Ira A. Fulton School of Engineering

Lung Cancer Detection Utilizing CT-Scan Images

Jay Jayani (1226668926)

Taraksinh Raulji (1222309972)

Tarpan Kothari (1223365767)

1. INTRODUCTION

With a mortality incidence of 1 in 6 persons in 2020, lung cancer will continue to be a leading cause of death globally. Sadly, the early detection of lung cancer symptoms results in a significant mortality rate compared to other cancers. Medical experts and researchers are working hard to advance the complicated and difficult process of early identification of lung cancer. Smoking, exposure to radon gas or secondhand smoke, and previous radiation therapy are all common risk factors for lung cancer. Long-lasting bloody coughs, discomfort in the chest and bones, shortness of breath, unexplained weight loss, and severe headaches are typical signs of advanced lung cancer. Lung cancer is often divided into four phases by medical professionals.

2. Motivation

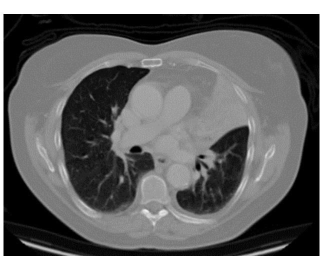
Nowadays, making an early diagnosis of lung cancer is challenging, and there is a chance of making a mistake. The authors of employed a multi-class SVM Classifier on CT scans to solve this problem and successfully detect lung cancer in its early stages. This strategy leveraging machine learning (ML) has demonstrated promise in resolving the problem of misclassification linked to conventional diagnostic testing. The use of ML algorithms should be seen as a tool to help and improve the diagnosis process rather than as a replacement for qualified medical personnel, it is crucial to remember.

3. METHODOLOGY

The description of the dataset used as the first step in this section's information about the proposed model's design. Pre-processing is required for data augmentation and to boost the model's performance because the dataset only has a few samples available for each category. In the same section, the rationale for creating the suggested model is explained. On top of that, the holdout dataset performance of the proposed model is assessed.

3.1 DATASET DESCRIPTION

The Kaggle dataset in the png/jpg format provided the CT scan dataset that was used in this analysis. There are four categories and a total of 1000 photos with a 224x224 pixel resolution are included. These categories are divided into three groups that represent various forms of lung cancer, and the fourth group includes typical CT scans. The following is an explanation of each category:



Adenocarcinoma

Large Cell

Squamous Cell

- Adenocarcinoma is the most common form of lung cancer and accounts for 30% of all cases, as well as 40% of non-small cell lung cancer cases.
- Large-cell undifferentiated carcinoma lung cancer is a highly aggressive type that grows and spreads quickly, and it can be found anywhere in the lung.
- Squamous cell carcinoma is typically found in the central region of the lung, where the larger bronchi join the trachea to the lung, or in one of the main airway branches.
- Normal cell category contains CT images with no evidence of lung cancer.

Category	Training	Validation	Testing	
Adenocarcinoma	195	23	120	338
Large Cell	115	21	51	187
Squamous Cell	155	15	90	260
Normal	148	13	54	215
	613	**72**	**315**	**1000**

Table 1: Dataset Split category-wise

3.2 PRE-PROCESSING

Pre-processing entails converting the raw input data into a suitable format that the model can utilize for training and testing. It is a crucial stage in any machine-learning effort. Pre-processing for CT scan pictures entails data augmentation and normalization. By generating additional data samples from the current ones, a technique known as data augmentation is utilized to enhance the size of the training data set. This is accomplished by subjecting the photos to random changes, such as rotating, shearing, zooming, flipping, etc. To carry out these modifications, one common Keras tool is the ImageDataGenerator module.

In the present case, we used the following hyper-parameters for data augmentation:

Augmentation Parameters	**Value**	**Descriptions**
rotation_range	15	Randomly rotate the images by up to 15 degrees
shear_range	0.2	Apply random shearing transformations to the images
zoom_range	0.2	Randomly zoom in or out of the images
horizontal_flip	True	Randomly flip the images horizontally
fill_mode	nearest	Fill any gaps that may appear in the images after transformation with the nearest pixel value.
width_shift_range	0.1	Randomly shift the images horizontally by up to 10% of their width.
height_shift_range	0.1	Randomly shift the images vertically by up to 10% of their height.

Normalization is another important pre-processing step that involves scaling the pixel values of the images to a range that is suitable for the model. In the present case, the authors have divided the pixel values by 255 to rescale them to the range [0,1]. This is a common normalization technique used in computer vision tasks.

3.3 MODEL MOTIVATION

In the preceding section, the authors conducted a literature study and discovered that complicated convolutional neural network (CNN) models have shown promising outcomes for identifying lung cancer. They researched many CNN models and evaluated their performance on the limited dataset, including DenseNet, MobileNet, ResNet, EfficientNet, VGG, Inception, and Xception.

Model	Loss	Accuracy	Recall	Precision
DenseNet169	**0.547894**	**0.853968**	0.93968	0.45963
DenseNet201	0.759781	0.765079	0.87619	0.57381
EfficientNetB4	0.995534	0.55873	0.82857	0.4579
InceptionV3	0.624704	0.793651	**0.94286**	0.44461
MobileNet	0.559497	0.825391	0.93651	**0.65121**
ResNet50	1.123878	0.514286	0.6	0.32643
VGG19	0.706449	0.72381	0.77778	0.58612
Xception	0.828387	0.669841	0.75238	0.59102

Table 2: Models' performance on the testing data

Here, we acknowledged the need for a classification model with high accuracy and recall in the area. After that we developed the models described in Table 2 with the structures shown in Figure 1 in order to accomplish this aim.

It appears that a study's authors trained a number of deep learning models on a training dataset and assessed their effectiveness on a testing dataset. They discovered that MobileNet and DenseNet169 both performed well on the provided dataset, with MobileNet displaying the best precision and DenseNet169 the best accuracy and loss values. Despite having the highest recall values, InceptionV3 underperformed MobileNet and DenseNet169 overall.

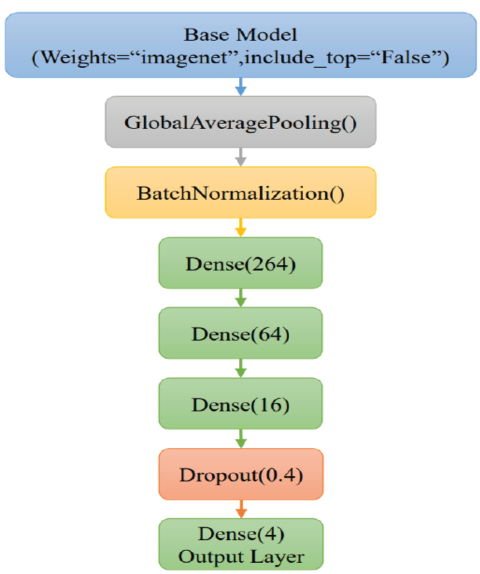


Figure 1: Architecture of the models

These results served as the basis for the creation of an ensemble model of MobileNet and DenseNet169 that outperformed them both separately. This is a common deep-learning technique where many models are combined to improve overall performance.

It's vital to understand that these findings are unique to the dataset utilized in the study and might not apply to other datasets. Also, the selection of assessment measures and the particular hyperparameters employed might have an impact on the outcomes. Before making any judgments, it is crucial to thoroughly assess how well deep learning models perform over a range of metrics and datasets.

4. PROPOSED CLASSIFICATION MODEL

DenseNet169 and MobileNet, two deep learning models, have been combined into an ensemble model to achieve superior performance. Using "imagenet" weights as its initialization, the option "include top" is set to False. The output of both models passes through two dense layers with 16 and 8 neurons each, a GlobalMaxPooling2D layer, a Flatten layer with a 20% dropout ratio, and a Dropout layer before being output. In order to determine the likelihood of each class, the outputs are then concatenated and sent into a final output layer of four neurons. The categorical cross-entropy loss function, Adam optimizer, accuracy, precision, and recall metrics are used in the model, which employs CT scan pictures as inputs. Early callbacks are also included in the model to minimize computation.

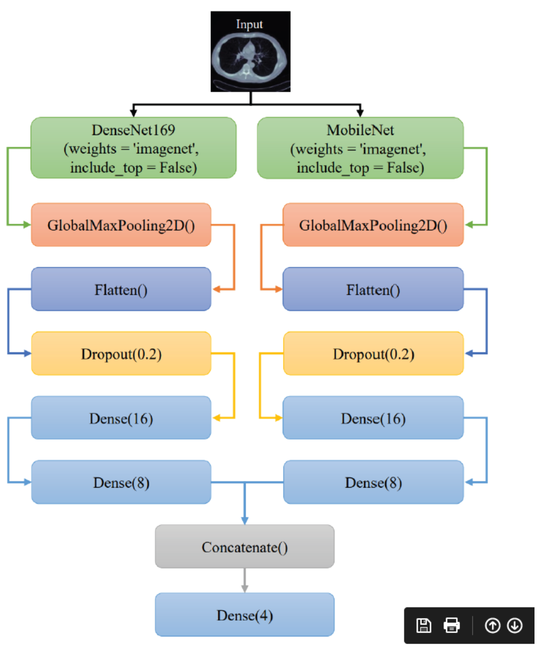


Figure 2: Proposed model

The performance metrics that were utilized to evaluate both the base models and the suggested arrangement are shown in Table 3. The table also includes information on the experiment's numerous hyperparameters. Setting up the experiment environment using the parameters from the respective column will allow you to reproduce the results shown in the table.

Parameters	**Values**
Programming Language	Python 3.8.0
Platform	Google Colab
Framework	TensorFlow
Batch size	8
Epochs	50
Optimizer	Adam
Loss	Categorical cross-entropy
Metrics	Accuracy, Recall, Precision
ModelCheckpoint	monitor = Validation Accuracy

Table 3: Performance Parameters

5. RESULTS

Using a dataset of lung cancer CT scans, the performance of the proposed model is assessed in this section using several base models. We examined the loss function, accuracy, recall, and precision values on the test dataset, which is the holdout dataset, to evaluate the performance of the suggested model. With the specified aims we decided to optimize the recall measure. They trained the model on the training dataset and verified it on the validation dataset to achieve this. In Figure 3, throughout 50 epochs, the training and validation datasets loss values and accuracy values are compared. For loss and accuracy on the corresponding subgraphs, the blue dot in Figure 3 reflects the best results obtained on the validation dataset. For the validation dataset, the greatest accuracy and lowest loss values were both 0.3280 and 91.67%, respectively.

Model	Loss	Accuracy	Recall	Precision
DenseNet169	0.54894	0.853968	0.93968	0.45963
MobileNet	0.559497	0.825397	0.93651	0.65121
InceptionV3	0.624704	0.793651	0.94286	0.44461
Proposed	**0.344487**	**0.873015**	**1**	0.58558

Table 4: Training-validation accuracy-loss comparison for the proposed model

The suggested model and the top-performing base models are contrasted in table 4. For the test dataset, the suggested model recorded the lowest loss value, which was almost 30% lower than the lonely-DenseNet169 base model. Moreover, the test dataset's total accuracy increased from 85.3968% to 87.3015%. Although the accuracy value stayed constant at around 0.58, the suggested model also had a perfect recall value of 1.0. In addition, the suggested model required more time to train than the basic models (15 seconds on average for each epoch).

Figure 3: Training-validation accuracy-loss comparison for the proposed model

While the suggested model was twice as large as the basic models, it was still possible to finish the training in 18 seconds for each epoch, which was seen to be a promising outcome. A recall is the most crucial performance statistic for a model in the medical field when it comes to identifying lung cancer in CT images. This is because it is more important to accurately identify all positive cases (lung cancer), even if it means labeling some negative instances (normal) as positive, since this may be further evaluated in later clinical analysis. While misclassifying a normal CT scan as lung cancer can be rectified later, misclassifying a lung cancer CT scan as normal can have serious ramifications for the patient's life. As a result, while accuracy and precision are less significant in this area, boosting recall values is essential.

6. CRITICIZATION

From this project, we found that improvement can be made in certain areas for the proposed model we got a good precision value, but we can get a better precision value by testing different models also, we can improve the precision by increasing the number of training epochs (use more resources). We can also improve the model of architecture to get a better value of the precision and hence can minimize the loss.

7. CONCLUSION

For the purpose of identifying lung cancer in CT scan pictures, the scientists used deep learning models. Even though they only tested their suggested model on a limited sample size of 1000 photos, it outperformed other models. We employed data augmentation during preprocessing to get around the problem of the tiny dataset. They did a thorough literature analysis of earlier techniques for lung cancer categorization using CT images before putting out their model. They also included a brief explanation of the dataset that was utilized and the preparation procedures that were carried out on it. Ultimately, using the architecture depicted in Figure 1, we trained fundamental models. After assessing how well each model performed on the test dataset, we decided to combine two models (DenseNet169 and MobileNet) to create their proposed model. Using callbacks, the suggested model was trained across 50 epochs. The proposed model performed better than the fundamental models separately in terms of recall, accuracy, and loss, according to the results.

REFERENCES:

1. A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbor classification
2. Lung cancer screening with low-dose CT scans using a deep learning approach.
3. Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN)
4. Automatic Lung Cancer Prediction for Chest X-ray Images Using the Deep Learning Approach
5. Chest CT-Scan images Dataset
6. WHO. 2022-11-15. WHO WHO-CancerJanee Alam, Sabrina Alam, and Alamgir Hossan. 2018. Multi- Stage Lung Cancer Detection and Prediction Using Multi-class SVM Classifie