in few-shot/zero-shot learning scenarios.

Keywords

neural-symbolic, AI, semantic web, knowledge graphs, CEUR-WS

1. Problem Statement

At the macro-level, our problem space is the subfield of artificial intelligence (AI) called neural-symbolic AI, which is concerned with integrating the learning capabilities of deep neural networks (NNs) with the knowledge representation and reasoning capabilities of the symbolic tradition of AI in order to get the best of both worlds. Combining the neural and symbolic traditions of AI is an open research challenge. As Valiant explains in [1], although learning and reasoning have been thoroughly studied (separately) by the two traditions of AI, the semantics of their models are hard to reconcile: neural AI learning models have a statistical character whereas symbolic AI knowledge representation and reasoning models have a logical character. Further, the distributed, sub-symbolic representations of knowledge (features) encoded by NNs differ dramatically from the symbolic representations of knowledge used in logic.

Our research explores neural-symbolic AI using Semantic Web (SW) ontologies and knowledge graphs (KGs) for symbolic knowledge representation and reasoning. Selecting SW technologies to represent symbolic AI is uncontroversial. In [2], Hitzler explains that the SW field has long had a strong association with the symbolic tradition of AI. Further, in a recent review of specific synergies between SW technologies and deep learning, Hitzler et al. [3] anticipate that hybrid, neural-symbolic systems that leverage OWL ontologies and KGs should be capable of delivering better performance and interpretability than deep learning alone.

At the micro-level, our problem space is the application area of visual relationship detection within images, using a small (5,000 images) dataset prepared for this purpose in 2016: the VRD dataset [4]. The dataset originators used crowd-sourcing to annotate each image with some number of visual relationships (VRs). A VR is a `(subject, predicate, object)` triple, where the `subject` and `object` are individual objects (represented by bounding boxes and class labels) and the `predicate` expresses some relationship between them. For example, `(person, ride, horse)` and `(horse, on, grass)` might be two VRs for a given image. The VR annotations refer to 100 common, everyday classes of object that broadly but sparsely span the material world: types of vehicle, furniture, appliance, device, clothing, sporting good, animal, plant, landscape feature, etc.. The 70 predicates (relationships) referred to in the VR annotations are primarily common spatial relations (above, below, behind, beside, on, in, ...) and common verbs (wear, hold, use, carry, drive, ride, eat, touch, kick, has, ...). The breadth and variety of object classes and predicates permitted us to design an ontology with rich class and property hierarchies for describing (what we call) this VRD-world domain. Our VRD-world ontology currently contains 239 classes and makes extensive use of object property characteristics such as domain/range restrictions, subPropertyOf, property equivalence, symmetry, inverses, etc.. The VR annotations also map nicely to RDF triples for populating a KG with facts (ABox data).

Deep learning is known to depend on big data for good performance, and to struggle to generalise and extrapolate to examples that lie outside the distribution of data seen during training. The small size of the VRD image dataset, and the long tail on the highly skewed frequency distribution of the VR annotations of the images, are likely to provoke these limitations of deep learning. Our research considers how to leverage the symbolic knowledge representation and reasoning capabilities of a KG hosting our OWL VRD-world ontology so as to overcome the limitations of deep learning and deliver superior performance at detecting VRs in images.

2. Importance

Our research has relevance and benefits for several groups: the AI community, industry, society generally and the SW community. Prominent voices from the AI community such as those of Marcus [5], Chollet [6] and Kautz [7] corroborate one another in arguing that, due to its limitations (like those just mentioned), deep learning alone, despite its spectacular achievements, will not lead to human-level, artificial general intelligence (AGI). Marcus [5] warns that the ill-advised hype around deep learning could lead to a 3rd AI winter. Kautz [7] speculates that the (current) 3rd AI summer *may* avoid succumbing to a 3rd winter, but only because of the momentum that now exists behind neural-symbolic integration. A consensus has emerged that the neural-symbolic road is the right one for AI today [6, 7, 5, 8]. By helping to show how symbolic background knowledge and reasoning can reduce deep learning's dependence on big data whilst boosting its ability to generalise, and by advancing understanding of neural-symbolic AI generally, our research directly contributes to taking AI further along the road toward human-level, AGI and to helping it avoid another AI winter.

AI has transformed many aspects of everyday life, in industry and for society generally, and in ways we all now take for granted. We all have expectations of continuing positive innovations and transformative effects from AI. Hence, as neural-symbolic AI research advances

AI along the road to AGI, industry and society generally will be impacted directly and benefit proportionately.

Finally, our research has the potential to show that OWL ontologies and KGs can be used to integrate neural, statistical learning with symbolic background knowledge and reasoning in concrete, tangible ways. In doing so, it can demonstrate that these SW technologies are exemplars of the types of symbol manipulating tools and abstraction and reasoning modules that Marcus [5, 8] and Chollet [6], respectively, call for being incorporated into hybrid, neural-symbolic systems in order for AI to advance along the road to AGI. Such a demonstration may shine a spotlight on SW technologies that helps to place them, and the SW community, at the centre of attention of neural-symbolic AI.

3. Related Work

Work relating to neural-symbolic AI in general can be reviewed in surveys such as [9, 10, 11, 12]. One prominent example is Logic Tensor Networks (LTN) [13, 14]. LTN is a fuzzy logic-based framework for training conventional NNs to satisfy logical constraints expressed as background knowledge over training data.

Work relating to neural-symbolic AI that uses SW technologies to represent the domain of symbolism is rapidly accumulating. Myklebust et al. [15] create a composite KG from disparate sources, use it to generate KG embeddings (via various models), and then use the KG embeddings to train a NN binary classifier to predict whether or not a property representing mortality risk should be present in the composite KG to link certain individual chemicals with certain individual species of organism. The theme of 'deep deductive reasoning' (training NNs to reason over SW knowledge bases and KGs) is progressively developed in [16, 17, 18]. The theme of using KGs to compensate for the lack of plentiful samples with which to train robust deep learning-based systems (in so-called few-shot and zero-shot learning scenarios) is studied in [19] and [20] and surveyed in [21]. Similarly, in [22], Wang et al. demonstrate that the structure of the class hierarchy of a cell ontology can be leveraged (as an undirected graph) to significantly improve the accuracy of deep learning-driven cell classification for cells whose types were unseen during training.

Work relating to using neural-symbolic AI for VR detection on the VRD dataset also exists. The original (2016) VRD paper by Lu et al., [4], does not mention neural-symbolic integration (which reflects how little traction this area of AI had just some years ago), and neither is it mentioned in the comprehensive survey of the area in [9]. But it should absolutely now be recognised as an early and innovative form of neural-symbolic AI because their system is a hybrid that includes a 'language module' trained on word embeddings of the (symbolic) VRD object class names. Donadello & Serafini [23] enumerate LTN negative domain/range knowledge constraints to train NNs to detect VRs in VRD images. Their approach, however, exposes a scalability limitation of LTN that we hope to show can be elegantly overcome by using KGs. Daniele & Serafini [24] test their KENN system on the VRD dataset.

4. Research Questions

RQ1: How can we combine learning and reasoning to get the best of both worlds? Here, 'learning' refers to the the statistical 'learning from data' capabilities of deep NNs, 'reasoning' refers to the symbolic knowledge representation and reasoning capabilities of OWL ontologies and KGs, and 'best of both worlds' refers to improved VR predictive performance. We hypothesise that each of the several distinct NN-KG integration (NN-KG-I) strategies that we have conceived (and which we describe shortly) will deliver VR predictive performance that is superior to whatever baseline performance our deep NNs are able to deliver by themselves. We aim to experiment with each of our NN-KG-I strategies individually, rank them, and explain their relative efficacy by analysing the nature of the interactions between deep learning and symbolic knowledge representation and reasoning that they exercise.

RQ2: Some of our NN-KG-I strategies will be compatible with one another. How will they perform when used in different combinations? We hypothesise that we will find the contributions to improved VR predictive performance (beyond the baseline) that they make when used individually are additive when used in certain combinations, but that when used in other combinations there are interesting interaction effects between the integration strategies which either amplify or diminish their collective effect on VR predictive performance. Analysis of the results of these experiments is expected to yield further insights into the nature of the interactions between deep learning and symbolic knowledge representation and reasoning.

RQ3: How best and to what extent can the symbolic knowledge representation and reasoning capabilities of OWL ontologies and KGs be leveraged by deep NNs to substitute for plentiful training data and enable robust generalisation on out-of-distribution examples? Within the VRD image annotations, many VR types have just a few training instances, and many have test instances but zero training instances. Hence, the VRD dataset lends itself to the examination of the sort of small dataset and few-shot/zero-shot learning scenarios in which deep learning alone struggles to perform well. We hypothesise that our NN-KG-I strategies will show that SW technologies such as our OWL VRD-world ontology, together with the reasoning capabilities of a KG, can be used to construct hybrid, neural-symbolic systems that out-perform deep learning alone in such settings.

5. Methods

The architecture of our baseline hybrid, neural-symbolic system consists of a NN pipeline and a KG populated with our VRD-world ontology. The NN pipeline consists of an object detection NN followed by a multi-label predicate prediction NN that takes ordered pairs of detected objects, plus geometric features derived from their bounding boxes (an idea borrowed from [23]), as input. Experimentation with our several NN-KG-I strategies for combining these neural and symbolic components will drive the exploration of our research questions. We describe two of our NN-KG-I strategies, denoted *S1* and *S2*, in some detail and mention others briefly.
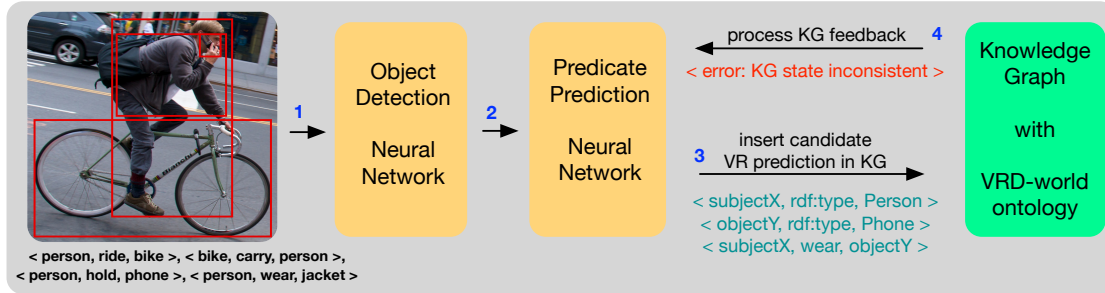
Figure 1: The baseline architecture of our hybrid, neural-symbolic systems illustrating NN-KG integration strategy *S1*. The KG deduces that **<person, wear, phone>** is a poor VR prediction (because it violates a range restriction of property **wear** — i.e. a phone is not a wearable thing).

S1: Strategy S1 is about leveraging a SW KG's ability to automatically enforce the semantic rules of an OWL ontology. The only predicted VRs that stand a chance of matching with annotated (ground truth) VRs are those that are semantically valid according to our VRD-world ontology. So it makes sense to help our predicate prediction NN learn the relevant semantic rules of our VRD-world ontology. This is analogous to a checkers-playing system learning the legal checkers moves. It does not fully solve the problem of finding the best move, but by narrowing the search space to the legal moves, learning to find the best move becomes significantly easier. One way to utilise strategy S1 is to take the VR predictions emitted by the predicate prediction NN during training, insert them into a KG populated with the VRD-world ontology, and then communicate any feedback from the KG regarding invalid VRs back to the NN (e.g. by penalising its loss function). This scenario is depicted in Figure 1.

There is also potential to utilise strategy S1 during inference. Despite having trained our predicate prediction NN as best we can to only predict semantically valid VRs, it may still sometimes predict semantically invalid VRs during inference on test set images. But the KG is an active agent capable of participating in determining the final predictions of the hybrid, neural-symbolic system. Specifically, we can use it as a final filter to suppress bad VR predictions. Rather than submit the VR predictions of the predicate prediction NN directly (on behalf of the hybrid, neural-symbolic system), we may first insert them into the KG and then submit only those that the KG does not flag as being semantically invalid.

Our strategy S1 is related conceptually to the approach taken in [23] (mentioned earlier) of enumerating (in code) LTN negative domain/range knowledge constraints to train NNs to detect VRs in VRD images. But by using a KG and ontology rather than hand-coded knowledge axioms, our strategy S1 will be shown to be more general, more scalable and more flexible than the LTN neural-symbolic framework. It is more general because a KG automatically enforces all semantic rules of an ontology, not just one particular category of domain rule. It is more scalable because a KG scales effortlessly to handle domains with any number of classes and properties, whereas the approach taken with LTN in [23] proved intractable given only the limited diversity of VRD object classes and predicates. Finally, it is more flexible because, as we have explained, a KG used for strategy S1 can participate not just in NN training but during NN inference as well.

S2: Strategy S2 involves using common-sense, Datalog-like *rules* to leverage and augment the reasoning capabilities of our KG. For example, a rule expressing the plausibility of VR pattern **(X, wear, Y)** can be described as follows:

```
wear(X,Y) :- Person(X), WearableThing(Y), ir(Y,X) ~ 1
```

Triples asserting the predicted classes of the detected objects X and Y will have been inserted into our KG before the rule is evaluated: (X, rdf:type, C_x), (Y, rdf:type, C_y). Determining if the first goal of the body of this rule is satisfied is accomplished by a KG query to check whether X is a member of class *Person*. Determining if the second goal is satisfied is accomplished similarly but depends on the KG having deduced whether Y is a *WearableThing* (which is not a VRD object class). The third goal represents a novel reuse of a bounding box geometric feature function (per [23]), the *inclusion ratio*. This goal is satisfied if the bounding box for Y is mostly enclosed within the bounding box for X (as would generally be the case when a person can plausibly be said to be wearing something).

We will define a collection (base) of such rules using the knowledge of the VRD-world domain gained from having analysed the VR annotations so as to design the ontology. The impact of strategy S2 will likely be proportional to the comprehensiveness of this rule base, but on a diminishing scale. Analysis of the frequency distribution of the VR annotation types will help us identify those rules likely to have the greatest impact. We can minimise the number of rules required to deliver a measurable performance effect by focusing on high-impact rules.

We have identified two approaches for implementing our common-sense rules. Approach S2-A is to use Python to build custom rules and a simple (non-recursive) rule engine component for evaluating them. This rule engine component would mediate interaction between the predicate prediction NN and the KG. The description of S2 given above presumes this approach. Approach S2-B is to define proper Datalog rules and to use a KG tool whose support for Datalog includes support for basic arithmetic functions sufficient for emulating bounding box geometric feature functions (such as the inclusion ratio).

A further approach to S2, S2-C, would share S2-A's Python-based rules and rule engine while allowing us to explore an additional dimension of NN-KG integration. This approach involves developing methods for transferring and representing the structure of the class hierarchy of our VRD-world ontology within supplementary layers (and their associated weight matrices) of our object detection NN. The objective is to enable the object detection NN to not just detect objects of base classes (e.g. jacket) but to perform the generalisations needed to convey all the parent classes for each object (e.g. clothing, wearable thing, etc..) as well. This way, our Python-based rule engine need not query the KG because any class membership information required to determine rule goal satisfaction will have been supplied by the NN.

S1 and S2 combined: NN-KG-I strategy *S1* should be good at identifying poor VR predictions (negative cases), while *S2* should be good at identifying plausible VR predictions (positive cases). So, in theory, they are complementary and may work well together, delivering a combined boost to VR predictive performance.

Some other NN-KG-I strategies: Another integration strategy involves using KG embeddings to train a NN to score the plausibility of VR predictions. This scoring NN would then be

used to help train the predicate prediction NN.

A further strategy involves leveraging the training set VR annotations (KG data). Each ordered pair of object classes will have some number of annotated VR instances involving some subset of the 70 predicates. These can be transformed into discrete probability distributions over the predicates. The VR predictions generated by our predicate prediction NN during training can similarly be transformed into corresponding discrete probability distributions. The dissimilarity of corresponding pairs of predicted and annotated VR probability distributions can then be measured (using a metric identified for this purpose), and these measures can be aggregated to produce a penalty term with which to augment the loss function of the predicate prediction NN. Tactics for leveraging the ontology so as to intelligently redistribute probability in the target distributions can also be explored to better facilitate few-shot and zero-shot learning.

Evaluation: The NN pipeline of our system architecture will be capable of delivering some measure of VR predictive performance on its own. This is the baseline performance measure against which all NN-KG-I strategies will be judged.

The authors of [4] and [23] measure VR predictive performance using a recall@N metric that measures recall globally, across all images. In addition to this global recall@N metric, we plan to use a per-image measure of recall@N that we average over the images. Basic recall@N (whether global or per-image and averaged), however, takes account only of the number of hits in the top N predictions. We have also designed a more sensitive metric that we call 'Mean Avg Recall@K top-N' that measures both the hit count and the positions of the hits within the top N ranked predictions.

As per [4] and [23], we plan to evaluate zero-shot VR predictive performance similarly to overall performance. The only difference is that when evaluating zero-shot performance, the annotated VRs participating in evaluation are limited to zero-shot VR instances — i.e. VR instances whose VR types are not represented within the training set VR annotations. We plan to evaluate few-shot VR predictive performance in the same way, but here the annotated VRs participating in evaluation will be limited to those for which the training set VR annotations contain only some small number of instances (1 to 5, say).

A key principle of our evaluation strategy is to keep the baseline architecture of our hybrid system unchanged across investigations of different NN-KG-I strategies. This will enable us to attribute changes in VR predictive performance to a NN-KG-I strategy alone. It will also best enable us to compare and rank our strategies in terms of VR predictive performance efficacy.

We use multiple metrics so our evaluation strategy cross-validates. If the multiple measures of performance corroborate one another, we can interpret the effect of a given NN-KG-I strategy with confidence. If not, this will signal the need for caution and investigation.

6. Preliminary Results

We are still assembling the infrastructure to enable experimentation, so we discuss preliminary results in the sense of things accomplished. The original, crowd-sourced VR annotations of the VRD dataset are full of inconsistencies and errors. For example, object class 'bear' refers to real bears and teddy bears; class 'plate' refers to dishware plates, license plates (on vehicles) and

baseball (home) plates; too many instances of VR pattern **(person, wear, Y)** have **Y** on a different person. Apart from making object detection and relationship prediction noisily problematic, the semantic variability of the object classes made precise ontology design infeasible. No class hierarchy felt credible, and few opportunities existed to define useful domain/range restrictions on object properties (VRD predicates). We therefore undertook a comprehensive VR analysis and customisation exercise to strengthen the semantic consistency of the VR annotations. In time, our VRD-world ontology, our customised VR annotations, our protocol for specifying VR customisations textually, and our code for applying them in an automated, repeatable fashion, will be made publicly available as a contribution to the AI and SW communities.

Object detection training and experimentation is underway. Our predicate prediction NN has been designed. Our evaluation metrics have been implemented and proof-of-concept testing confirms they behave as expected. Proof-of-concept exercises confirming the feasibility of NN-KG-I strategies *S1* and *S2* (described above) have been completed successfully. A customised binary cross-entropy loss function has been conceived for training our multi-label predicate prediction NN that provides parameters for influencing the loss attracted by predicted VRs that have no matching annotated VR. Many of these predicted VRs will be entirely plausible and yet will be treated as false positives simply due to the unavoidable sparsity and arbitrariness of the annotated VRs. We aim to explore the effect of influencing the magnitude of the loss attracted by such plausible false positive cases, based on judgements derived from our KG.

7. Reflection and Future Work

Reflection: First, our research is about combining the use of KGs with deep learning in hybrid, neural-symbolic systems. The application task of VR detection within images is simply a context for exploring combination/integration strategies. We believe our NN-KG-I strategies to be generic and widely applicable. However, we plan to continue looking for other dataset/ontology pairs with which to apply our strategies so as to further demonstrate their generality.

Second, as we have described, we chose to heavily customise the original, crowd-sourced VR annotations of the VRD images in order to enable the design of a precise ontology (and to correct egregious errors). A consequence of this choice is that we sacrifice the ability to directly compare the predictive performance results of our various hybrid VR detection systems with those of the systems of previous researchers (such as [4] and [23]). However, this sacrifice is justified by the fact that the purpose of our research is not to build a better VR detection system on the VRD dataset than others, it is to explore generic ways of combining KGs with deep learning that deliver performance superior to what deep learning can deliver alone.

Third, NN-KG-I strategies such as *S1* and *S2* that rely on real-time interaction with a KG are likely to increase NN training times considerably, particularly if the KG is out-of-process (even online) and accessed via a SPARQL endpoint. We do not, however, believe this consideration to be a major concern. In our case, the small VRD dataset (4,000 training images) means no issues should arise. More generally, we surmise that the growth rate of NN training times relative to dataset size, n, will be *linear* ($O(n)$ time) and that, on this basis, the computational complexity implications of real-time KG access should always be manageable. Further, tactics such as caching may well be exploitable to help keep KG access to a minimum.

Future work: Our research can readily extend in multiple directions. One direction is to pursue the goal of contributing to the development of a theory to help formalise the foundations of neural-symbolic AI, as advocated by van Harmelen in [25]. One such contribution involves positioning our NN-KG-I strategies within the schemes for categorising approaches to (and compositional patterns for) neural-symbolic integration proposed by others (e.g. [13, 26, 27, 7]). Where we find they do not fit comfortably, we might propose scheme/pattern refinements. We expect this task to be both challenging and rewarding given that several alternate categorisation/pattern schemes have been proposed and given that some of our multiple different strategies may well sit best at different positions within these different schemes. Another contribution involves taking our analyses of the interactions between deep statistical learning and symbolic knowledge representation and reasoning exposed by our NN-KG-I strategies to a deeper, more theoretical level. Another direction in which our research leads involves enhancing the interpretability of hybrid, neural-symbolic system behaviour by, for example, investigating methods for generating explanations of predictions for system users. Yet another direction involves exploring NN-KG-I strategies for the express purpose of extracting new knowledge from data to add to KGs (aka KG completion).

Acknowledgments

References

[1] L. G. Valiant, Three Problems in Computer Science, J. ACM 50 (2003) 96–99.

[2] P. Hitzler, A Review of the Semantic Web Field, Commun. ACM 64 (2021) 76–83. URL: https://doi.org/10.1145/3397512.

[3] P. Hitzler, F. Bianchi, M. Ebrahimi, M. K. Sarker, Neural-Symbolic Integration and the Semantic Web, Semantic Web 11 (2020) 3–11.

[4] C. Lu, R. Krishna, M. Bernstein, L. Fei-Fei, Visual Relationship Detection with Language Priors, in: European Conference on Computer Vision, 2016, pp. 852–869. URL: https://cs.stanford.edu/people/ranjaykrishna/vrd/.

[5] G. Marcus, Deep Learning: A Critical Appraisal, CoRR (2018). URL: http://arxiv.org/abs/1801.00631.

[6] F. Chollet, Deep Learning: Current Limits and What Lies Beyond Them, Presentation at RAAIS, 2018. URL: https://raais.co/speakers-2018.

[7] H. Kautz, The Third AI Summer, AAAI Robert S. Engelmore Memorial Lecture, Thirty-fourth AAAI Conference on Artificial Intelligence, New York, NY, 2020. URL: https://www.cs.rochester.edu/u/kautz/talks/, presentation slides and video.

[8] G. Marcus, The Next Decade in AI: Four Steps Towards Robust Artificial Intelligence, CoRR (2020). URL: https://arxiv.org/abs/2002.06177.

[9] T. R. Besold, A. S. d'Avila Garcez, et al., Neural-Symbolic Learning and Reasoning: A Survey and Interpretation, CoRR (2017). URL: https://arxiv.org/abs/1711.03902.

[10] A. d'Avila Garcez, M. Gori, et al., Neural-Symbolic Computing: An Effective Methodology for Principled Integration of Machine Learning and Reasoning, FLAP 6 (2019) 611–632. URL: https://collegepublications.co.uk/ifcolog/?00033.

[11] P. Hitzler, A. Eberhart, M. Ebrahimi, M. K. Sarker, L. Zhou, Neuro-Symbolic Approaches in Artificial Intelligence, National Science Review (2022).

[12] M. K. Sarker, L. Zhou, A. Eberhart, P. Hitzler, Neuro-Symbolic Artificial Intelligence: Current Trends, CoRR (2021). URL: https://arxiv.org/abs/2105.05330.

[13] S. Badreddine, A. d'Avila Garcez, L. Serafini, M. Spranger, Logic Tensor Networks, Artificial Intelligence 303 (2022) 103649.

[14] L. Serafini, A. S. d'Avila Garcez, Logic Tensor Networks: Deep Learning and Logical Reasoning from Data and Knowledge, in: Proceedings of NeSy'16, 2016.

[15] E. B. Myklebust, E. Jiménez-Ruiz, J. Chen, R. Wolf, K. E. Tollefsen, Prediction of Adverse Biological Effects of Chemicals using Knowledge Graph Embeddings, Semantic Web 13 (2022) 299–338. URL: https://doi.org/10.3233/SW-222804.

[16] F. Bianchi, P. Hitzler, On the Capabilities of Logic Tensor Networks for Deductive Reasoning, in: Proceedings of the AAAI-MAKE, volume 2350, 2019.

[17] M. Ebrahimi, A. Eberhart, F. Bianchi, P. Hitzler, Towards Bridging the Neuro-Symbolic Gap: Deep Deductive Reasoners, Appl. Intell. 51 (2021) 6326–6348.

[18] M. Ebrahimi, M. K. Sarker, F. Bianchi, et al., Neuro-Symbolic Deductive Reasoning for Cross-Knowledge Graph Entailment, in: Proceedings of the AAAI-MAKE, volume 2846, 2021. URL: http://ceur-ws.org/Vol-2846/paper8.pdf.

[19] Z. Chen, J. Chen, et al., Zero-Shot Visual Question Answering using Knowledge Graph, CoRR (2021). URL: https://arxiv.org/abs/2107.05348.

[20] Y. Geng, J. Chen, Z. Ye, Z. Yuan, W. Zhang, H. Chen, Explainable Zero-shot Learning via Attentive Graph Convolutional Network and Knowledge Graphs, Semantic Web 12 (2021) 741–765. URL: https://doi.org/10.3233/SW-210435.

[21] J. Chen, Y. Geng, et al., Low-Resource Learning with Knowledge Graphs: A Comprehensive Survey, CoRR (2021). URL: https://arxiv.org/abs/2112.10006.

[22] S. Wang, A. O. Pisco, A. McGeever, et al., Leveraging the Cell Ontology to Classify Unseen Cell Types, Nature Communications 12 (2021).

[23] I. Donadello, L. Serafini, Compensating Supervision Incompleteness with Prior Knowledge in Semantic Image Interpretation, in: IJCNN, IEEE, 2019, pp. 1–8.

[24] A. Daniele, L. Serafini, Knowledge Enhanced Neural Networks, in: PRICAI 2019: Trends in Artificial Intelligence, volume 11670, 2019, pp. 542–554.

[25] F. van Harmelen, Preface. The 3rd AI Wave is Coming, and it Needs a Theory, in: Neuro-Symbolic Artificial Intelligence: The State of the Art, IOS Press, 2021.

[26] M. van Bekkum, M. de Boer, F. van Harmelen, A. Meyer-Vitali, A. ten Teije, Modular Design Patterns for Hybrid Learning and Reasoning Systems, Appl. Intell. 51 (2021).

[27] F. van Harmelen, A. ten Teije, A Boxology of Design Patterns for Hybrid Learning and Reasoning Systems, J. Web Eng. 18 (2019) 97–124.