ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24 April 2015, i6doc.com publ., ISBN 978-287587014-8. Available from http://www.i6doc.com/en/.

Memory Transfer in DRASiW–like Systems

Massimo De Gregorio[1] and Maurizio Giordano[2]

1- Istituto di Cibernetica "Eduardo Caianiello" (ICIB – CNR)
Via Campi Flegrei, 34 – 80078 Pozzuoli (NA) – ITALY

2- Istituto di Calcolo e Reti ad Alte Prestazioni (ICAR – CNR)
Via P. Castellino, 111 – 80131 Naples – ITALY

Abstract. DRASiW is an extension of the WiSARD Weightless NN model with the capability of storing the frequencies of seen patterns during the training stage in an internal data structure called "mental image" (MI). Due to these capability, in previous work it was demonstrated how to reversely process MIs in order to generate synthetic prototypes from training samples. In this paper we show how DRASiW–like systems are able to transfer memory between different architectures while preserving the same functionalities.

1 Introduction

The knowledge from Artificial Neural Networks (ANN) that have learned one task can be reused on related tasks in a process that is called "transfer" [1]. In such cases, *knowledge transfer* or *transfer learning* between task domains would be desirable. In particular, *transfer learning* aims to extract the knowledge from one or more *source tasks* and to apply this knowledge to a *target task* [2].

The most important distinction between different types of transfer learning in ANN is *representational* versus *functional*. The former, is based on the idea to literally copy the trained network and train it on the new task; while, the latter does not involve the explicit assignment of prior task representation to a new task rather it employs the use of implicit pressures from supplemental training examples [3]. As reported in [1], source data is typically part of the training process in functional transfer if the source training data are available.

In this paper we propose a different approach to transfer learning in DRASiW–like systems. The new process, called *memory transfer*, is neither based on the availability of source training data nor on the aim of approaching related tasks. In particular, we focus the attention on how the functionalities of a DRASiW system can be transfer to another DRASiW system but with a totally different architecture. The aim of this work is to use the knowledge of the source system in order to create clones with different architectures but with the same functionalities. This is possible thanks to a particular characteristic DRASiW–like systems have: the capability of producing a pictorial representation of their stored information ("mental images").

In the next sections we are going to introduce the DRASiW–like systems and how to transfer the system functionalities between DRASiW systems through their mental images. Comparison of experimental results of DRASiW original *vs* cloned systems on four datasets, selected from UCI Machine Learning repository, shows low degradation of the proposed learning transfer process.

ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24 April 2015, i6doc.com publ., ISBN 978-287587014-8.
Available from http://www.i6doc.com/en/.

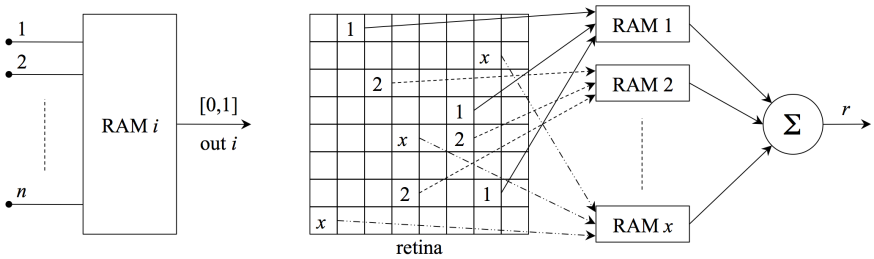

Fig. 1: RAM–neuron (left) and a WiSARD discriminator (right)

2 Weightless Neural Systems

Weightless Neural Networks (WNNs) [4][5], differently from classical Artificial Neural Networks, adopt a RAM–based model of neuron by which learned knowledge about a data domain is stored into RAM contents instead of computed weights of neuron connections. A RAM–based neuron receives an n–bit input that is interpreted as a unique address (*stimulus*) of a RAM location, used to access it either in writing (*learning*) or reading (*classification*) mode. WNNs have been shown to provide fast and flexible learning algorithms.

WiSARD systems are a particular type of WNN, that can be developed directly on reprogrammable hardware [6]. A WiSARD is composed of a set of classifiers, called *discriminators*, each one assigned to learn binary patterns belonging to a particular category/class. Therefore, a WiSARD has as many discriminators as the number of categories/classes it should be able to distinguish. The WiSARD is also called "multi–discriminator architecture".

Each discriminator is formed by a set of RAM nodes which store the information of occurrences of binary patterns during the learning stage. Given a binary pattern of size s, the so–called *retina*, it can be classified by a set of WiSARD discriminators, each one having x RAMs with 2^n locations such that $s = x \times n$. Since each RAM location is uniquely addressed by an n–tuple of bits, the input pattern can be partitioned into a set of n–tuples, each one addressing one location in exactly one RAM. n–tuples are pseudo–randomly selected and biunivocally mapped to RAMs (see right part of figure 1), in such a way that the retina is completely covered.

In order to train the discriminators one has to set all RAM locations to 0. For each training pattern, a 1 is stored in the memory location of each RAM addressed by the training pattern. During classification phase, the RAM contents addressed by the input pattern are read and summed by Σ. The number r thus obtained, called the *discriminator response*, is a sort of "similarity measure" of the input pattern with respect to the patterns in the training set (TS).

DRASiW [7] is an extension of the WiSARD model. During DRASiW training phase, memory locations accessed in write mode are incremented by 1 instead of set to 1. Thus, at the end of training, RAM contents store the number of

ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24 April 2015, i6doc.com publ., ISBN 978-287587014-8. Available from http://www.i6doc.com/en/.

occurrences (frequency) of a specific n–tuple of bits across training samples. In reading mode, the Σ adder of the DRASiW counts the number of non–zero values stored in the addressed RAM locations. Therefore a DRASiW system behaves like a WiSARD in the classification stage.

The DRASiW enhances the WiSARD with a backward–classification capability [8][9], i.e., each discriminator is able to produce representative examples of a class that have been learnt from trained patterns. In order to make this possible, RAM locations act as access counters, whose contents can be reversed to an internal "retina", where a "mental" image is produced, thus yielding a bidirectional structure. The "mental" image metaphor, associated with the internal "retina" metaphor, was originally explored in [10] in which the authors discuss the cognitive plausibilities related to these ideas.

3 From mental images to synthetic training set

Let us consider the mental image on the right side of figure 2. The image gray levels represent how many times the pixels were present in the instances of the original training set (some of these instances are shown on the left side of figure 2). The darker the pixel, more frequent is the bit associated to it. This is the only knowledge a mental image carries with it. In fact, we have no information about how the subpatterns (groups of correlated pixels) form the RAM addresses in each training sample. Hence, we expect that a memory transfer mechanism based on mental images will be characterize by a degradation in system performances although it will preserve the mental image.

The mental images can be depicted in 3D where the new dimension is represented by the pixel gray levels. Figure 3 shows the 3D representation (right side) of the mental image of class "2" (left side).

What we get is a sort of city downtown upper view where pixels are represented by skyscrapers. All the information is stored in the downtown except for the way it was built: we see the result but not the procedure to build it.

From the 3D representation of the mental image, we can create a *Synthetic Training Set* (STS) by considering the meaning of gray levels, that is how many times those pixels were black in the images forming the original TS. The highest gray level will represent the possible number of events, all the other bray levels will represent distinct favorable events. In order to reproduce a synthetic and plausible TS we randomly distribute the building floors along the size of the highest one. So doing, we generate a new downtown formed by skyscrapers

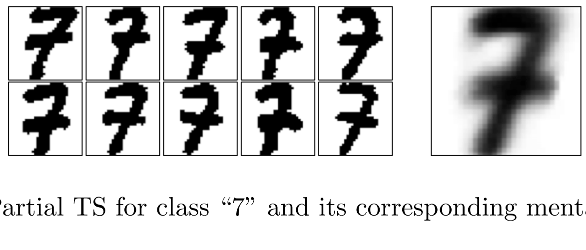

Fig. 2: Partial TS for class "7" and its corresponding mental image

ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24 April 2015, i6doc.com publ., ISBN 978-287587014-8. Available from http://www.i6doc.com/en/.

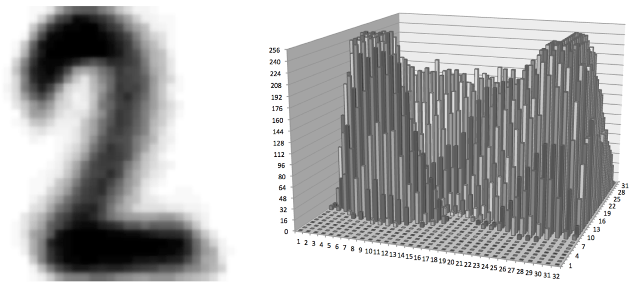

Fig. 3: Mental image and its 3D representation

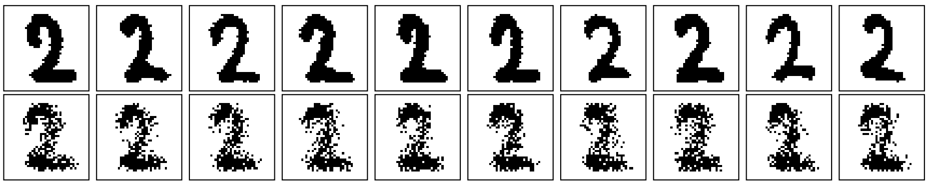

Fig. 4: Top: original training set (TS) – Bottom: synthetic training set (STS)

having all the same height but with real floors and missing floors from the ground to the top. At this point we start sinking the downtown floor by floor. In every sinking step, the patterns of the building floors are going to form the instances of the new STS (prototypes). In figure 4, prototypes of class "2" of the new STS generated by the mental image of figure 3 are reported. The STS is now used to train a new DRASiW–like system: the *clone* system.

The original system is mainly characterized by: a) the number of classes; b) the retina size; c) the n–bit addressing of RAMs. The number of classes fixes the number of DRASiW discriminators; the retina size together with the n–bit addressing determine the number of neurons for each DRASiW discriminator. The clone and the original system share only the number of discriminators and the capability of generating the same mental images. The retina and the n–bit addressing of the clone system do not depend on the original system.

4 Results

We evaluated the original and clone systems comparing them through their functionalities: *F–measures* (F_m) and mental images. The comparison was carried out on four different DataSets from the UCI ML Repository: *OptDigits*[1], *ImgSeg*[2], *Splice*[3], and *HAR*[4]. We chose these DataSets because of their differences in problem size and domain: 32×32 bitmap images (*OptDigits*); 19 numer-

[1]Optical Recognition of Handwritten Digits.

[2]Image Segmentation.

[3]Molecular Biology (Splice–junction Gene Sequences).

[4]Human Activity Recognition Using Smarphone.

ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24 April 2015, i6doc.com publ., ISBN 978-287587014-8. Available from http://www.i6doc.com/en/.

Trained On	OptDigits – retina=1024px			
TS	0.779 ± 0.014	0.924 ± 0.007	0.967 ± 0.003	0.977 ± 0.003
STS	0.738 ± 0.020	0.880 ± 0.007	0.922 ± 0.004	0.929 ± 0.004
Δ	*4.1%*	*4.4%*	*4.5%*	*4.8%*
	ImgSeg – retina=2432px			
TS	0.868 ± 0.015	0.899 ± 0.013	0.923 ± 0.012	0.951 ± 0.005
STS	0.850 ± 0.014	0.864 ± 0.010	0.874 ± 0.011	0.877 ± 0.015
Δ	*1.8%*	*3.5%*	*4.9%*	*7.4%*
	Splice – retina=30720px			
TS	0.380 ± 0.000	0.631 ± 0.010	0.778 ± 0.008	0.847 ± 0.011
STS	0.380 ± 0.000	0.601 ± 0.018	0.751 ± 0.012	0.788 ± 0.012
Δ	*0.0%*	*3.0%*	*2.7%*	*5.9%*
	HAR – retina=287232px			
TS	0.836 ± 0.005	0.853 ± 0.006	0.871 ± 0.006	0.892 ± 0.004
STS	0.823 ± 0.004	0.834 ± 0.005	0.843 ± 0.005	0.849 ± 0.004
Δ	*1.3%*	*1.9%*	*2.8%*	*4.3%*
	2 bits	4 bits	8 bits	16 bits

Table 1: F_m of the original and clone systems on different datasets

ical attributes associated to outdoor images (*ImgSeg*); 60–char strings representing gene sequences (*Splice*); 561 feature vectors with time and frequency domain variables (*HAR*). Furthermore, the retina size goes from 1024px in *OptDigits* to 287232px in *HAR*.

In order to test the systems (original *vs* clone), we generated 100 different retina–to-RAM mappings. For each mapping, we collected the F_m of the DRASiW systems trained on TS and STS, in different configurations (from 2 to 16 bits) and for each DataSet. Table 1 reports the F_m means and deviations and the differences (in percentage) between the F_m means of the original and clone systems (Δ).

Although the mental images do not bring with them the information about the correlation between subpatterns forming the original patterns, the performances of the clone systems are very comparable to those of the original systems. This loss of information and its consequences on the clone system performance become slightly more evident through increasing the number of bits and the retina size.

We also evaluated the systems in terms of the capability of generating plausible mental images: all the mental images generated by the clone systems are *exactly the same produced by the original systems*. This is due to the algorithm that generates mental images which does not depend on the original system architectures and configurations. Therefore, only upon the first cloning from the original system a degradation of F_m occurs: successive cloning processes from the first clone system do not imply further degradation of performance. The complete set of mental images generated by the DRASiW devoted to *OptDigits*[5] is reported in figure 5.

[5]We report only *OptDigits* mental images because their silhouettes are easily interpretable.

ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24 April 2015, i6doc.com publ., ISBN 978-287587014-8.
Available from http://www.i6doc.com/en/.

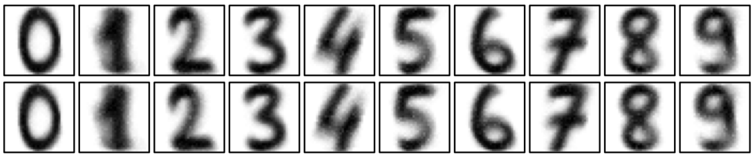

Fig. 5: Top: Original mental images – Bottom: Clone mental images

5 Conclusions

In this paper we have discovered and shown a new and interesting feature of DRASiW systems. A part from the new theoretical aspects related to Weightless Neural Systems, we have introduced a new methodology that can allow the use of the same DRASiW system in different applications and on various hardware configurations. For instance, smartphone applications in which we just have an empty and generic DRASiW system and use it just downloading, and hence, transferring the memory related to the new application domain.

In the wake of the results obtained, we are now facing the problem of memory graft in DRASiW systems. We are trying to merge the memories of two DRASiW systems in order to have just one system with new functionalities. In more details, we are addressing the problem of grafting the memory of a DRASiW system in part of the memory of another already trained system without catastrophic forgetting.

References

[1] L.Y. Pratt, J. Mostow, C.A. Kamm, and A.A. Kamm. Direct transfer of learned information among neural networks. In *Proceedings of AAAI–91*, pages 584–589, 1991.

[2] S.J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2010.

[3] D. Silver. *Selective Transfer of Neural Network Task Knowledge.* Ph.D. Thesis – University of Ontario, 2000.

[4] T.B. Ludermir, A. Carvalho, A.P. Braga, and M.C.P. Souto. Weightless neural models: a review of current and past works. *Neural Computing Surveys*, 2:41–61, 1999.

[5] I. Aleksander, M. De Gregorio, F.M.G. França, P.M.V. Lima, and H. Morton. A brief introduction to weightless neural systems. In *ESANN*, pages 299–305, 2009.

[6] I. Aleksander, W.V. Thomas, and P.A. Bowden. WISARD a radical step forward in image recognition. *Sensor Review*, 4:120–124, 1984.

[7] M. De Gregorio. On the reversibility of multi–discriminator systems. In *Technical Report 125/97*, Istituto di Cibernetica, CNR, 1997.

[8] C.M. Soares, C.L.F. da Silva, M. De Gregorio, and F.M.G. França. Uma implementação em software do classificador wisard. In *SBRN '98*, pages 225–229, 1998.

[9] B.P.A. Grieco, P.M.V. Lima, M. De Gregorio, and F.M.G. França. Producing pattern examples from "mental" images. *Neurocomputing*, 73(7–9):1057–1064, March 2010.

[10] E. Burattini, M. De Gregorio, and G. Tamburrini. Generation and classification of recall images by neurosymbolic computation. In *ECCM98*, pages 127–134, 1998.