Observations on semantic annotation of micrscope images for life sciences

Slob, J.; Kallergi, A.; Verbeek, F.J.; Marshall, M.; Scott; Burger, A.; ... ; Splendiani, A.

Observations on Semantic Annotation of Microscope Images for Life Sciences

Joris Slob, Amalia Kallergi, Fons J. Verbeek

Leiden Institute of Advanced Computer Science, Leiden 2333 CA, The Netherlands

keywords: ontology, Life Sciences, annotation, images, microscopy

1 Introduction

Microscopy images are important in the Life Sciences. These images not only vary in scale but also in modality, making it hard to do image registration. Ontology annotation helps us to link the images, but not all researchers in the Life Sciences are familiar with ontologies. The Cyttron Scientific Image Database for Exchange (CSIDx) presents different tools to help users annotate their images.

When taking microscopy techniques and methods into account, one cannot correlate images from the Life Sciences domain on a pixel by pixel basis. Even images with the same subject matter change so drastically with modality that it becomes impossible for classical image retrieval techniques to deal with this particular domain. The resulting images range from one-dimensional to more than four-dimensional artifacts. Without any added information it is impossible to discern the nature of these images. Semantic relations are needed to make the connections between the images. An important part of the semantics is not only the subject matter but also the modality of the image.

A good collection of ontologies can be found in the OBO Foundry. This consortium curates ontologies geared towards the Life Sciences and also provides mappings between different ontologies. Concepts from ontologies can be used to enrich images by associating each image with multiple ontology concepts. These concepts can come from many multiple ontologies. This allows maximum freedom for the user to annotate their images, while restricting them to a controlled vocabulary.

2 Implementation

To enable researchers in the Life Sciences to cooperate we present an annotated image database with a web front-end. This service, the Cyttron Scientific Image Database for Exchange (CSIDx), hopes to combine the general accessibility of a web based front-end with detailed knowledge enrichment.

Users of CSIDx can upload their images in any format to our service. We impose an annotation step for all submitted images, for reasons stated above. To get all relevant information attributed to images, we make use of an annotation process that consists of several steps. Because annotation takes a lot of time for the user, we try to automate this annotation process as much as possible. Sets

of annotations can be stored as templates that can be reused during annotation of other images.

In the CSIDx system we created templates for different types of microscopes. The user can choose an appropriate microscope template, for instance a generic electron microscope, and fill in the appropriate variables for this particular microscope. In this way we can recover information on how the image was generated. Staining and other preparation variables are considered part of the microscope setup.

It is a waste of valuable time if researchers have to search through all available ontologies every time they want to annotate an image. This is why we introduced the "My Terms" section of our website, where users collect the terms from semantic context drawn from OBO. In the annotation process the user can use these concepts to annotate their images without all the irrelevant terms. As another way to speed up the annotation process we have constructed a "meta-ontology", which is a collection of concepts from multiple ontologies that deal with a single topic. These concepts and their direct neighbours were extracted and put into a new knowledge structure which we call a meta-ontology.

Most researchers are not familiar with ontologies. To help them get a feeling of the structure of ontologies we added the ontology viewer in our interface. This helps the users understand the semantic context, by helping them visualize the relations of the concepts in ontologies.

Researchers want to capture relations that are not covered in ontologies. For this use-case we allow researchers to group images in datasets. We also allow users to link images to other databases, including PubMed. This allows scientists to link their images to their publications and in this way extra information can be looked up by following the link. In this way users can bring structure and meaning to their data in our system.

We wanted to provide an easy way to make use of the semantic content on a query result set. As a first view we implemented a tag-cloud based on ontology terms. In this way it becomes easy to explore image semantics or find more precise images for your query.

3 Conclusion

The CSIDx service strives to be a center for research image sharing with support for semantic annotation. Papers only present results in abstracted graphs and example images. CSIDx can make large volumes of images retrievable. This can promote better overview of the available imagery made by fellow researchers and hopefully stimulate cooperation.

4 Acknowledgements

This research was funded by the Cyttron consortium.