Proceedings of the 4th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/45/20241029

A bibliometric review of technical applications of deep learning in computer vision

Gengjia Weng

Nanjing University of Finance and Economics, 3 Wenyuan Road, Qixia District, Nanjing, Jiangsu Province, China

gengjiaweng@gmail.com

Abstract. Computer vision is a field of Artificial Intelligence (AI), which is the field of technology that allows computers and systems to capture meaningful information from visual inputs such as images, videos, etc. through external cameras, data, and internal algorithms, and to take action or provide recommendations based on the captured information. Common application scenarios of computer vision are image recognition, image classification, object detection, pose detection, image segmentation, etc. In the era of artificial intelligence, computer vision plays the role of a "perceptron", which is the "eye" of the artificial intelligence era, and provides the "planning" and "decision-making" for the artificial intelligence. "Decision-making" provides an effective source of information and information support, and promotes the breakthrough development of computer vision technology while realising the iterative update of AI technology. By 2022, the market value of this field reaches $48.6 billion. In this paper, a visual knowledge graph data analysis of relevant literature on Web of Science (WOS) and China National Knowledge Infrastructure(CNKI) was conducted through CiteSpace. Among them, it focuses on analysing the development trend of image restoration technology based on diffusion model, which provides some help for scholars to carry out research.

Keywords: Artificial Intelligence, Deep Learning, Computer Vision, Visual Analytics

1. Introduction

After the 21st century, unstructured data are rapidly squeezing the existence space of structured data and becoming the mainstream network information dissemination mode. Due to the multi-dimensional, large-scale data volume and other characteristics of unstructured data itself, the development of computer vision technology once received the limitations of the objective operating system conditions, and was unable to make feedback on the large amount of information received in a short period of time.

Similar to how the human visual system works, the computer vision system works by constructing a multi-layered neural network, with lower layers recognising primary image features, a number of underlying features making up a higher layer of features, and finally making a classification at the top layer through the combination of multiple layers.

Deep learning is a branch of machine learning, which is one of the major breakthroughs and research hotspots in machine learning in recent years. Deep learning is a branch of machine learning, an algorithm for learning representations of information using artificial neural networks as an architecture [1]. The benefit of deep learning is to replace manual feature acquisition with unsupervised or semi-supervised

DOI: 10.54254/2755-2721/45/20241029

feature learning and hierarchical feature extraction efficient algorithms [2]. In less than a decade, the accuracy of computer vision systems has increased from 50% to 99% [3] In 2006, Geoffery Hinton, a professor at the University of Toronto, Canada, and a titan in the field of machine learning, and his student Ruslan Salakhutdi-nov published an article in Science, a leading international academic journal [4], which proposed the idea of deep learning for the first time. Deep learning is an algorithm in machine learning based on learning by characterising data. An observation (e.g. an image) can be represented using a variety of ways, such as a vector of intensity values for each pixel, or more abstractly as a series of edges, a region of a particular shape, etc. Whereas it is easier to learn tasks from examples (e.g., face recognition or facial expression recognition) using certain specific representations [5]

In this paper, relevant literature from Web of Science (WOS) and China National Knowledge Infrastructure(CNKI) literature databases are imported into CiteSpace software, and through data processing and computation, the literature content is transformed into a visual mapping, and the diffusion model is datamined and analysed accurately for its application in vision and its advancement in the iterative aspects of image restoration techniques.

2. Research methodology and data sources

2.1. Research methodology

CiteSpace is a bibliometric modelling software used to analyze and visualize co-citation networks. CiteSpace supports many types of bibliometric studies, including institutional co-citation analysis, author collaboration network analysis, and the visualization of topic and field co-occurrences, using scientific mapping procedures to Help visualize and analyse the structure, dynamic patterns and trends in the field, enabling researchers to intuitively identify the evolutionary paths of disciplinary frontiers and classic foundational literature. The software was developed by Meichao Chen, Professor (tenured), School of Information Science and Technology, Drexel University, Philadelphia, PA, USA. This paper uses CiteSpace 6.2.R4 for data visualization and analysis, focusing on the development trend of image restoration techniques based on Diffusion model.

2.2. Data Source

Data collection was carried out through the Web of Science (WOS) database by entering the keyword "Artificial Intelligence in Computer Vision" and selecting "4.17 Computer Vision & Graphics" (search date: 13 July 2023, the same below); data collection through the China National Knowledge Infrastructure(CNKI)database, input the subject term "Artificial Intelligence in Computer Vision", and then according to the "Main Topic", "Computer Vision" is the main topic. "The data were collected from the China National Knowledge Infrastructure(CNKI)database. Finally, the articles from both types of sources were exported and converted into visual data for analysis using CiteSpace.

3. Analysis of research data

3.1. Time and volume of publications

In Web of Science (WOS), the earliest time of publication was in 2013, and the publication continued until 2023, with a total of 794 papers in the related field. 2020 ushered in the largest increase in the number of publications on the topic. The number of publications remained high and relatively stable from 2020-2022, and it can be predicted that the trend will continue in 2023. It is possible to create a new high.

In China National Knowledge Infrastructure(CNKI), the earliest publication was in 1991, and the publication continued until 2023, with a total of 1019 papers in the related field. 2018 ushered in the largest increase in the number of publications on the topic. 2018-2022 have maintained a higher and smoother number of publications, and based on the number of publications that are currently available in 2023. It can be predicted that 2023 will continue the trend.

DOI: 10.54254/2755-2721/45/20241029

According to the statistics of the annual number of publications, the overall trend is upward. The development of this field is good, and from about 2018, the attention and real application needs in the social and scientific research field are rising. It has become a common academic hotspot in this field at home and abroad.

3.2. Author mapping of publications

CiteSpace was used to visualise and analyse the authors of the literature related to the technical applications of deep learning in the field of computer vision, with the time span of "2001-2023" and "Author" as the mapping option. Finally, a network-like author visualisation map with Node=186, Density=0.0231 was generated (shown in Figure 1). In the graph, the larger font size of the author's name indicates the larger amount of publications, and the authors of collaborative research are distributed in blocks.

As can be seen from the figure, there is a rich number of scholars publishing in this field, but the gap between the number of publications of each scholar is not very obvious. Mutual communication between scholars and integration of academic exchanges in this field shows a good trend but still needs to be further developed.

Figure 1. Author mapping of publications related to "Deep Learning in Computer Vision".

3.3. Mapping of the country of the author of the publication

CiteSpace was used to visualize and analyze the issuing organizations of the literature in the field of diffusion model image denoising, with the time span "2001-2023" and "issuing organizations" as the mapping option, and a network-like visualization of the issuing organizations with Node=49, Density=0.0213 was generated (shown in Figure 2). The network visualization map of the countries where the authors of the articles are located can effectively reflect the countries that are more mature in the development of the field at present.

The field is popular and valuable, but international cooperation and exchange is low. There are barriers and geographical factors that affect the choice of partners among countries.

The developed countries have more research and advantages than the developing countries, except for China, the US, and India, which are the top three. They have technological and economic incentives for R&D in high-tech industries.

DOI: 10.54254/2755-2721/45/20241029

Figure 2. Country mapping of authors of publications related to "Deep Learning in Computer Vision".

4. Research hotspots and frontier analyses

4.1. Keyword co-occurrence mapping analysis

CiteSpace was used to visualise and analyse the keywords of the literature related to the technical applications of deep learning in the field of computer vision, and the time span was set as "2001-2023", with "keywords" as the mapping option. The keyword network visualisation map with Node=183, Density=0.0225 was generated (as shown in Figure 3). The font size in this network visualisation is positively correlated with the frequency of the word in the cited literature, and the thickness of the lines connecting the words is positively correlated with the degree of correlation between the words. Among them, the keyword "computer vision" and "artificial intelligence" appears 143 times in total. In addition, the keywords that appear more frequently are "deep learning", "machine learning", "natural language", "big data", "image processing", "vehicle initiative safety", etc. These high-frequency keywords reflect, to some extent, the main research directions of the technical application areas of deep learning in computer vision and the connection of various academic subfields from 2001 to 2023. It can also be seen that the research on the technical application of deep learning in the field of computer vision has been deepening and expanding, and the academic field has a relatively large academic achievement system, with the development results of various subcategories communicating and connecting with each other, and further optimizing on the basis of the mutual support of each other in the technical field. In the future computer vision field, it is expected that convolutional neural network and recurrent neural network (RNN) based on deep learning will become very popular network models, and will make better breakthroughs and progress in more application research [6].. Currently, research on combining deep learning and reinforcement learning is still in its infancy, but some research work in this area has already achieved good performance on multi-object recognition tasks [7] and learning for video games [8], which is one of the reasons for the excitement of many researchers in related fields.

DOI: 10.54254/2755-2721/45/20241029

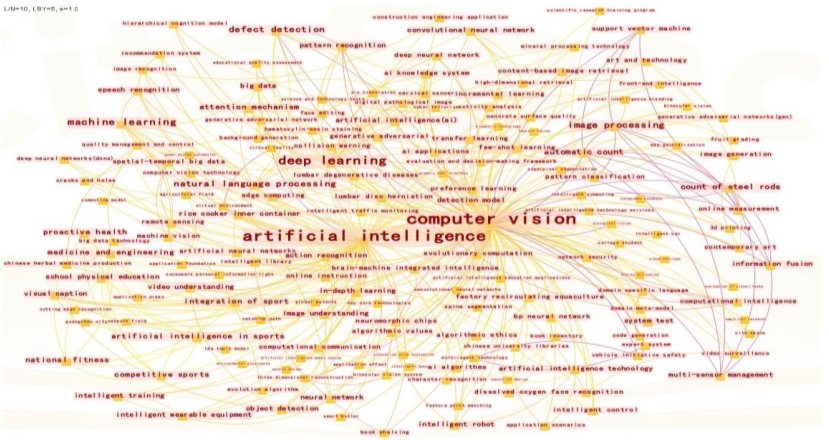

Figure 3. Keyword co-occurrence mapping of literature related to "Deep Learning in Computer Vision"

4.2. Keyword clustering mapping analysis

Using CiteSpace to visualise and analyse the keywords of the literature related to the technical application areas of deep learning in the field of computer vision, the keywords are clustered, and the five categories with the highest number of keywords are screened for the keyword clustering atlas production (shown in Figure 4). Modularity Q is 0.562, indicating that this keyword clustering map has a significant network-like structure (the value is greater than 0.3). The ordinal numbers in the plots are negatively correlated with the frequency of occurrence of the keywords in this category within the selected literature. The keyword clustering order is artificial intelligence, computer vision, deep learning, natural language processing, automatic count. thus, it can be seen that the development direction of the technical application field of deep learning in the field of computer vision has diversified characteristics, but to a certain extent also has a certain degree of concentration, and the degree of development of subfields is more mature. Overall, the field still has a large space for development and good research prospects.

Figure 4. Keyword clustering map of literature related to "Deep Learning in Computer Vision".

4.3. Keyword emergence mapping analysis

Keyword emergence refers to a large change in the frequency of occurrence of a keyword within a specific period of time, so CiteSpace was used to analyse the keywords in the keyword visualisation process in the literature related to the technological applications of deep learning in the field of computer vision. The time span is set as "2001-2023" and the number of generated samples is 9 (as shown in Figure 5). The field has been evolving rapidly with frequent technical updates since its emergence. The field has expanded from image recognition and processing to neural language, neural structure network, and other cross-disciplinary areas. The future research will focus on the integration of basic disciplines and the cross-field development of cutting-edge areas.

DOI: 10.54254/2755-2721/45/20241029

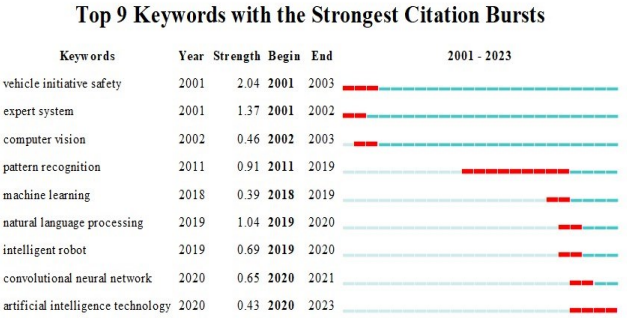

Figure 5. Keyword mutation mapping of literature related to "Deep Learning in Computer Vision".

5. Conclusion

Based on the high-quality literature about "technical application of deep learning in computer vision" in Web of Science (WOS) and China National Knowledge Infrastructure(CNKI)literature databases, this paper scientifically applies CiteSpace software to analyse the data of visual knowledge mapping.

This field is a hot topic with a growing number of publications since 2018. The authors are diverse and prolific, but more development and collaboration are needed. The research is distributed among many countries, but the international cooperation is low. Developed countries dominate the field, except for China, the US, and India, which are the top three. Future research will focus on the integration of basic disciplines and the cross-field development of cutting-edge areas

.Although CNNs have made great progress compared with traditional machine learning methods in areas such as computer vision, there are still many issues that require further research. These hyperparameters, such as learning rate, filter size, step size, and number, are all very internally dependent, and any small adjustment may have a large impact on the final training results [9].

In summary, deep convolutional neural network has the characteristics of local connection, weight sharing and multi-level network structure, which can extract the feature information from the image more effectively, and promote the development of research technology in the field of computer vision, but it is not comparable to the human brain system [10].

References

[1] Schmidhuber, J. Deep Learning in Neural Networks: an Overview. neural Networks. (2015) Neural Networks, 61: 85-117. arXiv:1404.7828 . doi:10.1016/j.neunet.2014.09.003 , https://doi.org/10.1016/j.neunet.2014.09.003

[2] Song, H. A.; Lee, S. Y. Hierarchical Representation Using NMF.(2013) ICONIP 2013: Neural Information Processing, volume 8226, pp 466–473, doi:10.1007/978-3-642-42054-2_58, https://link.springer.com/chapter/10.1007/978-3-642-42054-2_58

[3] 7 Breathtaking Practical Examples of Computer and Machine Vision, Bernard Marr, Forbes, 8 April 2019

[4] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J].(2006), Science., volume 313, Issue 5786 pp. 504-507, https://www.science.org/doi/abs/10.1126/science.1127647

[5] Glauner, P. Deep Convolutional Neural Networks for Smile Recognition (Dissertation). Imperial College London, Department of Computing. (2015). arXiv:1508.06535, https://doi.org/10.48550/arXiv.1508.06535

[6] Hongtao Lu,Qinchuan Zhang, A review of research on deep convolutional neural networks in computer vision Data Acquisition and Processing. (2016), 31(01), TP391.41;TP183, 10.16337/j.1004-9037.2016.01.001 Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China,

DOI: 10.54254/2755-2721/45/20241029

https://kns.cnki.net/kcms2/article/abstract?v=ZPxNsP0eklExo0nyL3aI0lQEz-GlU5t4lNhhGzL_DajcDC6AbiWfvIjBTrvAaoj4nsLacMET9aie_gaO1Q25K-nImHtEnAiRo0RoSnE2hKvdiw6GYY__FmqbdI1PYM1zb0Z2JEC0OuvD3lvJmxZ9jA==&uniplatform=NZKPT

[7] BaJ, Mnih V, Kavukcuoglu K. Multiple object recognition with visual attention [EB/OL].(2014), https://doi.org/10.48550/arXiv.1412.7755

[8] MnihV, KavukcuogluK, SilverD, etal. Human-level control through deep reinforcement learning [J]. (2015) Nature, 2015, 518 (7540): 529 -533. https://www.nature.com/articles/nature14236/?source=post_page

[9] Chao Chen, Feng Qi, A review of the development of convolutional neural networks and their applications in the field of computer vision, School of Management Science and Engineering, Shandong Normal University, Shandong, China

[10] Guohe Li, Yinghan Qiao, Weijiang Wu, Yifeng Zheng, Yunfeng Hong, Xiaoming Zhou, Deep Learning and Its Application in the Field of Computer Vision, (2019) Application Research of Computers / Jisuanji Yingyong Yanjiu, Dec2019, Vol. 36 Issue 12, p3521-3564. 10p. DOI:10.19734/j.issn.1001-3695.2018.08.0604 https://web.p.ebscohost.com/abstract?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=10013695&AN=140259632&h=VZ867FiB2hdDpvTLpHtaAScJ%2f4X8PKPO1PeHouBoX1%2f41YY8DAJGQQQSgLA3WO1Rf%2faFrSor%2bp0l6dsLyH6yAg%3d%3d&crl=c&resultNs=AdminWebAuth&resultLocal=ErrCrlNotAuth&crlhashurl=login.aspx%3fdirect%3dtrue%26profile%3dehost%26scope%3dsite%26authtype%3dcrawler%26jrnl%3d10013695%26AN%3d140259632