Traffic Sign Recognition with VG-RAM Weightless Neural Networks

Mariella Berger, Avelino Forechi, Alberto F. De Souza, Jorcy de Oliveira Neto,
Lucas Veronese and Claudine Badue
Departamento de Informática
Universidade Federal do Espírito Santo
Vitória - ES, Brazil
email: mberger@lcad.inf.ufes.br

***Abstract*—Virtual Generalizing Random Access Memory Weightless Neural Networks (VG-RAM WNN) is an effective machine learning technique that offers simple implementation and fast training and test. In this paper, we present a new approach for traffic sign recognition based on VG-RAM WNN. We evaluate its performance using the German Traffic Sign Recognition Benchmark (GTSRB), a large multi-class classification benchmark. Our experimental results showed that our VG-RAM WNN architecture for traffic sign recognition was able to rank at 4th position in the GTSRB evaluation system, with a recognition rate of 98.73%, and was overcome by only one automatic approach.**

***Keywords*-Traffic Sign Recognition, VG-RAM Weightless Neural Networks, German Traffic Sign Recognition Benchmark.**

I. Introduction

Automatic traffic sign identification has many practical applications, such as traffic sign regulation, driver assistance and automated intelligent driving, and has been a challenging and active research topic in computer vision in the last years. However, the identification of traffic signs with large variations in visual appearance, due to deterioration, illumination changes, partial occlusions, rotation, weather conditions, etc., remains still a challenging problem.

The problem of traffic sign identification can be formulated as follows: given an image of a scene, identify one or more traffic signs in the scene using a database of traffic signs. The solution typically involves segmentation of traffic signs from scenes (traffic sign detection), feature extraction from the traffic sign regions and recognition. In this paper, we examined the recognition part of the identification problem only.

Traffic sign recognition is a multi-class classification problem with unbalanced class frequencies. The challenge lies in the fact that, even though there is a wide range of variations between classes in terms of color, shape and presence of pictograms or text, there exist classes very similar to each other (see Figure 1).

Many methods have been used to tackle the problem of traffic sign recognition, such as those based on support vector machines [1], [2], color distance transform [11], committee of convolutional neural networks [3], [4], and random forest of trees and kd-trees [5]. Virtual Generalizing Random Access Memory Weightless Neural Networks (VG-RAM WNN) [6] is an effective machine learning technique that offers simple implementation and fast training and test. In this paper, we present a new approach for traffic sign recognition based on VG-RAM WNN. In previous works [7], [8], we evaluated the performance of VG-RAM WNN on face recognition. Our experimental results showed that VG-RAM WNN are robust enough to handle various occlusions and illumination conditions, showing better performance than many well known techniques [8]. This has motivated us to use VG-RAM WNN for traffic sign recognition.

Figure 1. Samples of very similar traffic sign classes.

We evaluate the performance of VG-RAM WNN on traffic sign recognition using the German Traffic Sign Recognition Benchmark (GTSRB) (http://benchmark.ini.rub.de) [9], [10], a large multi-class classification benchmark. We chose the GTSRB dataset because we were interested in comparing our experimental results with those submitted to the GTSRB evaluation system, which includes the results yielded by the works presented in [1], [2], [11], [3], [4], [5]. Our experimental results showed that our VG-RAM WNN architecture for traffic sign recognition was able to rank at 4th position in the GTSRB evaluation system with a recognition rate of 98.73%. Even retaining the 4th position in the GTSRB rank, the VG-RAM WNN performance has been overcome by only one automatic approach.

This paper is organized as follows. After this introduction, in Section 2, we introduce VG-RAM WNN and, in Section 3, we describe how we have used them for traffic sign recognition. In Section 4, we present our experimental methodology and experimental results. Our conclusions and direction for future work follow in Section 6.

II. VG-RAM WNN

RAM-based neural networks, also known as n-tuple classifiers or weightless neural networks, do not store knowledge in their connections but in Random Access Memories (RAM) inside the network's nodes, or neurons. These neurons operate with binary input values and use RAM as lookup tables: the synapses of each neuron collect a vector of bits from the network's inputs that is used as the RAM address, and the value stored at this address is the neuron's output. Training can be made in one shot and basically consists of storing the desired output in the address associated with the input vector of the neuron [12].

In spite of their remarkable simplicity, RAM-based neural networks are very effective as pattern recognition tools, offering fast training and test, in addition to easy implementation [6]. However, if the network input is too large, the memory size becomes prohibitive, since it must be equal to 2^n, where n is the input size. Virtual Generalizing RAM (VG-RAM) Weightless Neural Networks (WNN) are RAM-based neural networks that only require memory capacity to store the data related to the training set [13]. In the neurons of these networks, the memory stores the input-output pairs shown during training, instead of only the output. In the test phase, the memory of VG-RAM WNN neurons is searched associatively by comparing the input presented to the network with all inputs in the input-output pairs learned. The output of each VG-RAM WNN neuron is taken from the pair whose input is nearest to the input presented—the distance function employed by VG-RAM WNN neurons is the *hamming distance*. If there is more than one pair at the same minimum distance from the input presented, the neuron's output is chosen randomly among these pairs.

Table I shows the lookup table of a VG-RAM WNN neuron with three synapses (X_1, X_2 and X_3). This lookup table contains three entries (input-output pairs), which were stored during the training phase (entry #1, entry #2 and entry #3). During the test phase, when an input vector (input) is presented to the network, the VG-RAM WNN test algorithm calculates the distance between this input vector and each input of the input-output pairs stored in the lookup table. In the example of Table I, the *hamming distance* from the input to entry #1 is two, because both X_2 and X_3 bits do not match the input vector. The distance to entry #2 is one, because X_1 is the only non-matching bit. The distance to entry #3 is three, as the reader may easily verify. Hence, for this input vector, the algorithm evaluates the neuron's output, Y, as class 2, since it is the output value stored in entry #2.

Table I
VG-RAM WNN NEURON LOOKUP TABLE.

Lookup Table	X1	X2	X3	Y
entry #1	1	1	0	class 1
entry #2	0	0	1	class 2
entry #3	0	0	1	class 2
	↑	↑	↑	↓
input	1	0	1	class 2

III. TRAFFIC SIGN RECOGNITION WITH VG-RAM WNN

Our VG-RAM WNN architecture for traffic sign recognition has a single two-dimensional array of $m \times n$ neurons, N, where each neuron, $n_{i,j}$, has a set of synapses, $W = (w_1, w_2, ...w_{|w|})$, which are connected to the network's two-dimensional input, Φ, of $m \times n$ inputs, $\varphi_{k,l}$ (Figure 2). The synaptic interconnection pattern of each neuron $n_{i,j}$, $\Omega_{i,j,\sigma^2}(W)$, follows a two-dimensional Normal distribution with variance σ^2 centered at φ_{μ_i,μ_j}; i.e., the coordinates k and l of the elements of Φ to which $n_{i,j}$ connects via W follow the probability density functions:

$$\omega_{\mu_i,\sigma^2}(k) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(k-\mu_i)^2}{2\sigma^2}} \tag{1}$$

$$\omega_{\mu_j,\sigma^2}(l) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(l-\mu_j)^2}{2\sigma^2}} \tag{2}$$

where σ^2 is a parameter of the architecture.

The position (μ_i, μ_j) of Φ (the center of the synaptic interconnection distribution of neuron $n_{i,j}$) is given by the inverse log-polar transform of the position (i, j) of the array of neurons as:

$$\mu_i = \frac{w}{2} + d\cos(\theta) \tag{3}$$

$$\mu_j = \frac{w}{2} + d\sin(\theta) \tag{4}$$

where

$$d = \frac{w}{2}\left(\frac{\alpha^{\left|\frac{i-\frac{m}{2}}{\frac{m}{2}}\right|} - 1}{\alpha - 1}\right) \tag{5}$$

$$\theta = \begin{cases} \pi(\frac{3n}{2} - \frac{j}{n}) + \frac{\pi}{2n}, & \text{if } i \leq \frac{m}{2} \\ \pi(\frac{3n}{2} + \frac{j}{n}) + \frac{\pi}{2n}, & \text{if } i > \frac{m}{2} \end{cases} \tag{6}$$

where α represents the concentration factor of the log-polar distribution and is also a parameter of the architecture.

Figure 3 illustrates this synaptic interconnection pattern, that mimics that observed in many classes of biological neurons [14], and is created when the network is built and does not change afterwards.

VG-RAM WNN synapses can only read bits from the input. Thus, in order to allow our VG-RAM WNN to deal with images, in which a pixel may assume a range of different values, we use minchinton cells [15]. In the proposed VG-RAM WNN architecture, each neuron's synapse, w_t, forms a minchinton cell with the next, w_{t+1} ($w_{|W|}$ forms a minchinton cell with w_1). The type of the minchinton cell we have used returns 1 if the synapse w_t of the cell is connected to an input element, $\varphi_{k,l}$, whose value is larger

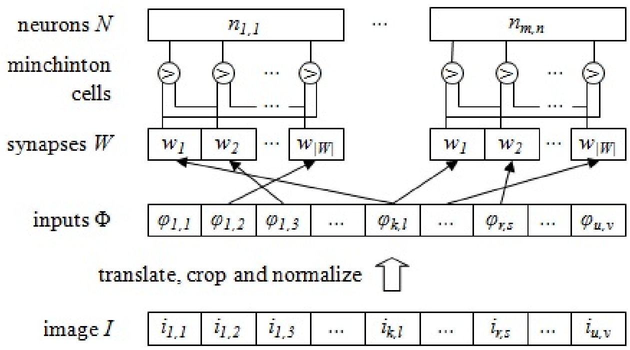


Figure 2. Schematic diagram of our VG-RAM WNN architecture for traffic sign recognition.

Figure 3. Neuron synaptic interconnection with the network input.

than the value of the element $\varphi_{r,s}$ to which the synapse w_{t+1} is connected, i.e., $\varphi_{k,l} > \varphi_{r,s}$; otherwise, it returns zero (see the synapses w_1 and w_2 of the neuron $n_{m,n}$ of Figure 2).

The input traffic sign images, I, are transformed before being copied to the networks input, Φ. The input images are first resized to 50×50 pixels, because the VG-RAM WNN architecture requires that all inputs have the same size (the sizes of the images in our benchmark varies from 15×15 to 250×250 pixels). Second, the input images are segmented in the RGB color space in three new images: red, green and blue images. Third, the segmented images are translated to compensate the decentralization of the region of interest. Fourth, they are cropped to keep only the region of interest. Finally, their contrast is normalized, using the Contrast-Limited Adaptive Histogram Equalization (CLAHE) [16], to minimize illumination variations.

During training, the pixels of the image I_x of a traffic sign t are copied to the VG-RAM WNN's input Φ and all $n_{i,j}$ neurons' outputs are set to the class identifier $c_t \in C = \{c_1, \ldots, c_{|C|}\}$, associated with the image of the traffic sign t ($|C|$ is equal to the number of known traffic sign classes). All neurons are then trained to output this class identifier with this input image. This procedure is repeated for all traffic sign images I_x in the training data set. During testing, each traffic sign image I_y is copied to the VG-RAM WNN's input Φ. Then, weights are assigned to neurons and their outputs are computed. After that, the number of votes to each class is computed as the sum of the weights associated with the neurons outputting that class. Finally, the network output is given by the class c_t with the largest number of votes.

The reason for assigning weights to neurons is as follows. Even after cropping the traffic sign images, they still present part of the background scene. This impacts negatively the VG-RAM WNN performance, because the neurons's synapses connected to the background regions would collect non-relevant information, which might generate ambiguous classification results. To minimize this effect, we assigned a distinct weight to each neuron.

Larger weights were attributed to neurons connected to regions near to the image center, while smaller ones to neurons monitoring regions near to image borders, since background regions typically appear on image corners. Consider the two-dimensional array of $m \times n$ neurons of our VG-RAM WNN architecture. The weights of these neurons follow an one-dimensional Normal distribution with variance σ_w^2 and mean $\frac{m}{2}$, where σ_w^2 is a parameter of the architecture. The same weight is attributed to all neurons along the same column.

Figure 4. Gaussian neuron weights.

IV. Experimental Results

To evaluate the performance of VG-RAM WNN on traffic sign recognition, we used the German Traffic Sign Recognition Benchmark (GTSRB) dataset (http://benchmark.ini.rub.de) [9], [10] The GTSRB contains 39,209 images in the training dataset and 12,630 images in the test dataset. Image sizes vary from 15×15 to 250×250 pixels.

For tuning the parameters of our VG-RAM WNN architecture, we generated a training and a test subset. For that, we randomly selected from the GTSRB training dataset 860 images for the training subset and 430 for the test subset.

Our VG-RAM WNN architecture has six parameters: (i) the number of neurons; (ii) the number of synapses per neuron; (iii) the size of the network input; (iv) the variance, σ^2, of the two-dimensional Normal distribution associated with the synaptic interconnection pattern of neurons; (v) the concentration factor, α, of the log-polar transform associated with the synaptic interconnection pattern of neurons; and (vi) the variance, σ_w^2, of the one-dimensional Normal distribution associated with the neuron's output weight. We tested our network with: (i) number of neurons equal to 34×18, 51×27

and 68×36; (ii) number of synapses per neuron equal to 32, 64, 128, 256 and 512; (iii) size of the network input equal to 50×50; (iv) σ^2 equal to 1, 3, 5, 7 and 9; (v) $\alpha = 2$; and (vi) σ_w^2 equal to 3.8, 4.7 e 5.7 for 34×18 neurons; 7.6, 9.4 e 11.3 for 51×27 neurons; and 11.3, 14.2 e 17 for 68×36 neurons (we did not vary the size of the network input and α to reduce the parameter search space).

The results of the experiments we carried out to tune the parameters of our VG-RAM WNN architecture show that the best configuration has around 51×27 neurons, 256 synapses per neuron, $\sigma^2 = 5$, and $\sigma_w^2 = 9.4$.

Figure 5 presents part of the results of the parameter tuning experiments. This figure shows the recognition rate (y axis) as a function of the number of synapses per neuron (x axis) and σ^2 (lines) for the best configuration of the number of neurons and σ_w^2 (51×27 neurons and $\sigma_w^2 = 9.4$). As Figure 5 shows, the performance (in terms of recognition rate) of the VG-RAM WNN architecture improves with σ^2; however, as σ^2 increases, the performance deteriorates. In the one hand, for smaller σ^2 values, the synaptic interconnection distribution of neurons is concentrated on a smaller region of the network input, which limits the amount of available information for neurons. In the other hand, for larger σ^2 values, the synaptic distribution is dispersed across a larger region of the network input and neurons may lose discriminative regions.

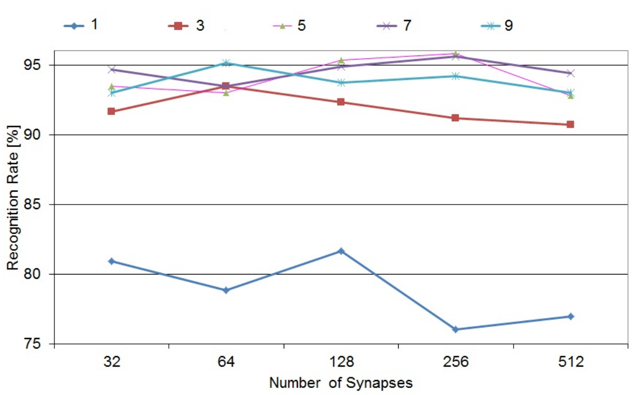
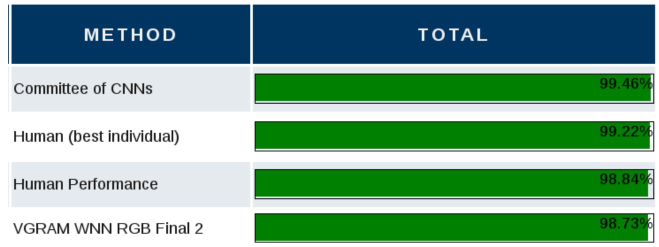


Figure 5. Recognition rate (y axis) as a function of the number of synapses per neuron (x axis) and σ^2 (lines) for the best configuration of the number of neurons and σ_w^2 (51×27 neurons and $\sigma_w^2 = 9.4$).

We submitted the results of the best VG-RAM WNN architecture configuration to the GTSRB evaluation system on September 15th 2012. Our VG-RAM WNN architecture for traffic sign recognition was ranked at the 4th position with a recognition rate of 98.73%. Figure 6 shows the result of this submission.



METHOD	TOTAL
Committee of CNNs	99.46%
Human (best individual)	99.22%
Human Performance	98.84%
VGRAM WNN RGB Final 2	98.73%

Figure 6. Result of the submission of our results to the GTSRB evaluation system on September 15th 2012. Our VG-RAM WNN architecture was ranked at the 4th position.

V. Conclusions and Future Work

In this paper, we presented a new approach for traffic sign recognition based on Virtual Generalizing Random Access Memory Weightless Neural Networks (VG-RAM WNN). Our experiments showed that VG-RAM WNN can be employed for traffic sign recognition with good accuracy (recognition rate of 98.73%). Even retaining the 4th position in the GTSRB rank, the VG-RAM WNN performance was overcome by only one automatic approach.

The main advantage of the VG-RAM WNN against other neural network approaches employed for traffic sign recognition is its simple implementation and fast training and test.

A direction for future work is to evaluate the performance of VG-RAM WNN on traffic sign recognition in Brazil's road environment.

VI. Acknowledgements

We would like to thank Conselho Nacional de Desenvolvimento Científico e Tecnológico-CNPq-Brasil (grants 552630/2011-0, 308096/2010-0, and 314485/2009-0) and Fundação de Amparo à Pesquisa do Espírito Santo-FAPES-Brasil grant 48511579/2009) for their support to this research work.

References

[1] S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Gil-Jimenez, H. GomezMoreno, F. Lopez-Ferreras, "Road-sign detection and recognition based on support vector machines", *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 264–278, Jun. 2007.

[2] R. Timofte, K. Zimmermann, L. Van Gool, "Multi-view traffic sign detection, recognition, and 3D localisation", *Workshop on Applications of Computer Vision*, pp. 1–8, Dec. 2009.

[3] P. Sermanet, Y. LeCun, "Traffic sign recognition with multi-scale Convolutional Networks", *The 2011 International Joint Conference on Neural Networks*, pp. 2809–2813, Aug. 2011.

[4] D. Ciresan, U. Meier, J. Masci, J. Schmidhuber, "A Committee of Neural Networks for Traffic Sign Classification", *The 2011 International Joint Conference on Neural Networks*, pp. 1918–1921, Aug. 2011.

[5] F. Zaklouta, B. Stanciulescu, O. Hamdoun, "Traffic sign classification using k-d trees and random forests", *The 2011 IEEE international joint conference on neural networks*, pp. 2151–2155, 2011.

[6] I. Aleksander, "From WISARD to MAGNUS: A Family of Weightless Virtual Neural Machines", *RAM-Based Neural Networks*, World Scientific, pp. 18–30, 1998.

[7] A. F. De Souza, C. Badue, F. Pedroni, E. Oliveira, S. S. Dias, H. Oliveira, S. F. Souza, "Face Recognition With VG-RAM Weightless Neural Networks", *Lecture Notes in Computer Science*, vol. 5163, no. 1, pp. 951–960, 2008.

[8] A. F. De Souza, C. Badue, F. Pedroni, S. S. Dias, H. Oliveira, S. F. Souza, "VG-RAM Weightless Neural Networks for Face Recognition", *Face Recognition*, pp. 171–186, InTech, 2010.

[9] J. Stallkamp, M. Schlipsing, J. Salmen, C. Igel, "The German Traffic Sign Recognition Benchmark: A multi-class classification competition", *The 2011 International Joint Conference on Neural Networks*, pp. 1453–1460, Aug. 2011.

[10] J. Stallkamp et al., "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition", *Neural Networks*, 2012.

[11] A. Ruta, Y. Li, X. Liu, "Real-time traffic sign recognition from video by class-specific discriminative features", *Pattern Recognition*, vol. 43, pp. 416–430, Jan. 2010.

[12] I. Aleksander, "Self-Adaptive Universal Logic Circuits", *IEEE Electronic Letters*, pp. 231–232, 1966.

[13] T. B. Ludermir, A. C. P. L. F. Carvalho, A. P. Braga, M. D. Souto, "Weightless Neural Models: A Review of Current and Past Works", *Neural Computing Surveys*, vol. 2, pp. 41–61, 1999.

[14] E. R. Kandel, J. H. Schwartz, T. M. Jessell, *Principles of Neural Science*, Prentice-Hall International Inc, 2000.

[15] R. J. Mitchell, J. M. Bishop, S. K. Box, J. F. Hawker, "Comparison of Some Methods for Processing Grey Level Data in Weightless Networks", *Based Neural Networks*, World Scientific, pp. 61–70, 1998.

[16] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization", *Graphics Gems IV*, Academic Press, pp. 474-485, 1994.