Multi Features based Fruit Classification Using different Classifiers

D.O.I - 10.51201/12503
https://doi.org/10.51201/12503

[1] Dr. Suryakanth B Ummapure[*] , [2] Shashikiran M Hanchinal

[1]Department of Computer Science ,Govt. First Grade College Shahapur, Karnataka, India

[2]Department of Computer Science, Govt. College Autonomous Kalaburagi, Karnataka, India

Abstract: *The fruits available naturally will be having different colors and shape in appearance. Humans can identify the type of fruit by seeing their shape and color without any difficulty. Here a practical approach has been offered in this paper to classify the fruit images based on the Color and Shape of the fruit. Five thousand images were taken from the standard Fruit-360 dataset for the experiment; the dataset contains Apple, Banana, Cherry, Grapes, and Mango. The color_moment and shape of the fruits were considered to extract the features from different fruit images. In this proposed work, three feature vectors are constructed. In the color_moment feature extraction, here statistical features such as mean and standard deviation of three-color channels (RGB) are computed. The binarized images of fruits were used to extract shape-based features, and a multifeatured vector consisting of color_moment and shape features were used. The SVM, MLP, and RF classifiers are used for the classification process. The recognition accuracy of 99.98% has been achieved using the combined feature vector (multifeatured vector) and RF classifier. This paper's contribution is that the color_moments feature extraction is carried out directly on fruit images without using any pre-processing techniques such as gray-scale and binary conversion on fruit images.*

Keywords: SVM, MLP, RF, Shape features, Color moments.

I. Introduction

Since ancient days, the fruit has been an essential food for humans. Fruits are an excellent source of essential vitamins and minerals, and they are high in fiber. Fruits provide a wide range of health-boosting antioxidants and flavonoids.

The fruit attributes can benefit various fields, particularly in fruit products' mass processing, with a system classifying different fruit and identification[1]. Fruits are a well-known

Corresponding Author: [*]Dr. Suryakanth B Ummapure

source of vitamins, vitamin C, and vitamin A deficiencies recognized as their function to avoid. People who use fruits and vegetables as part of a balanced eating routine are at a decreased risk of certain chronic diseases. Images are an essential source of knowledge and data in the agricultural sciences. Hence, the need arises to classify the fruits for the further process.

Without having prior knowledge, one cannot proceed with further work. Hence, any researcher needs to research a particular research area further. A brief overview of the methods and techniques available in the literature for fruit classification is described below.

Various fruit extraction and fruit classification algorithms were developed and tested [2]. Here a few papers are found in the literature related to the fruit classification problem [1] have reported the fruit types identification by employing the k-NN with SVM classifiers, they have extracted the second-order statistical features and achieved 96.3%, 93.8%, 25%, 83.8%, 90%, and 95% identification accuracy, respectively. The performances of fruit images using various classifiers are reported by [3], considering the area, color, equvidiameter, centroid, zone, perimeter, size, and roundness features. The random forest has secured the highest of 70.8, 75%, and 83.33%, among other classifiers. The CNN based fruit classification is presented by [4], they have considered the CNN features sets, and they have reported the 99.79% recognition accuracy. [5], have given the fruit classification based on the random forest algorithm, they have considered the apples, Strawberry, and oranges fruits. To extract the features, they have considered the Color, Shape, and SIFT after extracting these features fed to the Random Forest machine learning algorithm to obtain the recognition accuracy. Combining many features and classifiers based on fruit and vegetable classification from images has been presented in [6]. A review has been made to classify fruits by using the image processing methods given in [7]. The color features based apple fruit recognition system was proposed by in [8], they have captured the fruit along with the tree, and they have developed the system to early predicting models of apple fruit. In [9], apple fruit disease identification work has been carried out using the K-means color image segmentation technique. The orange fruit classification has been proposed in [10], the experiment has used the five data mining algorithms; J48, CART, BFT, LMT. The fruit grading system based on the machine vision and review have given in [10], they have comprehensively discussed the texture features; LBP, SURF, HOG, and so on with ANN and

CNN. The work proposed in [11], have based on south Indian fruit automatic classification, has been reported, and they employed the DTNB classifier with statistical analysis.

In this proposed work, three feature vectors are constructed. The color_moment feature extraction, here statistical features such as mean and standard deviation of three-color channels (RGB) are computed, the binarized images of fruits were used to extract shape-based features and a multifeatured vector consisting of color_moment and shape features were used and the rest of the paper is described as follows. The methodology is explained in section-II. Experimental results are described in section-III, and the conclusion is given in section-IV.

II. Methodology

Very few attempts have been made in the literature to classify the fruits based on color, Shape, and combination of color and Shape. Given this, the proposed method is planned towards a practical approach to classifying the fruit images based on color, Shape of the fruit, and the concept of feature hybridization. Five thousand images were considered for the experimentation collected from the standard Fruit-360 dataset [https: //www. kaggle.com /moltean/fruits].The dataset consisting of images of five different groups, namely Apple, Banana, Cherry, Grapes, and Mango, are considered for the experimentation. The color_moment and shape of the fruits were considered to extract the features from different fruit images. The purpose of this proposed work is to build a simple automatic system for fruit classification that recognizes the fruit from a set of images.

For fruit classification, three different approaches were considered viz., Color_moment feature extraction, Shape feature extraction, and fruits classification using multiple features. To extract the color moments features [12][13][14], are treated as useful components for this proposed experiment. Color is the most significant feature to make recognition of the fruit type. This is based on the color, one can easily recognize the type of fruit. The color_moment feature extraction is simple, and the state-of-the work method is used. The statistical standard deviation and mean features were extracted from the input image. To extract these features, the RGB color components were extracted from the color image of fruits. One each color that is for Red channel mean and standard deviation, Green channel mean and standard deviation, and Blue mean and standard deviation are computed. By this the features vector of size 6 is constructed.

The shape is the essential feature of any object for its recognition visually. A blind person can also recognize the type of fruit by holding in hand based on shape like Apple, Banana, Cherry, etc. Hence, the Shape-based feature approach can be used to recognize any type of fruit like banana and Apple, which are dissimilar in shape. The shape features are extracted [15][16] from the fruit images. To extract the shape features, the binarization process is incorporated here. The input image is converted to a binary image. From the set of binary images, area, eccentricity, extent, orientation, and perimeter-based shape feature are extracted to create a feature vector of size 5.

The fruit classification can be carried out by using the combination of different features, i.e., two discriminative features like color and shape, which can be combined into one feature vector. The combined feature vector is outperformed in recognition accuracy. To enhance the recognition accuracy, the color_moment feature vector and shape feature vectors are combined to construct a single feature vector of size 11, finally submitted to classifiers to classify fruit images. MLP, SVM, and RF classifiers classify the three feature vectors viz., Color_moment, Shape and Combined feature vectors.

For the classification of fruit images, three classifiers are used here i.e., SVM, MLP and RF classifiers. The Weka tool has been employed to get classification results. From the weka tool we have considered SMO-based SVM, MLP, and RF classification algorithms.

The SVM classifier [25] is a widely used supervised learning algorithm that implements the sequential minimal optimization algorithm from J. Platt [25] to train a support vector classifier. This implementation substitutes all missing values globally and translates nominal attributes into binary ones. It classifies the binary class problem (one to one). In this proposed work, the multiclass problem is undertaken, and it works in (One to many). Multilayer Perceptron (MLP) is a feed-forward ANN class with three layers, i.e., an input layer, an output layer, and a hidden layer. This classifier uses backpropagation to learn to classify instances with a multilayer perceptron [26]. Using a simple heuristic, the network can be constructed. During training time, the network parameters can also be controlled and updated. The nodes in this network are all sigmoid. Random Forest (RF), as its name suggests Random Forest consists of a large number of individual decision trees that act as an ensemble [27] [28]. Any single tree in the random forest spits out a class prediction, and the class with the most votes becomes the prediction of the model.

III. Results and Discussion

This section performs features and classifiers. The following tables and figure shows the detailed results. The fruit recognition system is challenging for computer vision, which can achieve human-level recognition [17]. For the experimentation, the fruits-360 standard dataset is considered [4][18]. Following figure 1 shows the sample images of fruits.

Figure 1: Sample fruit images

Here Table 1 represents the recognition accuracy for the five-fruit classification using Support Vector Machine (SVM), Random Forest (RF), and Multilayer Perceptron (MLP).

Table 1: Average recognition of the five types of fruits using color_moment features

Color_moment Features (6 Features)					
Sl. No.	Classifier	Rec. Acc.	Precision	Recall	F-Measure
1	SVM	85.62%	0.859	0.856	0.854
2	MLP	98.30%	0.984	0.983	0.983
3	RF	99.94%	0.999	0.999	0.999

Table 2: Confusion matrix for the five types of fruits using SVM based on color_moment Features

Confusion matrix for SVM					
Fruit Type	Apple	Banana	Cherry	Grapes	Mango
Apple	985	0	11	0	4
Banana	1	999	0	0	0
Cherry	132	0	652	50	166
Grapes	0	0	122	763	115
Mango	2	0	114	2	882

Table 3: Confusion matrix for the five types of fruits using MLP based on color_moment features

Confusion matrix for MLP					
Fruit Type	Apple	Banana	Cherry	Grapes	Mango
Apple	1000	0	0	0	0
Banana	0	1000	0	0	0
Cherry	0	0	999	0	1
Grapes	35	0	46	916	3
Mango	0	0	0	0	1000

Table 4: Confusion matrix for the five types of fruits using RF-based on color_moment features

Confusion matrix for RF					
Fruit Type	Apple	Banana	Cherry	Grapes	Mango
Apple	1000	0	0	0	0
Banana	0	1000	0	0	0
Cherry	0	0	999	0	1
Grapes	0	0	0	998	2
Mango	0	0	0	0	1000

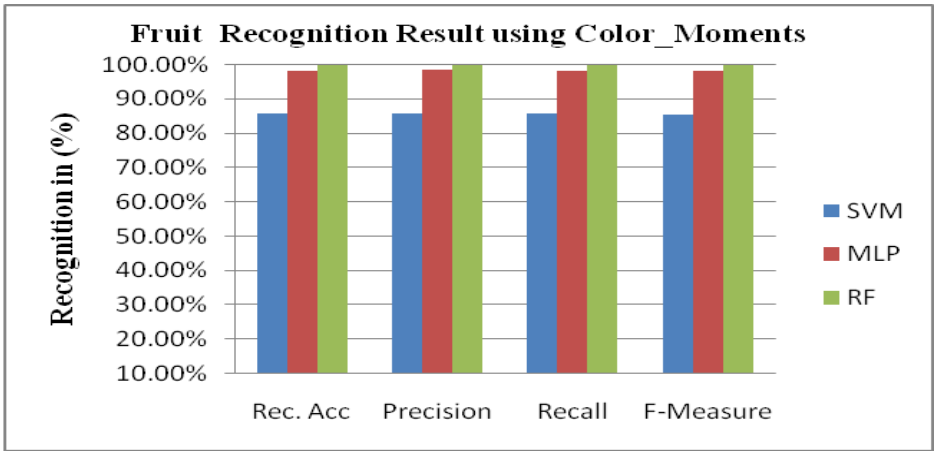

Figure 2: Average Recognition Result based on SVM, MLP, and RF classifiers using color_moments features.

From the above tables and figure, it is seen that the Random forest classifier has performed well as compare to other classifiers by giving the recognition accuracy of 99.94%, Precision of 0.999, Recall 0.999, and F-Measure as 0.999.

Table 5: Average recognition of the five types of fruits using color_moments features.

Shape Features (5 Features)					
Sl. No.	Classifier	Rec. Acc	Precision	Recall	F-Measure
1	SVM	51.54%	0.525	0.515	0.474
2	MLP	62.08%	0.648	0.621	0.617
3	RF	91.36%	0.914	0.914	0.914

Table 6: Confusion matrix for SVM classifier.

Confusion matrix for SVM					
Fruit Type	Apple	Banana	Cherry	Grapes	Mango
Apple	315	137	49	144	355
Banana	65	874	36	3	22
Cherry	187	22	116	54	621
Grapes	133	5	117	304	441
Mango	0	0	20	12	968

Table 7: Confusion matrix for MLP classifier.

Confusion matrix for MLP					
Fruit Type	Apple	Banana	Cherry	Grapes	Mango
Apple	408	9	180	215	188
Banana	47	898	20	21	14
Cherry	25	15	494	113	353
Grapes	82	8	167	475	268
Mango	6	0	133	32	829

Table 8: Confusion matrix for RF classifier.

Confusion matrix for RF					
Fruit Type	Apple	Banana	Cherry	Grapes	Mango
Apple	899	5	30	29	37
Banana	15	978	0	3	4
Cherry	25	4	908	20	43
Grapes	43	3	39	863	52
Mango	15	0	32	33	920

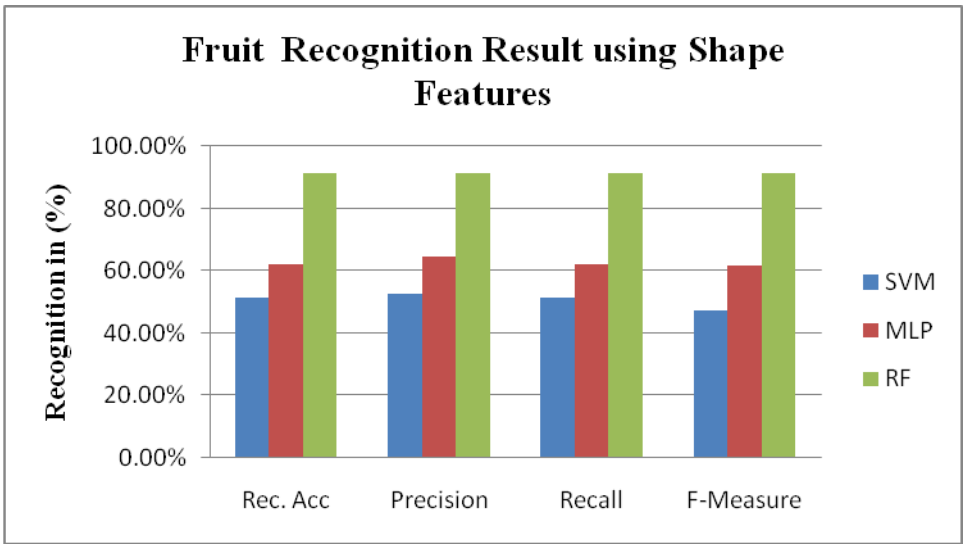

Figure 3: Average Recognition Result based on SVM, MLP, and RF classifiers using shape features.

The shape features performance is described in this section by applying the three types of classifiers: SVM, MLP, and RF. Among these three classifiers, Random Forest has the highest recognition accuracy of 91.36% with Precision of 0.914, Recall of 0.914, and F-Measure of 0.914.

Table 9: Average recognition of the five types of fruits using Hybrid features.

Hybrid Features (11 Features)					
Sl. No.	Classifier	Rec. Acc	Precision	Recall	F-Measure
1	SVM	91.56%	0.784	0.884	0.831
2	MLP	99.62%	0.996	0.996	0.996
3	RF	99.98%	1.000	1.000	1.000

Table 10: Confusion Matrix for the SVM classifier.

Confusion matrix for SVM					
Fruit Type	Apple	Banana	Cherry	Grapes	Mango
Apple	940	0	35	0	25
Banana	0	1000	0		0
Cherry	40	0	884	68	8
Grapes	5	0	137	831	27
Mango	0	0	71	6	923

Table 11: Confusion Matrix for the MLP classifier.

Confusion matrix for MLP					
Fruit Type	Apple	Banana	Cherry	Grapes	Mango
Apple	998	0	2	0	0
Banana	0	1000	0	0	0
Cherry	0	0	1000	0	0
Grapes	4	0	10	983	3
Mango	0	0	0	0	1000

Table 12: Confusion Matrix for the RF classifier.

Confusion matrix for RF					
Fruit Type	Apple	Banana	Cherry	Grapes	Mango
Apple	1000	0	0	0	0
Banana	0	1000	0	0	0
Cherry	0	0	999	1	0
Grapes	0	0	0	1000	0
Mango	0	0	0	0	1000

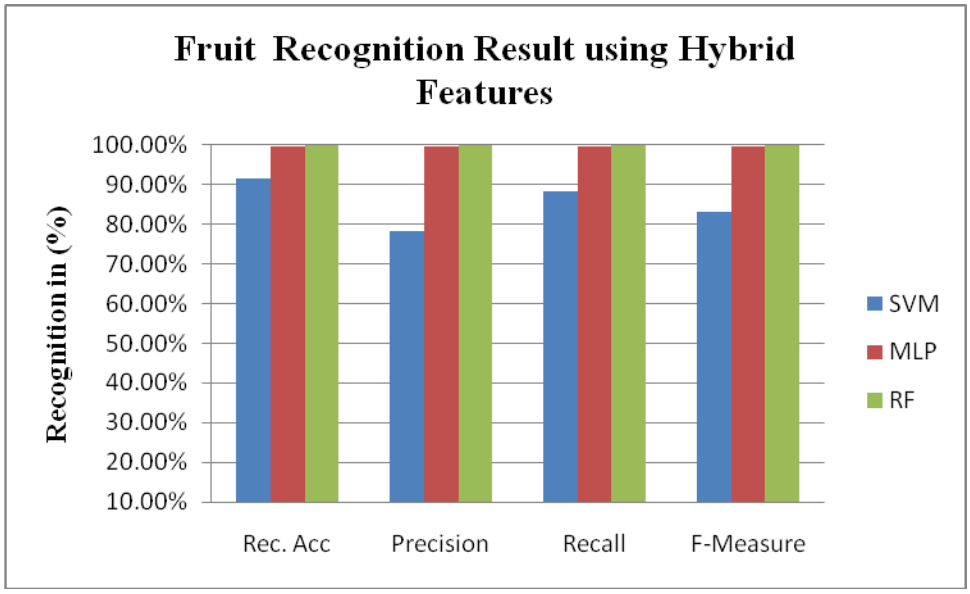

Figure 4: Average Recognition Result based on SVM, MLP, and RF classifiers using Hybrid features.

The maximum recognition nearer to 100% recognition accuracy has been reached by efficient utilization of hybrid feature vector contains color_moments and shape features. The Random Forest classifiers have outperformed all other classifiers. By giving the recognition accuracy of 99.98, and Precision, Recall and F-Measure by giving 1.000. Hence, it is noted that this combined/hybrid feature has given maximum recognition, and it is the contribution of this paper. Table 13 shows the comparative analysis.

Table 13: Comparative analysis of fruit classification

Reference	Features	Classifier	Average Accuracy
Dae Gwan Kim [19]	Intensity texture features	Discriminant analysis	96.00%
Proposed Method	Color_moments and Shape Features	Random Forest	99.98%

The above table has shown the comparative study of the proposed method with other methods, and the proposed method has given higher recognition accuracy compared with other work.

IV. Conclusion

Sometimes the one feature descriptor is not sufficient to capture the classes' dissimilarity when it comes to solving complex classification issues. Efficient and successful fusion policies are, therefore required, and it becomes essential. Hence, this paper successfully utilized the two feature vectors with the fused feature vector. The multiple classifiers have been considered to obtain the highest recognition accuracy. The combined feature vector with the RF classifier has received 99.98% recognition accuracy. More features may be considered in future work, and an attempt has been to apply deep learning techniques on the fruit images.

References

[1] A. Nosseir and S. E. Ashraf Ahmed, "Automatic classification for fruits' types and identification of rotten ones using k-NN and SVM," Int. J. online Biomed. Eng., vol. 15, no. 3, pp. 47–61, 2019, doi: 10.3991/ijoe.v15i03.9832.

[2] S. K. Behera, A. K. Rath, A. Mahapatra, and P. K. Sethy, "Identification, classification & grading of fruits using machine learning & computer intelligence: a review," J. Ambient Intell. Humaniz. Comput., no. Kondo 2010, 2020, doi: 10.1007/s12652-020-01865-8.

[3] C. C. Patel and V. K. Chaudhari, Comparative Analysis of Fruit Categorization Using Different Classifiers. Springer Singapore.

[4] S. Sakib and Z. Ashrafi, "Implementation of Fruits Recognition Classifier using Convolutional Neural Network Algorithm for Observation of Accuracies for Various Hidden Layers," pp. 8–11, 1980.

[5] H. M. Zawbaa, M. Hazman, M. Abbass, and A. E. Hassanien, "Automatic fruit classification using random forest algorithm," pp. 164–168, 2014.

[6] A. Rocha, D. C. Hauagge, J. Wainer, and S. Goldenstein, "Automatic fruit and vegetable classification from images," vol. 70, pp. 96–104, 2010, doi: 10.1016/j.compag.2009.09.002.
[7] S. P. Deenan and J. Satheeshkumar, "A Study on Image Processing Methods for Fruit Classification A Study on Image Processing Methods for Fruit," no. December 2012, 2016.
[8] "RECOGNITION ALGORITHMS FOR DETECTION OF APPLE FRUIT IN AN ORCHARD FOR EARLY YIELD PREDICTION Rong Zhou ."
[9] C. Pl, DEFECT IDENTIFICATION IN THE FRUIT APPLE USING K-MEANS COLOR IMAGE SEGMENTATION ALGORITHM," no. August 2017, 2019, doi: 10.26483/ijarcs.v8i8.4735.
[10] J. P. Mercol, J. Gambini, and J. M. Santos, "Automatic classification of oranges using image processing and data mining techniques," XIV Congr. Argentino Ciencias la Comput. XIV Argentine Congr. Comput. Sci. (CACIC 2008), pp. 1–12, 2008.
[11] S. Naik and B. Patel, "Machine Vision based Fruit Classification and Grading - A Review," Int. J. Comput. Appl., vol. 170, no. 9, pp. 22–34, 2017, doi: 10.5120/ijca2017914937.
[12] M. Sahana and H. B. Anita, "Automatic Classification of South Indian Regional Fruits using Image Processing," Indian J. Sci. Technol., vol. 10, no. 13, pp. 1–4, 2017, doi: 10.17485/ijst/2017/v10i13/110462.
[13] R. S. Choras, "Image Retrieval Using Color, Texture and Wavelet Transform Moments," Adv. Pattern Recognit., doi: https://doi.org/10.1142/9789812772381_0042.
[14] T. Weng, Y. Yuan, L. Shen, and Y. Zhao, "Clothing image retrieval using color moment," Proc. 2013 3rd Int. Conf. Comput. Sci. Netw. Technol. ICCSNT 2013, pp. 1016–1020, 2014, doi: 10.1109/ICCSNT.2013.6967276.
[15] S. Y. Ying Deng, "An image retrieval approach based on annular color moments and energy moments features," 2010.
[16] G. Stylianou and G. Farin, "Shape Feature Extraction," pp. 269–281, 2003, doi: 10.1007/978-3-642-55787-3_16.
[17] Y. Mingqiang, K. Kidiyo, and R. Joseph, "A Survey of Shape Feature Extraction Techniques," Pattern Recognit. Tech. Technol. Appl., 2008, doi: 10.5772/6237.
[18] S. R. Dubey and A. S. Jalal, "Application of Image Processing in Fruit and Vegetable Analysis : A Review," vol. 24, no. 4, pp. 405–424, 2015, doi: 10.1515/jisys-2014-0079.
[19] H. Mureşan and M. Oltean, "Fruit recognition from images using deep learning," arXiv, no. June 2018, 2017, doi: 10.2478/ausi-2018-0002.
[20] D. G. Kim, T. F. Burks, J. Qin, and D. M. Bulanon, "Classification of grapefruit peel diseases using color texture feature analysis," Int. J. Agric. Biol. Eng., vol. 2, no. 3, pp. 41–50, 2009, doi: 10.3965/j.issn.1934-6344.2009.03.041-050.
[21] S. Rege, R. Memane, M. Phatak, and P. Agarwal," 2D Geometric shape and color recognition using digital image processing," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 2, no. 6, pp. 2479-2481, June 2013.
[22] Gode, C.S., Ganar, A.N.: Image retrieval by using colour, texture and shape features. Int. J.Adv. Res. Electr. Electron. Instrumentation Eng. 3(4) (2014)
[23] Arivazhagan, S., Shebiah, R.N., Nidhyanandhan, S.S., Ganesan, L.: Fruit recognition using color and texture features. J. Emerg. Trends Comput. Info. Sci. 1(2) (2010)

[24] Dubey, S.R., Jalal, A.S.: Fruit and vegetable recognition by fusing colour and texture features of the image using machine learning. Int. J. Appl. Pattern Recognition, 2(2) (2015)

[25] J. Platt: Fast Training of Support Vector Machines using Sequential Minimal Optimization. In B. Schoelkopf and C. Burges and A. Smola, editors, Advances in Kernel Methods - Support Vector Learning, 1998.

[26] Zhen Zhang, Weimin Shao and Hong Zhang, "A learning algorithm for multilayer perceptron as classifier, " IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339),Washington,DC,USA,1999,pp.1681-1684vol.3, doi: 10.1109 /IJCNN .1999. 832627.

[27] Prinzie A., Van den Poel D. (2007) Random Multiclass Classification: Generalizing Random Forests to Random MNL and Random NB. In: Wagner R., Revell N., Pernul G. (eds) Database and Expert Systems Applications. DEXA 2007. Lecture Notes in Computer Science, vol 4653. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-74469-6_35

[28] Denisko D, Hoffman MM (February 2018). "Classification and interaction in random forests". Proceedings of the National Academy of Sciences of the United States of America. 115 (8): 1690–1692. doi:10.1073/pnas.1800256115.