Identification of Tea Leaf Based on Histogram Equalization, Gray-Level Co-Occurrence Matrix and Support Vector Machine Algorithm

Yihao Chen(✉)

Department of Informatics, University of Leicester, Leicester LE1 7RH, UK
yc306@le.ac.uk

Abstract. To identify tea categories more automatically and efficiently, we proposed an improved tea identification system based on Histogram Equalization (HE), Gray-Level Co-Occurrence Matrix (GLCM) and Support Vector Machine (SVM) algorithm. In our previous project, 25 images per class might be not enough to classify, and a small size of dataset will cause overfitting. Therefore, we collected 10 kinds of typical processed Chinese tea, photographed 300 images each category by Canon EOS 80D camera, and regarded them as a first-hand dataset. The dataset was randomly divided into training set and testing set, which both contain 1500 images. And we applied data augmentation methods to augment the training set to a 9000-image training set. All the images were resized to 256 * 256 pixels as the input of feature extraction process. We enhanced the image features through Histogram Equalization (HE) and extracted features from each image which were trained through Gray-Level Co-Occurrence Matrix (GLCM). The results show that the average accuracy reached 94.64%. The proposed method is effective for tea identification process.

Keywords: Tea leaves identification · Data augmentation · Histogram Equalization (HE) · Gray-Level Co-Occurrence Matrix (GLCM) · Cross-Validation (CV) · Support Vector Machine (SVM)

1 Introduction

Tea is a popular drink all over the world. There are a variety of functions and effects, such as promoting the absorption of iron in the body, helping to regulate cholesterol levels in the body, promoting metabolism, and even inhibiting the occurrence of cancer [1]. The record of the United Nations Food and Agriculture Organization and Wang' s researches on Chinese tea culture illustrate [2] that in China alone people grow, ferment and bake black tea, green tea, oolong tea, scented tea and other varieties of tea. The production process of tea is extremely complicated. Different production processes, time periods and fermentation levels will affect the properties of tea products. Moreover, there are many varieties of tea leaves, and even the tea products made with the same leaf through different processes have different characteristics, colors and fragrances [3].

Y.-D. Zhang et al. (Eds.): ICMTEL 2020, LNICST 326, pp. 3–16, 2020.
https://doi.org/10.1007/978-3-030-51100-5_1

Over the years, the computer processing and analysis of data more and more accurate and efficient, related algorithms are also improving. In the field of machine learning and artificial intelligence, image processing is one of the most studied topics, such as medical image analysis, biological image classification, etc. These advances are attributed to the formation of image analysis algorithm, the improvement of classifier and the development of convolutional neural network. The work of tea classification has been improved a lot, but they are based on different standards, original data types and algorithms, and the results are quite different. Since the appearance of processed tea leaves is quite different from that of fresh tea leaves, it can be further explored to use tea products as the research object. If the algorithm and results of this project can be applied to industrial production process or application development of mobile terminals, then this project might have a very good prospect.

Previously, Wu, Chen (2009) [4] proposed a tea identification method using discrete cosine transform (DCT) and least squares-support vector machine (LS-SVM) based on multi-spectral digital image texture features. They used a multispectral digital imager to take images of different bands of tea leaves and combine them into a new image of one object. They trained the dataset with 80 filters and classified it with a classifier. Although the accuracy of this method can reach a considerable 100% due to the maturity of the filter and classifier design, it has the disadvantage that is requiring the use of additional scanning imagers. It means that the acquisition of data sets will cost too much.

Similarly, Zhang and HE (2014) [5] designed a system for identifying green tea brands based on hyperspectral imaging techniques and the Gray-level Co-occurrence Matrix (GLCM) and LS-SVM classifier. In this system, the dataset and the algorithm are perfect, and the results are accurate. However, the collection of the data set still needs to be finished by peripheral devices. Zhao, Chen (2006) [6] used near-infrared spectroscopy to identify green tea, Oolong tea and black tea quickly. By collecting the different spectral features of each tea in the near-infrared spectroscopy, they used these features as input to the support vector machine (SVM) classifier. Their study was more than 90% accurate.

All the above methods have achieved good recognition results, but they needed to collect data with additional equipment, which leads to the acquisition of data sets that require material and time. Therefore, we tried to use the image recognition technology to identify the common tea product image. The advantage of this approach is that it is simpler, more efficient and less costly without expensive equipment and cumbersome processes. For example, Borah, Hines (2007) [7] proposed an algorithm for image texture analysis based on wavelet transform. The technique directly processes the feature information of a group of images and then analyzes it. Yang (2015) [8] created a system to identify green teas, Oolong teas and black teas based on wavelet packet entropy and fuzzy support vector machine. It was an efficient method and the accuracy was over 97%. Tang (2018) [9] experimented that tested Convolutional neural network provides us with an inspiration. In our previous study [10], we used the Gray-Level Co-Occurrence Matrix (GLCM) model and k-nearest neighbors algorithm to extract features from images which were taken by ourselves. Moreover, we got a good result that showed the accuracy was over 86%. However, one of the disadvantages is that this project is not automatic enough because we just called in the classification function in MATLAB software to test the initial performance of GLCM on this set of images.

In the previous study, since GLCM with fewer parameters and less operation can effectively extract features from images, we continued to use this model in this study. GLCM is a commonly used image processing technology, which can reflect the joint distribution between two pixel points with different spatial position relations [11]. In the feature extraction step, MATLAB was used to run GLCM algorithm to extract feature dataset, and the input images were digitized to obtain the feature matrix. Each of the characteristic values in these matrixes was a sorted input. After the training and testing of the classifier and the verification of the multi-fold cross-validation process, the accuracy of the test data set can be finally calculated. In this process, using GLCM algorithm to extract features was a key step.

The aim of this study is to design a tea identification system based on traditional computer vision methods using our own dataset and to test the classification performance of a combination of Histogram Equalization (HE), Gray-Level Co-Occurrence Matrix (GLCM) and Support Vector Machine (SVM).

In order to test the classification performance of our model, this study used the pictures of ten kinds of tea, which were, respectively, Anji White tea, Chrysanthemum Tea, Huangshanhoukui Green Tea, Jasmine Tea, Longjing Green Tea, Tawthron Black Tea, Tieguanyin, Xiangxin Black Tea, Yixing black tea and Zhengshanxiaozhong Black Tea, as the original data. For instance, there are three green teas which have similar features, such as leaf shape and color. However, there is a slight whitening change during the growth of white tea, and this feature still exists in the processed tea.

2 Methodology

2.1 Data Acquisition

In our previous study, the high-resolution tea pictures we photographed and collected by ourselves stored enough useful information, such as leaf texture feature, pixel value. Thus, we continued to use the similar set of pictures as the data set in this study. The categories of these teas are, respectively, Anji White tea, Chrysanthemum Tea, Huangshanhoukui Green Tea, Jasmine Tea, Longjing Green Tea, Tawthron Black Tea, Tieguanyin, Xiangxin Black Tea, Yixing black tea and Zhengshanxiaozhong Black Tea (10 categories tea). The tea leaves were photographed in two different shades of light, with half in the sun light and the other half in a warm LED light of 8 W. It was obvious to find that the tea images in these two environments were different, which are shown in Fig. 1. To reduce the influence of the reflected light of the background, we placed all the tea leaves on dark background. Finally, 300 high-quality images were selected for each category of tea, and the images from different environments were selected randomly.

The images we selected were approximately 4 MB in size and had a pixel value of about 4000 * 4000. These images contain enough tea information, such as color, leaf shape, texture, quantity, etc. Generally, image recognition only needs a small size to

(a) Anji white tea in the warm LED light. (b) Anji white tea in the sun light.

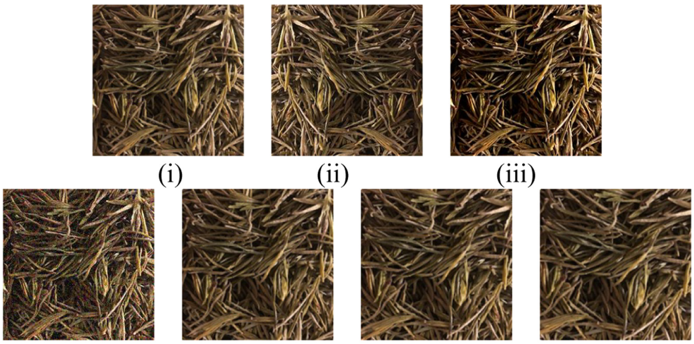
Fig. 1. The positions of Anji white tea

remain the program efficiency, so we resized these images to 256 * 256 pixels in batch through MATLAB software.

2.2 Data Augmentation

In our last experiment, Although the program can be run in the previous experiment, the average accuracy is about 65% [10]. The insufficiently sized dataset may be a main reason for the low accuracy. Besides, overfitting might be caused by a small number of samples [12]. The raw data sets are divided into two parts randomly, which are used as training sets and test sets respectively. 1500 samples are in the training set, and rest of them are tested as the testing set. We used data augmentation technology to rotate, mirror, crop and other operations on the acquired images [13].

Since our data set was small before, we applied this technique to increase the training data set and make the data set diverse. Since the larger dataset can be show a better performance of classification, we augmented the training set from 1500 images to 9000 images. Some samples are shown in Fig. 2: (i) is the raw image, (ii) is mirrored, (iii) is added zero-mean, Gaussian white noise with variance of 0.01, (iv) is enhanced contrast in a new range of intensity which is from 0.1 to 0.9, and (v), (vi) and (vii) are rotated 5°, 10°, 15° clockwise, respectively. These images were stored uniformly in a folder.



(i) (ii) (iii)

(iv) (v) (vi) (vii)

Fig. 2. Proposed samples of Anji White Tea

2.3 Feature Extraction

To reduce the influence of different intensities of light and unify the overall grayscale of the images, we introduced the method of HE. It can enhance the local contrast without affecting the overall contrast to present details of images. The process of HE is as following: (1) The frequency of occurrence of each gray level of the histogram is counted and normalized to [0,1]. The probability of a pixel with a grayscale of i is

$$p_x(i) = \frac{n_i}{n}, i \in 0, 1\ldots, L-1\,, \tag{1}$$

where n represents the number of times a certain grayscale occurs, L represents gray level.

(2) Accumulate normalization histogram:

$$cdf_x(i) = \sum_{j=0}^{i} p_x(j). \tag{2}$$

To summarize, the gray value is:

$$h(v) = round\left(\frac{cdf(v) - cdf_{min}}{(M\times N) - cdf_{min}} \times (L-1)\right), \tag{3}$$

where *round()* represents the returning the integral value that is nearest to the argument. M and N represent the number of pixels in length and width of the image.

GLCM represents the correlation between different gray pixels in the image. Texture feature extraction [14] is divided into four parts: extraction of gray image, gray level quantization, calculation of eigenvalue, and generation of texture feature maps [14]. Compared with wavelet transform, GLCM can better describe spatial relations. By extracting the original three-channel image and then processing the gray level, we can obtain the required gray level image. This was useful for determining the grayscale value of each pixel in the image [15].

Suppose that f(x, y) presents a two-dimensional image, the size of which is M × N. Then, relative frequencies in this image P(i, j) will satisfy the following condition:

$$P(i,j) = \#\{(x_1, y_1), (x_2, y_2) \in M \times N | f(x_1, y_1) = i, f(x_2, y_2) = j\}, \tag{4}$$

where #(x) presents the number of elements in the set.

The adjacent relations between the gray values of different pixels, the features, can be extracted, and then stored in one matrix [16–19]. The gray-level co-occurrence matrix can be obtained through this step. Through this process, we can represent larger color or gray images by images with smaller pixel features. If the gray value of pixels in the image are similar, the features in the matrix will be concentrated on the main diagonal. As shown in Fig. 3, a window which is set as 1 × 2 moves from the first two blocks to the end in horizontal direction, and the value of the step distance is 1.

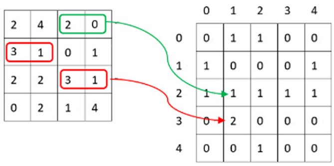


Fig. 3. A sample of GLCM

Many statistics can be used to describe texture features in GLCM, among which the more typical ones are contrast, correlation, energy and homogeneity. Contrast directly describes the brightness difference between pixels in an image. Mathematically,

$$CON = \sum_{i}^{k} \sum_{j}^{k} |i-j|^2 \cdot G(i,j), \tag{5}$$

where $G(i,j)$ represents the relative frequencies, k represents the maximum number of pixels.

Correlation indicates the similarity of features in horizontal and vertical directions. It can be calculated as,

$$COR = \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{(i \cdot j) \cdot G(i,j) - U_i U_j}{S_i S_j} .(S_i S_j \neq 0), \tag{6}$$

where

$$U_i = \sum_{i-1}^{k} \sum_{j=1}^{k} i \cdot G(i,j), \tag{7}$$

$$U_j = \sum_{i=1}^{k} \sum_{j=1}^{k} j \cdot G(i,j), \tag{8}$$

$$S_i^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} G(i,j)(i - U_i)^2, \tag{9}$$

$$S_j^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} G(i,j)\left(j - U_j\right)^2. \tag{10}$$

The ASM (Angular second moment) energy reflects the stability of the grayscale change of the image texture. The higher the energy value is the more uniform the texture will be. Homogeneity describes the local stability of the image. Mathematically,

$$ASM = \sum_{i}^{k} \sum_{j}^{k} (G(i,j))^2 \tag{11}$$

$$HOM = \sum_{i}^{k} \sum_{j}^{k} \frac{G(i,j)}{1+|i-j|} \tag{12}$$

In this project, we scanned the images in 0°, 45°, 90°, 135° respectively, so the feature map is a 4-channel matrix. It represents the preprocessed 256 * 256-pixels image as an 8 * 8 * 4-pixels image, and it is also an 8 * 8 * 4-matrix. Each number in the matrix represents the overall features and local features of the image. The four features mentioned above were also utilized. Different images of one same tea have similar texture characteristics, color and structure. Therefore, the feature matrix will have similar overall structure and specific feature values. In the future, we shall test other feature extraction methods, such as convolutional neural network [20–24].

2.4 Classification

Support vector machine (SVM) [25] is a popular learning model and classifier recently. The standard support vector machine mainly solves the binary classification problem, but after the algorithm improvement, the multi-classification problem can also be solved by it. We do not use deep learning methods [26–31] or transfer learning methods [32–35] since our dataset is small.

The algorithm and steps of standard SVM can be used to construct multiple classification boundaries, so as to realize one-against-all multi-classification [36]. Although support vector machine cannot solve outliers and noise, the classic SVM is worth using because it can show a good performance in the small sample classifying experiments. In Wu, Chen (2009) [4]'s study, Least Squares-Support Vector Machine was applied well. In Yang (2015) [8]'s research, they used fuzzy support vector machines and got good results.

We suppose that there are a set of inputs called $\mathrm{X} = \{X_1, X_2, \ldots, X_N\}$ and a goal called $\mathrm{Y} = \{y_1, y_2, \ldots, y_N\}$. Each item of data contains the feature space $\mathrm{X}_i = \{x_1, x_2, \ldots, x_N\}$ and the binary goal $\mathrm{y} = \{-1, 1\}$. We suppose to regard the input feature space as a hyperplane of the decision boundary, and the goal set can be divided into positive and negative categories. The hyperplane can be represented as

$$\omega^T X + b = 0. \tag{13}$$

The point to the plane distance D can be represented as

$$D = y_i\left(\omega^T X_i + b\right), \tag{14}$$

where ω represents normal vector and b represents intercept to the hyperplane.

In fact, the decision boundaries used to separate the samples were created. As shown in Fig. 4, the positive class (Class B) corresponding to the boundary $y_i = +1$ is $\omega^T X_i + b \geq +1$. The negative class (Class A) corresponding to the boundary $y_i = -1$ is $\omega^T X_i + b \leq -1$. We can deduce this method into multiple classification problems. Two interval boundaries are calculated for every two samples in the M groups of input samples, so that we can get [M * (M – 1)]/2 boundaries totally. SVM has a good performance in both binary and multi-classification problems.

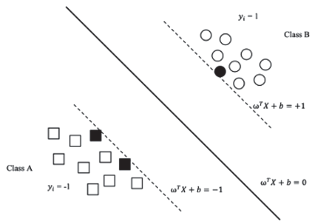


Fig. 4. SVM model

The original Support vector machine (SVM) is designed for the binary classification projects. However, in our study we have 10 classes. One solution to multi-classification is to regards one class as the positive set and the rest of the classes are negative set and every class should be looped to train the model, which is one-versus-rest (OVR).

3 Experiments and Results

This system was developed via the Grayscale Processing Toolbox in MATLAB R2018b. The programs were run on the HP laptop with 2.30 GHz i5–8300H CPU, 8 GB RAM, and Windows 10 operating system.

3.1 Data Acquisition

A set of raw images have been photographed by us via a digital camera with a 35-mm lens, which have 300 images in each class\category. Some samples are shown in Fig. 5. We randomly divided each category of images into two equal groups as testing set and training set, respectively.

Fig. 5. Some samples of tea leaf images

3.2 Data Augmentation

Firstly, we used MATLAB to resize the images because it was not necessary to scan a large image that is at 4000 * 4000-pixels. All images were processed into 256 * 256-pixels square images. The purpose of this preprocessing operation is to compress the image while retaining the information of the image itself to improve the efficiency.

We rotated the training images clockwise by following angles: 5º, 10º, 15º, mirrored each class of images separately, changed the contrast, added gaussian noise which are shown in Fig. 2. In order to utilize this dataset in our future work, we reserved this dataset and named it CTset, which shows in Table 1.

Table 1. Statistic of CTset

Training set	9000
Testing set	1500
Total number	10500

3.3 Histogram Equalization

Before extracting features, we first grayscaled the image. But the overall contrast of the image will be affected by many factors, such as lighting angle, light intensity, etc., so that the feature details in the grayscale image will not be clearly displayed. Therefore, HE was utilized to preprocess the grayscaled images. The output images contain 64-bins default. A sample of difference between one raw image and a processed image and their histograms are as showing in Fig. 6. (a) is an original grayscaled image. (b) is an image after histogram equalization. (c) and (d) are the histograms of (a) and (b) respectively.

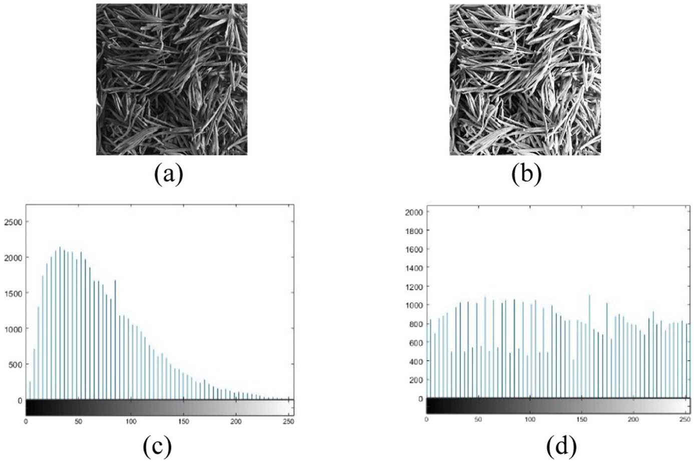


Fig. 6. A sample of histogram equalization

3.4 Feature Extraction

There is a grayscale processing toolbox which already contains the developed program command GLCM = graycomatrix (I) in MATLAB software, which can directly input and scan the input image. Based on our previous experimental results, we set the size of the output to 8 * 8 * 4. Regarding the obtained GLCM, most eigenvalues are distributed along the main diagonal of the matrix. One of the grayscale matrices in one channel is shown in Table 2.

Table 2. An 8 * 8 example of Anji white tea's GLCM in one channel

8591	2996	256	9	0	0	0	0
3027	11094	323	382	8	0	0	0
232	3550	7393	1968	1960	117	0	0
9	340	1973	2947	2953	465	3	0
4	341	1956	1490	2952	346	57	28
0	0	34	739	1028	597	208	5
0	0	25	12	194	130	98	40
0	0	0	7	5	8	18	11

Then, we combined all the training set matrixes and testing set matrixes of into one training matrix and one testing set respectively, the size of which are 9000 * 256 and 1500 * 256. They contain gray-level co-occurrence matrix features of the raw grayscaled matrix in which each row represents one image. We named them as A_TR and A_TE. We also extracted the contrast, correlation, energy values, and homogeneity of the matrix from the stats = graycoprops(A) command and stored them in 9000 * 16 and 1500 * 16 matrixes named B_TR and B_TE.

We artificially labeled each category of tea leaf from 1 to 10. The label matrix is called matrix C_TR and C_TE. The matrix A_TR and C_TR needed to be combined into an augmented matrix D_TR. And we combined B_TR and C_TR into E_TR to test the GLCM feature maps (same operation needs to apply in the testing set, D_TE and E_TE). The size of the D_TR and E_TR are 9000 * 257 and 9000 * 17. Through the above operations, we have used a matrix to represent all the data, where each element was the characteristic parameter of each image and each row represented the overall feature of each image.

Table 3. The classification accuracy with and without HE

Model	Average accuracy
GLCM+SVM	93.07%
HE+GLCM+SVM	**94.64%**

(Bold means the best.)

3.5 Classification

In order to reflect the advantages of the Histogram Equalization, we compared the classification results of our HE+GLCM+SVM model and GLCM+SVM model, which shows in Table 3. We ran the experiments for 10 times and obtained an average accuracy to avoid contingency. Table 3 shows that the proposed models achieves good results. The Histogram Equalization method benefits to our classification problem by increasing the average accuracy (about 1.57%).

Table 4. The comparison of GLCM features (Bold means the best.)

Classifier	Average accuracy
Contrast	**73.61%**
Correlation	65.93%
Energy	69.08%
Homogeneity	65.47%

We proposed to test the four main GLCM features separately to discover the most motivating feature for this problem, which are contrast, correlation, energy and homogeneity. The results are shown in Table 4. The contrast achieves the highest average classification accuracy in this study.

Table 5. The classification accuracy of common classifiers

Classifier	Average accuracy
KNN	94.38%
NBS	83.58%
DT	92.62%
RF	92.81%
HE+GLCM+SVM (Ours)	**94.64%**

(Bold means the best.)

We compared the results of this study with our previous experiments which is based on GLCM and K-Nearest Neighbor (KNN) [10], Naive Bayes Classifier [37], Decision Tree [38], and Random Forest [39]. The results on the testing set are listed in Table 5, each of which are tested in 10 times.

To summarize, the proposed HE+GLCM+SVM model demonstrates the best performance of classification, and the average accuracy reaches 94.64%. This model focuses more on the texture features through GLCM method and benefits from HE method by enhancing the feature details.

4 Conclusion

The aim of this project is to develop an automatic tea identification/classification system with a high accuracy. Compared with our previous work and some classic classifies, we have augmented the dataset and improved the model. It is conceivable that larger data sets can be tested to obtain better results and avoid overfitting. The system was created based on Histogram Equalization (HE), Gray-Level Co-Occurrence Matrix (GLCM) and support vector machine (SVM) to cope with the ten-class problem. The results showed that this model can successfully identify and classify different tea varieties, and the average accuracy reaches 94.64%.

In the future, we will collect more kinds of tea images or build a larger dataset, test other classifiers, and design a model based on Convolutional neural network (CNN) and other machine learning methods. Besides, our system has been proved to be successful only on our dataset, and it is not certain whether it is suitable for the common classification dataset. We shall collect more samples and classes of tea images and build a higher performance system.

References

1. Conney, A., et al.: Inhibitory effect of green and black tea on tumor growth. Proc. Soc. Exp. Biol. Med. **220**(4), 229–233 (1999)
2. Wang, L.: Tea and Chinese Culture. Long River Press (2005)
3. Zhang, L., et al.: Effect of drying methods on the aromatic character of Pu-erh Tea. **1**, 71–75 (2007)
4. Wu, D., et al.: Application of multispectral image texture to discriminating tea categories based on DCT and LS-SVM. Spectroscopy Spectral Anal. **29**(5), 1382–1385 (2009)
5. Zhang, H.-L., et al.: Identification of green tea brand based on hyperspectra imaging technology. Spectroscopy Spectral Anal. **34**(5), 1373–1377 (2014)
6. Zhao, J., et al.: Qualitative identification of tea categories by near infrared spectroscopy and support vector machine. J. Pharmaceutical Biomed. Anal. **41**(4), 1198–1204 (2006)
7. Borah, S., et al.: Wavelet transform based image texture analysis for size estimation applied to the sorting of tea granules. J. Food Eng. **79**(2), 629–639 (2007)
8. Yang, J.: Identification of green, Oolong and black teas in China via wavelet packet entropy and fuzzy support vector machine. Entropy **17**(10), 6663–6682 (2015)
9. Zhang, Y.-D., Muhammad, K., Tang, C.: Twelve-layer deep convolutional neural network with stochastic pooling for tea category classification on GPU platform. Multimedia Tools Appl. **77**(17), 22821–22839 (2018). https://doi.org/10.1007/s11042-018-5765-3
10. Chen, Y., et al.: Tea leaves identification based on gray-level co-occurrence matrix and K-nearest neighbors algorithm. In: AIP Conference Proceedings, p. 020084. AIP Publishing LLC (2019)
11. Benčo, M., et al.: Novel method for color textures features extraction based on GLCM. Radioengineering **16**(4), 65 (2007)
12. Tetko, I.V., et al.: Neural network studies. 1. Comparison of overfitting and overtraining. **35** (5), 826–833 (1995)
13. Tanner, M.A., et al.: The calculation of posterior distributions by data augmentation. J. Am. Stat. Assoc. **82**(398), 528–540 (1987)
14. Pagani, L., et al.: Towards a new definition of areal surface texture parameters on freeform surface: Re-entrant features and functional parameters. Measurement **141**, 442–459 (2019)
15. Nanni, L., et al.: Texture descriptors for representing feature vectors. Expert Syst. Appl. **122**, 163–172 (2019)
16. Bradley, P.S.: A support-based reconstruction for SENSE MRI. Sensors **13**(4), 4029–4040 (2013)

17. Wu, L.N.: Segment-based coding of color images. Sci. China Ser. F-Inf. Sci. **52**(6), 914–925 (2009)
18. Wu, L.N.: Pattern recognition via PCNN and tsallis entropy. Sensors **8**(11), 7518–7529 (2008)
19. Wu, L.N.: Improved image filter based on SPCNN. Sci. China Ser. F-Inf. Sci. **51**(12), 2115–2125 (2008)
20. Zhang, Y.-D., Jiang, Y., Zhu, W., Lu, S., Zhao, G.: Exploring a smart pathological brain detection method on pseudo Zernike moment. Multimedia Tools Appl. **77**(17), 22589–22604 (2017). https://doi.org/10.1007/s11042-017-4703-0
21. Cheng, H.: Multiple sclerosis identification based on fractional Fourier entropy and a modified Jaya algorithm. Entropy **20**(4) (2018). Article ID. 254
22. Zhang, Y.-D., Sun, J.: Preliminary study on angiosperm genus classification by weight decay and combination of most abundant color index with fractional Fourier entropy. Multimedia Tools Appl. **77**(17), 22671–22688 (2017). https://doi.org/10.1007/s11042-017-5146-3
23. Lu, S.: Pathological brain detection based on alexnet and transfer learning. J. Comput. Sci. **30**, 41–47 (2019)
24. Yang, J.: Preclinical diagnosis of magnetic resonance (MR) brain images via discrete wavelet packet transform with Tsallis entropy and generalized eigenvalue proximal support vector machine (GEPSVM). Entropy **17**(4), 1795–1813 (2015)
25. Parhizkar, E., et al.: Partial least squares- least squares- support vector machine modeling of ATR-IR as a spectrophotometric method for detection and determination of iron in pharmaceutical formulations. Iranian J. Pharmaceutical Res. **18**(1), 72–79 (2019)
26. Zhang, Y.-D., Dong, Z., Chen, X., Jia, W., Du, S., Muhammad, K., Wang, S.-H.: Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. Multimedia Tools Appl. **78**(3), 3613–3632 (2017). https://doi.org/10.1007/s11042-017-5243-3
27. Li, Z.: Teeth category classification via seven-layer deep convolutional neural network with max pooling and global average pooling. Int. J. Imaging Syst. Technol. **29**(4), 577–583 (2019)
28. Tang, C.: Cerebral micro-bleeding detection based on densely connected neural network. Front. Neurosci. **13** (2019). Article ID. 422
29. Jia, W., Muhammad, K., Wang, S.-H., Zhang, Y.-D.: Five-category classification of pathological brain images based on deep stacked sparse autoencoder. Multimedia Tools Appl. **78**(4), 4045–4064 (2017). https://doi.org/10.1007/s11042-017-5174-z
30. Chen, Y.: Cerebral micro-bleeding identification based on a nine-layer convolutional neural network with stochastic pooling. Concurrency Comput.: Practice Exp. **31**(1), e5130 (2020)
31. Wang, S.-H., Muhammad, K., Hong, J., Sangaiah, A.K., Zhang, Y.-D.: Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization. Neural Comput. Appl. **32**(3), 665–680 (2018). https://doi.org/10.1007/s00521-018-3924-0
32. Xie, S.: Alcoholism identification based on an AlexNet transfer learning model. Front. Psychiatry **10** (2019). Article ID. 205
33. Hong, J., Cheng, H., Zhang, Y.-D., Liu, J.: Detecting cerebral microbleeds with transfer learning. Mach. Vis. Appl. **30**(3), 1123–1133 (2019). https://doi.org/10.1007/s00138-019-01029-5
34. Jiang, X., et al.: Classification of Alzheimer's disease via eight-layer convolutional neural network with batch normalization and dropout techniques. J. Med. Imaging Health Inform. **10**(5), 1040–1048 (2020)

35. Govindaraj, V.V.: High performance multiple sclerosis classification by data augmentation and AlexNet transfer learning model. J. Med. Imaging Health Inform. **9**(9), 2012–2021 (2019)
36. Hsu, C.-W., et al.: *A* comparison of methods for multiclass support vector machines. IEEE Trans. Neural Netw. **13**(2), 415–425 (2002)
37. Rish, I.: An empirical study of the naive Bayes classifier. In: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, pp. 41–46. IBM New York (2001)
38. Safavian, S.R., et al.: A survey of decision tree classifier methodology. IEEE Trans. Syst. Man Cybern. **21**(3), 660–674 (1991)
39. Liaw, A., et al.: Classification and regression by randomForest. **2**(3), 18–22 2002