*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8297*
***Article Received:** 06 July 2023 **Revised:** 28 August 2023 **Accepted:** 12 September 2023*

# Hybrid Algorithm for the Detection of Lung Cancer Using CNN and Image Segmentation

**Drishti Singh, Jaspreet Singh**
School of Engineering and Sciences, GD Goenka University, Gurugram
drishti.190089@gdgu.org, jaspreet.singh@gdgu.org

**Abstract:** When it comes to cancer-related fatalities, lung cancer is by far the most common cause. Early detection is the key to a successful diagnosis and treatment plan for lung cancer, just as it is for other types of cancer. Automatic CAD systems for lung cancer screening using computed tomography scans primarily involve two steps: the first step is to detect all potentially malignant pulmonary nodules, and the second step is to determine whether or not the nodules are malignant. There have been a lot of books published recently about the first phase, but not many about the second stage. Screening for lung cancer requires a careful investigation on each suspicious nodule and the integration of information from all nodules. This is because the presence of pulmonary nodules does not always indicate cancer, and the morphology of nodules, including their shape, size, and contextual information, has a complex relationship with cancer. In order to overcome this problem, we suggest a deep CNN architecture that is different from the architectures that are commonly utilised in computer vision. After the suspicious nodules have been formed with the modified version of U-Net, they are used as an input data for our model. First, the suspicious nodules are generated with U-Net. To automatically diagnose lung cancer, the suggested model is a multi-path CNN that concurrently makes use of local characteristics as well as more general contextual characteristics from a wider geographical area. In order to accomplish this, the model consisted of three separate pathways, each of which used a different receptive field size, which contributed to the modelling of distant dependencies (short and long-range dependencies of the neighbouring pixels). After that, we concatenate characteristics from the three different pathways in order to further improve our model's performance. In conclusion, one of the contributions that we have made is the development of a retraining phase system. This system enables us to address issues that are caused by an uneven distribution of picture labels. The experimental findings from the KDSB 2017 challenge demonstrate that our model is more adaptable to the described inconsistency among the nodules' sizes and shapes. Furthermore, our model obtained better results in comparison to other researches.

**Keywords:** CNN, Image segmentation, CT scans, CAD systems.

## I. Introduction

Lung cancer is the leading cause of death associated with cancer in every region of the world (Rebecca et al.2018; WHO 2018). According to a research that was published in 2018 by the World Health Organization (WHO), lung cancer has been linked to an estimated 1.76 million deaths worldwide (WHO 2018). Lung cancer was responsible for an estimated 83,550 deaths in the United States of America alone in 2018, out of the total number of deaths that were caused by cancer (Rebecca et al. 2018). It is anticipated that the number will be significantly higher in emerging countries. The major types of cancer for the estimated mortality by sex in the United States of America are presented in Figure 1. Lung cancer is the leading type of cancer in both males and females. Early detection and treatment of lung cancer, like that of other malignancies, is your best option for beating the disease. To this aim, the fundamental and most important stage in the early diagnosis and treatment of lung cancer is identifying the lung in order to determine whether or not it is infected by cancer, with improved screening methods leading to improved patient outcomes. According to the findings of the national lung screening study (NLST), screening using a low-dose helical computed tomography (CT) scan reduced the death rate by 20% when compared to single view radiography in high-risk demographics (The National Lung Screening Trial Research Team 2011). However, there is a high incidence of false positive results when screening for lung cancer, which drives up costs and causes anxiety for patients (Patz et al. 2014). The use of computer-aided diagnosis (CAD) for lung cancer results in a higher attention paid to early screening and a lower rate of false positive diagnoses.

CAD of medical imaging using deep learning with CNN methods has achieved great success over the other state of the art for automated medical imaging applications (Alvarez et al. 2012). These methods, for instance, are able to detect skin cancer metastases (Liu et al. 2017), obtaining significantly better sensitivity performance than human. This success comes on the heels of CAD's great success in natural image recognition and classification. Also, Kingsley et al. 2017 tackled automatic lung cancer screening using 3D-CNN and found promising performance, although having certain shortcomings such as speed inference and memory efficiency. This study was published in 2017. However, the available automatic screening approaches for lung cancer have not proven results that are accurate and reliable enough for clinical

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8297*
*Article Received: 06 July 2023 Revised: 28 August 2023 Accepted: 12 September 2023*

usage for any of the several reasons described in this research. Detecting lung cancer can be difficult for a number of reasons, including the following:

- The intricate nature of the information that is contextualised within nodules. Because of the complex nature of the contextual information they include, the currently used algorithm for cancer screening has an issue with its accuracy, which makes it normally very difficult for doctors.
- The size and form that cancer nodules have by their very nature The morphology of cancer nodules, as well as the morphology of nodules in general, can vary from one CT scan image of a lung to the next. During the screening process, ambiguity will be caused for the radiologist and/or oncologist as a result of the inconsistent shapes and sizes of cancer nodules. As a consequence of this, radiologists and/or oncologists may advise patients to undergo additional procedures such as monitoring, blood testing, and biopsies; as a consequence of this, it becomes a burden for patients. Not only is it difficult for humans, but it is also challenging for an algorithm to locate the features of cancer with their particular sizes and shapes. The variation that was stated among the nodules is illustrated in figure 2.

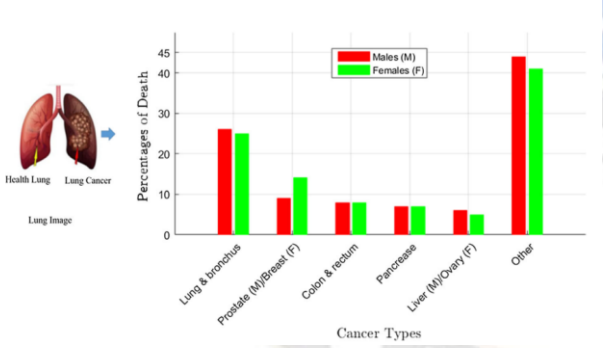


Figure 1 Cancer kinds that account for the greatest proportion of estimated fatalities in the United States in 2018 (Rebecca et al. 2018)

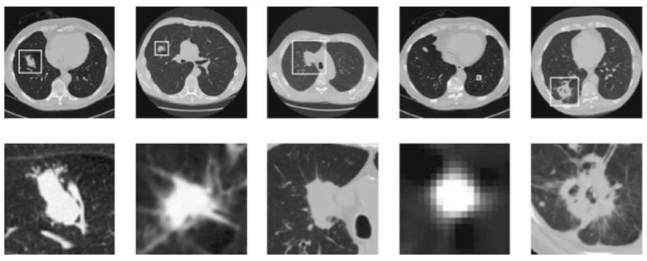

Figure 2 In the KDSB 2017 dataset, there are several examples of nodules of varying sizes and shapes. Topping: the entirety of the slice. At the bottom, the magnified versions of their respective images

The degree to which neighbouring pixels are accurately modelled can be used to differentiate between methods for detecting cancerous and noncancerous lung nodules using CNN. This is because the modelling of both long and short range dependencies is what determines the accuracy of model learning. Most of the time, CNN models that are used to identify lung cancer make use of receptive field sizes of 3x3, 5x5, or 7x7 in order to model dependencies in a single pathway ([5],[9],[20]). Although these methods provide significantly higher performance than the conventional methods for detecting lung cancer, they do not jointly incorporate the local and global range dependencies that are necessary for accurate diagnosis. By way of illustration, one can deduce from Figure 3 that the values gi1, gj1, and gk1 represent the corresponding outputs of the receptive fields Ri, Rj, and Rk when convolution is applied. The receptive field Ri is approximated by gi1, which means that all of the pixels of Ri are approximated by gi1 through the use of convolution. The same may be said about gj1 and gk1 as well. In light of this, the entirety of the feature maps are modelled. However, if we look at what gi1, gj1, and gk1 compute, we can see that they are significantly different from one another. This is because the receptive field size that was employed was different in each case. The optimal size is determined on the work at hand. As a result, it is essential to create an architecture that takes into account the influence of receptive field size on an average level.

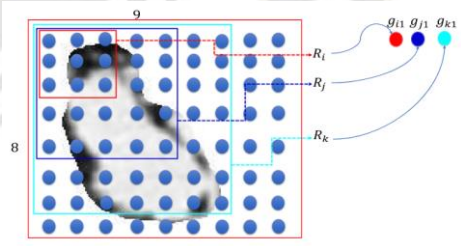


Figure 3 An illustration showing the different sizes of receptive fields. On the left is a feature map that has three separate receptive fields (Ri, Rj, and Rk), and on the right is the convolutional output that corresponds to those fields. The image in the backdrop is a sample of a nodule.

In addition, the current models are plagued by a lack of labelled data as well as an imbalance between the different classes of data. For instance, the malignancy to benign ratio in the CT scan used in the KDSB 2017 challenge dataset is 1:7. This indicates that there is a significant imbalance in the data. In this research, we offer a multi-path Convolution neural network which we refer as mp-CNN that, in general, we provide as a flexible and adaptive model that can be used to detect lung cancer. The purpose of this paper is to address the issues indicated above. Our strategy takes into account these significant difficulties by virtue of the way the model was developed and constructed. The architecture of the suggested multi-path CNN is comparable to

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8297*
*Article Received:* 06 July 2023 *Revised:* 28 August 2023 *Accepted:* 12 September 2023

that of the conventional CNN; nevertheless, there are structural differences between the two. It contains three different courses, and each of these three different paths is designed to mimic distant dependencies (pixels that are nearby). In general, our architecture is created in a cohesive shape so that it can deal with the numerous distant connections. The topic of unequal class representation is also addressed. The following is a condensed version of our work:

- We provide an original automatic multi-path architecture that can be utilised by CNN to detect lung cancer. It is developed to learn the varied size and shape characteristics of cancerous and non-cancerous nodules.
- We implement a technical strategy that includes a retraining phase so that we can address the critical issue of class imbalance. The contribution is broken down into its component parts in Section 3.1.
- In order to polish the model performance and improve its correctness, we additionally derived three routes by concatenation (and averaging).

The remaining parts of this work are structured as described below. In the second section, a quick review of relevant works is presented. In Section 3, we will present the dataset that was used as well as the pre-processing (segmentation of the lung, detecting the suspicious nodule and training U-Net). The proposed multi-path CNN is broken down into its component parts and detailed explanations of each path are provided in Section 4. The experimental results and discussions are presented in an organised manner in Section 5, and the study is brought to a close in Section 6.

## II. Related work

An automatic lung cancer detection algorithm can help a specialist better understand medical images, which enables medical image analysis to be performed with a higher level of sensitivity and specificity—something that is extremely important for patients. In[18], identification of lung cancer has significantly increased in last several years.

In general, there are many different techniques to image detection, particularly those that are devoted to lung cancer. These image detection methods can be grouped into two categories: those that are based on handmade models and those that are based on discriminative models. The painstakingly built models are founded on in-depth knowledge of the subject matter, specifically the existence of malignant and non-cancerous nodules. The appearance of the nodules is difficult to characterise, and the currently available handcrafted models typically misscreen as non-cancerous or cancerous nodules, which deviate from the normal one [3,12]. On the other hand, in contrast to models that are handmade, discriminative approaches make relatively minimal use of prior knowledge regarding the anatomy of the lungs. It concentrates the majority of its efforts on the extraction of low-level features, explicitly modelling the link that exists between these labels and features of the value that is delivered.

The more current discriminative models that were developed to identify a picture are presented here. The majority of these methods are geared at the detection of two-dimensional images. In the first phase of the Faster-RCNN algorithm [25], several proposed bounding boxes are generated, and the second phase involves the computation of a class decision. In more recent methods, there is only one stage, in which class probabilities and bounding boxes are instantly predicted [24]. In general, one-stage procedures are more efficient, although two-stage ones are superior in terms of accuracy. In the case of the detection of objects belonging to a single class, the second stage of the approaches consisting of two stages is no longer desired, and as a result, the approaches degenerate into methods consisting of a single stage.

An expansion of image detection using CNN from two dimensions to three dimensions is also being studied. Taking a CT image as an example, researchers [2],[5],[9] have used 3D CNN with multi-stage CNN and have produced relatively encouraging results. When it comes to our day-to-day work, this is a fairly common difficulty that arises when attempting to diagnose lung cancer. In addition, these methods do not take into consideration the issues that were discussed earlier in this article, whereas this paper addresses the issues that were discussed earlier.

## III. Dataset and Pre-processing

### 3.1 Dataset

KDSB-2017 competition served as the basis for our model's training (KDSB 2017). The dataset is in DICOM format, which includes patient data (as shown in Chart 1). It includes labelled information for 2101 patients, with a label of 0 indicating that the patient does not have cancer and a label of 1 indicating that the patient does have cancer. A CT scan will typically consist of between 100 and 400 images, each of which is a two-dimensional axial slice. The quantity of images that make up the CT scan will vary from patient to patient. These images of the 2D axial slice are used to train the model after some preprocessing has been performed on them.

We decided to use LUNA16 Challenge dataset because the Kaggle dataset on its own was not adequate to correctly categorise the validation set. The LUNA16 dataset includes 888 entries. The CT images that are included for each patient in the

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8297*
*Article Received: 06 July 2023 Revised: 28 August 2023 Accepted: 12 September 2023*

LUNA16 dataset are also 512 by 512 pixels, just like those that are included for the KDSB dataset.

We split the dataset into 2 parts which are validation set and training set of size 178 and 710 repectivily. We are able to segment potentially suspicious nodules on KDSB CT scan pictures with the assistance of the trained U-Net (which means U-Net is already trained with KDSB dataset). Following the approximation of the suspicious nodules regions, the resulting images are then separated into a test, validation and trained set so that our suggested architecture may be trained using them. Figure 6 presents the segmented results of U-examination Net's of the suspicious nodular samples.

The information that was collected by the KDSB in 2017 had a significant statistical imbalance, with 70 percent of the patients being cancer-free and rest are suffering from cancer. Choosing patches that are healthy are influential on the model, which will have a negative impact on the accuracy of the training for CNN models.

In order to get around this problem, we first constructed our dataset in such a way that both labels had the same probability. We execute data augmentation [10] on images with a label of 1 and training in order to make them equally probable. This allows us to make them equivalent in terms of their likelihood. Next, we take into account the imbalanced dataset and will do the retraining of outer most layer only (i.e., we maintain the weight of remaining layers same), this time using a label distribution that is more typical of the whole. We will have the most success if we proceed in this manner.

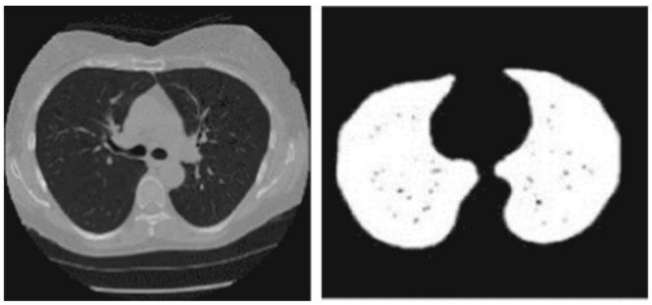

Figure 4 Input 2D CT slice (on the left), along with its threshold-based segmentation mask (right).

### 3.2 Data Pre-processing

#### 3.2.1 Lung Segmentation

In a CT scan of a lung, in addition to the lung itself, images of other tissues, such as bone, air, blood, and water, are also included. These compounds are not significant; yet, the fact that they are present has an effect on how well the model can characterise the nodules; hence, we have to rule them out.

The Hounsfield unit (HU) is the quantitative scale that is used for quantifying radio-density. Each CT scan of KDSB consists of numerous 2D axial slices with pixel values ranging from -1024 to 3071. These values correspond to the Hounsfield unit. Figure 5 depicts the distribution of pixel HU across various axial slices for the example patient photos that were used. The radio-densities, measured in HU, that are typical for the various materials seen in CT scans of the lung are listed in Chart 3.
Some methods of segmentation remove the mask of the lung in order to exclude such tissues, while leaving the detection stage open for all other types of tissue. The methods of thresholding (Alakwaa et al. 2017), clustering (Rao et al. 2016), watershed (Ronneberger et al. 2015), and k-means are frequently utilised by researchers. [Citation needed] (Gurcan et al. 2002). We filtered the 2D image with a Gaussian filter using thresholding, and then we normalised the pixel value to be between 0 and 1 by setting the threshold to a value of -600. This was done for each 2D slice. Figure 4 displays a lung segmentation mask that was constructed by thresholding, together with the original 2D CT slice that was taken of a sample patient.
We convert the image from HU to UINT8 as part of the process of getting the data ready for the network. After that, the raw data are trimmed within the range of
[-1024, 3071], and then they are linearly shifted to [0, 255]. (Fig. 5).

#### 3.2.2 Nodule Detection

In accordance with [22], we segregate suspicious nodules by training hybrid v.U-Net (we adjust some parameters of U-Net, such as model depth, kernel and input size) on LUNA 16 data. Chart 2 contains information on the changed U-parameter Net's settings.
During the training phase, the modified model receives an input of size 256 by 256 slices of 2D CT data, and the labels that correspond to those slices are generated by masking 256 by 256, with nodule pixels having a value of 1 and the rest having a value of 0.
An image with the same dimensions as an input is produced as a result of running this model. The likelihood that a given pixel is part of a nodule is represented by a number between 0 and 1 and is assigned to each pixel in the output. In order to make use of this information, the slice that belongs to label 1 of the softmax on the last U-Net layer is taken.
After that, the segmentation of the KDSB CT scan slices is carried out using the trained U-Net. These candidates vary in both size and shape, and because of this, we categorised them as test, validation and train sets in order to train the multi path Convolution Neural Network. (Chart 3)

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8297*
*Article Received: 06 July 2023 Revised: 28 August 2023 Accepted: 12 September 2023*

| Patient Id. | Cancer label |
|---|---|
| 0015ceb851d7251b8f399e39779d1e7d | 1 |
| 0030a160d58723ff36d73f41b170ec21 | 0 |
| 003f41c78e6acfa92430a057ac0b306e | 0 |
| 006b96310a37b36cccb2ab48d10b49a3 | 1 |
| 008464bb8521d09a42985dd8add3d0d2 | 1 |
| 0092c13f9e00a3717fdc940641f00015 | 0 |
| 00986bebc45e12038ef0ce3e9962b51a | 0 |
| 00cba091fa4ad62cc3200a657aeb957e | 0 |
| 00edff4f51a893d80dae2d42a7f45ad1 | 1 |
| 0121c2845f2b7df060945b072b2515d7 | 0 |
| 013395589c01aa01f8df81d80fb0e2b8 | 0 |
| 01de8323fa065a8963533c4a86f2f6c1 | 0 |
| 01e349d34c06410e1da273add27be25c | 0 |
| 01f1140c8e951e2a921b61c9a7e782c2 | 0 |
| .... | .... |
| .... | .... |

Chart 1 Some of the data from the KDSB 2017.

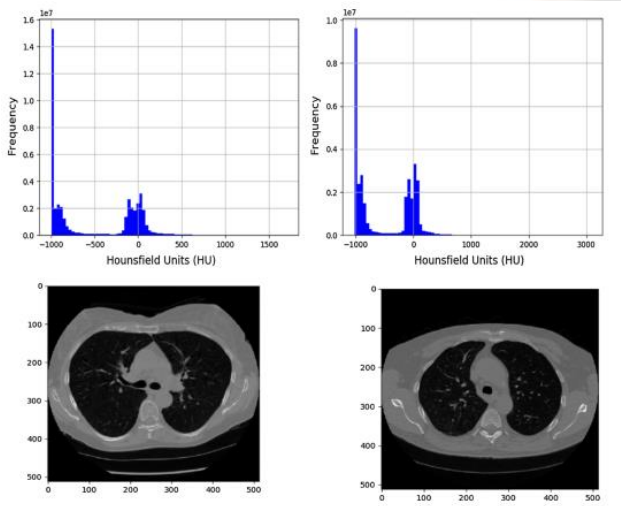


Figure 5 HU pixel distribution as well as example images obtained at various axial slices.

| Type | k. size | #ker. | Input |
|---|---|---|---|
| conv 1R | 3 × 3 | 32 | 256 × 256 × 1 |
| conv 1L | 3 × 3 | 32 | 256 × 256 × 32 |
| max p. | 2 × 2, stride 2 | | 256 × 256 × 32 |
| conv 2R | 3 × 3 | 80 | 128 × 128 × 32 |
| conv 2L | 3 × 3 | 80 | 128 × 128 × 80 |
| max p. | 2 × 2, stride 2 | | 128 × 128 × 80 |
| conv 3R | 3 × 3 | 160 | 64 × 64 × 80 |
| conv 3L | 3 × 3 | 160 | 64 × 64 × 160 |
| max p. | 2 × 2, stride 2 | | 64 × 64 × 160 |
| conv 4R | 3 × 3 | 320 | 32 × 32 × 160 |
| conv 4L | 3 × 3 | 320 | 32 × 32 × 320 |
| up conv 4R | 2 × 2 | | 32 × 32 × 320 |
| concat | conv4R, conv3R | | 64 × 64 × 320 |
| conv 5R | 3 × 3 | 160 | 64 × 64 × 480 |
| conv 5L | 3 × 3 | 160 | 64 × 64 × 160 |
| Up conv 5R | 2 × 2 | | 64 × 64 × 160 |
| concat | conv5R, conv2R | | 128 × 128 × 160 |
| conv 6R | 3 × 3 | 180 | 128 × 128 × 240 |
| conv 6L | 3 × 3 | 180 | 128 × 128 × 80 |
| up conv 6R | 2 × 2 | | 128 × 128 × 80 |
| concat | conv6R, conv1R | | 256 × 256 × 80 |
| conv 7R | 3 × 3 | 32 | 256 × 256 × 112 |
| conv 7L | 3 × 3 | 32 | 256 × 256 × 32 |
| conv8 | 3 × 3 | 3 | 256 × 256 × 32 |
| output | | | 256 × 256 × 2 |

Chart 2 Parameters for U-Net

| Substance | Radio density (HU) |
|---|---|
| Lung tissue | − 500 |
| Water and blood | 0 |
| Bone | 700 |
| CSF | 15 |
| Kidney | 30 |
| Liver | +40 to +60 |

Chart 3 Radio densities, measured in HU[11]

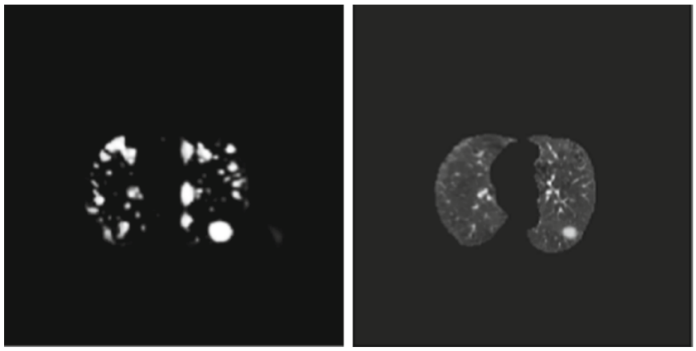

Figure 6 U-Net imaging of potentially malignant nodule samples

## IV. Proposed Mp- Convolution Neural Network method

### 4.1 Multi Path Convolution Neural Network architecture

Currently, CNN's workings points to a rather straightforward architecture consisting of one pile including multiple convolutional layers. This architecture is the one that is used the most frequently in actual implementations. However, one may create alternative architectures that have a chance of being suited for the problem that is currently being considered.

The variety in nodule size, shape, and the contextual factors that surround them are some of the primary obstacles that must be overcome in order to successfully diagnose lung cancer using CNN, as we have already indicated. Because of these obstacles, the currently available CNN-based lung cancer detection models have an issue with their level of accuracy. The receptive field size's ability to exert an influence on the modelling of distant dependencies is one of the factors contributing to the aforementioned difficulties.

In this study, we construct a Convolution Neural Network that has convolutional layers with multiple path. This means that the path takes into consideration different receptive field sizes, ranging from small to medium to big. These routes are referred to as 1$^{st}$ , 2$^{nd}$ , and 3$^{rd}$ path, respectively.

The first, second and third path has a receptive field of size 3by3, 5by5, and 7by7 respectively. One might wonder at this juncture if a receptive field size that is 3 by 3 or 7 by 7 would be more suitable. The answer is not objective; rather, it is contingent upon the activity. For study purpose, we will

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8297*
*Article Received: 06 July 2023 Revised: 28 August 2023 Accepted: 12 September 2023*

concentrate on certain receptive field sizes. In addition in the design we concatenate several feature maps that originate from last layer which is a convolutional layers. These feature maps assist the model in improving its ability to forecast. The forecast of a pixel's label being favourably influenced in a positive way by the following factors served as the impetus for the development of this architecture design.

- information regarding the surrounding neighbourhood pixels and the visual features of the area
- contextual information.

Additionally, the concatenation layer assists in determining significant features from the three pathways. This assists in the modelling of long-term and short-term relationships, which means that features located in the same areas are represented in a more accurate manner. Figure 7 is an illustration of the entire building, and Chart 4 lists the parameters of the architecture. This architecture is referred recognised as a multi-path Convolution Neural Network.

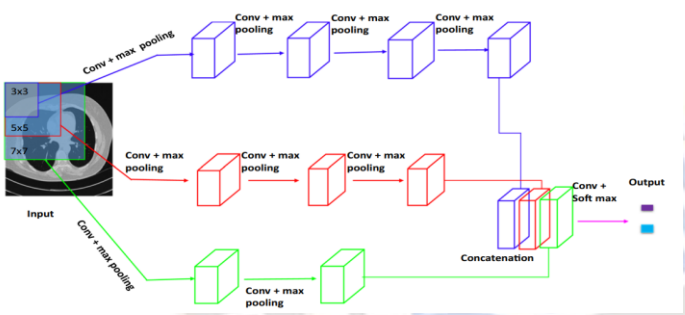


Figure 7 Architecture of a CNN with several pathways.

**Table 4** Multi-path CNN parameters setup

| Type | k. size | #ker. | act. | Input |
|---|---|---|---|---|
| *First path* | | | | |
| conv | 3 × 3 | 64 | relu | 128 × 128 × 1 |
| max p. | 2 × 2 | | | 64 × 126 × 126 |
| conv | 3 × 3 | 64 | relu | 64 × 125 × 125 |
| max p. | 2 × 2 | | | 64 × 123 × 123 |
| conv | 3 × 3 | 64 | relu | 64 × 122 × 122 |
| max p. | 2 × 2 | | | 64 × 120 × 120 |
| conv | 3 × 3 | 64 | relu | 64 × 119 × 119 |
| max p. | 4 × 4 | | | 64 × 117 × 117 |
| *Second path* | | | | |
| conv | 5 × 5 | 64 | relu | 128 × 128 × 1 |
| max p. | 2 × 2 | | | 64 × 124 × 124 |
| conv | 5 × 5 | 64 | relu | 64 × 123 × 123 |
| max p. | 2 × 2 | | | 64 × 119 × 119 |
| conv | 5 × 5 | 64 | relu | 64 × 118 × 118 |
| max p. | 2 × 2 | | | 64 × 115 × 115 |
| *Third path* | | | | |
| conv | 7 × 7 | 64 | relu | 128 × 128 × 1 |
| max p. | 2 × 2 | | | 64 × 122 × 122 |
| conv | 7 × 7 | 64 | relu | 64 × 121 × 121 |
| max p. | 2 × 2 | | | 64 × 115 × 115 |
| *Concatenation* | | | | |
| conv | 114 × 114 | 192 | s.max | 192 × 114 × 114 |
| Output | | | | 2 × 1 × 1 |

Chart 4 Mp- CNN setup for parameters

### 4.1.1 First path

This CNN path is the path with a reduced 3x3 receptive field size, and its purpose is to simulate the short-range dependencies between neighbouring pixels. It consists of four convolutional layers, the first three of which are followed by maximum pooling on a scale of 2x2 while the fourth layer is followed by maximum pooling on a scale of 4x4. The maximum pooling size of 4 by 4 is used because we want the spatial size of feature maps to be comparable to that of other pathways. It makes it possible for us to effortlessly combine many feature maps. As input, the first convolutional layer receives pre-processed 2D slices of lung pictures with the dimensions 128 by 128 by 1, and it filters the data using 64 filters with the dimensions 3 by 3. After max-pooling, the output of the first layer is used as input for the second layer's convolution operation. This operation uses 64 kernels, each of which is 3 by 3, and applies them to feature maps that are 64 by 125 by 125. In a similar manner, the third layer, after the max pooling step, uses the output of the second layer as an input and performs convolution using 64 filters of size 3 on feature maps that are 64 by122 by 122. Last but not least, the output of the third layer with a size of 64 by119 by 119 is used as the input for the fourth layer of this path, which then performs convolution using 64 filters with a size of 33. After 4x4 maximum pooling, it is abundantly clear that this layer has a 64 by114 by 114 output size. After that, the output of the third layer is combined with the output of the final layer of the remaining two routes via a process known as concatenation.

### 4.1.2 Second path

This is a path that has a medium receptive field size of 5 by 5, and its purpose is to mimic the average short-range and long-range dependencies of neighbouring pixels. This path is more effective at modelling the lung nodules that are medium in size and have a variety of shapes. On the other hand, the larger and smaller nodules are more effectively modelled by using a larger receptive field and a smaller receptive field, respectively. Concatenation of the results from the three different paths improves performance. The path consists of three convolutional layers, and after each of those layers comes a maximum pooling of 2 by 2. In a manner analogous to the first path, the first layer of this path accepts as an input pre-processed two-dimensional slices of lung cancer, and then applies sixty-four filters of size five by five to them. Following an application of max poling to the output of the layer that came before it, that output is then utilised as the input for the second convolutional layer. After applying 64 filters of size 55 to the input data, the system generates feature maps with a size of 64 by 119 by 119 (that is, 64 feature maps each with a spatial size of 119x119), and this happens after max pooling. The input data has a size of 64x123x123. The outputs of the previous layer are then taken into account by the subsequent layer, which is referred to as layer three, and another convolution operation with 64 filters of size 5x5 is performed on them.

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8297*
*Article Received:* 06 July 2023 *Revised:* 28 August 2023 *Accepted:* 12 September 2023

#### 4.1.3 Third path

Long-range dependencies are the focus of this route's investigation. It consists of two layers of convolutional processing, followed by maximum pooling on a size 22 grid. The first layer has inputs measuring 128 by 128 by 1, and it applies convolution to this input using 64 filters measuring 7 by 7, which results in outputs measuring 64 by 122 by 122 feature maps. After being subjected to the maximum pooling, it is then used as the input to the second layer. The identical action is carried out by the second layer, but this time with 64 filters that are each 7 by 7, and the output is 64 by 115 by 115. After the 2x2 max pooling step, its output is combined with the output of the final convolutional layer of the other two routes during the concatenation step.

#### 4.1.4 Concatenation

The output of the final convolution layer of each path is concatenated, and then this concatenation is convolved with 192 filters that have a size of 114 by 114, and then a soft-max function is used to predict the input. This helps to average out the effect that the size of the receptive field has on the output.
In most cases, the application of a multi-path model with concatenation is the most effective way to both harness the efficacy of CNNs and describe the relationships between neighbouring pixels in a more accurate manner for the detection task.
This kind of joint detection model requires far more processing power than a feedforward pass through a CNN would. When deciding whether or not to incorporate automatic lung cancer detection into our day-to-day operations, this is a very important step that needs to be taken into consideration.
In this section, we will discuss the CNN model, which simultaneously models remote relationships while maintaining the efficiency of CNNs. Since we would like the final prediction to be biased by the model's views about the labels of the neighbourhoods, we proposed adding the output probabilities of a third and second CNN as additional inputs to the layer of the first CNN. This was done because we would like the final prediction to be biased by the model's views about the labels of the neighbourhoods. Specifically, we accomplish this by chaining together many convolutional layers. In this particular instance, we join the final convolutional output layers from each path of the CNN together.

### 4.2 Training

Using the output of the CNN as a model for the distribution of detection labels, a popular training strategy is to either maximise the probability of all labels in the training set or minimise the negative log-probability. This is done by expressing the output of the CNN as a model (eq.6)

$$-\log P(Y|X) = \sum_{i,j} -\log(Y_{i,j}|X)$$

every lung slice that was given a label. In order to accomplish this, we make use of the stochastic gradient method, which involves selecting labels Yi,j multiple times at a random subset of patches within each lung, determining the mean negative log-probabilities for this mini batch of patches, and performing a gradient descent step on the CNN's parameters (i.e., the weight at all layers). Only a select few of the patches have had their versions updated by us. Because of this, we are able to avoid having to process the entire slice of the lung for each update while still providing sufficient updates for learning. This is a win-win situation. To be more specific, we implement this strategy by constructing a dataset consisting of mini batches of smaller slices of lungs image patches that are associated with the equivalent detection label as the target. We make use of the momentum strategy to improve the performance of our model because it has proven to be effective in the past (Krizhevsky et al. 2012). This tactic was utilised by us in (Eq.7 & Eq.8)

$$v_{i+1} = \eta * v_i - \gamma * \nabla W_i$$

$$W_{i+1} = W_i + v_{i+1}$$

where Wi stands for the parameters of the CNN at iteration I Wi is the gradient of the loss function at Wi, vi represents the integrated velocity with a starting value of zero, stands for the learning rate, and is the momentum coefficient. The beginning momentum will be equal to 0.5, and the ultimate momentum, will be equal to 0.9. Additionally, the learning rate of 0.005 is decaying exponentially with a factor of 0.1 for the decay.

## V. Experimental results and discussion

### 5.1 Experiment setup

We implement an experiment on one of the deep learning libraries known as tensor flow (Martin 2016) while using Kera as the backend. It is compatible with graphics processing units (GPUs), which significantly speed up the calculation of deep learning algorithms. Chart 4 provides information on a variety of our model's hyper-parameters, including filter size, maximum pooling size, architectural depth, and others. These particular hyper-parameters were selected because they allowed our models to achieve the highest possible scores on the validation set. We used a stride 1 for both the convolutional and the max-pooling layers of the model. This helps us to maintain the accuracy down to each individual pixel. According to the findings of our experiment, including additional convolutional and max-pooling layers on the third route does not improve the results. Furthermore, we discovered that including additional

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8297*
*Article Received: 06 July 2023 Revised: 28 August 2023 Accepted: 12 September 2023*

features did not result in any meaningful improvement. Because of this, we keep the layer depth of our model to a minimum. We randomly initialise the weights using the uniform distribution U(-0.005, -0.005), and with the exception of the softmax layer, where we initialise the weights to be the log of the label, biases are set to zero. Weights are initialised using this method.

### 5.2 Evaluation metrices

We compute specificity, S, recall, R, and accuracy, A, in equations 9, 10, and 11 respectively to determine how well our model performs:

$$S = \frac{tn}{tn + fp}$$

$$R = \frac{tp}{tp + fn}$$

$$A = \frac{tp + tn}{tp + fp + fn + tn}$$

where the number of true positives, the number of false positives, the number of false negatives, and the number of true negatives are denoted by the symbols tp, fp, fn, and tn, respectively. The results of these equations can be converted into a percentage by multiplying them by 100.

### 5.3 The Pre-process Stage

In the beginning, we trained our model using the unprocessed KDSB dataset, which did not include any nodule detection or lung segmentation. After that, we trained the mpCNN while excluding tissues like water, air, and bone; nonetheless, the results were still unsatisfactory. As a result, in order to enhance the performance of our model, we carried out segmentation and detection on the raw images. Following the application of thresholding to segment the lung, U-Net was utilised to search for potentially malignant nodules. The trained U-Net model was applied to the data in order to identify the region of the lung that would contain suspicious nodules (cancer or non-cancer nodules, respectively) (KDSB dataset was used). The experimental findings suggest that the performance of the multi-path convolutional neural network (mp-CNN) can be improved by using Thresholding and U-Net before directly deploying the network. The results of our model accuracy are displayed in Chart 5 and Figure 8 respectively, both with and without the pre-processing stage. The use of the pre-processing stage led to an improvement in the accuracy of the mp-CNN, as can be seen in the Chart. Additionally, one can see from Fig. 8 that the accuracy of the mp-CNN model improves with an increase in the number of training steps that are carried out when the pre-processing stage is carried out. mp-CNN that was trained with the pre-processing stage has an average accuracy that is 3.6 percentage points higher than mp-CNN that was not trained with the aforementioned pre-processing stage. Based on the results of the experiments, we discovered that the joint segmentation and detection pre-process that we utilised resulted in an improvement in the performance of the suggested mp-CNN.

| Model | Accuracy | |
|---|---|---|
| | With preprocessing | Without preprocessing |
| mp-CNNs | 0.878 | 0.673 |
| f-CNNs | 0.871 | 0.643 |

Chart 5 Accuracy of pre-processing stage

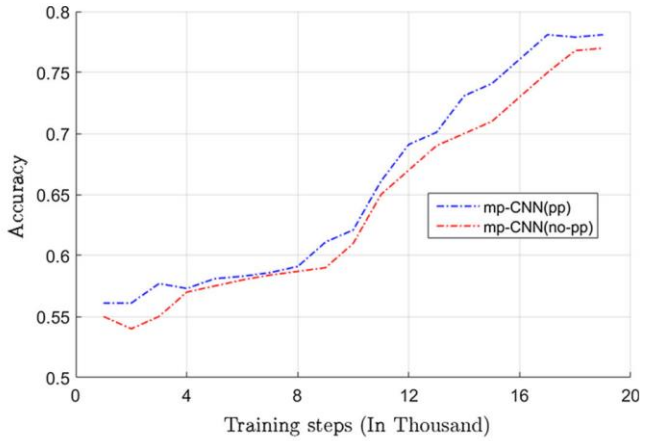


Figure 8 multiple path Convolution Neural Network model results.

### 5.4 The Multipath Architecture

As was indicated before, in contrast to the conventional CNN, the mp-CNN is comprised of three distinct channels. These pathways have been developed to more accurately simulate the local and global dependencies of the pixels that are adjacent to them. The first and second paths concentrate more on the particulars (local dependencies), whereas the third path is concerned with the information that is contextual (global dependency). The model operates more effectively as a result of our attention on both the background and the details. In order to have a clearer understanding of how the joint training of these pathways improves performance, we will first discuss the findings of each route individually, and then we will discuss the results of averaging the output of each pathway after it has been trained independently. Due to the fact that we additionally have to cope with the class imbalance of the labelled dataset, we retrain the model using the strategy that is outlined in Section 3.1. In order to understand how the two different phases of training interact with one another, we must first consider the outcomes of both of the training sessions individually, namely the outcomes of the initial training phase as well as the

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8297*
*Article Received: 06 July 2023 Revised: 28 August 2023 Accepted: 12 September 2023*

retraining phase. The CNN models that only contain the first path, the second path, and the third path are referred to by us as the f-CNN, s-CNN, and t-CNN models, respectively. Additionally, the CNN model that takes the average output of all three paths is referred to as a-CNN, while the CNN model that takes into account all possible paths is referred to as mp-CNN. A superscript's' indicates that the second phase of training has begun.

Chart 6 presents the quantitative findings that resulted from these changes. The results of the mp-CNN and f-CNN with single and retraining phase are included in the Chart, as well as the results of the s-CNN, t-CNN, and a-CNN with only retraining phase. It is clear from looking at the Chart that the initial option, which consisted of only one training phase, came in lowest place. However, adding a retraining phase to the model resulted in a significant improvement to its performance. This demonstrates the following:

- When compared to training each pathway separately and then averaging the results, collaborative training of all pathways yields superior performance.
- The retraining step contributes significantly to the overall improvement of the performance of the multi path Convolution Neural Network.

Indeed, the performance of a-CNNs is worse than that of f-CNNs due to our belief that t-CNN performance is poor. The mp-CNNs model is the one that performs the best overall.

Performance of our model are compared to the most recent and cutting-edge methodologies that have been published for detecting lung cancer demonstrated in Chart7. The Chart demonstrates that mp-CNNs fared better than any of the others. Results that our model has produced are comparable to those that were produced by the model that won the KBSB 2017.

| Methods | Recall | Specificity | Accuracy |
|---|---|---|---|
| mp-CNN$^s$ | 0.874 | 0.891 | 0.878 |
| f-CNN$^s$ | 0.862 | 0.900 | 0.871 |
| a-CNN$^s$ | 0.852 | 0.861 | 0.854 |
| s-CNN$^s$ | 0.827 | 0.831 | 0.828 |
| mp-CNN | 0.786 | 0.792 | 0.788 |
| t-CNN$^s$ | 0.764 | 0.772 | 0.766 |
| f-CNN | 0.670 | 0.693 | 0.676 |

Chart 6 Performance of multi path-CNN model and its several iterations

| Methods | Recall | Specificity | Accuracy |
|---|---|---|---|
| mp-CNN$^s$ | 0.874 | 0.891 | 0.878 |
| Kingsley et al. (2017) | 0.858 | 0.881 | 0.864 |
| Alakwaa et al. (2017) | 0.840 | 0.841 | 0.840 |
| Rao et al. (2016) | 0.815 | 0.801 | 0.811 |
| Huang et al. (2017) | 0.724 | 0.742 | 0.728 |

Chart 7 On the KDSB 2017 test set, a comparison of our models to the most advanced approaches currently available

### 5.5 Size and shape Effect

We assessed the performance of our model by selecting 50 large cancer nodules, 50 small cancer nodules, all of which had different forms from the KDSB 2017 challenge. The results that were obtained are presented in Chart 8. As can be seen from the Chart, our mp-CNNs model properly predicted 98 percent of the 50 larger cancer nodules that were picked, however it only accurately predicted 94 percent of the 50 smaller cancer nodules that were selected. In addition, the performance of our model is superior when applied to larger nodules as opposed to smaller nodules. Our mp-CNNs offers superior performance outcomes when compared to certain other approaches for detecting lung cancer [1][6][9][22]. These studies were published in 2017. In general, we have demonstrated that the proposed multi-parameter or path convolutional neural network diagnosis better as well as addresses the problem of variability.

| Methods | No. of cancer | Correctly predicted | Wrongly predicted |
|---|---|---|---|
| mp-CNN$^s$ | Larger (50) | 49 (98%) | 1 (2%) |
| | Smaller (50) | 47 (94%) | 3 (6%) |
| Kingsley et al. (2017) | Larger (50) | 44 (88%) | 6 (12%) |
| | Smaller (50) | 43 (86%) | 7 (14%) |
| Alakwaa et al. (2017) | Larger (50) | 43 (86%) | 7 (14%) |
| | Smaller (50) | 40 (80%) | 10 (20%) |
| Rao et al. (2016) | Larger (50) | 41 (82%) | 9 (18%) |
| | Smaller (50) | 37 (74%) | 13 (26%) |
| Huang et al. (2017) | Larger (50) | 40 (80%) | 10 (20%) |
| | Smaller (50) | 37 (74%) | 13 (26%) |

Chart 8 The performance of our model within the size and shape context, as well as its comparison with alternative techniques

## VI. Conclusion

In this research, we presented a strategy for the automatic identification of lung cancer that makes use of deep convolutional neural networks. We looked at a number of different architectures and examined how they affected the overall detection performance. The result that was conducted on Kaggle Dataset Science Bowl 2017 validates that we were successful in improving upon the existing methods in terms of specificity, sensitivity, and accuracy using our best model. Because both long and short range dependencies are modelled in an accurate manner, we are able to achieve a higher level of performance with the assistance of an innovative multi-pathway architecture. This architecture is able to model the lung nodules in a variety of sizes and shapes. The training technique

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8297*
*Article Received: 06 July 2023 Revised: 28 August 2023 Accepted: 12 September 2023*

consists of two stages, which enabled us to effectively and efficiently train CNNs despite the imbalance in the class distribution of labels. The speed of the resulting detection is attributed to the characteristics of Convolution Neural Network models and the GPU machine. In general, as compared to existing methods for detecting lung cancer using Convolution Neural Network models, our method is more flexible. It is also more able to manage variations in form and size of nodules, and it can better label class balancing.

## References

[1] Alakwaa, W., Nassef, M., & Badr, A. (2017). Deep convolutional neural networks for lung cancer detection. IJACSA, 8(8), 403–417.

[2] Alvarez, J. M., Gevers, T., LeCun, Y., & Lopez, A. M. (2012). Road scene segmentation from a single image. In Proceedings of the 12th European conference on computer vision (pp. 376–389).

[3] Clark, M., Hall, L., Goldgof, D., Velthuizen, R. P., Murtagh, F., & Silbiger, M. L. (1998). Automatic tumor segmentation using knowledge-based clustering. IEEE Transactions on Medical Imaging, 17, 187–201.

[4] Gurcan, M. N., Sahiner, B., Petrick, N., Chan, H. P., Kazerooni, E. A., Cascade, P. N., et al. (2002). Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computeraided diagnosis system. Medical Physics, 29(11), 25522558.

[5] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., et al. (2017). Brain tumor segmentation with deep neural networks. Medical Image Analysis, 35, 18–31.

[6] Huang, X., Shan, J., & Vaidya, V. (2017). Lung nodules detection in CT using 3D Convolutional neural networks. In International Symposium on IEEE (pp. 379–383).

[7] Jarrett, K., et al., (2009). What is the best multi-stage architecture for object recognition? In Proceedings 12th International Conference on IEEE Computer Vision (pp. 2146–2153).

[8] Kaggle Data Science Bowl (KDSB). (2017). https://www.kaggle.com/c/data-science-bowl-2017/data. Accessed 25 Mar 2018.

[9] Kingsley, K., Mathieu, R., Gaurav, M., Huiling, C., Jie, L., Babar, N., et al. (2017). Deep learning for lung cancer detection: Tackling the Kaggle data science bowl 2017 challenge, arXiv preprint arXiv:1705. 09435v1.

[10] Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). Image Net classification with deep convolutional neural networks. In NIPS.

[11] Lepor, H. (2000). Prostatic diseases. Philadelphia: W.B Saunders Company.

[12] Lin, D. T., & Yan, C. R. (2002). Lung nodules identification rules extraction with neural fuzzy network.International Conference on in Neural Information Processing, 4, 20492053.

[13] Liu, W., Anguelov, D., Erhan, D., Szegedy, C. Reed, S., Fu, C.-Y., et al. (2016). SSD: single shot multi box detector. In European Conference on Computer Vision (pp. 21–37).

[14] Liu, Y., Gadepalli, K., Norouzi, M., Dahl, G. E., Kohlberger, T., Boyko, A., et al. (2017). Detecting cancer metastases on giga pixel pathology images, arXiv preprint arXiv:1703.02442.

[15] Lung Nodule Analysis (LUNA). (2016). https://luna16.grand-challenge.org/. Accessed 2 Mar 2018.

[16] Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier non linearity improve neural network acoustic models. In Proceedings of ICML (Vol. 30).

[17] Martin, A. et al. (2016). TensorFlow: Large-scale machine learning on heterogeneous distributed systems, arXiv preprint arXiv:1603.04467.

[18] Maaroof Maaroof, M. A. ., & Naimi, S. . (2023). Developing Adaptive Project Construction Cost Control Using Multi-Nonlinear Regression Engineering Techniques. International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 579 –. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2729.

[19] Muhammad, R., Saeeda, N., & Ahmad, Z. (2017). Deep learning for medical image processing: Overview, challenges and future, arXiv preprint arXiv:1704.06825.

[20] Patz, E. F., Jr., Pinsky, P., Gatsonis, C., et al. (2014). Over diagnosis in low-dose computed tomography screening for lung cancer. Internal Medicine, 174(2), 269–274.

[21] Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Transactions on Medical Imaging,

[22] Ramachandran, P., Zoph, B., & Le, Q. V. (2017). Searching for activation functions, arXiv preprint arXiv: 1710.05941v2 [cs.NE].

[23] Kwame Boateng, Machine Learning-based Object Detection and Recognition in Autonomous Systems , Machine Learning Applications Conference Proceedings, Vol 3 2023.

[24] Rao, P., Pereira, N. A., & Srinivasan, R. (2016). Convolutional neural networks for lung cancer screening in computed tomography (CT) scans. In International conference on contemporary computing and informatics (pp. 489–493).

[25] Rebecca, L. S., Kimberly, D. M., & Ahmedin, J. (2018). Cancer statistics, 2018. CA: A Cancer Journal for Clinicians. https://doi.org/10.3322/caac.21442.

[26] Redmon, J., & Farhadi, A. (2016). Yolo: Better, faster, stronger, arXiv preprint arXiv:1612.08242.

[27] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in Neural Information Processing Systems (pp. 91–99).

[28] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation, arXiv preprint arXiv:1505.04597.

[29] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning representations by back-propagating errors. Nature, 323(6088), 533–536.

[30] The National Lung Screening Trial Research Team. (2011). Reduced lung-cancer mortality with low-dose computed tomographic screening. New England Journal of Medicine, 365(5), 395–409.

[31] WHO. (2018). http://www.who.int/news-room/fact-sheets/detail/cancer. Accessed 26 Oct 2018.