Tobacco Leaves Classification Based on Convolutional Neural Network

Wrriten by XMU Students
Weiyuan Huang*36920221153084, Zongyu Lan*36920221153088
Chuting Lin*30920221154276, Hao Li*36920221153092, Xiang Li*36920221153094
[1] Electronic Information Major, Department of Software Engineering, School of Information, Xiamen University
[1] Electronic Information Major, Artificial Intelligence Research Institute, Xiamen University

Abstract

China is the largest tobacco producing country and the largest tobacco consuming country in the world. As a major pillar of China's economy, the tobacco industry continues to provide tax revenue and employment, and its importance is self-evident. In the tobacco industry, tobacco grading is the most important, which is directly related to the quality of tobacco products and the economic benefits of tobacco farmers. At present, the most important tobacco leaf grading method in China's tobacco industry is manual sorting and grading through visual observation. This method is seriously affected by factors such as light, grader's proficiency, vision and even their emotional state. The accuracy is relatively unstable, and there are also shortcomings of high cost and low efficiency. In the context of economic globalization and the rapid development of the tobacco industry, the tobacco industry urgently needs a more efficient tobacco grading method. In recent years, with the rapid development of artificial intelligence, especially in the field of deep learning, the tobacco grading problem can be seen as a slightly complex image classification problem. We are trying to use trained convolutional neural networks to classify tobacco leaves in our data set and explore the effect of CNN's application on tobacco leaves grading.

Introduction

In the tobacco industry, tobacco grading is the most important part of the work. The tobacco leaves sent by the tobacco farmers to the purchase station will be graded by the grading personnel, and the purchase station will pay the tobacco farmers the corresponding purchase money according to the rating of these tobacco leaves. After that, these tobacco leaves will be made into different grades of cigarettes according to their respective grades and sold on the market. From this, we can see that the effect and rationality of tobacco grading not only determine the final economic benefits of farmers, but also play an important role in cigarette market prices.

Judging from the current tobacco leaf grading method adopted in China, it mainly depends on manpower. Specialized collectors grade the dried tobacco leaves according to the size, texture, shape, maturity, oil content and other indicators. This traditional method has the following pain points:

- The manpower grading is highly subjective, and the grading personnel basically make judgments based on experience. The accuracy and fairness of tobacco grading cannot be guaranteed in the absence of subsequent verification means. The grading standard is vague, and it is difficult to record orally as a guide for subsequent personnel grading, with poor reusability.
- The manpower grading is greatly restricted by objective factors, and the grading quality is unstable. The quality of tobacco grading is easily affected by the proficiency, vision and even emotion of grading personnel. In addition, light conditions also have a greater impact on tobacco grading. Therefore, there is still a lot of room to improve the stability of tobacco grading accuracy.
- The cost of manpower grading is high and the efficiency is low. Manpower grading has high requirements for the proficiency and professional knowledge of grading personnel, and training a grading personnel requires high time costs and money costs. However, manual tobacco grading is generally slow and inefficient.

In the context of today's economic globalization and the continuous improvement of the modernization level of various industries, the disadvantages of this traditional grading method, such as unstable effect, low qualification rate, high cost and slow speed, are becoming increasingly obvious. Obviously, the tobacco industry urgently needs a new method with high efficiency and high accuracy to grade tobacco.

In the context of today's economic globalization and the continuous improvement of the modernization level of various industries, the disadvantages of this traditional grading method, such as unstable effect, low qualification rate, high cost and slow speed, are becoming increasingly obvious. Obviously, the tobacco industry urgently needs a new method with high efficiency and high accuracy to grade tobacco.

In recent years, with the continuous development of artificial intelligence, people have made breakthroughs in computer vision, among which convolutional neural network is undoubtedly the most eye-catching one. It has achieved excellent results in target detection, image recognition and image classification, and has played a huge role in people's lives. It can automatically extract the features in the image, which is particularly suitable for image classification. Tobacco grading can be regarded as an image classification task. We naturally thought of using convolutional neural network to grade tobacco.

It is of great significance to conduct research on the effect of convolution neural network on tobacco grading. Compared with manpower, neural network can complete grading faster and more efficiently without fatigue. Once the research is successful, it will be of great help to the grading work of each purchase station and have a positive impact on China's tobacco industry. Although the effect of specific application is not yet clear, and human grading is still essential, the grading results given by neural network also have great reference value. The work of grading personnel can change from sorting all tobacco leaves to grading wrongly classified tobacco leaves in the grading results, and the workload is greatly reduced.

Related Work

At present, the most common tobacco grading method is manual grading. In recent years, with the development of artificial intelligence machine learning theory, many algorithms of machine learning for tobacco grading have been proposed.

In 2015, Yang Xiaojuan used image processing technology combined with an improved nearest neighbor classification method to achieve intelligent grading of tobacco leaves(Xiaojuan 2015). She used the nearest neighbor based backward fea-ture selection method to select the optimal feature set, and proposed an improved nearest neighbor classification method. The neighbor classification method improves the classification criteria and effectively improves the accu-racy of tobacco leaf classification.

In the same year, Wei Yangfan proposed a color fea-ture-based tobacco leaf classification algorithm(Yangfan 2014), which uses a uniform model as the standard for learning tobacco leaves of known grades. Tobacco leaves are graded ac-cording to the obtained distribution of color characteristics.

In 2017, Zhuang Zhenzhen proposed a tobacco leaf grading algorithm based on fuzzy pattern recognition(Zhenzhen 2016), de-veloped an automatic tobacco leaf classification software system, and obtained 93.02% grouping accuracy and 80.23% grading accuracy, reaching or even exceeding the level of manual grading.

At the same time, with the development of deep learn-ing, there are many studies based on neural networks.

In 2015, Guo Qiang proposed an improved algorithm to filter the tobacco leaf image(Qiang 2013), obtained the contour infor-mation of the tobacco leaf through the iterative threshold method and the contour extraction method, extracted the main characteristic parameters of the tobacco leaf, and screened out the appropriate tobacco leaf characteristics according to the tobacco leaf grading standard. Based on computer grading judgment, BP neural network is used to grade tobacco leaf images, and a large tobacco leaf grad-ing system is realized.

In 2017, Y. Sari and R. A. Pramunendar used a back-propagation neural network to classify tobacco leaves(Sari and Pramunendar 2017), which eventually achieved an accuracy of 77.50%.

In 2019, Luo Huixin proposed a tobacco leaf grading that combines online transfer learning and traditional feature extraction methods(Huixin 2019). Algorithm, in view of the prob-lem of few training samples and single category of tobac-co leaves, based on the AlexNet(Krizhevsky, Sutskever, and Hinton 2012) model to optimize pa-rameters to prevent over-fitting; delete the LRN layer and use the BN layer to optimize the network structure, and propose an active incremental learning method instead. The method optimizes the number of samples through mining in the complex sample, and combines the method of online migration. Experiments show that the effect of this fusion method is due to the traditional machine learn-ing tobacco leaf grading algorithm.

In the same year, Charlie S. Marzan proposed an auto-matic tobacco leaf grading algorithm based on image pro-cessing technology and convolutional neural network(Marzan and Ruiz 2019), which includes five parts: image acquisition, image pre-processing, tobacco leaf detection, tobacco leaf segmenta-tion, and tobacco leaf classification. The image processing technology separates 100% of the tobacco leaves from the images. Using a dataset that contains both segmented and unsegmented images, it takes an average of 7.43ms to classify a single tobacco leaf, and finally achieves an ac-curacy of 96.25%.

In 2020, Wang Shixin proposed a method based on convolutional neural network of -inception-V3(Shixin 2020). Instead of the fully connected layer , Wang used a global average pooling layer. The three-dimensional tobacco leaf data feature space can be mapped to the one-dimensional fea-ture space, which greatly reduces the tobacco leaf data feature parameters, improves the calculation efficiency, increases the stability of the model, and achieves an accu-racy rate as high as 97.21%.

In the same year, Siva Krishna Dasari proposed a tobac-co leaf grading algorithm based on a deep convolutional network(Dasari, Chintada, and Patruni 2018). Except for the input layer and the output layer, the neural network consists of three convolutional layers, three pooling layers and a fully connected layer. A dataset consisting of 120 tobacco leaf images was trained and fi-nally got 85.10% accuracy. These methods have achieved high accuracy, but there are also problems such as requiring large-scale data sets, long training periods, and too much training parameters. It is certain that the neural network can better handle the problem of tobacco leaf grading, and related research has the possibility of transformation and implementation.

Experiment

This project intends to explore and compare the effectiveness of AlexNet, ResNet(He et al. 2015), and ResNeXt(Xie et al. 2016) on tobacco picture classification based on a dataset of near-real scenes of tobacco leaves from relevant departments of tobacco industry. Based on the characteristics of the tobacco image pairs in the dataset, we propose a method to merge the three-channel RGB image tensor obtained from each tobacco leaf collected under two different conditions into a six-channel RGB image tensor input into a neural network for feature extraction, and compare it with a

Table 1: Hyperparameters

epochs	batch size	learning rate	weight decay	Momentum	dropout rate
200	32	0.001	0.0005	0.9	0.5

common network training method to explore a better method for tobacco leaf classification.

Dataset introduction

We used a small tobacco leaf dataset provided by the tobacco industry related department, which has a total of 1486 tobacco leaf image pairs (each image pair contains two images of the same tobacco leaf taken under two different lighting conditions), for a total of 2972 tobacco leaf images. The two images in Fig. 1 are identical in position and size, except for the different lighting conditions. There are differences in size between the different image pairs. There are 14 folders within the dataset, and each category folder contains several image pair folders. The differences between similar categories of tobacco within the dataset are very small and difficult to distinguish successfully even for professional tobacco sorters. Therefore, our task is actually image fine-grained recognition, which is a challenge for us.

Figure 1: Example of three similar categories.

Dataset processing

We first performed a simple denoising process on the dataset, a step that did not take us much time due to the high quality of the tobacco images in the dataset, and cropped the images with a relatively large percentage of blanks to make the tobacco part occupy as much image space as possible to help the network learn the tobacco features better. Since the image pairs are stored in folders, which is inconvenient to read, we wrote a python script to consolidate all the images in the folders into category folders, making sure that the category folders are composed of all images, and naming the images in each image pair so that they are organized in pairs. (Two images of the same tobacco leaf must be next to each other.)

For example, ”31.jpg” means that this image is the first image of the third tobacco leaf in this category. After that, we divided the training set and the test set, with 80% of the images as the training set with 2364 images and 20% of the images as the test set with 608 images, without a validation set. During the partitioning, we performed careful processing to ensure that the two adjacent images were moved together. After that, we traversed all the category folders by os.listdir, counted the number of images in each category and recorded the corresponding category labels of all images, which were stored in the list. Due to the small number of images in the dataset and the uneven number of images in each category, we wrote scripts to enhance the data, including horizontal or vertical flipping with 0.5 probability, mirror flipping, etc. For the sample imbalance problem caused by the small number of images in category C1 (only 30 images, while the other categories have at least 120 images), we used Gaussian blurring in addition to flipping to expand the number to 120 images.

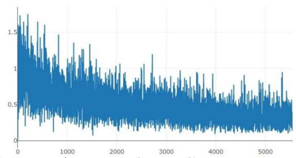
Implementation Details

We built models such as ResNet and wrote the training code by combining online materials and tutorials. The training was performed on an NVIDIA RTX 2080 on a cloud server. Each model was loaded with pre-trained weights provided on the official PyTorch website. Loading these weights trained on large datasets (e.g. ImageNet) can be a good way to compensate for the problem of too few images in our dataset. For each model, we used the same hyperparameters, as shown in Table.1(Srivastava et al. 2014).

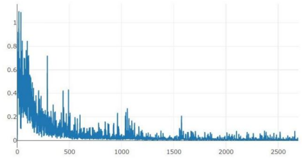
For the optimizer, we use the Adam optimizer with the stepLR stage-wise learning rate decay strategy. In the later stage when the model was difficult to converge, we switched to the SGD optimizer(Loshchilov and Hutter 2016) with momentum, together with the cosine annealing learning rate decay strategy with Warm Restarts to rapidly reduce the loss and accelerate the convergence of the model, and subsequent experiments also proved that such a combination was effective. We use Visdom to plot the variation curves of loss and accuracy on the training and validation sets. Each time when the curve has no more tendency to change, we end the training and save the trained model, and next time we will con-

Table 2: Model Accuracy

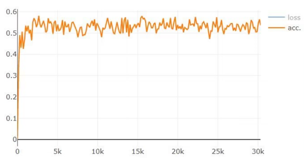
model	accuracy%
AlexNet	58.0%
ResNet-18	64.8%
ResNet-34	65.0%
ResNet-50	69.7%
ResNeXt-50	70.1%



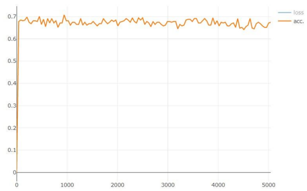
(a) Using Adam+StepLR, no convergence



(b) Convergence after using SGD + cosine annealing



(c) The effect of ResNeXt-50 using 3-channel image training



(d) The effect of ResNeXt-50 using 6-channel image training

Figure 2: Some curves of the process

tinue the training by modifying the hyperparameters based on the previous ones.

Our experiment was divided into three parts:

1. All 3-channel images are used for training, and each image is a separate training sample.

2. halve the dataset and use only the brighter half of the images for training and testing (the images in this condition have more distinct features).

3. Combine the two 3-channel images of each tobacco leaf into one 6-channel image as a training sample into the neural network for training.

All images are scaled to a size of 256×256 and center cropped to a size of 224×224 before being fed into the network, which is a common processing method.

Part I

Since the two images of the same tobacco leaf are treated as different samples in this method, the network does not work well to learn the features in both high light conditions and shaded conditions. The trained AlexNet model could only achieve 25% accuracy on the validation set, and the ResNet model did not work much better, achieving the best results with less than 40% accuracy for ResNet-34, which is not a feasible approach.

Part II

In the case that only images in the highlighted condition are selected for training and testing, at this point we are faced with an image classification task where features are harder to extract and categories are easily confused. We wrote python scripts to extract images with an even number of indexes in the training and validation sets to perform the experiments. With this approach, the situation improved and ResNet-50 achieved the best accuracy, 55%. Although we have expected that the accuracy of tobacco classification would not be too high, this result still does not satisfy us, and at the same time, this method does not take into account the characteristics of the dataset and wastes a lot of information.

Part III

In order to increase the amount of information that the network can learn, we combined two 3-channel images into a 6-channel image as a sample input to the network. To do this, we rewrote the Dataset class as well as the DataLoader class to allow the code to read two adjacent images at a time. Also, we changed the number of input channels of the first convolutional layer in the network to 6 and the dimension of the fully connected layer to 14, and removed the part of the first convolutional layer and the fully connected layer from the pre-training weights. The two images were merged together by torch.cat() after being transformed into a tensor. Afterwards, we ran the first test and, shockingly, the accuracy was only about 45%. We carefully investigated the reason for this and found that transforms can only process one image at a time, and the two images fed into the network were flipped differently and had no relationship. So we rewrote the trans-

Table 3: Prediction effect on C10 class pictures

	AlexNet	ResNet-18	ResNet-34	ResNet-50	ResNeXt-50
C01	1. 12e-05	8.46e-06	2. 23e-06	5. 08e-06	0. 000134
C02	6. 26e-05	2. 6e-06	6. 13e-06	1. 72e-07	8.18e-07
C03	0. 000279	0. 118	0. 00024	2. 25e-05	4. 73e-05
C04	0. 444	1. 13e-05	0. 000216	5e-05	2. 34e-05
C05	0. 000675	0. 000106	9. 76e-05	1. 72e-05	0. 000222
C06	0. 000684	0. 00324	0. 0007	3. lle-05	8. 79e-06
C07	3. 53e-05	2.08e-07	8.le-08	2.01e-07	3. 47e-07
C08	0. 0708	6. 55e-06	0. 00474	0. 000136	0. 000322
C09	0. 00137	9. 9e-07	0. 000211	6. 18e-06	1. 51e-05
C1O	0. 431	0. 879	0. 989	0. 999	0. 999
C11	0. 0394	3. 17e-05	0. 00514	0. 000815	5. 57e-05
C12	0. 00984	2. 22e-05	5. 21e-05	1. 66e-05	3. 16e-05
C13	0. 00169	4. 42e-08	1. 14e-07	1. le-07	5. 93e-07
C14	3. 47e-05	7. 89e-08	8. 83e-08	5. 3e-07	1. 15e-06
Result	C04	C10	C10	C10	C10

forms class to ensure that both images were processed the same way each time. After this step, the results were significantly improved and this is what we consider to be the best training method.

Result

Based on the best training method we explored, we trained five models and tested them, and finally ResNeXt-50 achieved the highest accuracy, 70.1%. The accuracy rates of the other models are shown in Table.2, and unlike what we expected, the results improved as the number of network layers increased, despite the smaller dataset.

Some test and training curves during the experiment are shown in Fig. 2. It can be seen that the decaying strategy of driving volume SGD with cosine annealing learning rate we used has a significant effect on model convergence, and the training effect is also better using 6-channel images compared with 3-channel images.

To verify the practical effects of the models, we explored the prediction results of five models for the same image, as shown in Table.3.

Conclusion

The main task of this paper is to explore the effectiveness of different models on tobacco classification based on a tobacco dataset close to real scenes, and to verify the effectiveness of our proposed 6-channel image training. The experiments show that all five models can have some results on tobacco classification, among which ResNeXt-50 achieves the highest accuracy, 70.1%. Combining two 3-channel images into one 6-channel image for training has a significant effect on tobacco classification, which can help the network learn features better and is well suited for this fine-grained recognition task. Tobacco classification is more complex than other image classification, which is a challenge for us, and we will try more models for exploration in the future to get better results.

References

Dasari, S. K.; Chintada, K. R.; and Patruni, M. 2018. Flue-Cured Tobacco Leaves Classification: A Generalized Approach Using Deep Convolutional Neural Networks.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2015. Deep Residual Learning for Image Recognition. *computer vision and pattern recognition*.

Huixin, L. 2019. *Research on Flue-cured Tobacco Classification Algorithm Based on Deep Learning*. Master's thesis, Beijing Jiaotong University.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. ImageNet Classification with Deep Convolutional Neural Networks. *Communications of The ACM*.

Loshchilov, I.; and Hutter, F. 2016. SGDR: Stochastic Gradient Descent with Warm Restarts. *arXiv: Learning*.

Marzan, C. S.; and Ruiz, C. R. 2019. Automated Tobacco Grading Using Image Processing Techniques and a Convolutional Neural Network. *International Journal of Machine Learning and Computing*.

Qiang, G. 2013. *The Research of Tobacco grading based on Image processing and Neural network*. Master's thesis, Central South University.

Sari, Y.; and Pramunendar, R. A. 2017. Classification Quality of Tobacco Leaves as Cigarette Raw Material Based on Artificial Neural Networks. *International Journal of Computer Trends and Technology*.

Shixin, W. 2020. A Tobacco Grading Algorithm Based on Convolution Neural Network. *Journal of Yunnan University for Nationalities (Natural Science Edition)*, 29(1): 65–69.

Srivastava, N.; Hinton, G. E.; Krizhevsky, A.; Sutskever, I.; and Salakhutdinov, R. 2014. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*.

Xiaojuan, Y. 2015. *Application of Tobacco Classification Based on K-nearest Neighbor*. Master's thesis, Zhengzhou University.

Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; and He, K. 2016. Aggregated Residual Transformations for Deep Neural Networks. *computer vision and pattern recognition*.

Yangfan, W. 2014. *Design and Implementation of a Tobacco Images Grading Algorithm Based on Color Characterist*. Master's thesis, Beijing Jiaotong University.

Zhenzhen, Z. 2016. *Method for Automatic Grading of Tobacco Based on Machine Vision*. Master's thesis, Southwest University.