Bern, 13.04.2021

u^b

Bern Data Science Day 2021 - April 23

Contribution - Region-based VQA in the medical domain

Sergio Tascon-Morales, Pablo Márquez Neila, Raphael Sznitman

AIMI Lab, ARTORG Center

{s.tasconmorales, pablo.marquez, raphael.sznitman} @artorg.unibe.ch

Abstract

Visual Question Answering (VQA) systems are models that, given an image and a textual question about it as input, can provide a textual answer as output. Current VQA systems in the medical domain have been following the progress of the general domain VQA systems, except for some aspects that are difficult to adapt to medical images such as object detection. Essentially, the nature of medical images differs notably from general images and, correspondingly, the types of questions that can be asked about a medical image are also different.

This work in progress addresses the question of how can VQA systems and corresponding datasets be built, that deal with questions about specific regions of an image. This includes questions about the contents of a single region (e.g. about the presence or absence of abnormalities within a region) as well as questions about the relationship between two different regions of an image (e.g. establishing similarities or differences between the contents of two regions). This task requires, besides the different levels of reasoning typically involved in VQA, the ability to integrate a description of the location of the region(s) into the model. Building a dataset for this task is particularly complex due to the limited amount of available annotated data. Furthermore, asking questions about regions introduces new parameters that affect the performance of the system, such as the amount of generated regions for each image, and their size.