

# Advanced Fuzzy Relational Neural Network

E. Di Nardo<sup>1,2</sup>, A. Ciaramella<sup>2</sup>

<sup>1</sup>Department of Computer Science, University of Milan, Milan 20122, Italy

<sup>2</sup>Department of Science and Technology, University of Naples Parthenope, Centro Direzionale Isola C4, I-80143, Napoli, Italy

## Abstract

Nowadays most of the researches aimed for studying artificial neural networks and in particular convolutional neural networks for the impressive results in several scientific fields. However, these methodologies need of post-hoc technique for improving their interpretability and explainability. In the last years, fuzzy systems are raising great interest for the simplicity to develop trustworthy and explainable systems. This work aims to introduce a fuzzy relational neural network based model for extrapolating relevant information from images data permitting to obtain a clearer indication on the classification processes. Encouraging results are obtained on benchmark data sets.

## Keywords

Deep Learning, Fuzzy Logic, Fuzzy Relational Neural Network, Computational Intelligence

## 1. Introduction

In recent years the Convolutional Neural Networks (CNNs) are playing a very important role for the results obtained in various scientific fields. The scientific community, however, under the guidance of the GDPR is experiencing considerable interest in eXplainable Artificial Intelligent (XAI). Techniques for XAI can be model agnostic (i.e. they can be applied to any AI algorithm), or model specific (i.e. they can be only applied to a specific AI algorithm). Moreover, they can be ante-hoc (transparent or “white box/glass box” approaches explainable by design or inherently explainable) or post-hoc (divided into global explanations or local explanations) explainability methods [1]. Ante-hoc methods are explainable by design and great interest in the last years is on Fuzzy Rule-based systems [2, 3, 4]. In this work we propose a fuzzy relational neural network that is based on a fuzzy inference scheme for the classification of images.

The paper is organized as follows. In Section 2 we present the proposed methodology and the used methods. Furthermore, in Section 3 the results of experiments on benchmark data are presented. Finally, the authors draw conclusions in Section 4.

---

*The 13th International Workshop on Fuzzy Logic and Applications (WILF 2021)*


✉ emanuel.dinardo@unimi.it (E. Di Nardo); angelo.ciaramella@uniparthenope.it (A. Ciaramella)

🌐 <https://sites.google.com/view/ciss-angelociaramella/home> (A. Ciaramella)

🆔 0000-0002-6589-9323 (E. Di Nardo); 0000-0001-5592-7995 (A. Ciaramella)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

## 2. Fuzzy Relational Neural Network model

Fuzzy Rule-based Systems (FRSs), are raising great interest in XAI in the last years as ante-hoc methodologies [1]. The main components of any FRS are the knowledge base (KB) and the inference engine module. The KB comprises all the fuzzy rules within a rule base (RB), and the definition of the fuzzy sets in the data base. The inference engine includes a fuzzification interface, an inference system, and a defuzzification interface [5, 4]. Fuzzy Relational Neural Network (FRNN) [6] is an adaptive model based on a FRS. FRNN can be developed with different norms and a backpropagation algorithm is used for learning. In this work we model local t-norms modifying the inner operation of convolution and replacing the linear combination provided by matrix multiplication with fuzzy operators. We define a receptive field that applies a triangular operation on a restricted area. As happens in convolution we have a kernel of size  $N \times M \times C_{in} \times C_{out}$  where  $N$  and  $M$  are the spatial dimensions,  $C_{in}$  is the number of input channels and  $C_{out}$  the number of output feature maps. Kernel slides over the image with a parametric step. Weights are initialized in range  $[0, 1]$  and constrained to be in the same interval after the optimization step using a scaling operation based on minimum and maximum values:

$$w_{out} = \frac{w - \min(w)}{\max(w) - \min(w)} \quad (1)$$

The network structure is composed by an input layer and a fuzzification layer where the membership function is just a scaling of the pixel value in range  $[0 - 1]$ . We compare the results by using one or two hidden layers. Next there is a defuzzification operation that is composed by a fully connected layer like in [6] and an output layer with a Categorical Crossentropy is used for classification. Architectures have been tested with and without a threshold activation function, a modification of leaky relu with a minimum boundary  $> 0$ . Networks are compared with equivalent CNN architectures.

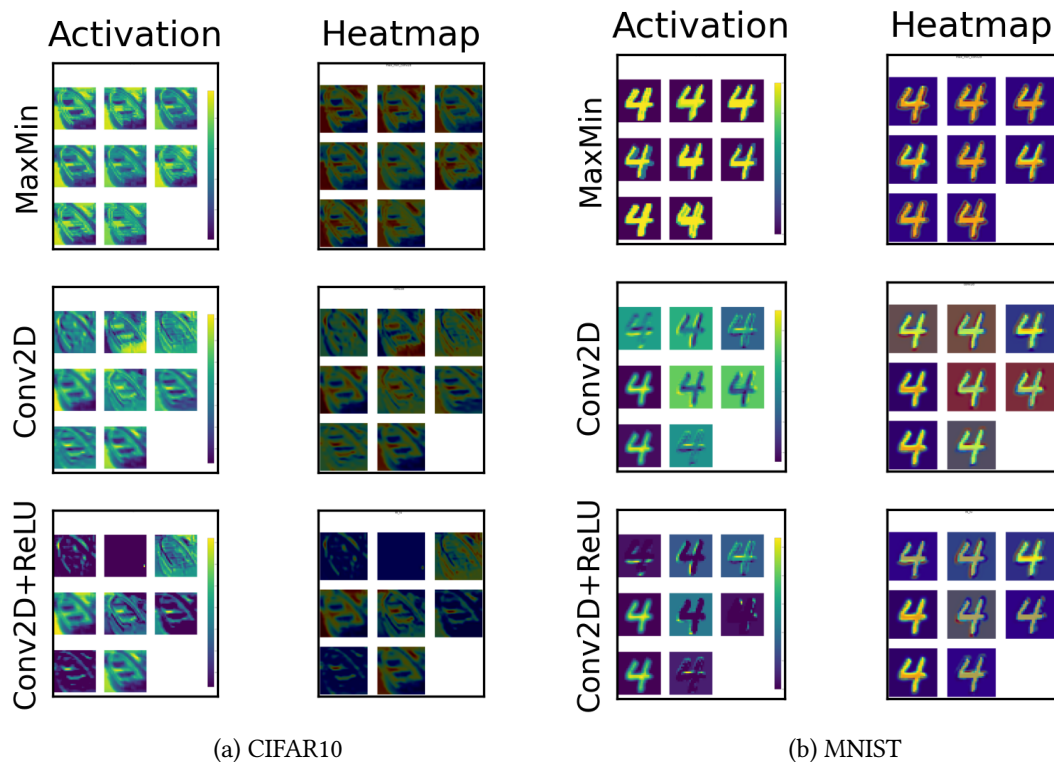
## 3. Experimental results

FRNN has been applied for images classification and the MNIST [7] and CIFAR10 [8] datasets are considered. Input images are scaled in range  $[0 - 1]$  as fuzzification step. The single hidden layer architecture uses a feature map of size 8, in the two hidden layers setup there are respectively feature maps of size 8 and 16. Weights are randomly initialized using a uniform distribution in range  $[0 - 1]$  in order to define a random degree order. Further weights are constrained to be in the same range after the backpropagation phase, because they have to define at any moment a data membership degree for all channels. There is a soft constraint that re-scales weights in the correct range without a hard clipping of the weights on boundaries, as in the gradient clipping case. All layers have a kernel size of 3 on spatial dimensions. In table 1 it is possible to check performance of CNN compared with fuzzy architectures. It is possible to observe that performances on MNIST are comparable but on CIFAR10, CNN outperform FRNN. However, observing the activations and the heatmaps (Fig. 1) of the models and some others visualizations based on GradCAM [9], Gradients\*Inputs [10] and Integrated Gradients [11] (Fig. 2, 3) it is possible to note that FRNN can explain more accurately the information used in

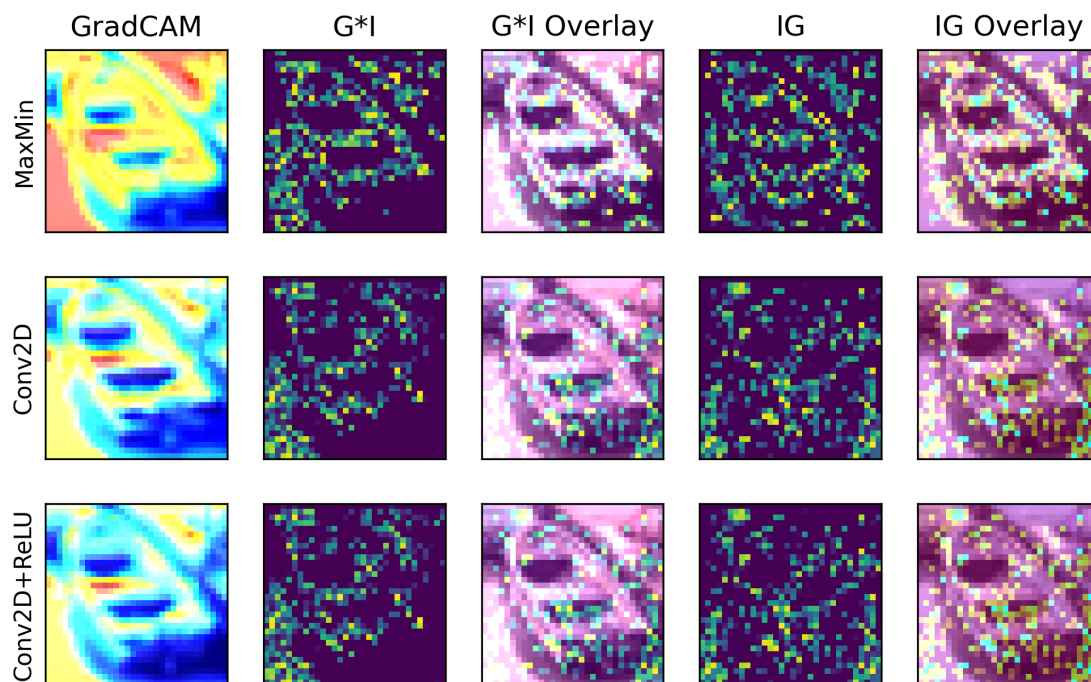
classification. GradCAM shows that conv2d and relu function cut-out important features of the object retaining the non relevant one, instead, FRNN, also if unrelevant area are present, is able to preserves the shape of the ships. It is more clear also observing the gradients. Both techniques have some noise, but the fuzzy module is able to focus more on the image subject. The same analysis is possible observing MNIST fig. 3 where gradients, the attention of the network, are located following the object shape. This statement is not valid for convolutional layer that shows a lot of noise and is unable to focus on the main subject.

Model	MNIST					CIFAR10				
	Training		Validation		Epochs	Training		Validation		Epochs
	Acc	Loss	Acc	Loss	—	Acc	Loss	Acc	Loss	—
Conv2D+ReLU	0.99	0.007	0.97	0.1	35	0.59	1.17	0.53	1.29	800
MaxMin	0.99	0.008	0.97	0.082	100	0.44	3.83	0.38	4.45	100
MaxMin (2L)	0.99	7.9e-4	0.98	0.09	90	0.36	9.23	0.35	11.7	260

**Table 1**  
Classification results of Convolutional Neural Networks and Fuzzy Neural Networks on MNIST and CIFAR10 datasets



**Figure 1:** Activations and heatmaps of layers on CIFAR10 and MNIST datasets

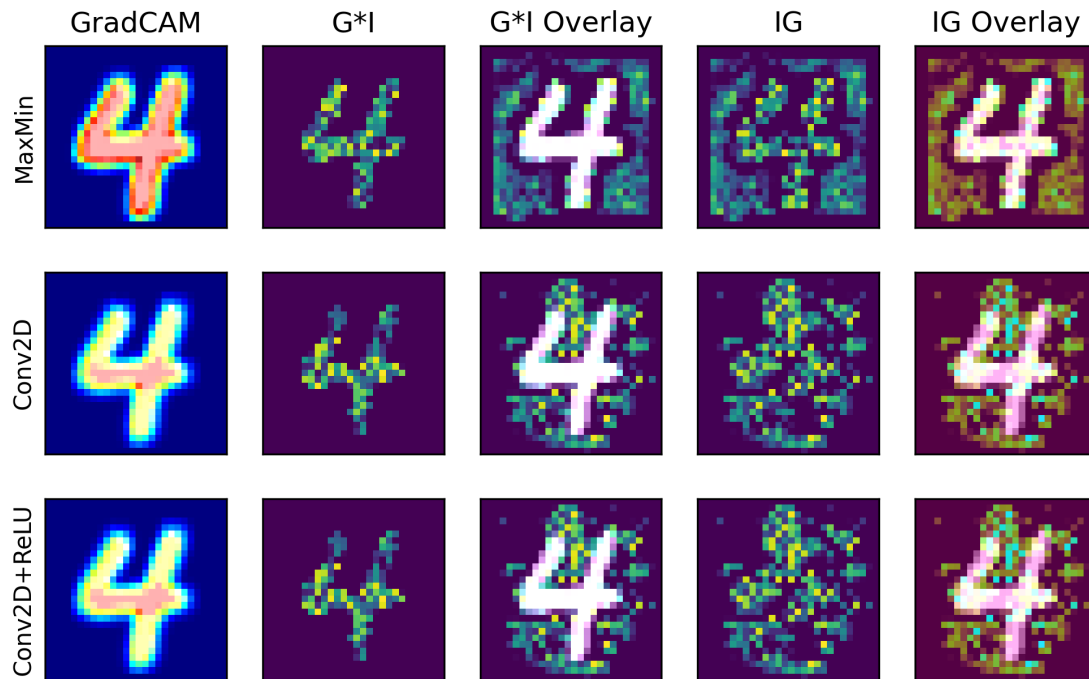


**Figure 2:** Comparison between MaxMin layer and Conv2D network (Conv2D and Conv2D + ReLU activation) on CIFAR10. G\*I is Gradients\*Inputs - IG is Integrated Gradients

## 4. Conclusions

In this work a fuzzy relational neural network based model for extrapolating relevant information from the image data has been introduced. From the preliminary results we observed that the model permits to obtain a clearer indication on the classification processes.

In the next future the authors will focus on further validations of the model from both theoretical and practical point of views.



**Figure 3:** Comparison between MaxMin layer and Conv2D network (Conv2D and Conv2D + ReLU activation) on MNIST. G\*I is Gradients\*Inputs - IG is Integrated Gradients

## References

- [1] A. Knapič, A. Malhi, R. Saluja, K. Främling, explainable artificial intelligence for human decision-support system in medical domain, arXiv (2021).
- [2] J. M. Mendel, P. P. Bonissone, Critical thinking about explainable ai (xai) for rule-based fuzzy systems, *IEEE Transactions on Fuzzy Systems* 14 (2019) 69–81.
- [3] C. Mencar, J. Alonso, Paving the way to explainable artificial intelligence with fuzzy modeling, volume 24, 2018, pp. 215–227.
- [4] F. Camastra, A. Ciaramella, V. Giovannelli, M. Lener, V. Rastelli, A. Staiano, G. Staiano, A. Starace, A fuzzy decision system for genetically modified plant environmental risk assessment using mamdani inference, *Expert Systems with Applications* 42 (2015) 1710–1716.
- [5] A. Ciaramella, R. Tagliaferri, W. Pedrycz, A. Di Nola, Fuzzy relational neural network, *International Journal of Approximate Reasoning* 41 (2006) 146–163.
- [6] A. Ciaramella, R. Tagliaferri, W. Pedrycz, A. Di Nola, Fuzzy relational neural network, *International Journal of Approximate Reasoning* 41 (2006) 146–163.
- [7] Y. LeCun, C. Cortes, MNIST handwritten digit database (2010). URL: <http://yann.lecun.com/exdb/mnist/>.
- [8] A. Krizhevsky, G. Hinton, et al., Learning multiple layers of features from tiny images (2009).
- [9] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-cam: Visual

- explanations from deep networks via gradient-based localization, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 618–626.
- [10] D. Baehrens, T. Schroeter, S. Harmeling, M. Kawanabe, K. Hansen, K.-R. Müller, How to explain individual classification decisions, *The Journal of Machine Learning Research* 11 (2010) 1803–1831.
- [11] Z. Qi, S. Khorram, F. Li, Visualizing deep networks by optimizing with integrated gradients., in: *CVPR Workshops*, volume 2, 2019.