UNIVERSITY OF COPENHAGEN
FACULTY OF HEALTH AND MEDICAL SCIENCES

Investigation of Coronary Angiographies of patients with Ischemic Heart Disease using deep learning

Christian Eschen[1], Karina Banasik[1], Anders Bjorholm Dahl[2] and Søren Brunak[1]
[1]Novo Nordisk Foundation Center For Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen
[2]Department of Applied Mathematics and Computer Science, Technical University of Denmark

Contact Information:
Christian Eschen
+45 20652915
christian.eschen@cpr.ku.dk

Background

Invasive coronary angiography (ICA) is considered the gold standard for diagnosis of ischemic heart disease (IHD). In ICA, a catheter is used to inject an iodine-based contrast agent intravenously into the coronary circulation system. While injecting the contrast agent, a video of real-time X-ray images are acquired using fluoroscopy. By inspecting this video, the cardiologist can detect blockages in the coronary arteries known as stenosis and it is possible to perform surgical intervention, which includes insertion of a stent or a bypass surgery.

Identification and quantification of stenosis and lesions in ICA is important for treatment, prognosis, and diagnosis of IHD. However, manually assessment of stenosis is not objective as the degree of percentage stenosis is often measured by "eyeballing", and it requires special expertise.

ICA videos are acquired as multiple video sequences with varying X-ray detector and source angles examining both left coronary artery (LCA) and right coronary artery (RCA). The videos are acquired in a non-standardized protocol, and the data requires manual annotations before it is suitable for further analysis.

This project aims to use deep learning to classify coronary angiographies for further analysis. The intended workflow is illustrated in Figure 1 below.

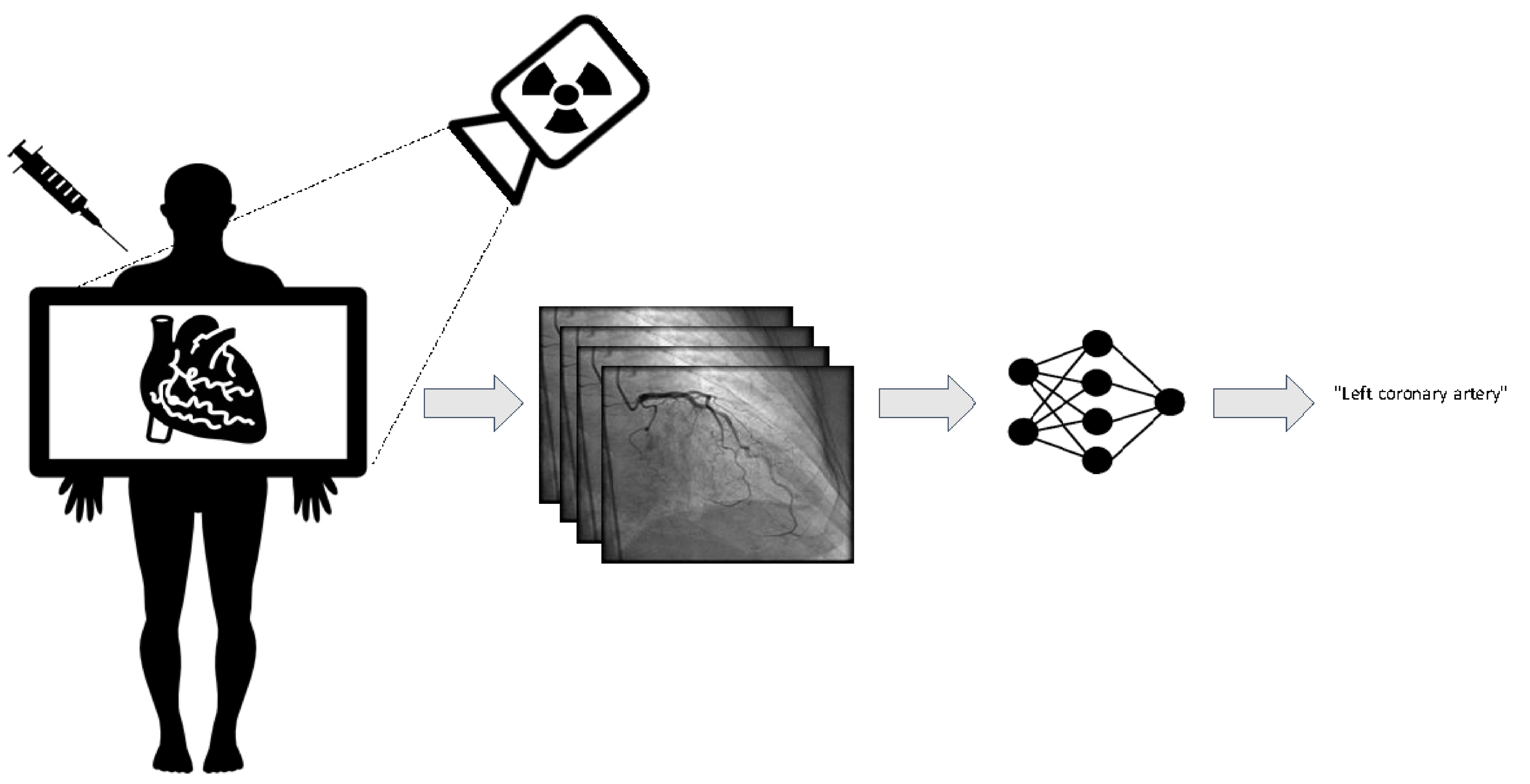

Figure 1: Illusration of workflow

Data

The data foundation available for this project consists of 2.7 million electronic health records (EHR). Within this cohort, a sub-cohort with approximately 36,000 patients underwent ICA is linked to the so-called PATS database. In the PATS database, the patient history data, the outcome of other kinds of examinations and the reported outputs from ICA are listed. Each examination of a patient consists of multiple videos from different anatomically positions. For approximately 26,000 of the patients who underwent ICA, we have longitudinal ICA information, meaning that these patients had more than one ICA examination. In Table 1 summary statistics for the raw ICA videos are presented.

	Examinations	Videos	Patients	Examinations pr patient	Videos pr patient
ICA raw	36 376	536 293	27 960	1.3	19.18

Table 1: Overview of data

Method

Data annotation

As each examination consists of multiple videos with labels only related to the examination and with unknown content on the individual video, we annotated 6625 videos. Using the annotations the aim is to build a model capable of classifying the coronary angiographies.
The videos are annotated by assigning a categorical value representing the overall semantic content to each video. The annotation scheme is presented in Table 2.

	Categories	Videos	Percentage
Left coronary artery	1	1397	0.35
Right coronary Artery	2	722	0.18
Left coronary artery + wire	3	687	0.17
"Bad data"	4	399	0.10
Right coronary artery + wire	5	358	0.09
Left coronary artery without contrast + wire	6	250	0.06
Right coronary artery without contrast + wire	7	214	0.05
Total		6625	1.00

Table 2: Annotation scheme

Supervised deep learning

We will utilize an 18-layer Resnet 3D model [1] as baseline for training a supervised ICA video classification model. The videos are resized to 224x224 in the spatial dimensions and random cropped to size 224x224x32 in the temporal dimension.
The annotated data above will be separated into a training set and a validation set with equal proportion of healthy patients and patients with IHD (see Table 3). As we do not know in advance the needed amount of labeled training data, we will iteratively increase the labeled training dataset size.
We expect to take data centric approach in which we only make minimal changes to model hyper parameters and instead change the training data.

	Examinations	Videos	Patients
Training	315	4476	251
Validation	160	2149	138

Table 3: Expected training and validation split.

Unsupervised deep learning

Since we have a lot of unannotated (about 500,000 videos), we plan to use unsupervised representation learning. Recently in 2020, there has been a breakthrough in un/self-supervised learning (e.g., [2] and [3]). It has been shown that it is possible to obtain comparable results on un/self-supervised learning with supervised learning on video classification tasks. We plan to reuse the preprocessing steps and the Resnet 3D model described in previous sections as backbone, and we plan to adapt the BYOL [4] or SIAM [5] framework as baseline.

Future work

- Can we predict percentage stenosis and anatomical position?
- Can we predict if it is a 1-vessel, 2-vessel or 3-vessel disease?
- Can we predict the SYNTAX score or similar grading score, which determines the complexity of the lesions?

References

[1] Tran, Du, et al. "A closer look at spatiotemporal convolutions for action recognition." *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 2018.
[2] Feichtenhofer, Christoph, et al. "A Large-Scale Study on Unsupervised Spatiotemporal Representation Learning." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.
[3] Qian, Rui, et al. "Spatiotemporal contrastive video representation learning." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.
[4] Grill, Jean-Bastien, et al. "Bootstrap your own latent: A new approach to self-supervised learning." *arXiv preprint arXiv:2006.07733* (2020).
[5] Chen, Xinlei, and Kaiming He. "Exploring simple siamese representation learning." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.