



Chinese Fingerspelling Recognition via Hu Moment Invariant and RBF Support Vector Machine

Ya Gao¹, Ran Wang¹, Chen Xue¹, Yalan Gao¹, Yifei Qiao¹,
Chengchong Jia¹, and Xianwei Jiang^{1,2}(✉)

¹ School of Mathematics and Information Science, Nanjing Normal University
of Special Education, Nanjing 210038, China

jxw@njts.edu.cn

² Joint Accessibility Key Laboratory, China Disabled Persons' Federation,
Nanjing 210038, China

Abstract. Sign language plays a significant role in smooth communication between the hearing-impaired and the healthy. Chinese fingerspelling is an important composition of Chinese sign language, which is suitable for denoting terminology and using as basis of gesture sign language learning. We proposed a Chinese fingerspelling recognition approach via Hu moment invariant and RBF support vector machine. Hu moment invariant was employed to extract image feature and RBF-SVM was employed to classify. Meanwhile, 10-fold across validation was introduced to avoid overfitting. Our method HMI-RBF-SVM achieved overall accuracy of $86.47 \pm 1.15\%$ and was superior to three state-of-the-art approaches.

Keywords: Hu moment invariant · RBF · Support Vector Machine · Chinese fingerspelling recognition

1 Introduction

Around the world, hundreds of thousands of people suffer from hearing impairment [1]. Deafness is a problem that affects people's lives in many aspects because all communication between individuals is through language. Nowadays, continuous advances in mobile technology and new forms of user interaction have made it possible to overcome the communication problems between deaf and healthy people. Even so, the use of sign language is necessary. Sign language (SL) refers to the communication tool that people use in the deaf environment to express a specific meaning according to certain grammatical rules in accordance with the gestures, movements, positions and orientations of their hands. Chinese sign language (CSL) can generally be divided into gesture sign language and fingerspelling language [2]. Gesture sign language is a major supplement of spoken language and plays an auxiliary role in communicating information and feelings with the outside world. Gesture sign language often uses both hands at the same time, supplemented by facial expressions, three main features were extracted from the gesture signature gestures, namely the position of the hand, the

direction of the hand signature, and the shape of the hand [3]. It emphasizes the situation, relatively complex. While the latter fingerspelling language emphasizes 30 basic finger languages, including 26 basic pinyin letters and 4 raised tongues, and expresses pinyin or some special meanings through their combination. Fingerspelling language is relatively simple and definite. In addition, the origin of gesture sign language is early, which has experienced three main stages, namely the germination of ancient times, the disunity of modern times and the perfection and unification of modern times. Fingerspelling language originated from the west, through the introduction and learning, gradually realize the Chinese practice. Currently, refinement and standardization are two basic tasks in the development of Chinese sign language.

Sign language recognition (SLR) is the use of computer technology to translate sign language information into text, natural language, audio and other information to facilitate understanding and communication [4]. Moreover, sign language recognition is a significant part of intelligent human-computer interaction. According to the data output, SLR technology can be divided into two categories: sensor-based sign language recognition technology and computer vision-based sign language recognition technology. The former employs wearable devices such as data gloves and EMG signal arms, while the latter mainly adopts depth cameras. Due to practical factors, sign language recognition based on computer vision is more popular. The features commonly used in sign language recognition are hand shape, direction, position and movement track. Most experimental studies have revolved around these features.

There are many methods of sign language recognition, one of which utilizes statistical analysis techniques to derive the various eigenvectors of a sample and then classify it. Hidden markov model (HMM) [5] is a typical representative of this. The second method employs template matching technology, that is, first construct a defined template, then match the original data with the template, and use the similarity as a reference to complete the identification. Another method uses the self-learning and organizational functions of the network to build on a new neural network. However, these methods have shortcomings. HMM relies on complex initialization process and huge computation workload. It requires the existence of successfully detected gesture region and gesture movement, which determines the robustness of the algorithm. The accuracy of traditional template matching methods is not satisfactory. New neural network techniques require large amounts of data and consume more training sessions. To solve these problems, a Chinese sign language recognition method based on wavelet entropy and support vector machine was proposed [6]. Gray-level co-occurrence matrix (GLCM) was also employed to identify sign language [7]. In addition, many experts try to use the depth information of gestures to solve the issue of recognition accuracy. Taiwan sign language recognition system based on Kinect extracted three main features from sign language gestures, namely gesture position, gesture direction and gesture shape [8]. The position of the hand was obtained by input sensor. HMM was introduced to determine the direction of the gesture, and a trained SVM was employed to identify the shape of the gesture, the recognition rate of this system reached 85.14% [9]. Convolutional neural network (CNN) and its variant were also tested to identify fingerspelling [10]. For the problem of hand posture recognition,

hand contour is usually adopted because of the simplification of real-time. All of the current related work can be divided globally into software-based and hardware-based solutions.

In this study, a suitable method for Chinese fingerspelling recognition via Hu moment invariant and RBF support vector machine was proposed. Hu moment invariant was employed to extract the image features and accelerate the training. RBF was employed to enhance classification due to its effective performance. Meanwhile, the experiment was carried out on 10-fold cross-validation to prevent overfitting.

2 Method

2.1 Subjects and Dataset

We acquired 1320 Chinese fingerspelling language samples to construct experiment materials, which were from 44 volunteers. Following 30 categories in Chinese fingerspelling language, every volunteer simulates action one time, thus totally 1320 images were gained. All these images were preprocessed by software and normalized to size of 256×256 . Figure 1 showed a sample from one volunteer.



Fig. 1. A sample of Chinese fingerspelling language from one volunteer

2.2 Hu Moment Invariant

Hu moment was proposed by Hu. M.K in 1962, which has translation, rotation and scale invariance. For images with gray distribution of $f(x, y)$, the $(p + q)$ order invariant moment is defined as:

$$M_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad p, q = 0, 1, 2, \dots \quad (1)$$

The central moment of order $(p + q)$ is:

$$C_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - x_1)^p (y - y_1)^q f(x, y) dx dy \quad p, q = 0, 1, 2, \dots \quad (2)$$

where moment center is (x_1, y_1) , and centroid of the component is as follows:

$$x_1 = M_{10}/M_{00} \quad y_1 = M_{01}/M_{00}$$

For the digital image, discretization means that in the discrete state, the formula of $(p + q)$ order ordinary moment and central moment of $f(x, y)$ is as follows:

$$M_{pq} = \sum \sum x^p y^q f(x, y) \quad (3)$$

$$C_{pq} = \sum \sum (x - x_1)^p (y - y_1)^q f(x, y) \quad p, q = 0, 1, 2, \dots \quad (4)$$

When the image changes, M_{pq} also changes. Although C_{pq} has translation invariance, it is still sensitive to rotation [11–13]. To address these problems, we normalize the central moment,

$$N_{pq} = C_{pq}/C_{00}^{1+p+q/2}, \quad p + q = 2, 3, 4 \dots \quad (5)$$

Directly represented by ordinary moments or central moments, features can't be translated, rotated, bonus scaled invariant at the same time. If the normalized center distance is used, the feature not only has translation invariance but also proportional invariance.

The Hu moment constructs seven invariant moments using the second and third order center distances. They keep the translation, zooming and rotation unchanged in consecutive pixels, making the image more efficient to optimize. This is the seven invariant moments of the Hu moment:

$$I_1 = N_{20}N_{02} \quad (6)$$

$$I_2 = (N_{20} + N_{02})^2 = 4N_{11}^2 \quad (7)$$

$$I_3 = (N_{30} + N_{12})^2 + (3N_{21} - N_{03})^2 \quad (8)$$

$$I_4 = (N_{30} + N_{12})^2 + (N_{21} - N_{03})^2 \quad (9)$$

$$I_5 = (N_{30} - 3N_{12})(N_{30} + N_{12}) \left[(N_{30} + N_{12})^2 - 3(N_{21} + N_{03})^2 \right] + (3N_{21} - N_{03})(N_{21} + N_{03}) \left[3(N_{30} + 3N_{12})^2 - (N_{21} + N_{03})^2 \right] \quad (10)$$

$$I_6 = (N_{20} - N_{02}) \left[(N_{30} + N_{12})^2 - (N_{21} + N_{03})^2 \right] + 4N_{11}(N_{30} + N_{12})(N_{21} + N_{03}) \quad (11)$$

$$I_7 = (3N_{21} - N_{03})(N_{30} + N_{12}) \left[(N_{30} + N_{12})^2 - 3(N_{21} + N_{03})^2 \right] - (N_{30} - 3N_{12})(N_{21} + N_{03}) \left[3(N_{30} + N_{12})^2 - (N_{21} + N_{03})^2 \right] \tag{12}$$

In short, the Hu moment is the center (x_1, y_1) of the image obtained from the ordinary moment M_{pq} , and then defines the center distance C_{pq} , normalizes the center distance to obtain the normalized center distance N_{pq} . In addition, together with the normalized second-order moment of the center distance and the third moments, seven invariant moments are formed.

2.3 Support Vector Machine

Among supervised classification methods, Support Vector Machine (SVM) [14] is the latest classification method based on machine learning theory. But it is a linear classifier which just uses one hyperplane to divide samples in D-dimension into two categories, thus it has a poor fitting effect [15–19]. For the purpose of improving the fitting effect, kernel-SVM was proposed. It is an innovation of SVM which can solve the problem of the disability of separating the practical data with complex distribution in supervised learning by combining several hyperplanes. Kernel-SVM extends original linear SVM to nonlinear SVM classifier which represents many machine learning algorithms by using the dot product between samples [20–24].

Based on the above theory, the linear function of SVM can be redefined as follows:

$$w^T x + m = m + \sum_{i=1}^n \gamma_i x^T x^{(i)} \tag{13}$$

Here, γ_i denotes coefficient vector, $x^{(i)}$ indicates training sample. At the same time, we can substitute $\varphi(x)$ with the output of the eigenfunction x , and introduce a kernel function to replace the dot product. The equation is as follows:

$$k(x, x^{(i)}) = \varphi(x) \cdot \varphi(x^{(i)}) \tag{14}$$

Where \cdot denotes the dot product. Thus, we can use the substitution function to predict.

$$f(x) = m + \sum_i \gamma_i k(x, x^{(i)}) \tag{15}$$

The function $f(x)$ is nonlinear and it is completely equivalent to preprocessing inputs using $\varphi(x)$ and then learning a linear model in the new transformation space. After transformation, the kernel-SVM is allowed to fit the maximum-margin hyperplane in a transformed feature space. The transformation can be nonlinear and the transformed space can be high dimensional, but after a series of transformations and replacements, although the classifier is a hyperplane in the high-dimensional feature space, it may be nonlinear in the original input space.

Based on the advantages that kernel trick has, it guarantees the effective convergence of optimization techniques for learning nonlinear models and the implementation of kernel function. $k(x, x^{(i)})$ is much more efficient than constructing $\varphi(x)$ and then computing the dot product, so we introduced the kernel function into the kernel-SVM.

2.4 Radial Basis Function

Among all of kernel functions, Gaussian kernel is the most popular used kernel which usually be utilized to solve two-dimensional time-independent convection-diffusive-reaction equations with inhomogeneous boundary conditions. It is also called as Radial Basis Function (RBF) [25], which can be expressed as follow:

$$k_G(x, x^{(i)}) = T(x - x^{(i)}; 0; \sigma^2 I) \quad (16)$$

Here T represents standard normal density. RBF means that its value decreased in the direction in which $x^{(i)}$ radiates outward from x . RBF also can be defined as the following formula in details.

$$k_G(x, x^{(i)}) = \exp\left(-\beta \|x - x^{(i)}\|^2\right) \quad (17)$$

Where β is a parameter which needs to be tuned.

To conclude, kernel-SVM has a plenty of advantages which are summarized as follows: (1) it guarantees the effective convergence of optimization techniques for learning nonlinear models; (2) the implementation of kernel function is much more effective than constructing and computing the dot product; (3) it has the faster speed to training samples and classifying samples, besides, it's memory requirement is relatively low; (4) it has less adjustable parameters; (5) convex quadratic optimization is utilized in training, which provides global and unique solution and avoids convergence to the local minimum. We do not use deep learning in this paper, since our dataset is small and not suitable for training a deep neural network model [26–30].

2.5 Implementation

We carried out the experiment on a platform of the personal computer installed Windows 10 operating system with Core i7 CPU and 16 GB memory. Overall accuracy (OA) was introduced to evaluate the results, which denotes the value of correct prediction over all test sets divide the total numbers.

In order to obtain the parameters of top optimal kernel-SVM, trial-and-error algorithm is commonly used. Meanwhile, 10-fold cross validation [31] is introduced to validate the performance. Cross validation is a practical method for cutting data samples into smaller subsets. In the 10-fold cross validation, the obtained data samples are randomly divided into 10 equal-sized subsamples. Among the 10 sub-samples, one sample is retained as validation data for testing the model, and the remaining 9 samples are used as training data. Then, the cross validation was repeated 10 times, and each sample was used as a test sample. The average of the 10 results is used as an estimate of

the accuracy of the algorithm. (See Fig. 2) The advantage of 10-fold cross validation is that all data samples are trained and verified, and the subset is verified only once, and the error is greatly reduced during the experiment.



Fig. 2. An illustration of 10-fold cross validation

3 Experiment and Results

3.1 Statistical Results of Proposed Method

In this experiment, we utilized this proposed method, called HMI-RBF-SVM, in which Hu moment invariant and radial basis function support vector machine were employed. The results of 10 × 10-fold cross validation are listed in Table 1. It can be observed that the highest accuracy in column Total (Overall Accuracy) achieves 88.79% that has been highlighted with bold font. Another marked number represents the highest accuracy of single run, which reaches 93.18%. Finally, the value of means and standard deviation based on 10 runs is 86.47 ± 1.15%, which indicates the results are effective and accepted.

Table 1. Results of 10 × 10-fold cross validation

Run	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Total
1	85.07	88.37	87.59	83.08	86.15	87.88	87.79	84.85	86.67	86.92	86.44
2	85.71	81.82	86.57	79.55	86.03	88.97	84.50	89.23	83.70	83.08	84.92
3	83.72	81.10	88.06	86.36	84.62	88.37	87.22	84.78	85.19	86.47	85.61
4	83.85	87.50	85.61	91.67	88.15	93.18	88.72	89.15	89.31	90.77	88.79
5	79.10	86.76	84.13	88.15	85.50	80.80	87.88	88.81	87.97	83.58	85.30
6	83.85	85.07	84.21	86.05	82.31	86.26	85.07	88.81	90.91	87.22	85.98
7	85.19	85.83	85.38	85.38	85.94	87.41	86.13	87.97	89.15	86.76	86.52
8	87.22	89.31	90.98	83.46	82.68	83.70	90.15	90.91	89.47	89.31	87.73
9	87.79	87.88	84.33	88.81	86.67	87.31	89.15	84.85	84.96	85.71	86.74
10	87.41	89.06	84.09	88.49	82.95	87.31	87.88	87.60	84.96	86.82	86.67
Mean ± SD	86.47 ± 1.15										

3.2 SVM Versus RBF-SVM

We compared RBF-SVM with traditional SVM and polynomial SVM. All the parameters were chosen by trial-and-error method. As can be seen from Table 2, Mean \pm SD of traditional SVM, polynomial SVM (PSVM) and RBF-SVM are $82.39 \pm 1.13\%$, $85.50 \pm 1.09\%$, and $86.47 \pm 1.15\%$, respectively. In term of accuracy, RBF-SVM gains about four percentage points ahead of traditional SVM and also gets about one point ahead of PSVM. It denotes that the higher accuracy is benefit from radial basis function which guarantees the effective convergence and speeds up the training.

Table 2. Comparison with other SVM variants

Run	SVM	PSVM	RBF-SVM (Ours)
1	83.18	86.06	86.44
2	82.42	87.12	84.92
3	83.71	85.45	85.61
4	82.27	85.76	88.79
5	80.00	86.06	85.30
6	82.35	83.64	85.98
7	82.58	83.71	86.52
8	81.59	85.15	87.73
9	83.94	85.91	86.74
10	81.89	86.14	86.67
Mean \pm SD	82.39 ± 1.13	85.50 ± 1.09	86.47 ± 1.15

3.3 Comparison of State-of-the-Art Approaches

In this experiment, three state-of-the-art approaches were compared with our HMI-RBF-SVM method. As can be seen from Table 3, WE-SVM [6], GLCM-MGSVM [7], and HMM-SVM [9] achieved overall accuracy of $85.69 \pm 0.59\%$, 85.3% and 85.14% , respectively. Different feature extraction methods, such as wavelet entropy, gray-level co-occurrence matrix, hidden markov model and Hu moment invariant were employed in these four approaches. Kernel SVM was applied in GLCM-MGSVM and HMI-RBF-SVM, and another two approaches utilized traditional SVM. Our method is superior others about one percentage point, which indicates that Hu moment invariant maintains image feature when extracting and radial basis function provides the effective convergence to enhance performance of classification.

Table 3. Comparison of four state-of-the-art approaches

Approach	Overall accuracy
WE-SVM [6]	85.69 \pm 0.59
GLCM-MGSVM [7]	85.3
HMM-SVM [9]	85.14
HMI-RBF-SVM (Ours)	86.47 \pm 1.15

4 Discussions

In our study, three advanced techniques: Hu moment invariant, RBF-SVM and 10-fold cross validation were applied. As the picture is identified by the feature quantity composed of the Hu moment invariant, the advantage is that the speed can be accelerated, and the disadvantage is that the recognition rate may drop relatively. Therefore, Hu moments invariant are generally used to identify large objects in an image. It describes the shape of the object better, but the texture features of the image cannot be too complicated. Thus, it will be relatively suitable for identifying the hand shape preprocessed. Due to excellent performance of RBF, the Gaussian kernel was chosen. It needs few parameters to tune and employs convex quadratic optimization to train. Particularly, kernel SVM provides unique and global solutions, preventing the convergence to local minima. When compared with traditional SVM, its superiority is highlighted distinctly. The application of 10-fold cross-validation is determined by its easy-to-use properties and the use of all data for training and validation, which will not improve the accuracy of the final classification, but will make the classifier reliable, so it can be generalized to other independent data sets. All these techniques contribute to improving of performance.

5 Conclusions

This study proposed a novel Chinese fingerspelling recognition via Hu moment invariant and RBF-SVM, carrying out on 10-fold cross validation. Hu moment invariant adapts to the feature traction of fingerspelling images, which accelerates the speed of recognition. Based on radial basis function, kernel SVM guarantees the effective convergence and improves the classification. 10-fold cross validation provides the sufficient confirmation of data samples, reducing the error. This approach gained overall accuracy of 86.47 \pm 1.15%, which indicates it has superiority in all four state-of-the-art approaches.

In the future, some contributions shall be tried as follows: (1) other advanced methods such as particle swarm optimization, principal component analysis, deep neural network and transfer learning may be applied in this theme. (2) We shall try to apply our approach to other fields.

Acknowledgement. This work was supported from Jiangsu Overseas Visiting Scholar Program for University Prominent Young & Middle-aged Teachers and Presidents of China, The Natural Science Foundation of Jiangsu Higher Education Institutions of China (19KJA310002), The Philosophy and Social Science Research Foundation Project of Universities of Jiangsu Province (2017SJB0668), The Natural Science Foundation of Jiangsu Province (16KJB520029).

References

1. Maharani, A., et al.: Hearing impairment, loneliness, social isolation, and cognitive function: longitudinal analysis using english longitudinal study on ageing. *Am. J. Geriatr. Psych.* **27** (12), 1348–1356 (2019)
2. Eifring, H.: Language contact across time: classical chinese on modern public signs. *J. Chin. Linguist.* **47**(2), 562–614 (2019)
3. Deriche, M., et al.: An intelligent arabic sign language recognition system using a pair of LMCs with GMM based classification. *IEEE Sens. J.* **19**(18), 8067–8078 (2019)
4. Kumar, E.K., et al.: 3D sign language recognition with joint distance and angular coded color topographical descriptor on a 2-stream CNN. *Neurocomputing* **372**, 40–54 (2020)
5. El Afia, A., et al.: Adaptive cooperation of multi-swarm particle swarm optimizer-based hidden Markov model. *Progress Artif. Intell.* **8**(4), 441–452 (2019)
6. Jiang, X., Zhu, Z.: Chinese sign language identification via wavelet entropy and support vector machine. In: Li, J., Wang, S., Qin, S., Li, X., Wang, S. (eds.) ADMA 2019. LNCS (LNAD), vol. 11888, pp. 726–736. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-35231-8_53
7. Jiang, X.: Isolated Chinese sign language recognition using gray-level co-occurrence matrix and parameter-optimized medium gaussian support vector machine. In: Satapathy, S.C., Bhateja, V., Nguyen, B.L., Nguyen, N.G., Le, D.-N. (eds.) *Frontiers in Intelligent Computing: Theory and Applications. AISC*, vol. 1014, pp. 182–193. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-9920-6_19
8. Kumar, P., et al.: A position and rotation invariant framework for sign language recognition (SLR) using Kinect. *Multimedia Tools Appl.* **77**(7), 8823–8846 (2018)
9. Lee, G.C., et al.: Kinect-based Taiwanese sign-language recognition system. *Multimed Tools Appl.* **75**, 261–279 (2016)
10. Jiang, X.: Chinese sign language fingerspelling recognition via six-layer convolutional neural network with leaky rectified linear units for therapy and rehabilitation. *J. Med. Imaging Health Inform.* **9**(9), 2031–2038 (2019)
11. Sun, P.: Pathological brain detection based on wavelet entropy and Hu moment invariants. *Bio-Med. Mater. Eng.* **26**(s1), 1283–1290 (2015)
12. Yang, J.: Pathological brain detection in MRI scanning via Hu moment invariants and machine learning. *J. Exp. Theor. Artif. Intell.* **29**(2), 299–312 (2017)
13. Hou, X.-X.: Alcoholism detection by medical robots based on Hu moment invariants and predator-prey adaptive-inertia chaotic particle swarm optimization. *Comput. Electr. Eng.* **63**, 126–138 (2017)
14. Tharwat, A.: Parameter investigation of support vector machine classifier with kernel functions. *Knowl. Inf. Syst.* **61**(3), 1269–1302 (2019)
15. 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)

16. Chen, S., et al.: Magnetic resonance brain image classification based on weighted-type fractional Fourier transform and nonparallel support vector machine. *Int. J. Imaging Syst. Technol.* **25**(4), 317–327 (2015)
17. 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
18. Yang, M.: Dual-tree complex wavelet transform and twin support vector machine for pathological brain detection. *Appl. Sci.* **6**(6) (2016). Article ID. 169
19. Chen, M.: Morphological analysis of dendrites and spines by hybridization of ridge detection with twin support vector machine. *PeerJ* **4** (2016). Article ID. e2207
20. 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)
21. 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. *Front. Comput. Neurosci.* **10** (2016). Article ID. 160
22. 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)
23. 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)
24. Chen, Y.: Pathological brain detection by wavelet-energy and fuzzy support vector machine. In: 8th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, pp. 409–412. IEEE (2015)
25. De Mulder, W., et al.: An interpretation of radial basis function networks as zero-mean Gaussian process emulators in cluster space. *J. Comput. Appl. Math.* **363**, 249–255 (2020)
26. Li, Z.: Teeth category classification via seven-layer deep convolutional neural network with max pooling and global average pooling. *Int. J. Imaging Syst. Technol.* (2019). <https://doi.org/10.1002/ima.22337>
27. Muhammad, K.: Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. *Multimedia Tools and Applications* **78**(3), 3613–3632 (2019)
28. Sangaiah, A.K.: Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization. *Neural Comput. Appl.* (2019). <https://doi.org/10.1007/s00521-018-3924-0>
29. Zhao, G.: Polarimetric synthetic aperture radar image segmentation by convolutional neural network using graphical processing units. *J. Real-Time Image Proc.* **15**(3), 631–642 (2018)
30. 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
31. Gilanie, G., et al.: Automated and reliable brain radiology with texture analysis of magnetic resonance imaging and cross datasets validation. *Int. J. Imaging Syst. Technol.* **29**(4), 531–538 (2019)