Graph Neural Networks for Lung Cancer Classification on 3D CT-scans

EECE 693

Bou Habib Philippe, El Sayed Youssef, Osman Sondos

ABSTRACT

Global statistics declared lung cancer as the leading cause of cancer related death and the most frequently occurring cancer [2]. To diagnose patients, doctors rely mainly on CT-Scans of a patient's chest to find cancerous nodules. In recent years, researchers have started developing computer-aided detection (CAD) systems to assist doctors during the analysis of medical images and through the detection process. While CNNs have been the focus in previous work, major limitations were presented including overfitting due to dataset imbalances and long training times. To resolve this issue, we propose a new model that does nodule segmentation from a patient's 3D CT scans followed by classification using a GIN graph classifier. We perform thorough pre-processing steps including lung segmentation and use U-Net++ to perform nodule segmentation, which are in turn transformed into a graph structure and fed into the GIN classifier. The GIN classifier resolves the issue of overfitting and dataset imbalances by leveraging neighbor-level relations. We were able to achieve an accuracy of 100% on our test set, providing a strong proof of concept and improvement to the state-of-art model.

I. INTRODUCTION

Lung cancer is a disease caused by the mutation of cells in the lungs. In general, these mutations happen after an individual inhales harmful and dangerous toxins. The leading cause of lung cancer is smoking, with 90% of lung cancer patients reporting being smokers [1]. Lung cancer is the leading cause of cancer-related death worldwide with 1.8 million deaths in 2020, and the second most commonly diagnosed cancer with 2.2 million new cases this same year (11.4% of cancers diagnosed) [2] It also has the lowest five-year survival rate with only 18.6% of patients surviving after five years of first diagnosis.

An early detection and diagnosis is key to improve a patient's chances of survival, but at the same time, it is significantly harder to detect early stages of lung cancer as it has less obvious manifestations [2]. Each year, both medical and technological specialists thrive to propose new methods of detection and diagnosis. The process relies mainly on imaging tests such as X-Ray images and CT scans of the chest to detect small cancerous nodules [3]. In practice, it is not as easy as it sounds, because early stage cancer nodules can be less than 10mm in diameter in large scan images. Therefore, the development of a state-of-the-art model that can reliably and accurately diagnose patients with lung cancer would provide a speedy and cost-effective solution to early lung cancer detection, potentially increasing a patient's survival chance.

Previous work has mostly relied on using Deep Convolutional Networks for pattern recognition and/or image classification. CNNs specifically are known to result in good accuracy in the area of medical imaging [19]. Most models that were built for lung anomalies detection relied on either CT scans or X-rays to train their models for malignancy detection [5]. It is interesting to work on CT Scans since they have high sensitivity [21]. In fact, most trained models exclusive for lung cancer detection used 2D [15] or 3D [18] CNNs using CT scans with max pooling, followed by different classifiers to produce an output [6]. They also relied on data augmentation and preprocessing of input images to extract features. Although multiple models reached decent levels of accuracy, the best models achieved a maximum accuracy of 87.5% [20], which is relatively lower than that of other lung disease detection models, such as COVID detectors that can achieve accuracies of 93.64% [20]. Moreover, a large number of false diagnoses were false positives [18]. In addition, some studies use fewer preprocessing steps than necessary in exchange for more efficient models. Thus, one of the possible limitations of current models is the lack of strong preprocessing and segmentation steps, which will allow better feature extraction and classification later on in the process. This can be considered one of the main causes yielding a relatively low accuracy and a high false positive rate. On the other hand, as efficient as they could be, CNNs present some limitations such as overfitting if the classes in the dataset are very imbalanced, which is the case in most lung cancer datasets [20].

Our approach is to build a binary classifier that would detect lung cancer from a patient's 3D CT-scans using Graph Isomorphism Networks (GINs). The GIN will be preceded by proper preprocessing and segmentation steps which will allow for a more accurate

classification by reducing the searching area for the classifier. These steps will include image preprocessing, lung segmentation, nodule segmentation using U-Net++, a state-of-the-art segmentation algorithm based on a CNN architecture, and steps to transform images into graph structures to input to the GIN. The GIN algorithm maps isomorphic graphs with the same embedding and leverages the neighbor-level relations in the graphs to offset the effects of an imbalanced dataset and avoid overfitting. Therefore, our goal is to develop a model that is both high in accuracy and efficient in detecting tumors in the lungs from a patient's chest 3D CT-Scan.

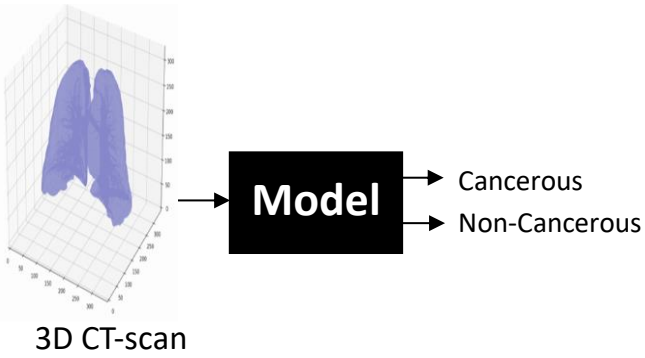


Figure 1. Simplified Black Box Diagram of our learning model

Moreover, if successful, this new method could have many applications in the detection of diseases that rely on imagery testing. We plan on implementing GNN, and more specifically GIN (Graph Isomorphism Network), along with thorough pre-processing and segmentation steps, to consider neighbor relations in scans to achieve higher accuracy and lower FP rate.

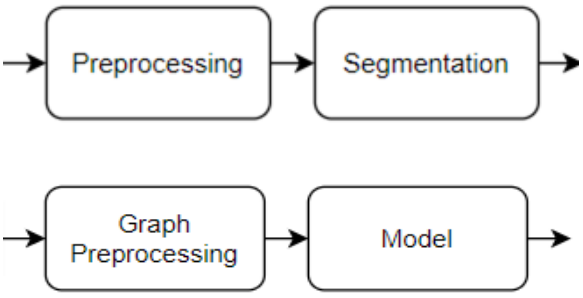


Figure 2. Summary of steps from intended model

As for our dataset, we will use the LIDI-IDRI database (Lung Image Database Consortium image collection) [4]. It contains 244,527 images of scans from 1010 lung cancer patients. For each patient, we will be using two modalities: CT (Computer Tomography), and DX (Digital Radiography- Digital X-Ray). The CT-scans are provided in slices, accompanied by the necessary annotation for each slice. Finally, each scan is accompanied by a detailed diagnosis corresponding to each patient ID.

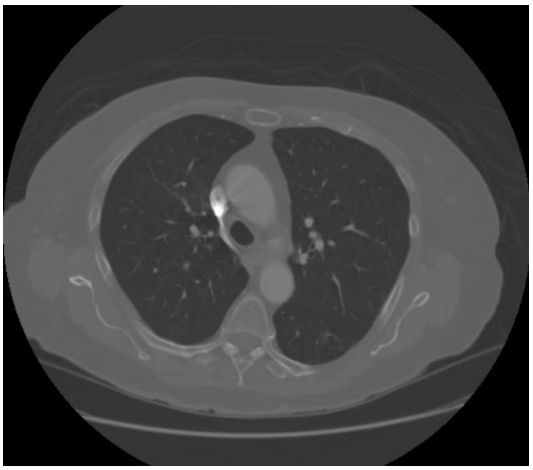


Figure 3. 2D CT scan slice from a lung cancer patient with a malignant nodule

II. SURVEY OF RELATED WORK

A. U-Net and U-Net++

U-Net is a convolutional network architecture mainly used for biomedical image segmentation. It is one of the best methods when it comes to microscopic units' segmentation, making it highly interesting for our project topic. U-Net++ is a nested U-Net architecture that differs mainly on skip pathways (convolution layers and dense skip connections) [7]. These changes can further improve gradient and semantic gaps between encoder and decoder which will yield better results.

Zhou et al. [8] present a comparison between the two architectures across multiple datasets, including the one we would be tackling later in this paper. They performed nodule segmentation on CT Scans, which are microscopic units. Their results show that Nested U-Net outperformed U-Net on all datasets, including the lung nodule segmentation where the difference was 77.21% in accuracy compared to 71.47% for U-Net (on a model with deep supervision).

B. Efficient-Net Benchmark

Efficient-Nets are a neural network architecture that uniformly scales all dimensions using a compound coefficient. It was proven to yield better accuracy and improves the efficiency of models. It mainly reduces the parameters used making it an effective compound scaling method.

Tan et al. [9] conducted several experiments using multiple scaled network architectures. They used the ImageNet pretrained weights to study ConvNet scaly and identify how balanced the dimensions of the network is. They show how compound scaling provides a good balance between width, depth, and resolution scaling which will lead to improvements in the model

performance. They concluded that the EfficientNet model can be very well scaled to perform a satisfying accuracy of 91.7% on the CIFAR-100 dataset and 98.8% on the Flowers dataset. Figure 4 shows how the proposed networks outperform conventional methods in terms of accuracy using ImageNet.

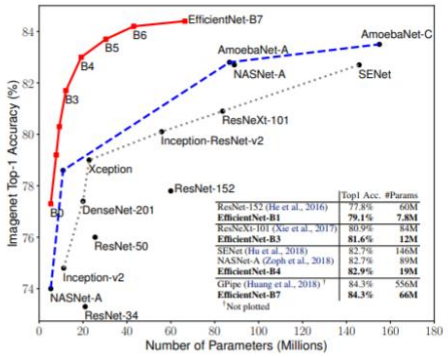


Figure 4. Model Size vs ImageNet Accuracy

C. Graph Neural Network

As mentioned in the introduction, a new method used to display relationship-features after data feature extraction is the use of GCN. This method is starting to be used in medical analysis to enhance the performance of normal CNNs.

With the rise of studies surrounding COVID-19, researchers such as Yu et al. [10] implemented a GCN based model called ResGNet to detect COVID-19 pneumonia from lung CT scans. They use the residual network ResNet101 to create ResNet101-C, which has enhanced feature extraction performance. ResNet101-C is reutilized to do feature extraction for the graph construction, contributing to the performance of GNN-based ResGNet-C. The paper focused on the benefits of GCN compared to CNN, especially on the fact that GCNs give weight to the underlying relationships between each element due to the fact that it considers each element as a node and its relationships to other nodes as edges. This is especially beneficial in image classification.

Zhou et al. [11] tackled colorectal cancer detection by transforming the H&E stained images into graphs using a cell-graph convolutional neural network. This emphasizes the connections between the nodes: having edges represent cellular interactions and nuclei represent the nodes themselves. They compared their results to other state- of-the-art models such as Context Aware CNN or CA-CNN and BAM, achieving better results in image accuracy proving the superiority of GCNs in medical imaging.

Zhang et al. [12] compared their model which combines GCN and CNN methods to help detect Breast Cancer using mammogram images. The model is a hybrid of an 8-layer CNN (including 2-fully connected dropout layers) followed by a two-layer GCN. They used GCN to represent relations between the features extracted from the data, followed by k-means clustering on each individual image-level representation. As a result, by comparing the hybrid model to other less advanced models, they realized that the most accurate model is the one they called BDR-CNN-GCR (batch normalization(B) , dropout (D), rank-based stochastic pooling(R)) achieving a sensitivity of 96.20±2.90%, a specificity of 96.00±2.31% and an accuracy of 96.10±1.60%. They envision working on other combinations of GCN and CNN and develop a deeper GCN with more than two layers to improve classification results.

D. Graph Isomorphism Network

GNNs proved themselves to be a powerful tool when it comes to capturing and representing neighboring schemes that will aid in ML tasks such as Classification. One special type of GNNs, the Isomorphism Networks, is a proposed state-of-the-art method when it comes to using GNNs.

Kim et al. [13] elaborated on the efficiency of these graphs. GINs are spatial-based convolutional GNNs considered as a dual representation CNN where the adjacent matrix is defined as a generalized shift operation. They explain how GINs can enhance learning non-linear mappings. They used these graphs to visualize the important areas of the brain by using mapping methods on trained GINs. They proposed a framework using Brain fMRI data to analyze the connectivity of the map to classifying the gender of the patients. They work reported an accuracy of 80.4% and an 84% precision after using a 20% sparsity of connected edges. They compared their results to a non-linear SVM method that achieved 73.82% accuracy, relatively less than their proposed method.

Saha et al. [14] proved how powerful GINs can be when paired with proper preprocessing steps. They used X-Rays and CT Scans to early detect COVID-19. They mapped the images into a graph classification problem by using edges found in the images. They show how GINs map any two graphs into different embeddings, isomorphic graph needs to be mapped into the same representation. Their pre-processing steps include edge detection, using Prewitt filter, and graph preparation where the edge maps are converted into graphs. Their method achieved 99-100% accuracy for all the different datasets used.

E. Previous Work in Lung Cancer Detection

C.1. Lung Nodule Classification using 2D CT Scans

Song et al. [15] worked on differentiating between benign and malignant lung nodules/tumors to ensure early diagnosis using 2D CT scans. They built three neural networks Models and compared their performance: CNN, DNN, and SAE (stacked autoencoder) The results showed that CNN achieved the highest accuracy(84.15%), sensitivity(83.96%) and specificity (84.32%) as opposed to DNN and SAE who performed poorly in precision and sensitivity. The next step is to work on improving the accuracy as well as the generalization of the task and apply the method using more advanced medical imaging tasks.

C.2. Lung Nodule Classification using 3D CNN

Zhang et al. [16] built a deep CNN to detect lung cancer nodules from 3D CT scans. The interesting part was at the preprocessing stage where image regions only containing lung tissue were isolated from the rest of the image in each CT slice. This was done by setting different threshold values of HU (Hounsfield unit) that can eliminate elements such as bones and air from the images. Using a 3D CNN, this paper managed to achieve higher accuracy, 83%, than radiologist assessment under efficiency circumstances. As for the limitations, it stated that it was constrained by not only the single source of inputs for both training and validation, but by the size of the dataset as well. Thus, key features might have been missed.

Jin et al. [17] addressed the impact of input volume size compared to accuracy when detecting early stages of lung cancer using deep whole volume features on 3D CT Scans. They achieved a maximum accuracy of 87.5% when the input volume was (128, 128, 20) in size. They also deduced that increasing the number of convolutional layers generated more complex feature maps which lead to increasing the overall score. The limits faced by the paper occurred when incorporating nodule features. I.e. Falsely labeling cancerous lungs with tiny nodules as non-cancerous, and non- cancerous lungs without any nodules as cancerous, leading to a significant false positive and false negative rates.

F. State of the Art

With the increase of non-Euclidian application domains, models like CNNs face several challenges regarding interdependencies between objects. This complexity level is solved using graph networks that extract the relevant features between nodes/edges. GNNs, specifically GINs, are currently under the spotlight in the deep learning world and are not leaving any time soon. With major performance advantages, these state-of-the-art models and their variants are at the top of the chain.

As Part E showed, the current base state-of-the-art algorithms used for this specific detection is a 3D CNN model based on CT-scans.

Some of the most relevant work is was done by Alakwaa's team [18]. They built a deep convolutional neural network to detect lung nodules in patients' CT scans, detecting the interest points employing U-Net architecture. Their work was conducted on the Kaggle Data Bowl 2017 dataset, which is a subset of the data we plan on using LIDC-IDRI. The output is fed into a 3D CNN classifier for binary classification. The preprocessing steps include segmentation (using U-Net) and nodule detection for each CT scan slice so that the 2D slices are stacked into a single 3D image. This was done to limit the searching space for the model since it was observed that the model performed poorly when fed with the whole images. Following that, we have normalization, down sampling, and zero centering steps that preceded the convolution layers. The model achieved 84.4% accuracy, with false positives of 11.9%.

Another notable work was done by Zheng et al. [19] where they propose a multi-planar lung nodule detection system that uses CNNs. Their model features the more advanced techniques proposed such as 2D U-Net++ and Efficient-Nets. They worked on the LIDC-

IDRI dataset. Their model consists of two steps where the first one does multi-planar nodule detection and the second works on false positive reduction. They implement U-Net++ for the first step where they would detect nodules on all sections of the planes. For the encoder, they used Efficient-Net (B4) classification (pre-trained on ImageNet) for feature extraction and classification. The obtained results were then fed into a 3F multi-scale dense training to reduce false positive rates. They achieved 89.3% sensitivity since their purpose was to reduce false negative rates.

The more advanced techniques using GIN that Saha et al. [14] implemented in detecting other diseases performed very well. Thus, we believe that applying similar techniques to detect lung cancer looked to be very promising. By providing locality, flexibility, shift-invariance, and compositionality, GINs extract this innate prior in imagery that other models fail to do.

What we propose in this paper is a GIN structure that takes as inputs segmented 3D CT-Scans slices. The model will implement sophisticated preprocessing techniques and will employ U-Net++ for the segmentation.

G. Qualitative Comparison

Although lung cancer detection has been a field that has been worked on, there were no works that achieved an impressive accuracy level which implies there still remains room for improvements. Based on works done using new, improved methods, we built a table to show the topics will be tackling when resolving our problem and compare it to all recent relevant works:

Topics	Segmenta-tion	U-Net / U-Net++	GNN	Dimensio-nality (3D)	Lung Cancer Detection
[8]	✓	✓	✗	✗	✓
[12]	✓	✗	✓	✗	✗
[14]	✗	✗	✓	✗	✗
[16]	✓	✗	✗	✓	✓
[18]	✓	✓	✗	✓	✓
[19]	✓	✓	✗	✓	✓
Our Work	✓	✓	✓	✓	✓

Figure 5. Comparative Table of previous work and our proposed contributions.

III. PROPOSED METHOD

The proposed method consists of three main steps, a pre-processing stage, a segmentation stage, and a classifier.

- *Pre-Processing Stage*

Starting with the pre-processing step, we input a patient's 3D CT-scan, composed of 100 to 200 2D DICOM (Digital Imaging and Communications in Medicine) slices stacked on top of each other. Each slice is taken through a pre-processing step that segments the lung regions and saves the images in addition to the corresponding mask images. Segmenting the lung regions corresponds to extracting only the lung tissue from the DICOM image and filtering out other components (bones, air…), reducing the searching area for the model. This is done by using a k-means clustering and morphological operations, sourced from different python libraries including the specialized pylidc library designed for our specific dataset, skimage, and scipy. We end up with two folders, one with the patients that are cancerous (have mask images) and another "Clean" folder with patients who are not cancerous (do not have mask images). These datapoints are split into train, test, and validation sub-folders.

Figure 6. One of our raw images, with the segmentation of the lungs, and the actual mask (Plotted on python)

- *Segmentation Stage*

The next step is to perform nodule segmentation on the pre-processed slices in order to identify areas that might include cancerous nodules, further reducing the searching area for the model. We do that by employing U-Net++, a segmentation algorithm that is based on a CNN architecture (convolutional neural network) and represents an improved version of the original U-Net.

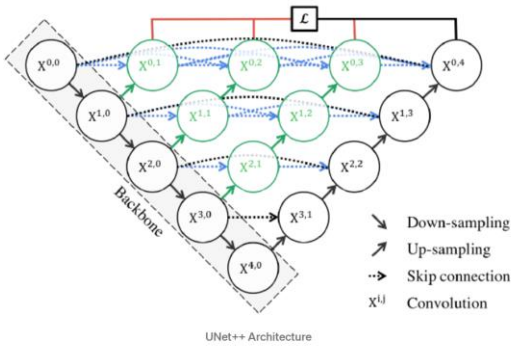


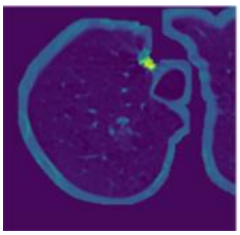
Figure 7. U-Net++ Architecture, Image from [7]

As in Figure 7, U-Net++ begins with an encoder subnetwork followed by a decoder subnetwork. The two subnetworks are connected using skip pathways and the use of deep supervision [7]. We train this network on our train dataset and evaluate it using the validation and test datasets. The output of the network is a mask image with the predicted areas of interest. To evaluate our network, we compare the original mask for each image to the predicted mask. Figure 8 illustrates an example:

Figure 8. Predicted Mask of our previous example(Plotted on python)

- *Graph Classification*

The next step is to prepare the segmented nodules to be fed into our Graph classifier. We start by cropping the areas of interest according to the output from the U-Net++ segmentation. For training, we label these cropped images as cancerous and non-cancerous.



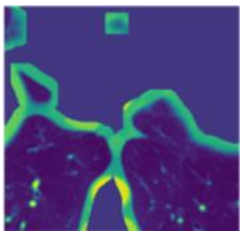


Figure 9. Nodule Segmentation for two example using U-Net++(Plotted on python)

We follow by transforming these images into graph structures, since the GIN needs a graph as input. This is done in two steps, first we detect the edges in each image, and then build the graph based on the edges. For edge detection, we transform our images into grayscale, and then apply a 3x3 Prewitt Filter to detect changes in intensity that denote an edge. If the intensity of the pixel is higher or equal than 128, then it is considered an edge. To build the graph, we consider each of those pixels with a grayscale intensity greater or equal to 128 as a node in the graph, with its grayscale intensity as node feature. Edges will be defined between the nodes of adjacent pixels. It is important to note here that each graph is built from a single image. By only considering image edges as nodes, the size of our dataset will be reduced significantly, substantially speeding up the training and testing process.

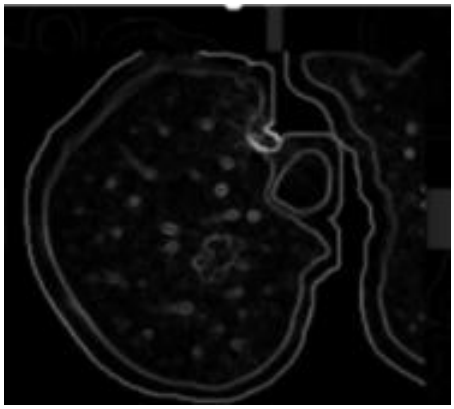


Figure 10. Greyscale image of example for Edge Detection(Plotted on python)

For our classifier, we employ a state-of-art GIN (Graph Isomorphism Network) that performs graph classification i.e., classifies each graph (in our case each graph represents an image of an area of interest) into cancerous or non-cancerous, or in other words whether this area contains cancerous nodules or not. The GIN we implement will creating embeddings for the nodes of the graph. Neighboring nodes will be encoded in a way where they are close in the embedding space. This is at the core of isomorphism. GIN follows the conditions set by the Weisfeiler-Lehman theorem: the embedding of graphs aggregates node features iteratively. This mapping must be injective (one-to-one mapping). Therefore, isomorphic graphs must share the same embeddings while non-isomorphic graphs do not.

Consequently, the model performs well on heavily imbalanced datasets.

The architecture proposed is the following. The images are transformed into graphs as mentioned above – graphs are fed to a sequence of Linear, ReLU, Linear which we will call 'GIN layer' – features extracted are fed into a ReLU, dropout and Normalization layers – output fed to another 'GIN layer' followed by the same three layers – new output fed into a third 'GIN layer' followed by ReLU, dropout, Mean-pooling, Linear, dropout and Linear layers – fully connected layer fed to logsoftmax for classification. Logsoftmax (Equation [1]) helps performance by better optimization of the gradient by heavily penalizing a model for an erroneous prediction [14].

$$logsoftmax(x_i) = log(e^{X_i} / \sum_j e^{X_j}) \quad [1]$$

Figure 11 and 12 summarizes the GIN structure.

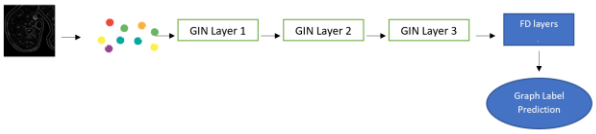


Figure 11. GIN structure summary

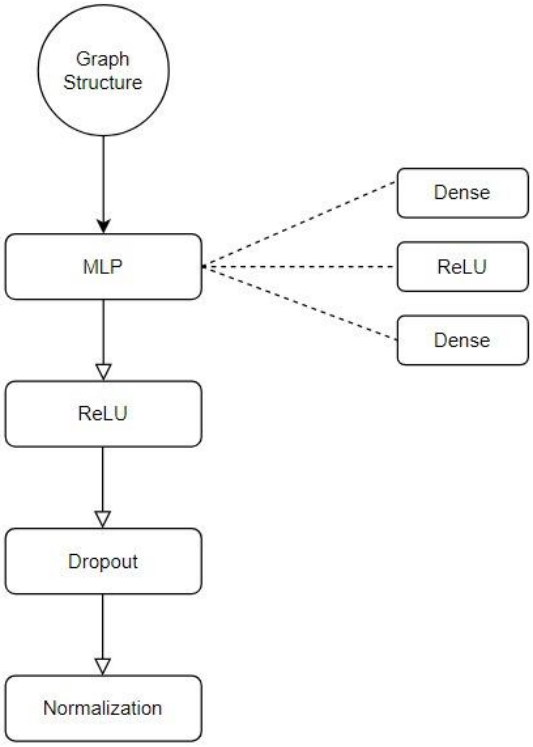


Figure 12. GIN layer architecture

Finally, we summarize the entirety of our proposed method in Figure 13.

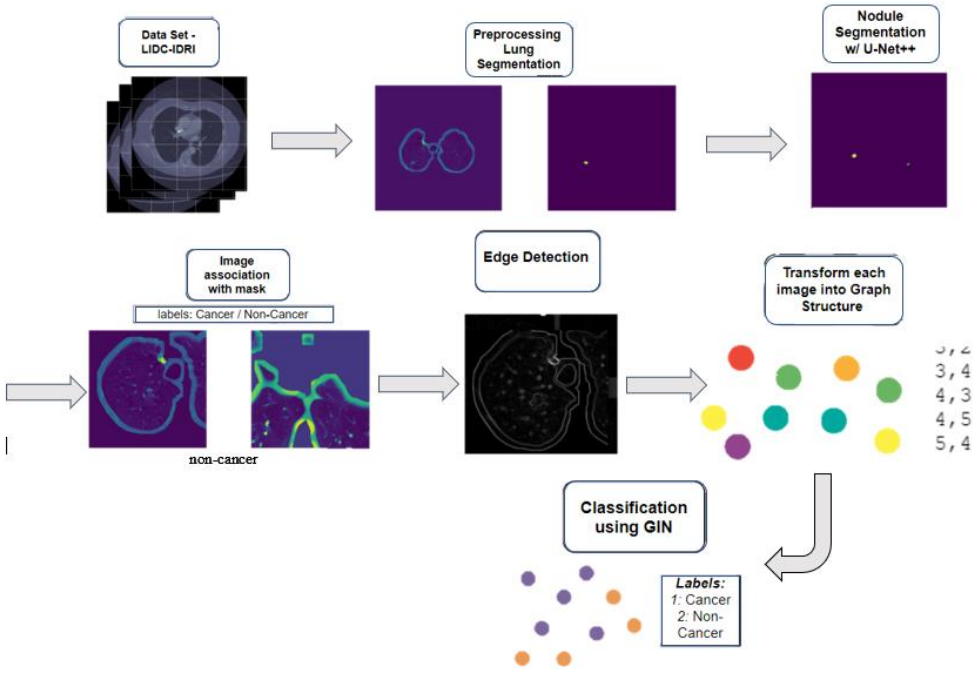


Figure 13. Summary of Proposed Method

IV. DATASET

This work will rely on the LIDC-IDRI (Lung Image Database Consortium image collection) database[4]. It contains 244,527 images of the 1010 cases. It includes 5 different image modalities: CT (Computer Tomography), DX (Digital Radiography- Digital X-Ray), CR (Computed Radiography), SR (SR Document), and SEG (Segmentation). It contains complete CT scans that form a 3D image for each patient, and accompanied by an XML file which records the annotations or segmentations of analysis conducted on the subjects by radiologists depending on the thickness of the spots or "lung nodules" on their lungs. This file will classify patients into cancerous/non-cancerous based on the professional's results. Other XML files also are downloaded which record the number of nodules and diagnosis of each patient. Our experiments were conducted on a total of 60 of these patients. Every patient had around 125 DICOM files each containing a CT Scan slice which amount to over 7500 slices as inputs.

V. EXPERIMENTS AND RESULTS

- *Planned Experiments*

As a first experiment, we would like to test out diverse segmentation algorithms on typical medical imagery. UNET and its variants are currently considered as state-of-the art networks in such a field. UNET itself has less parameters than its own variant UNET++, which is an important attribute to test on our dataset regarding accuracy and runtime efficiency.

Additionally, we will experiment with feeding the entirety of a 3D CT-scan of a patient as one form of input. Depending on the resulting outcome, we might alter this approach by feeding regions of interest into the model. This can be efficient when using preprocessing models such as U-Net (segmentation).

Next, we plan to set a benchmark CNN model for comparison at the end with our implementation using a GNN. EfficientNet has had recent success in the medical imagery field for AI. Different EfficientNets exist, from B0 to B7, with increasing parameters respectively. We will experiment with the different alterations and opt for the one with fastest training runtime and best accuracy result.

Finally, we ought to implement a GNN model that emphasizes neighboring relationships between nodes as opposed to the traditional image-level features extracted by CNNs. Graph similarities is a great method used to classify graphs. The term used when comparing similar graphs is Isomorphism. Thus we plan to transform the images into corresponding graphs and classify the latter accordingly.

- *Overview of target experiments and their objectives (experiment setup)*

As described in the section above, the target experiments we are aiming for are the following: Comparing the results of diverse state-of-the-art segmentation techniques such as UNET and UNET++, identify a benchmark model consisting of a CNN such as EfficientNet for comparison with our own model, and finally experiment with the GIN architecture for optimal results.

- *Details of experiments with analysis of results'*

- **Experiment 1**

For the first experiment, we contested UNET, a SOTA segmentation network for medical imagery that uses data augmentation with elastic deformations consisting of 7.76M trainable parameters, against its own variant UNET++ that has 9.04M learnable parameters. UNET solves the issue of small datasets with its augmentation techniques as well as reduces training time drastically. While UNET++, a more recent enhancement, reduces the semantic gaps between feature maps by connecting the encoder and decoder sub-networks using skip pathways. This architecture however opts for a higher number of parameters which increases the training time.

Since our dataset did not require immense training time due its relatively smaller size, UNET++ achieved better accuracy than UNET overall and was chosen as the segmentation algorithm to be used in the final model.

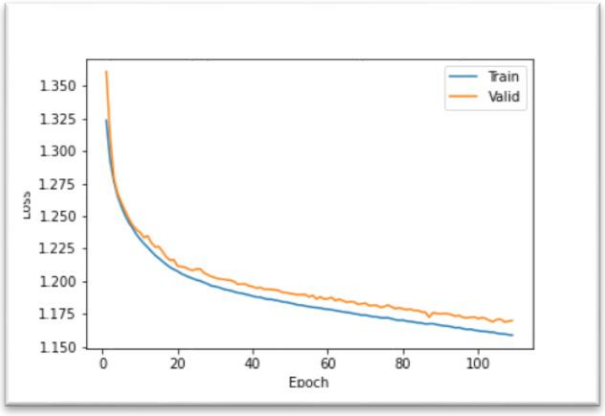


Figure 14. Our U-Net++ Loss Function in terms of epochs

After conducting our experiments, we notice that the results obtained using the techniques were slightly different:

	TP	FP	TN	FN	FP/scan
U-Net++	14	108	79	24	30.33
U-Net	10	1400	15	28	466.67

Figure 15. U-Net++ vs U-Net experimental results using metrics

- **Experiment 2**

As for the second experiment, we opted for a benchmark CNN model that is going to be used for comparison against our GNN. EfficientNet and its recent success in the biomedical field was the clever decision to make. This implementation is a new method of uniformly scaling the image width, height and resolution using a compound coefficient to achieve better accuracy. However, its main advantage is that it has orders of magnitude fewer parameters than other top CNNs. As mentioned above, our dataset does not require a very high number of parameters to train. B7 resulted in a training accuracy of 93% with test accuracy of 57% showcasing it has overfit the data with its 66M parameters. While B0 achieved overall better results: 97% train accuracy and 70% test accuracy on our dataset.

This result was expected as lung cancer dataset are known for their major imbalances. This model however performed better thanks to the data augmentation techniques that were done prior, solving this imbalance greatly. This accuracy is going to be contested with our own implementation using GIN.

- **Experiment 3**

For our last experiment, we built a Graph Isomorphism Network (GIN), as it has shown tremendous success in image classification tasks. Implementing it into our pipeline, GIN came at the end of our architecture, took the preprocessed segmented images (by UNET++) as input, and performed a binary classification task for labels: Cancerous or Non-Cancerous.

The problem of isomorphism can now be tackled, as isomorphic graphs have the same embedding, while others do not. Therefore, this model works perfectly for heavily imbalanced datasets such as ours. The model resulted in a perfect 100% accuracy. This is due many factors; the model properly segmented and preprocessed images which made the network able to properly understand the input graphs. Inputs are fed into sequential layers consisting of: linear, ReLU and another linear. The simplistic nature of the sequence of layers improves generalization. The introduction of ReLU layers introduced non-linearity and reduced the likelihood of vanishing gradients. Dropout regularization was additionally used to prevent overfitting by randomly ignoring neurons. This resulted in a reduced loss for each of the 10-fold validations.

Transforming the images into graphs after edge detection led to a total of 1089386 nodes for a total of 1058 areas of interest that might contain cancerous nodes solving the problem of data imbalance. Overfitting was reduced significantly again for the same reason. The confusion matrix plotted in Figure 16 portrays the results obtained with no false positives nor false negatives. We also plotted, in Figure 17, the ROC curve for verification purposes as well as the train/test accuracy vs. train-to-test ratio graphs

.

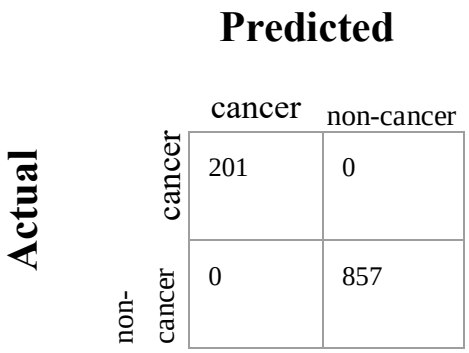


Figure 16. Confusion Matrix for our last experiment using GIN

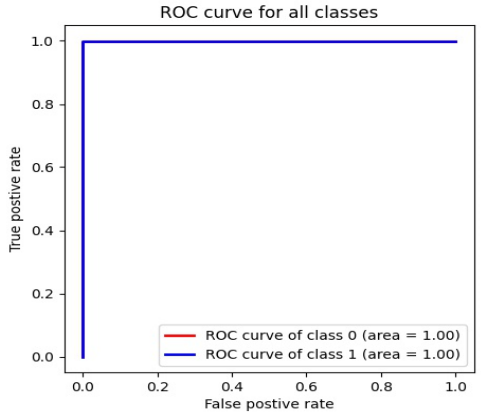


Figure 17. ROC curve for our last experiment

Our model converged using 10 epochs only. This showcases good optimization for the binary classification task.

❖ **Challenges Solved**

Our model showcased robustness and consistency that lacked in other networks such as CNNs. CNNs emphasized image-level features while our model put neighboring-level features and interconnections under the spotlight. Additionally, our benchmark model, EfficientNet, took orders of magnitude higher time to train and test than the GIN implementation.

The architecture presented solved the typical class imbalance problem that arises when working with lung cancer CT scans by transforming them into graphs with nodes and edges and classifying the graphs. Classification was based on the similarity or isomorphism of graph embeddings. This also resolved the problem of high training and testing runtime since dealing with nodes is much less complex than dealing with entire feature maps. In addition, data augmentation techniques through UNET++ helped reduce overfitting of the model by increasing the number of cancerous images. Our result abolished the false positive rate that was present in traditional CNNs and achieved perfect score.

VI. CONCLUSION

In conclusion, we have a presented a Lung Cancer Detection method that relies on Graph Isomorphism Networks (GIN) to perform nodule classification. The method includes multiple preprocessing and segmentation steps including nodule segmentation using U-Net++. The GIN algorithm maps isomorphic graphs with the same embedding and leverages the neighbor-level relations in the graphs to offset the effects of an imbalanced dataset and avoid overfitting. It also allows for a short training time due to the fact that only edges in cropped lung images were considered to build the graphs inputted to the GIN, significantly reducing the data load on memory. We experimentally validated the proposed model and showed that it outperforms all state-of-art methods when considering both accuracy and training time. Results showed that the model detects lung cancer with 100% accuracy on our test, far outperforming our state-of-art benchmark. Future work includes expanding our dataset to better generalize our method and extending its use to other medical-imagery related applications.

VII. REFERENCES

[1] "What Causes Lung Cancer," American Lung Association. [Online]. Available: https://www.lung.org/lung-health-diseases/lung-disease-lookup/lung-cancer/learn-about-lung-cancer/what-causes-lung-cancer#:~:text=Smoking%20is%20the%20number%20one,do%20for%20your%20lung%20health. [Accessed: 28-Feb-2021].

[2] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," American Cancer Society Journals, 04-Feb-2021. [Online]. Available: https://acsjournals.onlinelibrary.wiley.com/doi/full/10.3322/caac.21660. [Accessed: 28-Feb-2021].

[3] "Lung cancer," Mayo Clinic, 10-Oct-2020. [Online]. Available: https://www.mayoclinic.org/diseases-

conditions/lung-cancer/diagnosis-treatment/drc-20374627. [Accessed: 28-Feb-2021].

[4]LIDC-IDRI, https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI.

[5] K. J. Aditi Kulkarni, “Lung Cancer Detection Using Deep Convolutional Neural Network,” International Journal of Future Generation Communication and Networking. [Online]. Available: https://sersc.org/journals/index.php/IJFGCN/article/view/28246. [Accessed: 28-Feb-2021].

[6] K.-H. Yu, T.-L. M. Lee, M.-H. Yen, S. C. Kou, B. Rosen, J.-H. Chiang, and I. S. Kohane, “Reproducible Machine Learning Methods for Lung Cancer Detection Using Computed Tomography Images: Algorithm Development and Validation,” Journal of medical Internet research, 05-Aug-2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7439139/. [Accessed: 28-Feb-2021].

[7] S.-H. Tsang, “Review: UNet++-A Nested U-Net Architecture (Biomedical Image Segmentation),” Medium, 01-Oct-2019. [Online]. Available: https://sh-tsang.medium.com/review-unet-a-nested-u-net-architecture-biomedical-image-segmentation-57be56859b20. [Accessed: 27-Apr-2021].

[8] Zhou, Z., Siddiquee, M., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support : 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, S..., 11045, 3–11. https://doi.org/10.1007/978-3-030-00889-5_1

[9] Tan, M., & Le, Q.V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ArXiv, abs/1905.11946.

[10] Yu, Xiang & Lu, Siyuan & Guo, Lili & Wang, Shui-Hua & Zhang, Yu-Dong, ResGNet-C: A graph convolutional neural network for detection of COVID-19, Neurocomputing, 2020, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2020.07.144.

(https://www.sciencedirect.com/science/article/pii/S0925231220319184)

[11] Zhou, Yanning & Graham, Simon & Koohbanani, Navid & Shaban, Muhammad & Heng, Pheng-Ann & Rajpoot, Nasir. (2019). CGC-Net: Cell Graph Convolutional Network for Grading of Colorectal Cancer Histology Images. 388-398. 10.1109/ICCVW.2019.00050.

[12] Zhang, Yu-Dong & Satapathy, Suresh Chandra & Guttery, David S. & Górriz, Juan Manuel & Wang, Shui-Hua Improved Breast Cancer Classification Through Combining Graph Convolutional Network and Convolutional Neural Network, Information Processing & Management, Volume 58, Issue 2, 2021, 102439, ISSN 0306-4573, https://doi.org/10.1016/j.ipm.2020.102439.

(https://www.sciencedirect.com/science/article/pii/S0306457320309328)

[13] Kim, B. H., & Ye, J. C. (2020). Understanding Graph Isomorphism Network for rs-fMRI Functional Connectivity Analysis. Frontiers in neuroscience, 14, 630. https://doi.org/10.3389/fnins.2020.00630

[14] Saha, P., Mukherjee, D., Singh, P.K. et al. GraphCovidNet: A graph neural network based model for detecting COVID-19 from CT scans and X-rays of chest. Sci Rep 11, 8304 (2021). https://doi.org/10.1038/s41598-021-87523-1

[15] Song, QingZeng et al. “Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images.” Journal of healthcare engineering vol. 2017 (2017): 8314740. doi:10.1155/2017/8314740

[16] Zhang, Chao et al. “Toward an Expert Level of Lung Cancer Detection and Classification Using a Deep Convolutional Neural Network.” The oncologist vol. 24,9 (2019): 1159-1165. doi:10.1634/theoncologist.2018-0908

[17] T. Jin, H. Cui, S. Zeng and X. Wang, "Learning Deep Spatial Lung Features by 3D Convolutional Neural Network for Early Cancer Detection," 2017 International

Conference on Digital Image Computing: Techniques and Applications (DICTA), Sydney, NSW, Australia, 2017, pp. 1-6, doi: 10.1109/DICTA.2017.8227454.

[18] Alakwaa, Wafaa & Nassef, Mohammad & Badr, Amr. (2017). Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN). International Journal of Advanced Computer Science and Applications. 8. 10.14569/IJACSA.2017.080853.

[19] Zheng, S., Cornelissen, L. J., Cui, X., Jing, X., Veldhuis, R., Oudkerk, M., & van Ooijen, P. (2021). Deep convolutional neural networks for multiplanar lung nodule detection: Improvement in small nodule identification. Medical physics, 48(2), 733–744. https://doi.org/10.1002/mp.14648

[20] Alam N-A-, Ahsan M, Based MA, Haider J, Kowalski M. COVID-19 Detection from Chest X-ray Images Using Feature Fusion and Deep Learning. Sensors. 2021; 21(4):1480. https://doi.org/10.3390/s21041480

[21] Hope, M. D. et al. A role for CT in COVID-19? What data really tell us so far. Lancet (London, England) 395, 1189–1190 (2020).