

Internship opportunities for CS and AI students at IBM, in the first half of 2016

All 6 months and full time, either as thesis project or as internship with no particular academic deliverables (e.g. work experience). Start dates are to be agreed upon to suite everyone involved. Additional projects are also possible, for more information contact Zoltan Szlavik at zoltan.szlavik@nl.ibm.com.

Image classification (2 separate projects)

Theme: high accuracy image classification of a) medical, or b) macro-lens captured images. Suitable algorithms need to be developed for both domains, which will most likely be some versions of deep learning.

Various challenges such as image resolution will need to be addressed, and the classification is not only about finding a label for images, but also to identify focus areas within the image, and accurately labelling those.

Specific research questions can be identified based on discussions and mutual interest.

Both projects are connected to others focusing on gathering appropriate training data.

Image- and training-label gathering via crowdsourcing

Theme: in order to get sufficiently many and labelled images for automatic image classification (classification algorithm to be covered by another student project), we need to devise efficient and effective ways to get both the images and corresponding labels. This can be done via crowdsourcing, but the best way to do so it to be found out via this research.

Challenges: a) how do we take images of a certain subject (think images taken with a macro lens, exact industry is confidential but will be covered during discussions) so we can create many images for training, b) how do we make people annotate/label these images – both at image level and sub-image level (perhaps via a tailor-made way of gamification?)

Specific research questions can be identified based on discussions and mutual interest.

Labelled images are inputs to another student project focusing on machine learning on these images.

Bringing crowdsourcing into developer cloud (2 complementary projects)

Theme: this is an implementation-focused project (actually, two) that aims to take an existing crowdsourcing framework (CrowdTruth) and place it into the cloud (IBM Developer Cloud) as a service accessible via its API. Potential for running specific experiments (e.g. completing certain tasks using the crowd) once implementation is complete.

Challenges: How can we do this both in an effective and efficient way? How to use the latest algorithms and methods to achieve the goal?

Specific research questions (unless it's not a thesis project) are to be identified while studying both the source and target environments.