Bidirectional Pooling for Deep Neural Networks

Marilyn Bello[a1,2], Gonzalo Nápoles[3], Koen Vanhoof[2], and Rafael Bello[1]

[1]Department of Computer Science, Central University of Las Villas, Cuba
[2]Faculty of Business Economics, Universiteit Hasselt, Belgium
[3]Department of Cognitive Science & Artificial Intelligence, Tilburg University, The Netherlands

This work introduces a new neural network architecture that uses bidirectional associations-based pooling to extract high-level features and labels from multi-label data. Unlike the pooling approaches reported in the literature, our proposal does not require input data to have any topological properties as typically occurs with images and videos. The numerical results show that our bidirectional pooling helps reduce the number of problem features and labels while preserving the discriminatory power of the network.

Introduction

Pooling layers [1] help reduce redundancy and the number of parameters before building a multilayer (or recurrent) neural network that performs the remaining operations. Although these operators are able to deal with both single-label and multi-label classification problems (MLC) [2, 3], they are specifically aimed at reducing feature space. In the case of multi-label data, this should also be done in the label space. Despite their success, existing pooling operators [4] are focused on data with a well-defined structure (such as image and video) where the term *feature neighborhood* makes sense. However, while it is interesting to recognize faces or classify objects in images and videos, the truth is that there are other domains in which the data do not have a topological organization. In those cases, using standard pooling operators might have little sense, even when the problem at hand could benefit significantly from a deep learning solution.

The proposed network architecture

In [5], we proposed a bidirectional network composed of stacked association-based pooling layers to extract high-level features and labels in MLC problems with no specific topological organization. Unlike the classic use of pooling, this approach does not perform pooling over pixels but problem features or labels. The first pooling layer is composed of neurons denoting the problem features and labels (i.e., low-level features and labels). In contrast, neurons denote high-level features and labels extracted during the construction process in deeper pooling layers. Each pooling layer uses a function that detects pairs of highly associated neurons while performing an aggregation operation to derive the pooled neurons. Such neurons are obtained from neurons belonging to the previous layer to fulfill a certain association threshold. This model uses Pearson's correlation to estimate the association degree between two neurons. Overall, we compute the correlation matrix among features and labels and derive the degree of association of the pooled neurons from the degree of association between each pair of neurons in the previous layer. The pooling process is repeated over aggregated features and labels until a maximum number of pooling layers is reached.

Once the high-level features and labels have been extracted using the pooling operators, they are connected with one or several hidden processing layers. Finally, a decoding process [6] is performed to connect the high-level labels to the original ones by means of one or more hidden processing layers. Figure 1 depicts an example of this network architecture resulting in five high-level neurons that emerge from the association-based pooling layers. Hidden neurons in these hidden layers are equipped with can use any transfer function such as ReLU, sigmoid or hyperbolic tangent.

Numerical simulations

The performance of our model is evaluated using several MLC problems. Overall, we study how the model performs in terms of accuracy and number of features and high-level labels. Table 1 reports the number of high-level features (#HLF), feature reduction percentage (%Red-F), the number of high-level labels (#HLL), label reduction percentage (%Red-L), the accuracy obtained by the network using the extracted features and labels, the accuracy using the original features and labels (baseline model), and the loss of accuracy with respect to the baseline model.

Table 1: Performance assessment of the bidirectional association-based pooling approach.

Dataset	#HLF	%Red-F	#HLL	%Red-L	Accuracy	Baseline	Loss
D1	43	40.28%	6	0%	0.815	0.823	-0.008
D2	24	91.84%	6	0%	0.913	0.915	-0.002
D3	44	57.28%	13	7.14%	0.798	0.80	-0.002
D4	17	96.85%	22	87.43%	0.988	0.987	0.001
D5	19	96.75%	29	87.22%	0.99	0.99	0
D6	19	96.71%	50	87.5%	0.995	0.995	0
D7	20	96.85%	35	87.23%	0.991	0.991	0
D8	14	96.82%	8	0%	0.915	0.918	-0.003
D9	53	87.95%	4	0%	0.837	0.866	-0.029
D10	49	88.86%	5	16.67%	0.805	0.794	0.011
D11	4	96.67%	7	93.07%	0.965	0.965	0
D12	78	96.37%	10	95.19%	0.99	0.99	0
D13	80	92.01%	14	50%	0.928	0.988	-0.06
D14	18	96.4%	3	96.3%	0.977	0.977	0
D15	9	92.97%	3	96.3%	0.977	0.977	0

From these results, we can observe that our proposal significantly reduces the number of features and labels

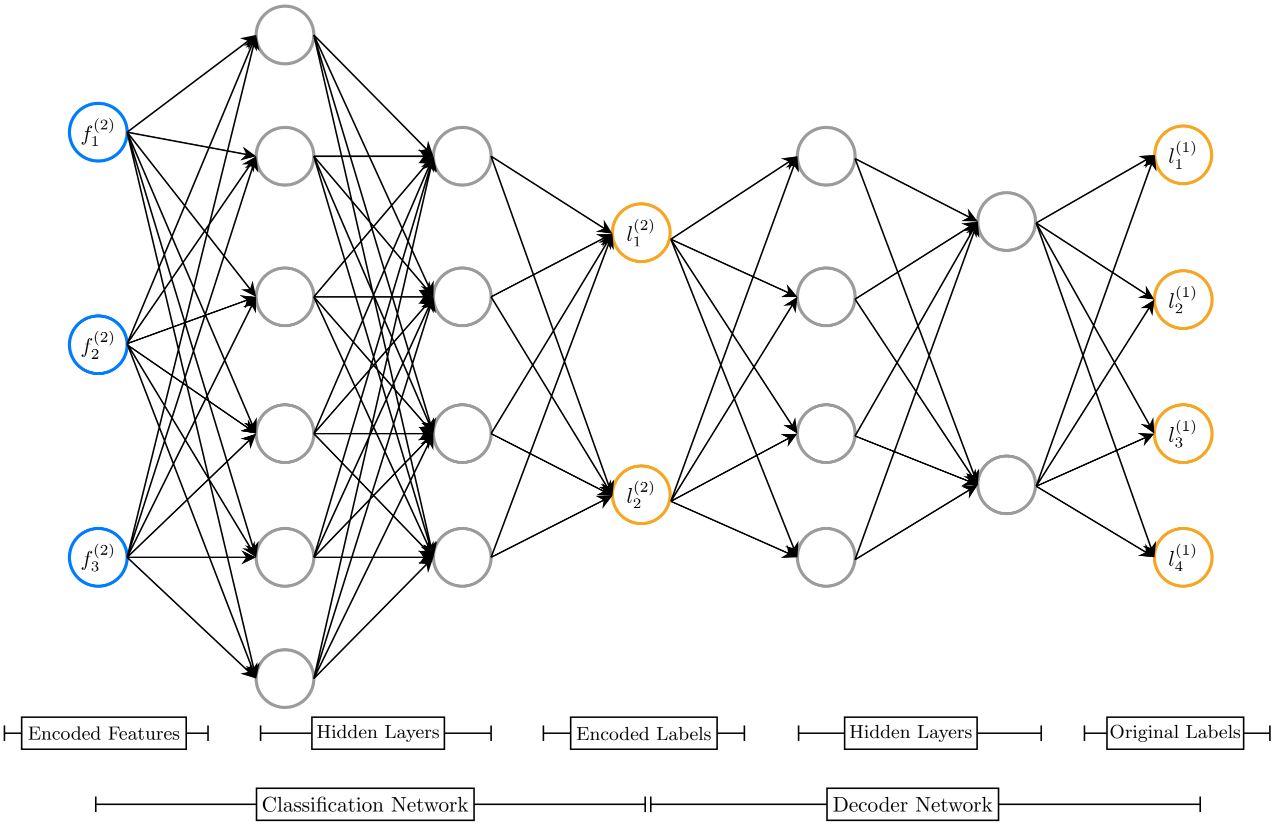

Figure 1: Neural network architecture involving three high-level features (resulting from the feature pooling step), two-high-level labels (resulting from the label feature step), four low-level labels and four hidden layers.

with a percentage reduction up to 96% and 87%, respectively. It is worth mentioning that the bidirectional association-based pooling reports a maximal accuracy loss of 0.06 for the $D13$ dataset. However, in some cases, we observed a small increase in the accuracy (e.g., dataset $D10$) even when our network was not conceived to increase the prediction rates but to obtain the same performance with smaller networks. Our proposal has no loss in accuracy for those problems having low variability in accuracy (i.e., datasets $D5 - D7$, $D11 - D12$, $D14 - D15$).

Conclusions

The numerical simulations have shown that our proposal is able to significantly reduce the number of parameters in deep feed-forward neural networks without harming their discriminatory power. Extracting high-level features and labels increases the possibility of building networks with more transparent inference models. For example, by using *post-hoc* interpretability techniques, we could shed light on the inner reasoning of the model when operating with high-level features. These techniques regularly have exponential algorithmic complexity, thus having networks with fewer parameters certainly helps reach this goal.

Notes

a. Email: mbgarcia@uclv.cu

References

[1] Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y., *Deep learning,* MIT press Cambridge (2016)

[2] Charte, F., del Jesus, M.J., Rivera, A.J., *Multilabel classification: problem analysis, metrics and techniques,* Springer (2016)

[3] Gibaja, E., Ventura, S., Multi-label learning: a review of the state of the art and ongoing research, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* **4** (2014) 411–444

[4] Yu, D., Wang, H., Chen, P. and Wei, Z., Mixed pooling for convolutional neural networks, *Lecture Notes in Computer Science,* **8818** (2014) 364–375

[5] Bello, M., Nápoles, G., Sánchez, R., Bello, R., Vanhoof, K., Deep neural network to extract high-level features and labels in multi-label classification problems, *Neurocomputing,* **413** (2020) 259–270

[6] Hinton, G.E., Salakhutdinov, R.R., Reducing the dimensionality of data with neural networks, *Science,* **313** (2006) 504–507