



Gingivitis Classification via Wavelet Entropy and Support Vector Machine

Cui Li^(✉) and ZhiHai Lu

School of Education Science, Nanjing Normal University,
Nanjing 210024, Jiangsu, China
{997268314, 1030609013}@qq.com

Abstract. Gingivitis is usually detected by a series of oral examinations. In this process, the dental record plays a very important role. However, it often takes a lot of physical and mental effort to accurately detect gingivitis in a large number of dental records. Therefore, it is of great significance to study the classification technology of gingivitis. In this study, a new gingivitis classification method based on wavelet entropy and support vector machine is proposed to help diagnose gingivitis. The feature of the image is extracted by wavelet entropy, and then the image is classified by support vector machine. The experimental results show that the average sensitivity, specificity, precision and accuracy of this method are 75.17%, 75.29%, 75.35% and 75.24% respectively, which are superior to the other three methods. This method is proved to be effective in the classification of gingivitis.

Keywords: Wavelet · Entropy · Support vector machine · Gingivitis · Classification

1 Introduction

Teeth are one of the most important parts of human mouth. There are many diseases of teeth. Gingivitis is a common disease of teeth. Gingivitis mainly refers to the acute or chronic inflammation on the gum. The common symptoms of gingivitis mainly include swelling, bleeding and pain. Some patients may also have local itching and bad breath. Plaque accumulation, lack of nutrients or improper brushing may cause gingivitis.

In recent years, many scholars have carried out research on gingivitis and achieved many new results. Li (2019) [1] identified gingivitis through the method based on gray level co-occurrence matrix (GLCM) and extremum learning machine (ELM). In the study, this method was used to identify 52 teeth images, and different information about teeth was obtained through segmentation and classification from teeth images. It is found that this method is superior to other techniques in average sensitivity, specificity, precision and accuracy. Li (2019) [1] studied a method to identify gingivitis based on contrast-limited adaptive histogram equalization (CLAHE), gray level co-occurrence matrix (GLCM) and extremum learning machine (ELM). By comparing 58 pictures of gingivitis with 35 pictures of healthy teeth, we found that compared with the most advanced method, this method has higher classification accuracy and more sensitive results. Supranoto, Slot (2015) [2] studied the effect of chlorhexidine denifrice or

gel versus chlorhexidine mouthwash on gingivitis by retrieving databases such as PubMed-MEDLINE. Through screening 2256 samples, 5 publications meeting the standards were obtained. The research found that: by using chlorhexidine denifrice or gel versus chlorhexidine mouthwash, we can effectively inhibit gingivitis. A randomized clinical trial was conducted by Sangeetha (2017) [3] to investigate the effect of triclosan containing tooth paste and conventional fluoride tooth paste on gingivitis. In the study, 56 children were randomly divided into two groups, experimental group used the triclosan containing tooth paste and control group used conventional fluoride tooth paste. The results showed that the experimental group was better than the control group in reducing the incidence of gingivitis. Triclosan containing tooth paste can inhibit gingivitis more effectively. Feng, Zhang (2015) [4] used wavelet energy (WavEnr) to identify brain images. Brown (2018) [5] used extreme learning machine (ELM) to identify gingivitis lesions of teeth images.

Through the above research results, we can find that the research of gingivitis mainly includes exploring the recognition methods of gingivitis, studying the methods of inhibiting gingivitis.

The main contribution of this study is to combine wavelet entropy (WE) and support vector machine (SVM) to propose a new method for gingivitis classification. wavelet entropy (WE) can not only get the image features, but also reduce the dimension of the features. As a powerful classifier, support vector machine (SVM) can achieve the image classification. We have achieved good results for combining WE with SVM.

The other parts of this paper are as follows: the second section describes the dataset, the third section briefly introduces the research methods, the fourth section is the experimental results and discussion, the fifth section is the conclusion of this study.

2 Dataset

In this study, we selected 5 patients with gingivitis from Nanjing Stomatological hospital to observe their gingivitis [1]. In this study, two digital single lens reflex (DSLR), i.e. A and B, were used to randomly select different teeth of each patient for image collection. A total of 170 teeth pictures were obtained, including 85 gingivitis pictures and 85 healthy teeth pictures. In the image, we mark three regions: near, middle and far. The field of view is 51–200 mm in diameter, and the voxel resolution is 0.1–0.39 mm.

In the study, we will manually adjust the length and width of the area of interest to make its appearance similar to the simulated image, in which we can clearly see the tooth area. The average length of the 12 bit image is 456732 and the average width is 567833. Figure 1 shows two sample images in the dataset. Figure 1(a) shows a gingivitis image, and Fig. 1(b) shows a healthy teeth image.



Fig. 1. Samples of our dataset

3 Methodology

3.1 Wavelet

Wavelet transform [6–10] is an improved frequency transform method based on Fourier Transform (FFT). Wavelet transform overcomes the defect that the unstable signal can't be processed in Fourier transform [11–15]. The infinite trigonometric function base is replaced by the finite attenuation wavelet base, which can not only capture the frequency of the unstable signal, but also locate the time.

The image is a two-dimensional matrix. We get four components (LL1, LH1, HL1, HH1) after wavelet transform. One component (LL1) represents the blurred image, and the other three components (LH1, HL1, HH1) represent the detailed image. Then we get another four components (LL2, LH2, HL2, HH2) of LL1 by double wavelet transform (2D-DWT). After the double wavelet transform (2D-DWT), we get an entropy value. The formula of wavelet is:

$$WT(\alpha, \tau) = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{\infty} f(t) * \varphi\left(\frac{t-\tau}{\alpha}\right) dt \quad (1)$$

Wavelet transform can effectively decompose images of different pixels and retain image information. But after decomposition, it contains too many image features, which not only takes up a lot of storage space, but also increases the calculation time [16–20]. So we need to reduce the dimension of the feature by introducing entropy.

3.2 Entropy

Entropy is the uncertainty degree of information that Shannon quoted from thermodynamics. Entropy is a measure of disorder, which is used to represent the average value of information of probability distribution [21–25].

Suppose a random variable X , the value of X in dataset D is $X = \{x_1, x_2, \dots, x_n\}$, $P(R)$ is the probability function, then the entropy is:

$$H(X) = E[I(X)] = E[-\ln(K(X))] \tag{2}$$

Where E is the expected value. If D is an infinite set, then the entropy of the random variable X is:

$$H(X) = - \sum_i P(x_i) \log_b P(x_i) \tag{3}$$

As shown in Fig. 2, after wavelet transform, the image is decomposed into seven components, and then the entropy of these seven components is calculated to get the eigenvector. Through the method of wavelet entropy, we can not only get the image features, but also effectively reduce the dimension of image features.

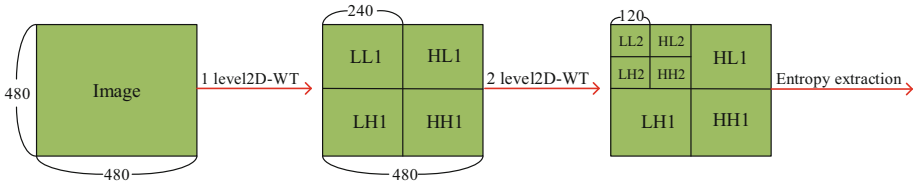


Fig. 2. Diagram of 2D-WT

3.3 Support Vector Machine

Like classified learning, the most basic idea of image classification is to find a partition hyperplane to separate different images. Support Vector Machine (SVM) is the latest classification method based on machine learning theory [26–29]. It is a two-classification model to find a hyperplane for image segmentation. The function of SVM is to help to build a hyperplane with a maximum interval. In support vector machine, a training sample set D is given.

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m), y_i \in \{-1, +1\}\}. \tag{4}$$

Based on the sample set D , a partition hyperplane is found to separate the categories of different samples. The linear equation of partition hyperplane is expressed as:

$$w^T x + b = 0 \tag{5}$$

Where w is the normal vector, which determines the direction of the hyperplane, and b is the displacement, which determines the distance between the hyperplane and the origin.

Supposing the hyperplane can classify the training samples correctly, for the training samples (x_i, y_i) , the following formula is satisfied:

$$\begin{cases} W^T x_i + b \geq +1 y_i = +1 \\ W^T x_i + b \leq -1 y_i = -1 \end{cases} \tag{6}$$

The formula is equivalent to:

$$y_i(W^T x_i + b) \geq +1 \tag{7}$$

The sample points which are closest to the hyperplane and meet the fomula of $y_i(W^T x_i + b) \geq +1$ are support vector.

From the above formula, we can get the interval:

$$\gamma = \frac{2}{\|W\|} \tag{8}$$

The idea of SVM is to maximize the interval, so the formula can be represented as:

$$\begin{aligned} & \max \frac{2}{\|W\|} \\ & s.t. y_i (W^T x_i + b) \geq +1 \end{aligned} \tag{9}$$

From the above formula, we can know that maximizing $\frac{2}{\|W\|}$ is equivalent to minimizing $\|W\|$, so the basic type of SVM can be expressed as:

$$\begin{aligned} & \min \frac{1}{2} \|W\|^2 \\ & s.t. y_i (w^T x_i + b) \geq 1, i = 1, 2, \dots, n \end{aligned} \tag{10}$$

The Lagrange multiplier method is used to solve the basic dual problem:

$$\begin{aligned} L(w, b, \alpha) = \min \frac{1}{2} \|W\|^2 - \sum_{i=1}^n \alpha^{(i)} [y^{(i)} (w^T \phi(x^{(i)}) + b) - 1] \\ s.t. y_i (w^T \phi(x^{(i)}) + b) \geq 1, i = 1, 2, \dots, n \end{aligned} \tag{11}$$

After deriving the w and b of the above formula, make equal to 0, and bring them into the Lagrange multiplier method, we can get:

$$\begin{aligned} L(w, b, \alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i x_j \\ s.t. \sum_{i=1}^n \alpha_i y_i = 0, \alpha_i \geq 0, i = 1, 2, \dots, n \end{aligned} \tag{12}$$

After solving the above problems, we can get the optimal classification function as follows:

$$f(x) = w^T x + b = \sum_{i=1}^n \alpha_i y_i x_i^T x + b \quad (13)$$

This WE and SVM combination belongs to traditional feature extraction + classifier combination. We do not use deep learning methods [30–38] because the small-size dataset.

3.4 10-Fold Cross-Validation

10-fold cross validation is a commonly used data accuracy test method in statistics. This method is to divide the data set into 10 groups as shown in Fig. 3, one group as the test set in turn, and the other nine groups as the training set for experiments. Each experiment will produce a result, and the average value of the 10 results is the accuracy value of the algorithm. In this study, 10-fold cross validation is used to verify the accuracy of image classification.

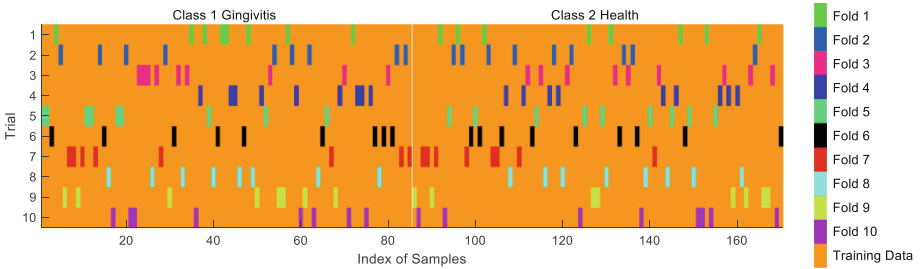


Fig. 3. Index of 10-fold cross validation

This 10-fold cross validation will be run 10 times, and we use those measures to performance the performance of proposed algorithm. There are six evaluation indexes in classification: sensitivity, specificity, accuracy, accuracy, F1 and MCC. Sensitivity is the proportion that the test is correctly recognized as positive; specificity is the proportion that the test is correctly recognized as negative. Precision is the ratio of the number of positive samples correctly predicted to the number of positive samples predicted. Accuracy represents the ratio of the number of correctly predicted samples to the total number of predicted samples. F1 is the harmonic average of precision and recall. MCC is essentially the phase between the observed value and the predicted value Relation number.

4 Experiment Results and Discussions

4.1 Statistical Results

In the first experiment, we set the decomposition level of WE as three, and the results are shown in Table 1. From Table 1, we can see that WE-SVM method has achieved good results as a whole. At the 6th run, the sensitivity, accuracy, F1 and MCC of the samples were all good, and the values were the highest; at the 8th run, the specificity and precision of the samples reached a peak with the values of 80.01% and 77.90% respectively, but the values of sensitivity, accuracy, F1 and MCC were relatively bad.

Table 1. Statistical analysis of our WE-SVM method.

Run	Sensitivity	Specificity	Precision	Accuracy	F1	MCC
1	72.98	78.82	77.49	75.90	75.14	51.91
2	77.63	72.95	74.14	75.29	75.85	50.64
3	72.90	76.47	75.55	74.71	74.17	49.44
4	76.47	71.76	73.03	74.11	74.71	48.28
5	77.63	70.60	72.51	74.12	74.98	48.35
6	77.66	76.47	76.74	77.06	77.20	54.13
7	72.95	77.63	76.55	75.29	74.70	50.65
8	70.57	80.01	77.90	75.29	74.05	50.80
9	77.63	71.73	73.38	74.71	75.42	49.50
10	75.30	76.44	76.23	75.88	75.76	51.76
Mean + SD	75.17 ± 2.49	75.29 ± 3.11	75.35 ± 1.85	75.24 ± 0.85	75.20 ± 0.88	50.55 ± 1.69

4.2 Optimal Decomposition

In this experiment, we compared the best decomposition levels. We suppose the decomposition level L vary from 1 to 4, and the corresponding results are shown below in Table 2.

Table 2. Results of different composition levels.

Decomposition level (L)	Sensitivity	Specificity	Precision	Accuracy	F1	MCC
1	71.41 ± 2.30	69.88 ± 2.53	70.39 ± 1.52	70.65 ± 1.16	70.85 ± 1.25	41.35 ± 2.32
2	73.52 ± 1.61	72.82 ± 1.93	73.04 ± 1.33	73.17 ± 1.02	73.26 ± 1.01	46.37 ± 2.06
3	75.17 ± 2.49	75.29 ± 3.11	75.35 ± 1.85	75.24 ± 0.85	75.20 ± 0.88	50.55 ± 1.69
4	73.52 ± 3.21	74.23 ± 4.28	74.22 ± 2.61	73.88 ± 1.40	73.77 ± 1.34	47.88 ± 2.78

(bold means the best)

In Table 2, we can see that when the value of L is 3, the values of sensitivity, specificity, precision, accuracy, F1 and MCC are the highest, so the effect of image at the third level is the best. we regard the third level as the best level of image decomposition.

4.3 Comparison with Other Classifiers

Decision tree (DT) and naive Bayesian classifier (NBC) are two classical classifiers. We compared using SVM with DT and NBC, at the condition of using 3-level decomposition. The results of DT and NBC are listed in Table 3. The different classifier comparisons are shown in Table 4.

From Table 3, it can be seen that DT and NBC are quite different in all aspects, of which the difference between the two methods in sensitivity and specificity is the most obvious. The highest value of DT is 78.82% and 76.47% respectively, NBC is 72.92% and 72.98%, DT is obviously superior to NBC. In Table 4, when DT, NBC and SVM are compared, SVM is superior to the other two methods in all aspects and the difference is large, which shows that SVM has obvious advantages as a classifier. So the SVM method used in this study is effective.

Table 3. Results using DT and NBC.

DT	Sensitivity	Specificity	Precision	Accuracy	F1	MCC
1	67.05	76.47	74.02	71.76	70.37	43.72
2	72.92	74.11	73.79	73.53	73.35	47.05
3	78.82	64.70	69.07	71.76	73.62	43.96
4	74.09	69.44	70.82	71.76	72.38	43.62
5	74.11	70.60	71.60	72.36	72.83	44.74
6	67.05	74.11	72.15	70.58	69.51	41.27
7	71.73	70.60	70.92	71.17	71.30	42.36
8	72.92	71.76	72.08	72.35	72.50	44.70
9	72.92	69.41	70.44	71.17	71.66	42.36
10	74.11	70.57	71.60	72.35	72.83	44.72
Mean + SD	72.57 ± 3.29	71.18 ± 3.08	71.65 ± 1.41	71.88 ± 0.78	72.04 ± 1.25	43.85 ± 1.54
NBC	Sensitivity	Specificity	Precision	Accuracy	F1	MCC
1	67.03	70.57	69.50	68.82	68.22	37.66
2	67.08	70.54	69.58	68.82	68.31	37.66
3	71.76	69.41	70.11	70.58	70.93	41.18
4	69.38	69.38	69.38	69.38	69.38	38.76
5	69.41	68.22	68.60	68.81	69.00	37.63
6	68.24	68.22	68.24	68.24	68.24	36.46
7	67.05	69.41	68.67	68.24	67.85	36.47
8	72.92	69.41	70.44	71.18	71.66	42.37
9	65.89	72.98	70.99	69.44	68.33	38.99
10	70.57	69.41	69.75	69.99	70.16	39.98
Mean + SD	68.93 ± 2.18	69.75 ± 1.31	69.53 ± 0.81	69.35 ± 0.93	69.21 ± 1.23	38.72 ± 1.86

Table 4. Comparison with different classifiers.

Classifier	Sensitivity	Specificity	Precision	Accuracy	F1	MCC
DT	72.57 ± 3.29	71.18 ± 3.08	71.65 ± 1.41	71.88 ± 0.78	72.04 ± 1.25	43.85 ± 1.54
NBC	68.93 ± 2.18	69.75 ± 1.31	69.53 ± 0.81	69.35 ± 0.93	69.21 ± 1.23	38.72 ± 1.86
SVM	75.17 ± 2.49	75.29 ± 3.11	75.35 ± 1.85	75.24 ± 0.85	75.20 ± 0.88	50.55 ± 1.69

4.4 Comparison to State-of-the-Art Approaches

We compared our WE-SVM method with three state-of-the-art approaches: WavEnr [4], ELM [5], GLCM [1]. The results are shown in Table 5. As shown in Table 5, when compared with the most advanced methods, we can find that the value of WE-SVM is higher than that of other methods in terms of sensitivity, specificity, precision and accuracy, which proves the effectiveness of WE-SVM and shows that WE-SVM is the optimal algorithm.

Table 5. Comparison with state-of-the-art approaches.

Approach	Sensitivity	Specificity	Precision	Accuracy
WavEnr [4]	62	68	65.96	65
ELM [5]	72	70	70.59	71
GLCM [1]	75	73	74	74
WE-SVM (Ours)	75.17 ± 2.49	75.29 ± 3.11	75.35 ± 1.85	75.24 ± 0.85

5 Conclusions

In this paper, a method of gingivitis classification based on wavelet entropy and support vector machine is proposed. By using this method, we can accurately classify the teeth pictures. Through the analysis of experimental data, WE-SVM can find that the sensitivity, specificity, accuracy and accuracy of the method are higher than 75%, which is more accurate and sensitive than the three most advanced methods.

However, due to the small number of samples in the database, it may cause over fitting phenomenon, which will be improved and avoided in future research, so that this research will be more helpful for dentists to carry out gingivitis testing.

References

1. Li, W.: A gingivitis identification method based on contrast-limited adaptive histogram equalization, gray-level co-occurrence matrix, and extreme learning machine. *Int. J. Imag. Syst. Technol.* **29**(1), 77–82 (2019)
2. Supranoto, S.C., et al.: The effect of chlorhexidine dentifrice or gel versus chlorhexidine mouthwash on plaque, gingivitis, bleeding and tooth discoloration: a systematic review. *Int. J. Dental Hygiene* **13**(2), 83–92 (2015)
3. Sangeetha, K.M.: Effect of triclosan containing tooth paste and conventional fluoride tooth paste on plaque and gingivitis: a randomized clinical trial. *Int. J. Sci. Study* **4**(10), 58–61 (2017)
4. Feng, C., et al.: Automated classification of brain MR images using wavelet-energy and support vector machines. In: Liu, C., Chang, G., Luo, Z. (eds.) *International Conference on Mechatronics, Electronic, Industrial and Control Engineering*, pp. 683–686. Atlantis Press, USA (2015)

5. Brown, M.: Gingivitis identification via grey-level cooccurrence matrix and extreme learning machine. *Adv. Soc. Sci. Educ. Humanit. Res.* **250**, 486–492 (2018)
6. Ramirez, J.: Unilateral sensorineural hearing loss identification based on double-density dual-tree complex wavelet transform and multinomial logistic regression. *Integr. Comput. Aided Eng.* (2019). <https://doi.org/10.3233/ICA-190605>
7. Gorriz, J.M.: Multivariate approach for Alzheimer's disease detection using stationary wavelet entropy and predator-prey particle swarm optimization. *J. Alzheimer's Disease* **65** (3), 855–869 (2018)
8. Li, Y.-J.: Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. *Multimedia Tools Appl.* **77**(9), 10393–10417 (2018)
9. Han, L., Identification of Alcoholism based on wavelet Renyi entropy and three-segment encoded Jaya algorithm. *Complexity*, **2018** (2018). Article ID. 3198184
10. Phillips, P.: Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm. *Neurocomputing* **272**, 668–676 (2018)
11. Farsi, C., et al.: Spectral triples and wavelets for higher-rank graphs. *J. Math. Anal. Appl.* **482**(2), 39 (2020). Article ID. 123572
12. Li, P., et al.: Pathological brain detection via wavelet packet tsallis entropy and real-coded biogeography-based optimization. *Fundamenta Informaticae* **151**(1–4), 275–291 (2017)
13. Li, Y.: Detection of dendritic spines using wavelet packet entropy and fuzzy support vector machine. *CNS Neurol. Disorders - Drug Targets* **16**(2), 116–121 (2017)
14. Nayak, D.R.: Detection of unilateral hearing loss by Stationary Wavelet Entropy. *CNS Neurol. Disorders - Drug Targets* **16**(2), 15–24 (2017)
15. Lu, H.M.: Facial emotion recognition based on biorthogonal wavelet entropy, fuzzy support vector machine, and stratified cross validation. *IEEE Access* **4**, 8375–8385 (2016)
16. Bertoluzza, S., et al.: Wavelets and convolution quadrature for the efficient solution of a 2D space-time BIE for the wave equation. *Appl. Math. Comput.* **366**, 21 (2020). Article ID. Unsp 124726
17. Gorriz, J.M., et al.: Wavelet entropy and directed acyclic graph support vector machine for detection of patients with unilateral hearing loss in MRI scanning. *Frontiers in Computational Neuroscience*, **10** (2016). Article ID. 160
18. Zhou, X.-X.: Comparison of machine learning methods for stationary wavelet entropy-based multiple sclerosis detection: decision tree, k-nearest neighbors, and support vector machine. *Simulation* **92**(9), 861–871 (2016)
19. Yang, M.: Dual-tree complex wavelet transform and twin support vector machine for pathological brain detection. *Appl. Sci.* **6**(6) (2016). Article ID. 169
20. Sun, P.: Preliminary research on abnormal brain detection by wavelet-energy and quantum-behaved PSO. *Technol. Health Care* **24**(s2), S641–S649 (2016)
21. Zhou, X., Zhang, G., Dong, Z., Wang, S., Zhang, Y.: Tea category classification based on feed-forward neural network and two-dimensional wavelet entropy. In: Xie, J., Chen, Z., Douglas, C.C., Zhang, W., Chen, Y. (eds.) *HPCA 2015*. LNCS, vol. 9576, pp. 48–54. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-32557-6_5
22. Mondal, H., et al.: Entropy generation of variable viscosity and thermal radiation on magneto nanofluid flow with dusty fluid. *J. Appl. Comput. Mech.* **6**(1), 171–182 (2020)
23. 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>
24. Cheng, H.: Multiple sclerosis identification based on fractional Fourier entropy and a modified Jaya algorithm. *Entropy*, **20**(4) (2018). Article ID. 254

25. Li, J.: Texture analysis method based on fractional fourier entropy and fitness-scaling adaptive genetic algorithm for detecting left-sided and right-sided sensorineural hearing loss. *Fundamenta Informaticae* **151**(1–4), 505–521 (2017)
26. Arahmane, H., et al.: Neutron-gamma discrimination based on support vector machine combined to nonnegative matrix factorization and continuous wavelet transform. *Measurement*, **149**, 9 (2020). Article ID. Unsp 106958
27. Chen, Y.: Pathological brain detection by wavelet-energy and fuzzy support vector machine. In: 8th International Symposium on Computational Intelligence and Design (ISCID), p. 409–412. IEEE, Hangzhou, China (2015)
28. Chen, M.: Morphological analysis of dendrites and spines by hybridization of ridge detection with twin support vector machine. *PeerJ*, **4** (2016). Article ID e2207
29. Liu, G.: Pathological brain detection in MRI scanning by wavelet packet Tsallis entropy and fuzzy support vector machine. *SpringerPlus*, **4**(1) (2015). Article ID. 716
30. Jiang, X.: Chinese sign language fingerspelling recognition via six-layer convolutional neural network with leaky rectified linear units for therapy and rehabilitation. *J. Med. Imag. Health Inform.* **9**(9), 2031–2038 (2019)
31. Govindaraj, V.V.: High performance multiple sclerosis classification by data augmentation and AlexNet transfer learning model. *J. Med. Imag. Health Inform.* **9**(9), 2012–2021 (2019)
32. Li, Z.: Teeth category classification via seven-layer deep convolutional neural network with max pooling and global average pooling. *Int. J. Imag. Syst. Technol.* (2019). <https://doi.org/10.1002/ima.22337>
33. Zhang, Y.D., Dong, Z., Chen, X., Jia, W., Du, Sidan, 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>
34. Wang, S.H., Sun, J., Phillips, P., Zhao, G., Zhang, Y.D.: Polarimetric synthetic aperture radar image segmentation by convolutional neural network using graphical processing units. *J. Real-Time Image Process.* **15**(3), 631–642 (2017). <https://doi.org/10.1007/s11554-017-0717-0>
35. Huang, C.: Multiple sclerosis identification by 14-layer convolutional neural network with batch normalization, dropout, and stochastic pooling. *Front. Neurosci.* **12** (2018). Article ID. 818
36. Pan, C.: Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU. *J. Comput. Sci.* **28**, 1–10 (2018)
37. Pan, C.: Abnormal breast identification by nine-layer convolutional neural network with parametric rectified linear unit and rank-based stochastic pooling. *J. Comput. Sci.* **27**, 57–68 (2018)
38. Zhang, Y.-D., Khan, M., 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>