ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

Enhancing classification performance at the RAM-neuron level

Gianluca Coda,[1] Alessandro De Gregorio,[2] Massimo De Gregorio,[1]
Antonio Sorgente[1] and Paolo Vanacore[3]

[1]Istituto di Scienze Applicate e Sistemi Intelligenti – CNR, Italy
[2]PRISCA Lab, DIETI – Università di Napoli "Federico II", Italy
[3]Area Territoriale di Ricerca Napoli 1 – CNR, Italy

Abstract.
Even though RAM-neurons may have different discriminative reliability, DRASiW treats their contribution uniformly. In this work, we introduce the new metric RDA (RAM Discrimination Amplifier) that assigns to each RAM a class-specific amplification (or reduction) factor. This factor is calculated from divergence metrics applied to the RAM address distributions. RDA preserves the weightless nature of DRASiW while improving the quality of decision. Experiments on 41 datasets show consistent *accuracy* and *f1-score* gains on different evaluation protocols.

1 Introduction

Weightless Neural Networks (WNNs) [1] are an alternative paradigm to traditional connectionist models [2] in which weighted synapses are replaced with memory-based units that store and retrieve binary patterns directly. These include WiSARD (Wilkie, Stonham and Aleksander Recognition Device) [3] which is one of the most investigated architectures because it is interpretable, trains fast and can be easily implemented in hardware.

Over the years, several extensions have been proposed to improve WiSARD's performance. The standard WiSARD performs classification by counting the number of RAM units (discriminators) that recognize the input pattern for each class. To handle ambiguous or overlap responses, the DRASiW (the first WiSARD extension) introduced the bleaching procedure [4, 5] that iteratively increases a threshold to select only those RAMs whose activation exceeds the threshold. Later, new variants such as b-bleaching [6], DAB and rDAB [7] have been introduced. Other works have focused on improving the input representation and mapping, showing how different encoding schemes [8, 9] and input-to-RAM allocation strategies (mapping) [10] affect classification accuracy. More recent works have explored hybrid symbolic-subsymbolic approaches [11, 12, 13] and models based on hierarchical decomposition method [11] exploiting internal representation.

All these improvements focus on maximising the contribution of RAMs in the search for optimal parameters, either in the decision-making process or in the structure. No approach explicitly quantifies the discriminating value of each individual RAM. In this work, we introduce the RAM Discrimination Amplifier

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

(RDA), which measures how class-selective each RAM is after training. The underlying idea is simple: although DRASiW treats all RAM units in the same way when aggregating them, in practice some RAMs can be more reliable than others. Some RAMs tend to activate specifically for one class, while others respond ambiguously to multiple classes, providing little discriminating information.

The most similar approach to ours is described in [14] where technique how to cut unnecessary RAMs is proposed. In [14], a method was proposed to evaluate the contribution of a single RAM based on its correct/incorrect activation in the dataset (empirical evaluation). The objective of that study is to prune the RAMs in order to reduce memory usage, which will also result in a loss of accuracy. In contrast, our approach aims to improve the performance of the model by introducing a measure that indicates how well a RAM discriminates between classes. This score is used to define a coefficient of amplification (or reduction) of the RAM contribution during the prediction phase.

With the proposed RDA approach, we quantify the discriminative behaviour of each RAM and assign a discrimination coefficient proportional to its ability to distinguish between classes. This mechanism does not change the network's learning mechanism: the DRASiW remains fully weightless and interpretable. What we introduce is a reliability based modulation applied during inference, which refines the decision process without altering the stored knowledge.

2 RAM Discrimination Amplifier (RDA)

The refined Dynamic Adaptive Bleaching (rDAB) [7] procedure extends the standard DRASiW model by exploiting the frequency information stored in each RAM cell. Instead of simply counting how many times a memory address was activated, each RAM accumulates the relative frequency of activation for each address, normalized with respect to the total number of training instances of the class associated to that RAM. As a result, the internal state of each RAM already encodes an empirical probability distribution of address activations, which can be interpreted as:

$$P_{j,c}(a) = \text{normalized content of RAM}_j \text{ at address } a \text{ for class } c. \tag{1}$$

This distribution $P_{j,c}$ can therefore be regarded as the statistical "signature" of class c in RAM_j.

Given these class-specific probability distributions, we define the value RDA as a measure of how well a RAM distinguishes between classes, based on the distance between its internal probability distributions. In general, the *discrimination score* (*ds*) between two classes c_1 and c_2 in a RAM_j can be expressed as:

$$ds(P_{j,c_1}, P_{j,c_2}) = \mathcal{D}(P_{j,c_1} \| P_{j,c_2}), \tag{2}$$

where $D(\cdot\|\cdot)$ denotes a generic divergence or distance measure between probability distributions (*e.g.*, KL, JS, Hellinger, Bhattacharyya, Wasserstein). In our experiments, we adopted, as distance measure between probability distributions

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

($\mathcal{D}$), the Jensen-Shannon divergence (equations 3 and 4) due to its symmetry, boundedness, and numerical stability, making it particularly suitable for discrete probability distributions derived from DRASiW RAMs. Therefore

$$JS(P_{j,c_1} \| P_{j,c_2}) = \frac{1}{2} KL(P_{j,c_1} \| M) + \frac{1}{2} KL(P_{j,c_2} \| M), \tag{3}$$

where $M = \frac{1}{2}(P_{j,c_1} + P_{j,c_2})$ denotes the average distribution, and KL denotes the Kullback-Leibler divergence defined as:

$$KL(P \| Q) = \sum_a P(a) \log \frac{P(a)}{Q(a)}. \tag{4}$$

From these distances, we derive for each RAM_j a class-specific score $DS_{j,c}$, as defined in Equations 5. This score quantifies how strongly RAM_j separates class c from all the others.

$$DS_{j,c} = \frac{1}{C-1} \sum_{c'=1 \ c' \neq c}^{C} JS(P_{j,c} \| P_{j,c'}). \tag{5}$$

Because the absolute scale of $DS_{j,c}$ depends on the dataset and on the number of classes, we convert the raw discrimination scores into a relative amplification factor, calibrated so that the average RAM has amplification equal to 1. To this end, we define the $RDA_{j,c}$ value as:

$$RDA_{j,c} = \frac{DS_{j,c}}{\frac{1}{n_{\text{RAM}}} \sum_{k=1}^{n_{\text{RAM}}} DS_{k,c}} \tag{6}$$

which expresses how much more (or less) discriminative RAM_j is compared to the average RAM for the same class c. Thus, RAMs with $RDA_{j,c} > 1$ are therefore more discriminative than average for class c, while those with $RDA_{j,c} < 1$ contribute less.

Finally, the $RDA_{j,c}$ values are used during prediction. Each RAM no longer contributes with a binary vote (0 or 1), but instead with a value scaled proportionally to its corresponding RDA value, thereby modulating its influence proportionally to its discriminative reliability.

3 Experimental Setup

For the evaluation of the proposed method, we performed tests on 41 public datasets for classification problems, using three distinct evaluation protocols. The goal is to measure the improvement, or gain, in model performance by evaluating the *f1-score* and *accuracy*.

The initial evaluation was performed on 26 classification datasets using two standard cross-validation protocols: 5-fold cross-validation (5cv) and 10-fold cross-validation (10cv). A further experiment (called TT) was performed using 15 additional datasets, which were divided into a dedicated training set and

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

Table 1: Results on 5cv and 10cv protocol (best results in bold)

	5cv				**10cv**			
	Differences		Gains		Differences		Gains	
Datasets	f1	acc.	f1	acc.	f1	acc.	f1	acc.
Alzheimer	0.0006	0.0000	0.0012	0.0000	0.0000	0.0118	0.0000	0.0267
Balance	0.0057	0.0016	0.0204	0.0141	**0.0252**	0.0000	0.0832	0.0000
Cleveland	0.0130	0.0033	0.0201	0.0079	0.0004	0.0033	0.0007	0.0080
Contraceptive	0.0035	0.0041	0.0075	0.0088	0.0058	0.0068	0.0124	0.0148
Ecoli	0.0058	0.0030	0.0211	0.0197	0.0000	0.0000	0.0000	0.0000
Glass	0.0083	0.0047	0.0292	0.0199	0.0069	0.0180	0.0234	0.0755
Hayes	0.0090	0.0062	0.0525	0.0312	0.0166	**0.0187**	0.0977	0.0937
Iris	0.0067	0.0067	0.2010	0.2000	0.0000	0.0000	0.0000	0.0000
Marketing	0.0002	0.0006	0.0002	0.0009	0.0002	0.0016	0.0002	0.0025
Mov_lib	0.0000	0.0000	0.0000	0.0000	0.0017	0.0028	0.0140	0.0244
Newthyroid	0.0155	**0.0093**	**0.5198**	**0.5000**	0.0000	0.0000	0.0000	0.0000
Optdigits	0.0002	0.0002	0.0112	0.0114	0.0000	0.0000	0.0000	0.0000
Page_blocks	0.0010	0.0009	0.0053	0.0259	0.0035	0.0002	0.0176	0.0055
Penbased	0.0003	0.0003	0.0342	0.0345	0.0000	0.0000	0.0000	0.0001
Satimage	0.0004	0.0005	0.0032	0.0052	0.0001	0.0000	0.0005	0.0000
Segment	0.0004	0.0004	0.0182	0.0182	0.0000	0.0000	0.0000	0.0000
Shuttle	0.0241	0.0002	0.3115	0.2745	0.0106	0.0001	**0.1580**	0.1463
Tae	0.0020	0.0067	0.0059	0.0201	0.0136	0.0133	0.0430	0.0431
Texture	0.0002	0.0002	0.0087	0.0096	0.0000	0.0000	0.0006	0.0000
Thyroid	**0.0275**	0.0069	0.1372	0.1701	0.0236	0.0075	0.1174	**0.1812**
Vehicle	0.0040	0.0048	0.0149	0.0184	0.0063	0.0059	0.0232	0.0227
Vowel	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0067	0.0000
Wine	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Wine_Red	0.0009	0.0012	0.0015	0.0040	0.0033	0.0000	0.0056	0.0000
Wine_White	0.0015	0.0010	0.0027	0.0033	0.0017	0.0004	0.0032	0.0014
Yeast	0.0062	0.0013	0.0149	0.0035	0.0000	0.0047	0.0000	0.0123
Max	0.0275	0.0093	0.5198	0.5000	0.0252	0.0187	0.1580	0.1812
Average	0.0053	0.0025	0.0555	0.0539	0.0046	0.0037	0.0234	0.0253

a separate test set. This protocol is crucial as it allows us to simulate real-world scenarios in which the model is tested on data never seen before during the training phase. Furthermore, to quantify the effectiveness of the proposed method, based on the *f1-score* and *accuracy* metrics, we evaluate the *difference* and the *gain* (or the *improvement ratio*). *difference* denotes the absolute change in a metric, while *gain* indicates the relative improvement in a metric. For a generic system S, they are defined as follows:

$$\textit{difference} = \text{Metric}_{S+\text{RDA}} - \text{Metric}_S, \quad \textit{gain} \quad = \frac{\textit{difference}}{1 - \text{Metric}_S}. \tag{7}$$

4 Results

The results consistently demonstrate that the proposed method yields positive gains in both *f1-score* and *accuracy* across all experimental protocols (see Tables 1 and 2). The TT experiment exhibited the largest average improvements, with a mean *gain* of approximately 9.4% in *f1-score* and 9.3% in *accuracy*. This substantial improvement suggests that the method is particularly effective in boosting performance on challenging datasets that are typically large and used with a fixed train/test split.

While the average improvement is significant, the method also demonstrated substantial gains on specific datasets, highlighting its effectiveness for certain classification problems. In the 5cv protocol, the maximum *f1-score gain* was observed on the *Newthyroid* dataset (*gain* ≈ 0.5198). In the TT protocol, sig-

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

Table 2: Results on TT protocol (best results in bold)

TT	Differences		Gains	
Datasets	f1-score	accuracy	f1-score	accuracy
Abalone	0.0010	0.0000	0.0011	0.0000
Arcene	0.0000	0.0000	0.0000	0.0000
Avila	0.0056	0.0113	0.0203	0.0354
Forest	0.0155	0.0123	0.0872	0.0784
HAR	0.0002	0.0000	0.0017	0.0000
Hill_valley	0.0168	0.0000	0.0337	0.0000
Img_seg	0.0024	0.0024	0.0330	0.0329
Madelon	0.0061	0.0067	0.0167	0.0183
Monks1	**0.0417**	**0.0417**	0.2953	0.2951
Monks2	0.0178	0.0162	0.2954	0.3043
Monks3	0.0116	0.0116	**0.3327**	**0.3333**
Occupancy	0.0076	0.0052	0.2402	0.2440
Pendigit	0.0003	0.0003	0.0126	0.0135
Statlog	0.0013	0.0000	0.0117	0.0000
Wilt	0.0047	0.0080	0.0203	0.0388
Max	0.0417	0.0417	0.3327	0.3333
Average	0.0088	0.0077	0.0935	0.0929

nificant gains were recorded on the *Monks3* (*gain* ≈ 0.3327) and *Monks2* (*gain* ≈ 0.2954) datasets. The overall reduction in average *gain* from 5cv (0.0555) to 10cv (0.0234) is expected, as 10-fold cross-validation typically provides a more stable, less variant, and thus more conservative estimate of performance than 5-fold cross-validation.

To assess the robustness of the results, a one-sample **t-test** was performed on the distribution of gains for each metric and protocol, testing the null hypothesis (H_0) that the mean *gain* is less than or equal to zero ($\mu \leq 0$) against the alternative hypothesis (H_a) that the mean *gain* is significantly greater than zero ($\mu > 0$). As shown in Table 3, the p-values for all six scenarios are significantly low. These results allow us to reject the null hypothesis at a typical significance level of $\alpha = 0.05$. This provides strong statistical evidence that the proposed method delivers a statistically significant positive improvement in both *f1-score* and *accuracy* across all three evaluated experimental protocols.

Table 3: Summary of average gain statistics

	5cv		10cv		TT	
	f1-score	*accuracy*	*f1-score*	*accuracy*	*f1-score*	*accuracy*
average gain	0.0555	0.0539	0.0234	0.0253	**0.0935**	**0.0929**
t-statistic	2.3786	2.4135	2.8215	2.7223	2.8677	2.8005
p-value	0.0253	0.0235	0.0092	0.0116	0.0124	0.0142

5 Conclusion

The experimental evidence, spanning 41 diverse datasets and three common evaluation protocols (5cv, 10cv, TT), strongly supports the effectiveness of the

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

proposed method: it consistently generated a positive mean *gain* in both *f1-score* and *accuracy* under all testing conditions; a one-sample **t-test** confirms that the observed performance gains are statistically significant ($p < 0.05$), validating the method's practical utility; the largest mean improvements were observed in the TT protocol, suggesting a strong benefit in scenarios where the generalization capability to a completely unseen test set is essential.

In conclusion, the new method is a robust and statistically validated method for significantly improving the performance of our system in classification tasks, particularly where model generalization is critical.

References

[1] 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.

[2] J.A. Feldman and D. H Ballard. Connectionist models and their properties. *Cognitive science*, 6(3):205–254, 1982.

[3] Igor Aleksander and Helen Morton. *An introduction to neural computing*. Chapman and Hall, 1990.

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

[5] C.M. Soares, C.L.F. da Silva, M. De Gregorio, and F.M.G. França. Uma implementação em software do classificador WiSARD. In *V Simpósio Brasileiro de Redes Neurais*, volume 2, pages 225–229, 1998.

[6] D.S. Carvalho, H.C.C. Carneiro, França F.M.G., and Lima P.M.V. B-bleaching: agile overtraining avoidance in the WiSARD weightless neural classifier. In *ESANN*, pages 515–520, 2013.

[7] G. Coda, M. De Gregorio, A. Sorgente, and P. Vanacore. Improving the DRASiW performance by exploiting its own "Mental Images". In *ESANN*, pages 363–368, 2023.

[8] Alan T., Bacellar L., Susskind Z., Villon L.A.Q., Miranda I.D. S., de Araújo L.S., Dutra D.L.C., Breternitz Jr. M., Lizy J., Lima P.M.V., and França F.M.G. Distributive thermometer: a new unary encoding for Weightless Neural Networks. In *ESANN*, pages 31–36, 2022.

[9] A. Kappaun, K. Camargo, F. Rangel, F. Firmino, P. M. V. Lima, and J. Oliveira. Evaluating binary encoding techniques for WiSARD. In *2016 5^{th} Brazilian Conference on Intelligent Systems (BRACIS)*, pages 103–108, 2016.

[10] M. Giordano and M. De Gregorio. An evolutionary approach for optimizing Weightless Neural Networks. In *ESANN*, pages 325–330, 2019.

[11] G. Coda, M. De Gregorio, A. Sorgente, and P. Vanacore. "Mental Images" driven classification. In *ESANN*, pages 197–202, 2024.

[12] P. Coutinho, H.C.C. Carneiro, D.S. Carvalho, and F.M.G. França. Extracting rules from DRASiW's "Mental Images". In *ESANN*, pages 529–534, 2014.

[13] B.P.A. Grieco, P.M.V. Lima, M. De Gregorio, and F.M.G. França. Extracting fuzzy rules from "mental" images generated by a modified WiSARD perceptron. In *ESANN*, pages 313–318, 2009.

[14] Z. Susskind, A.T.L. Bacellar, A. Arora, L. A.Q. Villon, R. Mendanha, L. Santiago de Araújo, D. L. C. Dutra, P. M.V. Lima, F. M.G França, I.D.S. Miranda, M. Breternitz Jr, and L.K. John. Pruning Weightless Neural Networks. In *ESANN*, pages 37–42, 2022.