Scholar Publishing Group
International Journal of Neural Network
https://doi.org/10.38007/NN.2020.010406
ISSN 2790-2692 Vol. 1, Issue 4: 42-51

Fruit and Vegetable Recognition and Nutrition Analysis Based on Multilayer Perceptron and Neural Network

Aditya Ijaz*

Tamale Technical University, Ghana

**corresponding author*

Keywords: Multilayer Perceptron, Neural Network, Fruit and Vegetable Recognition, Nutrition Analysis

Abstract: The current research on automatic recognition of vegetable and fruit images is more focused on a single fruit and vegetable image without background environment, and more features such as texture and color are extracted, and shallow learning technology is used to realize the recognition of vegetable and fruit images. The method cannot meet the classification and identification of various fruits and vegetables. The main purpose of this paper is to carry out research on fruit and vegetable identification and nutritional analysis based on the knowledge of multilayer perceptron and neural network (NN). In view of the above background, this research focuses on the research on IR of fruits and vegetables, using deep learning methods to improve the recognition rate of fruits and vegetables images, and constructing models through CNNs to achieve classification and recognition. Experiments show that the number of network layers has a greater impact on IR, and the deeper the network depth, the higher the IR rate. Considering that the increase in the number of network layers will affect the structural complexity of the network model and increase the amount of mathematical operations, reasonable network layer selection is particularly important in IR.

1. Introduction

The sorting and identification of agricultural products, as an important link in the production and operation of fruit and vegetable agricultural products, whether in the production and picking of fruits and vegetables in the early stage, or the wholesale and distribution of products in the later stage, is currently more of a manual identification and classification. This undoubtedly increases the cost of human and material resources. Compared with machine processing, manual identification has relatively low work efficiency, and it is easy to confuse similar vegetables, fruits and agricultural products through visual identification, and it is not conducive to standardized

Copyright: © 2020 by the authors. This is an Open Access article distributed under the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (https://creativecommons.org/licenses/by/4.0/).

classification and grading of products. This will affect the competitiveness of my country's fruit and vegetable agricultural products in the international market, and is not conducive to the import and export of agricultural products. At the same time, the consumption of operation and management caused by this link will directly increase the cost of products, which will reduce the profit margins of fruits, vegetables and agricultural products that are originally low in price, and reduce agricultural economic income. If things go on like this, it will restrict the development of the current agricultural economy, which is not conducive to the steady growth of the national economy [1, 2].

In related research, Karaki mentioned that multilayer perceptrons, recurrent NNs, convolutional networks and other types of NNs are now common [3]. NNs have hyperparameters, and Bayesian optimization is one of the methods used to tune the hyperparameters. Usually this technique treats the values of neurons in the network as a random Gaussian process. The experimental results of the multivariate normality test are reported and demonstrate that the neuron vectors are far from a Gaussian distribution. Panghal et al. introduced a faster method to train NNs to solve differential equations based on extreme learning machine algorithms [4]. Compared to traditional methods, the algorithm is much faster and also provides highly accurate results. The reliability of the method is tested by solving various cases of the hyperbolic telegraph equation.

In this study, considering that the traditional recognition technology is applied to the classification of fruits and vegetables, there are problems such as the lack of consideration of the complex background environment and light intensity and other parameters, and the application effect of the recognition results is not good. By integrating the NN model, a vegetable and fruit IR model is constructed to realize the classification and recognition of multiple varieties of vegetable and fruit images with complex backgrounds in the transaction environment in the natural environment. It has certain social value and research significance. The research first uses the camera and other tools to collect images of fruits and vegetables, and complete the image library of fruits and vegetables. Taking the vegetable and fruit image library as the main research sample, by modifying the network structure of the traditional LeNet-5 model, it is applied to the IR of the image library. By modifying the relevant parameters and methods of the model, we can deeply understand the influence of the parameters and methods of each network layer on the IR results.

2. Design Research

2.1. Comparison of Image Recognition (IR) Methods

Traditional feature extraction methods extract features manually, relying on one or several feature extraction algorithms to extract some basic image features from the original image according to specific rules [5-6]. Although this method can quickly extract the required features, these features may not have the best recognition effect. Because these feature extraction rules are all derived from experience, when faced with different background environments, the extracted features are good or bad [7-8]. Moreover, there is no unified strategy for how to combine these features to improve the IR effect. The IR process using the traditional feature extraction method is shown in Figure 1.

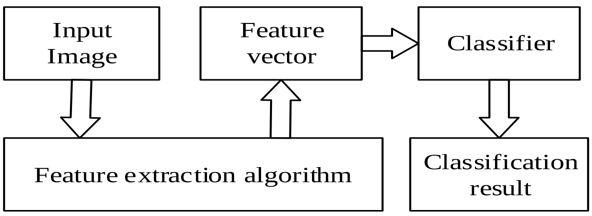

Figure 1. Workflow of traditional IR methods

This kind of IR method performed well when used for small and medium-scale IR in the past, but due to the limitations of manual feature extraction, it is not enough when faced with the recognition and classification of today's massive image data. With the development of deep learning, IR technology based on deep learning method gradually shows its advantages, and has achieved very good recognition results in many fields. Different from traditional IR methods, the extracted signs are constantly changing to ensure the best classification results [9, 10]. In other words, this approach allows the machine to learn which features to use, giving the machine the ability to "learn" without humans caring about how it is extracted Figure 2. Describes the IR process based on multilayer perceptrons and NNs [11, 12]. This IR method can not only reduce the difficulty of feature design, but also improve the accuracy of recognition, and because of the data-driven characteristics of deep learning, it can also show high performance in the face of large-scale image data [13, 14].

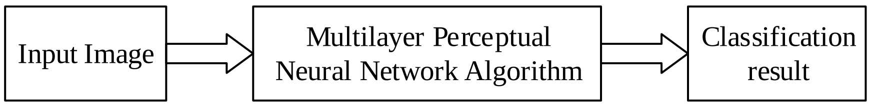

Figure 2. Workflow of IR method based on multilayer perceptron and NN

2.2. Multilayer Perceptron Network Model

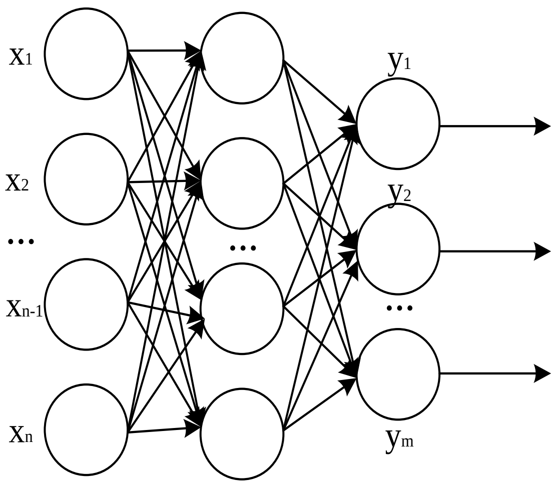

Figure 3. Multilayer perceptron network model

Figure 3 shows a one- or two-layer NN model (Multilayer Induction Machine) [15-16]. Its main structure includes an input layer, an output layer and an intermediate layer. In this case, both the output layer and the intermediate layer are computational layers. At this time, the input data of the intermediate node is calculated according to the data transmitted by the input layer, and the input data of the output layer is calculated according to the weight table of the output data of the upper layer (Layer layer).

At present, the classification of artificial NNs can be divided into two categories in terms of structure: feedforward NNs and feedback networks. Among them, the feedforward network is the process of processing the information in the network, which is to transmit from the input layer to each hidden layer and then calculate the output by the output layer. The input signal of each layer is the output signal of the previous layer, and there is no reverse transfer process in the transmission of the signal. The feedback network is based on the feedforward network. There is a transmission process between the upper and lower nodes. All nodes process the signal. The data flow is no longer a one-way transmission, but a process of signal feedback [17, 18].

2.3. Convolutional NN Structure

Convolutional neural network (CNN) is a deep NN designed and implemented based on multilayer perceptrons, which are influenced by the "local receptive field" of biological visual systems. Each layer of a CNN consists of multiple 2D layers, each of which contains multiple independent neurons. Neurons are the basic units that make up complex networks. Through literature research, it is found that although there are many network models based on network optimization, the basic structures of their network processes are very similar.

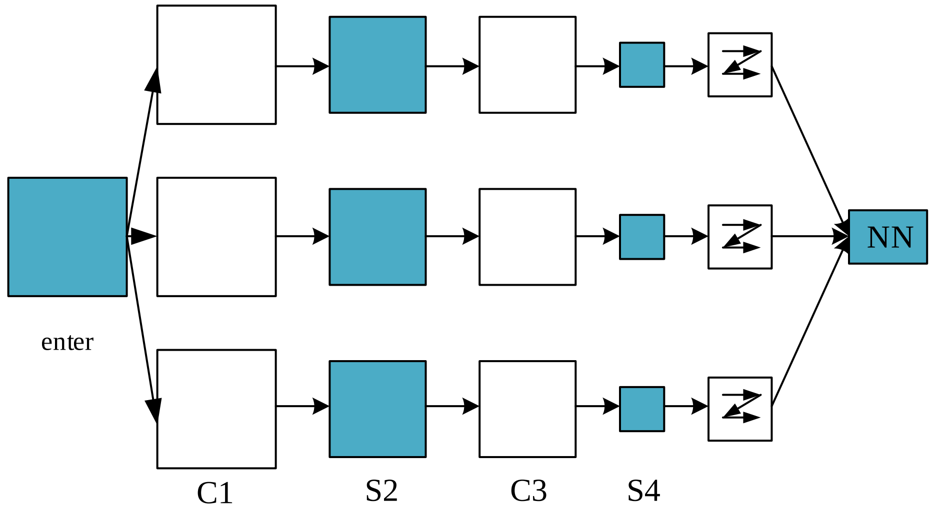

Figure 4. CNN structure model

As far as the above figure is concerned, the C layer is a convolutional layer, and the S layer is a downsampling layer.

The first layer of feature layer neurons is expressed as:

$$z_{i,j,k}^{l} = W_{k}^{lT} X_{i,j}^{l} + b_{k}^{l} \tag{1}$$

Among them, Wlk and blk are the weight vector and the k-layer filter bias parameter corresponding to the l-layer, respectively, and Xil,j is expressed as the input layer-related position feature value corresponding to the l-layer neuron. Through the convolution kernel function Wkl, the feature maps zli,j,k are generated.

For the neurons contained in each feature layer, it follows the basic neuron nonlinear triggering mechanism. If a() is used to represent its corresponding nonlinear activation function, then each neuron can be expressed as:

$$a_{i,j,k}^{l} = a(z_{i,j,k}^{l}) \tag{2}$$

At present, for neuron activation functions, it is mainly divided into three types: sigmoid type, ReLuctant type and Tanh type.

The formula for the parameter value of the downsampling layer:

$$y_{i,j,k}^{l} = pool(a_{m,n,k}^{l}), \forall (m,n) \in \Re_{ij} \tag{3}$$

where Rij represents the local leading position of the convolutional layer corresponding to the downsampling layer neuron (i, j).

After multiple layers of convolution and downsampling are performed on the input data, there are generally one or more fully connected layers, which are mainly used to output classification results. The final fully-connected layer produces the final classification result. Another classifier is the support vector machine classifier (SVM). If R is used to represent the parameters of the CNN, for an optimal classification model, the loss function is minimized by continuously optimizing the parameters. If x(n) is used to represent the nth input data, y(n) is used to represent the target label quantity, and o(n) is used to represent the output layer of the convolution function. Then the loss value of the CNN is expressed as:

$$\mathcal{L} = \frac{1}{N}\sum_{n=1}^{N} \ell(\theta; y^{(n)}, o^{(n)}) \tag{4}$$

3. Experimental Study

3.1. The Training Process of CNN

The training of the convolution model is a process of global optimization, and its training algorithm is similar to the BP algorithm, which mainly includes two stages: forward pass and backward pass. In the forward propagation stage, the image data is passed into the network model through the input layer, and after layer-by-layer data conversion, the final result is output. The backward propagation stage mainly compares the output result with the expected output, finds the error, transfers the error back to each hidden layer, adjusts the weight matrix according to the minimized loss function, and finds the optimal convolution model. The stochastic gradient descent method is a common method used to determine the optimal convolution parameter model. Specific steps are as follows:

(1) Select training samples. The network training samples are determined by randomly selecting a certain number of samples in the training sample set.

(2) Initialization parameters. The weights and thresholds of each network layer are set, and it is required to be set as a random parameter that is as close to 0 as possible, but not 0. Set the learning rate, training step size, and error range, or accuracy range.

(3) Extract the image input samples from the training samples, and calculate the output vectors of the intermediate layer and the output layer through the network.

(4) Obtain the error value between the output vector of the output layer and the target vector by operation.

(5) The error term of the hidden unit in the middle layer is derived from the final error value.

(6) From this, the adjustment amount of each connection weight and threshold is calculated, and then the parameter amount of the weight and threshold in the network is adjusted.

(7) Determine whether the error accuracy range is reached or the maximum number of training times is reached. If not, continue to repeat steps 3-6.

(8) Through training, the fitting of the relationship between the CNN and the image classification is finally realized, a stable NN classification model is obtained, and the relevant weights and thresholds are saved. The next time you train, you can call it directly.

3.2. Analysis Based on the Number of Network Layers

The network layer is the basis of the NN model. The current network layer has an accident of input and output layers, which are mainly divided into convolutional layers, pooling layers, fully connected layers, and so on. The traditional NN model mostly adopts the full connection method, which believes that the output of the current network layer is correlated with all the inputs of the previous network layer, which often leads to too many data parameters, too complicated network structure, and data operation problems such as large volume and overfitting of training results. Through the local connection and weight sharing unique to the convolutional layer, as well as the downsampling operation of the pooling layer, and finally using softmax through the full connection for image classification, although the number of network layers is increased, the number of relevant weights can be reduced. And the complexity of the network layer is greatly reduced, the occurrence of overfitting is less, and it can have better recognition efficiency and recognition effect.

In view of the influence of the number of network layers on the effect of IR of vegetables and fruits, based on the original vegetable and fruit classification model, the number of each network layer is modified, and the following models are obtained under the condition that other parameters remain unchanged:

Model 1: Remove the F6 fully connected layer of the model, and after the C5 network layer, directly input the parameters into the softmax value through the full connection to perform image classification.

Model 2: On the basis of the original model, one layer of convolution layer is reduced, that is, S4 is directly connected to F6, and no convolution operation is performed on it.

Model 3: Based on the original model, remove the two convolutional layers. After C1 is pooled and downsampled to obtain S2, it is directly connected to F6 to build a model and realize IR.

Model 4: Also based on the original model, while removing the two convolutional layers, a fully connected layer is removed, and the parameters of the S2 feature layer are directly connected to the softmax for image classification.

4. Experiment Analysis

4.1. Analysis of Classification Effect

Based on the above four models, plus the original original model, the fruit and vegetable images are classified. Its classification effect is as follows:

Table 1. The effect of different network layers on the IR effect

Model	Training set false positive rate %	Test machine false positive rate %
Base model	6.54	17.85
Model one	6.93	19.32
Model II	10.37	23.61
Model three	16.96	39.47
Model four	18.15	43.38

It can be seen from the results in Figure 5 that in the deletion of the network layer, the impact of the fully connected layer is smaller than that of the convolutional pooling layer. When a convolutional layer is deleted, the recognition effect becomes significantly worse. When the two-layer convolution pooling layer is removed, the error rate is more than double the recognition rate of the original model. It can be seen from this that the number of network layers of the CNN model has a greater impact on IR, and the deeper the network depth, the higher the IR rate. Of course, considering that the increase in the number of network layers will affect the structural complexity of the network model and increase the amount of mathematical operations, reasonable network layer selection is particularly important in IR.

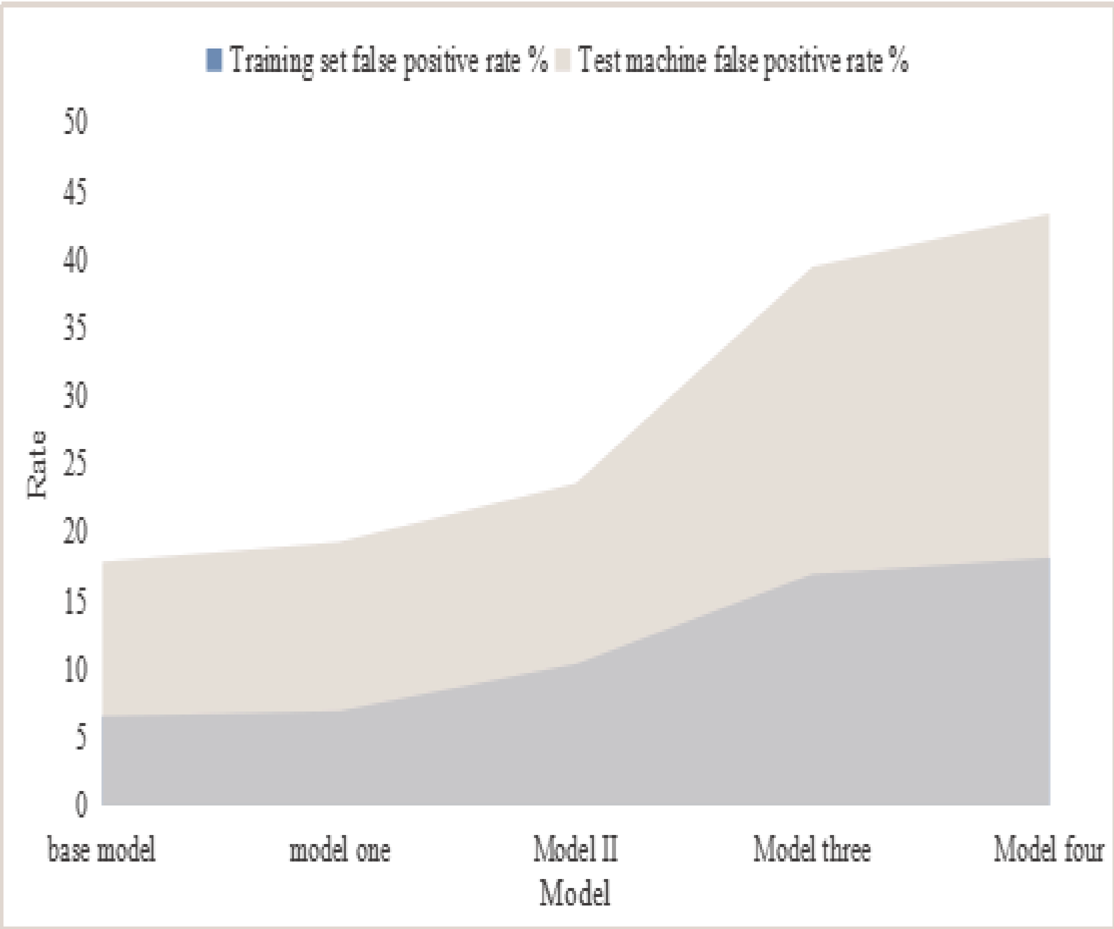

Figure 5. Analysis of the influence of the number of different network layers on the IR effect

4.2. Analysis of Recognition Effect

Aiming at the influence of the number of convolution kernels on the identification results of vegetables, fruits and agricultural products, the research is carried out by modifying the number of model convolution kernels while keeping other parameters unchanged. The specific number of convolution kernels is set as: 16-16-32; 16-32-32; 32-32-32; 32-32-64; 32-64-64; 32-64-120; 64-64-120. The model is used to train and identify the vegetable and fruit image database, and the experimental results are shown in Table 2.

Table 2. Effects of different kernel numbers on IR of fruits and vegetables

	Training set false positive rate %	Test machine false positive rate %
16-16-32	8.8	26.9
16-32-32	8.2	23.5
32-32-32	7.7	21.6
32-32-64	7.2	19.6
32-64-64	6.8	18.3
32-64-120	6.5	17.9
64-64-120	6.3	17.4

Figure 6. Analysis of the effect of different core numbers on IR of fruits and vegetables

In view of the influence of the size of the receptive field of the convolution kernel on the images of vegetables, fruits and agricultural products, without modifying other parameters, the receptive field of the model convolution kernel is modified to 2 × 2, 3 × 3, 5 × 5, and the IR of vegetables, fruits and agricultural products is carried out., the recognition effect is as follows:

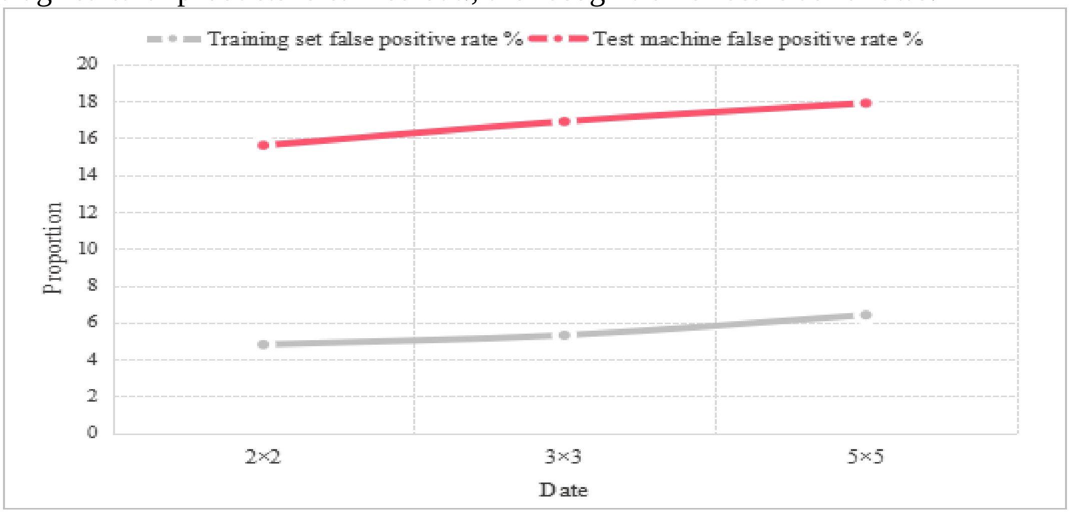

Figure 7. Effects of different receptive field sizes on IR of fruits and vegetables

It can be seen from the results in Figure 7 that by increasing the number of convolution kernels in the convolution layer, the fitting effect of the model can be improved to a certain extent, thereby improving the recognition rate of vegetables, fruits and agricultural products. But on the other hand, as the number of convolutional layers increases, the complexity of the network model increases, the amount of mathematical computation increases, and the memory usage increases. Once the size of the input image is large, it will affect the training and recognition efficiency of the model. By reducing the scope of the visual field of the convolution process (that is, the size of the convolution kernel) the recognition rate of fruits and vegetables can also be improved to a certain extent. In theory, the size of the convolution kernel of 1×1 can have a better recognition effect, but considering that the convolution kernel is too low, it will strengthen the influence of noise on IR to a certain extent, and at the same time increase the network complexity. Therefore, more network models use a 2×2 convolution kernel size for convolution operations.

5. Conclusion

As an important part of computer vision tasks, IR has always attracted the attention of scholars at home and abroad. Fruit IR technology can promote the development of contemporary smart agriculture and smart cities, and has high research value. At present, the commonly used fruit recognition methods cannot meet the needs of fast and accurate recognition in production and life, so it is necessary to study a new type of fruit IR technology. This paper proposes to use the multi-layer perceptron and NN technology, and use the CNN model to build a vegetable and fruit IR model, so as to realize the classification and recognition of multiple varieties of vegetable and fruit images with complex backgrounds in the transaction environment in the natural environment. On the basis of in-depth study of multi-layer perceptron and NN technology, the classical model is applied to fruit IR, and optimized and improved according to practical problems, so as to improve the performance of fruit recognition. It has certain social value and research significance.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Zouhri A, Amakdouf H, Mallahi M E, et al. Invariant Gaussian–Hermite Moments Based NNs for 3D Object Classification. Pattern Recognition and Image Analysis. (2020) 30(1):87-96. https://doi.org/10.1134/S1054661820010186

[2] Chishti M A, Al S. Assessing the Efficacy of Logistic Regression, Multilayer Perceptron, and CNN for Handwritten Digit Recognition. IJCDS Journal. (2020) 9(2):299-308. https://doi.org/10.12785/ijcds/090215

[3] Karaki Y, Ivanov N. Hyperparameters of Multilayer Perceptron with Normal Distributed Weights. Pattern Recognition and Image Analysis. (2020) 30(2):170-173. https://doi.org/10.1134/S1054661820020054

[4] Panghal S, Kumar M. A Multilayer Perceptron NN Approach for the Solution of Hyperbolic Telegraph Equations. Network: Computation in Neural Systems. (2020) 32(2-4):65-82. https://doi.org/10.1080/0954898X.2020.2015005

[5] Moshkbar B K. Bayesian Regularization of Multilayer Perceptron NN for Estimation of Mass Attenuation Coefficient of Gamma Radiation In Comparison with Different Supervised Model-Free Methods. Journal of Instrumentation. (2020) 15(11):P11019-P11019. https://doi.org/10.1088/1748-0221/15/11/P11019

[6] Balamurugan M. Multilayer Perceptron NN Analysis: Prediction of School Outcome from the Dimensions of School Process. Xi'an Jianzhu Keji Daxue Xuebao/Journal of Xi'an University of Architecture & Technology. (2020) XIII(7):59-67.

[7] Raziani S, Ahmadian S, Jalali S, et al. An Efficient Hybrid Model Based on Modified Whale Optimization Algorithm and Multilayer Perceptron NN for Medical Classification Problems. Journal of Bionic Engineering. (2020) 19(5):1504-1521. https://doi.org/10.1007/s42235-022-00216-x

[8] Verma P, Awasthi V K, Sahu S K. Classification of Coronary Artery Disease Using Multilayer Perceptron NN. International Journal of Applied Evolutionary Computation. (2020) 12(3):35-43. https://doi.org/10.4018/IJAEC.2020070103

[9] Zvarevashe K, Olugbara O O. Recognition of Speech Emotion Using Custom 2D-Convolution NN Deep Learning Algorithm. Intelligent Data Analysis. (2020) 24(5):1065-1086. https://doi.org/10.3233/IDA-194747

[10] Otoom A F. Software Effort Classification with Multilayer Perceptron NNs. International Journal of Advanced Trends in Computer Science and Engineering. (2020) 9(2):1678-1689. https://doi.org/10.30534/ijatcse/2020/119922020

[11] Vladimir G, Yury M, Natalya K. Pattern Recognition Techniques for Classifying Aeroballistic Flying Vehicle Paths. Neural Computing and Applications. (2020) 34(5):4033-4045. https://doi.org/10.1007/s00521-021-06662-8

[12] Asadi N, Scott K A, Komarov A S, et al. Evaluation of a NN With Uncertainty for Detection of Ice and Water in SAR Imagery. IEEE Transactions on Geoscience and Remote Sensing. (2020) (99):1-13. https://doi.org/10.1109/TGRS.2020.2992454

[13] Saha S, Dawson J, Murimi M, et al. Effects of a Nutrition Education Intervention on Fruit and Vegetable Consumption-Related Dietary Behavioural Factors among Elementary School Children: Health Education Journal. (2020) 79(8):963-973. https://doi.org/10.1177/0017896920944421

[14] Kojima G, Taniguchi Y, Urano T. Fruit and Vegetable Consumption and Incident Frailty in Older Adults: A Systematic Review and Meta-Analysis. The Journal of Frailty & Aging. (2020) 11(1):45-50. https://doi.org/10.14283/jfa.2020.32

[15] Aune D, Halvorsen R E, Molin M. Fruit and Vegetable Consumption and the Risk of Type 2 Diabetes: A Systematic Review and Dose–Response Meta-Analysis of Prospective Studies. BMJ Nutrition, Prevention and Health. (2020) 4(2):519-531. https://doi.org/10.1136/bmjnph-2020-000218

[16] Dumas B, Lee S H, Park S, et al. Associations Between Perceived Fruit and Vegetable Intake and Nutrition Messaging: Results from a Population-Based Survey Among US Adults. Current Developments in Nutrition. (2020) 4(Supplement_2):1364-1364. https://doi.org/10.1093/cdn/nzaa060_002

[17] Alnaim L, Taylor M, Gibbs H, et al. A Pilot Nutrition and Cooking Skills Program Improves Fruit Intake, Nutrition Knowledge, and Attitudes Toward Fruit and Vegetable. Current Developments in Nutrition. (2020) 5(Supplement_2):957-957. https://doi.org/10.1093/cdn/nzab051_001

[18] Azekour K, Bidi A, Bouhali E. Socioeconomic Characteristics and Fruit/Vegetable Intakes among Scholar Children in the Oasis of Tafilalet, Southeastern Morocco. Nutricion Clinica y Dietetica Hospitalaria. (2020) 39(3):24-28.