Proceedings of CONF-CDS 2026 Symposium: Computer Vision-Based Multimodal Cognitive Load Estimation for Adaptive Media Communication
DOI: 10.54254/2755-2721/2026.GL34284

Hardware Acceleration Technologies for Deep Learning and Reinforcement Learning of Mobile Robots

Jiayi Yan

School of CHIPS, Xi'an Jiaotong-Liverpool University, Suzhou, China
Jiayi.Yan24@student.xjtlu.edu.cn

Abstract. Micro-robots need to be in the real world and have a very small power supply; generally, the processors the paper currently has cannot meet the high calculation requirements for perception and decision-making in the algorithm. Therefore, in this paper, the paper will introduce the basic acceleration techniques for deep learning perception and reinforcement learning decision-making, as well as related contents from FPGA and ASIC architectures, such as quantization, pipelines, memory hierarchy optimisation, etc. According to the comparison of dedicated architectures and their hardware-software co-design, the power consumption per frame can be reduced to a few milliwatts and the throughput is more than 80 fps. The paper has also set the requirements for the precision of the fixed-point representation and the convergence speed of the algorithm to provide a Design Basis that is both accurate and economical in this paper. As mentioned in the paper, to solve the bottleneck of the "memory wall", in-memory computing and neuromorphic computing have been put forward as new directions to realise high power efficiency. The above technologies have formed a good hardware basis for the next generation of intelligent robots to achieve a high level of self-operation and extend the service life of robots.

Keywords: Mobile Robot, Reinforcement Learning, Hardware Acceleration Technology

1. Introduction

With the rapid development of intelligent manufacturing and service robots, mobile robots have been widely used in key areas such as logistics warehousing, disaster relief, and deep space exploration, becoming an important vehicle for improving social productivity and safety. Due to the complexity and frequent changes in the environment, at present, one of the main research directions for improving the ability of robots to navigate and make decisions independently in dynamic and unknown situations is in the field of robotics. Although deep learning and reinforcement learning algorithms can better solve the problem of high-dimensional perception data and complex strategies, they have large-scale exploration processes, many matrix multiplication and addition operations, and relatively high demands on memory bandwidth; thus, they do not meet the strict requirements for light design, miniaturisation and low power consumption of mobile robots. Therefore, the Design and Development of special acceleration chips for computationally intensive algorithms have been chosen as one of the main directions that research institutions and enterprises are pursuing to overcome the computing power bottleneck of mobile platforms [1].

DOI: 10.54254/2755-2721/2026.GL34284

Research is being carried out to solve the problem of computational bottleneck in the hardware acceleration of deep learning perception algorithms by means of network lightweighting and architecture optimisation. Li and others have developed an accelerator for feature extraction in visual localization and mapping based on convolutional neural networks [2, 3], etc. In order to improve the power efficiency and processing speed of the system, a parallel and reconfigurable CNN engine has been added, and in conjunction with multi-level storage organisation, the external memory access in this paper has been reduced to a certain extent. Given that neural networks are very expensive in terms of computation, Yin Zehua and others have been raising the issue of resource requirements for them and put forward an end-to-end fully pipelined SuperPoint feature extraction hardware accelerator [3]. At the bottom, in order to reduce the number of filters, the importance of convolutional layers at different depths in the network is taken into account for network lightweighting; thus, hardware resources are saved significantly and accuracy is maintained.

In the deployment of reinforcement learning decision-making algorithms, research focus has gradually shifted from pure algorithm optimization to "algorithm-hardware co-design." Zhuo Cheng et al.proposed a multi-core chip dynamic power management framework based on reinforcement learning [4], which directly performs intelligent scheduling at the chip level by introducing a deep Q-neural network, demonstrating the great potential of reinforcement learning in hardware-level power optimization. Meanwhile, Li Xiaoyu et al.studied the accelerated deployment strategy of deep reinforcement learning algorithms in map-free environments for robot chassis energy management and path planning [5], further promoting the application of reinforcement learning algorithms in low-latency decision-making on edge-constrained hardware.

In summary, current research on hardware acceleration for deep learning and reinforcement learning in mobile robots is rapidly evolving towards higher frame rates, lower power consumption, and greater hardware-software synergy.

Therefore, a comprehensive review and overview of dedicated hardware acceleration technologies for deep learning perception algorithms and reinforcement learning decision-making algorithms for mobile robots constitutes the core objective of the following discussion. In terms of content organization, the basic mechanisms of deep learning and reinforcement learning, as well as the performance differences of mainstream hardware acceleration platforms, will be clarified first. Second, the paper will introduce the main deployment architecture and typical application scenarios of convolutional neural network feature extraction at the perception end and reinforcement learning path planning at the decision end on dedicated chips in detail. Finally, briefly mention that the current deficiencies in fixed-point precision and hardware adaptability will be dealt with in the future, and then introduce the future development direction of energy-saving architectures, such as in-memory computing.

2. Basic concepts and hardware platform

2.1. Core algorithm definition

Deep Learning and Reinforcement Learning are often used together in the autonomous navigation of mobile robots for different reasons. Deep learning is employed in the perception stage of robots to process a large amount of unstructured data from sensors such as cameras and LiDAR, extract features, identify objects, perform semantic segmentation, and acquire an understanding of the surrounding environment; therefore, it is known as the "eyes" of the robot.

DOI: 10.54254/2755-2721/2026.GL34284

Reinforcement Learning is about decision-making and planning; it is the "brain" of the robot that learns how to choose actions based on environment state information provided by the perception layer and, through repeated trial and error, also learns via a reward function. Reinforcement learning is employed to solve the problems of path planning, dynamic obstacle avoidance and multi-machine cooperation in autonomous driving, and to find an optimal path in the whole world. The Perception layer is for collecting information on "where I am and what is around me", and the Decision-making layer answers the question of "where do I want to go and how do I get there?".

2.2. Hardware platform comparison

The speed and efficiency of the actual work of a mobile robot in operation are determined by the power of the computing platform. The running speed of some hardware platforms for the ORB-SLAM and neural network combined algorithm is not good at present.

ARM-based embedded CPUs are a kind of general-purpose computer; although versatile, they do not have strong processing power for complex vision algorithms, are usually relatively slow at about 7 frames per second, have a relatively high power consumption of about 257mJ per frame, and are therefore not suitable for the demands of real-time navigation. Some CPUs and GPUs in the NVIDIA Jetson series of embedded SoCs have been added to increase the processing speed of deep learning models to about 22 frames per second; However, the power consumption is about 10W, and thus the energy consumption per frame of about 455mJ may be too high a power draw for energy-constrained micro-robots. FPGAs can be customised to have a parallel pipeline structure that can achieve a processing speed of about 62 frames per second, and they have relatively low power consumption and a single-frame energy consumption of 4.6W and 74mJ, respectively; thus, they are very energy-efficient. At the same time, ASICs can achieve good energy efficiency for some algorithms; they have a processing speed of more than 80 fps and a power consumption of only a few milliwatts (e.g., 24mW - 243.6mW), and the energy consumption per single frame is as low as 0.1mW [6], so mJ is regarded as an excellent hardware solution to achieve the autonomy of nano and micro robots.

2.3. Acceleration principle

Hardware acceleration is to solve the instruction execution bottleneck of general-purpose processors by means of custom circuits. First of all, there will be optimisation in parallel computing and in the quantification and storage of data.

The first reason for the good performance of dedicated accelerators is that they have been optimised to some extent for the characteristics of algorithms and are closer to the hardware architecture; first of all, in terms of computing architecture, parallel computing and pipelined design can achieve spatial parallelism by using a large number of processing units to perform thousands of multiply-add operations simultaneously. All the pipelines can carry out the process of data acquisition, computation and output at the same time to increase the overall speed of the system and reduce processing lag. Second, for the representation of data, quantization and fixed-point processing are used to convert floating-point numbers into small-bit fixed-point numbers. Together with the pruning and low-bit quantization mentioned above, both the storage space and circuit area will be reduced in line with the requirements of power-limited mobile platforms, and the perception accuracy will still be good. Finally, in order to improve the efficiency of memory access, a hierarchical structure is employed to make better use of the on-chip cache and storage hierarchy, reduce the distance data needs to travel from the computing unit by frequently reusing data, etc.

DOI: 10.54254/2755-2721/2026.GL34284

Therefore, there will be little power consumption and a long-latency connection to the external storage. To make full use of all the internal registers, a fixed-output data-flow mode can be used; that is to say.

3. Analysis of deep learning acceleration technologies at the perception end

3.1. Deep learning perception acceleration strategy

For feature extraction algorithms (such as SuperPoint) used in mobile robot visual perception, the core challenge lies in the fact that the generation of high-dimensional descriptors involves pixel-by-pixel filtering and gradient calculation, which is computationally intensive and has irregular access patterns. Due to the lack of resources at the edge, all sorts of optimisations have been carried out to reduce the scope of the algorithm and the total hardware resources.

Filter pruning and lightweighting can be employed to determine how many convolutional layers at different depths are needed for various purposes, and then the number of network parameters can be reduced without affecting the accuracy of feature point detection. An entire pipeline is employed to combine the convolution operation and the post-processing module, such as non-maximum suppression and normalised exponential function, in order to reduce the amount of data that needs to be transmitted in the intermediate stage. Redundancy elimination is used to make full use of the sparsity of feature points; according to the selection logic of the data, descriptor calculation is skipped for non-feature points, and in order to reduce the computational burden of the post-processing module, approximation calculation methods are also employed.Corresponding FPGA-based acceleration has also been reported for ORB feature extraction in visual SLAM [7].

3.2. Analysis of the accelerator chip case

The reason for the slow processing speed of visual semantics and localisation in real time is that they cannot be accelerated effectively by memory. The visual CNN-SLAM processor designed by Li Z et al. in 2019 provides a typical solution to this problem [2].

The design firstly optimizes on-chip storage organization by constructing a multi-level storage structure and employing a refined data scheduling strategy. It will not need to access external memory often in the high-computation stage of bundle adjustment, so there will be no latency or high power consumption due to a bandwidth bottleneck. Second, a parallel reconfigurable engine will be employed in the hardware to extract features, and the number of features that need to be matched has been reduced by combining temporal pose prediction and address hashing. According to the results of the above experiment, the paper will reduce the number of redundant matching operations by 97%. Finally, after applying the optimisation method for fixed-point arithmetic and function reorganisation, the chip can run at 80fps with a resolution of 640x480, and it has very low power consumption of only 243.6mW; it has achieved an energy efficiency ratio of 3.6TOPS/W.Similar efficiency optimization strategies have also been explored for LSTM-based accelerators [8].

4. Reinforcement learning as a technology to accelerate decision-making

4.1. Design of RL inference acceleration engine

The application of reinforcement learning in mobile robot decision-making faces a significant computational bottleneck due to the high-frequency interaction between the agent and the

DOI: 10.54254/2755-2721/2026.GL34284

environment. Some research has begun to investigate how to combine Q-learning with strategy gradient methods, such as PPO, DDPG, etc. The inference engine will be carried out on the hardware accelerator.

On the one hand, it can compute the reward function and update the state of the environment very quickly at the hardware level; therefore, there is no need to consider the additional overhead of general-purpose processors for instruction decoding and frequent memory access. At the same time, the custom inference engine can also shorten the cycle delay of agent-environment interaction. Given that computers are relatively slow, the delay in information transmission must be small; otherwise, there will be a large error during movement in an abnormal situation.

4.2. An acceleration architecture for reinforcement learning

To speed up the response of robots in abnormal conditions, the acceleration architecture should be relatively light on hardware resources but still fast enough for the convergence of algorithms.

To reduce the size of the circuit and power consumption, a low-bit-width method can be used in the architecture design; one such way is fixed-point conversion. Although there will be some loss in accuracy and it may affect the convergence speed of reinforcement learning algorithms or result in decision bias, by carrying out a detailed floating-point to fixed-point loss analysis at the model development stage, the paper can maximise the use of hardware and still maintain good performance of the algorithm. To reduce the power consumption of the system and extend the operating life of the hardware more significantly, the paper can improve the acceleration architecture for robot path planning in real applications and maintain a good success rate of decision-making. This robot will not need a high-power computer to run for a long time in the energy-constrained edge environment, and it will still be able to work normally.

5. Outlook

5.1. Limitations analysis

Although there have been some improvements in energy efficiency, the technology of dedicated hardware acceleration is still lacking for practical application, and it is low in accuracy and does not have algorithm flexibility.

Due to the problem of imprecision, many accelerator designs are in the form of fixed-point and use low-bit quantization to save hardware resources and power. The above method will introduce quantization noise and thus reduce the accuracy of deep learning feature extraction; at the same time, it may also affect the stability of the convergence of the reinforcement learning algorithm. In the high-speed environment or when there are large changes in light, even a small computation error can accumulate and cause the robot to lose track of its position or make the wrong selection.

Flexibility Contradiction: Fixed Structure of ASIC vs. Algorithm Iteration; Although ASIC has low power consumption, it is difficult to change after the chip has been fabricated. In recent years, many deep learning and reinforcement learning algorithms have been developed at high speeds, and some of them are relatively new, such as CNNs and Transformers. There is a contradiction between the fixation of hardware and the flexibility of algorithms; thus, ASICs cannot be updated quickly, and although FPGAs are reprogrammable, they have lower power efficiency and larger area than ASICs [9].

DOI: 10.54254/2755-2721/2026.GL34284

5.2. Future development trends

In order to solve the problems of bottlenecks and power consumption in the current von Neumann architecture, hardware acceleration for future mobile robots will be based on an energy-saving computing model.

In the past, there were many data transfers between the storage and processing units of a computing system, and as a result, more power was consumed due to the "memory wall" problem. Technology for in-memory computation can be employed to bring computing logic closer to the storage medium for "data-as-computation", reduce energy consumption by not having to move data, and thus provide a new path to meet the high real-time requirements of mobile robot perception.

According to the pulse transmission method of the biological nervous system, neuromorphic chips (e.g., neuromorphic processors) process information by means of asynchronous spiking neural networks. There is a small amount of static power consumption for the processing of sparse data in the architecture. In recent years, neuromorphic computing-based autonomous navigation technology has developed, and NeuroSLAM is one such technology [10]; it can realise robot self-navigation and environmental exploration with the same performance as living organisms and has a very low power consumption.

6. Conclusion

Recently, in the field of research on hardware acceleration for deep learning and reinforcement learning, it has been determined that mobile robots are gradually departing from general-purpose computers and are now being employed on special architectures. Some optimisation measures and modifications that need to be made for the deployment of complex perception and decision-making algorithms on resource-constrained devices will be introduced in this paper.

Based on the above results, it is concluded that hardware-friendly algorithmic pruning and special data-path design can solve the problems of power wall and memory wall in mobile autonomy. CNN-SLAM processors are good examples of specialised accelerators that can operate in real time and have low power consumption compared to general-purpose GPUs. However, there are still problems of loss in accuracy after fixed-point conversion and inflexibility in ASIC design; thus, they have not been widely used.

As shown above, some achievements have been made; at present, a new type of nano-drones and microrobots can operate independently of cloud computing. Looking ahead, the combination of neuromorphic computing and in-memory computing architecture based on non-volatile memory has begun to show some results in reducing the energy consumption of biological systems and developing artificial intelligence. In the future, research should be conducted on reconfigurable accelerators that have good energy efficiency and can change according to the different demands of AI models at high speeds.

References

[1] Liu, B.Q., Shen, Z.X., Wang, J.P., et al. (2024) Overview of Simultaneous Localization and Mapping Dedicated Chips Research in Intelligent Robotics. Integrated Circuit and Embedded System, 24(11), 1-14.
[2] Li, Z., Chen, Y., Gong, L., et al. (2019) An 879GOPS 243mW 80fps VGA fully visual CNN-SLAM processor for wide range autonomous exploration. In: *2019 IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE, 134-136.
[3] Yin, Z.H. (2024) Design of Hardware Accelerator for Image Feature Extraction in Visual SLAM. Huazhong University of Science and Technology.

DOI: 10.54254/2755-2721/2026.GL34284

[4] Zhuo, C., Zeng, X.D., Chen, Y.F., et al. (2023) Multi-core Chip Dynamic Power Management Framework Based on Reinforcement Learning. Journal of Electronics & Information Technology, 45(01), 24-32.
[5] Li, X.Y., Zhang, J.H., Guo, X.G., et al. (2024) Reinforcement learning-based optimization algorithm for energy management and path planning of robot chassis. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 40(21), 175-183.
[6] Suleiman, A., Zhang, Z., Carlone, L., et al. (2019) Navion: A 2-mw fully integrated real-time visual inertial odometry accelerator for autonomous navigation of nano drones. IEEE Journal of Solid-State Circuits, 54(4), 1106-1119.
[7] Zhou, J.L. (2024) Hardware-accelerated Design for vSLAM Feature Extraction. University of Electronic Science and Technology of China.
[8] Chen, K., He, B., Teng, Z.H., et al. (2025) Design and implementation of high efficiency LSTM hardware accelerator. Electronics & Packaging, 25(9), 090302.
[9]Pu, S.H., Qing, X.G., Jiang, Y.Q., et al. (2025) V&V Feasibility Study of Hardware Design Based on Deep Learning. Process Automation Instrumentation, 46(09), 112-116+121.
[10] Yu, F., Shang, J., Hu, Y., et al. (2019) NeuroSLAM: a brain-inspired SLAM system for 3D environments. Biological Cybernetics, 113(5), 515-545.