Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Perspective:

Multiple knowledge representation for big data artificial intelligence: framework, applications, and case studies*

Yi YANG‡, Yueting ZHUANG, Yunhe PAN

College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China

E-mail: yangyics@zju.edu.cn; yzhuang@zju.edu.cn; panyh@zju.edu.cn

Received Sept. 30, 2021; Revision accepted Oct. 7, 2021; Crosschecked Nov. 22, 2021

https://doi.org/10.1631/FITEE.2100463

In this paper, we present a multiple knowledge representation (MKR) framework and discuss its potential for developing big data artificial intelligence (AI) techniques with possible broader impacts across different AI areas. Typically, canonical knowledge representations and modern representations each emphasize a particular aspect of transforming inputs into symbolic encoding or vectors. For example, knowledge graphs focus on depicting semantic connections among concepts, whereas deep neural networks (DNNs) are more of a tool to perceive raw signal inputs. MKR is an advanced AI representation framework for more complete intelligent functions, such as raw signal perception, feature extraction and vectorization, knowledge symbolization, and logical reasoning. MKR has two benefits: (1) it makes the current AI techniques (dominated by deep learning) more explainable and generalizable, and (2) it expands current AI techniques by integrating MKR to facilitate the mutual benefits of the complementary capacity of each representation, e.g., raw signal perception and symbolic encoding. We expect that MKR research and its applications will drive the evolution of AI 2.0 and beyond.

1 Multiple knowledge representation

In this section, we briefly revisit a few typical knowledge representations, followed by the introduction of the MKR framework (Pan, 2020b).

1.1 Revisiting knowledge representations

A single knowledge representation (KR) scheme usually emphasizes a particular aspect of transforming inputs into symbolic encoding or vectors. We first revisit two typical KRs, i.e., canonical knowledge representation and modern deep representation.

1.1.1 Canonical knowledge representation

Canonical knowledge representation models (e.g., generative representation, first-order logical representation, and procedural representation) take highly abstracted concepts as inputs and depict the causal relationships among them. Typical knowledge/information types abstracted/represented by these models include the following:

1. Declarative knowledge (also known as descriptive knowledge). This is usually expressed in declarative sentences and may include concepts and facts about the object of interest.

2. Procedural knowledge (also known as imperative knowledge). Depending on the specified task, this kind of knowledge consists of rules, strategies, procedures, and agendas. It is more of a knowledge

‡ Corresponding author
* Project supported by the National Key R&D Program of China (No. 2020AAA0108800)
ORCID: Yi YANG, https://orcid.org/0000-0002-0512-880X

representation of reasoning.

3. Heuristic knowledge. Heuristic knowledge includes rules of thumb based on experts, experiences, or other sources.

4. Structural knowledge. This describes the relationships between concepts and objects. A classical knowledge graph is a typical representation of structural knowledge.

1.1.2 Deep representation

The deep representation method takes raw signals (usually at relatively low abstraction levels, e.g., visual and auditory signals) as the input and encodes them in a feature vector through a DNN, such as deep convolutional neural networks (Krizhevsky et al., 2012; He et al., 2016) and Transformers (Vaswani et al., 2017). DNN representation is the major paradigm dominating big data AI research at present.

Deep representation is competent in perceiving unstructured data such as images, videos, audios, texts, or time sequence data for various tasks, e.g., classification and prediction. Compared with canonical representations, deep representation is more capable of uncovering and extracting information/knowledge from large volumes of data. However, unlike canonical representations, the current form of DNN algorithms may not work well for abstracting procedural knowledge and structural knowledge, thereby limiting the DNN reasoning capacity. In addition, a major weakness of DNNs is that, as criticized by researchers recently, the black-box nature of deep representation learning makes the output not explainable (Arrieta et al., 2020). This weakness severely limits the application of DNNs, especially when trustworthiness becomes a concern, e.g., in decision-making in medical scenarios.

1.1.3 Discussions

The aforementioned knowledge representations have limitations as a single measure to extract knowledge from input signals in many real-world problems. The canonical representation relies on pre-symbolized knowledge as a prerequisite. Although humans have accumulated abundant knowledge during the long history of civilization, there are still significant gaps between comprehensive cognition in the real world and the symbolic system derived from human knowledge. Computer-abstracted knowledge as vectors and symbols cannot, therefore, faithfully and fully extract all useful information for complete intelligent computing. Deep representation, on the other hand, can uncover perceptual and latent knowledge from data, but lacks common sense, reasoning capacity, and procedural and structural knowledge in its current form.

To comprehensively understand a concept, humans tend to combine multiple knowledge (Pan, 2020b) including intuitive perception, cognition, highly abstract knowledge, and logic. Leveraging the complementary benefits of many types of knowledge derived from different sources is a common approach in intelligent human activities, such as learning and decision making. This implies that MKR, which integrates multiple knowledge representations via appropriate mechanisms (Pan, 2020b), could be an option for advanced intelligent computing in the era of AI 2.0 and beyond.

1.2 MKR framework

MKR is aimed to acquire, represent, and manipulate knowledge at multiple abstraction levels, from different sources or derived by different approaches. Early examples can be found in the pioneering work of Pan (2020b). Fig. 1 illustrates the main features of MKR. It has the following connotations with possible extensions in characteristics.

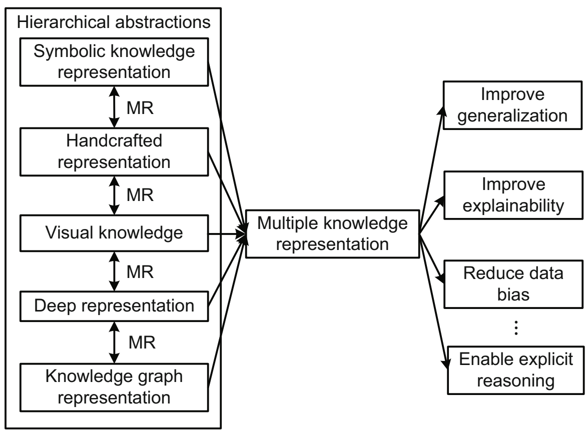

Fig. 1 An example of the multiple knowledge representation framework (MR: mutual reinforcement)

1.2.1 Multi-representation integration

MKR not only combines multiple knowledge representations, but compounds these presentations via an appropriate mechanism. Each knowledge has its own advantages. The main available knowledge representations for current MKR research include symbolic knowledge representation, knowledge graph representation, handcrafted feature representation, and deep representation. The quantity of human knowledge resides in the aforementioned four types of knowledge representations. Specifically, symbolic knowledge representation explicitly relies on expert-defined concepts and causal relationships (sometimes such relationships can be very complex). The knowledge graph consists of a collection of interlinked descriptions of entities and depicts only the relationship between entity pairs or entity chains. These two representations favor highly abstract concepts with emphasis on relationships (e.g., logical or semantic relationships). Handcrafted feature representation and deep representation are more of feature representation of data, and are superior at raw signal perception. The data-driven property of deep feature representation endows it with richer information that extends beyond any existing symbolic system. MKR is designed to leverage the advantages of each knowledge representation while suppressing their disadvantages, as discussed in Section 1.1. Note that new types of knowledge can be integrated in the future, e.g., visual knowledge (Pan, 2019, 2020a).

1.2.2 Multi-level knowledge abstraction

MKR encodes knowledge at different abstraction levels. In this subsection, the abstraction level refers to the degree to which trivial details are removed and elements of higher importance distilled.

In most cases, human perception and cognition is a shallow-to-deep procedure. For example, when recognizing an animal species, humans tend to first observe its appearance and sounds, as an intuitive perception. Information obtained by human senses, such as color, size, and the shape of the animal's teeth, provides more details. This information involves relatively low level abstraction. Higher-level abstractions, such as the living habits and the underlying taxonomy, can then be obtained. In this example, lower- and higher-level abstractions (e.g., animal appearance and habits) are indispensable and complementary. Integration of these abstractions yields a more comprehensive representation than using any one abstraction alone.

MKR enables knowledge fusion at multiple abstraction levels. Recently, a majority of AI applications is still targeted at relatively low abstraction levels. This is probably the reason why most recent AI research favors deep representations. However, there are also cases where high-level abstraction (e.g., common sense and logical reasoning) is also involved. MKR combines representations at different abstraction levels, therefore accommodating AI systems with more functions from perception, recognition to association, reasoning, and many more.

1.2.3 Multi-modal knowledge reinforcement

MKR enables multiple knowledge representations to reinforce each other through effective interactions and deep entanglement of multiple representations. MKR is not a simple combination of multiple representations. For example, in computer vision research, the feature representations, particularly the deep feature representation (He et al., 2020; Sun et al., 2020), are robust for machine perception. The deep feature contains information on visual details, compared with symbolic knowledge. On the other hand, symbolic knowledge can potentially improve the generalization capability of deep feature representation, which is a major weakness of deep learning. For example, entangling the symbolic knowledge that cars may be of different colors, and the visual feature of a red car, an AI model would easily detect a black car. Also, given the observation of cars with different colors, a symbolic system would become more confident that an artifact may be painted with different colors. Therefore, a key MKR research issue involves how MKRs should reinforce each other.

2 Applications and case studies

There is emerging research in areas that could be regarded as early attempts at MKR, in terms of either task objective or methodology.

1. Visual understanding (Pan, 2021). DNNs are powerful feature extractors. Structural information often provides sensible complementary cues to facilitate the understanding of visual contents. Structured representations are used in the process of structured visual understanding. For example, Xu et al. (2017) represented visual scenes as graphs containing objects, attributes, and relationships. The scene graph forms an interpretable and well-structured

representation of images. Researchers also used privileged information as auxiliary features to assist model training. Various types of privileged information can be exploited to facilitate learning. For example, Yan et al. (2016) proposed to exploit both visual and text features for active sample selection by taking text as privileged information. A typical method of integrating multiple cues is to simply apply late fusion. For example, fusion of the optical flow model and the RGB model has been widely used since the two-stream action recognition model was proposed (Simonyan and Zisserman, 2014; Zhu et al., 2021). Recently, Wang et al. (2020) considered a multi-stream framework for ego-centric action recognition. Multiple cues are adaptively integrated with a symbiotic attention mechanism. Mutual interactions are considered and intrinsic relations among these cues are explored. In multi-modal analysis, e.g., visual dialog, narrative structures need to be explored. Fan et al. (2020) proposed a dialog network to learn the contextual narrative structure. The network also transfers the knowledge from a sentence-level discriminator to guide the training of a generative model, which alleviates the problem of word-level overfitting and improves the semantic coherence.

2. Visual-knowledge-assisted computer graphics. Computer graphics studies the process of digitally synthesizing and manipulating visual content. Recently, generative adversarial nets (GANs) (Goodfellow et al., 2014) have provided an alternative for visual generalization. Later works extended GANs into text-to-image, image-to-text, text-to-video, and video-to-text generalizations. As a deep learning method, the GAN requires an astonishingly large amount of training data and lacks interpretability. In response to these weaknesses, some recent studies by Johnson et al. (2018) and Gogoglou et al. (2019) considered adding structured knowledge to better control the generative procedure. Specifically, Gogoglou et al. (2019) controlled the position, attributes, or category of the generated objects. Johnson et al. (2018) proposed to generate images from scene graphs, which enabled explicit reasoning about objects and their relationships.

3. Multimedia knowledge graph with abundant knowledge resources. The rapid development of the Internet has provided access to a large volume of multimedia data. To learn from these abundant knowledge resources, DBpedia (Auer et al., 2007), Wikidata (Vrandečić and Krötzsch, 2014), and IMGpedia (Ferrada et al., 2017) establish knowledge graphs. However, Internet data usually include significant noise and bias. Low-quality image, video, and audio data are much scarcer than text data, so it is critical to mine the cross-media relationship using MKR to improve the quality of multimedia knowledge graphs.

4. Neural-symbolic network. Several researchers have proposed to integrate DNNs and symbolic representations into a hybrid network, called the "neural-symbolic network." França et al. (2014) translated and encoded symbolic knowledge into network weights. Specifically, the neural-symbolic network uses background knowledge that is encoded as an initial propositional logic program to build a recurrent neural network. It also uses examples to apply standard back-propagation learning. This integration inherits both parallel learning from neural networks and explanatory power from propositional logic. Serafini and d'Avila Garcez (2016) proposed logic tensor networks, a uniform framework for integrating automatic learning and reasoning. Specifically, these networks implement "real logic" in DNNs to simultaneously benefit from the deductive reasoning of symbolic knowledge and data-driven machine learning.

Arguably, these works drew early attention to the combination of symbolic knowledge and deep feature representations. This combination may be viewed as a special and degraded case of MKR, because MKR generally has a broader and deeper research interest. In Table 1, we briefly compare the representations that are used in existing works. As explained in Section 1, MKR not only integrates multiple representations and different abstraction levels, but also seeks mutual reinforcement among components through deep entanglement.

Table 1 Representations used in a few recent works

Method	S	H	V	D	K
Scene graph (Xu et al., 2017)	✓	×	×	✓	×
IMGpedia (Ferrada et al., 2017)	×	✓	×	×	✓
LTN (Serafini and d'Avila Garcez, 2016)	✓	×	×	✓	×
MKR (this paper)	✓	✓	✓	✓	✓

S: symbolic knowledge representation; H: handcrafted representation; V: visual knowledge; D: deep representation; K: knowledge graph representation

3 A shift from deep learning representation to MKR in big data AI

Over the past decade, deep learning with big data has significantly advanced the development of AI, with profound impact in both academia and industry. Deep learning representation (DLR) has reshaped the AI research, and is also dominating many applications across domains, such as speech recognition, computer vision, natural language processing, and machine translation. However, the data-driven and black-box nature of DLR also results in a few problems and bottlenecks. MKR reinforces the strengths of different presentations, and will provide possible solutions to overcome the existing problems. In this section, using generalization and explainability as examples, we discuss how MKR could be used to advance the AI research dominated by deep learning. There could be more cases for which MKR could resolve the problems of any single representation.

3.1 MKR improves generalization

MKR improves generalization in two ways. First, it reduces the data bias, often by using symbolic knowledge, e.g., a knowledge graph. Second, it facilitates useful knowledge transfer from richly annotated data to poorly conditioned data or even totally novel data.

1. Data bias is a type of error in which certain elements of a dataset are more heavily weighted or represented than others. Data bias is a prominent challenge that hinders the generalization ability of data-driven AI algorithms. It not only compromises prediction accuracy, but at times may involve ethics and fairness issues. A well-known example is that the accuracy of a facial recognition algorithm has a strong bias on the color of skin. One solution could be injecting a symbolic knowledge representation into the deep learning representation to avoid color bias in face recognition. Another example is the work of Tang et al. (2020), which integrates structural representation and deep learning representation to relieve inference bias.

2. Knowledge transfer adapts previously learned knowledge to a new problem. The knowledge learned from previous tasks is relevant to the new problem, but is always different from previous domains. Knowledge transfer is an effective way of improving the model's ability to generalize. The domain gaps between the previous and new tasks are major problems to be confronted. MKR reinforces the symbolic knowledge representation and deep learning representation, making it capable of disentangling the domain gap into multiple underlying factors. Then MKR filters out irrelevant factors from the disentanglement outputs, and distills only useful knowledge for the new problem. For example, in the pedestrian detection task, a typical deep learning algorithm could be easily corrupted by clothing style changes. With symbolic knowledge that clothing style is an irrelevant factor, the detection algorithm will disregard clothing style information and thus improve robustness. Moreover, recent research shows that incorporating body structure representation in deep learning would improve the identification of persons by using models trained on holistic body images (Miao et al., 2021).

3.2 MKR improves explainability

Another limitation of deep learning is its "black-box" nature. Even the model designers cannot figure out why the algorithm arrives at a specific decision. Without explainability, AI accountability is limited, which in turn hinders the application of Al in certain domains where safety is a major concern (Amodei et al., 2016).

In contrast to a "black-box" AI system, a "white-box" AI system or explainable AI (XAI) system is a transparent model that can explain how an AI decision was made. For this purpose, an AI system should not rely solely on data, but must be elevated with human-understandable mechanisms, such as regularization mechanisms that reflect human knowledge. MKR provides a mechanism to entangle representations of data-driven knowledge and symbolic knowledge, making itself a highly attractive option for establishing XAI systems.

3.3 Changes made by MKR

We discuss several recently attainable changes for better generalization and explainability. First, we show that the AI research methodology could be changed by MKR. Then we discuss possible changes to applications that are used to assess investments.

3.3.1 Injecting symbolic knowledge into synthetic images for robustness

In computer vision, an approach for alleviating the burden of too many training data is to supplement the massive manually annotated data with synthetic data (de Souza et al., 2017; Veeravasarapu et al., 2017; Singh and Zheng, 2020). Using MKR in this process increases data diversity and injects some useful symbolic knowledge, which consequently benefits the AI models to be trained. Three examples are specified as follows:

1. By introducing climate and geography knowledge, the image generative model can simulate changes in the weather, scenery, and pedestrian clothing style. Using such synthetic data for training, the corresponding AI system gains rich visual knowledge from the image and the ability to deduce the effects of seasonal changes.

2. With knowledge of animal body structures and kinematics, the generative model will be able to simulate animals based on static shape and the dynamic postures of walking, running, and jumping. These synthetic data allow the AI system to learn the inherent relationships between different animals, their movements, and body structure from the video.

3. Employing the principle of light refraction and diffuse reflection, the image generation engine can simulate the color and morphology of various materials under different lighting conditions. This allows the AI system to overcome the domain adaptation challenge that is created by different lighting conditions in the current visual system.

3.3.2 Tackling automated student grouping for intelligent AI education

Automated student grouping is one of the major problems in intelligent AI education systems. The heterogeneous nature of educational data hinders the development of automated student grouping algorithms. Student grouping depends on multiple heterogeneous information in the form of audios, texts, videos, and structured tables. The information includes individual activities on the learning platform, cooperative dialogues among teammates, historical student–teacher interactions, and other information related to the students' learning experience. The quality of student grouping will rapidly influence the students' engagement, which has a high impact on groupwork communication, teaching management, etc. MKR enables a machine to automatically discover causal relationships and structural dependencies among heterogeneous cues, through its embedded graph operations and knowledge reinforcement capabilities.

For example, when applying MKR to automated student grouping, we will first map student–student interactions, student–teacher interactions, and a heuristic education–expert knowledge graph into a joint graph space. Second, a multi-knowledge reasoning procedure will be used to extract causal relations and abstract multi-level knowledge presentations, guided by the symbolic knowledge graph. Last, we will provide an explainable student–student relation graph where graph edges represent the grouping weights. Each grouping weight determines a statistical correlation between two students, which can be interpreted and explained via the students' nearest node from the education-expert knowledge graph. The final group recommendation can be readily visualized for teachers' decision making. In this case, MKR can better align heterogeneous data from student–student and student–teacher interactions. MKR will benefit online education systems in achieving automated student grouping and providing knowledge-driven learning experiences.

3.3.3 Sound explainability provides better FinTech for assessing investment opportunities

Robo-advisor is an important AI application in the financial field. It recommends investments according to a customer's investment interest, based on portfolio theory. Arguably, for personal investment advice, the major objective is to maximize capital benefits. In contrast, responsibilities for national investment extend beyond capital growth. These responsibilities include balanced regional development, regulating the poverty gap, environmental protection and sustainable development, and many other factors. To create responsible investment advice, one possible solution would be to include MKR. Specifically, one can use sociological knowledge (such as economics, politics, and geography) to make interpretable predictions and suggestions, and provide interactive investment information for decision makers. For example, to balance regional development, the robo-advisor needs to adjust its strategy according to the MKR of geographical differences among regions, industrial foundations, and even local residents. For appropriate and balanced development, the robo-advisor should employ MKR with information on biology and earth science to minimize the risk of environmental damage.

4 Conclusions

This paper introduces an MKR framework along with application examples and case studies. MKR is a new knowledge representation paradigm that learns from different abstraction levels, different sources, and different perspectives. These knowledge representations are deeply entangled with, and reinforced by, each other. Big data AI with MKR not only improves the accuracy of classical tasks, such as detection and recognition, but also equips an AI system with more features and functions, such as better generalization, explainable outputs, and stronger reasoning capacity. We expect that MKR will become a new tool of the AI 2.0 evolution and beyond.

Contributors

Yunhe PAN conceptualized the main idea and led the research. Yi YANG and Yueting ZHUANG surveyed the relevant materials. Yi YANG, Yueting ZHUANG, and Yunhe PAN had in-depth discussions; they drafted, revised, and finalized the paper.

Acknowledgements

The authors thank Drs. Yifan SUN, Linchao ZHU, Xiaohan WANG, and Yu WU for constructive discussions.

Compliance with ethics guidelines

Yi YANG, Yueting ZHUANG, and Yunhe PAN declare that they have no conflict of interest.

References

Amodei D, Olah C, Steinhardt J, et al., 2016. Concrete problems in AI safety. https://arxiv.org/abs/1606.06565v2

Arrieta AB, díaz-Rodríguez N, Del Ser J, et al., 2020. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inform Fus*, 58:82-115. https://doi.org/10.1016/j.inffus.2019.12.012

Auer S, Bizer C, Kobilarov G, et al., 2007. DBpedia: a nucleus for a web of open data. Proc 6th Int Semantic Web Conf and 2nd Asian Semantic Web Conf the Semantic Web, p.722-735. https://doi.org/10.1007/978-3-540-76298-0_52

de Souza CR, Gaidon A, Cabon Y, et al., 2017. Procedural generation of videos to train deep action recognition networks. IEEE Conf on Computer Vision and Pattern Recognition, p.2594-2604. https://doi.org/10.1109/CVPR.2017.278

Fan HH, Zhu LC, Yang Y, et al., 2020. Recurrent attention network with reinforced generator for visual dialog. *ACM Trans Multim Comput Commun Appl*, 16(3):78. https://doi.org/10.1145/3390891

Ferrada S, Bustos B, Hogan A, 2017. IMGpedia: a linked dataset with content-based analysis of Wikimedia images. Proc 16th Int Semantic Web Conf on the Semantic Web, p.84-93. https://doi.org/10.1007/978-3-319-68204-4_8

França MVM, Zaverucha G, d'Avila Garcez AS, 2014. Fast relational learning using bottom clause propositionalization with artificial neural networks. *Mach Learn*, 94(1): 81-104. https://doi.org/10.1007/s10994-013-5392-1

Gogoglou A, Bruss CB, Hines KE, 2019. On the interpretability and evaluation of graph representation learning. https://arxiv.org/abs/1910.03081

Goodfellow IJ, Pouget-Abadie J, Mirza M, et al., 2014. Generative adversarial nets. Proc 27th Int Conf on Neural Information Processing Systems, p.2672-2680.

He KM, Zhang XY, Ren SQ, et al., 2016. Deep residual learning for image recognition. IEEE Conf on Computer Vision and Pattern Recognition, p.770-778. https://doi.org/10.1109/CVPR.2016.90

He KM, Fan HQ, Wu YX, et al., 2020. Momentum contrast for unsupervised visual representation learning. IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.9726-9735. https://doi.org/10.1109/CVPR42600.2020.00975

Johnson J, Gupta A, Li FF, 2018. Image generation from scene graphs. IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.1219-1228. https://doi.org/10.1109/CVPR.2018.00133

Krizhevsky A, Sutskever I, Hinton GE, 2012. ImageNet classification with deep convolutional neural networks. Proc 25th Int Conf on Neural Information Processing Systems, p.1097-1105.

Miao JX, Wu Y, Yang Y, 2021. Identifying visible parts via pose estimation for occluded person re-identification. *IEEE Trans Neur Netw Learn Syst*, early access. https://doi.org/10.1109/TNNLS.2021.3059515

Pan YH, 2019. On visual knowledge. *Front Inform Technol Electron Eng*, 20(8):1021-1025. https://doi.org/10.1631/FITEE.1910001

Pan YH, 2020a. Miniaturized five fundamental issues about visual knowledge. *Front Inform Technol Electron Eng*, early access. https://doi.org/10.1631/FITEE.2040000

Pan YH, 2020b. Multiple knowledge representation of artificial intelligence. *Engineering*, 6(3):216-217. https://doi.org/10.1016/j.eng.2019.12.011

Pan YH, 2021. On visual understanding. *Front Inform Technol Electron Eng*, early access. https://doi.org/10.1631/FITEE.2130000

Serafini L, d'Avila Garcez A, 2016. Logic tensor networks: deep learning and logical reasoning from data and knowledge. Proc 11th Int Workshop on Neural-Symbolic Learning and Reasoning Co-located with the Joint Multi-conf on Human-Level Artificial Intelligence.

Simonyan K, Zisserman A, 2014. Two-stream convolutional networks for action recognition in videos. Proc 27th Int Conf on Neural Information Processing Systems, p.568-576.

Singh J, Zheng L, 2020. Combining semantic guidance and deep reinforcement learning for generating human level paintings. https://arxiv.org/abs/2011.12589

Sun YF, Cheng CM, Zhang YH, et al., 2020. Circle loss: a unified perspective of pair similarity optimization. IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.6397-6406.
https://doi.org/10.1109/CVPR42600.2020.00643

Tang KH, Niu YL, Huang JQ, et al., 2020. Unbiased scene graph generation from biased training. IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.3716-3725.
https://doi.org/10.1109/CVPR42600.2020.00377

Vaswani A, Shazeer N, Parmar N, et al., 2017. Attention is all you need. Proc 31st Int Conf on Neural Information Processing Systems, p.6000-6010.

Veeravasarapu V, Rothkopf C, Visvanathan R, 2017. Adversarially tuned scene generation. IEEE Conf on Computer Vision and Pattern Recognition, p.6441-6449.
https://doi.org/10.1109/CVPR.2017.682

Vrandečić D, Krötzsch M, 2014. Wikidata: a free collaborative knowledge base. *Commun ACM*, 57(10):78-85.
https://doi.org/10.1145/2629489

Wang XH, Zhu LC, Wu Y, et al., 2020. Symbiotic attention for egocentric action recognition with object-centric alignment. *IEEE Trans Patt Anal Mach Intell*, early access.
https://doi.org/10.1109/TPAMI.2020.3015894

Xu DF, Zhu YK, Choy CB, et al., 2017. Scene graph generation by iterative message passing. IEEE Conf on Computer Vision and Pattern Recognition, p.3097-3106.
https://doi.org/10.1109/CVPR.2017.330

Yan Y, Nie FP, Li W, et al., 2016. Image classification by cross-media active learning with privileged information. *IEEE Trans Multim*, 18(12):2494-2502.
https://doi.org/10.1109/TMM.2016.2602938

Zhu LC, Fan HH, Luo YW, et al., 2021. Temporal cross-layer correlation mining for action recognition. *IEEE Trans Multim*, early access.
https://doi.org/10.1109/TMM.2021.3057503