Our research will make use of publicly available VQA datasets such as GQA.

A potential solution is to drop this assumption and mine neuron activation patterns that identify the purpose of the model they try to explain, i.e., they are model-agnostic. In the context of VQA systems, a start in the definition of an interpretable space for explanations of VQA systems.

Other recent approaches rely on deep learning techniques. Being a problem on multimodal data, this implies to merge both images and questions into a common representation space. This is challenging because images and texts are very different data types, treated by means of different neural network architectures: CNNs (Convolutional Neural Networks) are the state of the art for image classification and representation, whereas text processing often resorts to RNNs (Recurrent Neural Networks). VQA solutions must orchestrate both technologies leading to systems that are extremely complex.

The complexity of existing VQA solutions makes the task of instruction and debugging very hard. In particular, neural networks are black-box models: one requires a significant amount of work and solid expertise to understand the inner-workings of the network. It becomes therefore very difficult to understand why a VQA system makes a mistake. Such a task, however, is vital for the progress of research in VQA.

Assignment

The main purpose of this post-doctoral fellowship is to port the principles of interpretable AI and ML to the domain of the visual question answering. Attaining such a goal requires us to overcome other challenges such as understanding what makes a good explanation in a multi-modal setting.

Main activities

The post-doctoral researcher in charge of this project will work on methods to explain the output of a VQA system in order to understand why it erred (or not).

In this regard, we aim at deploying post-hoc interpretability modules that can answer questions in natural language from the contexts of an image. VQA carries potential applications in multimodal information retrieval.

Current VQA solutions rely on deep learning techniques. Being a problem on multimodal data, this implies to merge both images and questions into a common representation space. This is challenging because images and texts are very different data types, treated by means of different neural network architectures: CNNs (Convolutional Neural Networks) are the state of the art for image classification and representation, whereas text processing often resorts to RNNs (Recurrent Neural Networks). VQA solutions must orchestrate both technologies leading to systems that are extremely complex.

The complexity of existing VQA solutions makes the task of instruction and debugging very hard. In particular, neural networks are black-box models: one requires a significant amount of work and solid expertise to understand the inner-workings of the network. It becomes therefore very difficult to understand why a VQA system makes a mistake. Such a task, however, is vital for the progress of research in VQA.

Our research will make use of publicly available VQA datasets such as GQA, VQA-E, C-VQA, and CLEVR.
Skills

We are searching for motivated candidates with a PhD degree in Computer Science and with competences in machine learning (preferably with focus on deep learning). Knowledge in data mining, e.g., sequence and itemset mining, will be also appreciated.

The candidate should be proficient in written and spoken English (at least B2 level according to the CEFR system).

Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Remuneration

Gross monthly salary (before taxes): 2653 €