

Research on Computer Vision Image Classification

Shiqin Wang

School of Computer Engineering, Anhui Sanlian University, Hefei, 230601, China

Abstract: *Computer vision image classification is an important research direction in the field of computer vision, which aims to classify images into different categories for more effective processing. With the continuous increase of image data, how to effectively classify and organize this data has become a persistent issue. This article introduces the research process, methods, applications, and future development trends of computer vision image classification in detail.*

Keywords: *computer Vision Image; classification; applications*

1. Introduction

In the digital era, with the development and application of information technology, there is a growing demand for computer vision technology. Computer vision is a technology to analyze and process images by computer, which can help people better understand and utilize image data. Therefore, the classification of computer vision images is a very important research task. At present, the classification of computer vision images is mainly divided into two types: binary classification and multi-category classification. Among them, binary classification refers to dividing images into two different categories, while multi-category classification refers to dividing images into multiple different categories. Both classification methods have their advantages and disadvantages, and a suitable classification method needs to be chosen for the specific problem. With the continuous increase of image data, how to effectively identify and classify images has become an important research topic. Computer vision is a branch of artificial intelligence that aims to use computers to simulate human visual systems, to obtain image information and classify, recognize, and understand it [1,2]. This article will review the research on computer vision image classification, with a focus on deep learning-based classification methods. The research on computer vision image classification can be traced back to the 1960s, when the focus was on image recognition and classification [3]. With the development of computer hardware and software, computer vision image classification has also been continuously improved. In the 1980s, the research focus of computer vision image classification began to shift towards machine learning [4]. With the emergence of deep learning algorithms, research on computer vision image classification has been further enhanced. In recent years, with the development of big data and cloud computing technology, research on computer vision image classification has been further accelerated [5,6]. Currently, computer vision image classification has been widely applied in various fields such as autonomous driving, medical image analysis, etc. [7]. As computer vision image classification technology continues to develop, researchers have also begun to focus on more complex issues such as recognition and classification of multiple objects in images, instance segmentation of images, etc. [8,9]. To solve these problems, researchers have developed many new algorithms, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), etc. In addition, research on computer vision image classification is constantly moving towards more intelligent and automated direction. Researchers are working to develop algorithms that can automatically learn features and achieve image classification, reducing the need for human intervention and improving the accuracy and efficiency of image classification. Overall, the research process of computer vision image classification has undergone a continuous evolution from recognition to classification, from manual feature extraction to automatic feature learning, and has achieved a technological leap while continuously improving computing power. At the same time, the research process of computer vision image classification is a constantly evolving process that expands the application fields of image classification technology while constantly improving computing power and algorithm efficiency. With the further development of technology, future research on computer vision image classification will continue to bring more amazing breakthroughs.

2. A Review of Computer Vision Image Classification Techniques

2.1. Basic concepts of computer vision image classification

Computer vision image classification is a technique that uses computers to recognize, classify and label images. It mainly relies on machine learning algorithms to identify objects in an image by analyzing the shape, color, texture, lighting and other features of the image. Computer vision image classification has a wide range of applications in various fields, such as: self-driving cars, image retrieval, surveillance systems, medical image analysis, etc. It can help computers to identify objects in images quickly and accurately, thus enabling automated image analysis and processing. Computer vision image classification technology continues to evolve and is becoming more and more pervasive. As the technology advances, it can play a greater role in handling complex image classification tasks and help people understand the information in images faster and more accurately.

2.2. General process of computer vision image classification

Computer vision image classification is a technique that uses computers to recognize, classify, and label images. It mainly relies on machine learning algorithms to identify objects in images by analyzing their shape, color, texture, lighting, and other features. Computer vision image classification has a wide range of applications in various fields, such as self-driving cars, image retrieval, monitoring systems, medical image analysis, etc. It can help computers quickly and accurately identify objects in images, enabling automated image analysis and processing. Computer vision image classification technology is constantly developing and becoming increasingly universal. With the advancement of technology, it can play a greater role in handling complex image classification tasks and help people understand the information in images more quickly and accurately.

Computer vision image classification is a technique that enables computers to recognize and classify objects in images. The general process of computer vision image classification is as follows:

Table 1: The general process of computer vision image classification

No.	General process
1	Data collection: Collect relevant image datasets for the task.
2	Data preprocessing: Preprocess the images, such as resizing, cropping, grayscale conversion, normalization, etc., for subsequent processing.
3	Feature extraction: Extract image features, such as edges, corners, textures, etc.
4	Feature selection: Select features based on importance and relevance to reduce computational complexity and improve classification accuracy.
5	Model selection: Select an appropriate classification model, such as SVM, neural networks, etc., based on task requirements.
6	Model training: Train the classification model using annotated datasets to optimize model parameters.
7	Model evaluation: Evaluate the trained model using a test set, such as calculating accuracy, recall, F1 value, etc.
8	Application deployment: Deploy the trained model to an application to classify new images.
9	Model optimization: Optimize the classification model based on actual usage, such as parameter tuning, data augmentation, etc., to improve classification accuracy and performance.
10	Transfer learning: Use a pre-trained model for transfer learning to reduce training time and improve classification accuracy for new tasks.
11	Deep learning: Use deep learning techniques, such as convolutional neural networks, recurrent neural networks, etc., to improve image classification accuracy and robustness.
12	Object detection: Extend image classification to object detection to detect and locate multiple objects in images.
13	Semantic segmentation: Extend image classification to semantic segmentation to recognize and segment different semantic regions in images.
14	Real-time processing: Optimize image classification algorithms to process and classify real-time video streams.
15	Application scenarios: Apply image classification technology to different scenarios, such as intelligent transportation, intelligent security, medical imaging, etc., to achieve more applications and value.
16	Weakly supervised learning: Use weakly labeled or unlabeled data for image classification learning, such as label noise robustness, semi-supervised learning, unsupervised learning, etc.
17	Multi-modal fusion: Extend image classification to multi-modal fusion, such as combining text, speech, sensors, and other information for classification.
18	Reinforcement learning: Use reinforcement learning techniques to optimize and self-learn image classification algorithms.
19	Federated learning: Use federated learning techniques to achieve distributed computing and data privacy protection.
20	Model interpretation: Analyze and interpret the decision-making process of the classification model to provide explanations and promote transparency.
21	Model ensemble: Integrating multiple classification models to improve classification accuracy and robustness.
22	Model compression: Compressing and optimizing classification models to achieve lightweight models and deploy them on resource-constrained devices.
23	Model evaluation and monitoring: Evaluating and monitoring classification models to discover problems and vulnerabilities, and improve the reliability and security of the models.

In summary, computer vision image classification is a constantly evolving and changing field, which involves the comprehensive application of many technologies and algorithms, and needs to be optimized and practiced in combination with specific application scenarios.

2.3. Evaluation criteria of computer vision image classification

Image classification is an important task in the field of computer vision. In order to evaluate the performance of a classifier, several factors need to be considered, such as accuracy, recall and F1-score metrics. Accuracy is the ratio of the number of data points correctly classified by the classifier to the total number of data points; recall is the ratio of the number of all data points correctly recognized by the classifier to all data points that can be recognized; and F1-score is a composite metric, which is a score value calculated by weighting the average of accuracy and recall. In addition, there are some other evaluation methods, such as ROC curve, Precision-Recall Curve, etc. These methods can be used to compare the performance differences between different algorithms. In addition to the above metrics, the robustness of the classifier needs to be taken into account. That is, whether the performance of the classifier remains stable when the input dataset changes. This can be assessed by testing different datasets or by using cross-validation methods. Also, it can be determined by comparing different models or algorithms to determine which is more suitable for a particular task.

3. Main methods of computer vision image classification

The classification of computer vision images is an important part of visual image recognition. To classify images, various algorithms and techniques are used to identify features in the image and classify them based on these features.

3.1. Statistical Methods

Statistical methods were an early approach to computer vision image classification. These methods typically use statistical features of the image, such as color, shape, and texture, to classify the image. These methods usually extract features from the image and use a classifier (such as SVM, decision trees, etc.) to classify the image. The advantages of statistical methods are their simplicity and high stability for a large amount of image data. However, the disadvantage of statistical methods is their inadequate classification performance for complex images because they only consider local features of the image.

Rule-based classification methods use a set of expert knowledge and hand-designed rules to classify images. The advantage of this method is that it can classify images quickly and is easy to implement, suitable for special cases of classification problems. For example, for face recognition, classification rules can be designed based on the shape and position of the face. However, the limitation of rule-based classification methods is that they are not applicable to all image classification problems, and it is difficult to completely cover hand-designed rules for some complex image classification problems.

Feature-based classification methods refer to extracting each part of the image, selecting useful features from them, and then classifying these features. The advantage of this method is that it can handle image details more finely and apply many image processing techniques to improve classification accuracy. For example, the method of feature extraction based on gray-level histogram is one of the commonly used image classification methods.

3.2. Deep Learning-based Methods

Deep learning-based methods are the latest approach to computer vision image classification. These methods typically include Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs). Compared to statistical methods, deep learning methods have better classification performance for complex images.

Convolutional Neural Network (CNN) is a commonly used computer vision algorithm, which uses convolution operations to extract features from images and classify them. The convolution operation is a method for extracting features from image data, which performs convolution with a fixed set of numbers (i.e. convolutional kernel) to obtain the features of the image. Compared with traditional manual design methods, convolutional neural networks have more powerful abstraction and generalization performance. The core component of a convolutional neural network is the convolutional layer. It can transform the input data and extract useful information. At the same time, the convolutional layer can

also realize the computation of nonlinear functions, which enhances the expressive power of the model. In addition, the convolutional neural network also includes a pooling layer and a fully connected layer to further improve the expressiveness of the model. In practical applications, convolutional neural networks are widely used for image recognition tasks. (As shown in figure 1)

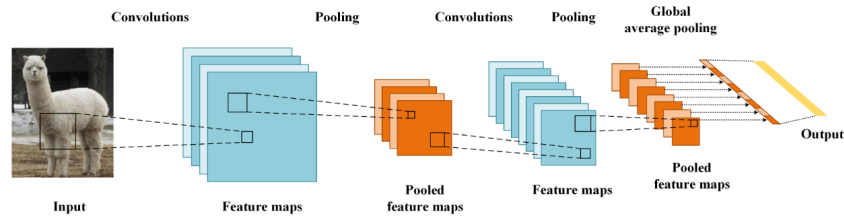


Figure 1: Convolutional Neural Network (CNN)

Recurrent Neural Network (RNN) is a computer vision algorithm used to recognize sequence data. In image classification, recurrent neural networks can be used to solve complex problems, such as face recognition, object detection, etc. The core components of recurrent neural networks are the convolution kernel and pooling operations, which are used for spatial representation and denoising filtering of the input image, respectively. The convolution kernel is a multidirectional graph structure that captures local features in an image and converts them into a high-dimensional spatial representation, while the pooling operation is a nonlinear transformation that reduces the size of the input image and preserves important feature information. Recurrent neural networks typically consist of one or more convolutional layers, one fully connected layer, and one or more pooling layers. Each convolutional layer spatially represents the input image and transforms it into a higher dimensional spatial representation; the fully connected layers are responsible for combining these spatial representations into a new spatial representation to obtain more accurate results; and finally, the pooling layers further reduce the size of the input image and retain important feature information. Thus, recurrent neural networks can perform the convolution-pooling operation by multiple cycles to achieve the image classification task.

4. Current status and trends in the application of computer vision image classification

4.1. Current status of computer vision image classification applications

With the continuous development of artificial intelligence technology, computer vision image classification application has gradually become one of the important areas of concern. At present, computer vision image classification is widely used in medical diagnosis, security monitoring, traffic management and so on. Among them, medical image analysis is one of the most prominent aspects of computer vision image classification applications. By automatically recognizing and classifying medical images, it can help doctors make accurate disease diagnosis results more quickly, thus improving patient treatment results and social benefits. In addition, computer vision image classification can also be used in scenarios such as vehicle detection and road environment monitoring in intelligent driving systems to achieve a safer and more reliable driving experience. In addition to the above application scenarios, computer vision image classification is also widely used in industrial automation production lines. For example, in the robot manufacturing process, computer vision image classification can be used to achieve rapid positioning and assembly operations of different parts; while in the food processing industry, computer vision image classification can be used for real-time assessment and control of food quality. In short, computer vision image classification, as an emerging technology, is expanding its application scope and playing an increasing role.

4.2. Future trends of computer vision image classification methods

Computer vision image classification technology is widely used in various fields, such as medical image analysis, video surveillance, computer graphics, intelligent transportation, etc. By using image classification technology, the management and analysis efficiency of image data can be greatly improved, and the accuracy of image analysis can be improved. In recent years, computer vision image classification methods have made significant progress.

Future trends in computer vision image classification methods:

(1) Further development of deep learning methods: deep learning methods have achieved significant results in the field of computer vision and will be further researched and improved in the future.

(2) Higher efficiency of image classification methods: with the improvement of hardware technology, image classification methods will become more efficient and real-time.

(3) Classification of multimodal data: computer vision image classification methods will not be limited to single image classification and will begin to consider multimodal data such as speech and image classification.

(4) Improvement of rule-based methods: rule-based methods will be improved to adapt to the constantly changing application needs.

(5) Classification of large-scale image data: with the continuous increase in data volume, classification methods will need to address the challenge of large-scale image data.

5. Conclusion

Computer vision image classification research is a very important and active field. With the continuous development of technology, image classification technology is also constantly improving. Currently, deep learning-based image classification technology has been widely used in multiple fields and achieved significant results. Therefore, we expect to continue researching and developing computer vision image classification technology in the future to better support image analysis and processing needs. In general, computer vision image classification is a very important research field that not only helps us better understand image data but also has important practical significance for multiple fields. Computer vision image classification methods will continue to develop in the future to meet the growing application demands.

References

- [1] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", *arXiv preprint arXiv: 1409.1556*, 2014.
- [2] R. Girshick, J. Donahue, T. Darrell, J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation", *arXiv preprint arXiv: 1311.2524*, 2013.
- [3] A. Krizhevsky, I. Sutskever, G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Advances in Neural Information Processing Systems 25*, 2012.
- [4] J. Deng, W. Dong, R. Socher, L. -J. Li, K. Li, L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
- [5] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, A. Zisserman, "The Pascal Visual Object Classes (VOC) Challenge", *International Journal of Computer Vision*, 88(2), 2010.
- [6] R. L. Duda, P. E. Hart, D. G. Stork, "Pattern Classification", 2nd ed., John Wiley & Sons, 2001.
- [7] C. M. Bishop, "Pattern Recognition and Machine Learning", Springer, 2006.
- [8] S. Haykin, "Neural Networks and Learning Machines", 3rd ed., Pearson, 2009.
- [9] S. Theodoridis, K. Koutroumbas, "Pattern Recognition", 4th ed., Academic Press, 2009.