

Characteristic Analysis and Application of the Optoelectronic Neural Network

Yuhan Chen*

Department of Communication Engineering, Fudan University, Shanghai, China

*Corresponding author: 20307130004@fudan.edu.cn

Abstract. In recent years, optoelectronic neural networks have garnered significant attention due to their potential to facilitate high-speed and efficient information processing by employing light-based signals to transmit and process data. This review provides an overview of the current state-of-the-art in optoelectronic neural networks, including their design principles, fabrication techniques, and applications. The article also presents five different methods for constructing optoelectronic neural networks, which offer insights into current ONN research and solutions to overcome the limitations of traditional neural networks. Furthermore, the review discusses three different applications of ONNs, including basic tasks such as data classification, speech recognition, and image recognition, as well as hardware accelerators and SNN algorithms for object detection. The promising potential of optoelectronic neural networks in transforming various fields, such as artificial intelligence, image recognition, and data processing, is also highlighted. As research in this area continues to advance, further breakthroughs in optoelectronic neural networks are anticipated.

Keywords: ONN, Optical Computing, Optical devices, Neural Network.

1. Introduction

Contemporary society has experienced a remarkable expansion of data, resulting in the indispensability of Artificial Intelligence (AI) in numerous application domains. Neural networks, as a critical component of AI, have achieved tremendous success in areas including speech recognition, image processing and natural language processing. A traditional electronic neural network, also known as an Artificial Neural Network (ANN), is a machine learning model that is inspired by the structure and function of the human brain. It consists of a large amount of interconnected processing nodes, or "neurons," which work together to process information and learn from data. However, traditional electronic neural networks are confronted with certain impediments when processing large-scale and high-speed data, requiring new technologies to improve the speed and efficiency of neural network.

An emerging method for neural network implementation is the application of optoelectronic technology, which exploits the distinctive properties of light to execute both computation and communication functions. Optoelectronic neural networks use light to carry information and perform computations, which offers several advantages over traditional electronic neural networks. In addition, optoelectronic neural networks have the potential to overcome the physical limitations of conventional electronic neural networks, such as bandwidth and latency.

The main advantages of optoelectronic neural networks compared to traditional electronic neural networks can be summarized as follows: 1) High-speed computation: Optoelectronic neural networks can perform parallel processing and data transmission at a high-speed using light signals, which enables faster and more efficient computation compared to traditional electronic neural networks. 2) Low power consumption: Optoelectronic neural networks consume less power than traditional electronic neural networks, which is beneficial for energy efficiency and scalability. 3) High precision: Optoelectronic neural networks can achieve high precision and accuracy due to the low noise and high linearity of optical components, which is essential for applications in fields such as medical diagnosis and autonomous driving. 4) Reconfigurability and scalability: Optoelectronic neural networks can be reconfigured and scaled to handle different types and amounts of data, making them more flexible and adaptable than traditional electronic neural networks. 5) Robustness to noise and

interference: Optoelectronic neural networks are less susceptible to noise and interference compared to electronic neural networks, making them more reliable for use in noisy environments.

In recent times, considerable attention has been devoted to the study of optoelectronic neural networks, resulting in substantial advancements in the development of fundamental components and their potential for diverse applications. For example, researchers have demonstrated the use of optical neurons and synapses to perform computations with light, as well as the integration of optoelectronic devices with traditional electronic circuits to create hybrid systems. These advantages make optoelectronic neural networks promising for applications in high-speed computing, large-scale data processing, and high-precision control. This review paper aims to provide a comprehensive introduction to optoelectronic neural networks by exploring their fundamental principles, device structures, and application domains.

In this review paper, an overview of optoelectronic neural networks, including the basic principles, building blocks, and applications was provided. The review commences by introducing the concept of optoelectronic neural networks and compare them with traditional electronic neural networks. This paper then describes the key building blocks of optoelectronic neural networks, including optical neurons, synapses, and interconnects. The challenges and opportunities in developing optoelectronic neural networks, such as the need for novel materials and devices, as well as the integration with existing electronic systems was also discussed. The subsequent section of the review highlights some of the recent progress in optoelectronic neural networks, including their application in various domains such as image recognition, speech processing, and pattern recognition. Furthermore, the review also discusses the potential impact of optoelectronic neural networks on future computing systems and the development of novel artificial intelligence applications.

2. Methods

2.1. Overview of the models

Due to the slowing down of Moore's Law in electronic computing the approaching physical limit of electronic transistors [1], it is difficult to further improve the computational speed of electronic neural networks. Optical computing has the advantages of large bandwidth, low power consumption, and low crosstalk, which is expected to greatly improve the computing speed. Therefore, optoelectronic neural networks have emerged as the times require. There are two main ways to construct it: one is to use a Mach Zehnder interferometer to construct an optoelectronic neural network [2]. The optoelectronic neural network constructed in this way has advantages such as high integration and strong reconfigurability, but requires photoelectric conversion, and is limited by the level of production technology, which is currently not suitable for large-scale neural networks; The other is a diffractive Optoelectronic Neural Network (DONN) based on a phase mask, which does not require photoelectric conversion processing. It constructs a large-scale neural network by directly inputting optical signals, thereby achieving more complex functions. Lin et al. proposed a photon depth neural network structure based on diffraction to achieve all-optical machine learning [3], and completed the classification of MNIST handwritten digital dataset and Fashion-MNIST dataset, where the size of each phase mask is $8\text{ cm} \times 8\text{ cm}$, with a layer spacing of 3 cm , and a mask made using 3D technology; Yan et al proposed a Fourier spatial diffraction depth neural network that performs saliency detection and high-precision target classification tasks at the speed of light [4]. By combining a dual $2f$ (f is the focal length) optical system to maintain spatial correspondence, it is helpful to complete the task of image-to-image mapping. However, the actual structure of these phase masks is too large for integration, and each pixel cannot be changed once training is completed, making them unsuitable for other datasets. Therefore, this optical computing architecture is not suitable for constructing neural networks with flexible network weights that are urgently needed in various fields.

2.2. Method to adjust dataset

In Zhao's research, the data sets utilized for training are the CelebA data set and the cat & dog data set, as depicted in Fig. 1. Each of the 202,599 face photos in the CelebA dataset, which includes 10,177 celebrity identities, has been annotated with a face annotation box, five persons face feature point coordinates, and forty attribute markers. A lot of face attribute recognition models are trained using this dataset, which is widely used and maintained by the Chinese University of Hong Kong. The cat & dog dataset training set includes 25,000 images, 12,500 images of cats, and 12,500 images of dogs. The processing steps of the network data are as follows. Input image preprocessing. Input size is 224×224 . The RGB image of 224×224 is preprocessed by subtracting the average value of RGB from each pixel in the training set [5].



Fig. 1 Dataset examples. (a) CelebA dataset; (b) cat and dog dataset

2.3. Hybrid photoelectric convolution with optimized diffractive optics

In order to construct a neural network with flexible and adjustable network weights, Chang et al. proposed a hybrid photoelectric convolution neural network with optimized diffractive optics [5]. By adding a layer of optical convolution operations before electronic calculations, that is, using a layer of phase masks, adding convex lenses before and after them, a $4f$ system is constructed to achieve convolution operations [6]. This can reduce energy consumption while maintaining network performance.

2.4. AlexNet based DONN

Colburn et al. proposed an AlexNet based DONN (AlexNet DONN) [7], which utilizes AlexNet convolutional neural network [8], spatial light modulator that only performs phase modulation, and complex amplitude modulation methods to achieve phase and amplitude adjustability, and finally completes the function of cat and dog dataset classification [9]. For AlexNet, the calculation time of the first and second layers of convolution is longer than that of the subsequent layers, accounting for 62.2% of the total time 7%, and the first layer processes image information using light as a carrier without requiring photoelectric conversion. Therefore, using optical computing to replace the first layer of electrical convolution operations is of great significance. Although they all overcome the problem of nonadjustable weights in all optical diffraction neural networks, they do not respond flexibly to situational information existing in the actual environment, making it difficult to meet complex and volatile needs, that is, they lack the ability to context dependent learning. Moreover, the AlexNet network structure has high memory and time consumption, multiple iterations Weak discrimination and multiple training parameters.

2.5. Prefrontal cortex of primates based OWM algorithm

In order to improve the adaptability of artificial intelligence systems to complex and volatile environments, Zeng et al. [10], inspired by the prefrontal cortex of primates, proposed an orthogonal weight modification (OWM) algorithm and a context dependent processing (CDP) module, using the ResNet network structure to extract the features of the CelebA dataset [11, 12]. However, this extraction method is also limited by the fact that Moore's Law is slowing down. Therefore, it is of great significance to use photoelectric mixing to accelerate it.

2.6. VGG 16-based DONN

Zhao et al. [13], proposed a VGG16-based DONN (VGG16-DONN) shown in Fig. 2, which uses optical diffraction layers to replace the first convolution layer in VGG16, which takes the most time to calculate. It not only solves the problem of slow computing speed of the traditional VGG16-based electronic convolutional neural network (VGG16-ECNN), but also solves the problems of excessive memory and time consumption, multiple iterations, weak discrimination, and multiple training parameters in AlexNet-DONN. This method implements the classification of CelebA data sets, with a classification accuracy comparable to that obtained by VGG16-ECNN; When classifying cat and dog datasets using this method, the classification accuracy is higher than that of AlexNet-DONN. Secondly, combining VGG16-DONN with the CDP module, using a hybrid photoelectric method, with the advantage of fast photon computation, can greatly accelerate the extraction process and achieve different outputs for the same input under different scenarios.

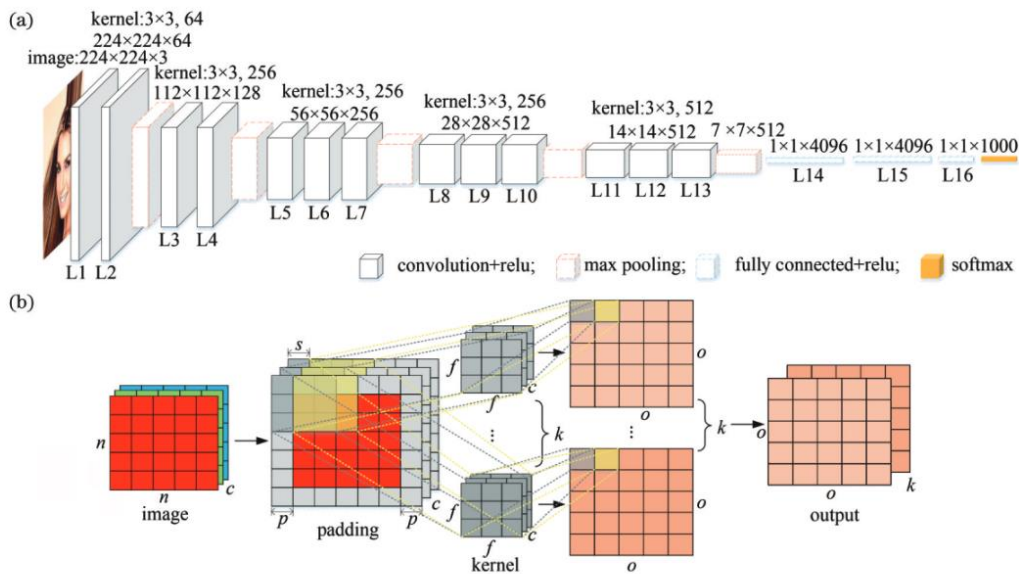


Fig. 2 Schematic diagram of VGG16-ECNN, and convolution process. (a) Schematic diagram of VGG16-ECNN; (b) convolution process

2.7. MRR and GST based Non-volatile Silicon Photonic(NVPP-CNN)

Guo et al. proposed a novel CNN hardware architecture based on MRR and GST (Non-volatile Silicon Photonic, NVSP CNN). Specifically, the accelerator uses an MRR with GST embedded at the top as the basic photon dot product operation unit and combines a balanced photodiode to subtract the optical power of MRR's Drop and Through port output values, simulating the case of negative weights. Compared with the DEAP-CNN structure, using GST for data storage utilizes its nonvolatile nature to achieve simultaneous optical storage and computation, avoiding potential latency issues in the memory computing separation architecture. In addition, this structure does not require external electrodes for modulation, reducing the area and power loss of the external electrodes.

3. Application and discussion

3.1. SNN

Optoelectronic neural networks (ONNs) offer several benefits, including low computational complexity and robustness to noise, making them applicable to various fields of information processing. SNN is suitable for hardware implementation. With the progress of optical device technology and process in the past decade, photonic SNN, which combines optical advantages to overcome electrical bottlenecks, is gradually developing. The article of Ke et al. introduced multiple optoelectronic SNN including Semiconductor Optical Amplifier implement solution [14], Electro Absorption Modulator implement solution, laser implement solution and so on.

3.2. ONN data classification

In Zhang T et al 's research [15], this paper proposes to use two typical neural evolutionary algorithms to determine the hyperparameter of ONN to optimize the weight in the connection. Simultaneously, the optical neural network integrated into the photonic chip will be applied to the classification task dataset of Iris plants, wine recognition dataset, and modulation format recognition. The calculation results demonstrate the accuracy and stability of the training algorithm based on traditional deep well evolution.

3.3. ONN based reinforcement learning mode

In Hu's article [16], he proposes a reinforcement learning model based on photonic neural networks. Compared to the decision tree array model proposed by Fulvio Flamini et al. [17], MZI has a significantly reduced number of uses and a compact overall structure. While fully utilizing the advantages of photonic neural networks such as high speed, low power consumption, and wide bandwidth, it can also successfully avoid the von Neumann bottleneck caused by memory separation in electronic neural networks in reinforcement learning applications.

4. Conclusion

This review provides an overview of the state-of-the-art in optoelectronic neural networks, encompassing their underlying design principles, fabrication techniques, and applications. This analysis reveals the potential of optoelectronic neural networks to transform various fields, including artificial intelligence, image recognition, and data processing. The article also discusses five different methods for constructing optoelectronic neural networks that enable readers to gain insights into current ONN research and solutions to overcome the limitations of traditional neural networks. Moreover, the review introduces three different applications of ONNs, which include basic tasks such as data classification, speech recognition, and image recognition, as well as hardware accelerators and SNN algorithms for object detection. However, despite these advancements, significant challenges must be addressed, including the need for improved integration with existing electronic systems, increased scalability, and the development of reliable and efficient light sources. In conclusion, optoelectronic neural networks hold great promise and are expected to revolutionize the way people process information as research in this field continues to progress.

References

- [1] Jang H, Liu C, Hinton H, et al. Optoelectronic Neural Networks: An Atomically Thin Optoelectronic Machine Vision Processor (Adv. Mater. 36/2020). *Advanced Materials*, 2020, 32(36): 2070275.
- [2] Xie Y. The development and application of photonic computer (in Chinese). *Science and technology and innovation*, 2020 (01): 62-63 + 66. DOI: 10.15913 / j.carol carroll nki kjycx. 2020.01.019.
- [3] Lin X et al. All-optical machine learning using diffractive deep neural networks. *Science*, 2018, 361(6406): 1004-1008.

- [4] Yan T, Wu J M, Zhou T K, et al. Fourier-space diffractive deep neural network. *Physical Review Letters*, 2019, 123(2): 023901.
- [5] Chang J L, Sitzmann V, Dun X, et al. Hybrid optical electronic convolutional neural networks with optimized diffractive optics for image classification. *Scientific Reports*, 2018, 8(1): 12324.
- [6] Ma S, Zeng C, Xu F. Optical image encryption and optical image encryption based on 4f system (in Chinese), Decryption simulation teaching. *University Physics Experiment*, 2018, 31(6): 39-45.
- [7] Colburn S, Chu Y, Shilzerman E, et al. Optical frontend for a convolutional neural network. *Applied Optics*, 2019, 58(12): 3179-3186.
- [8] Sarkar A, Maniruzzaman M, Alahe M A, et al. An Effective and Novel Approach for Brain Tumor Classification Using AlexNet CNN Feature Extractor and Multiple Eminent Machine Learning Classifiers in MRIs. *Journal of Sensors*, 2023, 2023.
- [9] Luo T, Yuan J, Chang J, et al. Resolution and uniformity improvement of parallel confocal microscopy based on microlens arrays and a spatial light modulator. *Optics Express*, 2023, 31(3): 4537-4552.
- [10] Zeng G, et al. Continual learning of context-dependent processing in neural network. *Nature Machine Intelligence*, 2019, 1(8): 364-372.
- [11] Hossain S, Lin X. Efficient Stereo Depth Estimation for Pseudo-LiDAR: A Self-Supervised Approach Based on Multi-Input ResNet Encoder. *Sensors*, 2023, 23(3): 1650.
- [12] Prabhu V U, Yap D A, Wang A, et al. Covering up bias in CelebA-like datasets with Markov blankets: A post-hoc cure for attribute prior avoidance. *arXiv preprint arXiv:1907.12917*, 2019.
- [13] Zhao X, Yang Z, Dai J, Zhang T, Xu K. VGG16 diffraction photon neural network and situation dependent processing (in Chinese). *Acta Optica Sinica*, 2022, 42(19): 151-161.
- [14] Curt Z, Peng C, Hu X, Xiao X. Research history and prospect of photonic pulse neural network (in Chinese). *Optical Communication Research*, 2023(01): 17-31. DOI:10.13756/j.gtxyj.2023.01.002.
- [15] Zhang T, Wang J, Dan Y, et al. Efficient training and design of photonic neural network through neuroevolution. *Optics Express*, 2019, 27(26): 37150-37163.
- [16] Hu F, reinforcement Learning Model Based on photonic neural networks and its application (in Chinese). Beijing university of posts and telecommunications, 2021. DOI: 10.26969 /, dc nki. Gbydu. 2021.002777.
- [17] Flanmini F, Hamann A, Jerbi S, et al. Photonic architecture for reinforcement learning. *New Journal of Physics*, 2020, 22(4): 045002.