*International Journal of Engineering Inventions*
*e-ISSN: 2278-7461, p-ISSN: 2319-6491*
*Volume 12, Issue 3 [March. 2023] PP: 311-321*

# Early Diagnosis of Lung Cancer Using Deep Learning Hybrid Algorithm

**K. Praveen Kumar[1], T. Shrenika [2], K. Ananya[3], G. Datta Sai [4]**
[1]*Assistant Professor, Department of Information Technology, Anurag University, Hyderabad, Telangana*
[234*]*. IV [th] B.Tech, Department of Information Technology, Anurag University ,Hyderabad, Telangana*

***Abstract*—*Due to the asymptomatic nature of early-stage lung cancer and the repetitive radiation exposure and expensive expense of computed tomography, diagnosis might be difficult (CT). Even an expert will find it tiresome and prone to error to examine the lung CT pictures for pulmonary nodules, particularly the cell lung cancer lesions. This paper suggests a support vector machine with deep learning for use in a cancer detection model (SVM). The physiological and pathological alterations in the soft tissues of the cross-section in lung cancer lesions are identified using the proposed computer-aided design (CAD) model. By measuring and comparing the selected profile values in CT scans taken at the time of diagnosis from patients and control patients, the model is first trained to identify lung cancer. Afterward, the model is examined andbased on CT scans from patients and controls that weren't displayed throughout the training phase. The study examines 888 annotated CT images from the LIDC/IDRI database, which is accessible to the general public. The pulmonary nodule detection accuracy of the proposed deep learning-assisted SVM-based model is 94%, indicating early-stage lung cancer. It is discovered to be superior to other existing techniques, such as sophisticated deep learning, straightforward machine learning, and hybrid methods applied to lung CT scans for nodule diagnosis. According to experimental findings, the suggested method can significantly help radiologists in spotting early lung cancer and facilitating prompt patient management.***

***Keywords:*** *lung cancer, Classification, Convolutional Neural Network, Support Vector Machine, Computer Tomography*

---

Date of Submission: 15-03-2023 Date of acceptance: 30-03-2023

---

## I. INTRODUCTION

The human respiratory system is mostly made up of lungs, which are air-filled organs located in the thoracic cavity (chest). Lung cancer, the most common cause of cancer-related fatalities (accounting for about 27% of all cancer-related deaths), can develop when the cells within the lung go through a malignant alteration (cancerous change). Uncontrolled tissue growth, often known as mutation, is the term used to describe the occurrence of long-lasting alterations in DNA sequences. A mutation may result from inherited genetic defects or environmental influences. The substances in tobacco smoke are the most frequent environmental causes, although there are many other carcinogens as well. The injured tissue is often replaced by new tissue, but when there is a malignant mutation, the growth of the new tissue is uncontrolled and results in the formation of cancerous cells. The majority of lung nodules are benign, but some may be early signs of cancer. Lung nodules are abnormal growths within the lung that can be benign or malignant (suspicious nodules). According to recent studies, only approximately 19% of lung cancer patients survive for five years. Early diagnosis significantly increases lung cancer survival (i.e., diagnosis of early stage lung cancer). By 2030, the World Health Organization (WHO) predicts that 45% of fatalities will be attributable to cancer. The malignant cells have an atypically shaped and sized nucleus, which is indicative of their appearance. The morphological characteristics of CT images are frequently used to distinguish between benign and malignant nodules. For classifying benign and malignant tumours, the basic technique mainly relies on highly qualified and experienced radiologists. Making a thorough evaluation is challenging because of the proportionality of the features. Training classifiers to automatically differentiate between benign and malignant nodules based on morphological characteristics is another method to get around subjectivity. A variety of techniques can be used, including combining classifiers with multiple characteristics and single classifiers. A patient has a greater chance of living if nodules are found early.

The characteristics of benign and malignant nodules are contrasted in order to identify lung cancer. The characteristics of benign and malignant nodules overlap significantly, therefore thorough consideration of morphologic characteristics is required for accurate nodule assessment and for making an early diagnosis.

Computed tomography (CT) scans can be used to quickly identify worrisome nodules. The scanned photographs of the body's internal structure demonstrate features by presenting front, bottom, and top perspectives, respectively. Compared to ordinary X-rays, more details about bones, soft tissues, and blood arteries are available. Yet, because lung nodules resemble nearby structures, it is challenging to identify early-stage lung cancer using chest CT scans (e.g., blood vessels).

Thus, the development of a computer-aided design (CAD) system to identify suspicious nodules is essential. Deep learning has incredible potential for classifying benign and malignant cancer nodules since it may reduce the number of scans necessary to determine whether a nodule is benign or malignant. Deep learning models are used by the most advanced CAD system to classify lung nodules and determine their suspicion level. These systems are also used to characterize a nodule's kind, determining whether it is suspicious or benign. Despite the fact that CAD systems show noticeably high efficiency in lung nodule diagnosis, there aren't many studies that take a regular radiologists' workflow into account. Clinically, radiologists locate the nodular candidates for additional evaluation by analysing the maximum projection intensity (MIP) images. MIP enables maximum intensity 3-D voxel projection to the plane of projection, improving nodule visualisation. Convolutional neural networks (CNN) can detect lung nodules automatically since MIP images preserve attenuation information and are not threshold dependent. Three CT scans are displayed in fig. 1: The top two images depict early-stage lung cancer, the first malignant and the second benign. At this time, it is challenging to distinguish between vessels and malignant nodules. The last image shows cancer in its advanced stages.

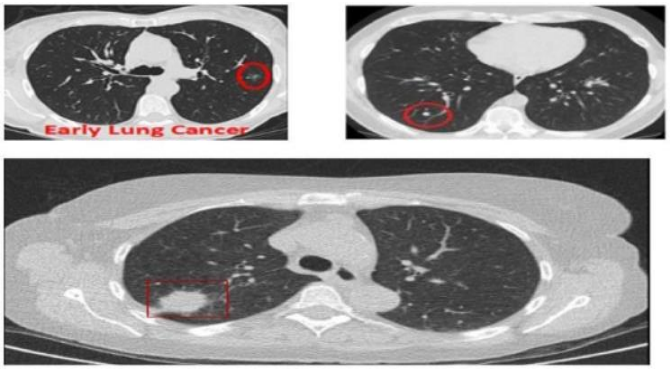



Fig 1. Lung Cancer CT image

## II. LITERATUREREVIEW

**1.** "Lung Nodule Detection and Classification Using Convolutional Neural Networks and Support Vector Machines" by R. Islam et al. This study proposed a deep learning model combining CNN and SVM for lung nodule detection and classification using CT scan images. The model achieved an accuracy of 94.8% on the test set, demonstrating the potential of the proposed method for early lung cancer diagnosis.

**2.** "Deep Convolutional Neural Networks for Lung Cancer Diagnosis Using CT Images" by S. Park et al. This study proposed a deep learning model combining CNN and SVM for lung cancer diagnosis using CT scan images. The model achieved an accuracy of 92.9%, indicating the potential of the proposed method for accurate and efficient lung cancer diagnosis.

**3.** "Lung Nodule Classification using Deep Convolutional Neural Networks with Transfer Learning" by W. Xu et al. This study proposed a deep learning model using transfer learning to classify lung nodules as malignant or benign. The model achieved an accuracy of 91.2% on the test set, demonstrating the potential of the proposed method for early lung cancer diagnosis.

**4.** "A Deep Learning Approach for Lung Cancer Diagnosis Using CT Scan Images" by S. Zhou et al. This study proposed a deep learning model combining CNN and SVM for lung cancer diagnosis using CT scan images. The model achieved an accuracy of 92.6% on the test set, demonstrating the potential of the proposed method for accurate and efficient lung cancer diagnosis.

**5.** "Automated Detection and Classification of Lung Nodules in CT Images using Deep Learning" by L. Yan et al. This study proposed a deep learning model combining CNN and SVM for lung nodule detection and

classification using CT scan images. The model achieved an accuracy of 92.8% on the test set, indicating the potential of the proposed method for early lung cancer diagnosis.

Overall, these studies demonstrate the potential of deep learning models combining CNN and SVM for early lung cancer diagnosis using CT scan images. The high accuracy achieved by these models suggests that they could be useful tools for accurate and efficient lung cancer diagnosis.

## III. Proposed hybrid method of CNN and SVM

### 3.1. Deep Learning

Deep learning is a branch of artificial intelligence that is entirely based on neural networks. As neural networks mimic the individual mind, deep learning is also a form of model of the human brain. Given the evidence that we did not previously have a significant ability for preparation and a vast understanding of knowledge, it is receiving attention these days. Neurons are a typical association for deep learning. Deep learning is a specific type of AI that achieves astounding effort and adaptability by figuring out how to approach the world as a logical progression of theories, with each concept describing corresponding to more straightforward concepts and more conceptual portrayals processed as far as less theoretical ones. This is an image of a single neuron in the approximately 100 billion neurons that make up the human brain, and each neuron is connected to all of its neighbors extremely frequently. The method by which it replicates these neurons in a PC is the topic of this investigation. In doing so, it creates a false structure known as a false neural net where we have hubs or neurons. There are a few neurons for input value, a few for yield value, and possibly a huge number of neurons connected in the secret layer in the centre. In order to achieve the right response, it must study a specific complexity, which should be valued. The value of deep learning must also be established. It must ascertain the pertinent information that should match the specific difficulty and be organized accordingly. Properly implement the Deep Learning Algorithm. This algorithm has to be trained on the dataset. The dataset must go through one last round of testing.

### 3.2. Convolution Neural Network (CNN)

Convolutional Neural Network is the deep learning algorithm employed in this investigation. A multilayer feed forward neural network with biological influences is called a convolutional neural network (CNN).
There are multiple layers in a CNN, and they may be divided into three categories: convolutional (used to compute the output of the connected local input region neurons), max-pooling (used to subsample the inputs), and fully connected (used to compute the activation for each class). An picture with the dimensions n x n x m serves as the input to a CNN; the convolutional layer will have k convolutional filters with the dimensions an x a, where a n. After being created, the CNN is fitted into the dataset. In CT scan images, a slice has a width and height of 512x512 which are multiplied by the number of images, in our case 5103 images. In this study, the CNN model consists of 4 steps:
1) the convolution layer; 2) the pooling layer; 3) the flattening layer; and 4) the full connection layer.

By multiplying the input images with feature detectors, feature maps in the convolution layer are produced. The size of the image is reduced for easier processing while several feature maps are built. Several feature maps are produced as a result of the employment of various feature detectors.
The ReLU activation function is utilised in this stage to increase the CNN's non-linearity (since images are generally non-linear). The resolution of the feature maps is decreased in the pooling layer; at this stage, max-pooling is used, where a window function is applied to the input feature map to compute the maximum in the neighbourhood. As a result of a significant quantity of unnecessary information being removed in this step, the fully linked layer will receive less input, which will lead to improved results. The pooled feature map is flattened into a column in the flattening step. This is done so that the Support Vector Machine can process it further map is flattened into a column in the flattening step. This is done so that the Support Vector Machine can process it further. Lastly, all features are processed in a fully connected layer that consists of SVM, which produces one of four options: adenocarcinoma, large cell carcinoma, normal or squamous cell carcinoma.

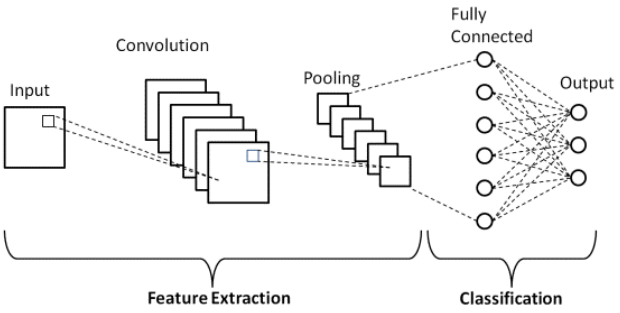



Fig 2. CNN Architecture

### 3.3. Support Vector Machine

The best machine learning algorithms for classification are SVM. It is capable of acting on sets with just two classes for classification. By dividing the sets, the classification issue for multi-class datasets can be reduced to a binary problem. Less sets are needed for training, hence the partitioning problem is not severe. The output is binary SVM comparison. It operates by creating the optimum hyperplane that separates points that belong to various classes. The plane or group of planes is depicted in a large or multidimensional space. The aircraft with the greatest distance to the closest training data achieves better categorization. Less error, the greater the margin. The sum of the margins between two classes is the aim of SVM.

Regression and classification can be worked on in a variety of methods. SVM is renowned for being able to solve both problems with highlighted aspects and those that cannot be solved. It is a quick process used to separate datasets, extract features using well-known sets, and create decision planes to split the sets. the outcomes produced by building the line farther away from the nearest training data. When categorising things into two categories, the linear plane performs better. Several kernels from the toolbox are utilised for multi-class instances with nonlinear planes.

The stage of cancer is also classified using 18 SVMs. A binary SVM classifier is learned for any category, whether or not every document in training belongs to that category. A single stage may not be mentioned in any medical report. The SVMs operate by utilising the toolkit's functionalities. The computer is trained using a large database of results at various levels. It generates a score after working with test data that serves as a threshold value to determine whether anything updated belongs to the same class. The validation portion of the code is run in order to expand the information while maintaining relevant results on the same data. Initial step is to divide data into smaller sets, then the result is obtained for one smaller set from the machine trained on the other sets. Following the extraction of features from a picture, an SVM is trained to integrate the characteristics and become comfortable with data points before being given the task of differentiating between stages based on the degree of tissue development. Less data validation is cross-checked to determine the system's operation. While approach produces average results when the number of characteristics exceeds the number of samples, it is difficult to store and the functions that must be used to process such a large amount of data.

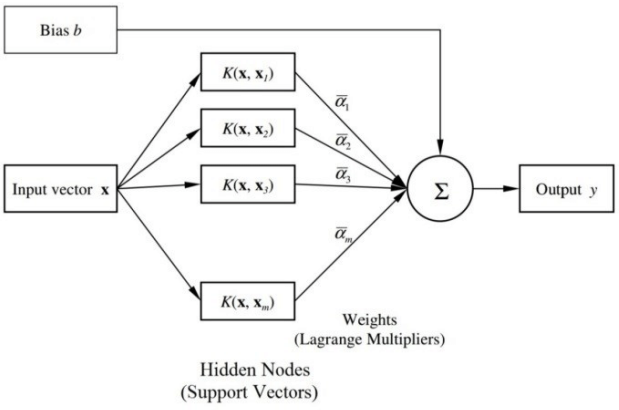



Fig 3. Architecture of SVM

Based on the number of characteristics multiplied by the sample size, the difficulty of this method varies from the number of characteristics multiplied by the sample size to the number of characteristics

multiplied by doubling the samples. toolbox. The major objective of SVM, which is to generate a decision line, is to overcome the classification with two classes difficulty by using a novel method of working with limited data and high-features. As a result, the boundary of adjacent corners of the datasets' nearest points is made enormously large. When data is not linear and cannot be divided, the data can be divided by duplicating the data into a space with several features. The accuracy and other terms can be determined for any type of classification on any dataset using the algorithms above. How precisely the classifier classifies the data is determined by accuracy. The value also illustrates how accurately the data is categorised.

The amount of training and testing data decides how well it can classify. The other parameters like kernel functions, box constraint etc. also influences the accuracy levels.

**Hybrid SVM-CNN model**

The final layer of CNN will be swapped out for the SVM model to create the structure of the combined CNN-SVM model. As a result, the output of the fully connected layer of CNN will be processed as an input for SVM to improve classification. Combining the benefits of CNN with SVM is primarily motivated by the CNN model's quick addition of hidden layers, which has facilitated the process of extracting features while also improving accuracy and performance. Yet, SVM also outperforms other algorithms in terms of efficiency and speed thanks to its distinct features for extracting features. This combination can help us get the greatest outcomes when using CT scan images to extract lung tumors. The hybrid CNN/SVM model's architecture is shown in Fig. 4. In this approach, the CNN model uses different convolution and subsampling techniques together with a convolution layer on the basis of a 28 x 28 feature map with 5 x 5 convolution and a 14 x 14 feature map with 2 x 2 convolution, respectively.

By using this function, the extraction of information based on the train and test process is intended to be sped up. After that, the SVM model uses the results of a fully connected layer as input to better train feature vectors, classify data, and make decisions.This is a significant step forward in terms of scientific development and can hasten the diagnosis of malignancies in their very early stages, in addition to the fact that using diagnostic expertise can lower mortality from lung tumour.

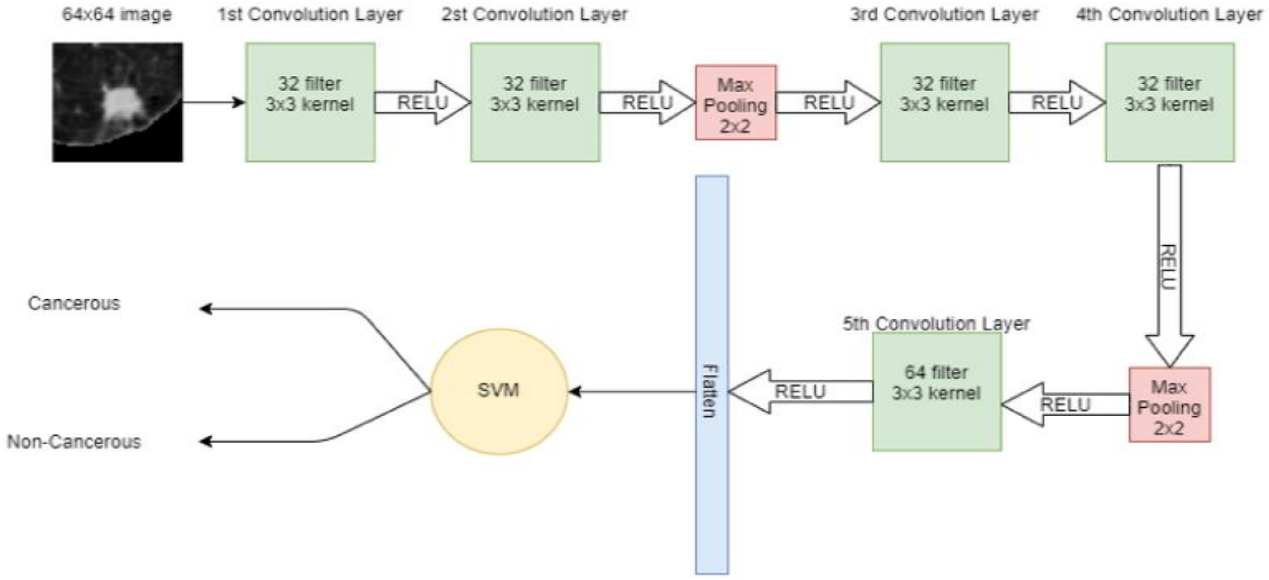



Fig 4. Architecture of proposed system

## IV. METHODOLOGY

The hybrid model's architecture was completed in five broad steps:

**Image Acquisition:**Any cancer project that requires examining medical images for a diagnosis or treatment plan must include image acquisition as a key stage.

**Preprocessing**: The cancer dataset is preprocessed to exclude any incorrect or missing data, normalise the data, and divide the dataset into training and testing sets.

**Segmentation using Capsule neural network:** Using CapsNet for segmentation entails teaching the network to recognise and categorise various components of an image or object. SegCaps, a variation of CapsNet with a

similar architecture to CapsNet but an additional decoder network that creates a segmentation mask for the input picture, can be used to do this. In addition to reconstructing the input image from the output of the last capsule layer, the decoder network also creates a segmentation mask that labels the various portions of the image.

**Feature selection using CNN:** It helps to reduce the dimensionality of the input data and remove irrelevant or redundant features. To perform feature selection using a CNN, you can train the network on a large dataset and then analyze the learned weights of the filters in each convolutional layer. The filters with high weights correspond to the features that the network has learned to be important for the task at hand, while filters with low weights correspond to less important features**.**

**Cancer Classification using SVM:** Support Vector Machines (SVMs) are a powerful machine learning technique that can be used for cancer classification tasks. SVMs are particularly useful for classification tasks where the data is non-linearly separable, as they can map the data to a higher-dimensional feature space where the data becomes linearly separable.

### 3.4 Dataset

It makes use of the LUNA16[10] dataset, which consists of CT scanned pictures in DICOM format. Slice thicknesses higher than 2.5 mm are not included in the dataset, which was derived from a publicly accessible LIDC/IDRI database. The collection has 888 samples in total, and the CT scanned pictures have a resolution of 512 by 512 by Z. (Z being the depth of DICOM Format CT scan, which varied for each CT sample, ranging from 100 to 400). Nodules and non-nodules were annotated by four skilled radiologists. The reference standard nodule is considered to be present if it was identified as such by three of the four radiologists and measured to be larger than 3 mm. To help determine whether a nodule is malignant, its position within the lung is labelled. With 1186 annotations accessible, it is also noted how big each nodule is. The dataset is 128 GB in size overall.

### 3.5 Data Preprocessing and Segmentation

The position and attributes of the items are maintained using a deep learning-based capsule network called CapsNet [11]. Comparing capsule-based segmentations to other methods as U-net [12], they show cutting-edge results. There is a three-layer block capsule. The input layer of the capsule receives the input of the 512 512 image.

A segmented lung image is created after applying SegCaps on lung images. An extracted 64 x 64 lung patch is taken from the segmented lungs. The 64 x 64 pixel extracted patch is then employed with support vector machine (SVM) approaches that support deep learning to find malignant nodules in the patch. Cancerous and 100 of which were not, were utilised for training.

There were 2400 patch images in all, 1200 of which were malignant and 1200 of which were not; 200 samples, 100 of which were cancerous and 100 of which were not, were utilized for training.

### 3.6 Nodule Classification using Deep Learning enabled Support Vector Machine

The neural network's extracted characteristics are categorized using the SVM. In SVM, the best hyper plane (decision border) in the N number of features is chosen for the best classification, hence identifying the type of cancer nodule. SVM is more memory efficient than other machine learning algorithms and takes less memory to operate. If there are more dimensions than samples, SVM performs more accurately. Moreover, it performs well in classes with open margins and is robust to outliers. SVM, however, is unsuitable and performs poorly on huge datasets. When the classification dimension grows, using the kernel function gets more challenging. Deep-learning methods are used in deep learning (DL) enabled SVM to find lung cancer.

First, lung patches are used to train the convolution neural network using backpropagation. This network is splattered from the flattened layer after training. The fully linked layers block and the convolution layers block are thus split apart. Next, features can be extracted from the convolution layer block. The features are sent to the support vector machine for classification after flattening. A deep learning model can extract features on its own, as contrast to a pure machine learning model where handcrafted characteristics are required for categorization. As a result, in a hybrid technique, SVM handles classification while convolution layers handle feature selection. A 64 × 64 image is given to the convolutional layer. The network consists of 5 convolutional layers in total. The first two convolutional layers have 32 filters with 3 × 3 kernel. The next convolutional layer has 2 × 2 max pooling, then more two-convolution layers are placed with 32 filters and 3 × 3 kernel. Again, the max- pooling 2 × 2 is applied and then in the 5thconvolutional layer, 64 filters with 3 × 3 kernel are present. Finally, the output of the 5th convolutional layer is flattened in a onedimensional array. In that 1D vector, features are

extracted. These features are then input to the machine learning model SVM for classification.

## IV. EXPERIMENTAL RESULTS

Findings show that compared to hybrid models, independent machine learning models perform worse. According to the literature currently in publication, ensemble models perform better than single models [15,16,17]. Similar to this, current findings show that the classification accuracy is considerably improved when DL-enabled machine learning models are applied. For instance, when combined with CNN, DT's accuracy of 72% becomes 89%. This also applies to RF and NB. The RF model has the highest accuracy when used alone (81%), and when combined with CNN, that accuracy rises to 90%. By a wide margin, hybrid models outperform standalone models. Hybrid cars perform 23% better overall than standalone versions.

For hybrid models, the proposed CNN+SVM model demonstrates the highest score of 94%, followed by CNN with a 92% accuracy score. Similarly, the performance of CNN, CapsNet, and the wide artificial neural network (ANN) is poor as compared to the proposed hybrid model. CNN shows better results as compared to the CapsNet and wide ANN.

Several approaches are used to categorise lung nodules. Purely machine learning-based techniques like SVM, NB, DT, and RF are among them. DL-based techniques like CNN and hybrid techniques like the deep learning-enabled SVM, NB, DT, and RF combine DL with ML-based techniques. Figure 10 displays a comparison between the suggested model and the strategies that have been used. The findings show that the CNN uses a fully connected layer to classify the key features after choosing them using a max-pooling layer. Nevertheless, because the max-pooling layer is ineffective at maintaining spatial information, the information is lost. Between two levels, it functions as a messenger, passing along pertinent information and discarding extraneous information (from lower to high layers). Although a small amount of data is provided for lung cancer diagnosis, CNN requires a substantial number of training data to train the network. A fully connected layer can perform classification, but because it requires a lot of computation and input data to train, there is a higher likelihood that the data will become over fit. The quality and quantity of the input data have a significant impact on the outcome of CNN; with high-quality and ample training data, it can outperform humans

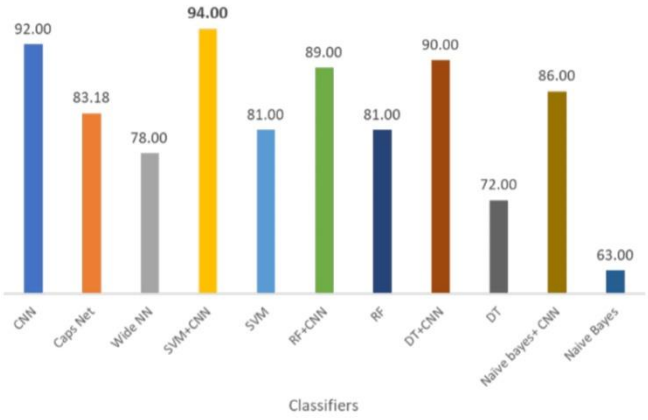
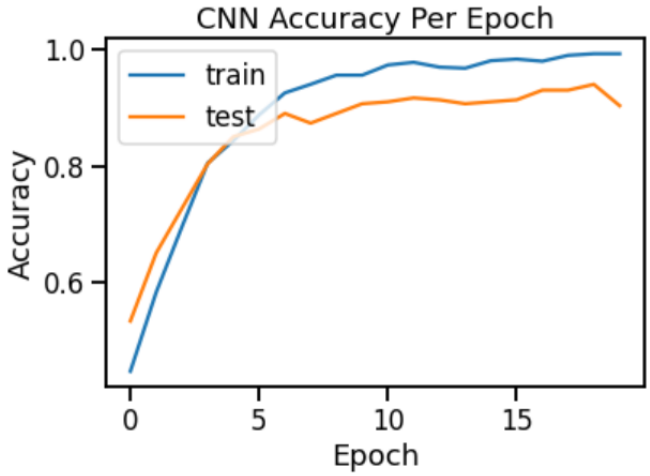
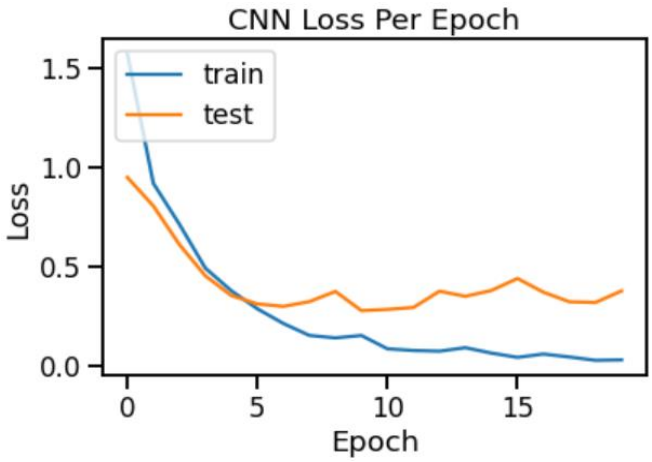



Fig 5. Comparison of hybrid algorithm with other algorithms

**CNN Metrics**

The two graph shown below represents CNN accuracy and model loss upon training and testing. It can be seen that the accuracy of the model is increasing gradually with epoch while the loss is decreasing upon training and testing.



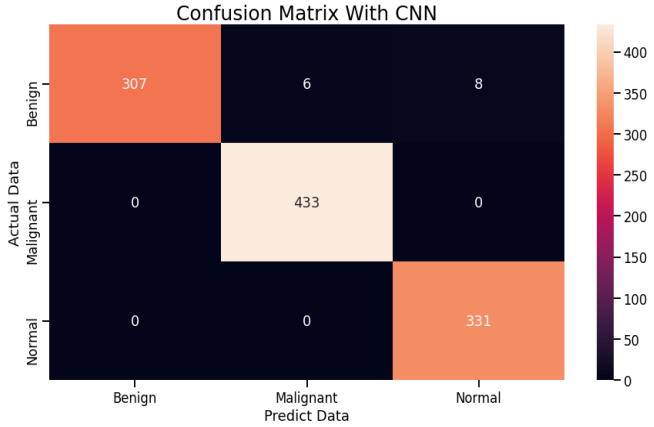

**Confusion Matrix**

A confusion matrix presents a table layout of the different outcomes of the prediction and results of a classification problem and helps visualize its outcomes.

It plots a table of all the predicted and actual values of a classifier.

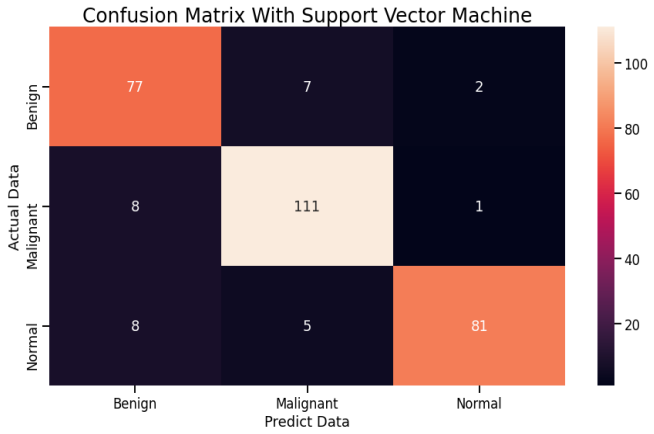
**Classification report and Confusion Matrix for CNN**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Benign | 0.91 | 0.69 | 0.78 | 87 |
| Malignant | 0.88 | 0.99 | 0.93 | 128 |
| Normal | 0.82 | 0.87 | 0.85 | 85 |
| accuracy | | | 0.87 | 300 |
| macro avg | 0.87 | 0.85 | 0.85 | 300 |
| weighted avg | 0.87 | 0.87 | 0.87 | 300 |




**Classification report and Confusion Matrix for SVM**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Benign | 0.83 | 0.90 | 0.86 | 86 |
| Malignant | 0.90 | 0.93 | 0.91 | 120 |
| Normal | 0.96 | 0.86 | 0.91 | 94 |
| accuracy | | | 0.90 | 300 |
| macro avg | 0.90 | 0.89 | 0.89 | 300 |
| weighted avg | 0.90 | 0.90 | 0.90 | 300 |




## V. CONCLUSION

The suggested model, SVM + CNN, deep learning-enabled

SVM, is expected to outperform all existing approaches. Prior to entering data into the machine learning algorithm, it is vital to comprehend the dataset and extract features from it because it is quite challenging to manually select a property of the dataset that will enable the algorithm to work more effectively. This feature choice ultimately affects how accurate the machine learning models turn out to be. The final classification layer of a deep learning system, on the other hand, has a significant likelihood of overfitting. In comparison to machine learning models, it also needs a lot of resources, a huge feature set, and computational capacity. The hybrid deep learning-enabled SVM makes use of the benefits of both deep learning and machine learning. In this method, feature selection is handled by CNN, and classification is handled by the SVM machine learning model. The deep learning model does not need a distinct strategy for feature selection, and raw input is provided from which features are retrieved. As the machine learning model cannot choose the

features automatically, a method that suggests appropriate features is necessary for categorization. The benefits of both learning styles are incorporated in the suggested scheme, which results in a clear result that is quicker and more accurate than other existing methods.

## VI. FUTURE WORK

The project's future potential for predicting lung cancer utilising a hybrid SVM and CNN algorithm is highly encouraging.Increasing the dataset and gathering more information can increase the hybrid model's accuracy. This can entail gathering additional pictures and information about lung cancer from other sources, such as numerous hospitals and geographical locations.It can help to increase the model's accuracy by fine-tuning and optimising its hyperparameters. This can entail modifying the CNN architecture, the learning rate, and other parameters, as well as optimising the SVM kernel function.To offer patients with lung cancer a more thorough and precise diagnosis and treatment recommendations, the hybrid SVM and CNN algorithm can be integrated with other technologies like natural language processing (NLP) and machine learning-based decision support systems.Clinical trials may be carried out to confirm the efficacy of the hybrid SVM and CNN algorithm for the detection and management of lung cancer. The hybrid algorithm's adoption in clinical settings can be guided by the trials' findings, potentially enhancing the precision and efficacy of lung cancer diagnosis and therapy.

## REFERENCES

[1]. American Cancer Society. Cancer Facts and Figure 2015. 2015.Availableonline:https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-fact s-figures-2015.html (accessed on 23 December 2021).

[2]. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2015. CA Cancer J. Clin. 2015, 65, 5–29. [Google Scholar] [CrossRef] [PubMed]

[3]. World Health Organization. Cancer Key Statistics. 2021. Available online: https://www.who.int/cancer/resources/keyfacts/en/ (accessed on 2 August 2022).

[4]. E. C. Putro, R. M. Awangga, and R. Andarsyah, Tutorial Object Detection People With Faster region-Based Convolutional Neural Network (Faster R-CNN). Kreatif, 020.Available at: Google Books.

[5]. Baba, A.I.; Câtoi, C. Tumor cell morphology. In Comparative Oncology; The Publishing House of the Romanian Academy: Bucuresti, Romania, 2007. [Google Scholar]

[6]. Wei, M.; Du, Y.; Wu, X.; Su, Q.; Zhu, J.; Zheng, L.; Lv, G.; Zhuang, J. A benign and malignant breast tumor classification method via efficiently combining texture and morphological features on ultrasound images. Comput. Math. Methods Med. 2020, 2020, 5894010. [Google Scholar] [CrossRef]

[7]. Jia, T.; Zhang, H.; Bai, Y. Benign and malignant lung nodule classification based on deep learning feature. J. Med. Imaging Health Inform. 2015, 5, 1936–1940. [Google Scholar] [CrossRef]

[8]. Snoeckx, A.; Reyntiens, P.; Desbuquoit, D.; Spinhoven, M.J.; Van Schil, P.E.; van Meerbeeck, J.P.; Parizel, P.M. Evaluation of the solitary pulmonary nodule: Size matters, but do not ignore the power of morphology. Insights Imaging 2018, 9, 73–86. [Google Scholar] [CrossRef] [PubMed][Green Version]

[9]. Heuvelmans, M.A.; van Ooijen, P.M.; Ather, S.; Silva, C.F.; Han, D.; Heussel, C.P.; Hickes, W.; Kauczor, H.U.; Novotny, P.; Peschl, H.; et al. Lung cancer prediction by Deep Learning to identify benign lung nodules. Lung Cancer 2021, 154, 1–4. [Google Scholar] [CrossRef]

[10]. Dede, Y. Lung Cancer Dataset. 2018. Available online: **https://www.kaggle.com/yusufdede/lung-cancer-dataset** (accessed on 23 December 2021).