Applied Soft Computing 25 (2014) 129–138

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

# A fast and efficient two-phase sequential learning algorithm for spatial architecture neural network

G. Yang [a,b], J.F. Qiao [a,c,*]

[a] College of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124, China
[b] School of Electrical and Electronic Engineering, East China Jiaotong University, Nanchang 330013, China
[c] Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China

ARTICLE INFO

*Article history:*
Received 22 February 2013
Received in revised form 16 May 2014
Accepted 12 September 2014
Available online 26 September 2014

*Keywords:*
Spatial architecture neural network (SANN)
Span-output-weight
Feedforward-output-weight
Sequential learning

ABSTRACT

Spatial architecture neural network (SANN), which is inspired by the connecting mode of excitatory pyramidal neurons and inhibitory interneurons of neocortex, is a multilayer artificial neural network and has good learning accuracy and generalization ability when used in real applications. However, the backpropagation-based learning algorithm (named BP-SANN) may be time consumption and slow convergence. In this paper, a new fast and accurate two-phase sequential learning scheme for SANN is hereby introduced to guarantee the network performance. With this new learning approach (named SFSL-SANN), only the weights connecting to output neurons will be trained during the learning process. In the first phase, a least-squares method is applied to estimate the span-output-weight on the basis of the fixed randomly generated initialized weight values. The improved iterative learning algorithm is then used to learn the feedforward-output-weight in the second phase. Detailed effectiveness comparison of SFSL-SANN is done with BP-SANN and other popular neural network approaches on benchmark problems drawn from the classification, regression and time-series prediction applications. The results demonstrate that the SFSL-SANN is faster convergence and time-saving than BP-SANN, and produces better learning accuracy and generalization performance than other approaches.



## 1. Introduction

Artificial neural networks (ANNs) are computational (or mathematical) models that are inspired by the structure and/or functional aspects of biological neural networks (BNNs) to solve actual real life problems. Since ANNs can "learn" from samples which much like human beings and other advantages, e.g., the capability of computing any arithmetic or logical function in principle, good at providing fast and close approximations of the correct answer, best at identifying patterns and trends in data, etc., ANNs have been widely used in a variety of real applications with great success in recent years [1].

Inspired by the connecting mode of neocortex neurons and the lateral inhibition mechanism, a novel ANN model, named spatial architecture neural network (SANN) [2] has been proposed in the previous work to solve real world regression and classification applications. Similar with the neuronal connect modes of excitatory pyramidal neurons and inhibitory interneurons within neocortex, the SANN has a multilayered structure. It consists of a number of neural layers of a similar structure cascaded one after another, which including feedforward and spatial span connections between different layers and also with recurrent lateral inhibitory connections among neurons in each hidden layer. One of the main structural features of SANN is the span connection between different neuronal types from any two non-adjacent horizontal layers, which corresponds to the canonical circuitry of pyramidal neurons in the upper layers II and III of neocortex. It means that these connections may form cortical columns and span all the thickness of the cortex [3]. And the second major structural principle of SANN is the introducing of lateral inhibitory connections between adjacent hidden neurons (within the same layer), which is used to organize inhibitory circuits and enhance the contrast of perception areas.

Although the structure of ANN defines its function, the learning algorithm is the other key issue, which can guarantee and maybe even improve the performance of network. A supervised error backpropagation-based (BP) learning algorithm suitable for the special structure of SANN guaranteed that SANN can be applied in

* Corresponding author at: College of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124, China.
Tel.: +86 10 67391631.
*E-mail addresses:* dr.hankyang@gmail.com (G. Yang), isi.bjut@gmail.com (J.F. Qiao).

http://dx.doi.org/10.1016/j.asoc.2014.09.012

real applications such as nonlinear function approximation, pattern recognition and time-series prediction problems. The SANN training utilizing the BP algorithm (which can be henceforth abbreviate as BP-SANN) can obtain a higher training accuracy and generalization ability than multilayer feedforward neural networks (MLFNNs) and other neural network methods. However, there is still a major unresolved problem: there is more time need for training SANN and sometimes the learning efficiency is not very optimistic because of the existing of spatial span connections from different layers. This work is focused on how to reduce the training time and improve the learning efficiency without weakening the generalization properties.

According to the connect characteristics of artificial neurons, the adjustable output weights (linking to the output layer) of SANN can be divided into two types: the span-output-weight and the feedforward-output-weight. Based on this concept, this work proposes a two-phase sequential learning strategy for SANN called span-output-weight and feedforward-output-weight sequential learning (SFSL) whose convergence speed of learning can be faster than BP-SANN while obtaining better generalization performance. Therefore, the proposed learning scheme tends to have less training time and have better generalization ability for BP-SANN.

The paper is organized as follows. Section 2 first gives a brief review of SANN including the topology and gradient-based learning algorithm. In Section 3, we present the proposed two-phase sequential learning scheme SFSL, and performance evaluation of the SFSL-SANN is shown based on the benchmark problems in the areas of classification, regression, and time-series prediction is shown in Section 4. Finally, discussion and conclusions based on this work are highlighted in Section 5.

## 2. Brief review of SANN

The basic concepts, topology and back-propagation based learning algorithm of SANN are briefly reviewed in this section to provide the necessary background knowledge of the proposed two-phase sequential learning strategy of SANN. A brief summary and mathematical description of SANN are given first.

### 2.1. Summary and mathematical description of SANN

SANN, an ANN model with spatial span connection and local lateral inhibition, has been developed based on the neuron connecting mode of neocortex and the information transformation mechanism of lateral inhibition. It originates from the ANN theory and cerebral cortex anatomical knowledge, especially the structure of neocortex. Different from the unified MLFNNs, SANN introduces the span connections between neurons from any two nonadjacent layers, and also adds lateral inhibitory connection between neurons from the same hidden layer. These two innovations are corresponding to the connection of pyramid neurons span all the thickness of cortex and the lateral inhibition of interneurons in the same layer, respectively. The structure of the $(L+1)$-layer SANN is shown in Fig. 1.

In the following description, $l \in [0, L]$ is called the layer index of the $(L+1)$-layer SANN. Where, $l = 0$ denotes the input layer, $l \in [1, L-1]$ indicates the $l$th hidden layer, and the output layer index is $L$. The neuron number of the $l$th layer in SANN is defined as $n_l$. So, if a problem has $n$ inputs and $m$ outputs was learned by a SANN, then we have $n_0 = n$ and $n_L = m$.

Without loss of generality, consider $N$ arbitrary distinct samples $(x_q, t_q)|q \in [1, N]$, where $x_q = [x_{q1}, x_{q2}, \ldots, x_{qn}]^T \in R^n$ is a $n \times N$ input vector and $t_q = [t_{q1}, t_{q2}, \ldots, t_{qm}]^T \in R^m$ is a $m \times N$ target vector.The

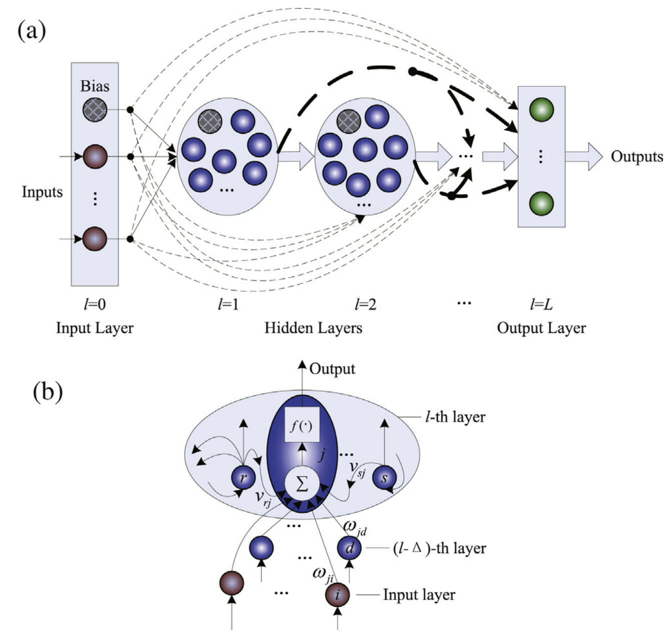


**Fig. 1.** (a) The structure of SANN and (b) the connecting mode of hidden neurons.

output of a $(L+1)$-layer SANN can be represented by

$$o_q = \sum_{k=1}^{m} \left\{ f^L \left[ \sum_{l=1}^{L-1} \left( \sum_{j=1}^{n_l} \omega_{kj} o_{qj}^l + b_l \right) + \sum_{i=1}^{n} \omega_{ki} x_{qi} + b_0 \right] \right\},$$
$$i \in [1, n], q \in [1, N] \quad (1)$$

where $x_{qi}$ is the $i$th input of the $q$th sample, $o_q$ is the output of the $q$th sample computed by SANN; $o_{qj}^l$ is the output of the $j$th hidden neuron from $l$th layer with respect to the $q$th input vector $x_{qi}$; $\omega_{kj}$ is the weight connecting the $j$th neuron of $l$th layer and the $k$th output neuron; $b_l$ is the threshold of the $l$th layer neuron; $f^l(\cdot)$ is the activation function of the $l$th layer neuron.

For the $j$th hidden neuron from $l$th layer with respect to the $q$th sample, the mathematical representation of $o_{qj}^l$ is given by

$$o_{qj}^l = f^l[x_{qj}^l - \tilde{x}_{qj}^l] = f^l \left[ \sum_{l=0}^{l-1} (\omega_{ij} o_{qi}^l) - \sum_{r=1, r \neq j}^{n_l} v_{rj} (o_{qr}^l - \theta_{rj}) \right] \quad (2)$$

where $x_{qi}^l$ is the inner product of input and weight, including signals received from the $(l-1)$th layer neurons and the spatial span connections from the range of $[0, l-2]$ layers, while $\tilde{x}_{qj}^l$ is the inhibitory input after lateral inhibited by surrounding neurons. $\theta_{rj}$ is the inhibiting threshold value of $j$th neuron due to $r$th neuron; $v_{rj}$ is the lateral inhibitory coefficient, which is a real number in [0, 1], with “0” meaning no inhibition and “1” meaning full inhibition.

### 2.2. Backpropagation-based learning algorithm of SANN

From the approximation point of view, a SANN with $L+1$ layers can approximate $N$ distinct training samples with any arbitrary error. The goal of SANN learning process is to find the best combination of connection weights and biases for achieving the minimum error between desired output and actual network output of