Proceedings of CONF-MLA Workshop: Mastering the Art of GANs: Unleashing Creativity with Generative Adversarial Networks
DOI: 10.54254/2755-2721/80/2024CH0068

A Review of Research on Image Classification Methods

Jiarun Wu

Qingdao Jiaozhou Yingzi Private School, Qingdao, China

wuj58046@gmail.com

Abstract. As one of the important research directions in the field of computer vision, image classification has a wide range of applications. Image classification is an important foundation for pattern recognition, machine learning, and artificial intelligence. Image classification generally includes three steps: region of interest selection, feature extraction, and classifier modeling. Among them, feature extraction of images is an important foundation for completing other tasks. In most pattern recognition scenarios, appropriate feature representation is a crucial step that directly affects the performance of the entire classification system. Among them, the most representative is the deep learning method that directly learns complex feature expressions from massive data. This article first reviews the research background and development history of image classification, then analyzes the application of image classification in different fields and lists case studies. It also outlines traditional image classification methods and mainstream image classification techniques in deep learning, including convolutional neural networks, model pre-training, and data augmentation. At the end of the article, the main challenges currently faced by image classification were analyzed, and future directions for improvement and technologies that can be combined were discussed. This review aims to provide researchers in image classification with certain research directions.

Keywords: Image classification, Deep learning, Convolutional Neural Network.

1. Introduction

In today's era of rapid technological development, image classification holds a pivotal position in the field of computer vision. It is like a magical key that opens the door to our effective understanding and management of massive image information. Yang et al. proposed a graph representation knowledge distillation method and tested it on 100 classification datasets published by the Canadian Institute for Advanced Research, greatly improving the ability of image classification [1].

Images, as an intuitive and rich information carrier, are everywhere in our lives. From the photos, the paper takes daily with our smartphones to the medical images in the healthcare field, from the license plate and traffic sign images in intelligent transportation systems to the numerous product images on e-commerce platforms, and then to the product inspection images in industrial production, images are omnipresent, and their information value is immeasurable [2,3]. The core goal of image classification is to automatically classify these complex images into pre-defined categories, thereby realizing the efficient organization and utilization of image information.

In the past, classifying images was an extremely arduous task that often required a large amount of human participation. Professionals need to rely on their experience and knowledge to judge the category of an image by observing its color, texture, shape, and other characteristics. However, this method is not

DOI: 10.54254/2755-2721/80/2024CH0068

only inefficient but also difficult to ensure accuracy. With the continuous progress of information technology, especially the rise of deep learning technology, image classification has ushered in a major transformation. Reference [4] proposed the combination of formal concept analysis and convolutional neural networks to enhance the interpretability of image classification, with high accuracy.

The second chapter of this article first introduces the application of image classification in different fields such as healthcare, security, and e-commerce. Then, the third chapter introduces some mainstream technologies of traditional learning and deep learning. The fourth chapter introduces the implementation of image classification methods. The fifth and sixth chapters respectively introduce the current challenges and future prospects and finally make a summary.

2. Application fields of image classification

2.1. Healthcare

Image classification technology assists doctors in disease diagnosis in the medical field. By analyzing medical images such as X-rays and CT scans, abnormal conditions such as tumors and fractures can be detected. This greatly improves the accuracy and efficiency of diagnosis and gains precious time for patient treatment. Deep learning performs well in medical image classification. This paper proposes a semi supervised framework for multi label medical image classification. The framework considers the distribution differences between labeled and unlabeled data, and introduces domain discriminators and feature extractors for adversarial learning, allowing the model to learn feature representations while maintaining consistency in the feature space. This method significantly improves the overall performance of multi label medical image classification [5].

2.2. Intelligent transportation

In intelligent transportation systems, image classification realizes license plate recognition and traffic sign classification. This is crucial for the development of autonomous driving technology. At the same time, it also improves the level of traffic management and ensures road traffic safety. Reference [6] conducted research on the hierarchical detection of crowd congestion inside buses based on convolutional neural network methods, and achieved excellent detection results for head targets of personnel. The results were used to calculate the number of people and crowd density, helping to detect traffic flow and evacuate traffic flow.

2.3. E-commerce

Classify product images to facilitate user search and recommend related products. While improving the user shopping experience, it also improves the operating efficiency of e-commerce platforms. Shen et al. proposed a product image classification method based on multi feature fusion according to the characteristics of product image data. At the same time, rich text content is extracted from the image, and contextual information is extracted through a language model to obtain text features. Utilizing the image feature extraction capability of convolutional neural networks to extract deep image features of product images, and exploring the internal relationship between the two features for classification, a WeChat mini program was developed to facilitate merchant product queries [7].

2.4. Industrial production

In industrial production, image classification can be used for detecting product defects and conducting quality control. Timely detection of product quality problems reduces production costs and improves production efficiency. Hu et al. proposed an automatic classification method for aviation fasteners based on image classification algorithms, designed a set of fastener image acquisition and automatic classification implementation schemes, and conducted evaluation experiments based on real industrial data. The evaluation experiments statistically analyzed the accuracy, recall, precision, and F1 score indicators of convolutional neural networks and Inception-v3 models; The experimental results show

DOI: 10.54254/2755-2721/80/2024CH0068

that the accuracy of the model is over 98%, which can effectively classify aviation firmware products [8].

3. Related technologies of image classification

3.1. Traditional methods

Based on manually designed features such as color, texture, and shape, combined with machine learning algorithms for classification. However, this method relies on artificial experience for feature design and has limited classification performance.

3.2. Deep learning methods

3.2.1. Convolutional Neural Network (CNN). Currently the mainstream method for image classification, it can automatically learn the features of images. It has strong feature extraction ability and classification performance and performs outstandingly in various image classification tasks. Reference [9] selected the VGGNet model for structural construction and model training through in-depth analysis of the performance characteristics of various convolutional neural network models and the practical work requirements of automotive product detection. After multiple meticulous parameter adjustments and optimizations, the accuracy and F1 score of the final model reached 95%.

3.2.2. Pretrained model. Using models trained on large-scale datasets, such as pretrained models on ImageNet. Fine-tuning on specific tasks can improve performance and reduce training time and data requirements. As reported in reference [10], pretrained models can be beneficial for image classification tasks.

3.2.3. Data augmentation. By performing random rotations, cropping, scaling and other operations on the original image, the diversity of data is increased and overfitting is reduced. As discussed in reference [11], data augmentation is an important technique in image classification.

4. Methods and implementation of image classification

4.1. Data collection and preprocessing

Collect a large amount of image data and perform cleaning, labeling and preprocessing to meet the model training requirements. Ensure data quality and accuracy and lay the foundation for subsequent training.

4.2. Model selection and training

Select the appropriate model architecture according to task requirements and use training data for training. Continuously adjust model parameters, optimize performance, and improve classification accuracy.

4.3. Model evaluation and optimization

Use the validation set to evaluate the trained model and optimize according to the results. For example, adjust hyperparameters and increase the amount of data to improve model performance and generalization ability.

4.4. Deployment and application

Deploy the trained model to practical applications to realize the image classification function. Ensure the stability and reliability of the model in practical applications.

DOI: 10.54254/2755-2721/80/2024CH0068

5. Challenges of image classification

5.1. Difficult data labeling

Data labeling in some fields is costly and difficult. It consumes a lot of manpower, material resources and time, limiting the application of image classification technology in these fields.

5.2. Model interpretability

The decision-making process of deep learning models is difficult to understand and lacks interpretability. There are certain risks when applied in key fields.

5.3. Adversarial attacks

Models are vulnerable to the influence of adversarial attacks, resulting in classification errors. Threatening the security and reliability of image classification.

6. Prospects of image classification

6.1. Multimodal fusion

Combine multimodal information such as images and text to improve classification accuracy and robustness. Give full play to the advantages of different modalities and provide richer information for image classification.

6.2. Self-supervised learning

Use self-supervised learning methods to reduce the dependence on a large amount of labeled data. Let the model automatically learn image features and improve generalization ability.

6.3. Model compression and acceleration

To meet the needs of practical applications, research model compression and acceleration technologies. Reduce computational complexity and storage requirements and improve operating efficiency.

6.4. Cross-domain learning

Solve the problem of data distribution differences in different fields and realize the generalization ability of the model. Enable the model to adapt to image classification tasks in different fields.

7. Conclusion

As an important task of computer vision, image classification has broad application prospects in many fields. Although facing some challenges, with the continuous development of technology, image classification will achieve more excellent results and bring more convenience to people's lives and work. In the future, technologies such as multimodal fusion, self-supervised learning, model compression and acceleration, and cross-domain learning will promote the continuous innovation and progress of image classification technology.

References

[1] Yang, C. G., Chen, L. M., Zhao, E. H. et al. (2024). Image classification method based on graph representation knowledge distillation. electronic journal, 1-13.

[2] Lai, Q., Vong, M. C., Yan, T., et al. (2024). Hybrid multiple instance learning network for weakly supervised medical image classification and localization. Expert Systems With Applications, 260125362-125362.

[3] Xiong, M., Luo, W. J., Ren. J., et al. (2024). Applying Deep Learning with Convolutional Neural Networks to Laryngoscopic Imaging for Automated Segmentation and Classification of Vocal Cord Leukoplakia.Ear, nose, & throat journal, 1455613241275341.

DOI: 10.54254/2755-2721/80/2024CH0068

[4] Khatri, M., Yin, Y., Deogun, J. (2024). Enhancing Interpretability in Medical Image Classification by Integrating Formal Concept Analysis with Convolutional Neural Networks. Biomimetics, 9(7): 421-421.
[5] Zhang, K., Liang, W., Cao, P., et al. (2024). Semi supervised multi label medical image classification model guided by label association. Small microcomputer system,1-9, http://kns.cnki.net/kcms/detail/21.1106.TP.20240927.0932.002.html.
[6] Zi, G. P. (2024). Research on Crowding Grading Detection Method in Intelligent Buses. Jilin University. DOI:10.27162/d.cnki.gjlin.2024.006982.
[7] Shen, Q. (2023). Research on Product Image Classification Method Based on Deep Learning. Tianjin Normal University. DOI:10.27363/d.cnki.gtsfu.2023.000359.
[8] Che, L. Z. (2023). Research on Key Technologies for Facial Privacy Protection and Anti tampering of Security Surveillance Images. Wuhan University of Science and Technology. DOI:10.27380/d.cnki.gwkju.2023.000386.
[9] Liu, D. W. (2024). Wang, X., Miu, J. M., Research on Optimization of Automotive Product Inspection Based on Convolutional Neural Networks. Commercial Vehicles. (02): 82-87.DOI:10.20042/j.cnki.1009-4903.2024.02.006.
[10] Donahue, J., Jia, Y., Vinyals, O., et al. (2013). DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. CoRR, abs/1310.1531
[11] Li, X. F., Gou, G. L., Han, Y. Q. et al. (2024). Small sample fine-grained image classification method guided by multi-scale knowledge for local enhancement. Journal of Intelligent Systems, 1-11. http://kns.cnki.net/kcms/detail/23.1538.TP.20240828.1351.030.html.