

Chapter 9

Locally Recurrent Neural Networks and Their Applications

Todor D. Ganchev
University of Patras, Greece

ABSTRACT

In this chapter we review various computational models of locally recurrent neurons and deliberate the architecture of some archetypal locally recurrent neural networks (LRNNs) that are based on them. Generalizations of these structures are discussed as well. Furthermore, we point at a number of real-world applications of LRNNs that have been reported in past and recent publications. These applications involve classification or prediction of temporal sequences, discovering and modeling of spatial and temporal correlations, process identification and control, etc. Validation experiments reported in these developments provide evidence that locally recurrent architectures are capable of identifying and exploiting temporal and spatial correlations (i.e., the context in which events occur), which is the main reason for their advantageous performance when compared with the one of their non-recurrent counterparts or other reasonable machine learning techniques.

INTRODUCTION

After the introduction of the simplified computational model of a neuron in the early 1940's (McCulloch & Pitts, 1943), various advanced models of neurons and derivative neural network structures were developed. Nowadays, i.e. sixty-five years later, there is a huge number of different neurons and immense number of neural networks, which differ by their building blocks, linkage, and structure. This abundance of architectures reflects the wide-spread use of neural networks on real-life problems, and the diversity of these problems attests to the inherent flexibility and enormous potential of neural networks.

DOI: 10.4018/978-1-60566-766-9.ch009

In the present chapter we focus our attention to one class of dynamic neural networks, namely the locally recurrent architectures that were introduced in the beginning of the 1990's, and that regained significant attention in the past few years.

The emergence of locally recurrent architectures contributed to significant advance in the theory and practice of neural networks. In the years following their introduction, numerous derivative architectures flourished, and the locally recurrent neurons become important building blocks in many original neural network designs, developed for the needs of particular real-world applications. In these applications, LRNNs were proved successful in dealing with spatial correlations, time sequences and structured data. Moreover, the engagement of LRNNs in real-world applications required with improved training strategies and variety of new architectures. The considerable research interest to LRNNs led to significant advance of the gradient-based learning techniques that had to be adapted for the needs of temporal sequences processing. Furthermore, numerous new training methods based on genetic algorithms, particle swarm optimization (PSO), and other evolutionary techniques appeared, diversifying the nomenclature of training methods and improving the applicability of LRNNs on specific domain-oriented tasks or in general.

Before proceeding with the main exposition, we would like to define few terms that we use through the text to avoid potential ambiguity:

First of all, we name *dynamic neural networks* these networks that incorporate dynamic synaptic or feedback weights among some or all of their neurons. These networks are capable of expressing dynamic behaviors.

Recurrent neural networks we label those architectures which incorporate feedback connections among the layers of the network, or those that do not have straightforward layered input-output architecture but instead the inputs flow forth and back among the nodes of the network.

Feedforward architectures that incorporate a layer of recurrent neurons, which possess feedbacks only from neurons belonging to the same layer, are referred to as *locally connected* recurrent neural networks. The locally connected architectures are also referred to as *non-local recurrent* structures.

Next, here we use the term *locally recurrent neural network* for any network that possesses neurons with local feedback connections. On the level of individual neurons, local recurrence is identified with the availability of one or more feedbacks that encompass one or more elements in the structure of a neuron. The following section discusses this matter in details.

Furthermore, on the level of neural networks (we assume multilayer neural networks) the local recurrent neurons might be organized in a feedforward, non-local recurrent, or global recurrent architectures, depending on the implementation of the network linkage. For further reference to terminology we direct the interested reader to the work of Tsoi and Back (Tsoi and Back, 1997), which offers a comprehensive overview of terms and architectures related to recurrent neural networks.

In the following sections we overview various neural network architectures that are based on locally recurrent neurons. We place the accent on detailing the structure of these neural networks and, therefore, the numerous training methods developed for these architectures are not covered in details. The interested reader shall follow the comprehensive references provided in order to obtain more information about the training methods used, and their performance on real-world problems.

In conclusion of this brief introduction, we would like to note that here we do not aim at exhaustive enumeration of all existing applications of LRNNs, nor do we intend to advocate or validate any LRNN-based solution against alternatives. Instead, we intend to list some examples of successful application of LRNNs on real-world problems, which illustrate the diversity and flexibility of these architectures, or which offer a perspective for prospective future applications. Thus, in the following sections, our

26 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/locally-recurrent-neural-networks-their/36986

Related Content

A General Knowledge Representation Model for the Acquisition of Skills and Concepts

Carlos Ramírez and Benjamín Valdés (2012). *International Journal of Software Science and Computational Intelligence* (pp. 1-25).

www.irma-international.org/article/a-general-knowledge-representation-model-for-the-acquisition-of-skills-and-concepts/88924

Improving Automated Planning with Machine Learning

Susana Fernández Arregui, Sergio Jiménez Celorrio and Tomás de la Rosa Turbides (2012). *Machine Learning: Concepts, Methodologies, Tools and Applications* (pp. 1355-1373).

www.irma-international.org/chapter/improving-automated-planning-machine-learning/56200

Applications of Data-Driven Modelling and Machine Learning in Control of Water Resources

D. P. Solomatine (2003). *Computational Intelligence in Control* (pp. 197-217).

www.irma-international.org/chapter/applications-data-driven-modelling-machine/6839

An Efficient Memetic Algorithm for Dynamic Flexible Job Shop Scheduling with Random Job Arrivals

Liping Zhang, Xinyu Li, Long Wen and Guohui Zhang (2013). *International Journal of Software Science and Computational Intelligence* (pp. 63-77).

www.irma-international.org/article/an-efficient-memetic-algorithm-for-dynamic-flexible-job-shop-scheduling-with-random-job-arrivals/88992

On Visual Semantic Algebra (VSA): A Denotational Mathematical Structure for Modeling and Manipulating Visual Objects and Patterns

Yingxu Wang (2009). *International Journal of Software Science and Computational Intelligence* (pp. 1-16).

www.irma-international.org/article/visual-semantic-algebra-vsa/37486