Research progress and trend analysis of computer vision based on cite space

Yuzhi Liu

1Guangdong University of Technology, China

Abstract. In recent years, with the rapid development of artificial intelligence and the proposal of efficient algorithms such as deep learning, the ability of computer vision to deal with complex problems has made a qualitative leap. Therefore, the study of computer vision is deeply expanding from academia to industry, and the application of industry continues to be enriched. It involves many disciplines in natural science and social science. The main applications and technologies include image classification, image recognition, deep learning, convolutional neural network, etc. In this paper, the bibliometrics software CiteSpace was used to conduct visualization processing based on the literatures in the core database of Web of Science with the time span from 2010 to 2020 and the subject words as computer vision. The current situation of scientific research cooperation is analyzed from the perspectives of countries, institutions, individuals and cited literatures themselves. The involvement of computer vision discipline is analyzed from the perspectives of natural sciences and social sciences in front, and analyzed from the funding institutions. Funded institutions are analyzed from the perspective of computer vision discipline. The current research hotspots and development frontiers are analyzed by using maps. It is found that the future computer vision will focus on deep learning, transfer learning and semantic segmentation. Finally, it points out that there are some problems in the current research, such as the lack of close scientific research cooperation, the lack of comprehensive discipline involved, and the difficulty of technology implementation. Three suggestions are put forward to solve these problems: (1) Holding trans-regional academic exchange meetings to promote the sharing of academic achievements. (2) Raising the attention to the partial humanities and promote the depth of disciplines. (3) Considering the cost and efficiency in practical application to accelerate the realization of scenario application.

1. Introduction
In recent years, with the rapid development of artificial intelligence, computer vision, as one of the branches, has been widely used in industrial manufacturing, Internet communication, smart city and other fields. In this background, countries around the world have designated relevant development strategies according to their own national conditions in order to further promote automated production and improve economic production efficiency. Germany, committed to intelligent production, has put forward the concept of "Industry 4.0", in which the use of computer vision and Internet cutting-edge technology is one of its development strategies to realize intelligent factory [1]. The United States, known as a world power, followed the tide of scientific and technological intelligentization and put forward the “National Strategic Plan for Artificial Intelligence Research and Development”. On the
road of artificial intelligence intellectualization, they vigorously develop computer vision technology, enhance the perception of intelligent system, and promote economic revitalization and industrial quality development [2]. At the same time, facing the increasingly mature artificial intelligence technology, China puts forward “2025 Made in China”, comprehensively promotes the strategy of manufacturing power, deeply integrates information, intelligence and industrial manufacturing, and further penetrates artificial intelligence into all fields [3]. On this basis, computer vision, as one of the popular applications of artificial intelligence, connects vision and artificial intelligence, which has become an effective means to obtain information intelligently. Because of the rapid ability to analyze and obtain information, computer vision is widely used in intelligent manufacturing, intelligent service and other fields, and plays a vital role in promoting the development of national productivity and economic structure adjustment. It has also gradually become the hot spot and frontier of researchers' discussion.

By simulating human vision, computer vision can quickly and accurately analyze and understand the acquired images. In 1982, Marr's book "Vision" marked the emergence of computer vision as an independent discipline. In the 1950s, computer vision began to have preliminary applications. At that time, the main applications were 2D image processing based on statistical model, such as micrograph processing, surface defects of parts, etc. By the 1960s, researchers could use computers to extract three-dimensional images from two-dimensional images and describe their spatial state. In the 1970s, MIT offered a course related to computer vision -- "Machine Vision". [4] The development of computer vision is very slow because the early technology is not mature and the limited application field. But with artificial intelligence booms, the development of computer vision ushered in the hope.

In the early 21st century, deep learning, machine learning and other efficient algorithms were proposed, which greatly improved the ability of computer vision to deal with complex problems, and the development of computer vision became fast. At present, it’s very broad for the application of computer vision and it attracts a large number of international scholars to study on it. Tadhg Brosnan pointed out that computer vision can be applied to image detection and classification, and these applications have made great contributions to food detection. He uses computer vision technology to classify vegetable images according to their hue and saturation, and then sort out the defective vegetables and fruits [5]. Ronald Poppe applied computer vision technology to target tracking. By modeling and analyzing the images transmitted by sensors and predicting the state of the next possible motion, it provided a better solution for scenes such as monitoring and human-computer interaction [6]. By studying the application of computer vision in image recognition, Komar proposed a deep neural network model. Combined with Caffe framework, this model has excellent performance in handwritten digital image recognition [7]. In addition, international researchers have also found that computer vision has a lot of interactions with other disciplines. Beijbom uses image understanding and other techniques to identify and annotate coral reefs on the seafloor, providing great help for biological research [8]. Nigel mentioned in the report that face recognition is closely related to education. Specifically, face recognition is used to recognize the status of students in class, and then targeted suggestions are put forward to improve the teaching method [9]. Weiming reported that there was a connection between computer vision and management. He pointed out that target tracking in computer vision can be applied to human flow statistics and remote personnel management, which provides great convenience for personnel management [10]. In China, there are also a large number of scholars continuing to explore the application fields and cross-disciplines of computer vision, found that computer vision has penetrated into all aspects. Zhang Qi applied computer vision to the field of image recognition. By optimizing and improving the convolutional neural network and modifying some layer structures of the neural network, he obtained an efficient model to identify liver cancer [11]. Bai Ziyi popularizes computer vision to face recognition in video by using deep learning and image fusion methods [12]. Bao Yuequan pointed out that there is a close relationship between computer vision and architecture, which is embodied in the fact that crack identification has made a great contribution to structural damage identification in architecture [13]. Bu Yingjia proposed an algorithm for license plate recognition based on computer vision, which achieved good results in areas
such as community access management and highway intersection fees [14], indicating that there is a certain connection between computer vision and management. In general, the fields and disciplines involved in the Computer Vision Institute in China and other countries are generally the same. However, foreign research is more in-depth, walking in the forefront of academic research, which is specifically reflected in the high frequency of foreign literature cited. However, Chinese researchers are more open-minded and adept at improving existing methods. For example, Zhang Qi summarized from a large number of foreign literatures and proposed a new optimization structure of convolutional neural network, whose model recognition rate was much higher than the previous traditional methods [11].

To sum up, computer vision technology has been widely used in image classification, image recognition, image understanding, face recognition, target detection, target tracking and other fields by virtue of its objectivity to information acquisition and recognition. Computer vision is closely related to the natural and social sciences, including biology, management, education, architecture and so on. It is contributed to many fields because of its wide application range, we need the statistical measurement analysis of the literature of the current study. The analysis of the current development of computer vision hotspots and speculate the future development of the frontier is important to promote the rapid development of national productivity, industrial structure transformation, development of intelligent manufacturing.

In this paper, based on the Web of Science core collection database and computer vision related literature, CiteSpace software was used to analyze the visual knowledge graph, revealing the hot spots, institutions, related disciplines, coupling between literature, current hot research directions and future research trends of computer vision research.

2. Analytical methods and data collection and processing

2.1. Analytical methods

Bibliometrics is a statistical method to analyze the structural characteristics and change rules of literature groups. It carries out quantitative research by digging the information of authors, key words and references, etc., so as to enable researchers to obtain the internal correlation from a large number of literatures and excavate the potentially valuable information [15].

CiteSpace is a software that enables quantitative analysis of the literature and visualizes the results of the analysis. It was developed by Dr. Chaomei Chen and Drexel University in the United States to help researchers quickly understand the current hot developments and future research trends in a particular discipline. CiteSpace is favored by a large number of researchers at home and abroad due to its powerful analytical capability and concise visual interface [16].

2.2. Data processing

Web of Science is a high-impact database with rich coverage of disciplines and a large number of collected literatures [16]. In order to ensure the authority of the results, the data in this paper were obtained from the core database of Web of Science. Firstly, this paper takes "computer vision" as the key word, downloads more than 40,000 papers from the Web of Science website, and names the format as "download.txt", and analyzes the scientific research cooperation and exchange of computer vision and the latest research progress. Then, in order to study the subject relationship involved in computer vision, this paper uses the keywords "TS = (Computer Vision AND Natural Science) OR (Computer Vision AND Social Sciences)" to screen the literature in the core database of Web of Science, downloads 1566 literatures, and makes the same processing. Due to incomplete literature collection in 2021 and to ensure the reliability of the results of the econometric analysis, the time span of the downloaded literature in this paper is from 2010 to 2020.
3. Analysis of scientific research cooperation
Research results are usually produced by a group of researchers from different countries and institutions. At the same time, an excellent academic achievement must also cite other people's literature as theoretical support. In order to deeply analyze the cooperation among researchers in the field of computer vision, this section will analyze the cooperation among scholars from the perspectives of countries, institutions, researchers and the cited literatures.

By using CiteSpace software to analyze the literature, the maps of collaboration between researchers from different perspectives were obtained. The node in the figure represents the sum of the number of corresponding attributes; the lines in the figure represent the connection between two types of nodes, and the thickness of the lines is proportional to the degree of cooperation. Between centrality refers to the number of shortest paths of other nodes through this node, which shows the degree of citation closeness between literatures. The greater the between centrality, the higher the importance of the node, the greater the influence and the greater the value of information transmission [17]. In addition, CiteSpace would divide the cooperative spectrum into multiple clusters, and the inner nodes of the same cluster were closely related, conversely, the correlation is lower.

3.1. Cooperation and exchange among countries, institutions and individuals
Figure 1a shows a country-level map of computer vision research collaborations. It can be intuitively seen from the figure that the network nodes of many countries represented by the United Kingdom, Germany, Spain, France, the United States, Canada, Italy, China and India are relatively prominent, and they have done a lot of research on computer vision. Among them, six countries, Germany, Spain, the United Kingdom, Italy, France and the United States, show a complex academic exchange network in the figure, and have a very close cooperation among them. The reason for the formation of a US-led cooperation and exchange circle can be seen from the relevant policies implemented by the United States. As early as the 1940s, in order to enhance its international status and comprehensive national strength, the United States implemented a technical assistance program and carried out scientific research exchanges and cooperation in Europe [19]. And it is one of the strategies of the United States to carry out scientific research cooperation with the modern western countries. With the arrival of Industry 4.0 and the rise of the tide of artificial intelligence, the United States proposed the further development of computer vision in the National Strategic Plan for Artificial Intelligence Research and Development [2]. In order to further strengthen its own strength, the United States will continue to conduct cooperation and exchanges with modern western powers to form a win-win academic exchange atmosphere.

It can also be found from the figure that Canada, China, India and Australia are loosely connected with other countries and far away from the academic exchange circle with the United States as the core, indicating that the scope of cooperation and exchange among them is not wide enough and the depth of communication is shallow. However, it is worth noting that although China and other countries do not carry out much scientific research exchanges in computer vision, China has the largest number of published literatures, which indicates that Chinese researchers have a high enthusiasm for computer vision research and close exchanges and cooperation within the country. In addition, the United States, the United Kingdom, France and Germany all have a large number of published literatures. In a word, the related research of computer vision has attracted the attention of many scholars.
Figure 1. Collaborative of computer vision research at the national level.

Figure 2 shows an agency-based computer vision research collaboration map. The resulting map is divided into two regions. At the top of the map, Chinese Academy of Sciences, leads domestic institutions such as Zhejiang University, Tsinghua University, have a lot of exchanges with academic institutions in neighboring Singapore, resulting in a confused and complex communication map. Among them, the relevant institutions are Singapore Nanyang University of Technology, National University of Singapore and so on. At the bottom of the map, the Massachusetts Institute of Technology (MIT) is the center, and Stanford University, Carnegie Mellon University, and the University of California, San Diego also have frequent collaborations. In general, there is close cooperation between many research institutions and the number of published papers is relatively objective.

Figure 2. Collaboration of computer vision research at the institutional level.
Table 1 shows the intermediate neutral status of the top 10 cited authors. Consider this in the context of Figure 2, although the Chinese Academy of Sciences published the most papers and showed the radiation effect on the surrounding institutions, the top 10 authors contained few Chinese researchers. This shows that foreign papers are more influential and have higher authority. However, Chinese literature is characterized by a large number of published papers but less influential, which may be because Chinese researchers have not done in-depth research on computer vision, and most of the studies only focus on the improvement of existing technologies while ignoring the depth of research. Chinese scholars should improve their academic innovation, actively solve problems from different perspectives, and strive to improve the influence of their research achievements.

Table 1. Top 10 authors of cited mediating centrality.

| Centrality | Author       |
|------------|--------------|
| 0.33       | Kohli P      |
| 0.32       | Vedaldi A    |
| 0.30       | Everingham M |
| 0.29       | Boykov Y     |
| 0.28       | Gonzalez RC  |
| 0.27       | Krizhevsky A |
| 0.27       | Quevedo R    |
| 0.25       | Cubero S     |
| 0.17       | He KM        |
| 0.17       | Uijlings JRR |

Figure 3 depicts the map of computer vision research collaboration at the individual level. As shown in the figure 3, DAWEN SUN as the core has published the largest number of papers in the cooperation field, which indicates that he and the researchers around him have frequent cooperation and communication and achieved good academic results. Professor DAWEN SUN is committed to research in the field of food and biology. He combines food safety with computer vision technology and uses image recognition technology to detect food quality, thus quickly classifying low-quality food and improving industrial production efficiency [20]. This also shows the wide application range of computer vision, has been deeply applied to the field of food safety. XUELONG LI and LEI ZHANG, who were second only to DAWEN SUN in the number of published papers, have collaborated with a large number of scholars, as shown in the figure. In addition, other scholars have formed their own cooperative networks and have academic exchanges within their own academic circles. However, it is a pity that there are fewer researchers in the academic frontier, and there is a lack of deeper and closer communication between them. This may be because different scholars have different research directions and fields, and the disciplines involved are also different, so there is still room for improvement in the cooperation between researchers.
Whether it is at the national, institutional or individual levels, the current state of cooperation shows a good situation. This shows that computer vision has attracted the attention of scholars from all levels, and has achieved a large number of research results in this field. However, there are still some problems in the current research exchange at all levels.

(1) At the national level: there are still a few countries lack of research on computer vision, and the cooperation with the current hot research countries is not close. The internal development of some countries is not balanced, and there is no attention paid to scientific research development or to the attention and investment of computer vision research. For example, the development of artificial intelligence in India is restricted by the unbalanced development of the country, which shows the low vitality of the development of artificial intelligence [21]. As a part of artificial intelligence, computer vision is also restricted.

(2) Institutional level: due to geographical conditions, language differences and other reasons, most of the institutions stay in the cooperation between their own countries or neighboring countries, and the cooperation between continents is not much. As can be seen from the figure, the academic exchange circle of Chinese institutions with Chinese Academy of Sciences as the core and the academic exchange circle of foreign institutions are separated, and the cooperation and exchange between them lack of depth and breadth.

(3) Personal level: the cooperation between individual researchers is not close enough and is limited to their own academic circle. Moreover, scholars at the forefront of academic cooperation are less, and they lack interdisciplinary and cross-disciplinary academic achievements sharing. If cutting-edge scholars actively collaborate, the number of larger nodes in the collaborative graph increases, their locations become more clustered, and the network becomes more complex and disorganized.

3.2. Academic literature exchange

Through CiteSpace software, references of the collected data were analyzed, and the resulting graph was shown in Figure 4. As can be seen from the figure, the literatures of He KM are most cited, and the literature of Ren Sq, Russakovsky O and Krizhevsky A and others are more cited. The academic achievements of these people are the basis of computer vision research. In addition, CiteSpace divided the analysis map into two areas, with the warmer areas of sharing academic results based on HE's literature at the top. The following is a literature cross-citation network led by Chang CC. In the above network, researchers' literatures are cited more frequently and their academic achievements are relatively authoritative. In the network below, researchers' literatures are rarely cited, but the network...
is relatively dense, and researchers in the network learn from each other. In general, researchers learn from each other, share high-quality academic results, and form a good trend of continuous innovation and common progress.

![Literature citation map.](image)

**Figure 4.** Literature citation map.

4. Research discipline analysis

4.1. Natural science and social science disciplines

The collected literature was analyzed with CiteSpace software, and the structure obtained was shown in Figure 5. Since computer vision is a branch of artificial intelligence, the most involved subjects are computer science and engineering. In addition, computer vision and many natural science disciplines have interaction, such as optics, geography, biology, chemistry, physics and so on; Interactions with social science disciplines include environmental science, psychology, education, agriculture, etc. Not only that, now there are more and more researchers promoting computer vision to more disciplines, so that computer vision for human services in all aspects. Guo proposed in his report that computer vision and architecture are closely connected, which can be specifically reflected in the fact that target detection and image understanding in computer vision can be used to warn workers whether there is danger in their current state [22]. Zhao pointed out in his research that computer vision is widely used in access control of face verification and law enforcement surveillance video to manage personnel [23], which shows that computer vision has a certain connection with management.
In general, computer vision is a highly comprehensive discipline, involving both natural science and social science. It covers a wide range of disciplines, which shows that computer vision provides great convenience in human life. However, the application of computer vision in social science is not comprehensive and deep enough. If computer vision can go deep into history and other humanities subjects, it will play an important role in promoting the development of these subjects and the improvement of computer vision technology. Therefore, it is of great value to study the current hot field of computer vision and its development frontier.

4.2. Funding analysis
In order to explore the disciplines involved in computer vision from the side, this section counted the institutions with the most funding from 2010 to 2020, as shown in Table 2. It can be seen that the National Natural Science Foundation of China (NSFC) has provided great support for the research of computer vision. In addition, computer vision has become the direction of NSFC with the largest amount of funding in computer-related fields in the past few years [24]. In addition, the National Science Foundation, Fundamental Research Funds for the Central University and China Postdoctoral Science Foundation all provide a lot of funds support for computer vision. It is worth noting that the Engineering and Physical Sciences Research Council mainly funds the research in engineering and physics, which also shows that computer vision is intersected with engineering and physics and is a field with strong intersecting with other disciplines.
Table 2. Top 5 institutions in terms of the number of grants.

| Frequency | Organization                                                      |
|-----------|-------------------------------------------------------------------|
| 128       | National Natural Science Foundation of China (NSFC)               |
| 37        | National Science Foundation (NSF)                                 |
| 18        | Fundamental Research Funds for the Central University             |
| 17        | Engineering and Physical Sciences Research Council                |
| 11        | China Postdoctoral Science Foundation                             |

5. Research progress and trend analysis of computer vision

5.1. Research progress analysis
Research progress is reflected by research hotspots, which refer to the topics on which academic research on some issues has been carried out and a large number of papers have been published in a certain field in a certain period of time. These topics represent current areas of focus for researchers. Therefore, the analysis of research hotspots is of great help to quickly understand the current main research direction. Dr. Chen Meichao proposed co-occurrence network based on bibliometric method to mine potential information in literature [25], while co-occurrence network based on keywords can reflect research hotspots in a certain field over a period of time. In this section, keywords of CiteSpace software papers are used for analysis to form a network of research hotspots with keywords as nodes to analyze the current research hotspots of computer vision. As shown in the figure, centered on the keyword Computer Vision, a large number of related algorithms, technologies and applications burst out.
Table 3 shows the top 10 hot keywords. As it can be seen from Table 3, besides computer vision, there are also deep learning and convolutional neural network, which are popular research points at present. This is because the rise of artificial intelligence drives the development of deep learning, and convolutional neural network algorithm and other efficient deep learning algorithms have been widely used in the field of computer vision. In addition, model and algorithm two keywords appear frequently in literature samples, because the technical core of computer vision is that algorithms play a role in frontier models, so as to realize the application of computer vision. It is worth noting that traditional applications of computer vision, such as classification, recognition and image processing, are still popular in recent years. While computer vision expands new fields, it also constantly improves and innovates existing applications.

Table 3. Computer vision key word frequency statistics table.

| Frequency | Keyword                                |
|-----------|----------------------------------------|
| 8313      | computer vision                        |
| 2385      | deep learning                          |
| 2023      | classification                         |
| 1592      | model                                 |
| 1572      | recognition                           |
| 1481      | system                                |
| 1453      | algorithm                             |
| 1385      | convolutional neural network           |
| 1207      | image processing                       |
| 1189      | segmentation                           |

5.2. Trend analysis

The research trend represents the latest research progress in a certain research field, which is involved in the literature actively cited by researchers recently, and plays an important role in promoting the development of this field [26]. CiteSpace was used to conduct keyword frontier analysis on the collected literatures, and the top 15 cited hot keywords from 2010 to 2020 were sorted out, as shown in Figure 7. The red ones represent the time period when the frequency of keyword usage bursts. The figure clearly shows that in the early stage, researchers' research on computer vision remained in
simple application scenarios such as human-computer interaction and camera correction, but with the passing of time, the research focus shifted to the application of gesture recognition, kinematics and other features extraction. At present, researchers show great interest in deep learning, convolutional neural networks, transfer learning, semantic segmentation and other applications. With the rise of deep learning technology, the learning advantages of deep neural network are fully explored. With the help of deep learning technology, computer vision is applied in a large number of scenes. He K proposed using residual learning framework to optimize the training of deep neural network, which greatly improved the accuracy and precision of image classification [27]. In Vijay's study, he mentioned a deep neural network structure based on codecs, which can perform semantic segmentation with efficient memory and time [28]. Rowley can better realize face detection by establishing a neural network filter and combining overlapping detection and arbitration [29]. In the future, computer vision will be further combined with deep learning and become a key development direction. In addition, transfer learning has gradually become the development frontier of computer vision due to its ability to overcome the problems such as the time-consuming network training of convolutional neural network. In addition, semantic segmentation complies with the development trend of the intelligent era, and is bound to become a research hotspot in the future when it is combined with emerging applications such as unmanned driving and robot navigation. The following paper will make a detailed analysis of these three research hotspots.

5.2.1. Deep learning. Deep learning builds a multi-level learning network structure by simulating the way the human brain recognizes external information. "Deep" is embodied in the multi-level network structure, while "learning" is embodied in learning the characteristics of information from low level to high level through the multi-level structure [30]. The common learning methods of deep learning include supervised learning and unsupervised learning. The convolutional neural network, which is widely used in computer vision, belongs to supervised learning. Convolutional neural network can effectively learn image feature information by using its own input layer, convolutional layer, pooling layer, full connection layer and output layer [31].
In the early studies, image recognition mostly relied on directional gradient histogram and directional gradient histogram, which could only capture a small amount of image information, and the recognition effect was very limited [33]. Until 2006, Hinton et al. found that deep neural network had great advantages in learning characteristics and training, thus triggering a boom in deep learning.[34] With the development of deep learning, problems that were difficult for computer vision to solve in the past have been presented with breakthrough solutions, such as greatly improving the accuracy and speed of image recognition. In 2012, A Krizhevsky et al. proposed a convolutional neural network based on deep learning, which refreshed the recognition record of data sets with large sample size, which also fully proved the advantages of deep learning in computer vision [35]. Deep learning has attracted the attention of a large number of researchers because of its ability to deal with problems that are difficult for human beings and its great commercial value. Industry leaders such as Stanford University, the world-renowned university, and Microsoft, the Internet giant, have also successively carried out research on deep learning [36].

In the future, deep learning will play an important role in how to quickly and accurately obtain the information you need in the era of information explosion. For example, deep learning is used to extract features of human faces in real-time urban monitoring and compare them with the face images of criminals, so as to quickly identify and locate them, which is of great practical significance to the security field. In general, the application of the combination of computer vision, such as face recognition, image recognition, target tracking, and deep learning to mine the value of information will be a hot research area in the future.

5.2.2. Transfer learning. In order to overcome the disadvantage that the mechanism of machine learning to generate data does not change with the environment, transfer learning is born. It can learn relevant information from the existing training results and apply it to the new training model to reduce the cost of training. The greater the similarity between the two fields of transfer learning, the better the sharing effect of transfer learning is [37].
Figure 9. Examples of transfer learning [38].

Although the arrival of deep learning has made a qualitative leap in the performance of computer vision technology, convolutional neural network has limited the performance of computer vision in image recognition due to large training sample size, long training time due to too many parameters, and problems such as over-fitting in the case of less sample training. The characteristics of transfer learning are the key to deal with these problems and break the limitation of image processing by computer vision [39].

The emergence of transfer learning greatly reduces the training time of the model, and it can also achieve better results in the case of some rare samples, so it is widely used in the application scenarios of computer vision. For example, a common application of transfer learning is the classification of medical images. In many cases, there are too few samples of training data due to the lack of pathological images for rare diseases and the fact that some diseases present vastly different clinical diagnostic images in patients with different constitutions. The transfer learning can learn relevant information from the trained medical image classification model, so as to face the classification problem under a small number of medical image training samples [40]. In the future, transfer learning will be applied to the field of computer vision recognition and classification with small or massive samples by relying on its fast learning and powerful ability to deal with rare samples. For example, in the field of unmanned driving, the actual traffic situation is affected by weather, pedestrians, vehicles, etc., which is full of complexity and uncertainty. However, if the characteristics of transfer learning are utilized to train and learn on the basis of existing traffic detection models, the training time can be reduced and more unknown and complex road conditions can be identified, which will be of great value to the advancement of unmanned driving technology.

5.2.3. Semantic segmentation. With the deepening of computer vision research, more and more scene applications need to use the knowledge expressed in images, such as measuring the size of human tissues and organs in medical images, software analysis of personal clothing matching, and so on [41]. The main task of semantic segmentation is to segment the image and label the things contained in the image [42]. It communicates information in images by connecting computer vision and other applications. Nowadays, semantic segmentation has become one of the common application fields of computer vision. In many scenarios, we can see related applications, such as human-computer interaction, image search engine, augmented reality and so on. Although traditional machine learning methods can deal with partial semantic segmentation problems, due to the emergence of deep learning, especially the advantages of convolutional neural network in image processing, the speed and precision of semantic segmentation have made a qualitative leap [43].
At present, in the era of rapid development of science and technology, more and more people pay attention to the application of automatic driving, unmanned aerial vehicle, intelligent robot, intelligent monitoring and so on. Semantic segmentation, as the carrier of these intelligent applications, will attract the attention of researchers and be deeply used in every field of life. For example, the semantic segmentation technology is applied to the mobile phone software, and the mobile phone camera can be used to query and analyze the scene seen in real time, so that people can obtain more detailed and accurate information from the scene in front of them.

6. Conclusions
Under the tide of artificial intelligence, the application of computer vision is deepening. Countries, institutions and individual researchers cooperate, communicate and share with each other, and the integration of computer vision is constantly strengthened, and combined with the disciplines in natural science and social science. Countries pay more and more attention to computer vision year by year, and provide a lot of funds support for computer vision. The current hot research fields of computer vision are deep learning, classification and so on. In the future, computer vision will focus on deep learning, transfer learning and semantic segmentation. However, there are still following shortcomings in the current research of computer vision.

(1) The research cooperation between some countries and researchers is not close enough. Although countries, institutions, individuals have academic exchanges in the field of computer vision research. However, the scope of communication is not wide enough. It is only limited to its own academic circle, and there is no more extensive communication and interaction.

(2) The subject involved is not comprehensive and in-depth. Although there are interactions between computer vision and natural science, social science, there are few cooperative studies in some humanities subjects in social science, such as history, economics and so on.

(3) It is difficult to realize the technology application. There have been a lot of researches on computer vision related technologies, but few of them are actually implemented due to factors such as time, money and cost, and the gap between actual scenes and theoretical scenes.

In order to make the computer vision better forward development, in view of these three deficiencies put forward the following suggestions.

(1) Promote exchanges and share results. Countries and institutions should hold cross-regional academic exchanges according to the actual situation, so that scholars in the same research field can share their research results and promote a good cooperative environment.

(2) Promote discipline interaction. Researchers should pay more attention to the more partial humanities, such as the use of computer vision image understanding techniques in history, better analysis of historical images, or use of image addition and other techniques to complete incomplete historical photographs.

(3) Combine theory with practice to accelerate the realization of scene application. For some of the excellent performance of the technology, they should consider the cost and efficiency in the practical
application, combined with the problems that may occur in the practical application of optimization, to achieve the optimal cost and efficiency, so as to better promote the development of computer vision.

References
[1] Posada J, Zorrilla M, Dominguez A, et al. Graphics and Media Technologies for Operators in Industry 4.0[J]. IEEE Computer Graphics and Applications, 2018, 38(5):119-132.
[2] Jiang Fengguang, Xiong Bolong, Zhang Chao. How to Achieve Strategic Breakthrough in Artificial Intelligence in China: A Comparison and Interpretation of Four Reports on Artificial Intelligence Development in China and the United States [J]. Research on Modern Distance Education, 2020, 32(01):3-11.
[3] CAI Zixing. Forty Years of Chinese Artificial Intelligence [J]. Science and Technology Review, 2016, 34(15):12-32.
[4] CHEN Dan. Development and Application of Computer Vision Technology [J]. Computer Knowledge and Technology, 2008, 4(35):2449-2450+2452.
[5] Tadhg Brosnan and Da-Wen Sun. Improving quality inspection of food products by computer vision—a review [J]. Journal of Food Engineering, 2004, 61(1): 3-16.
[6] Ronald Poppe. Vision-based human motion analysis: An overview [J]. Computer Vision and Image Understanding, 2006, 108(1): 4-18.
[7] Komar M, Yakobchuk P, Golovko V, et al. Deep Neural Network for Image Recognition Based on the Caffe Framework[C]// 2018:102-106.
[8] Beijbom O, Edmunds P J, Kline D I, et al. Automated annotation of coral reef survey images[C]// IEEE, 2015.
[9] Nigel Bosch, Sidney K. D'mello, Jaclyn Ocumpaugh, Ryan S. Baker, Valerie Shute. Using Video to Automatically Detect Learner Affect in Computer-Enabled Classrooms [J]. ACM Transactions on Interactive Intelligent Systems (TiiS), 2016, 6(2).
[10] Weiming Hu, Tieniu Tan, Liang Wang 0001, Stephen J. Maybank. A survey on visual surveillance of object motion and behaviors. [J]. IEEE Trans. Systems, Man, and Cybernetics, Part C, 2004, 34(3).
[11] Zhang Qi. Pathological image classification based on convolutional neural network [D].
[12] Bai Ziyi, Mao Yirong, Wang Ruiping. Review of video face recognition [J]. Computer Science, 2021, 48(03):50-59. (in Chinese)
[13] Bao Yuequan, Li Hui. Civil engineering in the era of artificial intelligence [J]. China Civil Engineering Journal, 2019, 52(05):1-11.
[14] The Bu Ying Family. Research on algorithm of license plate recognition system based on computer vision [D]. University of Electronic Science and Technology of China, 2015.
[15] Zheng Wenhui. A comparative study of bibliometrics and content analysis [J]. Journal of Information, 2006(05):31-33.
[16] Hou Jianhua, Hu Zhigang. Review and prospect of CiteSpace software application research [J]. Modern Information, 2013, 33(04):99-103.
[17] Qian G. Scientometric sorting by importance for literatures on life cycle assessments and some related methodological discussions [J]. International Journal of Life Cycle Assessment, 2014, 19(7):1462-1467.
[18] Chaomei Chen, Zhigang Hu, Shengbo Liu, Hung Tseng. Emerging trends in regenerative medicine: a scientometric analysis in CiteSpace [J]. Expert Opinion on Biological Therapy, 2012, 12(5).
[19] Miller, Clark A. "An Effective Instrument of Peace": Scientific Cooperation as an Instrument of U.S. Foreign Policy, 1938-1950. [J]. Osiris, 2006.
[20] Sun D W, Sun D W. Computer Vision Technology for Food Quality Evaluation [J]. Computer Vision Technology for Food Quality Evaluation, 2008:57–80.
[21] D Mcduie-Ra, Gulson K. The backroads of AI: The uneven geographies of artificial intelligence and development [J]. Area, 2020, 52(3).
[22] Guo Brian H. W., Zou Yang, Fang Yihai, Goh Yang Miang, Zou Patrick X. W. Computer vision technologies for safety science and management in construction: A critical review and future research directions [J]. Safety Science, 2021, 135.

[23] Zhao W Y, Chellappa R, Phillips P J, et al. Face Recognition: A Literature Survey [J]. ACM Computing Surveys, 2003, 35(4):399-458.

[24] Hou Sujuan, Zhao Ruizhen, Song Su. [J/OL]. Proceedings of the National Academy of Sciences of the United States of America. Journal of Software: 1-6[2021-03-18].https://doi.org/10.13328/j.cnki.jos.005849.

[25] Chen C, Yue C, Horowitz M, et al. Towards an explanatory and computational theory of scientific discovery [J]. Journal of Informetrics, 2009, 3(3):191-209.

[26] Chaomei Chen. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature [J]. Journal of the American Society for Information Science and Technology, 2006, 57(3).

[27] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition [J]. 2016.

[28] Vijay, Badrinarayanan, Alex, et al. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Scene Segmentation. [J]. IEEE transactions on pattern analysis and machine intelligence, 2017.

[29] Rowley H A, Baluja S, Kanade T. Neural Network-Based Face Detection[C]// Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 1996.

[30] Lecun Y, Bengio Y, Hinton G. Deep learning [J]. Nature, 2015, 521(7553):436.

[31] Dumoulin V, Visin F. A guide to convolution arithmetic for deep learning [J]. 2016.

[32] Ketkar N. Convolutional Neural Networks [J]. Springer International Publishing, 2017.

[33] Zhang Jianming, Zhan Zhicai, Cheng Keyang, Zhan Yongzhao. Research and development of deep learning [J]. Journal of Jiangsu University (Natural Science Edition), 2015, 36(02):191-200.

[34] Hinton G E, Salakhutdinov R R. Reducing the Dimensionality of Data with Neural Networks [J]. Science, 313.

[35] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks[C]// NIPS. Curran Associates Inc. 2012.

[36] Kai Y, Lei J, Chen Y, et al. Deep Learning: Yesterday, Today, and Tomorrow [J]. Journal of Computer Research & Development, 2013, 20(6):1349.

[37] Pan S J, Qiang Y. A Survey on Transfer Learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10):1345-1359.

[38] Zhuang F, Qi Z, Duan K, et al. A Comprehensive Survey on Transfer Learning [J]. 2019.

[39] Xie Xiaohong, Li Wentao. Research on image classification training method based on transfer learning [J]. Information and Computer (Theoretical Edition), 2020, 32(07):53-55.

[40] Apostolopoulos I D, Bessiana T. Covid-19: Automatic detection from X-Ray images utilizing Transfer Learning with Convolutional Neural Networks [J]. 2020.

[41] Chen Hongxiang. Image Semantic Segmentation Based on Convolutional Neural Network [D]. Zhejiang University, 2016.

[42] Wang P, Chen P, Yuan Y, et al. Understanding Convolution for Semantic Segmentation[C]// 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018.

[43] Garcia-Garcia A, Orts-Escolano S, Oprea S, et al. A Review on Deep Learning Techniques Applied to Semantic Segmentation [J]. 2017.

[44] Arbelaez P, Hariharan B, Gu C, et al. Semantic segmentation using regions and parts. IEEE, 2012.