Doctoral Dissertation
Doctoral Program in Computer Engineering (36^{th}cycle)

Robust machine learning models for high dimensional data interpretation

By

Manigrasso Francesco

******

Supervisor:
Prof. Fabrizio Lamberti

Refrees:
Prof. Annette ten Teije, Vrije Universitat (Referee)
Prof. Luciano Serafini, Fondazione Bruno Kessler (Referee)
Prof. Francesco Setti, Università di Verona
Prof. Luca Bortolussi, Università di Trieste
Prof. Paolo Garza, Politecnico di Torino

Politecnico di Torino
2024

Summary

The success of deep neural networks (DNNs) in analyzing multidimensional data, such as images, stems from their ability to model both linear and non-linear features, as well as the availability of large datasets. Despite these strengths, DNNs often lack key attributes such as robustness, control, and transparency, which are essential for real-world applications. To address these limitations, researchers are increasingly integrating symbolic knowledge representation and statistical reasoning to improve interpretability and reasoning capabilities.

The central aim of this Ph.D. research is to bridge the gap between symbolic knowledge representation and deep learning, focusing on neural-symbolic (NeSy) integration architectures. This approach enhances DNNs by incorporating prior knowledge into the learning process, allowing methods like fuzzy logic to be embedded as tensors within neural networks.

This dissertation contributes to the development of architectures designed to solve image-based tasks such as object detection and classification, embedding logical reasoning within image feature representations. The practical implementations explored introduce various inductive biases to align learning processes, making them applicable to real-world scenarios through diverse datasets.

Recent advances in NeSy architectures have garnered interest in the computer vision community due to their ability to combine neural networks and symbolic reasoning. These hybrid systems improve interpretability, reasoning, robustness, and reduce the reliance on labeled data by incorporating external knowledge bases. This study specifically explores the integration of Logic Tensor Networks (LTNs) into architectures for complex computer vision tasks. The work is divided into two main parts: object detection and zero-shot learning (ZSL).

The first research area focuses on improving object detection by integrating prior knowledge into deep learning models through the *Faster-LTN: a neuro-symbolic, end-to-end object detection architecture*. A pre-trained Convolutional Neural Network (CNN) is used to label images with bounding boxes that identify object locations and classes. Previous studies have shown that logic networks trained with tensor data from CNNs can enhance overall performance. This research presents a symbolic network that facilitates semantic

knowledge exchange through logical axioms, using distributed data from CNNs. It emphasizes end-to-end training, highlighting both computational challenges and the performance benefits of neuro-symbolic architectures, particularly under limited data conditions.

The second research field addresses Zero-Shot Learning (ZSL) by developing progressively complex architectures that integrate symbolic reasoning with deep learning models. The goal is to establish relationships between images and classes using semantic descriptions of the available classes, enabling the system to recognize previously unseen classes during inference. The first architecture, *PROTOtypical Logic Tensor Networks (Proto-LTN)*, combines NeSy frameworks with prototypical networks to ground abstract class concepts in a continuous embedding space, where prototypes are matched using Euclidean distance. This approach improves performance in scenarios where labeled data is limited. Following this, the *Fuzzy Logic Visual Network (FLVN)* introduces an end-to-end model capable of transferring image features into attribute spaces, incorporating new axioms to better manage unseen classes and expand prior knowledge during training. Finally, the *Fuzzy Logic Prototypical Network (FLPN)* enhances the previous models by employing an attention mechanism to extract attribute features. This architecture refines prototype matching through CNN- and transformer-based backbones, improving generalization to unseen classes and surpassing the performance of earlier approaches.

In conclusion, the dissertation presents the outcomes of integrating symbolic reasoning with deep learning to tackle complex computer vision tasks. The research highlights the advantages of using NeSy architectures to enhance transparency, reasoning, and robustness in image analysis. Future work will explore additional ways to refine these hybrid systems, pushing the boundaries of interpretability and learning efficiency in AI.