RADIAL BASIS NEURAL NETWORK WITH MULTIPLE CONNECTED WEIGHTS

Marakhimov Avazjon
Department of Information Processing and Control Systems, Tashkent State Technical University, Tashkent, Uzbekistan
Khudaybergenov Kabul
Department of Applied Informatics, Kimyo International University in Tashkent, Uzbekistan
E-mail: kabul.kudaybergenov@gmail.com

DOI: ***https://doi.org/10.5281/zenodo.19829106***

Abstract: *In this work, we propose a new type of radial basis neural network model where the connection between two units is not a single value but a set of values, which means multi-connected weights exist. In our model, each pair of units is connected by more than one link. These links mimic different neurotransmitters, and their number reflects the number of neurotransmitter types considered. Experimental tests on benchmark datasets from the Machine Learning Repository show that using radial basis with multiple weight connections improves performance over traditional neural networks. This method gives a new way to design and build artificial neural networks.*

Keywords: *radial basis function, radial basis neural network, neurotransmitter, multiple connections, weight, hidden layer, classification*

1. Introduction

In this paper, we propose a new type of neural network model, namely we call this as radial basis function neural network (RBF) with multi-connected weights (RBFMC) by enlarging the dimensions of connections between two units from one to several, where we use the principle of information communication between two neurons where introduced in [1], for instance, two neurons transfer data by generating multiple types of neurotransmitter signals rather than just a one. Every single connection in the high dimension of connections in RBFMC corresponds to various neurotransmitter types. In particular, the connection between two units for the new model is multidimensional, whereas it is one-dimensional for all neural network models that are currently in use [2,3].

2. RBF neural networks with multiple connection

This section presents the architectural framework of radial basis function (RBF) neural networks incorporating multi-connected weight configurations within the context of classification applications. Mostly, we choose Gaussian function $\varphi(\mathbf{x}, \mathbf{c}_m, \sigma_m)$ as an nonlinear activation function, which is defined below.

$$\varphi(\mathbf{x}, \mathbf{c}_m, \sigma_m) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}_m\|^2}{2\sigma_m^2}\right) \tag{2}$$

where $\mathbf{x} \in R^I$ is an input vector, $\mathbf{c}_m \in R^I$ – the centers, and σ_m – width of the $m-\text{th}$ hidden neuron. ||·|| - is the Euclidean norm. For simplicity and convenience, we denote all these parameters as $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_M]$ with $\boldsymbol{\theta}_m = [\mathbf{c}_m^T, \sigma_m]$.

Let us assume that we are given P pairs of training data samples $\{(\mathbf{x}_p, d_p)\}_{p=1}^{P}$, where $d_p \in R^K$ is the desired output values for the p-th input data point $\mathbf{x}_p$ and K is number of classes in the dataset. According to the traditional RBF neural network architecture which is presented in Fig. 1, the k-th output y_{pk} of RBF neural network model for input data $\mathbf{x}_p$ is computed as follows

$$y_{pk} = w_{0k} + \sum_{h=1}^{H} w_{hk}\varphi(x_p, \theta_h) \tag{3}$$

where w_{0k} - the bias parameter of the k^{th} output neuron, w_{mk} - is the weight parameter between the m-th hidden neuron and the kth output unit. In this case, the output for traditional RBF network is calculated by

$$Y = \Phi\mathbf{W} \tag{4}$$

where $Y = [y_1, ..., y_p]^T$ is a desired output with $P \times M$. $\mathbf{y}_p \in R^K$ - is a desired output vector. $\mathbf{x}_p$ - is a pth input data, $\Phi = [1, \varphi_1, \varphi_2, ..., \varphi_H]$ - is a matrix with $K \times (M+1)$ dimension.

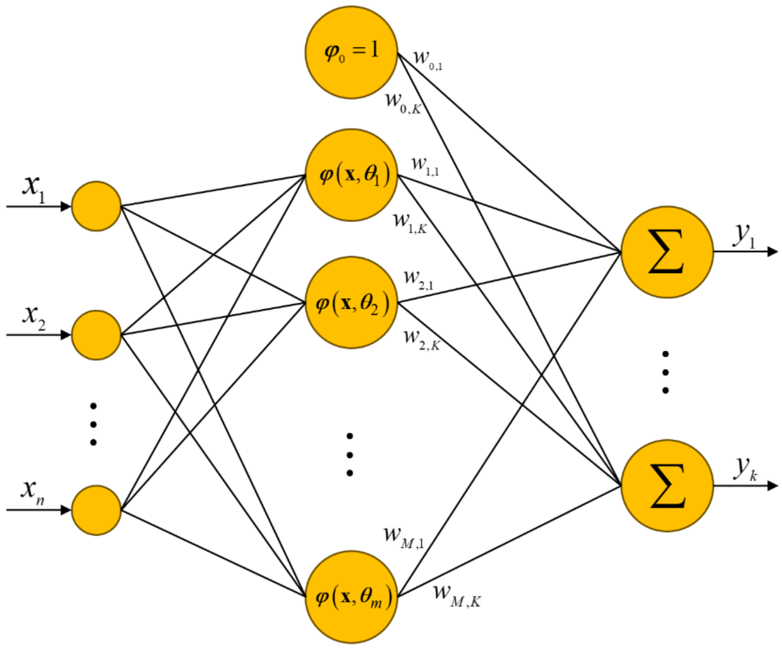

Fig. 1. The structure of a traditional RBF neural network (I-input, M-hidden neurons and K-outputs units)

Let us consider that we have multi-dimensional connections between the hidden layer and the output layer such as in Fig. 2(b).

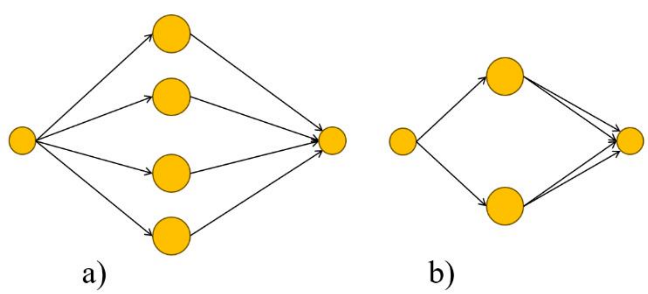

Fig. 2. (a) The structure of a traditional feed-forward neural network 1-4-1. (b) The structure of the proposed RBFMC 1-2-1.

Then, we can define a model for the RBFMC model as below

$$y_m = \sum_{m=1}^{M} \left(w_{0m} + \sum_{h=1}^{H} w_{h,m} \varphi(\mathbf{x}, \theta_h) \right) \tag{5}$$

where $h = 1,\ldots,H,$ is the number of connection dimensions in RBFMC where we extended up to 2 connections between hidden and output units. Here, we consider RBFMC model for classification tasks. These indicates that the model must have K outputs for K -classes and y_k - output prediction value for $k-th$ class. In Fig. 2(b), we have 2 connections between the hidden and the output units $(h=2)$.

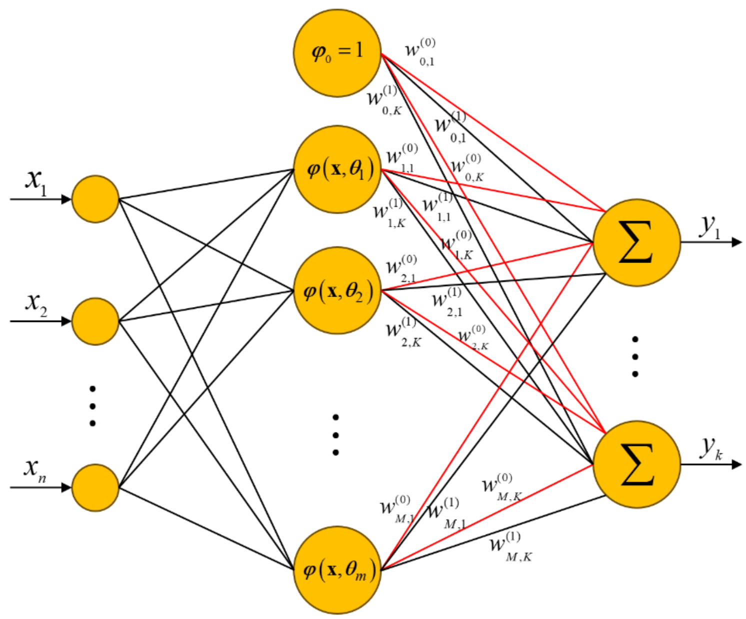

Fig. 3. The model structure of the proposed RBFMC with a two-dimensional connection.

Here, we extend the approximation properties of traditional RBF neural networks to RBFMC to classifier neural networks. On the k-dimensional space $\mathbb{R}^k$ consider a Euclidean norm

$$\|x\| = \sqrt{\sum_{i=1}^{k} |x_i|^2},\ x = (x_1, \dots, x_d). \tag{6}$$

Theorem 2.1: *Let radial basis function* $\varphi(\mathbf{x}, \mathbf{c}_h, \sigma_h)$ *be a non-constant, bounded, and monotone increasing function. Let* K *be a compact subset of* R^k, *and* $f(x)$ *be a real-valued continuous function on* K. *Then for arbitrary* $\varepsilon > 0$, *there exist real constants* $w_{i,j}, \theta_j$ *, which satisfies*

$$\left\| f(\mathbf{x}) - \sum_{k=1}^{K} \sum_{h=1}^{H} \left(w_{0,k}^{(h)} + \sum_{m=1}^{M} w_{m,k}^{(h)} \varphi\left(\mathbf{x}, \theta_m^{(h)}\right) \right) \right\|_C \le \varepsilon. \tag{7}$$

Proof. If $H = 1$, then the RBFMC model becomes a traditional RBF network model and its proof is trivial. We have to consider the case when $H > 1$ since our proposed RBFMC model has multiple connections between hidden unit and output unit.

According to the universal approximation ability of traditional RBF network, for any $f \in C(x,R)$, $\varepsilon > 0$, there exist $w_{m,1}, \theta_m \in R$ such that

$$f(x_1,\dots,x_n) = w_{0,1} + \sum_{m=1}^{M} w_{m,1}\varphi(\mathbf{x},\theta_m) \tag{8}$$

is an approximation of the function $f(\cdot)$.

For each $k = 1\dots K$ define projection map $\pi_k : \mathbb{R}^K \to \mathbb{R}$ as follows

$$\pi_k(x) = x_k,\; x = (x_1,\dots,x_d).$$

Then for each $k = 1\dots K$ a function $f \circ \pi_k : \mathbb{R}^K \to \mathbb{R}$ satisfies all conditions of Theorem 2 from [15]. Then by Theorem 2 for each $k = 1\dots K$ for any $f \in C(x,R^k)$, $\varepsilon > 0$, there exist $w_{m,1}, \theta_m \in R$ for $m \in 1,\dots,M,\; h \in 1,\dots,H,\; k \in 1,\dots,K,$ for which

$$\left\| \pi_k(x) - \sum_{h=1}^{H}\left(w_{0,k}^{(h)} + \sum_{m=1}^{M} w_{m,k}^{(h)}\varphi\left(\mathbf{x},\theta_m^{(h)}\right)\right)\right\| < \frac{\varepsilon}{k}. \tag{9}$$

Further

$$\left\| f(x) - \sum_{k=1}^{K}\sum_{h=1}^{H}\left(w_{0,k}^{(h)} + \sum_{m=1}^{M} w_{m,k}^{(h)}\varphi\left(\mathbf{x},\theta_m^{(h)}\right)\right)\right\| =$$

$$\sqrt{\sum_{i=1}^{K}\left|\pi_i\left(f(x)\right) - \sum_{h=1}^{H}\left(w_{0,k}^{(h)} + \sum_{m=1}^{M} w_{m,k}^{(h)}\varphi\left(\mathbf{x},\theta_m^{(h)}\right)\right)\right|} \overset{(8)}{<}$$

$$\sqrt{\sum_{i=1}^{k}\frac{\varepsilon^2}{k^2}} = \sqrt{k\frac{\varepsilon^2}{k^2}} < \varepsilon. \tag{10}$$

4. Computational experiments

We tested cross-validation for training on benchmark datasets. On each epoch, we use half of the samples in classes of face dataset with more than one training sample. The average accuracy of the proposed model and DeepID method is estimated to be 99.10% and 97.57%. In the meantime, we test our proposed model on the LFW face dataset. The comparison results are given below in Table 1.

Table 1. Comparison results on the LFW dataset with different classifiers.

Method	**Avg. accuracy**

DeepFace-ensemble	0.9724 ± 0.0022
DeepID	0.9735 ± 0.0025
ConvNet-RBM	0.9227 ± 0.0037
Proposed NN model	0.9910 ± 0.0025

Table 2. Recognition rates results of the different datasets.

Methods	**NN**	**SVM**	**Traditional RBF NN**	**Proposed RBFMC model**
ORL Face	92.50	94.00	95.12	98.25
LFW	95.25	95.50	96.92	99.10
FERET	86.11	74.07	90.81	97.80
SCface	91.50	93.40	93.79	97.30
AR Face	89.20	79.50	89.85	98.10
YALE	90.45	91.70	92.40	95.20

From Table 2, we can see that the proposed RBFMC architecture always has the highest recognition rates than other methods. Moreover, the proposed RBFMC model outperforms the traditional RBF network. Therefore, we can claim that our proposed RBFMC structure is much more effective compared to other algorithms.

4. Conclusion

In this paper, we introduced a new neural network model architecture according to research in the field of cognitive science, mutual neurotransmitters influence biological neurons; that is, between two neurons, they release several neurotransmitters of different values and send information to each other. Based on this biological approach, a new model of neural networks and new types of artificial neural networks are proposed by increasing the number of connected weights between two adjacent neurons.

References

1. Y. Yin, Z. Han, M. Jian, G. Wang, L. Chen, R. Wang. AMSUnet: A neural network using atrous multi-scale convolution for medical image segmentation. Computers in Biology and Medicine. Volume 162, August 2023, 107120, doi:10.1016/j.compbiomed.2023.107120.

2. M. Farrell, S. Recanatesi, E. Brown. From lazy to rich to exclusive task representations in neural networks and neural codes. Current Opinion in Neurobiology. 2023, Vol. 83, 102780, doi:10.1016/j.conb.2023.102780.

3. A. Apicella, F. Isgrò, A. Pollastro, R. Prevete. Adaptive filters in Graph Convolutional Neural Networks. Pattern Recognition. 2023, Vol. 144, 109867, doi:10.1016/j.patcog.2023.109867.