International Journal of Multidisciplinary Innovation and Research Methodology (IJMIRM)
Volume 2, Issue 1, January-March, 2023, Available online at: https://ijmirm.com

Lung Cancer Detection and Classification using Deep Learning Techniques

Dr. Tao Tian

Nanjing University, Nanjing, China

ABSTRACT

Better health outcomes have been achieved as a result of improvements in the ability to identify lung cancer. are utilised frequently in the field of medicine to detect lung tumours at an early stage. Deep learning models such as UNET, Efficient-net, Resnet, VGG-16, and others have been used in a number of studies to improve the accuracy with which lung cancer may be detected. This paper offers an approach that combines UNET and Efficient-Net neural networks for the purpose of lung nodule segmentation and classification. The goal of the algorithm is to improve the detection performance. In order to make use of the massive volume of CT scan pictures that do not have any pathological diagnoses attached to them, an approach that is feature-extraction based and semi-supervised is applied. A feature pyramid network (FPN) with the ResNet-50 model is used for feature extraction, and a neural network classifier is used for predicting unlabeled nodules. This setup allows for semi-supervised learning to take place. The skip-connections are the primary innovation that UNET brings to the table. These connections grant the decoder access to the features that the encoder learnt at different scales, which in turn enables accurate localization of lung nodules. An effective neural network for image classification is produced by Efficient-Net by employing scaling on all three dimensions of the network (depth, width, and resolution) in conjunction with a compound coefficient that scales all of the network's dimensions consistently. This work has been evaluated on the LIDC-IDRI dataset, which is available to the public, and it performs better than the majority of the existing approaches. Issues such as a high false-positive rate, small nodules, and a wide range of non-uniform longitudinal data are some of the problems that the suggested algorithm intends to address. The findings of the experiments indicate that this model has a better level of accuracy than earlier research, coming in at 91.67% of the time.

Keywords: **Cancer, CT, UNET, Efficient-Net, Feature, Accuracy.**

I. INTRODUCTION

Lung cancer is the second most lethal type of cancer in both men and women. In addition to this, the widespread spread of the coronavirus has led to an increase in the likelihood of death among lung cancer patients [1]. A lung tumour could be fatal if it is not diagnosed in a timely manner. Pulmonary nodules, also known as tiny cell growths in the lungs, can either be cancerous (also known as malignant) or noncancerous (also known as benign) [2]. Due to the high degree of similarity between early-stage cancer lung nodules and noncancerous nodules, a differential diagnosis that takes into account the locations, morphological characteristics, and clinical biomarkers of the nodules is required. According to the World Health Organisation (WHO), there were 2.21 million instances of lung cancer reported in 2020, and 1.80 million people lost their lives as a result of the disease [3].

When cells in the lung tissue continue to grow out of control, this can lead to the development of lung cancer. This leads to the development of a tumour. They can potentially expand to a wide variety of other regions of the body and interfere with normal breathing [4]. In order to identify lung malignancies at an early stage, diagnostic imaging procedures such as computed tomography, sputum cytology, chest X-rays, and magnetic resonance imaging (MRI) are utilised. In order for tumours to be found, they are first classified into one of two categories: benign tumours (those that do not cause cancer) and malignant tumours (those that do cause cancer) [5]. Image processing techniques have the potential to make manual analysis more accurate. There have been a number of studies that have proven the existence of diagnostic mistakes in clinical practises [6]. These mistakes can be traced back to a wide array of causes, which can be largely grouped into three categories: difficulties that are specific to the individual, the nodule, and the environment respectively.

A computed tomography (CT) scan will create images of the human body that are cross-sectional and very detailed. In contrast to a conventional x-ray, which only records one or two images, a computed tomography (CT) scan compiles a large number of images. A computer will eventually combine these images in order to make a slice of the part of the body that is being investigated. When compared to a CT scan, a traditional chest x-ray has a lower likelihood of detecting lung cancer [7]. In addition to this, it can show the size, form, and location of any lung tumours, as well as disclose any swollen lymph nodes that may be hiding cancer that has spread [8].

Because of breakthroughs in machine learning and deep learning in particular [9], the medical fields, specifically lesion categorization, tissue recognition, and segmentation, have experienced tremendous progress in recent years. [9] These advancements were made possible by machine learning. Deep learning is becoming increasingly prevalent in the process of analysing medical images as a result of advancements made in artificial intelligence. Deep convolutional neural networks present a strategy that can be put into practise for the purpose of obtaining characteristics that are both richer and more powerful [10]. When it comes to the diagnosis and classification of lung nodules, automated algorithms can offer a solution that is both quicker and more accurate, which ultimately leads to better patient outcomes [11].

II. REVIEW OF THE LITERATURE

A. The preparatory work

The accuracy and reliability of later studies could be jeopardised if large amounts of unprocessed data from CT scans contain noise, artefacts, and inconsistencies. These issues could be caused by the scans themselves. Amalorpavam et al. give a very extensive and systematic assessment of the many morphological processes that are currently used in digital image processing in their published work [12]. The transformation of images by means of mathematical morphology in order to standardise them for the sake of the desired analysis is the fundamental objective of this work. According to the findings of this research, a variety of techniques, including erosion, dilation, thresholding, and blurring, can be successfully implemented in the preliminary processing of pictures obtained from a CT scan. These techniques are utilised to improve the contrast between the different types of tissues, eliminate noise in the images, and eliminate any formations that are not nodules.

B. Segmentation

UNet is a semantic segmentation network that was built from the ground up on the framework of a fully convolutional neural network. Olaf et al. [13] were able to successfully simulate the convolutional network-based U-Net architecture, which was developed particularly for the purpose of the segmentation of pictures used in the field of biomedicine. This study utilises a conventional UNet design that consists of an encoder and a decoder, and it generated an average IOU (also known as "intersection over union") value of 92%.

A UNET-based technique was developed for segmenting lung nodules in CT images in a study that was conducted by Chen et al. [14]. The proposed approach involved two steps: for the first stage's initial segmentation, a pre-trained UNET model was utilised, and for the second stage's refinement, a 3D fully linked conditional random field (CRF) was used. Ronneberger and colleagues [15] created a U-Net network that was built on FCN for the purpose of segmenting lung lesions. U-Net integrates information of low and high resolution by skipping connections in the process of segmenting medical images with fuzzy boundaries. Information with a low resolution is utilised for target identification, whereas information with a high resolution is utilised for localising segmentation.

C. Extraction of Characteristics

In a feature pyramid network, a convolutional neural network (CNN) is utilised to construct a feature map by extracting features from an input image. These features are employed for classification purposes. Xiaolong Wang et al. [16] proposed

a technique for semi-supervised learning that makes use of both labelled and unlabeled data by employing a multi-level feature pyramid network (FPN). The FPN is used to extract features at multiple scales, and the method adds consistency regularisation to encourage the model to provide similar predictions on labelled and unlabeled data. These features can be extracted using the FPN. The suggested approach obtains the best reported classification error rate of 4.42% on the test set with just 4,000 labelled samples as the basis for the analysis. It also achieves an error rate of 3.57%, which is the best result to date, using 4,000 labelled samples and 50,000 unlabeled samples. This is the best result that has been achieved to date.

Feature pyramid networks with Relu Cascade were proposed by Guangrui Mu and his colleagues [17] for the purpose of CT pulmonary nodule detection. First, a detection network is trained using a small number of positive annotations (called nodules) and a large number of randomly chosen negative samples (called background). In FPN, the process of extraction is carried out independently on each feature map in order to generate feature maps at a variety of scales. Images are incorporated into each feature map. The approach that was utilised to link these networks together in a cascade is what gives the "Relu cascade" its distinctive quality.

D. Classification

Image classification is just one of the many computer vision jobs that has been revolutionised by a design of convolutional neural networks known as Efficient-Net. Agrawal et al., [18] provided a variety of deep learning based pre-trained CNN algorithms that can differentiate between benign and malignant brain tumour images. In order for them to finish the tasks, they made use of several distinct optimizers, particularly Adam, RMSprop, and stochastic gradient descent (SGD). The findings of their study showed that a well-tuned version of Alex Net may achieve particularly impressive results when faced with problems concerning medical imaging. They applied approaches for data preprocessing and augmentation with the goal of increasing the diversity of the data samples in order to reduce the amount of overfitting that occurred with the earlier models.

A three-dimensional deep convolutional neural network (also known as a 3D DCNN) with dense connections and shortcut connections was built by Hwejin Jung and colleagues [19] for the purpose of classifying lung nodules. Shortcut connections and dense connections are both effective solutions to the problem of the gradient vanishing because they make it possible for the gradient to move quickly and directly. This approach earned the highest possible CPM score, which was 0.920.

Divya et al., [20] provided approaches based on transfer learning that enhance current architecture in multi-class classification by depending on pretrained DCNN trained on ImageNet dataset. These methods were presented to improve existing architecture. The initial weights of the model are determined by transferring the pretrained weights of EfficientNetV2-B0, and the model is then fine-tuned so that it can differentiate between benign and malignant tissue samples in cancer cells and classify them.

In order to segment and categorise lung nodules in CT scan pictures, the purpose of this work is to develop and build a deep learning model that is founded on the UNET segmentation network and the Efficient-Net architecture. The work addresses a number of limitations, including the fact that improvements are needed in the detection rates of lung nodules, as the proposed models have only focused on training dataset image samples and also segmented anomalies as positive samples; an unequal number of benign and malignant lung nodules, resulting in an imbalanced dataset; an accurate diagnosis of the pathological type of lung cancer is crucial for effective treatment; and the performance of model relies heavily on the number of training dataset image samples. Additionally, the work addresses that an accurate

UNet is a semantic segmentation network that was built from the ground up on the framework of a fully convolutional neural network. The network is comprised of a total of 23 layers, which is a considerable reduction in comparison to the amount of layers present in other networks while yet preserving accuracy [21]. Extracting feature vectors from segmented images is accomplished through the usage of feature pyramid networks. For the purpose of an unlabeled nodule prediction job [22], these feature vectors are fed into a neural network classifier in the capacity of an input. In order to attain cutting-edge performance while remaining computationally inexpensive, a family of convolutional neural networks known as inexpensive-Net was developed. It is able to achieve this by achieving a balance between the network's depth, width, and resolution in order to maximise performance while staying within a set computational budget [23].

The LIDC dataset will be used to evaluate the proposed algorithm to determine whether or not it can correctly classify and identify specific characteristics. The results of the evaluation will be compared to conventional metrics such as accuracy, recall, and F1-score [24]. Accurate detection and segmentation of lung nodules is the goal of the approach that has been developed. The purpose of this project is to investigate, with the help of an ablation study, how the performance of the proposed lung nodule segmentation and classification algorithm [25] is affected by a variety of factors (hyper-parameters).

In order to investigate the possibilities of using semi-supervised learning and transfer learning to improve the performance of the proposed method on datasets with a limited number of labelled data, the goal of this project is:

[26]. Adam optimizer is introduced into the model so that the performance of the model may be stabilised and the influence of errors can be reduced [27].

III. PROPOSED METHODOLOGY

A. The Lung Image Database Consortium (abbreviated as LIDC-IDRI)

The LIDC dataset consists of 1018 cases, each of which was labelled in a joint effort by four radiologists. The National Cancer Institute was the driving force behind the collecting of these cases. The LIDC-IDRI dataset in the Cancer Imaging Archive (TCIA) contains 1018 clinical chest CT scans with lung nodules gathered from seven different institutions. These scans were published by Armato et al., 2011, Armato III et al., 2015b, and Clark et al., 2013. An attached XML file for each CT image that was examined for annotation by up to four radiologists contains specific information regarding the locations of the nodules as well as the nine sematic qualities of subtlety, sphericity, internal structure, margin, lobulation, and spiculation (Qin et al., 2019). Malignancy was also included among the nine sematic features. It is a database that is applied extensively and is devoted to the improvement of techniques that are currently being used for lung nodule segmentation. There are three distinct types of nodules, which can be broken down as follows: non-nodules (size >=3mm), nodules (size 3mm), and nodules (size >=3mm). Every one of the images will be in the DICOM format, and their dimensions will be 512 pixels on each side. The approximately 12000 labelled CT scan images and the 3000 unlabelled CT scan images that are used are selected based on the malignancy details that are provided in the metadata csv and xml files.

B. Architecture of the System

This section discusses the architecture, description, and algorithmic components of the many subsystems that will make up the proposed system. The suggested system's design is depicted in greater depth in figure 1, which may be found below. An approach has been proposed that uses the UNET model to segment the lungs in photos and uses Efficient-net architecture to classify the images. This algorithm can be found here. A previously unlabeled dataset is given a label by employing FPN with Resnet-50 serving as the backbone. A neural network classifier receives the extracted features from tagged pictures using FPN. These extracted features then serve as input for the neural network classifier. The classifier makes a guess as to the category of the unlabeled images, and then those images are included in the dataset to be used for training the model.

This model often entails a series of preprocessing procedures, such as gaussian blurring, thresholding, erosion, and contouring, in order to improve the quality of the images and segment the various objects within them. The thresholding function simplifies the image for further processing by converting it to black and white after the image has been smoothed out by the Gaussian blur function, which also reduces the amount of noise in the image. After then, contouring draws lines around the things in the image to separate them from the background. After the preprocessing of the photos has been finished, the model will extract features from the images using the FPN architecture. The Efficient-Net model is a cutting-edge example of deep learning architecture. It has demonstrated outstanding performance in the picture classification tasks for which it was designed, which makes it an effective instrument for object recognition.

The system will use the UNET model to separate the lung nodules from the surrounding tissue, and it will use the efficient-net architecture to determine whether or not the nodules are cancerous. Ultimately, this will allow the system to accomplish its purpose. The Feature Pyramid networks are utilised in order to carry out the feature extraction. The technology will also have the potential to be utilised in other medical imaging modalities, such as magnetic resonance imaging (MRI) and positron emission tomography (PET), for the purpose of detecting and classifying a wider range of abnormalities and disorders.

C. Module Design

There are five distinct components that make up the architecture. Preprocessing, augmenting, segmenting, feature extraction, and nodule classification are the respective modules that make up the system.

Pre-Processing

Because CT scans can be acquired by a variety of scanners in a variety of medical clinics without the use of identical acquisition protocols, the data pre-processing [28] step is essential to normalise the data in a way that enables the convolutional network to learn appropriate and meaningful features properly. This is necessary because of the fact that CT scans can be acquired by different scanners. Noise, a lack of apparent fluctuations in the greyscale border, and other characteristics that make it difficult to split directly obtained CT scans of the lungs are some of the factors that contribute to this difficulty. As a consequence, the pre-processing of the images is the initial task. It is vital to exclude specific regions that make it tough to separate in order to cut down on the background noise in the CT scans of the lungs.

In the meantime, the boundary area of the CT picture that changes has been highlighted in order to simplify the subsequent segmentation process. Sharpening is required for smooth operation. The Preprocessing Module, which can be seen represented by Figure 2, is made up of the Gaussian Blur, Techniques of Thresholding, Erosion, and Dilation, as well as Contouring.

Gaussian blur

The Gaussian blur [29] effect is a linear low-pass filter that determines the value of each pixel by applying the Gaussian function to the calculation. The pace at which this weight decreases is defined by a Gaussian function; hence, the term "Gaussian blur" was coined to describe this phenomenon.

Augmentation

Image data augmentation is a method that creates altered versions of the images that are contained inside a training dataset. This allows the size of the dataset to be artificially increased without the need for additional data collection. The label for each of the photographs will be the same, and it will be taken from the image that was used as the basis for their generation.

In order to make the model become more resilient to alterations in the orientation of the patient within the CT scan, flipping the images vertically and horizontally can help. This is due to the fact that CT scans can be performed in a variety of positions. By inverting the pictures, one can replicate these varying positions, which in turn helps the model generalise better to new data.

UNET segmentation

Applications that perform picture segmentation typically make use of the U-Net architecture, which is prevalent in the field of biomedicine. It has two sides, one of which contracts while the other side extends in a symmetrical manner. The contracting side is composed of convolutional, Rectified Linear Unit (ReLU), dropout, and pooling layer sections. The number of feature channels can be raised by a factor of two through the process of downsampling. The expanding side contains portions of the upsampling (transpose convolution) layer, as well as the convolutional layers, the dropout layer, and the concatenation with the relevant feature channel from the contracting side.

By utilising the connections that exist between the contracting and expanding sides of the network, it is possible for the network to retrieve the necessary features from the appropriate layer. The number of feature channels in each enlarged region is cut in half. This applies to all expanded regions. The number of feature channels is reduced by one-half at each segment on the expanded side of the graph. The final convolution layer contains a mapping of the feature vector to the anticipated number of class labels.

The encoder network in UNET is the component of the architecture responsible for contracting. It is made up of convolutional layers and max-pooling layers, both of which gradually lower the spatial resolution of the input image while simultaneously raising the number of feature channels. This particular path of contraction is intended to record the contextual information as well as the high-level characteristics of the image. The decoder network, on the other hand, is the component of the architecture that takes up the most space. It is made up of upsampling and convolutional layers, both of which steadily raise the spatial resolution of the feature maps while simultaneously lowering the total number of feature channels.

The high-resolution details of the image are reconstructed using this expansion path, which also helps to optimise the segmentation mask. Encoder and decoder networks are linked together by a bottleneck layer, which is responsible for maintaining the high-resolution characteristics of the image. The skip connections that exist between the networks of the encoder and the decoder help to combine the high-level information from the encoder with the low-level features from the decoder, which ultimately leads to a segmentation that is more accurate.

Feature extraction

Feature extraction via Feature Pyramid Networks (FPN) with ResNet as the backbone is a well-established method in the field of computer vision that may be utilised for a variety of applications including object identification, segmentation, and classification. A CNN's performance can be improved with the help of a feature extraction network called FP N. This network makes use of multi-scale feature maps, which gives the CNN the ability to recognise objects on a variety of various scales. On the other hand, ResNet is an example of a deep residual network architecture that has been demonstrated to perform better than conventional CNN architectures. This is because ResNet enables deeper network depths to be achieved without running into the issue of vanishing gradients.

ResNet is able to identify high-level features in the input image when it is employed as the backbone for FPN. This is due to the fact that ResNet provides a solid foundation for feature extraction. The feature maps that are found in ResNet's intermediate layers are combined by the FPN architecture to build a pyramid of multi-scale feature maps. Up sampling the feature maps at lower resolutions and then combining them with the feature maps at higher resolutions is part of the top-down pathway in FPN.

The prediction of unlabelled nodules:

The FPN was developed to solve the challenge of locating items in an image that range in size from very small to very large. The backbone network is often a deep convolutional neural network like ResNet, which is used to extract feature maps from the input image. These feature maps can then be used in further processing. The input image is processed by the first convolutional layer, which consists of 64 filters and has a kernel size of 5x5 pixels.

After being processed by a batch normalisation layer, a ReLU activation layer, and a max pooling layer with a kernel size of 3x3 and a stride of 2, the output of the first layer is then output. After the output of the first stage has been processed by the second stage, which is made up of three blocks of residual data that have 256 filters and a kernel size of 3x3 pixels.

- The output of the second stage is fed into the third stage, which has a kernel size of 3x3 and consists of 4 residual blocks with 512 filters each.

- The output of the third stage is fed into the fourth stage, which has a kernel size of 3x3 and consists of 6 residual blocks with 1024 filters each. The output of the fourth stage is fed into the fifth stage, which is comprised of three residual blocks with a total of twenty-four hundred and eighty-eight filters and a kernel size of three by three.

- In order to construct a feature pyramid, the output of each stage is routed through a feature pyramid network, also known as an FPN, module. Following that, the feature pyramid is utilised for Classification as well as other activities farther downstream.

Classifier:

The During the training process for the classifier model, two distinct class numpy array files are utilised. These files are normally produced by a training dataset. The feature vectors that were retrieved from benign photos are located in one of the array files, and the feature vectors that were extracted from malignant images are located in the other array file.

Nodule Classification

Efficient Net [31] is a design for convolutional neural networks as well as a method for scaling that applies a compound coefficient to scale all dimensions of depth, width, and resolution in a convolutional neural network equally. The Efficient Net scaling method, in contrast to the prevalent practise, which arbitrarily scales these factors, adjusts network breadth, depth, and resolution uniformly using a set of predetermined scaling coefficients.

A model that makes use of an efficient net architecture that offers a compound scaling approach (one that scales all of its depth, width, and resolution dimensions) can realise the maximum accuracy increases possible.The baseline network has a considerable influence on the degree to which models scale well. Efficient Net is capable of handling a wide variety of tasks relating to the classification of images. Because of this aspect, it is an effective model for the process of transfer learning. Compound Scaling is a technique that is utilised by Efficient-net. This technique involves scaling all three dimensions (depth, width, and picture resolution) of the network while ensuring that there is a balance between these aspects of the network. Efficient-net's architecture is made up of a variety of convolution and MBConvolution blocks, all of which are coupled to one another in order to generate feature maps from images.

MB Conversion:

The mobile inverted bottleneck convolution, abbreviated MBConv, is utilised in this designed architecture. Then, two more concepts, namely inverted residual connections and linear bottlenecks, are taken from MobileNet-V2, which is a second enhanced version of MobileNet.

Mobile inverted bottleneck convolution, also abbreviated as MBConv, is the fundamental constituent of the Efficient-Net model family. The MobileNet models serve as an inspiration for the ideas that underpin MBConv.

OPERATIONS:

The input features are subjected to a linear transformation as the initial step in the process, with the weights from the first layer serving as the determining factor. A weight matrix W1 with dimensions (input_dim, units) is applied to the input features X, which causes a multiplication to take place. This operation results in the creation of a matrix of activations, and its dimensions are "batch size" and "units."

The subsequent step is to apply a non-linear activation function to the activations, which is the following step in the process. The ReLU activation function is used in the example code. This function applies the element-wise function $f(x) = \max(0, x)$, where x is the input to the function.
The process is repeated for consecutive layers, and on each subsequent layer, an additional linear transformation is applied to the activations of the layer below it, which is then followed by an application of a non-linear activation function.

The last layer consists of a single unit and implements the sigmoid activation function. This function transfers activations to a value that ranges from 0 to 1, where 1 indicates the highest anticipated likelihood of belonging to the class.

During the training phase of the model, the backpropagation method is used to compute the gradients of the loss with respect to the weights of each layer. These gradients are then applied to the process of optimising the model, where they are used to update the weights.

These gradients are generated by applying the chain rule of calculus, which entails computing the derivative of the loss with respect to the output of each layer, and then propagating these derivatives backwards through the layers of the network. The end result is a gradient that represents the relationship between the loss and the output of each layer.

Using the features included within the input data, the model is able to learn how to correctly predict the label that should be applied to each sample that is fed into it.

RESULTS AND DISCUSSIONS

The Python programming language and the Anaconda Jupyter tool were utilised in the successful completion of the project. Python is one of the programming languages that is utilised for deep learning the most frequently because of its adaptability and the vast number of libraries that are readily available. Some of these libraries include TensorFlow, PyTorch, Keras, and Scikit-learn. The Anaconda distribution of the Python programming language comes with a plethora of packages and tools that are standard fare in the field of data science. The Anaconda distribution features Jupyter as one of its pre-installed packages.

Because graphics processing units (GPUs) are built to handle huge quantities of parallel computing, sophisticated neural networks can be easily trained on GPUs. The RAM of 16 gigabytes makes it feasible to handle enormous datasets, which is essential for the training of our deep learning models. In order to successfully run deep learning software and tools like TensorFlow and PyTorch, a robust and user-friendly environment is required, which is something that the Windows operating system can provide. For the purpose of training the model, a set of pre-processed CT images was provided to the U-Net. Normalisation was used as the starting point for the initialization of the weights, and the ReLU activation function was used. Both the learning rate and the batch size were set to 32. The learning rate was set at 0.001. For the purpose of calculating the loss function, the binary cross-entropy loss function was put to use. In comparison, just 20% of the data are used for validation while 80% of the data are used for training. The training of the model lasted for a total of 250 epochs. The best roll of the dice resulted in a value of 0.5009, while the average roll of the dice produced a value of 0.4273. It was discovered that the model suffered an average loss of 0.27.

The Adam optimizer is a technique that can be utilised as an alternative to the stochastic gradient descent method for deep learning models. This method deals with sparse gradients on noisy scenarios and offers optimisation by integrating the essential qualities of AdaGrad and RMSProp. Due to the fact that it is a non-linear activation function, the ReLU function is resistant to simultaneously activating all neurons. The U-Net model that's been suggested uses ReLU activation since it helps produce better results. If the result of the linear transformation is less than 0, the neuron will stop sending out signals and will become inactive. When the ReLU activation function was utilised alongside the Adam optimizer, one may say that the outcomes were satisfying.

The accuracy of a model is plotted on a graph called the training vs. validation accuracy graph, which illustrates how accurate a model is on both the training dataset and the validation dataset.

During the training and validation processes, a model will experience loss on both the training dataset and the validation dataset. This loss will be plotted on a graph called the training vs validation loss graph. During the training phase of the process, the model is optimised for performance on a training dataset in an effort to reduce the value of the loss function.

CONCLUSION AND FUTURE WORKS

The findings of the suggested model for lung nodule segmentation and classification utilising the UNet and Efficient-Net neural network have been encouraging in terms of their ability to detect and classify lung nodules in an accurate and time-saving manner. In terms of accuracy, recall, and loss, the suggested approach surpasses other methods that are considered to be state-of-the-art. The ablation study that was carried out as part of this project shed light on the contribution that each individual component of the suggested algorithm made to the performance as a whole.

In addition, the technique that was proposed has the potential to be applied to larger datasets and to be optimised for real-time processing. Both of these possibilities are exciting. It is possible that the successful adoption of this algorithm will contribute to the early detection and treatment of lung cancer, as well as increase the accuracy and efficiency of lung cancer screening programmes. However, the suggested approach still has some shortcomings, including the requirement for large annotated datasets as well as the possibility of overfitting.

The suggested algorithm has the potential to be enhanced by the incorporation of explainable artificial intelligence (XAI) techniques. These approaches would make the algorithm more interpretable and transparent. Incorporating approaches from XAI can make it possible to identify the features and patterns that the algorithm utilises to make judgements, which can then provide insights into the logic behind the algorithm's output. This can be especially helpful for applications in the medical field, where the decision-making process is required to be open to clinicians and patients and to be explained in terms that both can comprehend.

In order to increase the accuracy of lung nodule segmentation and classification, the suggested algorithm can be optimised for multimodal imaging, such as computed tomography (CT) and positron emission 40 tomography (PET) scans. In addition, the model that has been proposed has the potential to be expanded so that it may take into account additional forms of lung cancer, such as small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Therefore, work that will be done in the future can concentrate on establishing a more comprehensive algorithm that is able to effectively and efficiently detect and categorise a variety of different types of lung abnormalities.

REFERENCES

[1]. Liang M, Tang W, Xu DM, Jirapatnakul AC, Reeves AP, Henschke CI, Yankelevitz D, "Low-Dose CT Screening for Lung Cancer: Computer-aided Detection of Missed Lung Cancers. Radiology," 2016 Oct; 281(1):279-88.
[2]. Liang M, Tang W, Xu DM, Jirapatnakul AC, Reeves AP, Henschke CI, Yankelevitz D, "Low-Dose CT Screening for Lung Cancer: Computer-aided Detection of Missed Lung Cancers. Radiology," 2016 Oct; 281(1):279-88.
[3]. Juanyun Mai, Minghao Wang, Jiayin Zheng, Yanbo Shao, Zhaoqi Diao, Xinliang Fu, Yulong Chen, Jianyu
[4]. Xiao, Jian You, Airu Yin, Yang Yang, Xiangcheng
[5]. Liu K, Li Q, Ma J, Zhou Z, Sun M, Deng Y, Tu W, Wang Y, Fan L, Xia C, Xiao Y, Zhang R, Liu S, "Evaluating a Fully Automated Pulmonary Nodule Detection Approach and Its Impact on Radiologist Performance," Radiol Artif Intell. 2019 May 29; 1(3):e180084.
[6]. Wei Y, Shen G, Li JJ, "A fully automatic method for lung parenchyma segmentation and repairing. J Digit Imaging," 2013 Jun; 26(3):483-95.

[7]. G, Amalorpavam & T, Harish & Kumari, Jyoti & Mallaiah, Suresha, "Analysis of Digital Images Using Morphlogical Operations," International Journal of Computer Science and Information Technology, 2013.
[8]. Tran, Song-Toan, Ching-Hwa Cheng, Thanh-Tuan Nguyen, Minh-Hai Le, and Don-Gey Liu, "TMD-Unet: Triple-Unet with Multi-Scale Input Features and Dense Skip Connection for Medical Image Segmentation" Healthcare , 2021, no. 1: 54.
[9]. Ronneberger, Olaf & Fischer, Philipp & Brox, Thomas, "U-Net: Convolutional Networks for Biomedical Image Segmentation," LNCS. 9351. 234-241.
[10]. Guangrui Mu, and Yanbo Chen, "Relu Cascade of Feature Pyramid Networks for CT Pulmonary Nodule Detection," Machine Learning in Medical Imaging, 10th International Conference, MLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, 2019.
[11]. Agrawal, Anand, Ansari, and Mehrotra, "A transfer learning approach for AI-based classification of brain tumors," Machine Learning with Applications, 2020.
[12]. Jung, Hwejin & Kim, Bumsoo & Lee, Inyeop & Lee, Junhyun and Kang, Jaewoo, "Pulmonary Nodule Classification in Computed Tomography Image Using a 3D Deep Convolutional Neural Network," KIISE Transactions on Computing Practices, 2018.
[13]. Divya Anwesh Sahu, Nagaraju.Y., Sheela Rachel.K., and Venkatesh, "Histopathological Image Classification of Breast
[14]. C. Zhao, J. Han, Y. Jia, and F. Gou, "Lung Nodule Detection via 3D U-Net and Contextual Convolutional Neural Network," 2018 International Conference on Networking and Network Applications (NaNA), Xi'an, China, 2018.
[15]. Bianconi F, Fravolini ML, Pizzoli S, Palumbo I, Minestrini M, Rondini M, Nuvoli S, Spanu A, Palumbo B, "Comparative evaluation of conventional and deep learning methods for semi-automated segmentation of pulmonary nodules on CT," Quant Imaging Med Surg., 2021 July.
[16]. Kingma, Diederik and Ba, Jimmy, "Adam: A Method for Stochastic Optimization," International Conference on Learning Representations, 2014.
[17]. Jambek A.B, and Said K.A.M, "Analysis of Digital Images Using Morphological Operations," Journal of Physics Conference Series, 2021.
[18]. Ahmed Abou ElFarag, Nahla M. Ibrahim, and Rania Kadry, "Gaussian Blur through Parallel Computing," International Conference on Image Processing and Vision,Engineering, 2021.
[19]. Goutami Dey, Nilanjan Dey, Saurab Dutta, Sayan. C Payel Roy, and Ruben Ray, "Adaptive thresholding:
[20]. A comparative study," International Conference on Control Instrumentation, Communication and Computational Technologies (ICCICCT).
[21]. Haikel Alhichri, Nassim Ammour, Naif A. Alajlan, Asma S. Alswayed, and Yakoub Bazi, "Classification of Remote
[22]. Sensing Images using EfficientNet-B3 CNN Model with Attention," Advanced Lab for Intelligent Systems Research (ALISR), Computer Engineering Department, College of Computer and Information Sciences, King Saud University, Saudi Arabia.
[23]. Parikh, H. (2021), "Algae is an Efficient Source of Biofuel", International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 11.
[24]. Parikh, H. (2021), "Diatom Biosilica as a source of Nanomaterials", International Journal of All Research Education and Scientific Methods (IJARESM), Volume 9, Issue 11
[25]. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," IEEE Trans Med Imaging, 2016 May.
[26]. Tan, Mingxing and Le Quoc, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," International Conference on Machine Learning, 2019. Suji RJ, Bhadouria SS, Dhar J, Godfrey WW, "Optical Flow Methods for Lung Nodule Segmentation on LIDC-IDRI Images," J Digit Imaging, 2020 Oct.