Journal of Computer Technology and Applied Mathematics
Vol. 1, No. 1, 2024

Deep Learning Algorithms Based on Computer Vision Technology and Large-Scale Image Data

TIAN, Jingxiao [1*] QI, Yaqian [2] LI, Hanzhe [3] FENG, Yuan [4] WANG, Xiangxiang [5]

[1] San Diego State University, USA
[2] Baruch College, USA
[3] New York University, USA
[4] Duke University, USA
[5] University of Texas at Arlington, USA

***** TIAN, Jingxiao is the corresponding author, E-mail: tianjingxiao07@gmail.com***

Abstract: This paper explores the profound impact of deep learning technology on computer vision and its wide-ranging applications across multiple domains. From pedestrian detection to medical image analysis, deep learning's versatility is harnessed alongside traditional methods, ushering in a new era of efficiency and accuracy. Large-scale image storage solutions, coupled with advanced retrieval and processing techniques, ensure the seamless handling of vast datasets. In the medical realm, imaging technologies have undergone significant transformation, with medical image processing emerging as a crucial component in diagnostics and treatment planning. Leveraging innovative hardware and iterative algorithms, this field continues to evolve, promising enhanced capabilities in non-invasive diagnostics and personalized medicine.

Keywords: Computer Vision, Deep Learning, Image Processing, Medical Imaging.

DOI: https://doi.org/10.5281/zenodo.11072457

1 Introduction

In the past 20 years, with the rapid development of deep learning technology and the widespread popularity of hardware computing devices such as Graphics processing unit (GPU), deep learning technology has been applied to almost every field of computer vision, such as object detection, image segmentation, super-resolution reconstruction and face recognition. It has immeasurable commercial value and broad application prospects in image search, automatic driving, user behavior analysis, text recognition, virtual reality and lidar products [1]. Computer vision based on deep learning technology can also have a profound impact on other subject areas, For example, animation simulation and real-time rendering technology in computer graphics, microscopic image analysis technology in the field of materials, medical image analysis and processing technology, intelligent education for real-time evaluation of teachers and students' classroom performance and examination room behavior, intelligent system for analyzing athletes' competition performance and technical statistics.

Deep learning was introduced into the field of machine learning by Dechter [2] as early as 1986, and Artificial neural network (ANN) was introduced into the field of machine learning by Aizenberg et al. [3] in 2000 [4]. Deep learning methods consist of multiple layers and are used to learn data features with multiple levels of abstraction [4]. In the field of artificial neural networks, deep learning, also known as hierarchical learning [5], is a technique that regulates network activation by accurately assigning scores at different computational stages [4]. Deep learning often uses a variety of abstract structures to learn complex mapping relationships, such as ANN with hidden layer [6] proposed by Professor Bengio of the University of Montreal in 2009. Deep learning techniques can be viewed as representational learning, a branch of machine learning.

2 Related Work

2.1 Computer Vision

Computer Vision, also known as Machine Vision, is a discipline that lets machines learn how to "see" and is an important application field of deep learning technology, which is widely used in security, industrial quality inspection and automatic driving scenarios. Specifically, it is to let the machine to identify the object in the picture or video taken by the camera, detect the location of the object, and track the target object, so as to understand and describe the scene and story in the picture or video, in order to simulate the human brain visual system. Therefore, computer vision is also commonly referred to as machine vision, and the goal is to build artificial systems that can "sense" information from images or videos.

The development of computer vision begins with biological vision. [7-8]For the origin of biological vision, the academic community has not yet formed a conclusion. After decades of development, computer vision technology has been applied in many fields such as traffic (license plate recognition, road violation capture), security (face gate, community monitoring), finance (face payment, automatic ticket recognition of counters), medical (medical image diagnosis), industrial production (automatic detection of product defects), etc. Influence or are changing People's Daily life and industrial production mode.

Figure 1 Schematic diagram of pedestrian detection and NMS combined repeat area

In the future, with the continuous evolution of technology, more products and applications will emerge, creating greater convenience and broader opportunities for our lives.

Computer vision tasks depend on image features (image information), and the quality of image features largely determines the performance of the vision system. Traditional methods usually use SIFT, [9] HOG and other algorithms to extract image features, and then use SVM and other machine learning algorithms to further process these features to solve visual tasks. Pedestrian detection is to determine whether there are pedestrians in the image or video sequence and give accurate positioning. The earliest method is HOG feature extraction +SVM classifier, and the detection process is as follows:

1. Use the sliding window to traverse the entire image to obtain the candidate region

2. Extract HOG features of candidate regions

3. Use SVM classifier to classify the feature map (to determine whether it is human)

4. Duplicate areas appear when sliding Windows are used. Use NMS(non-maximum) to filter the duplicate areas

2.2 Large-scale image storage

1 Database

The first thing to consider when storing large-scale image data is the choice of database. In the traditional relational database, because it is designed to deal with structured data, it is not suitable for unstructured image data. However, NoSQL databases (e.g. MongoDB, CouchDB) [10] are a good choice for storing image data due to their good support for unstructured data. In addition, there are several databases specifically designed to store large-scale image data, such as Google's Bigtable, Facebook's Cassandra, and others.

2 Storage Format

In addition to the choice of database, we also need to consider the storage format of the image data. A common practice is to store image data as binary large objects [11] (BLOBs). This method can store image data directly in the database, but because of the lack of structure, it is often difficult to efficiently query and process. Another approach is to convert the image data into some feature vector (for example, features extracted using a deep learning model) and then store those features. Although this method requires additional processing steps, it can greatly improve the efficiency of retrieval and processing.

3 Distributed Storage

For very large image data sets, stand-alone storage often cannot meet the demand. In this case, we need to use distributed storage systems, such as Hadoop's [12] HDFS and Google's GFS. The distributed storage system can distribute data to multiple machines, which improves the storage capacity and data read and write speed. At the same time, they also provide data redundancy and fault tolerance mechanisms, so that data can still be stored safely in the face of hardware failure.

2.3 Retrieving large-scale image data

1 Feature Extraction

Retrieving large-scale image data requires extracting features from images first. These features can range from basic color, texture, and shape information to advanced features extracted by deep learning models. By extracting features, we can transform image data into a structured form for efficient retrieval.

2 Index Construction

After extracting the features, we need to build an index to speed up the retrieval. A common practice is to use data structures such as kd trees or ball trees to build indexes. These data structures can divide the feature space into multiple regions and then achieve efficient retrieval through fast region lookup. In addition, there are several indexing methods specifically designed for high-dimensional data, such as locally sensitive hashing (LSH) and product quantization [13] (PQ).

In the face of large-scale image data, precise retrieval is often impractical because it requires scanning the entire dataset. To solve this problem, we can use Approximate Nearest Neighbor Search. Approximate retrieval does not guarantee finding the nearest neighbor, but it can greatly reduce the computational effort to achieve efficient retrieval on large data sets. Common approximate retrieval algorithms include locally sensitive hashing (LSH), product quantization (PQ), and embedded vector indexing (IVF).

2.4 Processing large-scale image data

1 Distributed Processing

For large-scale image data, a single machine is often unable to meet the processing needs, so we need to use distributed processing technology. MapReduce is a widely used distributed processing framework that can split large-scale computing tasks into smaller tasks and then run them in parallel on multiple machines. In addition, there are several distributed processing frameworks designed specifically for image processing, such as Apache Hadoop, Spark, and Flink.

2 Parallel computing

In addition to distributed processing, we can also use parallel computing to speed up image processing. On the CPU, we can use multithreading technology to achieve parallel computing. On the GPU, we can use parallel computing frameworks such as [14] CUDA or [15] OpenCL to achieve large-scale parallel processing. Through parallel computation, we can greatly improve the speed of image processing.

3 Deep Learning

Deep learning has become a mainstream method for processing image data. With deep learning, we can build complex models for image classification, object detection, semantic segmentation and other tasks. For large-scale image data, we can use distributed deep learning frameworks such as TensorFlow, PyTorch, and MXNet for efficient training and reasoning.

3 Methodology

The rapid evolution of medical imaging technology over the past century has revolutionized diagnostic practices, offering non-invasive solutions that are now integral to modern healthcare systems. At the forefront of this progress lies the interdisciplinary domain of medical image processing, which has seen remarkable advancements. Today, these systems boast enhanced spatial and intensity resolution, coupled with accelerated acquisition times, resulting in vast quantities of high-fidelity image data. However, the interpretation of this data remains a critical challenge, necessitating sophisticated processing techniques for accurate diagnosis.

This paper delves into the pivotal realm of medical image processing within the framework of computer vision technology. It navigates through various imaging modalities, contextualizing their significance within the broader landscape of healthcare diagnostics. Furthermore, the paper elucidates the prevalent challenges and emerging trends shaping this field, providing valuable insights for researchers and practitioners alike.

3.1 Medical image processing

When it comes to building the field of medical image processing, there are a number of concepts and approaches that focus on different aspects at the core of the field. In simple terms, these aspects can be divided into three main processes: image formation, image calculation, and image management. First, image formation refers to the process of how medical image data is obtained. This involves using different types of medical imaging devices (such as X-rays, CT scans, MRI, etc.) to capture images of the inside of the human body. This process involves the process of converting physical information of the body into digital images, usually including the steps of data acquisition, pre-processing and digitization. [16] Secondly, the computational phase of the image refers to how to process and analyze the medical image data that has been obtained to extract useful information from it. This may involve various image processing techniques such as image reconstruction, filtering, segmentation, feature extraction, etc., so that doctors or researchers can better understand the image and make an accurate diagnosis or analysis.

Finally, image management refers to how to efficiently store, retrieve and transmit large amounts of medical image data. This includes the use of specific data storage systems and standards such as [17] PACS and DICOM to manage the storage and access of image data, as well as the use of compression and streaming technologies to efficiently handle the transmission and sharing of image data. As shown in Figure 2, the field of medical image processing covers three key processes of image formation, calculation and management, which together constitute the core content of medical image processing.

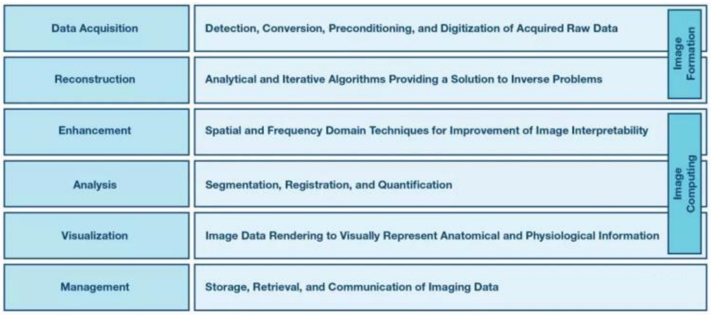

Figure 2 Structural classification of subject types in medical image processing

The process of image formation consists of two phases: data collection and image reconstruction. These two steps are used to solve the inverse math problem. [18] The purpose of image computation is to improve the interpretation of reconstruction images and extract relevant clinical information. Finally, image management involves compressing, saving, retrieving, and transferring captured images and derived information.

3.2 Image data acquisition

Essentially, when it comes to medical imaging, each technique focuses on capturing specific physical properties of the body's internal organs. These properties serve as the foundation for all subsequent steps in image processing.

Different imaging methods rely on distinct physical principles to gather this information. For instance:

- In digital radiology (DR) [19] or computed tomography (CT), the focus is on the energy carried by incident photons.

- Positron emission tomography [20] (PET) measures both the energy and detection time of photons.

- Magnetic resonance imaging (MRI) hones in on parameters within the radio frequency signals emitted by excited atoms.

- Ultrasound, on the other hand, centers on echo parameters.

So, while the methods may vary, they all aim to extract essential physical data about the body's internal structures for further analysis and diagnosis.

However, regardless of the type of imaging mode, the data acquisition process can be divided into the detection of the physical quantity, the conversion of the physical quantity into an electrical signal, the pretreatment of the obtained signal, and the digitization of the physical quantity. A general illustration (see Figure 3) shows how all of these steps apply to most medical imaging modes.

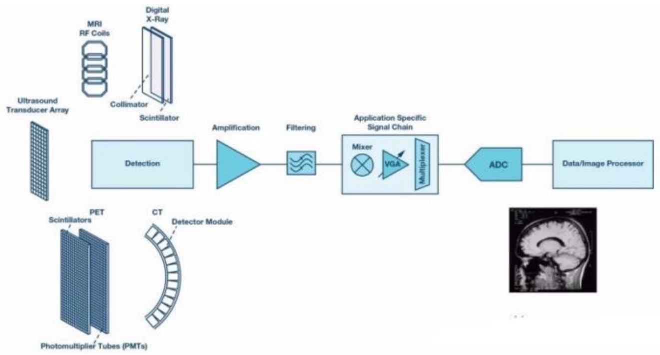

Figure 3 Medical data acquisition architecture

3.3 Image management

In the final stage of medical image processing, managing the acquired information is crucial, encompassing storage, retrieval, and transmission using various technologies. Standards and techniques have been established to address these aspects, such as the Medical Imaging Technology Image Archiving and Transmission System (PACS) for storage and access, and the Medical Digital Imaging and Communication (DICOM) standard for image storage and transmission. Additionally, specialized techniques like image compression and streaming aid in efficient management.

Despite its conservative nature, the field of medical imaging continuously evolves, driven by the complexity and challenges it faces across various scientific disciplines. Trends in medical image processing highlight innovations in image acquisition hardware, leading to improved data quality and enriched information content. This includes advancements in integrated front-end solutions, enabling faster scan times and higher resolutions, as well as the development of combined imaging systems like ultrasound/mammography or CT/PET.

Another significant trend involves the adoption of fast and efficient iterative algorithms for image reconstruction, which enhance image quality in PET, reduce X-ray dose in [24[CT scans, and facilitate compression detection in MRI. Moreover, there's a shift towards data-driven signal models to address inversion problems based on limited or noisy data, replacing manually defined models for better solutions. Key research areas in image reconstruction encompass system physical modeling, signal models, optimization algorithms, and image quality assessment methods.

3.4 Results of large-scale image processing

Analog Devices offers a range of advanced solutions to meet the high demands of medical imaging. These solutions cover a variety of requirements in data acquisition electronics design, such as dynamic range, resolution, accuracy, linearity, and noise reduction. Here are a few examples:

1. ADAS1256: This is a highly integrated analog front end with 256 channels designed for digital radiological imaging (DR) applications. It ensures optimal performance when capturing medical images.

2. ADAS1135 and ADAS1134: These multi-channel data acquisition systems have excellent linearity and are ideal for enhanced computed tomography (CT) [26] applications to improve image quality.

3. AD9228, AD9637, AD9219 and AD9212: These multi-channel ADCs are optimized for superior dynamic performance and low power consumption, making them ideal for meeting the requirements of positron emission tomography (PET) [27].

4. AD9656: This is a pipeline ADC with excellent dynamic performance and low power consumption, ideal for magnetic resonance imaging (MRI) applications.

5. AD9671[28]: Designed for medical ultrasound applications, this integrated receiver front end is an excellent choice with low cost, low power consumption and a small package.

To sum up, medical image processing is a highly complex field involving multiple disciplines, including mathematics, computer science, physics, and medicine. This paper aims to present a simplified but well-structured presentation of key themes, trends, and challenges in the field. In this context, the data acquisition process is particularly important, as it defines the initial quality of the raw data used in subsequent stages of the medical image processing framework.

4 Conclusion

Computer vision and large-scale image storage and processing have significantly impacted various fields, including medical imaging. Computer vision, with its roots in biological vision, has evolved using deep learning technologies and traditional methods like HOG and SVM for tasks like pedestrian detection. Large-scale image storage employs databases like NoSQL [29] and distributed storage systems like HDFS for efficient management. Retrieving large-scale image data involves feature extraction and index construction, utilizing techniques like LSH and PQ. Processing large-scale image data requires distributed and parallel computing, often employing frameworks like MapReduce and deep learning models.

In the medical field, advancements in imaging technologies have revolutionized diagnostics. Medical image processing involves image formation, computation, and management. Image acquisition, utilizing various modes like MRI and PET, requires precise data detection and conversion. Image management involves standards like DICOM and techniques for storage and transmission. Despite being a conservative field, medical imaging continuously evolves, benefiting from hardware innovations and iterative algorithms for image reconstruction. Analog Devices offers solutions for high-quality data acquisition in medical imaging, emphasizing the complexity and interdisciplinary nature of the field.

In conclusion, the advancements in computer vision, large-scale image storage, and processing have significantly impacted various sectors, including medical imaging, leading to improved diagnostics and treatments.

Acknowledgments

We sincerely thank the authors of this article Cui, Z., Lin, L., Zong, Y., Chen, Y., & Wang, S. For your hard work and outstanding contributions. Their research results provide valuable exploration and inspiration for the progress of gene editing technology. We would also like to thank them for their breakthrough work on the application of deep learning technology in gene editing, which provides an important reference for researchers in related fields. Finally, we would also like to thank them for their relentless pursuit of scientific research and their concern for human health, which will help drive progress and innovation in the biomedical field. The deep learning techniques they studied were applied in the field of gene editing to improve the accuracy of CRISPR-Cas9 editors. By using deep learning technology, the research team has successfully developed an efficient gene editing method, providing new ideas and methods for the further development and application of gene editing technology.

Funding

Not applicable.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author Contributions

Not applicable.

About the Authors

TIAN, Jingxiao

Electrical and Computer Engineering, San Diego State University, CA, USA.

QI, Yaqian

Quantitative Methods and Modeling, Baruch College, NY, USA.

LI, Hanzhe

Computer Engineering, New York University, NY, USA.

FENG, Yuan

Interdisciplinary Data Science , Duke University , North

Carolina, USA.

WANG, Xiangxiang

Computer Science, University of Texas at Arlington, Arlington, TX, USA.

References

[1] Choudhury, M., Li, G., Li, J., Zhao, K., Dong, M., & Harfoush, K. (2021, September). Power Efficiency in Communication Networks with Power-Proportional Devices. In 2021 IEEE Symposium on Computers and Communications (ISCC) (pp. 1-6). IEEE.

[2] Akbar, A., Peoples, N., Xie, H., Sergot, P., Hussein, H., Peacock IV, W. F., & Rafique, Z. . (2022). Thrombolytic Administration for Acute Ischemic Stroke: What Processes can be Optimized?. McGill Journal of Medicine, 20(2).

[3] Xu, J., Wu, B., Huang, J., Gong, Y., Zhang, Y., & Liu, B. (2024). Practical Applications of Advanced Cloud Services and Generative AI Systems in Medical Image Analysis. arXiv preprint arXiv:2403.17549.

[4] Yu, L., Liu, B., Lin, Q., Zhao, X., & Che, C. (2024). Similarity Matching for Patent Documents Using Ensemble BERT-Related Model and Novel Text Processing Method. Journal of Advances in Information Technology, 15(3).

[5] Zhang, Y., Liu, B., Gong, Y., Huang, J., Xu, J., & Wan, W. (2024). Application of Machine Learning Optimization in Cloud Computing Resource Scheduling and Management. arXiv preprint arXiv:2402.17216.

[6] Gong, Y., Huang, J., Liu, B., Xu, J., Wu, B., & Zhang, Y. (2024). Dynamic Resource Allocation for Virtual Machine Migration Optimization using Machine Learning. arXiv preprint arXiv:2403.13619.

[7] Liu, B. (2023). Based on intelligent advertising recommendation and abnormal advertising monitoring system in the field of machine learning. International Journal of Computer Science and Information Technology, 1(1), 17-23.

[8] Che, C., Zheng, H., Huang, Z., Jiang, W., & Liu, B. (2024). Intelligent Robotic Control System Based on Computer Vision Technology. arXiv preprint arXiv:2404.01116.

[9] Zheng, H., Xu, K., Zhou, H., Wang, Y., & Su, G. (2024). Medication Recommendation System Based on Natural Language Processing for Patient Emotion Analysis. Academic Journal of Science and Technology, 10(1), 62-68.

[10] Ding, Weike, et al. "Automated Compatibility Testing Method for Distributed Software Systems in Cloud Computing." (2024).

[11] Biology-based AI Predicts T-cell Receptor Antigen Binding Specificity. (2024). Academic Journal of Science and Technology, 10(1), 23-27. https://doi.org/10.54097/wy28c490

[12] Wu, Binbin, et al. "Enterprise Digital Intelligent Remote Control System Based on Industrial Internet of Things." (2024).

[13] Zhou, Tong, et al. "Image retrieve for dolphins and whales based on EfficientNet network." Sixth International Conference on Computer Information Science and Application Technology (CISAT 2023). Vol. 12800. SPIE, 2023.

[14] Sun, Y. (2024). TransTARec: Time-Adaptive Translating Embedding Model for Next POI Recommendation. arXiv preprint arXiv:2404.07096.

[15] Pan, Y., Wu, B., Zheng, H., Zong, Y., & Wang, C. (2024, March). THE APPLICATION OF SOCIAL MEDIA SENTIMENT ANALYSIS BASED ON NATURAL LANGUAGE PROCESSING TO CHARITY. In The 11th International scientific and practical conference "Advanced technologies for the implementation of educational initiatives"(March 19–22, 2024) Boston, USA. International Science Group. 2024. 254 p. (p. 216).

[16] Li, X., Zheng, H., Chen, J., Zong, Y., & Yu, L. (2024). User Interaction Interface Design and Innovation Based on Artificial Intelligence Technology. Journal of Theory and Practice of Engineering Science, 4(03), 1-8.

[17] Wu, Y., Jin, Z., Shi, C., Liang, P., & Zhan, T. (2024). Research on the Application of Deep Learning-based BERT Model in Sentiment Analysis. arXiv preprint arXiv:2403.08217.

[18] Shi, C., Liang, P., Wu, Y., Zhan, T., & Jin, Z. (2024). Maximizing User Experience with LLMOps-Driven Personalized Recommendation Systems. arXiv preprint arXiv:2404.00903.

[19] Zhou, Y., Tan, K., Shen, X., & He, Z. (2024). A Protein Structure Prediction Approach Leveraging Transformer and CNN Integration. arXiv preprint arXiv:2402.19095.

[20] Yu, D., Xie, Y., An, W., Li, Z., & Yao, Y. (2023, December). Joint Coordinate Regression and Association For Multi-Person Pose Estimation, A Pure Neural Network Approach. In Proceedings of the 5th ACM International Conference on Multimedia in Asia (pp. 1-8).

[21] Xu, K., Zhou, H., Zheng, H., Zhu, M., & Xin, Q. (2024). Intelligent Classification and Personalized Recommendation of E-commerce Products Based on Machine Learning. arXiv preprint arXiv:2403.19345.

[22] Zhou, H., Xu, K., Bao, Q., Lou, Y., & Qian, W. (2024). Application of Conversational Intelligent Reporting System Based on Artificial Intelligence and Large

Language Models. Journal of Theory and Practice of Engineering Science, 4(03), 176-182.

[23] Li, L., Xu, K., Zhou, H., & Wang, Y. (2024). Independent Grouped Information Expert Model: A Personalized Recommendation Algorithm Based on Deep Learning.

[24] Che, C., Lin, Q., Zhao, X., Huang, J., & Yu, L. (2023, September). Enhancing Multimodal Understanding with CLIP-Based Image-to-Text Transformation. In Proceedings of the 2023 6th International Conference on Big Data Technologies (pp. 414-418).

[25] Huang, Zengyi, et al. "Research on Generative Artificial Intelligence for Virtual Financial Robo-Advisor." Academic Journal of Science and Technology 10.1 (2024): 74-80.

[26] Huang, Zengyi, et al. "Application of Machine Learning-Based K-Means Clustering for Financial Fraud Detection." Academic Journal of Science and Technology 10.1 (2024): 33-39.

[27] He, Zheng, et al. "Application of K-means clustering based on artificial intelligence in gene statistics of biological information engineering."

[28] Ni, Chunhe, et al. "Enhancing Cloud-Based Large Language Model Processing with Elasticsearch and Transformer Models." arXiv preprint arXiv:2403.00807 (2024).

[29] Xu, J., Zhu, B., Jiang, W., Cheng, Q., & Zheng, H. (2024, April). AI-BASED RISK PREDICTION AND MONITORING IN FINANCIAL FUTURES AND SECURITIES MARKETS. In The 13th International scientific and practical conference “Information and innovative technologies in the development of society”(April 02–05, 2024) Athens, Greece. International Science Group. 2024. 321 p. (p. 222).