Object detection in Lung Cancer CT scans using deep learning and transfer learning

Dmytro Pravdyvets

Health Sciences

Sergi Martínez Maldonado

Ferran Prados Carrasco

25/06/2023

Universitat Oberta de Catalunya

C) Copyright

SUMMARY OF THE FINAL PROJECT

Title of the project:	**Anomaly detection using deep learning in Lung Cancer CT scans**
Author name:	Dmytro Pravdyvets
Project supervisor:	Sergi Martínez Maldonado
Coordinating professor:	Albert Solé Ribalta
Date of submission (MM/YYYY):	06/2023
Name of the degree:	Máster en Ciencia de Datos
Topic of the final project:	Health Sciences
Language:	*English*
Keywords:	Lung cancer, Attention, Deep learning

Abstract

Lung cancer is a leading cause of cancer-related deaths worldwide. Computed Tomography (CT) scans are a primary imaging modality used for the diagnosis and treatment of lung cancer. However, interpreting and analyzing CT scans can be challenging due to the complex and heterogeneous nature of lung cancer. Attention maps have emerged as a promising tool to improve the interpretability of CT scans by highlighting regions of interest.

Attention maps are a visualization technique that uses deep learning algorithms to identify and highlight regions of an image that are important for a given task. In the case of lung cancer CT scans, attention maps can be used to identify regions of interest such as nodules, masses, or other abnormal tissue structures. By providing a visual representation of these regions, attention maps can assist radiologists in making more accurate diagnoses and treatment decisions.

Several studies have demonstrated the potential of attention maps in improving the accuracy and efficiency of lung cancer diagnosis. For example, attention maps have been used to detect lung nodules with high sensitivity and specificity, outperforming traditional detection methods. They have also been

used to predict the malignancy of nodules, with promising results.

Despite the potential benefits, attention maps still face several challenges in their application to lung cancer CT scans. One of the main challenges is the need for large datasets with annotated regions of interest. Another challenge is the potential for overfitting, where the attention maps may highlight features that are not clinically relevant.

We believe that attention maps hold great promise for improving the interpretability of lung cancer CT scans and in this study we are looking to optimize them to overcome the difficulties and limitations that might occur in the context of lung cancer.

Index

List of figures

1. Introduction

Lung cancer is a type of cancer that starts in the cells of the lungs, usually in the lining of the bronchi or the alveoli. It can develop in any part of the lungs and can spread to other parts of the body through the lymphatic system or the bloodstream.

The most common cause of lung cancer is smoking tobacco, either through cigarettes, cigars, or pipes. Exposure to secondhand smoke, radon gas, asbestos, and other environmental pollutants can also increase the risk of developing lung cancer. In rare cases, genetic mutations or family history can increase the likelihood of developing lung cancer, even in people who have never smoked. Treatment options for lung cancer include surgery, radiation therapy, chemotherapy, targeted therapy, and immunotherapy, depending on the type and stage of the cancer. Early detection through screening with CT (computed tomography) scans can improve treatment outcomes and increase the chances of survival.

CT scans are used to detect lung cancer because they provide detailed images of the lungs and can detect small nodules or masses that may be cancerous. CT scans use X-rays and computer technology to create cross-sectional images of the body, which can reveal abnormalities in the lungs that may be indicative of cancer. Early detection of lung cancer through CT scans can improve treatment outcomes and increase the chances of survival.

It is also important to notice that lung cancer is a major problem worldwide. Based on the latest epidemiological data[1][2][3] we can see that:

1. According to the World Health Organization (WHO), lung cancer is the most commonly diagnosed cancer worldwide, accounting for 11.6% of all cancer cases in 2020.
2. In 2020 the total number of new lung cancer cases was 2.2 million with a total of 1.8 million deaths.
3. The prevalence of lung cancer varies greatly between nations, with low- and middle-income nations often having the highest rates.
4. Recent years have seen a decline in lung cancer incidence and mortality rates in high-income countries, primarily as a result of smoking rates falling and improvements in early identification and treatment.
5. In contrast, lung cancer rates are still rising in low- and middle-income nations as a result of a high smoking prevalence, exposure to air pollution, and other risk factors.
6. Almost 85% of cases of lung cancer are caused by tobacco use, making it the world's most common cause.
7. The presence of air pollution, occupational exposure to carcinogens like asbestos, and a family history of the disease are additional risk factors for lung cancer.
8. Depending on the stage at diagnosis, lung cancer has a different prognosis, with early-stage illness having a better prognosis than

advanced-stage disease. Lung cancer five-year survival rates are typically lower in low- and middle-income nations than in high-income nations.

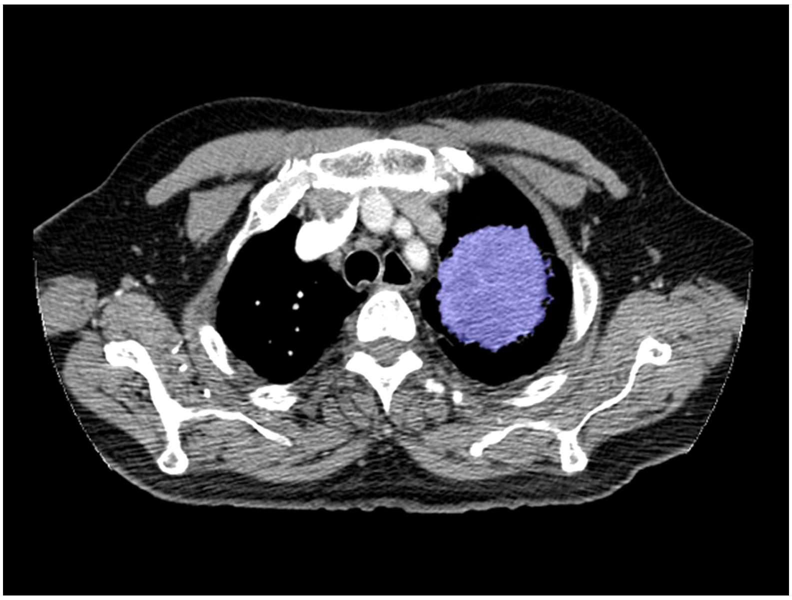

Figure 1: CT scan example.

1.1. Context and motivation

Attention maps on CT scans are relevant because they can help identify regions of interest within the image data. Computed tomography (CT) scans create detailed images of the inside of the body by taking multiple X-ray images from different angles and using a computer to reconstruct a 3D image. These images contain a wealth of information and attention maps can be used to highlight areas that are important for further analysis or diagnosis.[4][5]

Attention maps in CT scans are typically generated using deep learning techniques, where a neural network is trained to identify regions of interest based on patterns in the image data.[5][6] The resulting attention map highlights these regions by assigning higher values to the pixels that are most important to the network segmentation or classification task.[7] For example, in the case of lung cancer diagnosis, care maps can help identify suspicious regions that may contain tumors or other abnormalities.

Fortunately there is a large amount of public data on CT scans. This allows us to create a model that will perform well at detecting anomalies in new images that the model has not previously seen.[8]

1.2. Goals

In this project we want to develop a tool that makes it easier for radiologists and other medical professionals to focus their attention on specific regions of interest within the CT scan, potentially improving the accuracy and efficiency of their diagnoses. Additionally, attention maps can be used to guide automated image analysis and segmentation algorithms, making these processes more reliable and accurate. Overall, care maps are a powerful tool for analyzing and interpreting CT data and are an important part of modern medical imaging.

The main objectives of the work are:

1. Obtain an unbiased set of public data that is as diverse as possible to cover as much variability as we can.
2. Create a framework to modify and parse image data into a format that a Deep Neural Network can ingest it.
3. Create and train a DNN model that is capable of identifying abnormalities in CT scans for lung cancer.
4. Learn how to interpret the results in a clinical context in a way it is readable for a clinician.

1.3. Sustainability, diversity, and ethical/social challenges

This section should assess the positive/negative impact of the project in the following dimensions. It is not required to reach a positive impact in any/all dimensions, but it is necessary to consider and discuss whether there is an impact or not from the beginning of the project.

Sustainability

The use of attention neural networks in CT scans has the potential to reduce the ecological footprint of medical imaging by improving the accuracy and efficiency of the diagnostic process.

Firstly, attention neural networks can help reduce the number of CT scans needed to make a diagnosis by improving the accuracy of image analysis, thereby reducing the need for follow-up scans. This can lead to a reduction in energy consumption and resource usage associated with producing and analyzing additional images.

Secondly, attention neural networks can also help reduce the amount of radiation exposure associated with CT scans. By identifying regions of interest within the image data, attention maps can help guide radiologists to focus on specific areas, reducing the need for scanning the entire body. This can lead to a reduction in the amount of radiation exposure, which is important for both patient safety and reducing the environmental impact of medical imaging.

Finally, attention neural networks can potentially reduce the depletion of raw materials associated with the production of CT scanners and related equipment.

By reducing the need for follow-up scans and unnecessary tests, attention neural networks can help extend the lifespan of existing equipment, reducing the need for new equipment and thereby reducing the demand for raw materials.

Overall, the use of attention neural networks in CT scans has the potential to improve the sustainability and reduce the ecological footprint of medical imaging by improving accuracy, reducing the need for unnecessary tests, and potentially reducing the amount of radiation exposure and waste associated with the diagnostic process.

From a computational point of view, we aim to work locally on a laptop with a small GPU unit. Using a small graphics unit (GU) can be a sustainable choice because it consumes less energy and produces less heat than larger graphics units. The use of smaller GUs can also extend the lifespan of electronic devices by reducing wear and tear on other components, such as the power supply and cooling system. In addition, smaller GUs may require fewer raw materials to manufacture, reducing the overall environmental impact of the production process. By reducing energy consumption, extending device lifespan, and minimizing the use of raw materials, using a small GU can help reduce the ecological footprint of electronic devices and promote sustainability.

Ethical behavior and social responsibility

Using CT scans of confidential patients can be ethical if certain measures are taken to protect patient privacy and ensure that the benefits of using the scans outweigh the potential risks. The use of CT scans can be critical in the diagnosis and treatment of medical conditions, and withholding scans from patients could potentially harm their health. However, it is important to obtain informed consent from patients before using their scans and to take steps to ensure that the scans are stored and transmitted securely to protect patient privacy. Additionally, using de-identified scans for research purposes can help ensure patient confidentiality while still allowing for the advancement of medical knowledge. Overall, the ethical use of CT scans of confidential patients requires balancing the potential benefits against the risks to patient privacy and well-being.

Taking into account that we are using a public database where the patients have consented to the use of the data and cannot be traced back from that data we think there is no potential harm to those individuals from this project.

Diversity, gender and human rights

Diversity, gender, and human rights considerations in public CT scan data involve ensuring that the data used for research is representative of different populations and does not perpetuate discrimination or biases. This includes collecting data from diverse populations to ensure that research findings apply to all individuals, regardless of gender or race. Additionally, it is important to ensure that CT scan data is obtained and used ethically, with respect for human rights and privacy. This includes obtaining informed consent from individuals and protecting their personal information. By promoting diversity, gender equality, and respect for human rights, public CT scan data can be used to improve medical understanding and promote health equity for all. The public dataset we are using comes only from China so the diversity is not as big and we would like it to be, but we also have to keep in mind that the scale of the project is on a level of a master thesis and not on a level of a commercial product.

1.4. Approach and methodology

The use of care maps in the medical field is a field that has already been explored, so in this project we want to optimize the development of our model using the information already published on previous models, but we do not want to reuse existing models, but to create a new one. in order to facilitate the interpretation for personnel who do not have so much knowledge of the computer world.

1.5. Schedule

For this work we will use public data from: [A Large-Scale CT and PET/CT Dataset for Lung Cancer Diagnosis (Lung-PET-CT-Dx)](). Once we get and process the images we are going to use Python as the main language. Libraries like PyTorch to create the model and train it. We are going to use one of the available Python data visualization libraries like Matplotlib or Seaborn to visualize the resulting attention maps.

Obtaining and processing the images should not require much time. In this way, where we are going to focus more in the time dedicated to the thesis, it will be in the improvement of the model and the facilitation of the interpretation of the results.

The work schedule can be seen here:

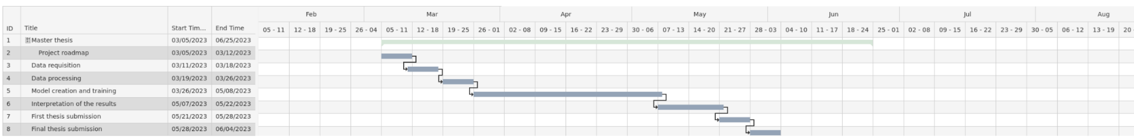

Figura 2: Gantt diagram.

The planning consists of separating the project into 3 large blocks with 2 to 3 sub-blocks. Chapter 1 will consist of project planning and data collection. We want to minimize the time dedicated to this block since they are preliminary works to the important part that is the next chapter. Chapter 2 will consist of the creation of the anomaly detection model. In this chapter, the first block will be dedicated to the creation of a CNN (Convolutional Neural Network) model with attention layers that will be used to create the attention map that will be used in subblock two, where we will improve the model and seek to have maximum interpretability of the results. Finally, chapter 3 will focus on the writing of the thesis and the defense. We want to have all the analyzes already done before we start writing the document.

We have also taken into account that some parts of the project might not go as planned. We are aware that there are other public CT scan datasets available online in case the performance we obtain with this dataset is not good enough or if we want to have an external training/validation dataset to benchmark our model on. Also we will try to maintain the model small so it can be trained locally so we stay inside of our sustainability plan. Lastly we have looked into the state of the art models in the area so we expect to try different architectures for our DNN model to see how the performance and the readability of the outputs change depending on the architecture.

1.6. Summary of the outputs of the project

The main output of the project is a computational model that is capable of ingesting CT scan images and translating into an attention map that highlights the abnormalities of the scan, if there is any. This map can later be used by a trained radiology technician to evaluate the results by a human being.

Ideally this model is a tool that can be used as an aid in the detection and evaluation of the results of a CT scan to reduce the number of following scans and increase the sustainability as explained in **1.3 Sustainability, diversity, and ethical/social challenges.**

2. State of the art

2.1 State of the art models and approaches

In medical imaging, attention maps are frequently utilized to assist in highlighting the regions of interest in the images. Attention maps can assist in locating abnormal regions on CT scans that may necessitate additional examination.

For the purpose of creating attention maps for CT scans, a number of cutting-edge models have been developed. The Attention Guided Deep Multiple Instance Learning (AGD-MIL) framework, which was proposed by Zhang et al., is one such model. 2020 [9]. Their new framework proposes a new framework called AGD-MIL for medical image analysis using attention-guided deep multiple instance learning. The framework integrates the attention mechanism into a deep multiple instance learning architecture to effectively identify the regions of interest in medical images. The model is trained on a large dataset of CT scans and tested on various segmentation tasks. Results show that AGD-MIL outperforms several state-of-the-art models and demonstrates its efficacy in accurately identifying regions of interest in medical images, making it a promising approach for computer-aided diagnosis and treatment.

The Attention U-Net (AU-Net), developed by Oktay et al., is a new model for generating attention maps in CT scans. 2018 [10]. This paper showcases a new deep learning architecture called Attention U-Net for accurate segmentation of the pancreas in CT scans. The architecture combines the popular U-Net architecture with an attention mechanism to focus on regions of interest and improve the accuracy of segmentation. The model was trained and tested on a dataset of abdominal CT scans and achieved state-of-the-art performance on pancreas segmentation tasks. The results demonstrate the efficacy of the Attention U-Net architecture in accurately identifying the pancreas in medical images, which could have significant clinical applications for computer-aided diagnosis and treatment.

Chen et al.'s Region-Attention Network (RAN) is a third model for creating CT scan attention maps. 2019 [11]. The RAN architecture combines both global and local information to accurately identify regions of interest in mammograms. The model was trained and tested on a dataset of mammography images and achieved state-of-the-art performance on breast mass segmentation tasks. The results demonstrate the efficacy of the RAN architecture in accurately identifying breast masses, which could have significant clinical applications for computer-aided diagnosis and treatment of breast cancer.

In general, these cutting-edge models show that attention maps are effective at accurately identifying CT scan regions of interest, which can boost medical diagnosis and treatment accuracy and efficiency. But as can be seen most of the papers are used in areas outside of lung cancer and not as many cutting edge architectures have been used in that area.

One of the most famous studies regarding CT scans and lung cancer is Wang J et al [12]. In their 2019 study, Wang et al. proposed a lung cancer detection model that uses a co-learning approach to combine chest CT images and clinical demographics. The model is based on a deep learning framework that uses convolutional neural networks to extract features from the images and a support vector machine to classify the cancer status. The authors evaluated the performance of the model on a dataset of 1017 patients and achieved an AUC of 0.97, which outperformed other state-of-the-art models of the time. The results suggest that the co-learning approach can effectively integrate multiple sources of information to improve the accuracy of lung cancer detection.

Image diagnostics is a field that is growing constantly with new research, new ideas and techniques. Looking more specifically into lung cancer we can see that a lot of progress has been made since Wang J et al publication has been released[13]. More specifically we can see that algorithms that are based on segmentation of the image tend to have very good performance while being easier to understand and transfer the results into diagnostics and clinical interpretation[14].

Very interesting study comes from Sagayam, K.M. *et al.*[15] where they make the model “pay attention” to specific areas of the CT scan by using means of image segmentation and instead of working with the entire image at a time, splitting the scan into segments showed a better overall performance. A pyramid attention network between the encoder and decoder architecture is a very interesting and novel approach to attention in CT scans that we think can be used in many other ways to boost the resulting accuracy even more. The authors also conducted ablation studies to evaluate the contribution of the different components of the proposed model. The results showed that the pyramidal attention mechanism and the multi-scale feature fusion module were the most significant contributors to the overall performance of the model. This shows us that good image representation and feature selection is essential to obtain a well performing model.

It is also important to remember that as any other Machine Learning field, cancer detection with deep learning can suffer from the same data based issues and biases as the rest of the fields. Ren Ge and colleagues *et al.* [16] have shown how a pre-trained model can be transferred and fine tuned using external datasets and image manipulation to significantly improve the performance of the model. Another interesting aspect of the paper is the usage of 3D scans as the input and evaluation of the 3 separated images and using this to calculate different types of correlations between the results and the information given by each of the 3 images for each lobe of the lungs. The results suggest that different lobes of the lungs provide different information and that the usage of all of the globes is the optimal approach. In summary the study shows that the transfer learning framework synthesized CT-based perfusion images that demonstrated a strong voxel-wise correlation and function-wise similarity with the CT perfusion images for lung cancer patients. Suggesting that deep learning methods have great potential for providing functional information on a regional basis, which could be useful for functional lung avoidance radiation therapy.

In this thesis we aim to find a new approach that has been tested in other areas of CT scans and clinical imaging and see how it performs in lung cancer.

3.Study design and implementation

3.1 Image annotation

Image annotation in computed tomography (CT) scans is an essential step in the process of training deep learning models. In the context of radiology, accurate annotation of structures and pathologies on CT scans is crucial for disease detection and diagnosis. Advances in deep learning techniques have revolutionized the field of radiology and have enabled significant improvements in the accuracy and efficiency of medical image interpretation.

CT annotation refers to the task of identifying and labeling regions of interest within images, such as tumors, lesions, or anatomical structures. This implies precise segmentation of these regions, which can be challenging due to the complexity and variability of shapes and textures on CTs. To address this challenge, deep learning algorithms such as convolutional neural networks (CNNs) are commonly used to perform automatic or computer-assisted annotation.

The key advantage of using deep learning techniques in CT annotation is its ability to learn complex features and patterns directly from the data. These models can learn to recognize subtle features and differences in images, improving annotation accuracy. In addition, deep learning approaches allow for precise semantic segmentation, making it easier to identify and quantify structures of interest.

However, CT notation also presents unique challenges. CT scans typically have a large number of 3D slices, requiring precise annotation in multiple planes to fully capture the information. In addition, annotations must consider anatomical variability between patients and possible deformations due to patient position or breathing.

In the context of this project, the annotations come in a **.xml** file where the field **"size"** defines the 4 coordinates that define the bounding box. The coordinates are **xmin, ymin, xmax, ymax.** These 4 numbers define the location of the box where cancer has been identified by a group of radiologists. Each DICOM image is linked to an annotation file through a **"dicom_num"** field. Unfortunately, some DICOM images did not have linked annotation files and the same stays true the other way around, the files with missing counterparts were excluded from the study.

The previously mentioned coordinates of the bounding box are of high interest in our context since we designed our algorithm in such a way that it can predict the 4 values and define the bounding box for unlabeled images.

3.2 Convolutional neural network

Convolutional Neural Networks (CNNs) have revolutionized the field of lung cancer detection in computed tomography (CT) scans. These powerful deep learning models are highly effective in identifying subtle features and patterns associated with lung cancer, leading to significant improvements in diagnostic accuracy.

The key to the success of CNNs in CT lung cancer detection lies in their ability to automatically learn relevant features directly from image data. These features may include abnormal shapes, irregular densities, or suspicious nodules in the lungs. Through multiple layers of convolution and pooling, CNNs can extract discriminative features from CT images and generate activation maps that highlight areas of interest.

Lung cancer detection on CT with CNN generally involves a binary classification approach. After feature extraction, a fully connected layer is used to perform the final classification. These CNNs are trained on large annotated CT data sets, where regions of interest have been previously marked by expert clinicians. The training process involves adjusting the weights and parameters of the CNN to minimize the difference between the model predictions and the expert annotations.

In the context of our project, we have used a relatively simple CNN that consists of 2 convolutional layers, one pooling layer and 2 linear layers that output the 4 coordinates of the bounding box that is later compared to the manually annotated one to check the performance of the model.

The architecture of the model can be seen here:

```
class SimpleCancerCNN(nn.Module):
    def __init__(self):
        super(SimpleCancerCNN, self).__init__()
        self.conv1 = nn.Conv2d(1, 16, kernel_size=3, stride=1, padding=1)
        self.conv2 = nn.Conv2d(16, 32, kernel_size=3, stride=1, padding=1)
        self.pool = nn.MaxPool2d(kernel_size=2, stride=2)
        self.fc1 = nn.Linear(32 * 128 * 128, 256)
        self.fc2 = nn.Linear(256, 4)

    def forward(self, x):
        x = self.pool(torch.relu(self.conv1(x)))
        x = self.pool(torch.relu(self.conv2(x)))
        x = x.view(-1, 32 * 128 * 128)
        x = torch.relu(self.fc1(x))
        x = self.fc2(x)
        return x
```

As input, this model takes a 512 by 512 matrix that represents the DICOM image that has been transformed previously into a grayscale. This input is processed by the neural network and as output, we obtain the 4 values of the bounding box. Ideally the area of the original bounding box and the predicted one overlap perfectly or almost perfectly.

The dataset contains 4 different subtypes of lung cancer, those being: 'A' was diagnosed with Adenocarcinoma, 'B' with Small Cell Carcinoma, 'E' with Large Cell Carcinoma, and 'G' with Squamous Cell Carcinoma. Due to hardware limitations, we were unable to use the entire dataset to train the model, so we decided to work on a subset of 1000 images of subtype “B”.

For training we made a random subset of 1000 DICOM images of subtype “B” so in this way, we do not bias the dataset for specific patients or tumor locations. We used 800 images for training and 200 images to test the model.

The model has been trained for 1000 epochs using ADAM optimizer and Mean Square Loss. ADAM optimizer is commonly used in deep learning due to its adaptive learning rate and momentum. It efficiently updates model parameters during training, leading to faster convergence and better performance. MSELoss measures the mean squared difference between predicted and target values, making it suitable for our use case.

The evolution of loss can be seen here:

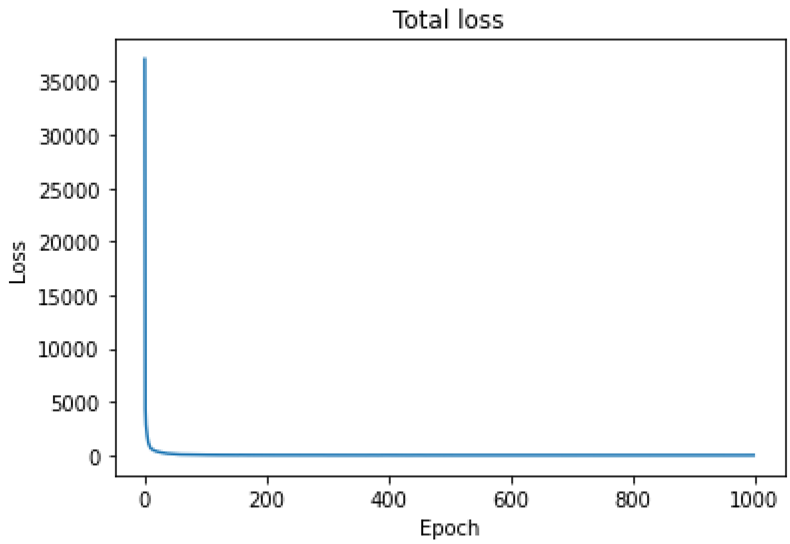

Figure 3: Model loss

Here we can see the total loss of the model during the 1000 epochs. As can be seen, the initial loss is very big since the model assigns randomly the bounding boxes, making it hard to see the evolution of the model post epoch 1. For this, we have made the same visualization using the loss values of epochs after epoch 100:

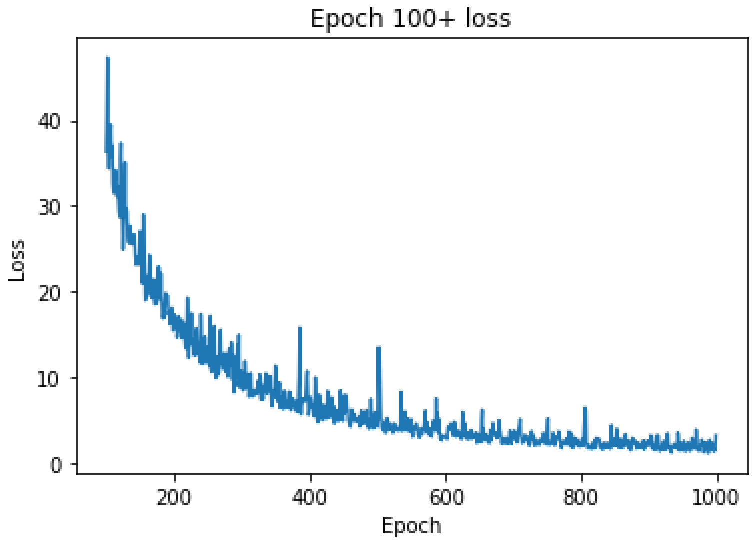

Figure 4, Epoch 100+ loss

As can be seen, the loss goes down over time, not perfectly steady, but it does reach almost 0, implying that the performance of the model is quite good. We can see some oscillation in the loss, when the loss function is oscillating during gradient descent, it indicates that the optimization process is struggling to find the global minimum. This oscillation occurs because the gradient descent algorithm overshoots and undershoots the minimum in successive iterations. Nevertheless, due to a good learning rate, the model can successfully learn to find tumors in CT scans based on the characteristics of the image.

Lastly, as the last step of the study, we have added Sobel’s edge detection algorithm. Sobel's edge detection algorithm is a commonly used technique to identify and highlight edges in images. This algorithm is based on the convolution of the original image with two filters, one to detect vertical changes and another to detect horizontal changes in pixel intensity. Combining these two answers gives an approximation of the magnitude of the gradient at each point in the image, revealing the location of the edges.

In the context of object detection in computed tomography (CT) scans for lung cancer, Sobel's algorithm is interesting because of its ability to highlight regions of interest, such as lung nodules. These nodules often present well-defined borders on CT images, and accurate detection of these borders is critical in identifying and diagnosing lung cancer.

By applying Sobel's algorithm to CT images, the edges of lung nodules can be detected, which aids in the segmentation and precise localization of these regions. This is essential to make an early and accurate diagnosis, as well as to follow the progression of the disease.

In addition, Sobel's algorithm is computationally efficient and relatively simple to implement, which makes it suitable for application in real-time medical image processing. By highlighting the edges, this algorithm provides important visual information and serves as a crucial step in the automatic detection of lung nodules on CT.

Here we can see an example of an image from the test set of subtype “B”:

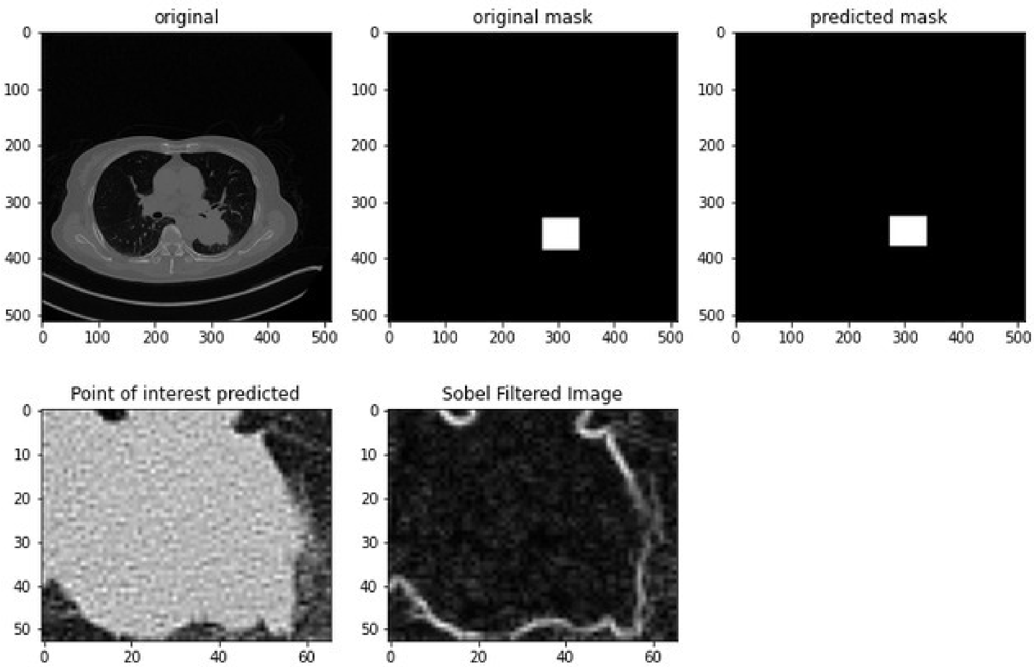

Figure 5, Result “B”

In order we have: 1. The CT scan itself, 2. The area annotated by the radiologist, 3. The area predicted by the model, 4. Zoom in into the predicted bounding box 5. A cleansed outline of the tumor using the Sobel edge detection algorithm.

As can be seen, the model can detect the area of interest and using Sobel edge detection algorithm we can clear noise for noise CT scans to get a clear outline of where the tumor is located and its shape.

As in any other object detection model, we have used Intersection over Union (IoU) as a metric to verify the results. In the case of the test dataset, the IoU was 0.8, not perfect but a good starting point as an assistance tool for clinicians.

3.3 Transfer learning

Transfer learning is a technique in machine learning where knowledge learned from a previously trained model is used to solve a similar problem. Instead of training a model from scratch, features learned from a pre-existing model are taken advantage of and transferred to a new model.

This allows for faster and more efficient training, especially when you have small data sets. Transfer learning is useful in various tasks, such as object recognition, image classification, and natural language processing, and has been shown to significantly improve model performance.

In the context of this study, as mentioned before, the model was trained using only 1 subtype of lung cancer. We think it is interesting to see how well this model transfers to other lung cancer subtypes. For this, we used the pre-trained model and ran it on 1000 images for each of the other cancer subtypes “A”, “G”, “E”.

The results were not very good, in many cases the model was not able to identify the region of the tumor and the respective IoUs were small:

Subtype	IoU
A	0.07
G	0.15
E	0.17

Here are visual examples of the results:

A:

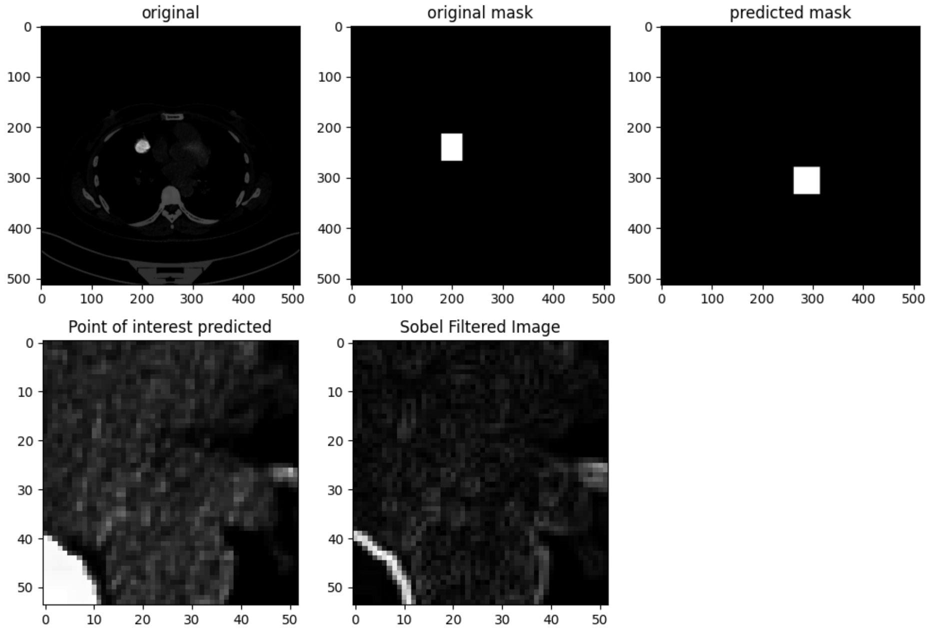

Figure 6, No fine-tuning “A”

G:

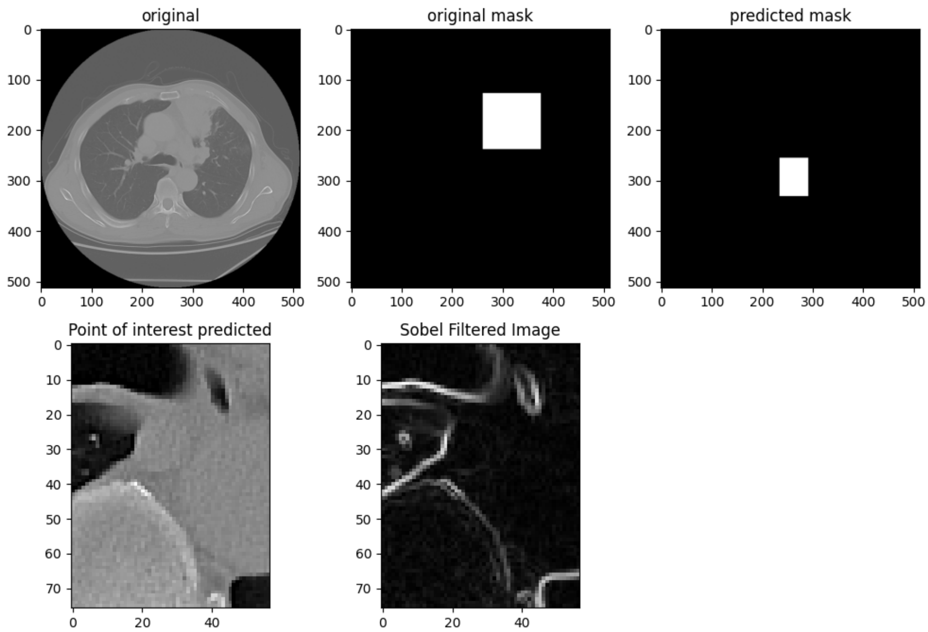

Figure 7, No fine-tuning “G”

E:

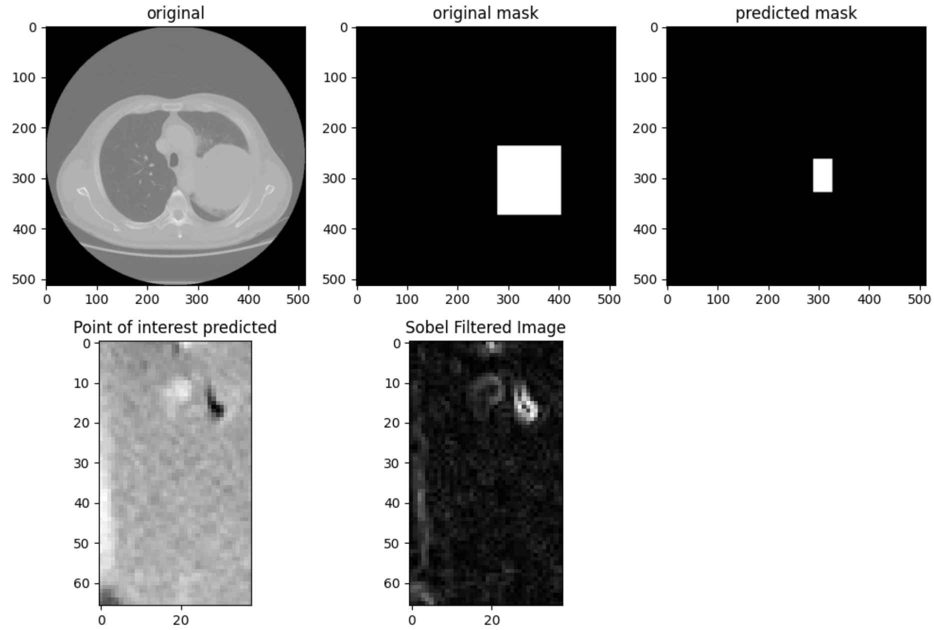

Figure 8, No fine-tuning “E”

As can be seen, the model either does not find the location of the tumor or does not identify the limits on the bounding box correctly. For this, we thought that it might be interesting to do some fine-tuning on the unseen datasets.

3.4 Fine-tuning

Fine-tuning for unseen data is a process in machine learning in which a pre-trained model is specifically adapted or "tuned" for a new data set that has not been used during its initial training. This involves taking a model previously trained on a large and diverse data set and then continuing its training on a more specific and task-relevant data set. By finely tuning the model, you want to improve its ability to generalize and make accurate predictions on the new, unseen data. During fine-tuning, you typically fit model weights and parameters using a more limited training data set and allow the model to adapt to the unique features and patterns present in that new data.

Fine-tuning is especially useful when you have a limited amount of data for the task at hand, as it allows you to take advantage of the knowledge and representations learned by the pretrained model in other domains or related tasks. By fine-tuning a model, a balance can be achieved between prior knowledge and adaptation to new data, which can lead to better performance and greater generalizability to the specific context of unseen data.

In the context of this study, we have trained the pre-trained model on a subset of unseen data for a small number of epochs to see how much it can boost the performance. For this, we used 1000 images of subtypes “A” and “G” since the number of available images was large, but for subtype “E” we used the entire dataset of 201 images. For each subtype, we randomly split the data using an 80/20 split and trained the model for 100 epochs to see how much this can increase the IoU of the entire dataset and of the test dataset.

The evolution of loss for each subtype can be seen here:

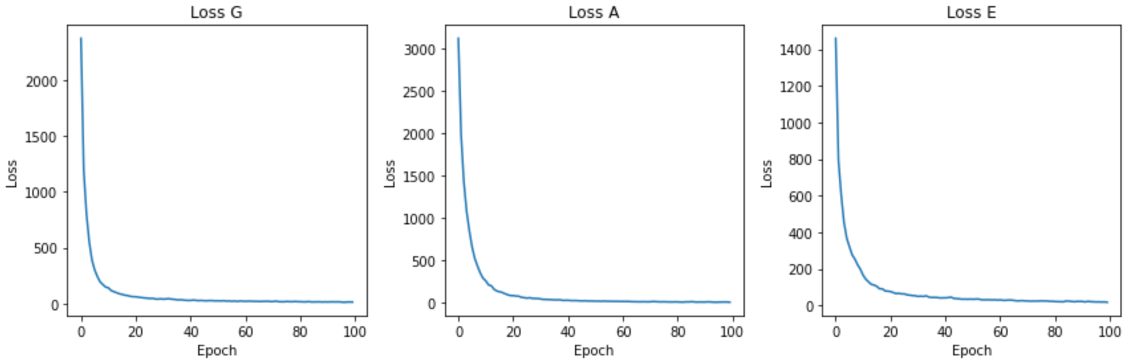

Figure 9, Loss fine-tuning

Since the loss in epoch 1 is so big, we showcase the same plot after epoch 10:

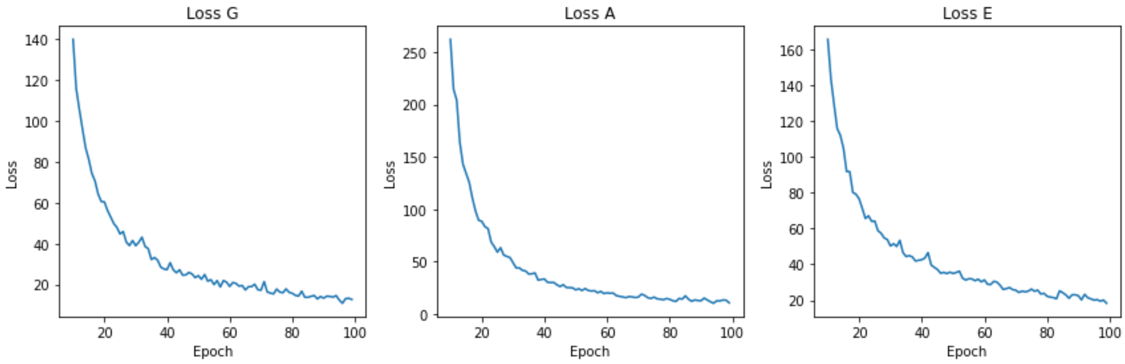

Figure 10, Epoch 100+ loss fine-tuning

This translates into the model giving a higher IoU score for each subtype. In this case, we can see a clear increase in IoU:

Subtype	Entire dataset	Test subset	Old IoU
A	0.73	0.67	0.07
G	0.77	0.73	0.15
E	0.8	0.73	0.17

As can be seen, by training the model pre-trained on the “B” subtype we can decrease the time needed to get a higher value of IOU. We consider that any IOU over 0.7 is a good result in the context of our study.

To visually see the improvements here we have some examples of the fine tuned models:

A:

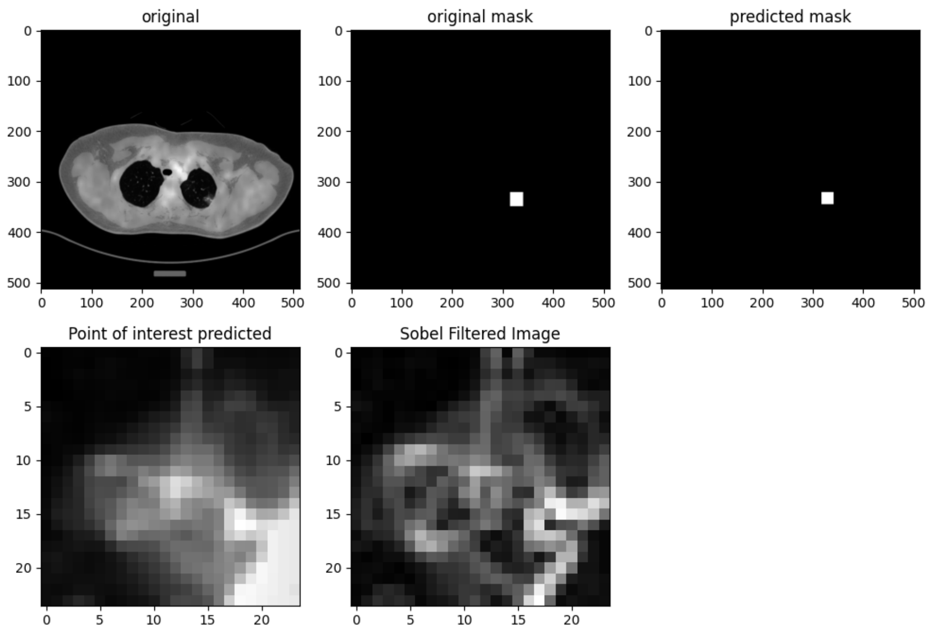

Figure 11, Fine-tuning “A”

G:

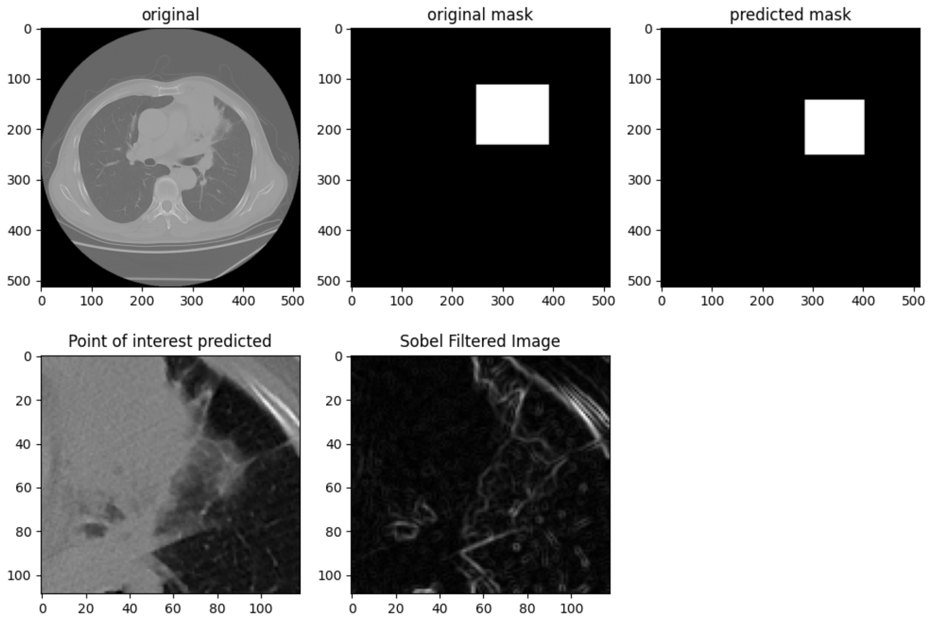

Figure 12, Fine-tuning “G”

E:

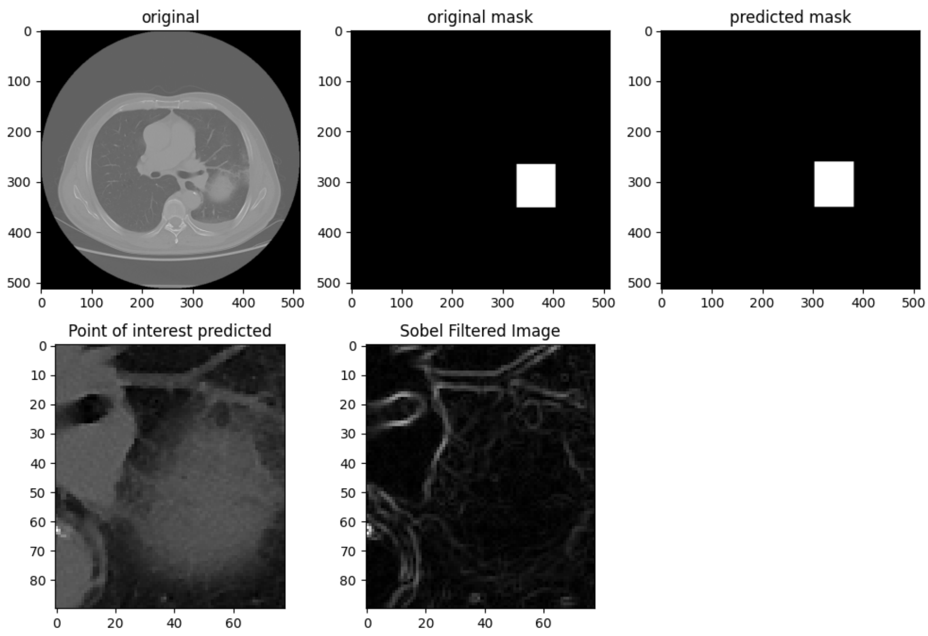

Figure 13, Fine-tuning “E”

3.5 Possible improvements

There are different ways this study can be improved:

1. Overcoming hardware limitations: This study has been performed on an MSI laptop with the following characteristics: **32GB Ram, 11th Gen Intel® Core™ i7-1185G7 @ 3.00GHz × 8, NVIDIA Corporation / NVIDIA GeForce GTX 1650 Ti with Max-Q Design/PCIe/SSE2.** We did not have access to cloud services, thus we had hardware limitations and time restrictions. With the ability to have a better GPU and more RAM, it would be possible to use the entirety of the “B” subtype and/or expand the model to let it learn on the entire dataset and see how this affects the performance.

2. Model complexity: Due to hardware limitations, the complexity of the model was also restricted since bigger models would not have enough memory to run, or it would take too long to train them on smaller batches. It is also worth mentioning that a much more complex model does not directly translate into much better performance, so it would be interesting to find the “sweet spot” of model complexity, training time, and performance.

3. Identifying the exact cancer shape instead of the bounding box: It would be interesting to see if by looking at the weights of the model and adding attention it would be possible to let the model identify the exact area of the tumor instead of the bounding box.

4. Add labeling into the algorithm: It would be interesting to add labeling of the CT scan as another output of the model and see if it can identify: 1) If the image comes from a patient with cancer 2) Where is that cancer located. Computing a ROC curve, contingency table and other typical metrics per cancer subtype would also be interesting to see which one is the hardest to identify.

Bibliography

Here we have attached all the scientific publications in the field that we consider as the basis on which we are going to build our work.

1. Alberg AJ, Brock MV, Samet JM. Epidemiology of Lung Cancer: Looking to the Future. J Clin Oncol. 2005 Sep 20;23(27):3175-85. doi: 10.1200/JCO.2005.09.163. PMID: 15955903.
2. de Groot PM, Wu CC, Carter BW, Munden RF. The epidemiology of lung cancer. Transl Lung Cancer Res. 2018 Jun;7(3):220-233. doi: 10.21037/tlcr.2018.05.06. PMID: 30050761; PMCID: PMC6037963.

3. Malhotra J, Malvezzi M, Negri E, La Vecchia C, Boffetta P. Risk factors for lung cancer worldwide. Eur Respir J. 2016 Nov;48(5):889-902. doi: 10.1183/13993003.01128-2016. Epub 2016 Aug 18. PMID: 27539869.
4. Kermany DS, Goldbaum M, Cai W, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell. 2018;172(5):1122-1131. doi:10.1016/j.cell.2018.02.010
5. Wang G, Li W, Aerts HJWL, et al. Radiomics and Deep Learning: Competitors or Allies in Cancer Imaging Research? Radiology. 2019; 293(2): 349-350. doi: 10.1148/radiol.2019190086
6. Wang S, Summers RM. Machine learning and radiology. Med Image Anal. 2012;16(5):933-951. doi:10.1016/j.media.2012.02.005
7. Chen C, Dou Q, Chen H, et al. Multi-scale attention convolutional neural network for Alzheimer's disease diagnosis using MR images. Pattern Recognit. 2018;80:25-34. doi: 10.1016/j.patcog.2018.02.017
8. Choy G, Khalilzadeh O, Michalski M, et al. Current Applications and Future Impact of Machine Learning in Radiology. Radiology. 2018;288(2):318-328. doi:10.1148/radiol.2018171820
9. Zhang, J., Cai, Z., Wu, Y., Wang, Y., He, Y., & Cai, D. (2020). Attention Guided Deep Multiple Instance Learning for Medical Image Analysis. IEEE Transactions on Medical Imaging, 39(3), 719-729.
10. Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M. P., Misawa, K., ... & Rueckert, D. (2018). Attention U-Net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.
11. Chen, L. C., Yang, Y., Wang, J., Xu, W., & Yuille, A. L. (2019). Region-Attention Network for Breast Mass Segmentation in Mammography. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 154-162). Springer, Cham.
12. Wang J, Gao R, Huo Y, Bao S, Xiong Y, Antic SL, Osterman TJ, Massion PP, Landman BA. Lung Cancer Detection using Co-learning from Chest CT Images and Clinical Demographics. Proc SPIE Int Soc Opt Eng. 2019 Mar;10949:109491G. doi: 10.1117/12.2512965. PMID: 31602088; PMCID: PMC6786775.
13. Wang, L. Deep Learning Techniques to Diagnose Lung Cancer. *Cancers* **2022**, *14*, 5569. https://doi.org/10.3390/cancers14225569
14. Said, Y.; Alsheikhy, A.A.; Shawly, T.; Lahza, H. Medical Images Segmentation for Lung Cancer Diagnosis Based on Deep Learning Architectures. *Diagnostics* **2023**, *13*, 546. https://doi.org/10.3390/diagnostics13030546
15. Bruntha, P.M., Pandian, S.I.A., Sagayam, K.M. *et al.* Lung_PAYNet: a pyramidal attention based deep learning network for lung nodule

segmentation. *Sci Rep* **12**, 20330 (2022). https://doi.org/10.1038/s41598-022-24900-4
16. Ren Ge, Li Bing, Lam Sai-kit, Xiao Haonan, Huang Yu-Hua, Cheung Andy Lai-yin, Lu Yufei, Mao Ronghu, Ge Hong, Kong Feng-Ming (Spring), Ho Wai-yin, Cai Jing. A Transfer Learning Framework for Deep Learning-Based CT-to-Perfusion Mapping on Lung Cancer Patients. Frontiers in Oncology 12 2022.10.3389/fonc.2022.883516