

Review

Not peer-reviewed version

Zero Shot Learning Recent Advances in Robotics

[Sunil Singh](#) *

Posted Date: 20 June 2023

doi: 10.20944/preprints202306.1353.v1

Keywords: zero-shot learning; robotics



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Zero Shot Learning Recent Advances in Robotics

Sunil Singh

University of Pune

Abstract: Zero-shot learning ability robotics enables robots to generalize their knowledge and perform tasks for which they have not been explicitly trained. This ability is extremely helpful because in classic learning approaches, robots are trained on specific tasks and require extensive data and labeled examples to perform accurately. However, in real-world scenarios, robots often encounter new or unseen tasks that were not part of their training data. We briefly survey works in this area to give the reader a perspective about progresses in this area.

Key words: zero-shot learning; robotics

Introduction

Despite significant and remarkable benefits, deep learning has a drawback related to the requirement of a large amount of labeled data points. However, a solution to this limitation is zero-shot learning, which allows for the learning of new tasks or recognition of new objects without direct training on them (Pourpanah et al. 2022; Wang et al. 2019; Xian, Schiele, and Akata 2017). Zero-shot learning achieves this by utilizing transfer learning (Feng and Zhao 2020) and semantic relationships (Zhang and Saligrama 2015), enabling the generalization of knowledge obtained from previously learned tasks or objects to effectively perform in novel situations.

In traditional deep learning approaches, a significant amount of labeled data is necessary for training the model to perform specific tasks. This requirement can be burdensome as collecting and labeling large datasets can be time-consuming and expensive. Zero-shot learning presents an alternative approach that mitigates this need for exhaustive training data. In zero-shot learning, the focus shifts from training the model explicitly on each task or object to leveraging transfer learning and semantic relationships. Transfer learning allows the model to utilize knowledge gained from previously learned tasks or objects and apply it to new and unseen ones. By transferring this knowledge, the model can effectively generalize and adapt its understanding to novel situations.

Semantic relationships play a crucial role in zero-shot learning. They provide a way to establish connections between different tasks or objects based on their inherent similarities or attributes. By understanding these relationships, the model can make intelligent inferences and predictions about unseen tasks or objects. This process is particularly useful when faced with limited or no labeled data for new tasks or objects, as the model can rely on the knowledge it has acquired from related tasks or objects. By combining transfer learning and semantic relationships, zero-shot learning empowers deep learning models to handle new tasks or recognize new objects without the need for direct training on them. This approach opens up possibilities for faster and more efficient adaptation to new scenarios, as well as reducing the dependency on extensive labeled datasets.

Zero-shot learning offers the ability for models to effectively classify instances that belong to classes not encountered during the training phase. This capability becomes particularly valuable in situations where new classes emerge over time or when it becomes impractical or costly to gather labeled data for all possible classes. Moreover, zero-shot learning addresses the challenge of extensive annotation efforts by leveraging auxiliary information (Fu et al. 2015; Rostami, Isele, and Eaton 2020) or semantic embeddings (Ren et al. 2023; Zhang and Saligrama 2015) that provide descriptions of the relationships between different tasks. Consequently, the need for creating labeled datasets is significantly reduced, leading to notable time savings.

One of the key advantages of zero-shot learning is its scalability and flexibility in handling a vast number of classes (Rohrbach, Stark, and Schiele 2011). Rather than training separate models for each individual class, a single model can be trained to recognize and classify a large number of classes

by leveraging semantic relationships. This approach proves highly beneficial in scenarios where the number of classes is too extensive to feasibly collect training samples for each specific class. By utilizing zero-shot learning techniques, models can adapt to new classes without requiring explicit training on them. This adaptability is crucial in dynamic environments where new classes continually emerge. Instead of starting from scratch and training a new model each time a new class appears, zero-shot learning allows existing models to leverage their knowledge of related tasks or objects to recognize and classify instances from unseen classes. This adaptability enables the model to stay up-to-date with the evolving nature of the data it encounters. Additionally, zero-shot learning addresses the challenges posed by the practical limitations of data collection. Collecting labeled data for every possible class can be resource-intensive and time-consuming. Zero-shot learning reduces this burden by capitalizing on auxiliary information or semantic embeddings that capture the relationships between tasks or objects. By leveraging this information, the model can generalize its understanding and make accurate predictions for new and unseen classes without requiring extensive labeled data.

Zero Shot Learning in Robotics

The capability of zero-shot learning also extends to the domain of robotics, enabling robots to adapt and learn rapidly, thereby enhancing their versatility and flexibility in dynamic environments ([Abderrahmane et al. 2018](#); [Cui et al. 2022](#)). By incorporating zero-shot learning techniques, robots can acquire an understanding of the attributes and characteristics of new tasks or objects based on their prior knowledge. This understanding can then be utilized to successfully complete the task at hand or recognize the object in question ([Thomason and Knepper 2017](#)). The application of zero-shot learning expands the range of tasks that robots can effectively handle without the need for extensive retraining or human intervention, ultimately bolstering their autonomy and problem-solving capabilities in various robotic applications.

In dynamic environments, robots often encounter situations where they need to perform tasks or recognize objects that were not part of their initial training ([Lesort et al. 2019](#)). In such cases, zero-shot learning becomes invaluable as it allows robots to leverage their existing knowledge to quickly adapt and learn new tasks or objects ([Wei et al. 2020](#)). This adaptability significantly reduces the time and effort required to retrain robots for specific tasks, making them more efficient and adaptable in ever-changing environments. Zero-shot learning empowers robots to understand the underlying concepts and shared attributes between known and unknown tasks or objects. By recognizing these relationships, robots can make intelligent inferences and predictions, enabling them to successfully perform tasks or identify objects that were not part of their initial training dataset. This ability to transfer knowledge and generalize understanding plays a vital role in enhancing the problem-solving capabilities of robots.

Moreover, zero-shot learning minimizes the need for human intervention in the retraining process ([Kankuekul et al. 2012](#); [Rezaei and Shahidi 2020](#)). Instead of relying on humans to manually update the robot's knowledge base for new tasks or objects, zero-shot learning allows robots to autonomously adapt and learn based on the semantic relationships and transferable knowledge they have acquired. This autonomy reduces the dependency on human experts, streamlining the learning process and enabling robots to quickly adapt to new scenarios. The versatility and flexibility provided by zero-shot learning have profound implications across a wide range of robotic applications. For instance, in industrial automation, robots can seamlessly handle new manufacturing tasks without requiring extensive reprogramming. In robotic perception, robots can recognize and understand novel objects in dynamic environments, enabling them to perform complex manipulation tasks. Additionally, in interactive robotics, zero-shot learning allows robots to learn and understand new commands or gestures from human users, facilitating more natural and intuitive human-robot interaction ([Madapana and Wachs 2019](#)).

Zero Shot Learning for Reinforcement Learning

Reinforcement learning (RL) offers significant advantages for robotics due to its ability to learn through interaction and adapt to dynamic environments (Kober, Bagnell, and Peters 2013). RL allows robots to learn through trial and error by interacting with their environment. Instead of relying solely on pre-programmed instructions, RL agents can explore different actions and receive feedback in the form of rewards or penalties. This iterative learning process enables robots to discover optimal strategies and adapt their behavior based on the outcomes of their actions. RL also excels in environments where the dynamics or task requirements may change over time. Robots equipped with RL algorithms can continuously update their policies based on new experiences, allowing them to adapt to unforeseen circumstances or variations in the environment. This adaptability is crucial in real-world scenarios where the robot must cope with changing conditions or interact with humans and other dynamic agents.

RL also allows robots to make complex decisions by optimizing long-term cumulative rewards. In tasks with high-dimensional state and action spaces, RL algorithms can handle the complexity and provide solutions by learning effective policies. This capability enables robots to navigate complex environments, manipulate objects, perform delicate tasks, and make intelligent decisions in real-time (Kober, Bagnell, and Peters 2013).

Zero-shot learning (ZSL) and reinforcement learning (RL) can complement each other in certain scenarios (Xian et al. 2018). An important limitation of RL is that it typically requires a large amount of interaction with the environment to learn effective policies. This can be time-consuming, costly, and even risky in real-world robotics applications. By combining ZSL and RL, it is possible to address certain challenges in robotics. ZSL can enable the transfer of knowledge from one robot to another, even if they have different physical configurations or sensory capabilities (Xian et al. 2018). By leveraging shared information or prior knowledge, the target robot can benefit from the experiences and policies learned by a source robot. In certain situations, the environment or task requirements may change slightly, and retraining an RL agent from scratch might not be feasible. ZSL techniques can assist in adapting the existing policies or fine-tuning them with limited or no additional data, reducing the time and effort required for adaptation (Xian et al. 2018). Finally, ZSL can help bridge the gap between language and robotic actions. By associating textual descriptions or natural language instructions with robot behaviors, ZSL can facilitate human-robot communication, enabling robots to understand and execute commands in a zero-shot manner (Xian et al. 2018).

Zero Shot Learning in Imitation Learning

When it comes to imitation learning, the objective is for a robot or agent to acquire a policy through reinforcement learning (RL) by imitating human demonstrations (Osa et al. 2018). The conventional approach to imitation learning involves utilizing a dataset consisting of labeled demonstrations, enabling the robot to replicate the observed behavior. However, a drawback of the traditional imitation learning paradigm lies in its dependency on having demonstrations available for every conceivable action or task that the robot might encounter. In practical terms, this requirement can prove to be impractical or even unattainable, particularly in environments characterized by complexity and diversity. This is where zero shot Learning enters the picture.

Zero shot learning provides a solution to the limitations of traditional imitation learning. In zero shot Learning, the aim is to enable robots or agents to perform tasks that have not been explicitly demonstrated to them. Instead of relying solely on labeled demonstrations, zero shot Learning leverages the power of transfer learning and generalization to equip robots with the ability to extrapolate their learned behaviors to novel situations and tasks. By employing zero shot Learning, robots can effectively bridge the gap between the known and unknown, allowing them to operate in real-world scenarios where the availability of labeled demonstrations may be limited or non-existent. This approach empowers robots to adapt and learn from a smaller set of demonstrations, while still being capable of applying their acquired knowledge to perform a broader range of tasks.

Zero shot learning introduces a valuable capability for the agent to extend its acquired behavior to unfamiliar actions or tasks, leveraging auxiliary information or prior knowledge (Jang et al. 2022). This approach empowers the agent to establish associations and comprehend the relationships among various actions or tasks, guided by their underlying semantic or conceptual similarities. By incorporating zero shot learning into the domain of imitation learning, the agent gains the ability to imitate demonstrated behaviors pertaining to known actions or tasks, while also expanding its knowledge to encompass previously unseen actions or tasks (Pan et al. 2020). This is made possible by leveraging shared semantic information or prior knowledge, acting as a bridge between the familiar and the unknown actions.

The utilization of zero shot learning in imitation learning enables the agent to learn from a limited set of labeled demonstrations for known actions or tasks, while also benefiting from the transfer of knowledge to perform novel actions or tasks. By leveraging the common semantic attributes or conceptual connections between the known and unknown actions, the agent can generalize its learned behaviors to unobserved scenarios, adapting its actions based on the available prior knowledge. This way, the agent acquires a more comprehensive and flexible understanding of the task domain, extending its capabilities beyond what has been explicitly demonstrated.

For example, let's say an agent has learned to imitate human demonstrations of various grasping tasks. With zero shot learning, the agent can generalize its grasping skills to perform a new, unseen grasping task by leveraging the semantic relationships between known grasping tasks and the new task. By understanding the common underlying principles or concepts of grasping, the agent can transfer its knowledge and adapt its grasping policy to the new task without explicitly being shown a demonstration for that specific task (Du et al. 2022).

Conclusions and Discussions

We briefly delved into the concept of zero shot learning and its application in the field of robotics, particularly when reinforcement learning is the learning mechanisms. Zero shot learning offers a compelling solution to the limitations that traditional learning methods present for robots, particularly when it comes to adapting and learning in dynamic environments. By leveraging the principles of transfer learning and semantic relationships, robots can extend their knowledge from previously learned tasks or objects to effectively handle new and unfamiliar ones.

The integration of zero shot learning into robotics brings forth a multitude of exciting possibilities across various domains of robotics. Industries employing automation can greatly benefit from this approach, as robots equipped with zero shot learning capabilities can swiftly adapt to changing manufacturing processes or new product lines. Moreover, in the realm of robotic perception, zero shot learning empowers robots to recognize and understand objects or scenarios they have never encountered before, expanding their understanding and enabling them to navigate previously unexplored environments. Additionally, in interactive robotics, zero shot learning allows robots to quickly grasp and respond to user commands or interact with humans in a more natural and intuitive manner, broadening adoption of deep learning in robotics. Zero shot learning can also facilitate human robot interaction and having humans in the learning loop.

Nevertheless, despite its numerous advantages, zero shot learning in robotics also poses certain challenges that need to be addressed. One key challenge lies in accurately modeling the semantic relationships between different tasks or objects, particularly when RL tasks are involved, ensuring that the transferred knowledge is relevant and applicable to new situations. Additionally, efficiently transferring knowledge from one domain to another remains an ongoing area of research, as robots must be able to leverage their existing knowledge effectively while adapting to novel contexts. Moreover, handling ambiguous or unseen situations presents another hurdle, as robots need to possess the capability to make informed decisions even in unfamiliar scenarios.

To overcome these challenges, further research and development efforts are necessary. Advancements in modeling semantic relationships, developing efficient transfer learning algorithms,

and enhancing robots' ability to handle uncertainty and ambiguity are crucial for realizing the full potential of zero shot learning in robotics. By continuously pushing the boundaries of knowledge and innovation, zero shot learning is poised to play a pivotal role in shaping the future of robotics, developing autonomous robots, enabling robots to tackle complex and ever-changing real-world challenges with efficiency, adaptability, and intelligence.

References

- Abderrahmane, Z.; Ganesh, G.; Crosnier, A.; and Cherubini, A. 2018. Haptic zero-shot learning: Recognition of objects never touched before. *Robotics and Autonomous Systems* 105:11–25.
- Cui, Y.; Niekum, S.; Gupta, A.; Kumar, V.; and Rajeswaran, A. 2022. Can foundation models perform zero-shot task specification for robot manipulation? In *Learning for Dynamics and Control Conference*, 893–905. PMLR.
- Du, M.; Lee, O. Y.; Nair, S.; and Finn, C. 2022. Play it by ear: Learning skills amidst occlusion through audio-visual imitation learning. *arXiv preprint arXiv:2205.14850*.
- Feng, L., and Zhao, C. 2020. Transfer increment for generalized zero-shot learning. *IEEE Transactions on Neural Networks and Learning Systems* 32(6):2506–2520.
- Fu, Y.; Hospedales, T. M.; Xiang, T.; and Gong, S. 2015. Transductive multi-view zero-shot learning. *IEEE transactions on pattern analysis and machine intelligence* 37(11):2332–2345.
- Jang, E.; Irpan, A.; Khansari, M.; Kappler, D.; Ebert, F.; Lynch, C.; Levine, S.; and Finn, C. 2022. Bc-z: Zero-shot task generalization with robotic imitation learning. In *Conference on Robot Learning*, 991–1002. PMLR.
- Kankuekul, P.; Kawewong, A.; Tangruamsub, S.; and Hasegawa, O. 2012. Online incremental attribute-based zero-shot learning. In *2012 IEEE conference on computer vision and pattern recognition*, 3657–3664. IEEE.
- Kober, J.; Bagnell, J. A.; and Peters, J. 2013. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research* 32(11):1238–1274.
- Lesort, T.; Lomonaco, V.; Stoian, A.; Maltoni, D.; Filliat, D.; and Díaz-Rodríguez, N. 2019. Continual learning for robotics. *arXiv preprint arXiv:1907.00182* 1–34.
- Madapana, N., and Wachs, J. 2019. Database of gesture attributes: Zero shot learning for gesture recognition. In *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 1–8. IEEE.
- Osa, T.; Pajarinen, J.; Neumann, G.; Bagnell, J. A.; Abbeel, P.; Peters, J.; et al. 2018. An algorithmic perspective on imitation learning. *Foundations and Trends® in Robotics* 7(1-2):1–179.
- Pan, X.; Zhang, T.; Ichter, B.; Faust, A.; Tan, J.; and Ha, S. 2020. Zero-shot imitation learning from demonstrations for legged robot visual navigation. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 679–685. IEEE.
- Pourpanah, F.; Abdar, M.; Luo, Y.; Zhou, X.; Wang, R.; Lim, C. P.; Wang, X.-Z.; and Wu, Q. J. 2022. A review of generalized zero-shot learning methods. *IEEE transactions on pattern analysis and machine intelligence*.
- Ren, W.; Tang, Y.; Sun, Q.; Zhao, C.; and Han, Q.-L. 2023. Visual semantic segmentation based on few/zero-shot learning: An overview. *IEEE/CAA Journal of Automatica Sinica*.
- Rezaei, M., and Shahidi, M. 2020. Zero-shot learning and its applications from autonomous vehicles to covid-19 diagnosis: A review. *Intelligence-based medicine* 3:100005.
- Rohrbach, M.; Stark, M.; and Schiele, B. 2011. Evaluating knowledge transfer and zero-shot learning in a large-scale setting. In *CVPR 2011*, 1641–1648. IEEE.
- Rostami, M.; Isele, D.; and Eaton, E. 2020. Using task descriptions in lifelong machine learning for improved performance and zero-shot transfer. *Journal of Artificial Intelligence Research* 67:673–704.
- Thomason, W., and Knepper, R. A. 2017. Recognizing unfamiliar gestures for human-robot interaction through zero-shot learning. In *2016 International Symposium on Experimental Robotics*, 841–852. Springer.
- Wang, W.; Zheng, V. W.; Yu, H.; and Miao, C. 2019. A survey of zero-shot learning: Settings, methods, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)* 10(2):1–37.
- Wei, K.; Deng, C.; Yang, X.; et al. 2020. Lifelong zero-shot learning. In *IJCAI*, 551–557.
- Xian, Y.; Lorenz, T.; Schiele, B.; and Akata, Z. 2018. Feature generating networks for zero-shot learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 5542–5551.

Xian, Y.; Schiele, B.; and Akata, Z. 2017. Zero-shot learning-the good, the bad and the ugly. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4582–4591.

Zhang, Z., and Saligrama, V. 2015. Zero-shot learning via semantic similarity embedding. In *Proceedings of the IEEE international conference on computer vision*, 4166–4174.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.