A SMOKING RELATED CANCERS [LUNG & LIVER] SEGMENTATION USING UNET

Department of Computer Science
Mansoura University, FCIS
Mansoura

Proposal Report

Feb 25, 2023

Abstract: This project is going to handle cancers and tumors which have a great relation to smoking, of course lung cancer will be the forefront cancer to be studied and modeled as it's directly related to smoking as nearly 9 out of 10 deaths of the lung cancer are mainly caused by smoking or secondhand smoke exposure, we will study one another cancer related to smoking which is liver cancer as smoking affects the risk of liver cancer and it becomes high likelihood to be happened. Segmentation is considered to be used to localize and annotate every single defected pixel, all of that will be modeled using U-Net architecture which is used for semantic segmentation to ease the process of analyzing and reading the input image. Replacing doctors is not our objective, we are targeting a huge portion of people in the world which are smokers, and they are not just millions at the meantime but billions (**1.1 billion**). Being diagnosed by one the mentioned cancers or even both is high likely to happen, therefore our application here is going to serve them, they are going to answer some related questions about different symptoms to indicate whether he's likely to have that cancer or not, then make a specific kind medical imaging like MRI and process it in the segmentation model to indicate it's actually a tumor or not and localize it if exists. Mobile Application will be developed to send inputs in the form of either plain text or images and receive outputs in the same format as well, that generic mobile application will be handled using Flutter framework technology with quite good and simple UI to entice potential patients to use it.

Keywords: Machine Learning, Computer Vision, Sematic Image Segmentation, CNN, Neural Networks, U-Net Architecture, Tensor Flow, Keras, Flutter, dart, Python, MRI, .NET.

Team Members	
Name	**Track**
AbdelRahman Ibrahim Ibrahim Karawia	**Artificial Intelligence & Flutter**
Eman Samy Mahmoud Mohamed Ramadan	**Artificial Intelligence**
Samir Magdy Isaac Yakoub	**Artificial Intelligence & Backend**
Haidy Essam Abdel-Kader	**Artificial Intelligence**

1. MOTIVATION

Cancer is a **leading** cause of death worldwide, accounting for nearly 10 million deaths in 2020, or nearly one in six deaths, but the patient can handle that cancer or at least limit its symptoms with early detection which is the main key here. The number of cases for every cancer or tumor changes significantly but its not all about the number of each case yearly, it's about the death rate for each one as well, we are going to introduce the most dangerous cancers and tumors.

Table 1. Cancer New Cases and Deaths

New Cases		Cancer Deaths	
Breast	2.26 million	**Lung**	**1.80 million**
Lung	**2.21 million**	Colon and Rectum	916,000 deaths
Colon and Rectum	1.93 million	**Liver**	**830,000 deaths**
Skin	1.2 million	**Breast**	685,000 deaths

Lung Cancer is considered as the leading cause of cancer deaths worldwide with 1.8 million every year as an average. Smoking is the main role here, if we are going speak about lung cancer, smoking is the first cause popped to our mind as smokers have the greatest likelihood risk of lung cancer and this risk increases with different factors like time span of smoking and number of cigarettes smoked.

Liver Cancer, we can imagine how critical it is, even it's not the most cancer with deaths, but according to the number of new cases which is not that high like breast cancer of lung cancer however, liver cancer has that high number of deaths every year **Table 1**, and here's the point, we should not consider only the cause of cancer but also the after getting that cancer what makes it deadly, or what's increasing the probability of ending up dead, cigarette smoking has been linked to a higher risk of liver cancer depends on the number of cigarettes smoked per day and timestamp of being a smoker.

Therefore, it's important to develop a full system, begins with questioning the potential patient some related questions and according to that symptoms, we move to the next step according to the results of the first step, which is receive specific type of medical imaging based on the case and process it through our model to end up with an image full of information and translate all of that to be just a text for the normal person to understand. It could be of great help to automate the detection and locate tumor tissue cells and to speed up the process. This way one would be able to overcome the dependence on the pathologist which would be especially useful in regions where no experts are available but that's not the main aim of the project which will be discussed latter.

2. METHODS

2.1 Dataset

We are going to take advantage of a well-known challenge first introduced in 2018 and there was a lot of competitors which is "**Medical Segmentation Decathlon**" and its aim to utilize semantic segmentation algorithms but in 3-D to make a model handle many tasks for different organs of the human being as Medical imaging is missing a fully open source and comprehensive benchmark. The MSD data set is publicly available under a Creative Commons license CC-BY-SA4.0, allowing broad (including commercial) use. The training data used in this study is available at https://medicaldecathlon.com.

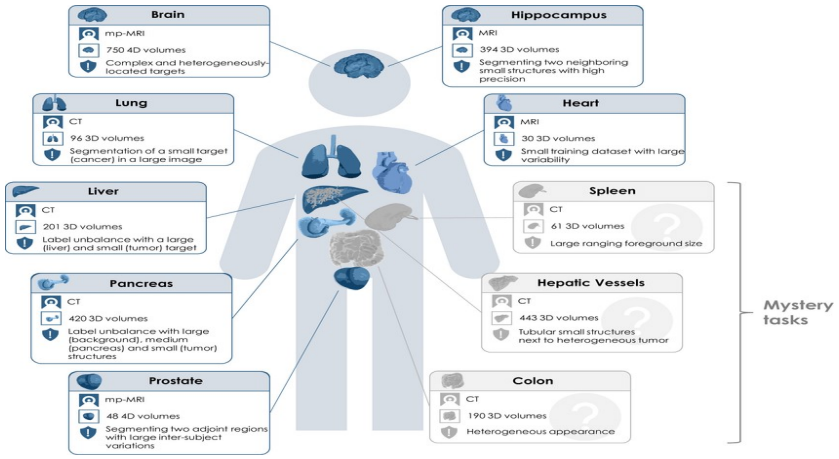

Figure 1. Medical Segmentation Decathlon Tasks

2.2 Models

Having such a model designed for biomedical applications is a great help for our problem and it is a convolutional neural network **U-Net**, which mainly used for high complex computer vision tasks and it's designed to solve the task of **image segmentation**.

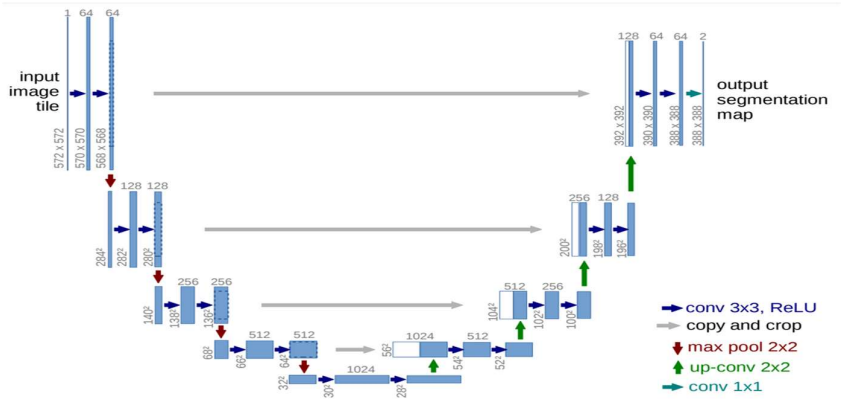

Figure 2. U-Net Architecture

2.3 Flutter Framework

Flutter is a cross-platform framework that targets developing high-performance mobile applications. Flutter was publicly released in 2016 by Google. Besides running on Android and iOS flutter applications also run on **Fuschia.** Flutter is chosen as Google's application-level framework for its next-generation operating system. Flutter is exceptional because it is dependent on the device's **OEM** widgets rather than consuming web views. Flutter uses a high-performance rendering engine to render each view component using its own. This provides a chance to build applications that are as high-performance as native applications can be. In view of architecture, the engine's C or C++ code involves compilation with **Android's NDK** and **LLVM** for iOS respectively.

3. Related Papers

1- Antonelli, M., Reinke, A., Bakas, S. et al. **The Medical Segmentation Decathlon**. Nat Commun 13, 4128 (2022). https://doi.org/10.1038/s41467-022-30695-9

"International challenges have become the de facto standard for comparative assessment of image analysis algorithms. Although segmentation is the most widely investigated medical image processing task, the various challenges have been organized to focus only on specific clinical tasks. We organized the Medical Segmentation Decathlon (MSD)—a biomedical image analysis challenge."

2- Primakov, S.P., Ibrahim, A., van Timmeren, J.E. et al. **Automated detection and segmentation of non-small cell lung cancer computed tomography images**. Nat Commun 13, 3423 (2022). https://doi.org/10.1038/s41467-022-30841-3

"Detection and segmentation of abnormalities on medical images is highly important for patient management including diagnosis, radiotherapy, response evaluation, as well as for quantitative image research. We present a fully automated pipeline for the detection and volumetric segmentation of non-small cell lung cancer (NSCLC) developed and validated on 1328 thoracic CT scans from 8 institutions. Along with quantitative performance detailed by image slice thickness, tumor size, image interpretation difficulty, and tumor location."

3- Shimazaki, A., Ueda, D., Choppin, A. et al. De**ep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method**. Sci Rep 12, 727 (2022). https://doi.org/10.1038/s41598-021-04667-w

"There are two main methods for detecting lesions using DL: detection and segmentation. The detection method is a region-level classification, whereas the segmentation method is a pixel-level classification. The segmentation method can provide more detailed information than the detection method. In clinical practice, classifying the size of a lesion at the pixel-level increases the likelihood of making a correct diagnosis. Pixel-level classification also makes it easier to follow up on changes in lesion size and shape, since the shape can be used as a reference during detection. It also makes it possible to consider not only the long and short diameters but also the area of the lesion when determining the effect of treatment."