Current Research and Applications of Combining Computer Vision (CV) and Machine Learning (ML) at Home and Abroad

CiteSpace-based Integration of CNKI and Google Scholar's Public Research Results for 2019-2024

Xinyue Zhang*

Hebei University of Technology, Tianjin China

*602858601@qq.com

Abstract. Computer vision provides the theoretical and algorithmic basis of image and scene analysis for machine vision, and machine vision provides the sensor model, system structure and means of implementation for the realisation of computer vision. The two share a set of theoretical systems, only the direction of development is different, machine vision focuses on practical applications in the industrial field, while computer vision focuses on the research of theoretical algorithms. Therefore, the cross research of computer vision and machine learning has become one of the important directions in the current artificial intelligence research field. This paper first defines the concepts of artificial intelligence, computer vision and machine learning, and further illustrates the strong correlation between computer vision and machine learning in terms of methodology and research logic. Then, it further uses bibliometrics to integrate the research results of Knowledge and Google Scholar based on CiteSpace for the years 2019-2024 to reveal the relationship between the two in scientific practice in a realistic way. Finally, the article conducts the qualitative analysis of this paper to summarise the current state-of-the-art research and applications at the intersection of computer vision and machine learning and their problems. The integration of previous research will facilitate the academic community's understanding of important nodes and fundamental issues in the discipline's research process, and more purposeful planning of future research.

Keywords: computer vision; machine learning; artificial intelligence; Bibliometrics; theoretical applications.

1 Introduction

According to the conjecture and definition of Claude Elwood Shannon, the founder of information theory, in 1948, computer vision, as a scientific discipline, 'investigates related theories and techniques in an attempt to build artificially intelligent systems capable of obtaining "information" from images or multidimensional data' [1]. As

E. P. H. Lau et al. (eds.), *Proceedings of the 2024 3rd International Conference on Information Economy, Data Modelling and Cloud Computing (ICIDC 2024)*, Advances in Computer Science Research 114,
https://doi.org/10.2991/978-94-6463-504-1_7

mentioned above, because of their complementary methodological and scientific perspectives and their applications, machine semester and computer vision have become a current research hotspot in the last decade. The field of computer vision is distinguished by its diversity and imperfections, and machine learning is a complex practical methodology, so the research and application of computer vision and machine learning are always full of opportunities and challenges, and the scope, innovations and limitations of their current research should be discussed.

Relevant research originated in the 1970s and 1980s, and there were already 1,040,000 publications 14 years ago, mainly exploring the interaction methods, theoretical innovations, development possibilities and ethical issues of computer vision and machine learning. In the last decade, the number of articles has remained stable at 1,300,000, but the main research direction has been biased towards the practice and application of the intersection technology between the two (Table 1). Domestic research in China started late, but has developed rapidly and has become a current research hotspot. Bibliometrics of all 2202 studies in the cross-field of computer vision and machine learning retrieved by CNKI found that the research trend in this field is on the rise, with some related studies being advanced in 2006, and the period of 2016-2018 was a significant rise, in which the annual average of 200 publications was rapidly increased. The research heat in this area has continued at a high level since 2019, reaching an average of 350 publications per year (Figure 1) [2].

Table 1. Features of dataset extracted from SIGCOMM and INFOCOM (self-generated)

Attribute name	**Type of attribute**	**Count**	
		SIGCOMM	**INFOCOMM**
Number of articles	Numerical	1962	6745
Number of authors	Numerical	4196	8415
Names of authors	String	4196	8415
Number of institutes	Numerical	576	1678
Names of institutes	String	576	1678
Number of references	Numerical	46809	142487
References from last 10 years	Numerical	33650	105753
Citations of articles	Numerical	76594	210555
Number of participating countries	Numerical	57	70
Names of participating countries	String	57	70

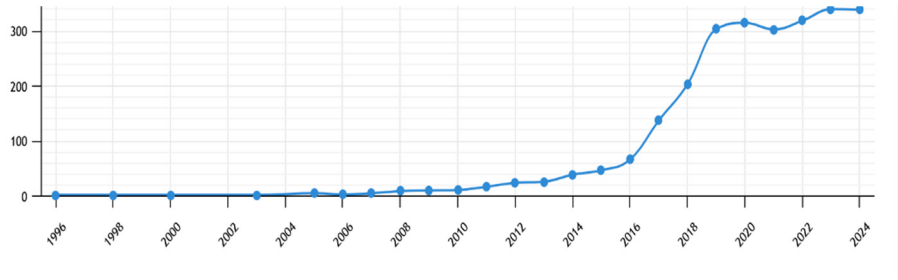

Fig. 1. Visualisation of the amount of research literature published on the intersection of computer vision and machine learning on the Knowledge Network, 1996-2024 (self-generated)

2 Computer vision and Machine Learning

2.1 Strong Correlation Between Artificial Intelligence, Computer Vision and Machine Learning in Real-World Studies

Artificial Intelligence (AI) is a broad field that involves computers modelling various aspects of human intelligence, including learning, reasoning, perception, understanding, planning and action. Artificial Intelligence aims to give computers human-like intelligence to solve a variety of complex problems. Computer Vision (CV) is a branch of Artificial Intelligence that focuses on allowing computers to take information from an image or video and understand its meaning. Computer Vision uses techniques such as image processing, image analysis and machine learning to enable computers to recognise, track and interpret objects and scenes in images. Machine Learning (ML) is one of the core components of Artificial Intelligence, which is a technique that uses data and experience to train machines to improve their performance. Machine Learning uses algorithms to allow computers to learn from data and automatically discover patterns and regularities in order to predict and analyse unknown data. Computer vision and machine learning are closely related in many ways. Image processing and analysis techniques commonly used in computer vision can be optimised and improved by machine learning algorithms. For example, tasks such as classifying, recognising and tracking images using machine learning algorithms can automatically improve the performance of computer vision systems. At the same time, computer vision also provides a large number of datasets and application scenarios for machine learning to test and improve the performance of its algorithms. In summary, artificial intelligence is a broad field that covers several disciplines and technologies. Computer vision, as a branch of AI, focuses on image and video processing and analysis. Machine learning, on the other hand, is an important means of achieving AI, which can improve the performance and accuracy of computer vision and other fields. In practice, these three fields often intersect and merge with each other to solve a variety of complex problems. By gaining a deeper understanding of their differences and connections, they can be better applied to solve practical problems [3].

Theoretically, machine learning (ML) belongs to the discipline of methodology, computer vision (CV) is the application of the scene. To solve problems in CV, many

methods in ML are used, which are not only limited to CV, but also can be applied to data mining, financial analysis, artificial intelligence, etc. As a concrete example, a common problem in CV is visual tracking of moving targets. Among the methods for solving this CV problem, there is a large class of methods belonging to the discriminative type, i.e., the tracking problem is regarded as a binary classification problem between the target and the background on each image frame [6]. The tracking process trains the classifier by continuously collecting and updating positive and negative samples, and the one with the largest classifier response from the current target to be selected is selected as the tracking result for that frame. The binary classification problem here is typical of the methods used in ML, e.g., Support Vector Machine (SVM), naive bayes, etc.

From the published research results, computer vision is also strongly correlated with machine learning. In this paper, the content relevance of 200 papers in the field of machine learning and computer vision published on Zhi.com and Google Scholar in the last five years was analysed using CiteSpace, and the statistical model used for bibliometrics was as follows:

$$g^2 \leq k\sum_{i\leq g} c_i, k \in Z^+ \tag{1}$$

$$\mathbf{ccv}\ (i,j) = \frac{cc(i,j)}{\sqrt{c(i)*c(j)}} \tag{2}$$

Taking computer vision as the core independent variable, the research development of 200 Chinese literatures can be found that the core content of computer vision research in machine learning, 128 (64%) of the articles are research in the cross direction of machine learning and computer vision, while the field is densely associated with artificial intelligence and deep learning. Computer vision ethics, image recognition and 3D pose recognition, on the other hand, are the main research branches in this cross-cutting area [4].

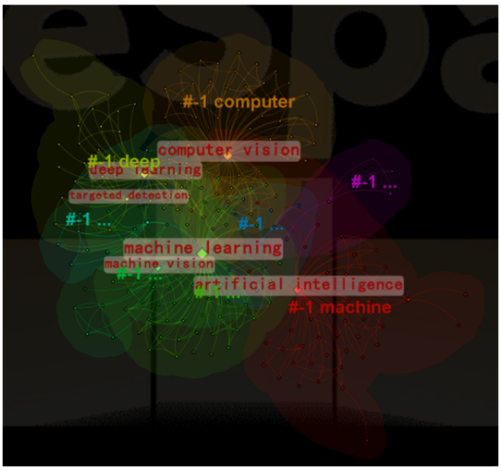

Fig. 2. Measurement and Relational Analysis of Two Hundred Relevant Literature in CNKI (Chinese, self-generated)

2.2 Research and Application of Combining Machine Learning and Computer Vision at Home and Abroad

The study of network image recognition methods for computer vision systems based on machine learning is an important research direction in this cross-cutting field. Image recognition algorithms typically use machine learning and deep learning models to identify objects by analysing each individual pixel in an image [5]. The image recognition algorithm is fed with as many labelled images as possible in an attempt to train the model to recognise objects in the image. The dataset is first collected and labelled with images. Once the images are labelled, they are fed into the neural network to train on the images. After the model has been trained on the dataset, it is fed a "test" dataset containing unseen images to validate the results. The model will use the knowledge learnt from the test dataset to predict the object or pattern present in the image and try to recognise the object. As mentioned above, a large number of scholars are currently conducting research in this direction [6]. Domestic research in computer theory has focused on the enhancement of image recognition technology. More typically, in 2023, Zhu took the network image recognition method of computer vision system as the research object, and explained the existing machine learning-based recognition method from the aspects of design principle, algorithm design and processing flow. Based on the principle of the original algorithm, the structural balanced network is introduced, distinguishing the pixels as nodes, taking the grey scale matrix of the image as the connection weight matrix, distinguishing the connection weights with the help of Hadamard product transform, extracting the feature parameter, traversing the grey scale image, and weighting the feature parameter, and ultimately realising the goal of improving the recognition method. Overseas mainly doing ethical research in computer vision and machine learning, key examples include important books by Shalev-Shwartz, S. and Ben-David, S in 2014 and a new paper by Mintarya in 2023 [7]. Simon also systematicly introduced a new computer vision model (Figure 3):

$$\begin{aligned} \mathbf{Pr}\,(\boldsymbol{x}) &= \int \mathbf{Pr}\,(\boldsymbol{x},\boldsymbol{y})\boldsymbol{dy}, \\ \mathbf{Pr}\,(\boldsymbol{y}) &= \int \mathbf{Pr}\,(\boldsymbol{x},\boldsymbol{y})\boldsymbol{dx}. \end{aligned} \tag{3}$$

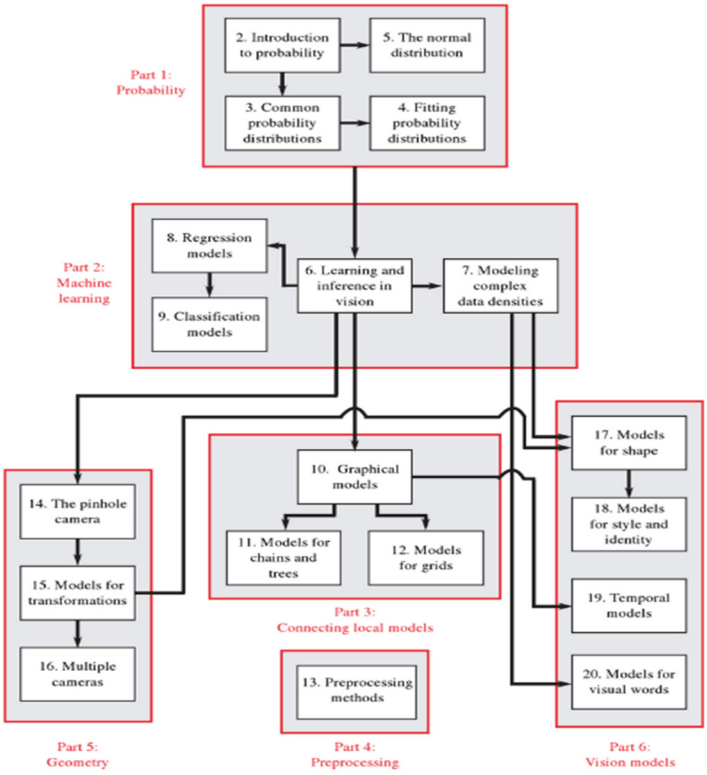

Fig. 3. Core Structure of the introduction of Simon's new CV model and theory (Simon)

These researches are centred on image recognition and have made important contributions to human-machine inspection and item classification. However, they still remain at the 2D level, lacking in motion capture and behaviour prediction. On the other hand, the corresponding attitude capture systems have been updated and applied in this cross-cutting area, the first uses being in the sports and industrial sectors. Tian Haonan and Wang Jun presented a study on the application of OpenPose and machine learning in football training in the new Society of Sports Science [4], based on which Zhu Haoyu,Xu Guodong further investigated the application of pose recognition techniques in tennis [5]. These are able to examine the behavioural postures of the target characters in different activities for more scientific sports training and game judgments. Overseas research seems to be developing even more rapidly and the functionality has been expanded into the engineering industry. Khan, M.R in 2024 proposed a model for assessing and testing the stability of worker's working posture and building erection in construction works, which will effectively ensure the labour safety of workers and the quality of the building. These studies have expanded the field of machine learning and artificial intelligence to 3D, which will open up possibilities for its wider and higher dimensional use.

3 Conclusion

By interpreting the current research of CNKI and Google Scholar, it is possible to deeply realise the importance of the development of the integration of machine learning and computer vision for the development of AI technology. With the gradual upgrade of 2D vision to 3D vision, 3D vision perception industry enterprises are favoured. In the future, China's machine vision industry will flourish towards 3D machine vision, machine learning, deep learning, and deepening downstream penetration. In addition, the rebound of traditional industrial products, the concept of artificial intelligence is strong will also bring new life to the machine vision industry. Further integration and research on computer vision and machine learning will contribute to China's future technological and economic development.

References

1. ZHU Qinghua. Research on network image recognition method for computer vision system based on machine learning. Information Recording Materials,2023,24(11):161-163.
2. TIAN Haonan,WANG Jun. Research on the application of OpenPose and machine learning in football training. Chinese Society of Sports Science. Abstracts of the Thirteenth National Sports Science Conference,2023:3, 88-173.
3. ZHU Hao-Yu,XU Guo-Dong. Application of posture recognition technology in tennis//Chinese Society of Sports Science. Abstract collection of papers of the 13th National Sports Science Conference,2023:3. 188-222.
4. FU Yongmin,FAN Lei,et al. Cascade detection of tobacco impurity images based on computer vision and machine learning. Journal of Light Industry,2023,38(04):113-121.
5. Khan, A.I. and Al-Habsi, S., 2020. Machine learning in computer vision. Procedia Computer Science, 167, pp.1444-1451.
6. Dhanya, V.G., Subeesh, A., Kushwaha, N.L., Vishwakarma, D.K., Kumar, T.N., Ritika, G. and Singh, A.N., 2022. Deep learning based computer vision approaches for smart agricultural applications. Artificial Intelligence in Agriculture, 6, pp.211-229.
7. Shalev-Shwartz, S. and Ben-David, S., 2014. Understanding machine learning: From theory to algorithms. Cambridge university press.