International Journal of Research Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 01 Issue 01
February 2014

A Survey of Deep Learning Techniques for Mobile Robot Applications

[1] Prasad Layam,
[2] Dr. Vitushi Sarma
[1]Research Scholor, [2] Professor, ECE, University of Allahabad

Abstract: Advancements in deep learning over the years have attracted research into how deep artificial neural networks can be used in robotic systems. It is on this basis that the following research survey will present a discussion of the applications, gains, and obstacles to deep learning in comparison to physical robotic systems while using modern research as examples. The research survey will present a summarization of the current research with specific focus on the gains and obstacles in comparison to robotics. This will be followed by a primer on discussing how notable deep learning structures can be used in robotics with relevant examples. The next section will show the practical considerations robotics researchers desire to use in regard to deep learning neural networks. Finally, the research survey will show the shortcomings and solutions to mitigate them in addition to discussion of the future trends. The intention of this research is to show how recent advancements in the broader robotics field can inspire additional research in applying deep learning in robotics.

Index Terms: Deep learning, robotic vision, navigation, autonomous driving, deep reinforcement learning, algorithms for robotic perception, Semi-supervised and self-supervised learning, Deep learning architectures, multimodal, decision making and control.

INTRODUCTION

Deep learning is defined as the field of science that involves training extensive artificial neural networks using complex functions, for example, nonlinear dynamics to change data from a raw, high-dimension, multimodal state to that which can be understood by a robotic system. However, deep learning entails certain shortcomings which affect physical robotic systems whereby generation of training data in overall is costly and therefore sub-optimal performance in the course of training poses a risk in certain applications. Yet, even with such difficulties, robotics researchers are searching for creative options, for instance, leveraging training data through digital manipulation, automated training and using multiple deep neural networks to improve the performance and lower the time for training.

One type of layer that demands specific mention is convolutional layers. Unlike traditional layers that are fully connected, convolutional layers apply the same weights in order to operate in all the input space. This brings about a significant reduction of the overall number of weights in the neural network which is specifically vital with images that normally compose of hundreds of thousands and millions of pixels that require processing. It should be noted that processing these kinds of images which have fully connected layers would need over 100K2 to 1M2 weights which connect to each layer which makes it entirely impractical.

The inspiration of convolution layers came from cortical neurons within the visual cortex which only respond to stimuli in a receptive environment.

It is therefore on this premise that deep learning has been identified as being effective in managing multimodal data generation in robotic sensor applications. These applications include integration of vision and haptic sensor data, incorporating depth data and image information from RGB-D camera data. Due to the extensive number of meta-parameters, deep neural networks have evolved somewhat a reputation of being challenging for non-experts to be used effectively. However, such parameters also avail significant flexibility which is a vital factor in their general success. Therefore, training deep neural networks needs the user to be able to develop at least an elementary level of familiarization with many concepts. Specifically, applying these techniques will help in tacking advanced object recognition challenges and reduced the extent of the entire changes as well.

DEEP LEARNING FOR ROBOTIC PERCEPTION

Although current trends are more leaning to deep and big models, a simplified neural network with just a single hidden layer and a basic sigmoid shaped activation function will train faster and provide a baseline that is used to give meaning to any deeper model improvements. When we use deeper models, Leaky Rectifiers are able to normally promote faster training by lowering the impact of the diminishing gradient challenge and improving accuracy through using simplified monotonic derivatives. Furthermore, since models with additional weights have increased flexibility to over fitting training data, regularization is a vital technique in training the best model.

The difference between deep learning and machine learning is that deep learning place emphasis on the subset of machine learning resources and method and uses them to solve any difficulties that need "thought" whether human or artificial. Deep learning is also introduced as a means of making sense of data with the use of multiple abstraction layers. In the course of the training process, deep neural networks are able to learn the means of discovering useful patterns to digitally represent data such as sounds and images. This is specifically why we observe more advances in the areas of image recognition and natural language processing originating from deep learning. It is on this backdrop that deep learning has taken the forefront position in helping researchers develop breakthrough methods to the perception capabilities of robotics systems. In more simplified language, perception refers to the functionality of robots being able to detect its surroundings. It is therefore heavily reliant on multiple sources of sensory information. However, with traditional robot technology extracting data from raw sensor by using rudimentary constructed sensors these old methods were limited by constraints of adapting to generic settings.

However, certain challenges remain unresolved in these robotic systems particularly the areas of perception and intelligent control. Some of these challenges are reflected in the process needing a lot of data to be able to train and teach algorithms progressively. Large datasets are required to ensure machines deliver the desired outcomes. In the same way as the human brain needs rich experience to learn and deduce information, artificial neural networks also need abundant amounts of data. This means for more powerful abstractions, more parameters are required and hence more data. Another challenge is the tendency of over fitting in neural networks whereby in certain cases, there is a sharp distinction between an error within a training set and that encountered into new untrained datasets. This problem arises when many models make the relative number of parameters fail to reliably perform. Therefore, the model only memorizes training examples and fails to learn generalization of new situations and new datasets.

DEEP LEARNING FOR ROBOTIC CONTROL AND EXPLORATION

Realizing the benefits of autonomous robot exploration presents robotics researchers with many applications of considerable community and financial impact. Robotics research relies on perfect knowledge and control of the environment. The problems related to unstructured environments are an outcome of the high-dimensional state space as well as the inherent likelihood in mapping sensory perceptions on particular states. It should be noted that the high dimensionality of the state space is representative of the most basic difficulty since robots leave highly controlled environments of a laboratory and enter into unstructured surroundings. For example autonomous unmanned aerial vehicles used deep learning to classify terrain and solve any exploration shortcomings by generating control commands for its human operator so as to adapt to a certain tradeoff. The major hypothesis of this approach is therefore for mobile robots to succeed in unstructured surroundings such that they can carefully choose assignment specific attributes and identify the relevant real-time structures to lower their state space without impacting the performance of their exploration objectives. Robots perform assignments by exploring their surroundings. As such, given our focus on autonomous mobile exploration, we shall direct most attention to exploration in service of movement, that is to say collision-free movement for end-effector placement. The challenge of generating such movement is an example of the problem faced in motion planning. Motion planning for robotic systems with many levels of freedom is computationally challenging even in environments that are highly structured due to the increased-dimensional configuration space.

DEEP LEARNING IN ROBOTIC ROBOTIC NAVIGATION AND AUTONOMOUS DRIVING

Using deep learning to attain autonomous driving assignment is not a perfectly controlled and modeled task as most people think. Instead, it needs optimal perceptual capabilities. The process of perception of a robots environment and interpreting the information it acquires allows it to understand the condition of its surroundings, devise plans to change the state and observe how its actions impact its environment. In unstructured environments, recognition of objects has been proven to be highly challenging. With immense volumes of sensor data and increased variation of objects within similar object categories, for example, a paved and unpaved road as well as recognizing objects. Deep learning uses machine learning motion capturing abilities as well as optimal perceptual functionalities. This is a prerequisite of several vital applications for robots for instance flexible manufacturing, planetary exploration, collaboration with human experts and elder care. The challenge of driving in an environment includes problems of movement of the robot in navigating varied obstacles by

pushing and pulling. Even in structured environments, automated driving is difficult due to the complexities of the related state space. Autonomous driving in unstructured environments faces many challenges which do not exist in structured environments. In unstructured environments, object attributes needed for driving cannot be defined as priori. Information concerning objects has to be gained through sensors even though these are normally ambiguous and therefore introduce uncertainty and avail information that is redundant.
However, in practical terms, it is impossible to avail autonomous driving with full priori models in the actual world. However, models that are perfect are not a prerequisite for successful autonomous driving. Robot mobile driving can be guided using existing structures in the world and in most cases those which are easy to perceive. As such, by leveraging this structure, the complexity of autonomous driving in unstructured environments is lowered significantly. Similarly, understanding the intrinsic measures of freedom of objects in an environment is also able to lower the complexity of autonomous driving in unstructured surroundings.

SEMI-SUPERVISED AND SELF-SUPERVISED LEARNING FOR ROBOTICS

Imitation based learning is a promising approach to tackling the difficult robotic assignments, for instance, autonomous navigation. However, it needs human supervision to oversee the process of training and sending correct control commands to robots without feedback. It is this form of procedure that is prone to failure and high cost. Therefore, in order to lower human involvement and limitmanual data labeling of autonomous robotic navigation using imitation learning, the techniques of semi-supervised and selfsupervised learning can be introduced. It should be noted however that these techniques need to operate according to a multi-sensory design approach. The solution should comprise of a suboptimal sensor policy founded on sensor fusion and automatic labeling of states the robot could encounter. This is also aimed at eliminating human supervision in the course of the learning process . Furthermore, a recording policy needs to be developed to provide throttling of the adversarial impact of too much data being learned from the suboptimal sensor policy. As such, this solution will equip the robot with the capability of achieving near-human performance to a large extent in most of its assignments. It is also capable of surpassing human performance in situations of unexpected outcomes for instance hardware failure or human operator error. Furthermore, the semi-supervised method can be considered as a solution to the problem of track classification in congested environments such as a room. This problem entails object classification undergoing segmenting and tracking without using class models.

MULTIMODAL DEEP LEARNING METHODS FOR ROBOTIC VISION AND CONTROL

The environment of a robot can be controlled too many levels. In principle, less constrained environments are more difficult to perceive. In the real-world and its unstructured and dynamic surroundings such as vegetation landscape and terrain, the perception of a mobile robot needs to be capable of navigating this unknown environment by using sensor modalities. More so, even without the introduction of uncertainty, sensors in themselves are

International Journal of Research Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 01 Issue 01
February 2014

ambiguous. For example, a lemon and a soccer ball can look similar from a certain perspective. In addition, a cup could be invisible in case the cupboard is shut and it can be challenging to tell the difference between a remote control and cell phone is they are both facing down. These factors are all contributive to the challenges of perceiving the state of the environment. Furthermore, for example, advances in face recognition normally operate under the assumption concerning the position and orientation of the individual in the image.
Therefore, in order to tackle perception in unstructured surroundings, robots need to be able to lower the state space that requires being analyzed. to provide facilitation of certain perceptual assignments by limiting uncertainty and as such lowering the dimensionality of the state space. For instance, in order to compute the distance of objects in an environment, robots need to relate depth to visual information. This is normally done with the use of a stereo vision system and solving the correspondence challenge between two static 2D images. In addressing the correspondence problem, however, it is complicated due to noise, many likely matches and the uncertainty in calibrating the camera. On the other hand, in a system capable of the capture of at least three view angles in one image, this lowers the state space by reducing a multi-sensor system to a single sensor.

APPLICATION OF DEEP MODELS TO PROBLEMS IN VISION AND ROBOTICS

The preceding overview of machine learning applications in robotics will highlight five major areas where considerable impacts have been made by robotic technologies currently and in the development levels for long-term use. However, by no means inclusive, the aim of this summarization is to provide the reader with a preview of the form of machine learning applications in existence within robotics and motivate the desire for extended research in such and other fields. The growth of big data which is to day visual information provided on the internet with the inclusion of annotated images and video has pushed forward advancements in computer vision which has in turn assisted in extending machine-based learning systems to prediction learning methods such as those presented by research at Carnegie Mellon. This presentation involved unveiling offshoot examples such as the anomaly detection using supervised learning have been applied in building structures with the capability of searching and assessing damages in silicon wafers with the use of convolutional neural networks. In addition, extrasensory technologies for instance lidar, ultrasound, and radar such as those developed by Nvidia as also propelling the creation of 360- degree vision-based systems for autonomous vehicles and UAVs. Imitation learning which is closely associated with observational learning is also a field categorized by reinforcement learning or the difficulties of gaining an agent to act towards maximizing rewards.

BENEFITS AND DRAWBACKS OF DEEP LEARNING TO BE APPLIED IN MOBILE ROBOTS

The gains of deep learning as a component of the wider family of machine learning techniques founded on representations of learning data in opposition to assignmentparticular algorithms through supervised, unsupervised and semi-supervised learning allows for structured on the interpretation of information processing and patterns of communications which can be viewed as trials at defining a relation between multiple stimuli and related neuronal responses. Deep learning architectures for instance deep neural, deep beliefs as well as recurrent neural networks have been utilized in arenas inclusive of computer vision, natural language processing, social network filtering, speech recognition, bioinformatics and audio recognition. In these mentioned fields, deep learning architecture has produced outcomes in comparison to an in certain case more advanced to human expertise. Furthermore, deep learning algorithms utilize a cascade of many nonlinear processing unit layers for extraction of features and transformation with particular layers applying the output from the past layer as input.

The drawbacks of deep learning in applied robotics is that the storage of the agent data using replay memory does not allow for re-batching or sampling at randomly from varied time-stages. As such, memory aggregation in this approach lowers non-stationary and eroded updates while in simultaneously limiting the techniques to off-policy reinforcement learning algorithms.

CONCLUSION

Deep learning is set to transform the arena of artificial intelligence as well as represent a measure in the direction of developing autonomous systems with an increased scope of perceiving the visual world. Presently, deep learning is allowing scaling of challenges that were traditionally intractable for instance learning to directly play video games for pixels. Furthermore, deep learning algorithms are also utilized in robotics to foster the capability of control functionality for robots indirectly learning from cameral inputs in the actual world. It is on this premise that the survey above illustrates the major advances and approaches of reinforcement learning in regard to the main streams of value and policy-driven methods as well as associated coverage of central algorithms in deep learning for instance deep network, asynchronous advantage actor-critic as well as trust region policy optimization. Furthermore, the research survey has highlighted the gains of deep neural networking with emphasis on visual perception through deep learning. One of the core objectives of the discipline of artificial intelligence it the production of completely autonomous agents that are able to interact with their settings to learn optimal behavior and demonstrate improvements with time by trial and error regiments. It is therefore on this premise that creating artificial intelligence systems that are responsive and with the capability of learning has long been an elusive challenge. However, hope is found in the principled mathematical framework of deep learning with utilizes experience driven autonomous learning to apply a functional approximation to represent learning attributes of deep neural networks to overcome these challenges.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.

[2] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[3] J. S. Esteves, A. Carvalho, and C. Couto, "Generalized geometric triangulation algorithm for mobile robot absolute selflocalization," in *Industrial Electronics, 2003. ISIE'03. 2003 IEEE International Symposium on*, vol. 1. IEEE, 2003, pp. 346–351.

[4] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 2015, pp. 689–692.

[5] A. Eitel, J. T. Springenberg, L. Spinello, M. Riedmiller, and W. Burgard, "Multimodal deep learning for robust rgb-d object recognition," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015, pp. 681–687.

[6] Y. Yang and Y. Li, "Robot learning manipulation action plans by" watching" unconstrained videos from the world wide web." 2015.

[7] D. C. Cires¸an, U. Meier, L. M. Gambardella, and J. Schmidhuber, "Deep, big, simple neural nets for handwritten digit recognition," *Neural computation*, vol. 22, no. 12, pp. 3207–3220, 2010.