ITM Web of Conferences **70**, 04036 (2025) https://doi.org/10.1051/itmconf/20257004036
DAI 2024

Research on the Application of Deep Learning Methods in the Field of Image Classification

Yuhui Peng*

School of Computer Science and Technology, Hainan University, Hainan, 570000, China

Abstract. With the rapid development of image classification technology, it has become a current research hotspot to apply image classification technology to various fields and to improve the accuracy and efficiency of image classification technology in various fields. In the field of fruit classification and textile, the application of image classification technology has been widely concerned. This paper reviews the current research status of image classification models, focusing on the application of DenseNet-201, Xception, MobileNetV3-Small and ResNet-50 models in the fruit field. The application of deep learning methods such as Convolutional Neural Network, Recurrent Neural Networks and Long Short-Term Memory in image classification is also discussed. In this paper, it is concluded that these models have achieved high accuracy in fruit classification and the textile field, especially the combination of CNN, RNN and LSTM deep learning methods for feature fusion can enhance the accuracy and robustness of the model. In addition, this paper also discusses the limitations of the current research and makes some suggestion.

1 Introduction

With the rapid progress of image classification technology, it plays an extremely important role in various fields. In numerous image classification tasks, such as the development of automatic supermarkets, the application of image classification has significantly improved the efficiency of automatic recognition of goods. By capturing the image of the product through the camera, and then using the image classification technology for identification and classification, it can realize fast and accurate commodity identification and settlement, which considerably improves the shopping experience and the efficiency of the supermarket operation. In industrial quality inspection, smart agriculture, intelligent security and other scenarios, image classification technology is also playing an increasingly crucial role. It can help us identify and locate various problems more accurately, improve production efficiency and product quality, and also provide new possibilities for smart agriculture to help us realize precision agriculture and intelligent management. In recent years, the progress of deep learning technology has brought new possibilities to image classification tasks, greatly improving the accuracy and efficiency of image classification.

* Corresponding author: 20203105078@hainanu.edu.cn

Among the deep learning models, the Convolutional Neural Network (CNN) has achieved remarkable results. When dealing with image classification tasks, CNN effectively extracts image features and realizes spatial convolution operation through the convolutional layer, Pooling layer and Fully connected layer. CNN has high accuracy and efficiency in processing image data, and has been widely used in image classification, object detection, semantic segmentation and other tasks. In addition to CNN, there are also many deep learning methods, including Recurrent Neural Networks (RNNS) and Long Short-Term Memory Networks (LSTMS) have also achieved certain results in the field of image classification. RNN is suitable for processing sequence data, such as speech, text, and time series, and it can classify input data by capturing long-term dependencies in sequence data. LSTM is an improved version of RNN, which can better capture the long-term dependencies in the sequence and avoid the explosion and disappearance of gradients.

In the field of fruit classification, DenseNet-201, Xception, MobileNetV3-Small, ResNet-50 and other models have achieved high accuracy. These models perform well in fruit classification tasks, which provides strong support for the development and application of image classification technology in practical application fields. In addition, in the study of Gang Xue et al., a fruit image classification method based on hybrid deep learning is used, which uses convolutional autoencoder to pre trained images and attention-based DenseNet to extract features of images, which improves the performance of CNN model in fruit classification problems [1]. Roshani Raut et al. proposed a machine learning-based fruit classification framework, which uses the pre trained Efficient-Net model to achieve excellent results by fine-tuning the epoch parameter and learning rate parameter [2].

In the field of textiles, Probabilistic Neural Network (PNN) has also shown certain classification abilities. Although there are still limitations in the recognition of some categories, PNN has high performance in dealing with uncertainty and randomness, which provides a new possibility for the application of image classification technology in the textile field. Rehan Ashraf et al. proposed a GoogleNet image processing technology based on cnn to judge fabric defects by identifying the appearance shape, size and position of defects on the fabric, which is better than Bayes, SVM and other methods [3].

In this paper, the above image classification models will be introduced in detail, and their applications in fruit classification and textile fields will be analyzed. At the same time, this paper will discuss the possible problems and limitations of these models in practical applications, and propose suggestions for further research discussion. Through the research of this paper, the paper expects to play a positive role in promoting the development and application of image classification technology.

2 An Introduction to Image Classification Models and Algorithms

This article will mainly introduce several common image classification models from the perspectives of Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM). First, this article will discuss the application of deep learning models such as CNN, RNN and LSTM in image classification, as well as their advantages and limitations when dealing with different types of data. Then, Convolutional Neural Network models such as DenseNet-201, Xception, MobileNetV3-Small, and ResNet-50 will be introduced, which have shown promising performance in image classification tasks. Commonly used datasets in Fruit domain applications include Fruits-360 and Fruit Recognition dataset, where the Fruits-360 dataset is a large-scale fruit recognition dataset. The dataset contains 360 different kinds of fruits covering various colors, shapes, and sizes. The Fruits-360 dataset has a total of 91,120 color images, 70% of which are used for training and 30% for testing. Each image has a corresponding label to be used for classification tasks. The Fruit Recognition dataset is a smaller dataset in fruit recognition, which contains 10 categories of fruits with a total of 650 color images. The main feature of this dataset is that the image quality is high, and each image has been strictly screened to ensure the clarity of the fruit and the accuracy of recognition. The fruit images in the dataset cover different angles, lighting conditions, and background environments to simulate the recognition challenges in real-world scenes.

As a deep learning model, CNN is mainly applied to image classification, object detection and semantic segmentation. Through structures such as the Convolutional layer, Pooling layer and Fully connected layer, it effectively extracts local features and spatial information of the image, and performs spatial convolution operations between neurons [4]. CNN has high accuracy and efficiency in processing image data.

As a traditional deep learning model, RNN has superior performance in processing sequence data such as speech, text and time series [5]. It contains a hidden layer to capture long-term dependencies in sequence data. RNN predicts the output at the current time by updating the hidden state at the previous time in the sequence one by one. However, RNN is prone to problems such as gradient disappearance and gradient explosion when dealing with long sequences, which makes the model difficult to train.

LSTM is an improved RNN model, which is used to solve the problem of gradient disappearance and gradient explosion when RNN deals with long sequences [6]. By introducing a gate mechanism and a unit structure, LSTM can effectively capture the long-term dependencies in the sequence and avoid the explosion and disappearance of gradients. LSTM has good performance in tasks such as speech recognition, natural language processing and time series prediction.

DenseNet-201 is a Dense Convolutional Network with dense connections and feature maps generated layer by layer. By connecting the output of each layer with the input of all previous layers, it forms a dense feature map, which reduces the information loss in the feature transfer process and improves the efficiency of feature transfer and reuse.

Xception is a Convolutional Neural Network designed to improve the accuracy of image classification. It uses a technique called Xception Pooling to aggregate the features of different samples of the same class, resulting in a richer and more accurate feature representation. In Xception, each Convolutional layer is followed by an Xception Pooling layer and a ReLU activation function to aggregate the features of different samples of the same class and suppress the features related to other classes. This technique can effectively reduce overfitting and improve classification accuracy.

The MobileNetV3-Small model is a lightweight Convolutional Neural Network model suitable for resource-constrained environments such as mobile devices and embedded systems. By using techniques such as depthwise separable convolutions, locality-sensitive hashing, and skip connections, it reduces the complexity and the number of parameters of the model while maintaining good performance. The MobileNetV3- Small model is more suitable for resource-constrained environments by briefly explaining the use of lightweight techniques such as depthwise separable convolution.

The ResNet-50 model is a variant of Deep Residual Network (ResNet), which has a deep network structure and excellent performance. It solves the problems of gradient disappearance and gradient explosion in deep networks by using techniques such as Residual Block and Residual Path, and improves the training stability and performance of the network. In ResNet-50, the deeper network is decomposed into smaller residual blocks, which makes the network easier to train and adapt to various tasks.

Probabilistic Neural Network (PNN) is a type of neural network based on a probabilistic model, it has high performance in dealing with uncertainty and randomness problems. Compared with traditional neural networks, PNN introduces the concepts of probability distribution and conditional probability, which can better describe the probabilistic characteristics of input data [7]. The main components of PNN include the Input layer, Hidden layer and Output layer. During training, the network learns probability distributions and conditional probabilities by maximizing the likelihood estimation. Specifically, the network uses the maximum likelihood estimation method to adjust the network parameters, so that the network output probability distribution matches the actual data distribution, and can deal with uncertainty and randomness.

3 Application of Image Classification Algorithms

3.1 In the field of fruit

In the field of fruit classification, image classification and deep learning algorithms can improve classification accuracy. Farsana Salim et al. show that in the test of Fruit-360 dataset, when the training test image size is specified to be less than the default value size, the accuracy of DenseNet-201, MobileNetV3-Small, and ResNet-50 are 97.33, 95.65, and 98.36, respectively, while the accuracy of Xception is only 84.34, which indicates that when dealing with small-size images, the accuracy of DenseNet-201, MobileNetV3-small, and ResNet-50 are 97.33, 95. 65, and 98.36, respectively [8]. DenseNet-201, MobileNetV3-Small, and ResNet-50 are more suitable for object detection than Xception. When the default image size of each model is used for training and testing, the accuracy of DenseNet-201, Xception, and MobileNetV3-Small is 99.87, 98.94, and 99.93, respectively, maintaining a high level, and the accuracy of ResNet-50 is 100 [8]. This shows that each model performs best when using the appropriate image size, and ResNet-50 is even more error-free. The test results on the Fruit Recognition dataset show that the accuracy of DenseNet201, Xception, MobileNetV3-Small and ResNet-50 is 99.13, 97.73, 62.73 and 76.47, respectively [8]. The accuracy of DenseNet-201 and Xception is much higher than that of MobileNetV3-Small and ResNet-50, which indicates that when more realistic factors such as lighting and shadow are added to the image, DenseNet-201 and Xception are more suitable for detection. In general, when the test image is larger, there will be more important feature information in the image, and the accuracy of the model after training is greatly improved. The accuracy of DenseNet-201 and Xception can be maintained at a high level after training by adding factors such as lighting conditions and background environment in the image, and DenseNet-201 and Xception are more suitable for more realistic image detection.

In the model mentioned by Harmandeep Singh Gill et al., deep learning methods (CNN, RNN and LSTM) are combined for use. CNN is used for image feature extraction, and RNN is used for feature labeling and combination to solve the problem that CNN cannot separate the unique fruit features [9]. The LSTM method is used for fruit classification to solve the gradient and descent problems encountered by RNNs in classifying different features. The images after feature fusion are tested. The experiments show that under the target of detecting only varieties (such as guava and orange), the overlap degree between the image and the actual situation is 0.25, 0.5, or 0.75. The accuracy of the model detection is about 0.85 [9]. When distinguishing red, green, and golden apples, the test accuracy is higher when the overlap is 0.25 and 0.75 than when the overlap is 0.5, and the highest is when the overlap is 0.75, which indicates that the more real the test image is, the higher the accuracy of the model in distinguishing different color apples [9]. In general, the model can improve the accuracy of the model by combining CNN, RNN, and LSTM methods for feature fusion, and it is also more suitable for distinguishing apples of different colors in more realistic images.

https://doi.org/10.1051/itmconf/20257004036

3.2 In the field of textile

Similarly, in the textile field, deep learning methods and image classification algorithms can improve the classification efficiency of textile color, type, etc. Rocco Furferi et al. mentioned that by using a PNN network to associate training set elements with target elements, the network can take as input any vector of four elements (consisting of a normalized number) pixels and then output what is defined as the most likely class or even two or more classes that are most likely to be classified [10]. Experiments show that when distinguishing between brown, red, blue, green, gray, and black, the reliability of brown, red, blue, and green is higher, while when distinguishing gray and black, the model shows a significant deficiency and the reliability is lower [10]. Compared with the Computer Vision-Based Method, only gray and black have low reliability in the experiments of distinguishing different colors. In a word, this method has low-cost requirements and has a certain degree of usability for scenarios with low-complexity datasets.

4 Suggestion

Combined with the above-mentioned, Convolutional Neural Networks (CNN) has achieved remarkable results in image classification tasks. DenseNet-201, Xception, MobileNetV3-Small, ResNet-50 and other models have achieved high accuracy in the field of fruit classification. In the field of textile, Probabilistic Neural Network (PNN) has also shown certain classification ability.

For the application of image classification in different fields, the suggestions of this paper are as follows. Firstly, to improve the performance of the model in a specific domain, it is recommended to use a dedicated dataset for that domain for training and testing. Secondly, according to the task requirements and computing resources, the appropriate model is selected. Then, in the resource-constrained environment, the lightweight MobileNetV3-Small model can be selected. And in tasks with high performance requirements, models with superior performance such as ResNet-50 can be selected. Then, combining the advantages of different models can improve the performance of the model in a specific domain. For example, in fruit classification tasks, the combination of CNN, RNN, and LSTM methods for feature fusion can improve the accuracy of the model. Finally, the model is tuned to improve the performance for specific tasks and domains.

The limitations are also obvious. In terms of datasets, deep learning models usually require a large amount of training data, and obtaining high-quality data may be difficult in some domains. In terms of complexity, the overfitting problem can be more severe as the model complexity increases. To solve this problem, techniques such as regularization and Dropout can be tried. In terms of computational resources, deep learning models may require high computational resources during training and prediction, which may limit their application in some fields. In terms of model interpretability, deep learning models are usually difficult to interpret, which may affect their practical application in some fields.

5 Conclusion

This paper mainly introduces the application of several commonly used image classification models in the field of image classification, including DenseNet-201, Xception, MobileNetV3-Small, and ResNet-50. In addition, this paper also discusses the application of deep learning methods such as Convolutional Neural Networks (CNN), Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) in image classification, and evaluates the performance of these models through experimental results of fruit classification and textile field.

This paper concludes that the four models DenseNet-201, Xception, MobileNetV3-Small and ResNet-50 have high accuracy in general fruit classification tasks, especially when using images of appropriate size for training and testing. However, in the fruit recognition test with more realistic images, only DenseNet-201 and Xception maintain high accuracy. In addition, the method of combining CNN, RNN and LSTM for feature fusion can improve the accuracy and robustness of the model. In the textile domain, Probabilistic Neural Network (PNN) has also shown some classification ability, although there are still limitations in the recognition of some categories.

Despite the progress made in the current research, there are still some aspects that can be further investigated. Firstly, more deep learning models and algorithms can be explored to improve the accuracy and efficiency of image classification. Secondly, knowledge from other fields, such as natural language processing and image generation, can be combined to further improve image classification models. In addition, new data augmentation techniques and model interpretation methods can be explored to enhance the robustness and interpretability of the model. Finally, image classification models can be applied to a wider range of domains, such as medical image recognition, intelligent transportation, and UAV navigation.

References

1. G.Xue, S. Liu, & Y.Ma.Complex Intell. Syst. 9, 2209–2219 (2023).
2. R. Raut, A. Jadhav, C. Sorte and A. Chaudhari, "Classification of Fruits using Convolutional Neural Networks," 2022 Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), Bhilai, India, 2022, pp. 1-4
3. R.Ashraf, Y.Ijaz, M.Asif. Mathematical Problems in Engineering,2573805, 16, (2022).
4. S.Y. Huang, W.J. An, D.S. Zhang, N.R. Zhou, Optics Communications, 533, 129287, (2023).
5. J. Guo, Q. Zhang, Y. Zhao, H. Shi, Y. Jiang and J. Sun, "RNN-Test: Towards Adversarial Testing for Recurrent Neural Network Systems," in IEEE Transactions on Software Engineering, vol. 48, no. 10, pp. 4167-4180, 1 Oct. (2022).
6. S. Bian, H. Li, C. Wang, C. Song and Y. Tang, "MSBF-LSTM: Most-significant Bit-first LSTM Accelerators with Energy Efficiency Optimisations," 2023 IEEE 31st Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM), Marina Del Rey, CA, USA,218-218, (2023)
7. S. Guan, Q. Fang and T. Guan, "Application of a Novel PNN Evaluation Algorithm to a Greenhouse Monitoring System," in IEEE Transactions on Instrumentation and Measurement, 70, 1-12, (2021).
8. F. Salim, F. Saeed, S. Basurra. Electronics 12, 3132(2023).
9. H. S. Gill, G. Murugesan, A. Mehbodniya, G. S. Sajja, G. Gupta, A. Bhatt, Computers and Electronics in Agriculture,211,107990, (2023).
10. R. Furferi, M. Servi, Appl. Sci. 13, 2464(2023).