An Intelligent Lung Cancer Prediction System Using Efficient Net B7, ResNet, and VGG 16

Satyaki Mandal, Ankan Chakraborty

Abstract – *Lung cancer is one of the most lethal cancers in the world. Lung cancer cannot be avoided, although its risk can be lowered. Early diagnosis, screening, and accurate customized therapy are critical for patient recovery. Histopathological pictures of biopsied tissue from possibly diseased parts of the lungs are used by medical practitioners to make diagnoses. Obtaining an accurate categorization from a biopsy picture is a difficult undertaking since the pathologist must be familiar with the details of both normal and diseased cells. Manually identifying malignant cells from tiny biopsy photos is a time-consuming operation that needs years of skill. Many often, the diagnosis of lung cancer is incorrect and time-consuming. AI technology can help by relieving the burden on busy doctors and identifying patterns and traits that the human eye cannot see. In this article, we present a histopathological technique for predicting tumor histology in non-small cell lung cancer (NSCLC) from invasive standard-of-care biopsy imaging data. We trained and validated convolutional neural networks (CNNs) on 15,000 histopathological images from benign, adenocarcinoma (ADC), and squamous cell carcinoma (SCC) (SCC). The CNNs predicted tumor histology with a 98 percent accuracy (F1 score = 0.97). We also found that combining a stacked ensemble of current pre-trained image classification models such as ResNet, VGGNet, and EfficientNet with a multinomial Logistic regression classifier produced equivalent, if not superior, discriminative performance with an accuracy of above 99 percent. In diverse test sets, our best-performing CNN performed as a robust probabilistic classifier. Deep learning-based radionics may detect histological phenotypes in lung cancer. It has the ability to supplement existing procedures and act as a diagnostic tool for physicians.*

Keywords – lung, cancer, histopathological, diagnosis

1. INTRODUCTION

Cancer is a condition that causes certain cells in the body to grow uncontrolled and spread to other regions of the body. Cancer may begin practically anyplace in the human body, which contains billions of cells. Human cells normally divide and proliferate to generate new cells as needed by the body. Cells die as they age or become damaged, and new cells replace them. This ordered process can sometimes break down, allowing aberrant or damaged cells to grow and reproduce when they should not. Tumor, which are masses of tissue, can grow from these cells. Tumor can be malignant or benign (benign).

Cancerous tumor infiltrate neighbouring tissues and can move to distant parts of the body to generate new tumor (a process called metastasis). Malignant tumor are another name for cancerous tumor. Many cancers create solid tumor, while blood malignancies, such as leukaemia, do not. Benign tumor do not spread or infect surrounding tissues. Benign tumor seldom recur after removal, although malignant ones do. However, benign tumor can grow to be extremely enormous. Some, such as benign brain tumor, can produce significant symptoms or even be fatal.

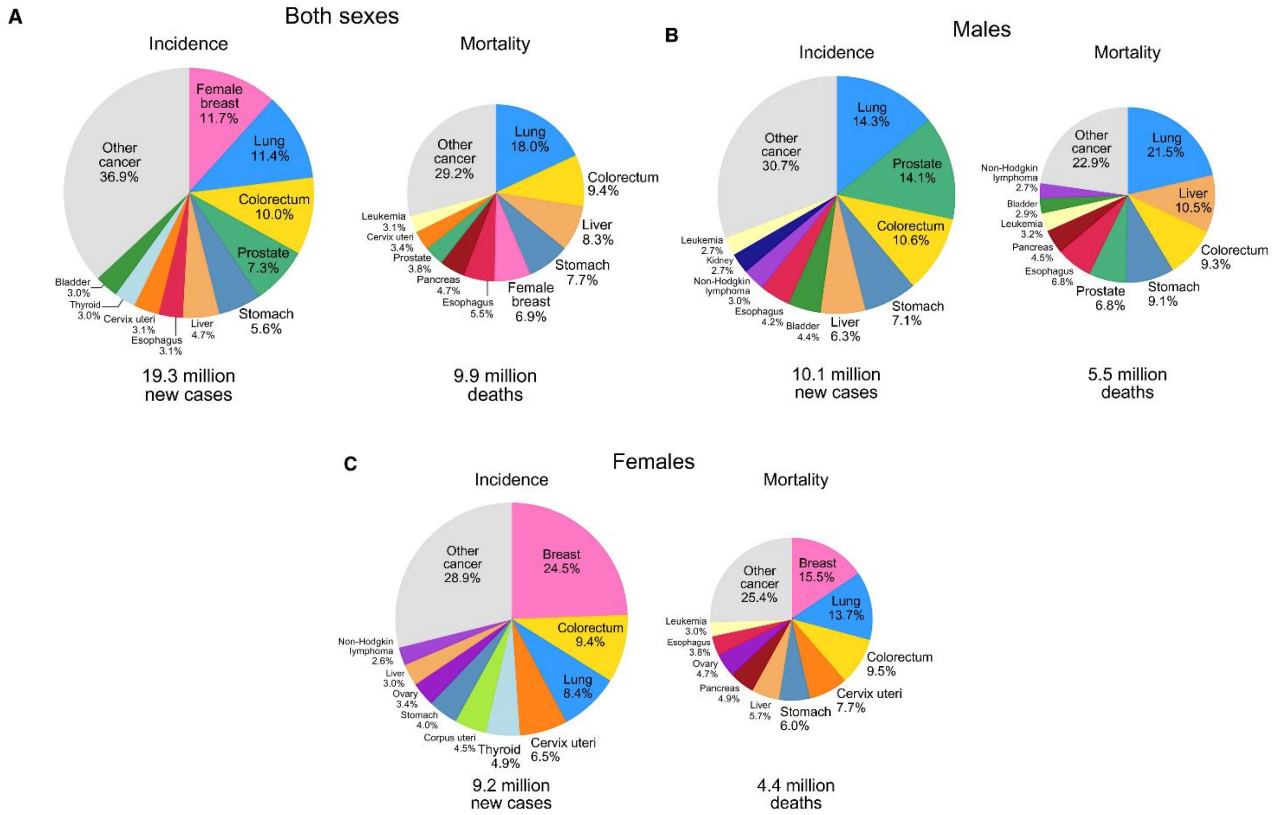
Cancer is a genetic illness, which means that it is caused by abnormalities in the genes that regulate how our cells behave, particularly how they grow and divide. Cancer-causing genetic alterations can occur as a result of:

- Errors that occur during cell division
- Toxic environmental chemicals such as cigarette smoke components and UV radiation from the sun produced DNA damage,
- handed down genetically from our parents.

Normally, the body eliminates damaged DNA cells before they become cancerous. Our bodies' capacity to do so, however, weakens as we age. This increases the chance of cancer later in life. Each person's cancer has its own unique collection of genetic changes. As cancer spreads, further changes will occur. Distinct genetic changes may exist in different cells within the same tumor.

Lung cancer is a leading cause of cancer-related death. The infection might start in the windpipe, bronchial airway, or lungs. It is caused by uncontrolled cell growth and spread from the lungs. People who have pulmonary illness, such as emphysema, and a history of chest pain are more prone to develop lung cancer. Tobacco usage, which includes cigarettes and vapes, is the biggest cause of lung cancer. Even non-smokers can get lung cancer from radon gas, air pollution, and toxins in the workplace, such as asbestos, diesel exhaust, or certain other substances. Primary lung cancer develops in the lungs, whereas secondary lung cancer begins in the lungs and spreads to other parts of the body. The size of the tumor and the extent to which it has spread indicates the stage of cancer. An early-stage cancer is one in which cancerous tissue or a malignant tumor has been diagnosed in the lung, whereas advanced cancer has spread to surrounding tissue or another part of the body (metastasis). An improved understanding of risk factors can benefit the prevention of lung cancer..

The pie charts below depict the top 10 cancer types in terms of estimated cases and deaths for men and women worldwide, both collectively and individually, with NMSC included in the other group. More than 60% of newly diagnosed cancer cases and more than 70% of cancer mortality in both sexes are attributed to the top 10 cancer types. Female breast cancer is the most common cancer (11.7 percent of all cases), followed by lung cancer (11.4 percent), colorectal cancer (10 percent), prostate cancer (7.3 percent), and stomach cancer (7.3 percent) (5.6 percent). Lung cancer is the leading cause of cancer death (18.0 percent of all cancer fatalities), followed by colorectal (9.4%), liver (8.3%), stomach (7.7%), and female breast (6.9%). In terms of incidence and mortality, lung cancer is the most prevalent cause of cancer death in men, followed by prostate and colorectal cancer. In terms of incidence and mortality, breast cancer is the most common cancer diagnosed in women and the leading cause of cancer death, followed by colorectal and lung cancer. [54]



Because of their similar symptoms, lung cancer is sometimes mistaken for Tuberculosis. Pathologists and other medical professionals must spend time diagnosing lung cancer and its forms. When cancer types are misdiagnosed, it leads to ineffective therapy and may result in patient death. Early detection using machine learning methodologies is the key to increasing survival rates, and if we can utilize this to make the diagnosis process more efficient and effective for radiologists, it will be a huge step toward the goal of improved early detection.

The two main types of lung cancer are non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). These variants flourish and spread in various ways. They are typically addressed differently.

- *Non-small cell lung cancer (NSCLC)*

  Non-small cell lung cancer accounts for approximately 85%–90% of all lung cancers. There are three types of cancer. They are classed based on the type of lung cell that started the malignancy and how the cells look under a microscope. They do differ somewhat. They do, however, have a similar prognosis and are frequently treated in the same manner:

  (a) **Adenocarcinoma:** The most common kind of NSCLC. It is the most common kind of nonsmoker lung cancer. It is, however, more prevalent among smokers or ex-smokers. It frequently develops on the lungs' outer edges. It progresses more slowly than other types of lung cancer.

  (b) **Squamous cell carcinoma (*epidermoid carcinoma*):** This kind of NSCLC is more common among smokers or former smokers. These malignancies usually begin in the centre of the lungs, near the major airways (the bronchi).

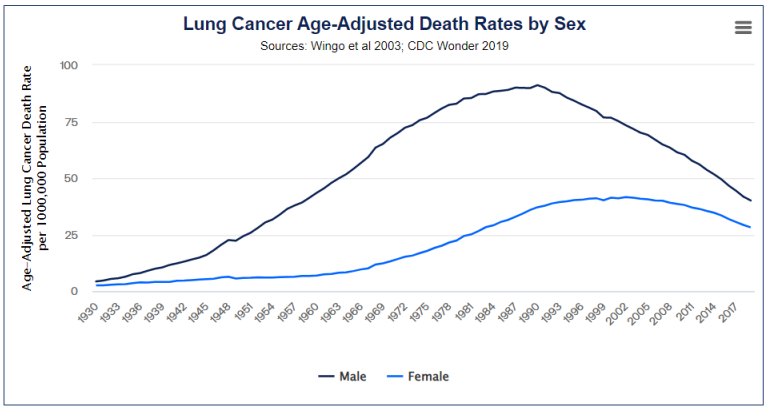
  (c) **Large cell carcinoma:** This is the least prevalent kind of NSCLC. It grows rapidly and spreads to other organs. This may make treatment more difficult.

- *Small cell lung cancer (SCLC)*

  Small cell lung cancer affects 1 in 10 to 3 in 20 lung cancer patients. Another name for it is oat cell cancer. It is almost entirely found among smokers. It spreads more quickly than non-small cell lung cancer. It frequently spreads to other parts of the body at an early stage.

Due to their relative higher occurrence amongst patients, in this paper we have only considered Adenocarcinoma and Squamous Cell Carcinoma for further study.

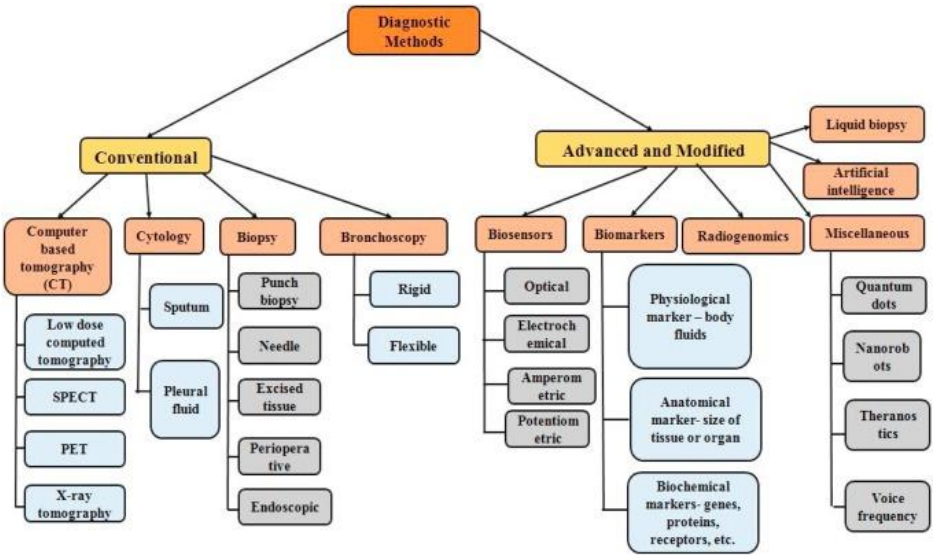
Lung cancer affects both men and women, accounting for more than 25% of all cancer mortality. Smoking is responsible for around 80% of lung cancer fatalities. Lung cancer killed 139,682 people in 2019, accounting for 23% of all cancer mortality. Men died in 74,909 instances, while women died in 64,773 cases. In 2019, males were 42 percent more likely than women to die from lung cancer (40.1 per 100,000 people) (28.2 per 100,000).



CURRENT TRENDS

In India, it is frequently misdiagnosed as TB. Tuberculosis, on the other hand, is uncommon in Western countries such as the United States, the United Kingdom, and Canada. The biggest disadvantage of lung cancer is that symptoms appear at very late stages when the tumor has progressed extensively. As a result, extremely sophisticated methods are used for lung cancer detection, staging, and therapy.

The optimal approach for screening any illness, particularly cancer, must have certain desirable qualities, including little danger to patients, high sensitivity, few false positives, patient compliance, and cost-effectiveness. Because current procedures have low patient compliance and greater costs, the focus of research is turning toward creating more accessible and less expensive ways for cancer screening. The numerous LC diagnostic procedures are illustrated here.



Imaging sets (x-ray, CT scan), Sputum cytology, and tissue samples (biopsy) are utilized to look for malignant cells and rule out other possible illnesses. Experienced pathologists must review microscopic histopathology slides during the biopsy to

establish the diagnosis and characterize the types and subtypes of lung cancers. An overview of the benefits and drawbacks of lung cancer diagnostic procedures is provided below.

S.no.	*Techniques*	*Advantages*	*Limitations*
1.	CT Scan	Painless and well-established technique	Radiation, allergic reaction to contrast material and misinterpretation
2.	Sputum cytology	Minimally invasive and highly sensitive for centrally located carcinoma of lungs	Produces false positives, induction of sputum causes discomfort among patients
3.	Pleural fluid cytology	Misdiagnosis can be avoided, better than radiology	Highly invasive
4.	Biopsy	Detailed and precise data is obtained	Highly invasive which causes discomfort and it is a time-consuming process
5.	Bronchoscopy	Efficient detection of specific tumor tissue	Invasive and lack of expertise can cause scratches in the laryngeal region and airways
6.	Biosensors and biomarkers	Highly sensitive and fast response time	Sample preparation and high expertise in analyzing
7.	RG	Non-invasive and precise	Time-consuming
8.	AI	Fast, efficient, specific and can be used for developing personalized medicines	Expertise required and costly
9.	LB	Minimally invasive, patient compliance and used along with genomic sequencing for developing precision medicine.	Requires highly sensitive analytical techniques, lower molecular abundance, and specificity
10.	Miscellaneous (QD, nanobots, theranostics and voice frequency)	Patient compliance and improved efficiency	Cost of implementation is high.

Robust technology to aid doctors is widely desired in order to avoid Lung Cancer and analyse the early stages of this illness. Engineering solutions may be used to process medical field data using image processing, machine learning, and artificial intelligence approaches in order to detect and diagnose lung disease. Various Neural Network and Machine Learning techniques must be used to pre-process and train medical field data, including X-rays and Computed Tomography (CT) scan images, on the input dataset. Scientists are investigating Lung Disease Diagnosis (LDD) prototypes to improve disease detection processes in the early stages of lung cancer, which will help practitioners or physicians.

AI IN DETECTION

In the later stages of the disease, when treatment is more challenging, around 70% of lung cancers are found, which contributes to the low 5-year survival rate. Initial symptoms of lung cancer are frequently everyday complaints that are simple to brush off as unimportant, including a persistent cough or fatigue.

To keep up with rising demand, there are not enough radiologists available to review lung images. There are a lot of people and a lot of CT scans. Due to their large workloads, overworked radiologists may make errors. Radiologists may readily overlook tiny malignant tumors due to the limits of human vision. For instance, up to 35% of lung nodules are missed during the first examination. AI can help on both counts by taking some of the weight off of overworked experts and identifying lung spots that are invisible to the human eye.

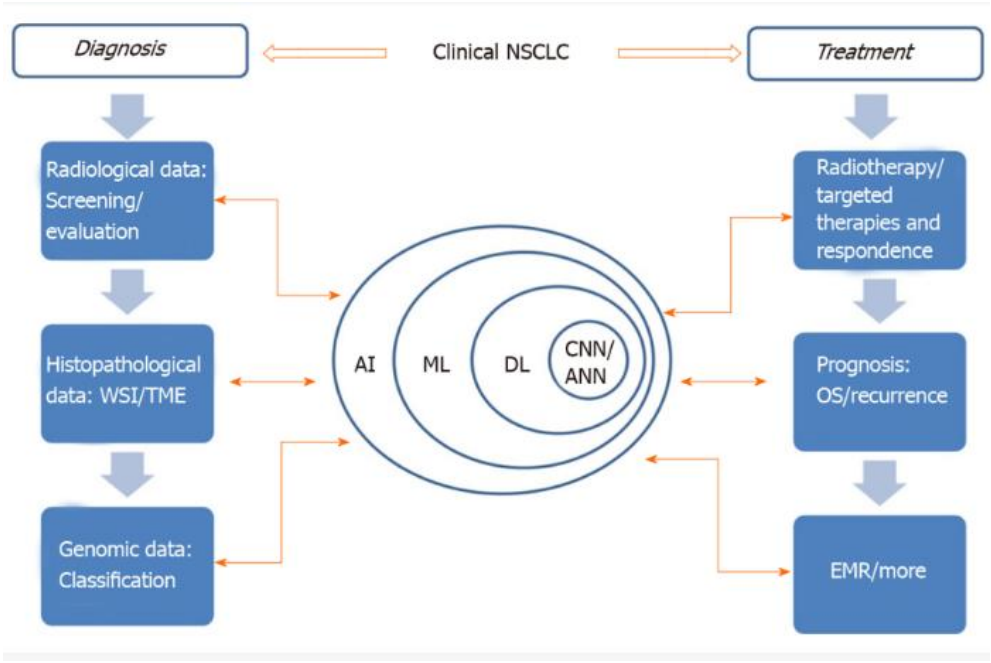
Modern AI systems are based on the idea of deep learning. Instead of looking for tumour features already given by a programmer, deep-learning algorithms comprehend what a tumour is from real-world examples. Researchers feed hundreds of lung CT pictures from individuals with and without cancer into the programmes. The devices can now visualise a lung cancer nodule's appearance. The more training photos the system views, the more accurate it will get at differentiating lung tumours from benign blotches. And they accomplish it with greater precision than earlier, non-AI systems. Some deep-learning algorithms also gauge doctors' trust in their judgement, which might aid them in making better treatment choices.

At the Feinberg School of Medicine at North Western University in Chicago, Illinois, Mozziyar Etemadi and his team of researchers conducted one such study [1]. Using a database of more than 40,000 CT scans, including scans performed before patients were given a lung cancer diagnosis, the researchers trained the system. The professionals gave the computer

instructions throughout this training phase regarding which early-stage scans identified cancerous regions and which did not. The computer gradually mastered the visual characteristics that identified malignant from benign patches, and it improved at identifying early signs of cancer. The system's accuracy is increased since it can analyse a whole 3D CT scan rather than just a series of 2D slices. Additionally, 3D scans provide more diagnostic details about non-tumor properties such blood vessels. Places far from the tumour start to appear in the 3D volume.

The model that detects nodules that indicate early lung cancer has been built by researchers in another deep-learning AI study led by Ulas Bagci, a computer-vision specialist at the University of Central Florida in Orlando, and his colleagues. In contrast to radiologists' typical accuracy rate of 66%, the computer effectively recognised tiny particles of malignancy on CT scans around 95% of the time. [1]

Deep-learning systems' ability to expedite population screening, which tries to find lung cancer much sooner, is one of its main advantages. There is strong evidence that these programmes would be effective.



Clinical non-small cell lung cancer and the use of artificial intelligence The two-way arrows signify the process of learning and the use of AI in different fields: ANN: Artificial neural network; CNN: Convolutional neural networks; DL: Deep learning; AI: Artificial intelligence; EMR: Electronic medical record; ML: Machine learning; NSCLC: Non-small cell lung cancer; OS: Overall survival time; TME: Tumor microenvironment; WSI: Whole slide image. [2]

FEATURE SELECTION

Since our dataset does not have features labels, we had to opt for unsupervised feature selection methods to achieve our goal. This included,

- **Variance Threshold Filter method:** A general method for choosing features is the variance threshold. All characteristics whose variance falls below a certain level are removed. It eliminates all zero-variance features—that is, features whose value is the same across all samples—by default. Although one of the limitations of filter techniques is that they do not take into consideration the relationship between feature variables or between feature and target variables, we presume that features with a larger variance may contain more important information.

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Variance of variable x

For the pre-trained strong sub-learners and for the CNN model built by us, the variance threshold was set to be greater than or equal to 0.98 i.e. 98%.

- **LASSO Regularization (L1) Embedded method**: Regularization includes either imposing a penalty equivalent to the magnitude of the coefficient in absolute terms or merely capping the size of the coefficients. modify the different machine learning model parameters in order to reduce the model's flexibility and hence prevent over-fitting. The penalty for regularising a linear model is placed on the coefficients that multiply each predictor. Lasso or L1 regularisation, among other types, has the capacity to zero out certain coefficients. As a consequence, the model can do without that feature.

  When there are many features, L1 regularisation is the preferable method since it offers sparse solutions. Because features with zero coefficients may be eliminated, we even get the computational gain. Lasso Regression is the name of the L1 regularisation technique-based regression model.

$$Cost = \sum_{i=0}^{N} (y_i - \sum_{j=0}^{M} x_{ij} W_j)^2 + \lambda \sum_{j=0}^{M} |W_j|$$

L1 regularization

The characteristics with the strongest correlation to the target types were then chosen using the LASSO approach, which minimises absolute shrinkage and selection. Although Lasso regression has an L1 penalty, it is quite similar to Ridge regression. The main benefit of Lasso regression is the effective, automated feature selection it does. In order to choose the best features, Lasso regression will attempt to reduce the coefficient of the less significant feature to 0. If two features are linearly linked, their simultaneous existence will enhance the value of the cost function.

$$\frac{1}{2N_{training}} \sum_{i=1}^{N_{training}} (y_{real}^{(i)} - y_{pred}^{(i)})^2 + \alpha \sum_{j=1}^{n} |a_j|$$

Lasso Regression

In order to improve the statistical model's predictability and interpretability, LASSO applies both variable selection and regularisation. [3], [4] These criteria describe the lasso estimate:

$$RSS(D) = D_0 + \sum_{j=1}^{N} \left(y_j - D_0 - \sum_{i=1}^{M} x_{ij} D_j \right)^2$$

Least squares estimation in which we pick coefficient D= (D0, D1…, DN) to minimize the residual sum of squares:

$$D^{lasso} = D_0 + \sum_{j=1}^{N} \left(y_j - D_0 - \sum_{i=1}^{M} x_{ij} D_j \right)^2 + L \sum_{j=1}^{M} |D_j|$$

The concept of using Lasso regression for feature selection is thus pretty straightforward: we fit a Lasso regression on a scaled version of our dataset and we only take into account those features that have a coefficient different from 0. For the proper kind of Lasso regression, it goes without saying that we must first calibrate the hyper parameter.

The Lasso (L1) Regularization method is applied in the Logistic Regression meta learner for the Ensemble model.

CLASSIFICATION

Classification with deep convolutional neural networks

The feature extraction and image classification processes in this exploratory investigation employed CNNs. We used a transfer learning strategy while also creating our own model to solve the issue caused by the lack of curated medical data and

histopathology imaging data often found in everyday clinical practice. On our imaging data, visual recognition is performed by tuning robust models that are successful at completing other computer vision tasks. ResNet50V2, InceptionV3, VGG16, VGG19, EfficientNetB7, Xception, InceptionResNetV2, ResNet101, ResNet101V2, ResNet152, ResNet152V2, DenseNet121, DenseNet169, DenseNet201, MobileNet, and MobileNetv2 neural network architecture, evaluated after being pre-trained on a sizable natural image dataset (ImageNet). Of which only EfficientNetB7, ResNet50, ResNet101, ResNet152 and VGG16 were further studied, due to their better performance over others. An ensemble of these models formed the basis of our primary classification approach; whereas, the secondary model was studied directly on the chosen dataset. We assessed the networks for the pre-trained models by fine-tuning the pooling, fully connected layers, and hidden layers. Convolutional, fully connected, and pooling layers were adjusted in the secondary CNN model. Iteratively investigating hyperparameter optimization The pre-models received 256x256 pixel picture patches as input. Compared to the secondary model, which used grayscale pictures as its only input, the primary model featured three input channels that were each fed one of the RGB channels of the image (i.e., model inputs were stacked greyscale images of each channel). A group of pictures from either adenocarcinoma (ADC), squamous cell carcinoma (SCC), or benign histology for both models were used for fine-tuning across 50 epochs. The final prediction (softmax) layer was set to three. Using the confusion matrix and other performance indicators described in the model evaluation section, the predictive performance of the models was assessed.

Feature-based analysis and classification

When combined with supervised machine learning classifiers like support vector machines (VM), random forest, and logistic regression classifier, CNN-derived feature maps outperform the original CNN in classification tasks. [5],[6],[7] In contrast to manually created radionics features, features from CNNs maintain global spatial information using convolutional kernel operations on the input picture. [8] They benefit from this in terms of fine-grained recognition, domain adaptability, contextual recognition, and texture attribute recognition. [8]CNNs also rely less on human curation, which lessens bias. With discriminative machine learning classifiers, such as Logistic Regression, classification performance must be optimized. Despite having broad criteria, it may perform badly if certain traits greatly depart from a normal distribution.

CONTRIBUTION

We have been successful in training and testing over 17 different pre-trained models on the dataset and amongst these we have hand-picked the top performing models i.e. VGG16, EfficientNeTB7, ResNet101, ResNet50, ResNet152. Employing these, as individual strong sub-learners, we were able to develop a fast and accurate stacked ensemble model, which upon further testing, correctly identified benign lung tissue from cancerous ones. The model achieved the distinction between adenocarcinoma and squamous cell carcinoma, with an accuracy of over 90%.

Furthermore, we also developed a CNN model, with 4 convolutional layers, one fully connected layer, and four hidden layers with dropouts embedded in between. This model achieved an accuracy score of over 97%. Similar to the ensemble model, this was also able to accurately distinguish benign tissue from malignant ones.

This model was tested with a couple of different image processing techniques. The image segmentation process did not contribute any significant positive effect on the original results. It was determined that the process, upon segmenting, resulted in images, that lacked extractable features, and thus on averaging the results of the segmented parts, images were at times falsely classified. Thus this procedure was not incorporated in the CNN model.

However, with the Grey Scale Characteristics (GSC) image filter, the model achieved its final results. The image is split into its respective RGB channels and their RGB pixel values were interpolated from 0-255 to 0-1, followed by splitting them into B bins. Integral histograms are calculated for each of these bins which forms the feature matrix that is inputted into the model's CNN layers.

Organization of the Paper

Section 2 of this paper deals with the various pre-existing systems which are currently being used in the detection of lung cancer. Followed by that, Section 3 explains the overall architecture of the model in a nutshell. Then Section 4 provides a detailed description of the proposed work including the various aspects which went into the creation of the model. Section 5 portraits the various steps which went into the training and the testing phase of the model and also depicts the results obtained from the model. This section also deals with the comparison between the existing and the developed model. Section 6 ensembles the conclusion and the future scope, this project can move ahead with. Wrapping it up is Section 7 which contains the detailed description of the various papers and journals which were referred whilst the project development was underway.

2. LITERATURE SURVEY

The Lung Disease Diagnosis (LDD) literature review comprises a study of numerous articles on Lung Cancer that use image processing techniques on Computed Tomography (CT) X-ray images and histopathology image data.

Early works in the field involve the use of artificial learning to detect the presence of cancerous cells which are normally aided by the examination of the performance of various pre-existing classification algorithms using Support Vector Machines and the Naïve-Bayes method.

In order to lessen the problems CAD systems encounter in the identification of lung cancer, authors Xiaodan Chen, S. Feng, and D. Pan (2015) applied an adaptive threshold algorithm, also known as mathematical morphology and Watershed algorithm. They enhanced the picture quality by applying the gradient enhancement approach and a Gaussian filter to remove noise in the original Computed Tomography image (CT). Following that, the OTSU approach was used to segment the photos. Finally, the Watershed method was used after executing the tests in a CT, and the scientist was successful in demonstrating how lung parenchyma was segregated correctly. [14]

The model proposed by Jin, X., Zhang, Y., & Jin, Q. recommended the use of CNN (Convolution Neural Networks) as a classification approach in the Computer-Aided Diagnostic (CAD) system to diagnose lung cancer. This model achieved 85% accuracy, 83 % sensitivity, and 87 % specificity. The recommended method has the benefit of applying a circular filter in specified regions during extraction, which helps to reduce the total cost of the detection and training phases, however the accuracy is still insufficient. [15]

The researchers of this article have combined the kernel graphs cut algorithm and mathematical solution, and this algorithm is also compared with K-means algorithm and OTSU's maximum between-cluster variance algorithm. The researchers have applied the kernel graph cuts method for image segmentation from Computed Tomography. Xinyan Li S.F., Daru Pan. (2016). It is abundantly obvious from the researchers' comparison that the accuracy of K-means and OTSU approaches is considerably superior than that of the suggested algorithm. [16]

In 2019, Ahmed, Amer, and F.E.Z. Chadi created an early-stage optimising method for applying deep learning to find lung nodules. Prior to using transfer learning for feature extraction, the authors first pre-processed the input picture by increasing the image contrast for low dose scans. The genetics algorithm GA is then used in the third step to the deep learning features collected as well as the training subset of data to obtain significant pulmonary nodules from the best characteristics recovered. Using SVM Support Vector Machine and the ECLAP online lung image database, the categorization of chosen features for diagnosing lung cancer was the last stage. This prototype's testing results showed 92.5% detection accuracy, 90% sensitivity, and 95% specificity. [17]

Lung cancer was identified using a novel method by Suren Makaju et al. (2018). Instead of the Gabor filter, the researchers utilised the Median filter and the Gaussian filter to pre-process computed tomography (CT) images. The pictures were then segmented using the Watershed method. The authors could recognise marked cancer nodules up until this stage. Additionally, the investigators retrieved variables including eccentricity, perimeter and area, pixel mean intensity, Centroid, and diameter. The classifier was then employed for the SVM-based classification of cancer nodules after being trained to do so. For this, the authors used the Lung Image Database Consortium (LIDC). The accuracy of identifying lung cancer and categorizing it as benign or malignant was enhanced by the writers of this paper. There is still room for improvement in improving the accuracy and detection of different stages of lung cancers such as Stage-I, Stage-II, Stage III, and Stage IV, which can really help the medical field diagnose lung cancer and make appropriate treatment decisions for the patient. The accuracy of lung cancer detection has increased in this paper from an earlier detection value of 86.6 % to 92 %. [18]

Zhuoqi and S.H. Idiopathic. (2019) investigated CNN structure selection and effort comparison, used Fast RCNN and Faster RCNN for candidate detection, residual learning neural network for increasing network depth, transfer learning for initialising parameters, and network parameter iterative method for optimising training. Comparatively, all of these techniques are used to evaluate and find lung nodules. Although neural network transfer learning techniques produce respectable results, there are a few other strategies that may be developed to produce even better outcomes. This article outlined the domain of future work for categorization enhancement by illustrating the efficacy of the cutting-edge deep learning method. [19]

In order to simplify classification calculation, Pankaj Nanglia, Sumit Kumar, and colleagues developed a unique hybrid approach called Kernel Attribute Selected Classifier, combining SVM with Feed-Forward Back Propagation Neural Network. They created three fundamental classification techniques, the first of which is pre-processing the dataset. The second block involves the SURF technique for feature extraction, followed by a genetic algorithm for optimization, and then FFBPNN for classification. The total accuracy of the proposed method is 98.08 percent. [20]

Researchers Ruchita Tekade and Prof. DR. K. Rajeswari studied the idea of lung nodule detection and malignancy degree prediction using lung CT scan images. The datasets used in this experiment were LIDC IDRI, LUNA16, and Data Science Bowl 2017. The experiment was run on a CUDA-capable GPU Tesla K20. The dataset was assessed using an artificial neural

network with the aim of extracting characteristics and classifying the data. They used the 3D multigraph VGG-like architecture to detect lung nodules and forecast malignancy level, and the U-NET architecture to segment lung nodules from CT scan images. The best results were obtained when these two techniques were combined. This approach has an accuracy of 95.66%, a loss of 0.09, a dice coefficient of 90%, and a log loss prediction of 38%. [21]

The research of Moffy Vas and Amita Dessai was focused on the classification of cancerous and non-cancerous lung images. Unwanted areas of the lung CT picture were removed during pre-processing, according to their recommended approach. They used a middle filter to decrease the sounds of salt and pepper. Precise lung segmentation and tumour zone identification are made possible by a mathematical morphological technique. The segmented region's energy, correlation, variance, homogeneity, differential entropy, correlation information measure, and contrast were extracted and fed to the classification algorithm's feed forward neural network with back propagation. The technique looks for the lowest error function when using the weight space gradient descent approach. The weights are mixed around in order to minimise the error function. The accuracy during training was 96%, while the accuracy during testing was 92%. The specificity was 97.1%, while the sensitivity was 88.7%. [22]

A. Thirach, S. Marukatat, W. Ausawalaithong, and T. From chest X-ray pictures gathered from several data sources, Wilaiprasitporn utilised deep learning in combination with a transfer learning technique to predict lung cancer. An picture of size 224x224 was produced using a single sigmoid node and a 121-layer densely linked convolutional network (DenseNet-121). The suggested model obtained 74.436 percent mean accuracy, 74.969 percent mean specificity, and 74.6815 percent mean sensitivity for multiple image source datasets. [23]

Deep Convolutional Neural Network (DCNN) was utilized on cytological images by T. Atsushi, T. Tetsuya, K. Yuka, and F. Hiroshi to automate the categorization of lung cancer kind. They included pictures of adenocarcinoma, squamous cell carcinoma, and small cell carcinoma in their collection. With three convolution and pooling layers, two fully connected layers, and a dropout of 0.5, the DCNN architecture produced a somewhat poor total accuracy of 71.1 %. [24]

W. Rahane, H. Dalvi, Y. Magar, A. Kalane, and S. Jondhale proposed employing image processing and machine learning (Support Vector Machine) for lung cancer diagnosis on computed tomography (CT) images. Binarization, noise reduction, and other image processing techniques were used. The size, perimeter, and eccentricity of the segmented picture region of interest were input to the support vector machine (SVM) model as features. [25]

In order to identify lung cancer using whole-slide histopathology images, M. Šarić, M. Russo, M. Stella, and M. Sikora developed CNN architectures based on VGG and ResNet. The results were compared using the receiver operating characteristic (ROC) plot. Patch level accuracy for ResNet50 and VGG16 was 0.7205 and 0.7541, respectively, respectively. The authors cited the high pattern variability across many slides as the cause of the offered models' poor accuracy. [26]

The authors S. Sasikala, M. Bharathi, and B. R. Sowmiya suggested using CNN to CT scan pictures to identify and categorise lung cancer. They worked using MATLAB, and during the first training stage, they extracted useful volumetric characteristics from the input data before classifying it. With 96 percent accuracy, their suggested technique could discriminate between malignant and non-cancerous cells. [27]

Manikandan T (2017), in his paper, stated that the CAD (Comp. Assisted Diagnosis) model can be helpful to resolve lung cancer recognition in its early stages using CT images. He directed that the major four steps are there to detect the nodule in a lung cancer patient. The author suggested that there are certain challenges will be faced by medical professionals while diagnosing the cancer nodule is complex from standard CT images, so these four stages will be impactful to detect. In the second step, a suspected lung cancer nodule is segmented using various segmentation techniques, including multiple thresholding, optical thresholding, global thresholding, active contour method, morphological method, watershed segmentation, shape-based method, and template matching. In the third step, features (2-D features), including shape-size features, geometric features, greyscale features, gradient features, and statistical features are extracted. As stated by the author, there is still much room for improvement in the sensitivity, specificity, and accuracy of the current diagnosis method. [28]

In their work, T Aggrawal et al. (2015) suggested a technique for classifying cancer nodules as well as a typical lung structure. In order to distinguish between the normal and cancerous nodule structures, the inventor of this model first processed the lung nodule using segmentation, extracting geometric information from the nodule, and combining these geometric data with the LDA classification algorithm. The suggested approach attained 84 percent accuracy, 97.14 percent sensitivity, and 53.33 percent specificity. The suggested model found the cancer nodule, although there is still opportunity for accuracy enhancement. There is opportunity for improvement in accuracy because this model only employed a basic segmentation method. [29]

P. B. Sangamithraa and Govindaraju, S. (2016) proposed a model in which CT images are first pre-processed for noise removal, the tumour is then segmented from the lungs using the fuzzy k-means algorithm, the segmentation result is then improved using the K-means approach, and finally, some features of the lungs, including entropy, correlation, homogeneity, SSIM, and

PSNR, are extracted from the CT images Although this model's accuracy was 90.7%, Support Vector Machine and other new categorization approaches may help to increase it. [30]

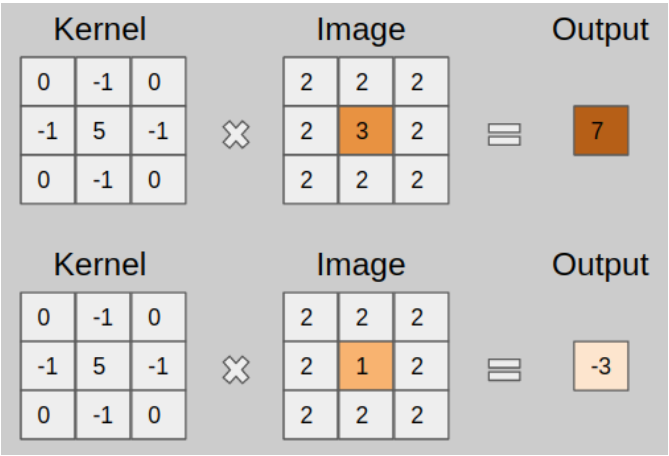
A.M.R.W. and Allison The suggested model's creators employed deep learning to improve lung cancer diagnostic precision by utilizing a CNN network with different pre-processing techniques. The combination of two networks with different sorts of inputs and the use of a voting mechanism make the recommended approach unique compared to others. The steps in this model use a Gaussian filter to pre-process computed tomography images before using them as a training dataset for CNN networks. Next, smoothed and unsmooth images are combined in the second stage, and in the third stage, a convolution filter is used to send the feature to the maximum pooling layer. the output of the third stage will be the input to the soft Each CNN network will be constructed in this manner, and the voting mechanism will produce the output in the end. The accuracy reached by this model was 97.5 %, with a false positive rate of less than 10%; the authors intend to improve the accuracy by modifying the voting method. [31]

In this study, Yu, KH., Zhang, C., Berry, G. et al. extracted 9,879 quantitative image features from 2,186 haematoxylin and eosin stained histopathology whole-slide images of lung adenocarcinoma and squamous cell carcinoma patients from The Cancer Genome Atlas (TCGA) and 294 additional images from Stanford Tissue Microarray (TMA) Database. They then used regularised machine-learning methods With the use of the TMA cohort, they verify the framework for predicting survival (P = 0.036 for both tumour types). Our findings imply that computer generated image characteristics can forecast lung cancer patients' prognoses and hence support precision oncology. The histopathological pictures of different organs may be analysed using our approaches. Using tools from the open microscope ecosystem, all pictures taken at x40 magnification were tiled. Our automated workflow ignored areas with relatively low cellularity, such as alveolar gaps, and chose the ten densest tiles from each picture for additional examination in order to pinpoint areas with pathological alterations. Moreover, we obtained 294 tissue microarray pictures from the Stanford Tissue Microarray (TMA) Database40, with one representative histopathology image being chosen by pathologists for each of the 227 patients with lung adenocarcinoma and 67 patients with lung squamous cell carcinoma. [32]

3. SYSTEM ARCHITECTURE

15000, 768x768 pixel histopathological images from the Lung and Colon Cancer dataset, were reshaped to into 256x256 pixel images, to serve as input to the models.

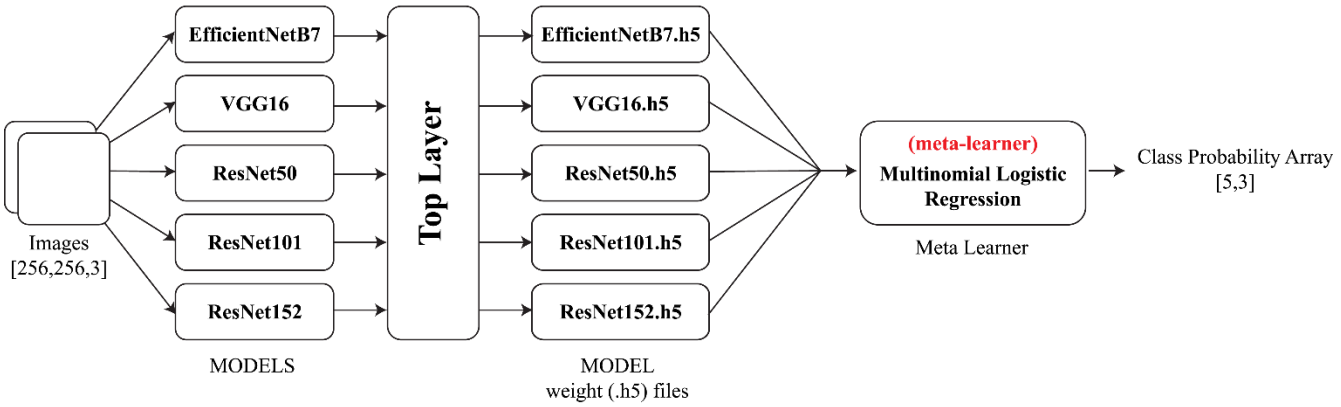
The primary model consisting of ResNet50, ResNet101, ResNet152, EfficientNetB7 and VGG16, stacked on each other, take the reshaped images as input, and passes them through a variety of convolutional filters (kernels), extracting various features from the images. A kernel is a matrix, which is slid across the image and multiplied with the input such that the output is enhanced in a certain desirable manner. A filter however is a concatenation of multiple kernels, each kernel assigned to a particular channel of the input. Filters are always one dimension more than the kernels. For example, in 2D convolutions, filters are 3D matrices (which is essentially a concatenation of 2D matrices i.e. the kernels). So for a CNN layer with kernel dimensions h*w and input channels k, the filter dimensions are k*h*w.



Deep CNNs allow us to learn these kernel values that can extract latent features in place of manually created kernels for feature extraction. [11]

For a particular image, individual models generate a 1x3 matrix of probabilities, indicating, the probable change of the image to belong to that particular class. The dataset is split into an 80-20 ratio i.e. 12000 images for training and 3000 for testing. Thus for every image 5 models would generate a 5x3 matrix of probabilities. Consequently, for 12000 images, the matrix

shape would become [12000x5x3]. This three-dimensional matrix is then reshaped to [12000,15], consolidating all predictions about an image into one array. This is then passed onto the Multinomial Logistic regression meta-learner. L1 regularization is applied to prevent the model from overfitting the data.



Model Architecture

For the secondary CNN model, the reshaped 256x256 pixel images were taken as input and were split into their fundamental channels (here, RGB). The pixel values of these channels were then normalized to lie between (0,1) i.e. converted to greyscale, followed by splitting the grayscale images into several bins. Each of them is then converted into an integral histogram
Which become the feature matrices to be inputted into the model. Just like the primary model, the feature matrices are now passed through various kernel filters to extract and select notable features in the images.

Both models have a similar output format. They generate a 1x3 matrix of probabilities which indicate the chances of a particular image belonging to a particular class. The highest value in this array along with its index gives us the required class. Class. 0 being for adenocarcinoma, 1 for benign and 2 for squamous cell carcinoma.

4. PROPOSED WORK

A comparative analysis of 18 image classifiers was conducted with respect to a simple CNN model, which served as a control. In parallel, we endeavoured towards making a CNN model, capable enough to classify cancerous from benign tissue (specifically into 3 classes, namely *adenocarcinoma*, *squamous cell carcinoma* and benign), with accuracies comparable to Tensorflow's pre-trained models. These models are, ResNet50, ResNet50V2, InceptionV3, VGG16, VGG19, EfficientNetB7, Xception, InceptionResNetV2, ResNet101, ResNet101V2, ResNet152, ResNet152V2, DenseNet121, DenseNet169, DenseNet201, MobileNet, and MobileNetv2.

The top-performing pre-trained models are then plugged into a stacked ensemble model, aiming for better model metrics than individual sub-learners. The top models are found to be ResNet50, ResNet152, ResNet101, VGG16 and EfficientNetB7.

4.1 Background

A. Feature Extraction

A step in the dimensionality reduction process, feature extraction breaks down a large collection of raw data into smaller, more manageable categories to make processing simpler.

A convolution in the context of a convolutional neural network is a linear operation that, like a standard neural network, incorporates the multiplication of a set of weights with the input. The multiplication is carried out between an array of input data and a two-dimensional array of weights, known as a filter or a kernel, because the technique was designed for two-dimensional input. Convolution computation layer is the most important layer in CNN structure. The convolution layer consists of numerous convolution filters, sometimes referred to as convolution kernels, that may extract different properties from the input picture data. Convolution between an image and a convolution kernel may be explained as follows:

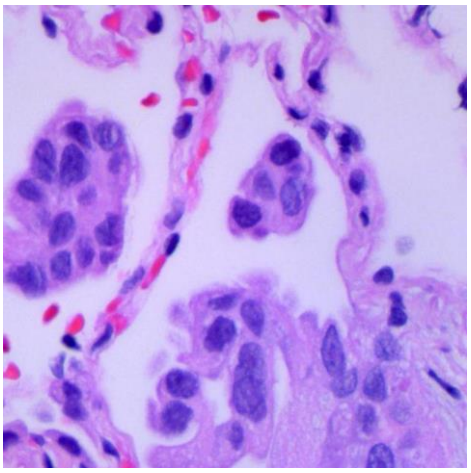
$$S(i,j) = (k * I)(i,j) = \sum_{m} \sum_{n} I(i+m, j+n) K(m,n)$$

The filter is smaller than the input data, and the method of multiplication is a dot product between a patch of the input and the filter that is the same size as the filter. The element-wise multiplication of the input and filter's filter-sized patches results in a dot product, which is then added to form a single value. Since it generates a single value, the process is referred to as the "scalar product." The filter is intentionally smaller than the input because doing so enables the same filter (set of weights) to be multiplied by the input array many times at different locations on the input. To each overlapping component or filter-sized patch of the incoming data, the filter is successively applied from left to right, top to bottom.

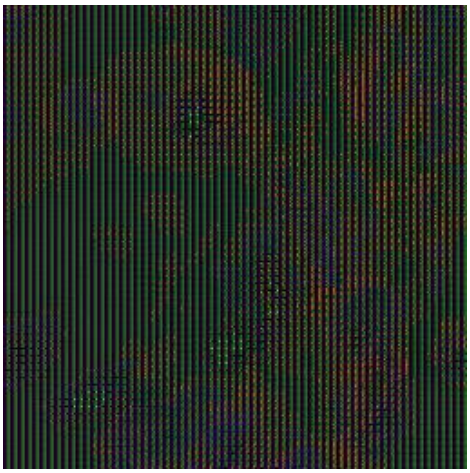
This methodical application of the same filter on an image is a potent idea. Applying the filter systematically across the whole input image enables the filter to find the desired feature wherever it may be in the image if it is intended to recognise a certain type of feature in the input. Translation invariance, the name of this quality, denotes the broad interest in the presence of the characteristic rather than its location.

When using pre-trained models, the neural network's feature extraction section is already populated with pre-trained weights for the ImageNet dataset. Even though filters are only weights, their unique two-dimensional structure gives the weight values spatial relationships, making it useful to plot each filter as a two-dimensional image.

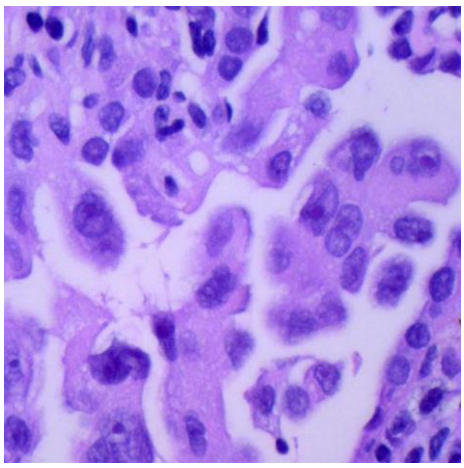
Upon passing through the feature extraction layers of the pre-trained models, such images were re-generated from the filtered arrays



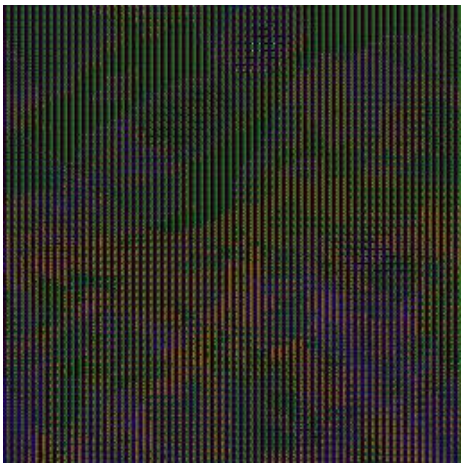
Original Tissue Image



Post filtering re-formed image from the filtered arrays



Original Tissue Image



Post filtering re-formed image from the filtered arrays

In case of the secondary CNN model, Grey Scale Characteristics (GSC) filter is employed. The image processing process had the following steps,

- The inputted RGB image of shape (256,256,3) is segmented into its fundamental Red, Green and Blue channels, each with a shape of (256,256,1).

- The pixel values of each of these channels are then normalized to lie between 0 and 1, i.e. converted to grayscale (256,256,1).

- Therefore, the image is broken into N channels (here 3 i.e. RGB). Now each channel is broken into B bins. Followed by computing the integral histograms of each bin.

- Red, green, and blue (if using RGB) or hue, saturation, and value (if using HSV) colour channels are each separated into equal-sized ranges. A 3D "bin" is created by every combination of the three channels' ranges. If we opted to divide the blue channel into two bins, for instance, our ranges in the blue channel would be: $0 \leq B \leq 0.50 \leq B \leq 0.5$ and $0.5 < B \leq 10.5 < B \leq 11$.

- The recursive propagation method known as the integral histogram operates in Cartesian spaces and may be expanded to any dimensions data space and any tensor representations. It is made, pixel-by-pixel creating an image's histogram and recording the histogram at each stage. Consider scanning pixels from left to right and top to bottom as an example.

- The integral histograms are then inputted to the CNN layers as a feature vector.

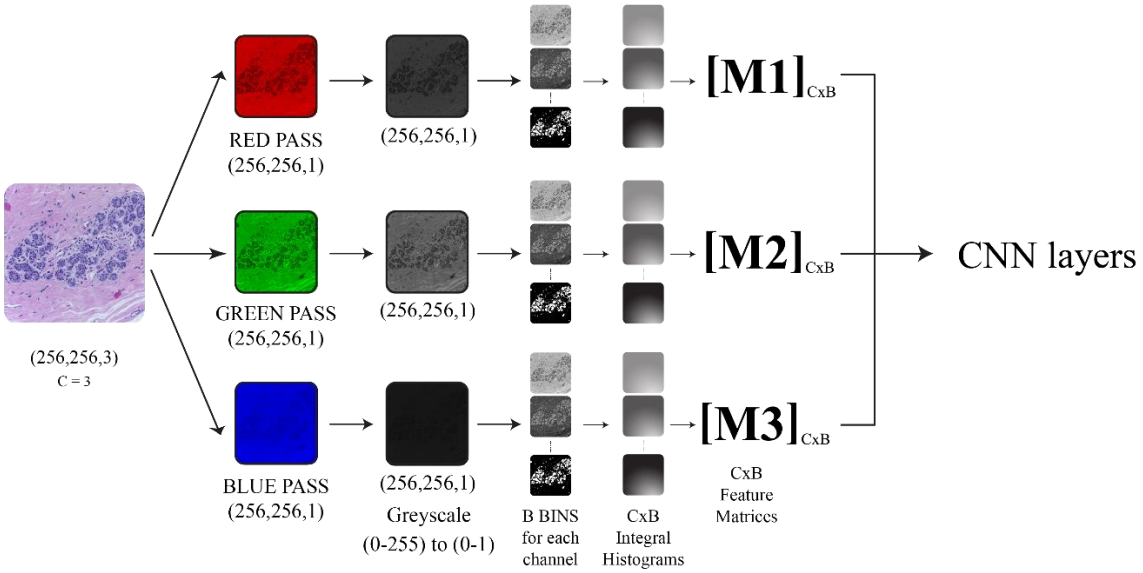


Image Processing

Clear Peugeot aspects in medical images are the grey mean, grey variance, and grey histogram entropy. They are shown in the formulae below,

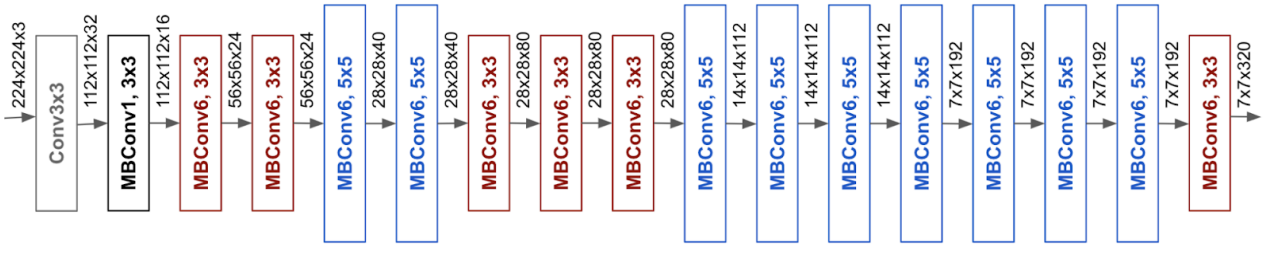
$$G_m = \frac{\sum_{i=0}^{M}\sum_{j=0}^{N} I(i,j)}{M*N} \quad \text{(grey mean)}$$

$$G_n = \sum_{i=0}^{M}\sum_{j=0}^{N}(I(i,j) - G_m)^2 \quad \text{(grey variance)}$$

$$H = -\sum_{i=0}^{L-1} p(i)\log_2 p(i) \quad \text{(grey histogram)}$$

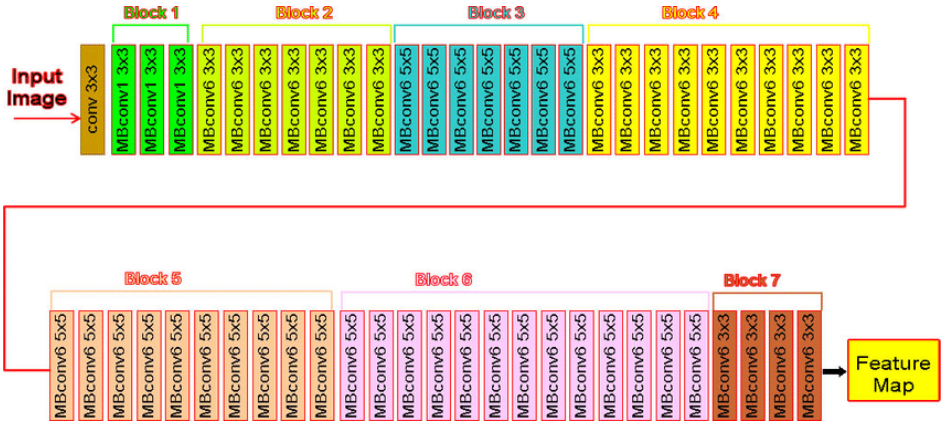
B. EfficientNetB7

The effectiveness of model scaling depends on the baseline network. By using a baseline network and the AutoML MNAS framework to conduct a neural architecture search, which maximises both accuracy and efficiency, its performance is improved (FLOPS). The end result uses mobile inverted bottleneck convolution (MBConv), like MobileNetV2 and MnasNet, but is much larger because to a larger FLOP budget. Afterward, we expand the fundamental network to create the EfficientNets family of models.



The baseline architecture of EfficientNet-B0 is simple and clean, making it easier to scale and generalize

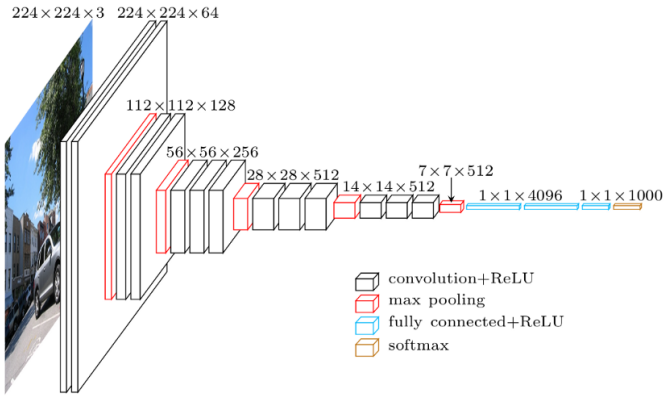
AutoML MNAS created a baseline network called EfficientNet-B0, and then scaled it up to create Efficient-B1 through B7. In fact, although being 8.4x smaller than the best current CNN, our EfficientNet-B7 achieves new state-of-the-art 84.4 % top-1 and 97.1 % top-5 accuracy.



EffieientNetB7 Architecture

C. VGG16

VGG16 is a convolution neural net (CNN) architecture that won the 2014 ILSVR (ImageNet) competition. It is regarded as one of the best vision model architectures to date. The most distinctive aspect of VGG16 is that it focused on having convolution layers of a 3x3 filter with stride 1 and consistently used the same padding and max-pool layer of a 2x2 filter with stride 2. Convolution and max pool layers are always arranged in this way throughout the design. Finally, it has two FC (fully connected layers) and a softmax for output. The 16 in VGG16 relates to the fact that it has 16 weighted layers. This network contains over 138 million parameters and is rather large. The pre-trained network can classify images into 1,000 distinct item categories, including keyboards, mice, pencils, and other animals. As a result, the network has learned intricate feature representations for a wide range of images. The network takes 224 by 224 image input.

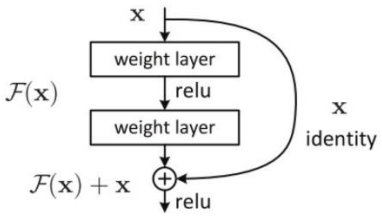


Architecture of VGG16

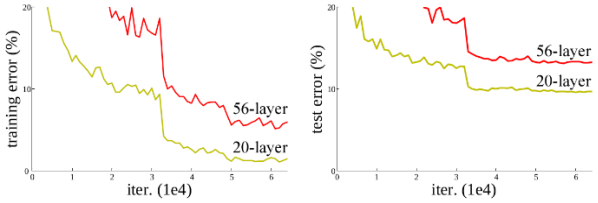
D. Residual Neural Network

Deep Learning and Computer Vision have seen a number of advancements. These models, particularly with the advent of extremely deep Convolutional neural networks, aided in achieving state-of-the-art results on issues such as image identification and picture classification. As a result, deep learning architectures developed deeper and deeper (adding more layers) over time to handle more and more complicated problems, which also aided in increasing classification and recognition task performance and making them more resilient.

However, when additional layers are added to the neural network, training gets increasingly challenging, and the model's accuracy begins to saturate and subsequently decreases, due to the vanishing gradient problem. Here comes the ResNet to save us from that predicament and assist in resolving this issue, with its Residual blocks, offsetting the vanishing gradient by allowing your network to skip through layers of it feels they are less relevant in training.



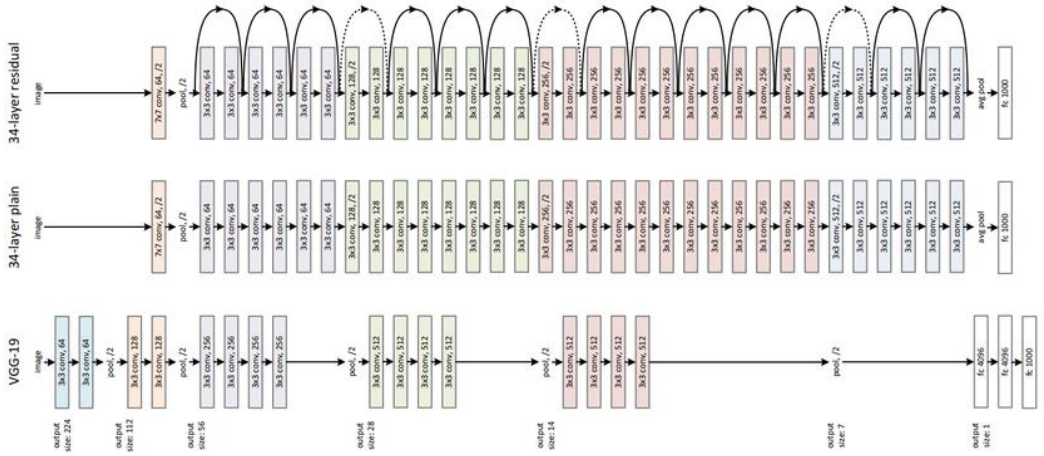
A Residual Block



Vanishing/Exploding Gradient problem

we can see on the left and the right that the deeper model is always producing more error, where in fact it shouldn't have done that.

Shaoqing Ren, Kaiming He, Jian Sun, and Xiangyu Zhang introduced the Residual Network (ResNet), one of the well-known deep learning models, in their study. "Deep Residual Learning for Image Recognition" was the title of the article in 2015 [55]. One of the most well-liked and effective deep learning models to date is the ResNet model. The 34-layer simple network with shortcut and skip links is part of the design, which is based on the VGG-19. The design becomes the residual network as a result of these skip connections or residual blocks, as seen in the picture below.



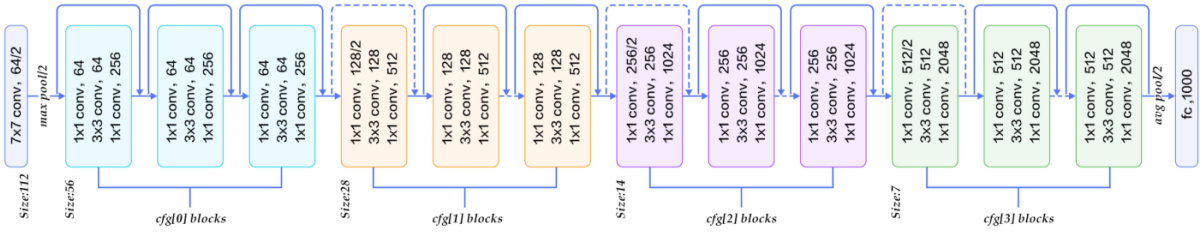
Base Architecture of ResNet with Residual Blocks as compared to VGG19

layer name	output size	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2		
conv2_x	56×56	3×3 max pool, stride 2		
		$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax		
FLOPs		3.8×10^9	7.6×10^9	11.3×10^9

Architecture summary of ResNet50, ResNet101 and ResNet152

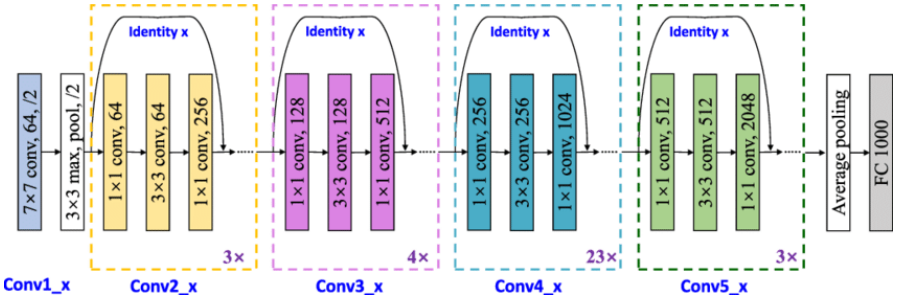
The sole difference between ResNets is the amount of layers of a certain type of convolution. As shown in the table above, the layer sequences for ResNet50, ResNet101, and ResNet152 are [3, 4, 6, 3], [3, 8, 36, 3], [3, 4, 23, 3], and [3, 8, 36, 3], respectively.

- **ResNet 50**
  ResNet50 is a ResNet model version having 48 Convolution layers, 1 MaxPool layer, and 1 Average Pool layer. The number of floating point operations are 3.8 x 10^9.



Architecture of ResNet50

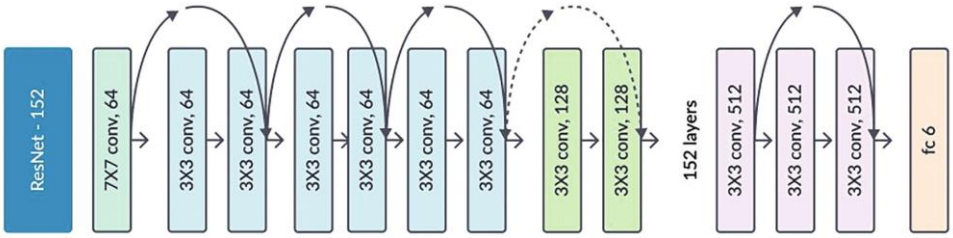
- **ResNet101**



Architecture of ResNet101

- **ResNet152**

  In comparison to VGG-16, GoogleNet (Inception-v1), and PReLU-Net, the top-5 error rate drops to 5.71 % when the network depth is extended to 152 layers. When employing 10-Crop Testing + Fully Conv with Multiple, ResNet-152 has an error rate of 4.49 %.



Architecture of ResNet152

4.2 Proposed Model

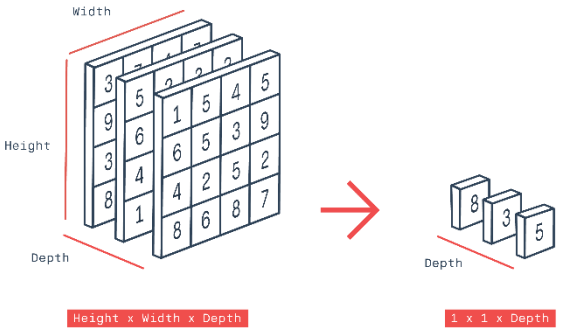
4.2.1 ***Ensemble Model***

Individual pre-trained models from Tensorflow, were initialized with the weights from ImageNet dataset. The Lung and Colon Cancer Histopathological Image dataset is broken into Training and Validation set in 8:2 ratio, giving us 12000 images for training and 3000 for testing purposes.

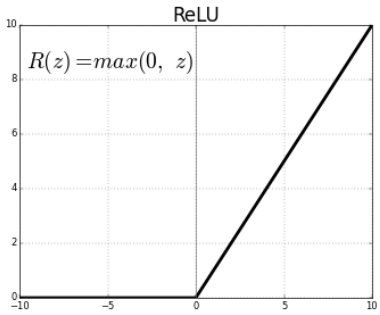
Feature Extraction with Data Augmentation

We used the pre-trained model by adding our Dense layers to end of convolutional base. This allowed us to use data augmentation. But however, this method is slower compared to if the pre-trained models were used as is. But we hope that using data augmentation will prevent overfitting and improve the accuracy.

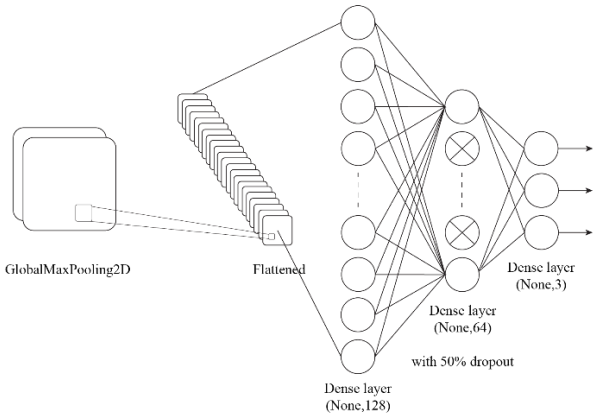
Each model was fitted with a top layer consisting of GlobalAveragePooling2D layer to average the tensors, followed by a flattened layer to convert all the resultant 2-Dimensional arrays from pooled feature maps into a single long continuous linear vector; followed by a fully connected layer. This was followed by a couple of Dense layers with an output shapes of (None, 128) and (None, 64), both accompanied by the 'ReLU' linear activation function, with a 50% dropout layer in between. The output shape was set at (None,3), corresponding to the 3 classes we wish to classify the data, with the Softmax activation function.



Global Max Pooling 2D



ReLU activation function



Top Layer

The models are compiled with the loss function 'categorical-cross-entropy' with accuracy as the metric of choice. Individual images were of the shape 768x768 pixels, which were reshaped to 256x256 pixels. The batch size was set to 15 and epochs at 50. The input shape was set to (256,256,3), for the RGB channel. The output layer has output shape 3, corresponding to the three classes, and a Softmax activation function

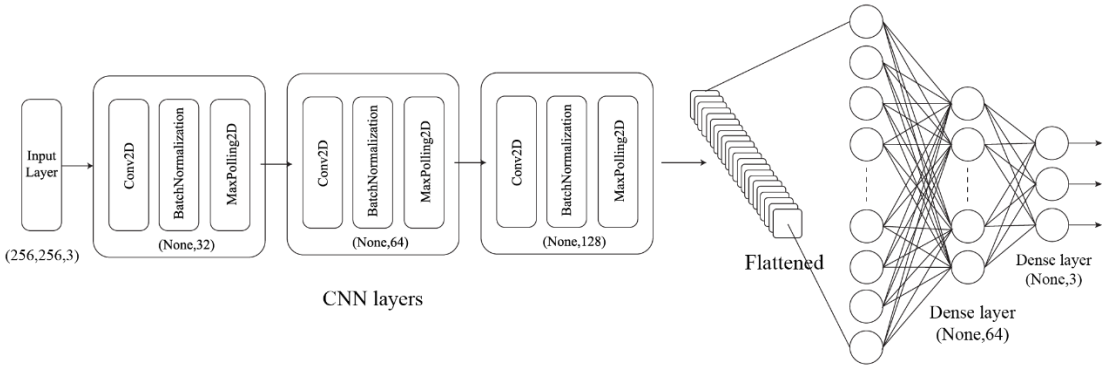
$$Loss = -\frac{1}{output\ size}\sum_{i=1}^{output\ size} y_i\ .\log \hat{y}_I + (1 - y_i)\log(1 - y_i)$$

Categorical Cross Entropy function

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

Softmax activation function

To contrast these pre-trained models, we constructed a simple CNN model, with typical architecture and weights to serve as a control. It is made of a Sequential model made with three, 2D convolution layers, with output shapes (None,32), (None,64), and (None,128), with the first layer having an input shape of (256, 256, 3). Each one of these layers has (3,3) strides of convolution with ReLU activation function. Followed by, a Batch Normalization and a 2D Max Pooling layer. The classification layers included a fully connected layer, followed three dense layers of output shapes (None,128), (None,64) and (None,3). The first two with ReLU activation function and the output dense layer with Softmax activation function.

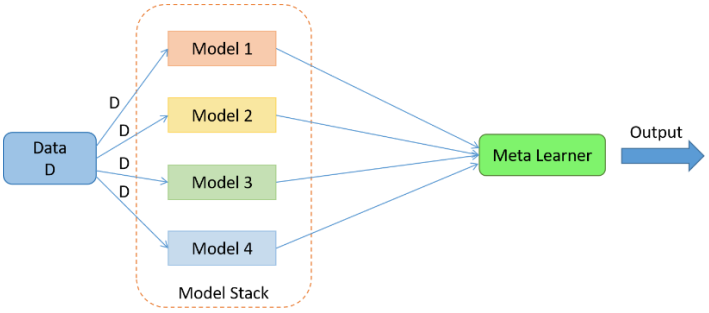


Simple CNN

$$\hat{x}^k = \frac{x^{(k)} - E[x^k]}{\sqrt{Var[x^{(k)}]}}$$

Batch Normalization function

The individual learners showed validation accuracies of over 99.5%. These being strong learners, normal ensemble processes such as bagging and boosting was not considered, and a stacking ensemble model was implemented. The meta learner in this case was chosen to be a Multinomial Logistic Regression Model with a 'lbfgs' solver, and the penalty parameter set to L1 regularization. The 'lbfgs' approximates second derivative matrix updates via gradient evaluations. It saves memory by just storing the most recent changes, however with big data sets, it isn't quite as quick.



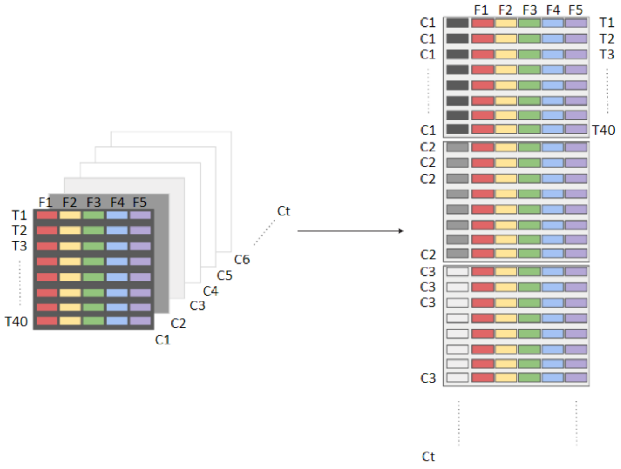
Stacked Ensemble Model

$$P = \frac{e^{a+bX}}{1 + e^{a+bX}}$$

Logistic Regression

To fit the model, we created a training dataset for the meta-learner. By giving each sub-model instances from the test set and gathering the predictions, we created a training dataset. In this situation, every model will produce three predictions for every example, along with probability that an example may fall into one of these three classifications. Thus, the test set's 3,000 samples will provide five arrays of the form [3000, 3].

Each new set of predictions may be stacked to create a three-dimensional array with the form [3000, 5, 3]. We need a few samples with some characteristics as input for the new model. Given that we have five separate sub-models and that each of them provide three predictions for each example, the sub-models would yield 15, i.e.(3x 5) characteristics for each example. To train a meta-learner and flatten the final two dimensions, we may convert the [3000, 5, 3] shaped predictions from the sub-models into a [3000, 15] shaped array.

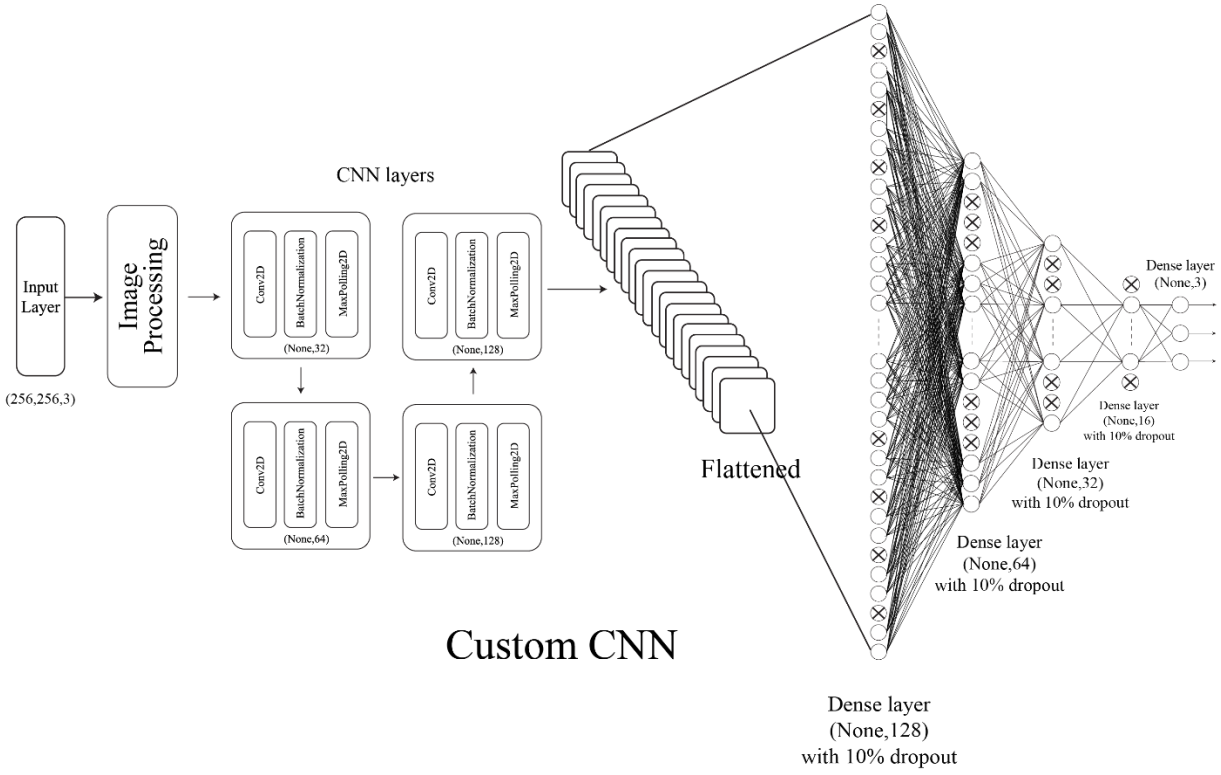


Reshaping a 3-dimensional array to a 2-dimensionl array

4.2.2 ***Custom CNN Model***

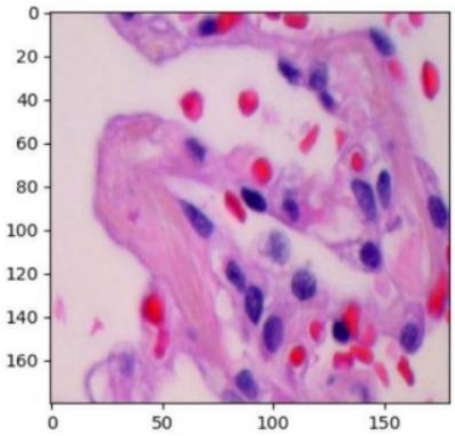
In parallel, we also developed a CNN model from scratch which gave us a validation accuracy of 97.73%. Its construction is as follows.

- For the classification, a linear stack of CNN layers with kernel filtering, max pooling, and fully connected layers was used.
- The neural network is constituted of four 2D convolution layers with input shapes 16, 32, 64 and 128 respectively, with 3*3 kernel matrix and ReLu activation function. Each of these layers are followed by a Batch Normalization and 2*2 Max Pooling layer.v
- The input shape is taken to be 256, 256, 3
- A fully connected layer is followed by 4 hidden layers, with shapes (None, 128), (None, 64), (None, 32), (None, 16) with ReLu activation, and 10% dropout.
- The output layer has output shape of (None,3), corresponding to the three classes, and a Softmax activation function
- The Adam optimizer with Categorical Cross entropy as the loss function and accuracy as the metric of choice is used to compile the model

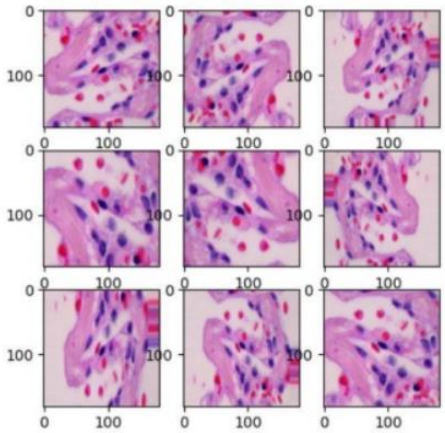


To further, optimize the output, the following image segmentation process was used,

- The data set contains images whose size is 768x768 pixels
- Every image is broken down into 9, 256x256 pixel images belonging to the same class.
- Each of these individual images is flipped horizontally and vertically such that every 256x256 image is now a set of 3 images, all belonging to the same class, and representing the same area but in a different perspective.
- Thus the dataset which originally had only 12000 images for training, now has 12000*9*3 = 324000 images for training.
- A similar process is applied for the testing/validation purpose.



Original Image



Segmented Image

Although this approach did increase the accuracies slightly for some models such as ResNet50 and ResNet101, for the majority, it did not. It was determined that this effect was probably due to the loss of features, in some images when they were broken into the 9 separate parts. Due to its failure to generate any significant change in the original result, it was not considered in the final testing phase.

5. RESULT AND DISCUSSION

5.1 Experimental setup

The model training and validation along with all exploratory data analysis were performed on a Kaggle notebook, with the help of a 16 GB RAM unit and a GPU accelerator, consisting of an NVIDIA TESLA P100 GPU unit. The procedures were coded in Python with no specific environment preferences.

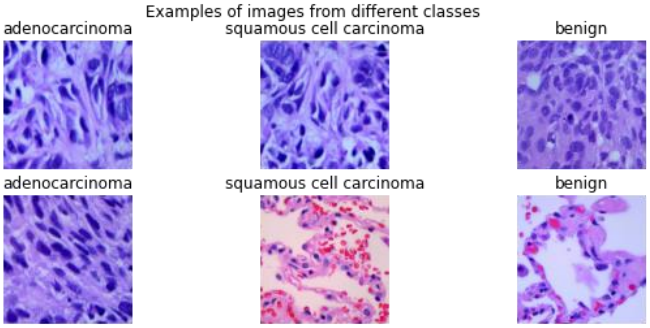
5.2 Dataset

The Lung and Colon Cancer Histopathological Image dataset is a collection of 25,000 histopathology images for lung and colon cancer organized into 5 categories. Every image is 768 by 768 pixels in size and is in the jpeg format.

The images were generated from an original sample of HIPAA compliant and validated sources, consisting of 15000 total images of lung tissue (5000 benign lung tissue, 5000 lung adenocarcinomas, and 5000 lung squamous cell carcinomas) and 10000 total images of colon tissue (5000 benign colon tissue and 5000 colon adenocarcinomas) and augmented to 25,000 using the Augmentor package.

The collection includes five classes, each containing 5,000 images:

- Lung benign tissue
- Lung squamous cell carcinoma
- Lung adenocarcinoma
- Colon benign tissue
- Colon adenocarcinoma



In this article, we have only considered the images of the lung cancer tissue images and their associated benign tissue samples.

Evaluation Parameters

Out of the 18 different models, compared against one another against an 80%-20% training-validation split, i.e. 12000 and 3000 images respectively in each section. Only the top 5 models were selected for further processing. This selection was based on the model's overall training and validation accuracies and errors. The models with the most consistent maximum accuracies and minimum losses across epochs were considered for the ensemble process. These were, EfficientNetB7, VGG16, ResNet101, ResNet152 and ResNet50.

The Ensemble model was evaluated based on its accuracy, precision, recall, and overall F1 score, based on its confusion matrix.

$$precision = \frac{TP}{TP + FP}$$

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

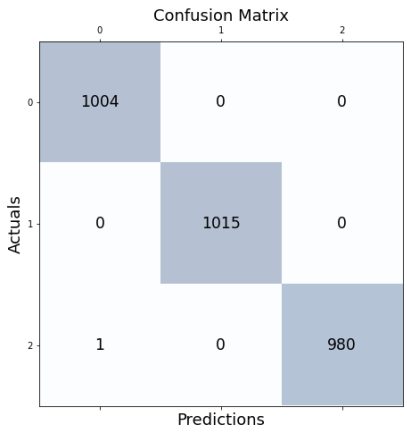
$$recall = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$F1 = \frac{2 * precision * recall}{precision + recall}$$

5.3 Experimental Results

The stacking model may be used to generate predictions on new data after being fitted. To do this, create an input dataset for the meta-learner using the sub-models, and then use that dataset to generate a prediction. The accuracy, precision, recall, and F1-score values of the model metrics are obtained by comparing these predictions to the dataset's known and anticipated outputs.



Where 0 represents Adenocarcinoma, 1 represents benign and 2 represents Squamous cell carcinoma tissues respectively.

This gives us the following scores

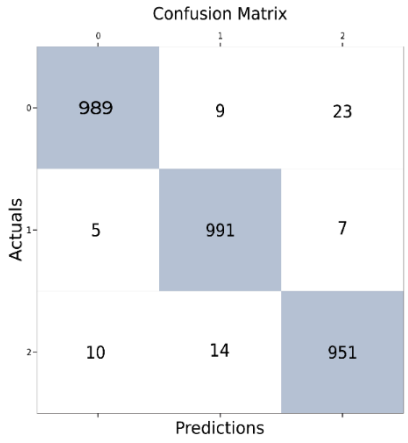
- Accuracy: 0.9996666667
- Precision: 0.9996666666666667
- Recall: 0.9996666666666667
- F1: 0.9996666666666667

Class	Accuracy	Precision	Recall	F1 Score
1	99.97%	1.0	1.0	1.0
2	100.0%	1.0	1.0	1.0
3	99.97%	1.0	1.0	1.0

While testing it was found that the model was extremely accurate in distinguishing between cancerous and benign cells/tissues. However, upon testing with randomly collected images from different sources, we found that the model was only about 90% accurate in distinguishing between adenocarcinoma and squamous cell carcinoma cells/tissues.

In contrast, the model developed by us showed an overall accuracy of 97.73% with the following metric scores per class.

Class	Accuracy	Precision	Recall	F1 Score
1	98.43%	0.97	0.99	0.98
2	98.83%	0.99	0.98	0.98
3	98.23%	0.98	0.97	0.97

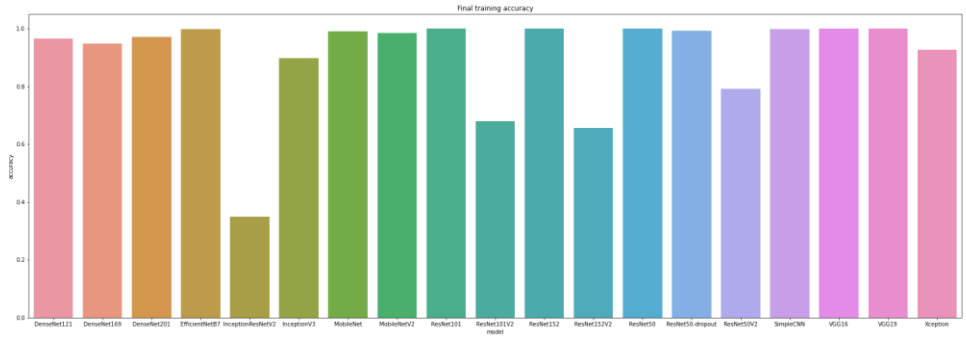


However, even in the case of this custom CNN model, we found that similar to the ensemble model, it also struggled with distinguishing between adenocarcinoma and squamous cell carcinona tissues but was quite accurate in identifying benign and cancerous lung tissue.

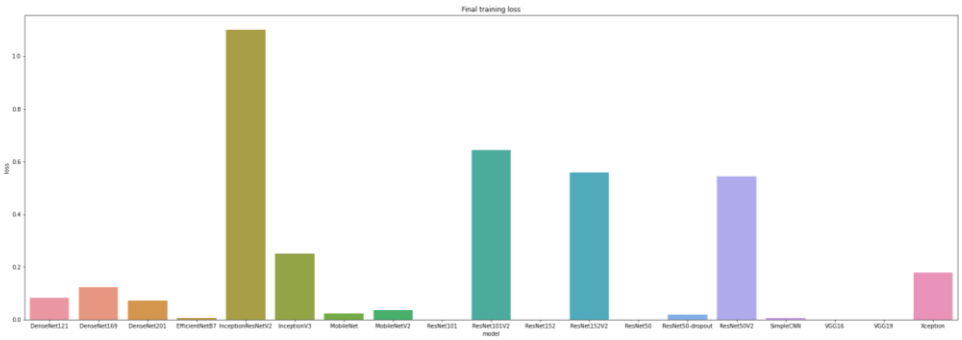
5.4 Comparative Analysis

From the initially available models from Tensorflow, we selected the top 17 and compared them with a simple CNN model as a control.
This simple CNN model was made of a Sequential model made with three, 2D convolution layers, with output shapes (None,32), (None,64), and (None,128). Each one of these layers has (3,3) strides of convolution with the ReLU activation function. It is followed by, a Batch Normalization and 2D Max Pooling layer.

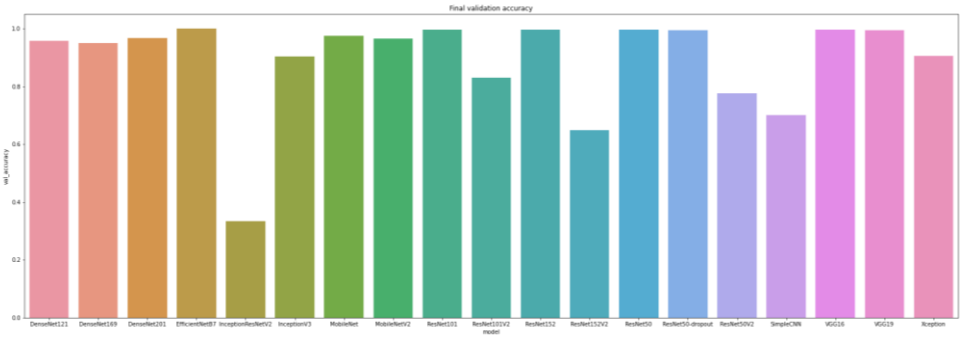
The accuracies and losses of each of these models were compared against each other.
These include, ResNet50, ResNet50-dropout, ResNet50V2, InceptionV3, VGG16, VGG19, EfficientNetB7, Xception, InceptionResNetV2, ResNet101, ResNet101V2, ResNet152, ResNet152V2, DenseNet121, DenseNet169, DenseNet201, MobileNet, MobileNetv2, SimpleCNN



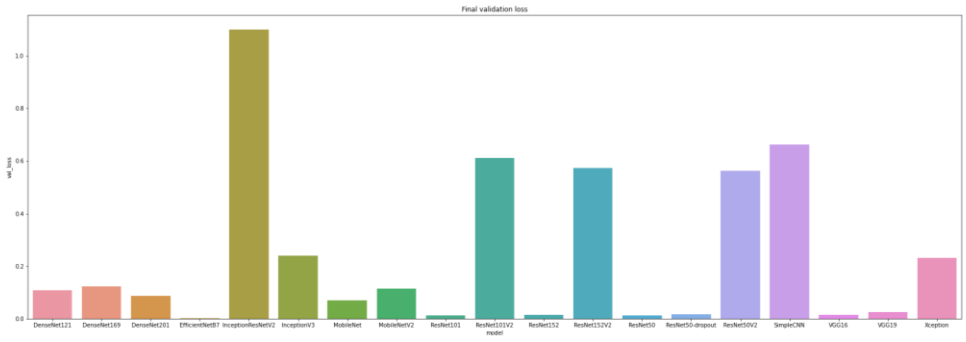
Final Training Accuracies of all pre-trained Models



Final Training losses of all pre-trained Models

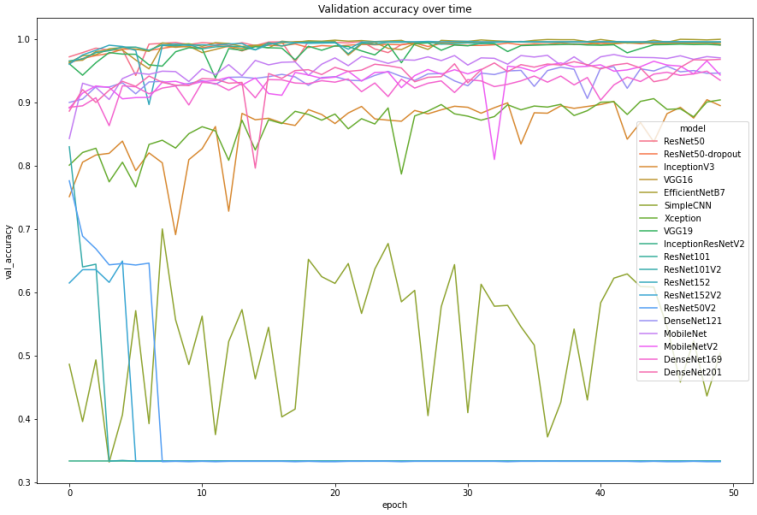


Final Validation accuracies of all pre-trained Models



Final Validation losses of all pre-trained Models

Comparing the Validation and Training accuracies and losses over 50 epochs

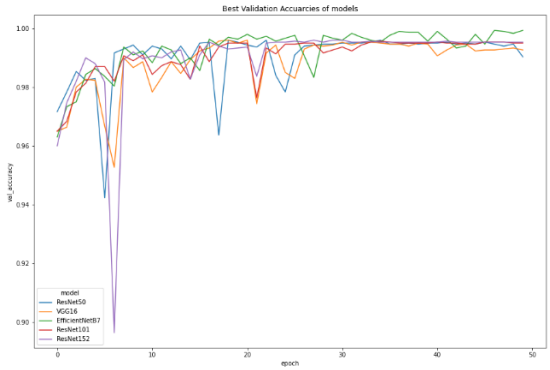


Validation accuracies of all 18 models over 50 epochs

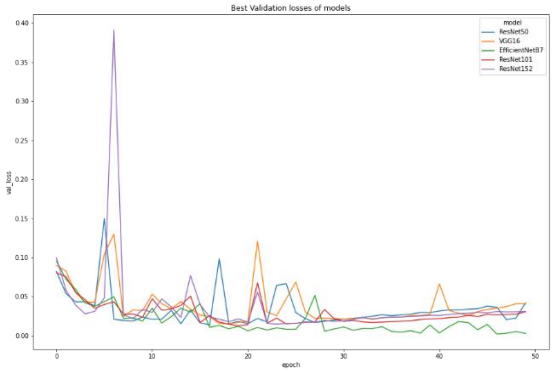
We see that the entire region is separated into 4 sections. One with accuracies below 70%, one with accuracies between 70% and 90%, one with accuracies within 80% and 95%, and one other section with accuracies above 99.5%. For our final consideration we take into account only those models whose validation accuracies are above 99.5%.

As we can see, fine-tuned models are better than simple CNN and rich reach high accuracy very quickly. InceptionV3 is better than Simple CNN but significantly less accurate than other fine-tuned models. The picture with a self-written CNN is quite chaotic. EfficientNetB7 showed the best accuracy among fine-tuned models, but this difference is not so significant (since all fine-tuned models except InceptionV3 have validation accuracy > 0.99). The fine-tuning approach is better than a simple self-built CNN model. Among fine-tuned models, InceptionV3 is not good, however. On some other hand, a ResNet50-dropout will probably work better since it should be less overfitted.

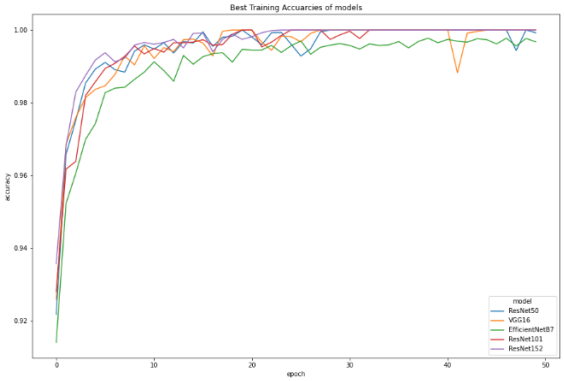
The top 5 models were found to be, ResNet50, ResNet101, ResNet152, VGG16 and EfficientNetB7



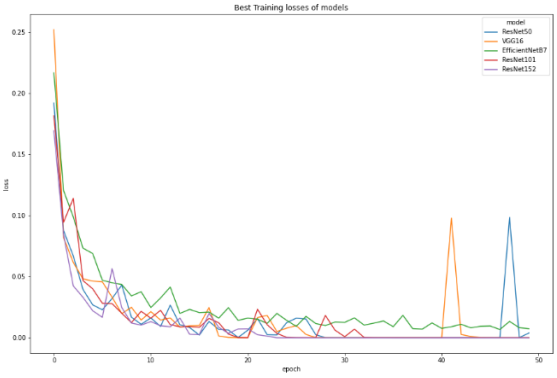
Validation accuracies of the top 5 models across 50 epochs



Validation losses of the top 5 models across 50 epochs



Training accuracies of the top 5 models across 50 epochs



Training losses of the top 5 models across 50 epochs

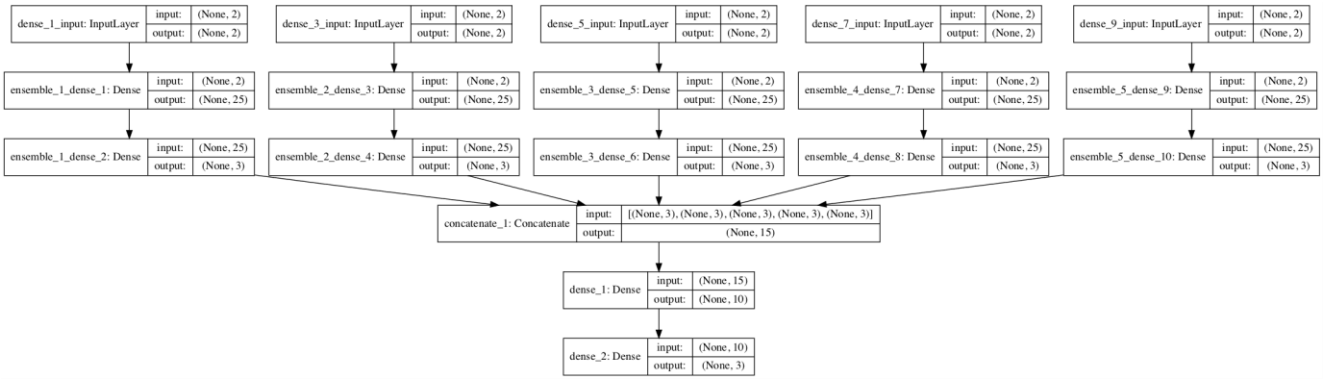
6. CONCLUSION AND FUTURE WORK

This study shows lung cancer detection utilizing histopathology scans. Two models were implemented. A main stacked ensemble model and a secondary CNN (convolutional neural network) were used to categorize pictures into three categories: benign, adenocarcinoma, and squamous cell carcinoma. The primary model attained an overall accuracy of more than 99%, while the secondary model reached an accuracy of roughly 97%. To assess model performance, precision, F1-score, and recall were computed, and a confusion matrix plot was created. On the one hand, the major ensemble model had accuracy and F1 scores over 0.99, but the secondary model had a precision of 0.97 and an F1 score of 0.98.

Using routinely collected invasive histopathological pictures from biopsies, we explored the efficacy of CNNs in predicting malignant histology in patients. CNNs were proven to be successful at natural image identification tasks and may be used to differentiate between benign and malignant lung histology as well as the most frequent histopathologic subgroups in NSCLC. CNNs can recognize small changes in pictures to predict phenotypes when given enough labeled samples. We were able to limit the chance of over-fitting by using pre-trained models. Our best-performing model was able to discriminate between malignant and benign histology with ease. It detected squamous cell carcinoma (SCC) with better specificity than sensitivity when compared to adenocarcinoma, suggesting greater potential in computer-assisted diagnosis and utility as a screening tool. AI-assisted tumor histology prediction has the potential to improve pathologists' accuracy and efficiency, resulting in considerable cost and time savings.

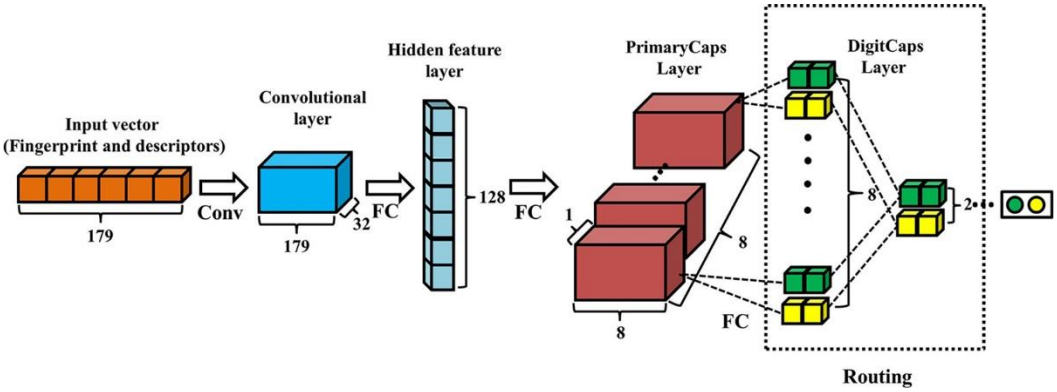
Future work:

Integrated Stacking Model: When using neural networks as sub-models, it may be desirable to use a neural network as a meta-learner. Specifically, the sub-networks can be embedded in a larger multi-headed neural network that then learns how to best combine the predictions from each input sub-model. It allows the stacking ensemble to be treated as a single large model. The benefit of this approach is that the outputs of the sub-models are provided directly to the meta-learner. Further, it is also possible to update the weights of the sub-models in conjunction with the meta-learner model.



Validation of our methodology using histopathology images from an independent tissue microarray database obtained from clinics and medical colleges should be done.

To improve the network's performance in situations when it struggles, class activation maps are evaluated. For class invariants, the CNN employs max-pooling, although class equivalence is required for biological pictures. The main difficulty with this approach is that a considerable amount of important data is lost, and CNN is also a poor model of the human visual system. Its importance suggests that capsule neural networks may train on far less input and generate results that are more accurate. Capsule Networks (CapsNet) are networks that may retrieve geographical information and additional significant aspects in order to combat the information loss experienced during pooling operations. Let's examine the distinctions between a capsule and a neuron. As an output, Capsule provides us with a vector that has a direction. For instance, if you rotate a picture, the vector will travel in the same direction, but a neuron's output is a scalar quantity that conveys no information about direction.



Convolution-capsule network architecture (Conv-CapsNet). 179 components make up the input's one-dimensional vector. 32 1-inch filters make up the convolution layer. Each layer has 128 nodes for the hidden feature layer and 64 nodes for the PrimaryCaps layer. Two iterations of dynamic routing were used, using a weight matrix of 8x8x2 between the PrimaryCaps and DigitCaps layers. [10]

REFERENCES

1. Deep learning delivers early detection. Nature, 18 November, 2020, by Elizabeth Svoboda. DOI: https://doi.org/10.1038/d41586-020-03157-9
2. Liu Y. Application of artificial intelligence in clinical non-small cell lung cancer. Artif Intell Cancer 2020; 1(1): 19-30 [DOI: 10.35713/aic.v1.i1.19]
3. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction 2nd Edition. (2009). DOI: https://doi.org/10.1007/978-0-387-84858-7
4. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning with Applications in R. (2013). 10.1007/978-1-4614-7138-7
5. Notley, S., & Magdon-Ismail, M. Examining the use of neural networks for feature extraction: a comparative analysis using deep learning, support vector machines, and k-nearest neighbor classifiers. arXiv [csLG] (2018). http://arxiv.org/abs/1805.02294
6. Setiono, R. & Liu, H. Feature extraction via Neural networks. In Feature extraction, construction and selection: a data mining perspective (eds Liu, H. & Motoda, H.) 191–204 (Springer, Boston, MA, 1998).
7. Hall, P., Park, B. U. & Samworth, R. J. Choice of neighbor order in nearest-neighbor classification. Ann. Stat. 36, 2135–2152 (2008).
8. Djuric, U., Zadeh, G., Aldape, K. & Diamandis, P. Precision histology: how deep learning is poised to revitalize histomorphology for personalized cancer care. NPJ. Precis. Oncol. 1, 22 (2017). DOI: 10.1038/s41698-017-0022-1
9. Yu, KH., Zhang, C., Berry, G. et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun 7, 12474 (2016). https://doi.org/10.1038/ncomms12474
10. Wang Y, Huang L, Jiang S, Wang Y, Zou J, Fu H and Yang S (2020) Capsule Networks Showed Excellent Performance in the Classification of hERG Blockers/Nonblockers. Front. Pharmacol. 10:1631. doi: 10.3389/fphar.2019.01631
11. https://towardsdatascience.com/types-of-convolution-kernels-simplified-f040cb307c37
12. Deep Residual Learning for Image Recognition. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. https://doi.org/10.48550/arXiv.1512.03385
13. R. P.R., R. A. S. Nair and V. G., "A Comparative Study of Lung Cancer Detection using Machine Learning Algorithms," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), 2019, pp. 1-4, doi: 10.1109/ICECCT.2019.8869001.
14. Chen, S. Feng, and D. Pan. An improved approach of lungs image segmentations based on watershed algorithms, 7th International Conference on Internet Multimedia Computing & Service - ICIMCS '15. 2015. Zhangjiajie, Hunan, China: ACM New York, NY, USA ©2015.
15. Jin, Zhang, Y., & Jin. Pulmonary Nodules Detections Based on CT Image Using Convolutional Neural Networks, 9th International Symposium on Computational Intelligence and Design, 2016
16. Xinyan Li, S.F., Daru Pan. Enhanced lungs segmentations in chests CT images based on kernel graphs cut, International Conference on Internet Multimedia Computing and Service. 2016. Xi'an, China: ACM New York, NY, USA ©2016.
17. Elnakib, A., H.M. Amer, and F.E.Z. Abou-Chadi, Computer aided detection systems for early cancerous pulmonary nodule by optimizing deep learning feature, 8th International Conference on Software and Information Engineering
18. Makaju, S., et al., Lungs Cancer Detections using CT Scan Image, Procedia Computer Science, 2018. 125: p. 107-114. Armato, McLennan, Gray, Michael, Charles, Reeves, Anthony P., Clarke, Laurenc, Data from LIDC-IDRI. The Cancers Imaging. 2015.
19. Zhuoqi Sun, S.H. Idiopathic Interstitial Pneumonias Medical Image Detections Using Deep Learning Technique: A Survey, ACM Southeast Conference (ACMSE 2019), April 18-20, 2019, : ACM, New York, NY, USA

20. Pankaj Nanglia,Sumit Kumar, Aparna N. Mahajan, Paramjit Singh, Davinder Rathee,"A hybrid algorithm for lung cancer classification using SVM and Neural Networks",The Korean Institute of Communication and Information Science(KICS),2020.
21. Ruchita Tekade, Prof. DR. K. Rajeswari, "Lung Cancer Detection and Classification using Deep Learning" at the Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), IEEE, 2018
22. Moffy Vas, Amita Dessai, "Lung Cancer detection system using lung CT image processing", IEEE,2017
23. Ausawalaithong, Worawate & Thirach, Arjaree & Marukatat, Sanparith & Wilaiprasitporn, Theerawit. (2018). Automatic Lung Cancer Prediction from Chest X-ray Images Using the Deep Learning Approach. 1-5. 10.1109/BMEiCON.2018.8609997.
24. Teramoto, Atsushi & Tsukamoto, Tetsuya & Kiriyama, Yuka & Fujita, Hiroshi. (2017). Automated Classification of Lung Cancer Types from Cytological Images Using Deep Convolutional Neural Networks. BioMed Research International. 2017. 1-6. 10.1155/2017/4067832.
25. Rahane, Wasudeo & Dalvi, Himali & Magar, Yamini & Kalane, Anjali & Jondhale, Satyajeet. (2018). Lung Cancer Detection Using Image Processing and Machine Learning HealthCare. 1-5. 10.1109/ICCTCT.2018.8551008.
26. Hatuwal, Bijaya & Thapa, Himal. (2020). Lung Cancer Detection Using Convolutional Neural Network on Histopathological Images. International Journal of Computer Trends and Technology. 68. 21-24. 10.14445/22312803/IJCTT-V68I10P104.
27. S. Sasikala, M. Bharathi, B. R. Sowmiya. "Lung Cancer Detection and Classification Using Deep CNN." (2019).
28. Thiyagarajan, Manikandan. (2019). A Computer-Aided Diagnosis System for Lung cancer detection. 10.35940/ijitee.A1081.1191S19.
29. Aggarwal, Taruna & Furqan, Asna & Kalra, Kunal. (2015). Feature extraction and LDA based classification of lung nodules in chest CT scan images. 1189-1193. 10.1109/ICACCI.2015.7275773.
30. P. B. Sangamithraa and S. Govindaraju, "Lung tumour detection and classification using EK-Mean clustering," 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2016, pp. 2201-2206, doi: 10.1109/WiSPNET.2016.7566533.
31. Zhou, A.M.R.W. Deep Learning for Categorizations of Lungs Cancer CT Image, IEEE/ACM International conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE). 2017. Philadelphia, PA, USA: IEEE.
32. Yu, KH., Zhang, C., Berry, G. *et al.* Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun* **7,** 12474 (2016). https://doi.org/10.1038/ncomms12474
33. Svoboda E. Artificial intelligence is improving the detection of lung cancer. Nature. 2020 Nov;587(7834):S20-S22. doi: 10.1038/d41586-020-03157-9. PMID: 33208974.
34. Liu Y. Application of artificial intelligence in clinical non-small cell lung cancer. *Artif Intell Cancer* 2020; 1(1): 19-30 [DOI: 10.35713/aic.v1.i1.19] [Cited by in CrossRef: 2] [Cited by in F6Publishing: 2] [Article Influence: 1.0]
35. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction 2nd Edition. (2009)(no citation)
36. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning with Applications in R. (2013)
37. Yu, KH., Zhang, C., Berry, G. *et al.* Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun* **7,** 12474 (2016). https://doi.org/10.1038/ncomms12474
38. https://arxiv.org/pdf/1512.03385v1.pdf
39. Djuric U, Zadeh G, Aldape K, Diamandis P. Precision histology: how deep learning is poised to revitalize histomorphology for personalized cancer care. NPJ Precis Oncol. 2017 Jun 19;1(1):22. doi: 10.1038/s41698-017-0022-1. PMID: 29872706; PMCID: PMC5871847.
40. Notley, S., & Magdon-Ismail, M. (2018). Examining the Use of Neural Networks for Feature Extraction: A Comparative Analysis using Deep Learning, Support Vector Machines, and K-Nearest Neighbor Classifiers. *ArXiv, abs/1805.02294*.
41. Setiono, R., Liu, H. (1998). Feature extraction via Neural networks. In: Liu, H., Motoda, H. (eds) Feature Extraction, Construction and Selection. The Springer International Series in Engineering and Computer Science, vol 453. Springer, Boston, MA. https://doi.org/10.1007/978-1-4615-5725-8_12
42. Hall, P., Park, B. U., & Samworth, R. J. (2008). Choice of Neighbor Order in Nearest-Neighbor Classification. *The Annals of Statistics*, *36*(5), 2135–2152. http://www.jstor.org/stable/25464706
43. Wang Y, Huang L, Jiang S, Wang Y, Zou J, Fu H and Yang S (2020) Capsule Networks Showed Excellent Performance in the Classification of hERG Blockers/Nonblockers. *Front. Pharmacol.* 10:1631. doi: 10.3389/fphar.2019.01631
44. Barron, Jonathan & Malik, Jitendra. (2010). Discovering Efficiency in Coarse-To-Fine Texture Classification.
45. https://ai.googleblog.com/2019/05/efficientnet-improving-accuracy-and.html#:~:text=EfficientNet%2DB0%20is%20the%20baseline,than%20the%20best%20existing%20CNN. (*Hongkun Yu, Ruoming Pang, Vijay Vasudevan, Alok Aggarwal, Barret Zoph, Xianzhi Du, Xiaodan Song, Samy Bengio, Jeff Dean, and the Google Brain team.*)
46. https://venturebeat.com/2019/05/29/googles-efficientnets-is-faster-at-analyzing-images-than-other-ai-models/
47. He, Kaiming et al. "Deep Residual Learning for Image Recognition." *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016): 770-778.
48. Ramagiri, Dr. Venkata & Lakshmi, D. & Sunitha, K.V.N.. (2012). Feature Extraction Methods For Color Image Similarity. Advanced Computing : an International Journal. 3.
49. https://machinelearningmastery.com/convolutional-layers-for-deep-learning-neural-networks/

50. Tianming Song & Xiaoyang Yu & Shuang Yu & Zhe Ren & Yawei Qu, 2021. "Feature Extraction Processing Method of Medical Image Fusion Based on Neural Network Algorithm," Complexity, Hindawi, vol. 2021, pages 1-10, October.
51. https://www.cancer.gov/about-cancer/understanding/what-is-cancer
52. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021 May;71(3):209-249. doi: 10.3322/caac.21660. Epub 2021 Feb 4. PMID: 33538338.
53. Prabhakar B, Shende P, Augustine S. Current trends and emerging diagnostic techniques for lung cancer. Biomed Pharmacother. 2018 Oct;106:1586-1599. doi: 10.1016/j.biopha.2018.07.145. Epub 2018 Jul 28. PMID: 30119234.
54. Erratum: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Volume 70Issue 4CA: A Cancer Journal for Clinicians pages: 313-313 First Published online: April 6, 2020. Freddie Bray BSc, MSc, PhD,Jacques Ferlay ME,Isabelle Soerjomataram MD, MSc, PhD,Rebecca L. Siegel MPH,Lindsey A. Torre MSPH,Ahmedin Jemal PhD, DVM. https://doi.org/10.3322/caac.21492
55. He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778) 2016