Machine learning to automate image analytics in image guided procedures Jan2020

Background and problem description

SINTEF is a broad, multidisciplinary research organisation with international top-level expertise in the fields of technology, the natural sciences, medicine and the social sciences. We conduct contract R&D as a partner for the private and public sectors, and we are one of the largest contract research institutions in Europe.

The Medical Technology research group at SINTEF focuses on the use of ultrasound imaging, image processing and navigation technology for guiding of medical procedures. Both diagnostics and therapy.

1. Machine learning to automate workflow of lung cancer diagnostics

In the field of lung medicine, more specifically lunger cancer diagnostics and therapy (surgery), we have several projects running. The main goal is higher success rate in bronchoscopic lung lesion diagnostics. Some of the lesions (typically found in CT scans) are located peripherally in the lungs and challenging to reach with current bronchoscopy techniques. Better preoperative planning and intraoperative guidance will improve the situation, and lead to correct treatment at an earlier stage for the patients.

The current activities, relevant for several master students, revolves around trying to automate preoperative and intraoperative image processing;

i) Segmentation of structures like lung, lesion/tumour, vessels, airways, particularly from CT images (usually stacks of several hundred images)
ii) Co-registration (matching) between different imaging modalities like CT, MR, PET, ultrasound
iii) Use of the live video stream from the bronchoscopy camera to estimate and track the position of the bronchoscope tip in relation to the patient's CT images (video—to-CT registration for position tracking)

These tasks, and others in the projects, are well suited for machine learning.

A software framework and a lab setup has already been established at the research group. The open source software platforms CustusX, www.custusx.org, and FAST, https://www.eriksmistad.no/fast/, are used in the project.

Data

We have already acquired datasets, including patient data. In addition, for machine learning purposes, we use open source data sets. These two sources of data can be found here:

i) One of our own data set published at Dryad, from segmentation study in Trondheim, bronchoscopy patients: https://datadryad.org/resource/doi:10.5061/dryad.mj76c
ii) Open access data set at LIDC-IDRI: https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI
    a. Reference: Armato et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules

on CT scans. Medical Physics, 38(2):915-931. DOI: 10.1118/1.3528204, url: https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.3528204

iii) Open access data set, approx. 15.000 CT scans, and ~450 with digital pathology image:
   a. https://cdas.cancer.gov/nlst/ More information about this data set can be given on request. We have these data physically available at the research group

iv) Other open available data sets will be considered continuously by the project team, including acquiring larger data sets retrospectively from patient population in mid-Norway.

Using data from patients at St. Olavs, enrolled in approved studies, are usually de-identified, and a subset is annotated to outline the extent of the structures of clinical interest.

Challenges

The aims and challenges of the proposed master's thesis (2-3 of them, possibly in collaboration at the same time) are:

i) Use machine learning / neural networks to automatically identify the structures of clinical relevance/interest for automating planning and guidance of bronchoscopy and lung tumour surgery. Use vessels and airways to create optimal path to target structures.

ii) Use neural networks to automatically classify tumours from CT and PET images into malignancy grade. It may also be of interest to include other sources of information like histopathology results from tumours (digital pathology).

iii) Video to CT registration, which may be challenging due to varying quality of live video stream over time because of mucus and blood, in addition to movement and patient motion (breathing, heartbeat).

Thesis possibility

The thesis projects have a general time frame of 3-12 months and can also be adapted to be started as a Project thesis. In addition, it might be possible with a summer job with the research group related to the same projects. Contact person and supervisor at SINTEF is Chief Scientist Thomas Langø (thomas.lango@sintef.no).