



Fruit Image Recognition Based on Census Transform and Deep Belief Network

Qi Xin¹, Shaohai Hu^{1(✉)}, Shuaiqi Liu^{2,3(✉)}, Hui Lv⁴,
Shuai Cong^{5(✉)}, and Qiancheng Wang²

¹ College of Computer and Information,
Beijing Jiaotong University, Beijing 100044, China
shhu@bjtu.edu.cn

² College of Electronic and Information Engineering, Hebei University,
Baoding 071000, China
shdkj-1918@163.com

³ Machine Vision Technology Creation Center of Hebei Province,
Baoding 071000, China

⁴ Beagledata Technology (Beijing) Co., Ltd., Beijing 100089, China

⁵ Industrial and Commercial College, Hebei University,
Baoding 071000, Hebei, China
congshuai@hbu.edu.cn

Abstract. Fruit image recognition plays an important role in the fields of smart agriculture and digital medical treatment. In order to overcome the disadvantage of the deep belief networks (DBN) that ignores the local structure of the image and is difficult to learn the local features of the image, and considering that the fruit image is affected by the change of illumination, we propose a new fruit image recognition algorithm based on Census transform and DBN. Firstly, the texture features of fruit images are extracted by Census transform. Secondly, DBN is trained by Census features of fruit images. Finally, DBN is used for fruit image recognition. The experimental results show that the proposed algorithm has a strong feature learning ability, and the recognition performance is better than the traditional recognition algorithm.

Keywords: Fruit image recognition · Deep belief network · Census transform

1 Introduction

Nowadays, the sharp increase in the amount of image data makes the number of images in general image recognition tasks become larger and larger, which also makes it difficult for traditional methods to meet people's needs. As a new subject in machine learning, deep learning has a lot of achievements in various fields. Compared with the artificial feature extraction method, the data features acquired through deep learning model are more representative of the rich inner information of big data and good features can be learned automatically without manual feature extraction. Therefore, deep learning is the future and will receive more attention in big data analytics [1–3].

Fruit image recognition plays an important role in the fields of smart agriculture and digital medical [4, 5]. With the rapid development of smart agriculture and digital medical in recent years, fruit image recognition has attracted more and more researchers' attention. In order to meet the needs of large-scale and efficient fruit recognition and classification, researchers began to identify fruit image with different algorithms. For example, in [6], the authors proposed a fruit classification based on six layer convolutional neural network and the classification accuracy is higher than that of traditional single feature. In [7], the authors present an automatic fruit recognition system for classifying and identifying fruit types and which is capable of automatically recognize the fruit name with a high degree of accuracy. And in [8], the authors proposed Kiwifruit recognition method at night based on fruit calyx image, whose recognition rate reached 94.3%. And in [9], the authors proposed a fast and accurate object recognition method especially for fruit recognition to be used for mobile environment. They combined color, shape, texture and intensity into their associated code fields to generate an object code that could be used as a search key for the feature database. And in [10], the authors given a fruit recognition method via image conversion optimized through evolution strategy principal component analysis and achieved a better recognition effect through the pretreatment, training and recognition of fruit images, with an average recognition rate of over 92%.

In general, the process of fruit image recognition system mainly focuses on pretreatment and feature extraction. In this kind of recognition system, fruit image acquisition is mostly conducted by placing the collected fruit in a strictly defined background in order to ensure that the recognition system is less interfered by the outside world, so as to improve the recognition accuracy of the system. However, the image in the actual environment is easily affected by the factors such as illumination change, fruit reflection and shielding, which in vary degrees impact the recognition accuracy of fruit image. In fruit recognition system, fruit features mainly include odor, color, shape and texture. While in the process of growth, different environment will lead to difference in shapes, sizes and colors. In addition, natural light intensity and shadow will also be different when fruit images are collected, which will affect the accuracy of image recognition. What's more, the complexity of the color and texture features of fruit images also makes the recognition more difficult. Therefore, better recognition algorithms are needed to solve this problem.

As a representative method in the deep learning, DBN is quite different from the previous algorithms in terms of training method structure. By adopting the idea of layered training, the training speed of DBN is greatly improved [11, 12]. In addition, the idea of layering also increases the system's ability to express complex functions. DBN usually takes pixel-level images as input and extracts the abstract features of the input images from bottom to top, from simple to complex, which is a process of automatic mining useful information in the data. However, general pixel-level images are easily affected by illumination and other factors, which affects the extraction of essential features of input samples in DBN. In order to improve the fruit image recognition performance of DBN under different illumination, we propose a new method combining Census transform with DBN to extract the texture features of images through Census transform to eliminate the influence of ununiform lighting on the recognition results.

2 Fundamentals of Census Transform

Census transform [13] is a locality-based non-parametric transformation algorithm, which is mainly used to characterize the local structure features of images and can well detect the edge and corner features of images. It uses the relationship between the gray value of the neighborhood pixel and center pixel as similarity base to eliminate the influence of ununiform lighting. And, the implementation of non-parametric transformation is simple and real-time, which has been widely popularized and applied in engineering implementation. The algorithm flow is shown in Fig. 1.

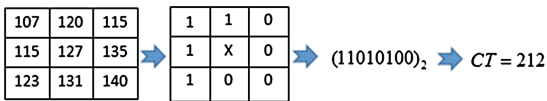


Fig. 1. The flow of Census transform algorithm

In Fig. 1, a pixel matrix is usually taken as the mask, whose central gray value is 127 and the gray value of each neighborhood pixel is shown in Fig. 1. We can get an 8-bit binary bit string through sequentially combine the eight digits obtained by the gray value comparison (from top to bottom, from left to right). Convert the binary bit string into a decimal number, and the gray value (whose range is [0, 255]) of the center pixel after Census transform is calculated.

Census transform replaces the original gray value of a pixel with its Census transform value, making these Census transform values mutually constrain and correlate [14]. These correlations make the Census transform value imply global feature information by passing local potential edge information to each other. In addition, the transformed values store information in a certain order, preserving the texture structure information between the local neighborhoods of the image, so that the global and local features of the image are not damaged, and the transformed features are easy to distinguish [15]. Therefore, we extract fruit features with the above algorithm.

3 Deep Belief Networks

Deep belief network [16] is a deep learning model with efficient learning algorithm proposed by Hinton. It combines unsupervised and supervised machine learning models and has become the main framework of deep learning algorithms since then. Deep belief network is superimposed by several restricted Boltzmann machines (RBM). Restricted Boltzmann machine [17] (RBM) is developed from a generated random neural network Boltzmann machine based on the principle of statistical mechanics. RBM consists of two layers that are visible layer and hidden layer. There is no connection within the layer and full connection between layers in the connection of neurons.

As shown in Fig. 2, the entire network of RBM is a bipartite graph, where $v = (v_1, v_2, \dots, v_{n_v})^T$ and $h = (h_1, h_2, \dots, h_{n_h})^T$ represent the state vector of the visible layer and the hidden layer, v_i represents the state of the i -th neuron in the visible layer, h_j represents the state of the j -th neuron, n_v and n_h represent the number of neurons contained in the visible layer and hidden layer, respectively. Let $a = (a_1, a_2, \dots, a_{n_v})^T \in R^{n_v}$ and $b = (b_1, b_2, \dots, b_{n_h})^T \in R^{n_h}$ respectively represent the bias vectors of the visible layer and the hidden layer, a_i represents the bias vector of the i -th neuron in the visible layer and b_j represents the bias vector of the j -th neuron in the hidden layer. And $W = (w_{i,j}) \in R^{n_h \times n_v}$ represents the weight matrix between the hidden layer and the visible layer and $w_{i,j}$ represents the connection weight between the i -th neuron in the hidden layer and the j -th neuron in the visible layer.

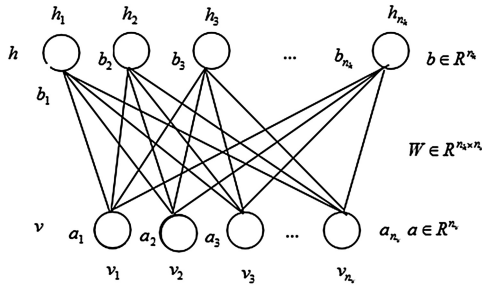


Fig. 2. RBM model

Even if the parameters of the model w_{ij} , a_i and b_j can be obtained through training, we still cannot effectively calculate the distribution determined by these parameters. However, the special structure based on the RBM model, that is, the visible layer and the hidden layer, are conditionally independent. When the state of all neurons on the visible layer is given, the probability that a certain neural unit on the hidden layer is activated (that is, the value is set to 1) is

$$P(h_j = 1|v, \theta) = \sigma(b_j + \sum_i v_i w_{ij}) \tag{1}$$

where $\sigma(\cdot)$ is the Sigmoid activation function and it is defined as

$$\sigma(x) = \frac{1}{(1 + \exp(-x))} \tag{2}$$

Similarly, we can get

$$P(v_i = 1|h, \theta) = \sigma(a_i + \sum_j w_{ij} h_j) \tag{3}$$

The relationship between the values of all nodes in a layer and the values of a single node is:

$$P(h|v) = \prod_{j=1}^{n_h} P(h_j|v); P(v|h) = \prod_{i=1}^{n_v} P(v_i|h) \quad (4)$$

The main task of RBM learning is to get the value of the optimal parameter $\theta = \{a_i, b_j, w_{ij}\}$ to fit the distribution of the given data sample. Assuming the training data sample to be $S = \{v^{(1)}, v^{(2)}, \dots, v^t\}$, we usually obtain the RBM parameters $\theta = \{a_i, b_j, w_{ij}\}$ by maximizing the RBM logarithm likelihood $L(\theta)$ on the training data sample. The specific formula is as follows:

$$\theta^* = \arg \max_{\theta} \sum_{i=1}^t \log P(v^{(i)}|\theta) \quad (5)$$

In order to obtain the optimal parameters θ , we use the stochastic gradient ascent method to find the maximum of $\sum_{i=1}^t \log P(v^{(i)}|\theta)$.

4 Fruit Image Recognition Based on Census Transform and DBN

The parameter setting rules adopted in this paper are as follows: the size of the fruit image in the training and test data is 32×32 , and the pixels of image is directly taken as the input of DBN, and the input layer of DBN is defined as 1024 units. In order to improve the computational efficiency and reduce the sampling error of the gradient estimation, it is necessary to use batch learning and divide the data set into small batches containing dozens or hundreds of samples in advance. Due to the limited data samples, each batch of data in this chapter is set to 10 samples. The fine-tuning learning rate of RBM is set as 0.01 and the number of iterations is set as 100. The activation function adopted by the neural network is Sigmoid function. As DBN can only be used for feature learning instead of classification function, we connect a softmax regression classifier at the last layer to classify the abstract feature attributes (network weight optimality of the hidden layer) that DBN eventually learns. The experiment shows that when the number of DBN layers is 2 and the number of hidden layers is 500, the identification accuracy of the whole model is the highest.

Based on the advantages of Census transform and deep belief network, we propose a fruit image recognition algorithm based on Census transform and deep belief network. The flow chart of the algorithm is shown in Fig. 3.

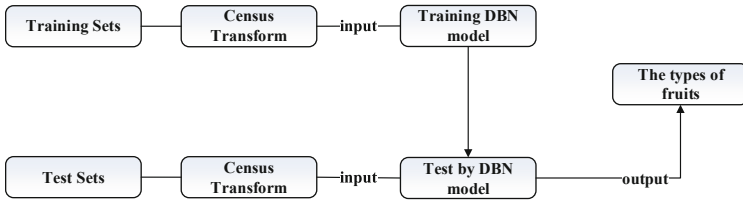


Fig. 3. The flow chart of the proposed method

The detail steps of the proposed algorithm are as following:

Firstly, the texture features of fruit images are extracted through Census transform to eliminate the influence of light changes on feature extraction. Census transform can effectively retain the local features of the image, which makes deep belief network effectively learn the local features of the image and reduce the learning of adverse feature description to deep belief network;

Secondly, the obtained Census features of the fruit images were used to train the depth belief network and obtain the relevant parameters of the network;

Finally, the deep belief network is used for fruit image recognition.

5 Experimental Results

Since the fruit images collected in our paper were collected under three different lighting environments and the fruit reflected light under different lighting conditions, the recognition results of DBN network were affected by ununiform lighting. In order to eliminate the influence of ununiform lighting, Figure 4 is an example of the illumination invariance of the fruit image obtained by our fruit data. Although there are illumination variations in Fig. 4(a) and Fig. 4(b), the renderings obtained by Census transform are almost the same in Fig. 4(c) and Fig. 4(d).

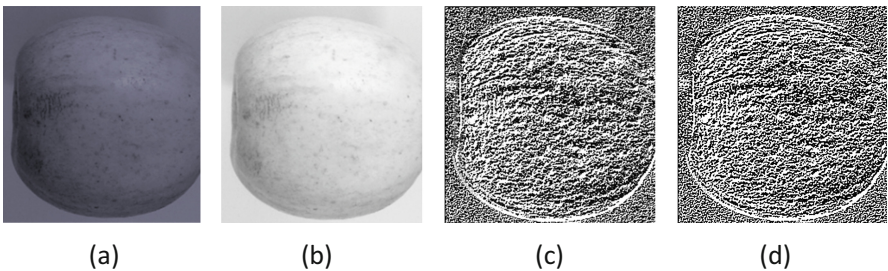


Fig. 4. The performance of Census transform on illumination invariance

In order to better illustrate the effectiveness of the proposed algorithm which named Census+DBN, different recognition algorithms are used for fruit image recognition,

such as CNN proposed in [6], DBN proposed in [17] and Census+Softmax proposed in [20]. The experimental results are shown in Table 1.

Table 1. The performance comparison of different fruit image recognition methods

Method	CNN	DBN	Census+Softmax	Census+DBN
Recognition accuracy rate	<i>0.94</i>	<i>0.95</i>	<i>0.90</i>	<i>0.98</i>
False recognition rate (%)	<i>0.84</i>	<i>6.88</i>	<i>13.4</i>	<i>0.17</i>

It can be seen from Table 1 that CNN and Census+DBN have achieved good recognition results in fruit image recognition, with recognition accuracy rates 0.94 and 0.98, while the false recognition rates are 0.84% and 0.17% respectively, which indicates that Census+DBN is as robust to light as CNN. Compared with the recognition results of Census+Softmax method, it shows that the deep learning method has a strong feature learning ability and the recognition performance is better than the traditional recognition algorithm. In the two deep learning methods adopted in this paper, although CNN achieves a recognition accuracy rate of 0.94 and false recognition rate of 0.84%, the training time of CNN is too long. One iteration of CNN spends 160 s without GPU acceleration in our experiment. Although convergence of CNN can be achieved with a few iterations, it cannot meet the requirements of practical application. Although DBN takes 2.57 s for one iterate, illumination variation will affect DBN's feature extraction of images and the recognition performance of the algorithm. In this paper, we effectively eliminate the influence of uniform lightning of the image by using Census transform. At the same time, the recognition performance is better than CNN and the iteration time is 1.75 s since the Census transform is a high-speed transform without parameters. So, the proposed fruit image recognition by combining Census transform and DBN can satisfy the requirement of real time application.

6 Conclusion

In this paper, we propose a new fruit recognition algorithm by combing Census transform and DBN. Census transform is adopted to extract texture features from fruit images and the extracted features are taken as the input of DBN. The feasibility and effectiveness of the method are proved through experiments. However, in our method, we do not consider other features of fruits images such as color, shape, texture and intensity. In the future, we can add these features to improve the performance of our algorithm.

Acknowledgments. This work was supported in part by National Natural Science Foundation of China under Grant 61572063 and 61401308, Natural Science Foundation of Hebei Province under Grant F2020201025, F2016201142, F2019201151 and F2018210148, Science Research Project of Hebei Province under Grant BJ2020030 and QN2017306, Foundation of President of Hebei University under Grant XZJJ201909. This work was also supported by the High-Performance Computing Center of Hebei University.

References

1. Liu, S., Wang, J., Lu, Y., et al.: Multi-focus image fusion based on residual network in non-subsampled shearlet domain. *IEEE Access* **7**(4), 152043–152063 (2019)
2. Liu, S., Liu, T., Gao, L., et al.: Convolutional neural network and guided filtering for SAR image denoising. *Remote Sens.* **11**(6), 702–720 (2019)
3. Chen, Y., Hu, X., Fan, W., et al.: Fast density peak clustering for large scale data based on kNN. *Knowledge-Based System*, 104824 (2019)
4. Altaheri, H., Alsulaiman, M., Muhammad, G., et al.: Date fruit dataset for intelligent harvesting. *Data Brief* **26**, 104514 (2019)
5. Soltan, M., Elsamadony, M., Mostafa, A., et al.: Nutrients balance for hydrogen potential upgrading from fruit and vegetable peels via fermentation process. *J. Environ. Manage.* **242**, 384–393 (2019)
6. Lu, S., Lu, Z., Aok, S., et al.: Fruit classification based on six layer convolutional neural network. In: 2018 ICDSF, China, Shanghai, pp. 1–5. IEEE (2018)
7. Zawbaa, H.M., Hazman, M., Abbass, M., et al.: Automatic fruit classification using random forest algorithm. In: 2014 14th International Conference on Hybrid Intelligent Systems, Arab, Kuwait, pp. 164–168. IEEE (2014)
8. Fu, L., Sun, S., Vázquez-Arellano, M., et al.: Kiwifruit recognition method at night based on fruit calyx image. *Trans. Chin. Soc. Agricult. Eng.* **33**(2), 199–204 (2017)
9. Kim, J., Vogl, M., Kim, S.D.: A code based fruit recognition method via image conversion using multiple features. In: 2014 International Conference on IT Convergence and Security (ICITCS), China, Beijing, pp. 1–4. IEEE (2014)
10. Vogl, M., Kim, J.Y., Kim, S.D.: A fruit recognition method via image conversion optimized through evolution strategy. In: 2014 IEEE 17th International Conference on Computational Science and Engineering, China, Chengdu, pp. 1497–1502. IEEE (2014)
11. Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. *Science* **313**(5786), 504–507 (2006)
12. Dunjko, V., Briegel, H.J.: Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Rep. Prog. Phys.* **81**(7), 074001 (2018)
13. Zabih, R., Woodfill, J.: Non-parametric local transforms for computing visual correspondence. In: Eklundh, J.-O. (ed.) *ECCV 1994*. LNCS, vol. 801, pp. 151–158. Springer, Heidelberg (1994). <https://doi.org/10.1007/BFb0028345>
14. Lai, X., Xu, X., Lv, L., Huang, Z., Zhang, J., Huang, P.: A novel non-parametric transform stereo matching method based on mutual relationship. *Computing* **101**(6), 621–635 (2019). <https://doi.org/10.1007/s00607-018-00691-3>
15. Lai, X., Xu, X., Zhang, J., et al.: An efficient implementation of a census-based stereo matching and its applications in medical imaging. *J. Med. Imaging Health Inform.* **9**(6), 1152–1159 (2019)
16. Mureşan, H., Oltean, M.: Fruit recognition from images using deep learning. *Acta Universitatis Sapientiae, Informatica* **10**(1), 26–42 (2018)
17. Ahmad, M., Ai, D., Xie, G., et al.: Deep belief network modeling for automatic liver segmentation. *IEEE Access* **7**, 20585–20595 (2019)
18. Le Roux, N., Bengio, Y.: Representational power of restricted Boltzmann machines and deep belief networks. *Neural Comput.* **20**(6), 1631–1649 (2008)

19. Larochelle, H., Bengio, Y., Louradour, J., et al.: Exploring strategies for training deep neural networks. *J. Mach. Learn. Res.* **10**, 1–40 (2009)
20. Cornejo, J.Y.R., Pedrini, H.: Automatic fruit and vegetable recognition based on CENTRIST and color representation. In: Beltrán-Castañón, C., Nyström, I., Famili, F. (eds.) *CIARP 2016*. LNCS, vol. 10125, pp. 76–83. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-52277-7_10