Transfer learning with GoogLeNet for detection of lung cancer

Muayed S AL-Huseiny, Ahmed S Sajit
College of Engineering, University of Wasit, Iraq (UoW), Iraq

Article Info

ABSTRACT

The use of computer algorithms has gained momentum in filling/assisting roles of specialists especially in early diagnosis scenarios. This paper proposes the employment of deep neural networks (DNN) to detect images with malignant nodules of lung computed tomography (CT). The method includes subjecting input images to a simple and fast pre-processing which isolates regions of interest (ROI), that’s the lungs dominated area, ridding the images of other surrounding tissues and artifacts. Centered and size normalized images are then fed to a deep neural network for training and validation. In this work transfer learning is used to readjust GoogLeNet DNN to learn this medical data. This includes allowing final layers of the DNN to evolve while restricting deep layers. In this setting, a rough, unprocessed dataset, the IQ-OTH/NCCD lung cancer dataset was used to train/validate the proposed algorithm. Experimental results show that this algorithm scores 94.38% accuracy, which outperforms benchmark method previously used with this dataset.

Keywords:
Artificial intelligence
Computer vision
CT scans
Deep neural network
Lung cancer
Transfer learning

This is an open access article under the CC BY-SA license.

1. INTRODUCTION

Lung cancer is considered among the fiercest forms of diseases across the world. World health organization (WHO) estimates that more than 1.76 million perish annually due to this type of cancer, [1], [2]. One of the most important tools for the diagnosis of lung cancer is computed tomography (CT). This screening method, however, requires the skills of a trained specialist, something not always available [3]. The use of the digital processing in medicine which started in early sixties of last century offers reducing processing time and improvement in specificity and sensitivity [4]. Also, the steady increase in the scale and number of medical data modalities is considered as a challenge for health communities. Computer algorithms are the prime candidates in providing the capacity to ingest and analyse these data. The medical deployment of these algorithms have seen relentless proliferation in recent years [5]. These methodologies are further specialised according to their final produce into computer aided detection (CADe) and diagnosis (CADx) systems [6]. The adoption of CAD systems en masse relies on achieving high sensitivity, cutting false positives down and increasing processing speed while maintaining affordable costs [7].

Computerised medical data analysis usually includes pre-processing, segmentation, feature extraction and classification [8]-[11]. The nature of tissues and organs of human body such as those of lungs makes isolation of cancer nodule a demanding chore for radiologists with failure rates of about 35% [12], [13]. Specifically, the lack in homogeneity of lung tissue and abundance of blood vessels and bronchioles complicates the task of segmentation [14], [15] which is the initial step in detection and diagnosis [15]. The authors in some studies [16] developed a support vector machine (SVM) based algorithm where genetic algorithms were used to select the optimal set of features, the achieved accuracy was 93.66%, another
instance of this algorithm which uses neural networks is claimed to have achieved accuracy of 95.87%. Some researchers [17] ensembled neural networks guided by fuzzy logic clustering to segment lung cancer images on a moderate dataset, this was used along with some physical signs and symptoms in the classification process. Other researchers [18] also used neural networks with self-organizing maps to develop their CAD system gaining accuracy of 90.63%. Different approach was taken in [19], the focus of their study was on the fate of patients of lung cancer by trying to predict mortality rates, their stated accuracy was 96%.

With the introduction of numerically feasible deep neural networks a new wave of CAD based lung cancer detection systems were proposed with success rates comparable to those of human experts. The authors of [20] used three types of classifiers, convolutional neural network (CNN), deep belief network (DBN), and stacked denoising autoencoder (SDAE) on the LIDC/IDRI dataset, where, the nodules scans are initially segmented manually, they are then downsampled to 52X52 before feeding the features to the classifiers. The work in [21] propose an interesting 3D based features, the median intensity projection (MIP) derived multi-view data, the features are then used to train a CNN to extract lung nodules whose output is fed to a Gaussian process regression (GPR) to score the degree of malignancy. Microscopic images are used in [22] to train a CNN for automatic detection of lung cancer. The algorithm proposed in [23] uses MIP 3D features proposed in [21] with the same dataset to train a powerful CNN, the GoogLeNet [24] in transfer learning framework. Different approach is followed in [25] where a short pipeline is utilized by replacing segmentation by thresholding, U-Net is then used to select images with nodule candidates, these are classified by using 3D CNNs, Vanilla and GoogLeNet, Data Science Bowl [26] and lung nodule analysis 2016 [27] datasets were used to train and test the proposed framework. The algorithm in [28] suggests deploying the AlexNet CNN directly to classify malignancy in lung CT scans of the IQ-OTH/NCCD [3] dataset, while the method produces good results, there are still issues regarding the ability of such approach to produce similar results with different datasets or with augmentation of the same dataset usually employed to improve the reliability of algorithms and to remedy the issues of overfitting.

As such experimentation and development are still viable for CAD systems to gain maturity as these are still vulnerable to high false-positive rates compared to specialist [12]. In this study a framework is presented to serve in the direction of diagnosing lung cancer in CT scan images. The rest of this paper includes: a detailed description of the methods used; the design of the experiments; discussion of the results; and finally, the concluding remarks.

2. RESEARCH METHOD

The proposed framework consists of a set of pre-processing steps, followed by the setup of a machine learning model. The model is trained and validated by using image dataset. The various stages of the framework are presented subsequently:

2.1. Pre-processing

The main purpose of this stage is to help the learning model focus its attention on the most relevant field rather than capturing unrelated tissues, markers and other clutter. The outcome of this stage is a region of interest (ROI) containing centred and size normalised lungs region, this is summarized:

a) Texture analysis: to achieve this, Gabor filter [29] is applied. These filters are bandpass filters normally utilized in image processing to emphasize regions with similar texture [30] as shown in 2nd column of Figure 1, This is the case in human organs. Thus, the response of the filter to an input image \( I(x, y) \) is computed in (1):

\[
g(x, y) = \exp \left( -\frac{(x^2 + y^2)^2}{2\sigma^2} \right) \cos \left( 2\pi \frac{x}{\lambda} + \psi \right)
\]

where \( x = x \cos \theta + y \sin \theta \), and, \( y = x \sin \theta + y \cos \theta \). Given that \( \lambda \) is the sinusoidal wavelength, \( \theta \) is the angle of the normal to the wave of a Gabor function, \( \psi \) is the phase shift, \( \sigma \) is the Gaussian standard deviation and \( \gamma \) is the aspect ratio [31]. These parameters were selected empirically to maximize the response of the filter in the lungs’ tissue regions.

b) Morphological operation: the output of the previous step is first, threshold-ed to convert the image into black and white mask. Image dilation followed by erosion are then used to fill holes generated due to filtering and to produce homogeneous regions. The largest region which represents the surrounding of patient body is then removed to produce the lungs mask depicted in 3rd column of Figure 1.

c) ROI extraction: lungs mask is multiplied by the original image. Bounding box is then calculated which frames lungs’ extremities. Finally, images are size normalised. Example ROI images are shown in last column of Figure 1.
2.2. Transfer learning

It is an acceptable approach in artificial intelligence (AI) (particularly, in machine learning (ML)) as with other fields of science that knowledge and experience gained in solving certain problem is utilized in analyzing and tackling other related (and even possibly unrelated) problems and challenges. This in the research of AI and ML is called transfer learning [32]. The motive for such trend is efficiency, where often, limited computing resources deter research. Utilizing previously trained models help bridge some gaps, yet, careful choice of pretrained model is necessary to ensure fine tuning its performance towards the new domain produce feasible results compared to initializing the entire process of model training. Figure 2 details this process.

Initial layers which are designed to fire on low level features are left intact as it is assumed that knowledge gained in this area is useful in the transferred domain as well. Later layers which learn domain specific (high level) features are replaced and tuned via the subsequent training process. The overall training time and parameter adjustment is significantly reduced as most of the model is inherited from the original domain training [34]. Since classification tasks of everyday ordinary objects is in essence synonymous to that used in medical application such as the classification of lung nodules in terms of extracting and learning some low-level attributes, like edges, texture, and color information. Hence, the utilization of pre-trained parameters facilitates learning or high-level feature in order to cut training time and improve model performance [23]. This notion is particularly exercised here by repositioning the parameters of a well-known CNN modeled around ImageNet objects [33] to the domain of lung cancer data.

![Image](image_url)

Figure 1. Three steps of pre-processing are shown for two randomly selected input images, each input image and the subsequent preprocessing are depicted on a row. Column-wise, input images are in (i); texture analysis in (ii); morphological operations in (iii); ROI extraction in (iv)

![Diagram](diagram_url)

Figure 2. Block diagram showing the process of transfer learning of a DNN trained on images from ImageNet dataset [33], the network is modified to learn/classify patterns from different domain (adapted from [34])
2.3. GoogLeNet

GoogLeNet is a convolutional neural network; its architecture is shown in Figure 3. The model of this network was reportedly developed to account for high feature representations by using million everyday object images included in a huge dataset, the ImageNet dataset [33]. It has the capacity to classify patterns of around 1000 images. It utilizes 12 times less parameters than Alexnet [35]. Similar to other neural networks employed in computer vision applications, this model accepts images as input and produces labels of one of its learned classes together with the level of confidence as output [34]. The architecture of GoogLeNet is built of 22 layers including 9 inception modules. The modified inception module shown in Figure 4 uses learnable filters [24] with sizes ranging from (1x1) to (5x5) to perform convolution in parallel which helps capture feature of different levels of details [35].

Figure 3. The architecture of GoogLeNet [24] with modified inception module (diagram partitioned to fit)

Figure 4. The components of the modified inception module used in GoogLeNet architecture [24]
2.4. Experiment design

In this study we compiled the publicly available (IQ-OTH/NCCD) lung cancer dataset of CT images [3] which was used to train/verify the proposed learning model. The dataset was recently gathered in two specialist hospitals in Iraq during the last quarter of 2019. The dataset is focused mainly on categorizing cases in one of three modalities of lung cancer: normal (NRM), benign (BGN) and malignant (MLG), all performed by oncologists and radiologists. This repository includes about 1190 CT scan images with dimensions 512*512. The images come from 110 cases decimated into: 40 cases classified as MLG; 15 cases classified as BGN; the remaining 55 cases were classified as NRM. The technical details regarding the imaging process are listed in [3]. Quick review of this dataset shows that this moderate sized collection has some irregularities, such as unusual pose and presence of noise. This makes the images interesting to use to measure the reliability of learning models. The user interface (UI) of MATLAB© deep learning toolbox showing the training parameters of GoogLeNet by using the IQ-OTH/NCCD dataset as shown in Figure 5.

![Figure 5](image-url)  
Figure 5. The user interface (UI) of MATLAB© deep learning toolbox showing the training parameters of GoogLeNet by using the IQ-OTH/NCCD dataset

3. RESULTS AND DISCUSSION

In this research transfer learning of a pretrained deep neural network, the GoogLeNet was employed to develop a learning model of the IQ-OTH/NCCD lung cancer dataset. Experimental results show that the trained model has gained an overall accuracy of 94.38% on the validation data. Figure 6 as shows the progress of the training and validation over 1788 iteration divided into 12 epochs. In the top plot which presents the accuracy, it is noticed that the algorithm starts settling within the vicinity of its final value halfway at around 900 iterations. It is also noticed that the algorithm keeps suffering from fluctuations in the training process. These fluctuations are attributed to the nature of the data whose imaging conditions can be described as non-uniform. Additionally, the augmentation used with the data is complicating this issue. Nevertheless, it is assumed that such conditions are the norm when a machine learning model is deployed to real problems where data conditions are far from perfect. Bottom row of Figure 6 show the loss associated with the training/validation procedure. Figure 7 presents example images classified by the proposed algorithm.

The results also show that the trained model achieved 95.08% sensitivity and 93.7% specificity. Since the primary aim of classification models is to boost the sensitivity which manifests the success or hit rate, while preserving the specificity which signifies true negative rate [36]. The achieved values of these measures presented here reflect the reliability of the proposed framework (ROI extraction followed by deep learning) in classifying similar modality of data compared to the same algorithm used without pre-processing also implemented by the authors, which scores around 70% accuracy. The confusion matrix of the discussed results is presented in Table 1. It’s worth referring that for the sake of simplicity of presentation of the confusion matrix both classes of BGN and NRM are considered same category here. This conforms to the initial aim of this paper that is to design a model capable of detecting malignancy in lung CT scans where other non-malignant issues are not further detailed.
Transfer learning with GoogLeNet for detection of lung cancer (Muayed S AL-Huseiny)

---

Transfer learning with GoogLeNet for detection of lung cancer (Muayed S AL-Huseiny)

---

Figure 6. Progress of training/validation of the proposed algorithm generated by the MATLAB® deep learning toolbox over 12 epochs and 1788 iterations: (i) accuracy; (ii) loss

Figure 7. Four randomly selected example CT scan images shown above each is the classification label with its probability as produced by the proposed algorithm. NRM refers to Normal, BGN refers to Benign, and MLG refers to Malignant

Table 1. Confusion matrix

<table>
<thead>
<tr>
<th>Actual class</th>
<th>Predicted class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-malignant (positive)</td>
<td>Non-malignant (negative)</td>
</tr>
<tr>
<td>Malignant</td>
<td>TP=116</td>
</tr>
<tr>
<td>Non-malignant</td>
<td>FP=8</td>
</tr>
</tbody>
</table>
It's worth noting that the time consumed by the training and validation stages is 57 minutes and 41 seconds, while the time for stamp for pre-processing step was less than 2 seconds run on a single core of an Intel Core-i3 processor powered MacBook Pro machine clocked at 2.5 GHz with 16 GB ram, as mentioned in Table 1. These outcomes were also compared to those recorded by the benchmark algorithm originally proposed for this dataset. In [3] the authors employed SVM to perform classification duties. The proposed work also performs better than the algorithm in [28] which employs AlexNet to classify the images of IQ-OTH/NCCD dataset.

In Table 2, it is evident that the proposed algorithm comparatively outperforms in all three performance measures. It is clear that the proposed algorithm also outperforms other approaches implemented by using datasets different from the current one. Although, care should be taken not to draw solid conclusions regarding the performance of this algorithm compared to the other algorithms unless tested comparatively on equal grounds.

### Table 1. Comparison of performance metrics

<table>
<thead>
<tr>
<th>Research Method</th>
<th>Dataset</th>
<th>Epochs</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM [3]</td>
<td>IQ-OTH/NCCD</td>
<td>/</td>
<td>89.033%</td>
<td>93.66%</td>
<td>89.88%</td>
</tr>
<tr>
<td>MIP-CNN [23]</td>
<td>LIDC/IDRI</td>
<td>100 train./300 val.</td>
<td>84%</td>
<td>78%</td>
<td>81%</td>
</tr>
<tr>
<td>GPR [25]</td>
<td>Vanilla Data Science</td>
<td>/</td>
<td>59.3%</td>
<td>76.1%</td>
<td>70.5%</td>
</tr>
<tr>
<td></td>
<td>GoogLeNet Bowl [26]/LUNA16 [27]</td>
<td>/</td>
<td>77%</td>
<td>74.1%</td>
<td>75.1%</td>
</tr>
<tr>
<td>Multi-classif [20]</td>
<td>BN DAE</td>
<td>LIDC/IDRI</td>
<td>100</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>TumorNet [21]</td>
<td>LIDC/IDRI</td>
<td>10,000 iter.</td>
<td>/</td>
<td>/</td>
<td>82.47%</td>
</tr>
<tr>
<td>Microscop. [22]</td>
<td>Proprietary 60000</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>71%</td>
</tr>
<tr>
<td>AlexNet [28]</td>
<td>IQ-OTH/NCCD</td>
<td>100</td>
<td>95.71%</td>
<td>95%</td>
<td>93.54%</td>
</tr>
<tr>
<td>Proposed</td>
<td>IQ-OTH/NCCD</td>
<td>12</td>
<td>95.08%</td>
<td>93.7%</td>
<td>94.38%</td>
</tr>
</tbody>
</table>

### 4. CONCLUSION

In this paper a learning algorithm is presented to model publicly published medical data. The data, IQ-OTH/NCCD lung cancer dataset, includes a collection of CT scans of lung region for subjects with probable lung cancer. The algorithm proposes to use computationally inexpensive pre-processing process to extract lungs region and remove remaining clutter and irrelevant surroundings. The proposed framework also explores the benefits of using pretrained deep neural networks in a transfer learning approach to reduce time and resources. This is achieved by retraining GoogLeNet to accommodate for the features of this medical data. Empirical outcomes show that the proposed algorithm yields higher accuracy at 94.38% compared to the algorithm suggested originally with this data set which scores 89.88%. The proposed algorithm also outperforms other algorithms implemented on different datasets. Furthermore, other performance metrics such as sensitivity and specificity were computed. These metrics also support the above verdict which entails that this approach is suitable for reliably classifying medical imagery. The merger of suitable pre-processing with convolutional neural networks in this fashion has been validated by using the same deep learning model trained here without utilizing ROI extraction (pre-processing), the later variant resulted in much lower accuracy at 70%. The presented approach, also hints at the prospects of expanding the utilization of pretrained models as a means of compensating for lack of computational resources in medical facilities and academic institutions while in the same time prove concepts of software engineering such as software reusability.

### REFERENCES


Transfer learning with GoogLeNet for detection of lung cancer (Muayed S AL-Huseiny)
BIOGRAPHIES OF AUTHORS

Muayed S AL-Huseiny received his BSc and MSc degrees in Computer and Software Engineering from AL-Mustansiriya University, Iraq, in 2002 and 2005 respectively. He received PhD in Electronics and Computer Science from the University of Southampton, UK, in 2012. He received several research grants as well as study scholarship. He is an alumnus of the Fulbright Visiting Scholars program in the University of Central Oklahoma, USA, in 2014. Dr. AL-Huseiny is currently a faculty member and Dean of College of Computer Science and Information Technology in Wasit University, Iraq. His research interests include computer vision, machine learning, biomedical engineering, IoT, and cryptography. He published several research papers in these fields and authored a book in cryptography. His current focus is on medical informatics and developing wearables for medical diagnosis by using machine learning.

Ahmed S Sajit is a faculty member at Wasit university, Iraq; he received his BSc in Control and Computer Engineering from the University of Technology, Iraq, in 2010. He finished his MSc in Computer Engineering at California State University in Fullerton, USA, in 2017. He was granted a nationally prestigious scholarship from the Iraqi prime minister’s office in 2010. Ahmed is currently pursuing his PhD in phoronics and Laser applications. His research interests include: Computer Architecture, Holograms, Neural Networks, AI, Algorithms, and Phase Change Material Memory.