iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** IV **Month of publication:** April 2020

DOI: http://doi.org/10.22214/ijraset.2020.4037

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue IV Apr 2020- Available at www.ijraset.com

Prediction of Lung Cancer from Pre-processed CT Scan Image using CNN

Goutam R[1], Sachitha S[2], Rakshith Thorvath[3]
[1, 2, 3]*Computer Science Department, Atria Institue of Technology University*

Abstract: ***Lately, computerised picture processing is generally utilised for the clinical treatment grouping and conclusion. Lung cancer is the most driving reason for death throughout the world these days. In light of the signs and side effects it can't be finding and treatment arranged at the beginning time. Anyway it very well may be recognised through the side effects like coughing out blood and chest torment, the stages and hazard elements of the malignant growth can't be distinguished through the symptoms. The CT scanned lung pictures ought to be engaged with picture order processing for prior prediction of stages and treatment finding. In existing, machine learning treatment characterisation should be possible through the SVM classification. If there should arise an occurrence of enormous arrangement of images for processing tests, this won't be in precise way and it has less accuracy due to inappropriate component extraction strategies. In this manner the performance of the order depends on the segmented highlights acquired in before areas. The extricated fine-grained training data through deep learning are used for the grouping utilizing Convolution Neural Network (CNN). Right now, propose a novel structure to order both little cell and enormous cell lung cancer and foresee its type and treatment utilizing CNN. It is additionally focuses on the preprocessing and segment procedures to achieve the exactness in expectation. The test results in Python - TensorFlow with Kaggle picture dataset show that contrasted with best in class of arrangement and forecast strategies, the proposed plan can get a lot higher exactness in type expectation and treatment determination.***
Keywords: SVM classification, Convolution Neural Network, Deep learning.

I. INTRODUCTION

Right now, the early grouping and forecast assumes a fundamental job in the clinical conclusion and treatment expectation dependent on the complexity. Indeed, the complexities and furthermore the phases of the malignant growth cells are likewise anticipated these days through the execution of the machine learning and deep learning techniques. The malignancy is the most treacherous disease in the clinical field, since it ought to be identified at before stage to lessen the unpredictability of the finding and treatment. This work is focused on the most essential lung cancer diagnosis and treatment grouping through the deep learning procedures. Here the machine learning techniques like Naïve Bayes and Support Vector Machine characterizations are generally utilized for the usage of the stage grouping and treatment expectation. The exactness of the outcomes acquired through the machine learning techniques should be improved for additional innovation upgrades. The Convolutional Neural Networks can be utilized for the execution of the deep learning grouping and stage identification.In this proposed work, the lung cancer distinguishing proof, order of the illness and the stage forecast can be performed through the usage of the Convolutional Neural Networks with the wide scope of test samples. In this way the effectiveness of the proposed model can be recognized through the confusion matrix and exactness dependent on implementation.

II. BACKGROUND KNOWLEDGE

The a portion of the area information ought to be required to get the subtleties and portrayal of the issue and the critical thinking procedures. This can be clarified in a word as following.

A. Lung Cancer Detection and Prediction

Lung cancer is a sort of malignant growth that starts in the lungs. Your lungs are two light organs in your chest that take in oxygen when you breathe in and discharge carbon dioxide when you breathe out. Lung cancer is the main source of disease causing deaths in the United States, among both men and women. Lung cancer asserts a bigger number of lives every year than do colon, prostate, ovarian and bosom diseases consolidated. Individuals who smoke have the most serious danger of lung cancer, however lung disease can likewise happen in individuals who have never smoked. The danger of lung disease increments with the timeframe and number of cigarettes you've smoked. On the off chance that you quit smoking, considerably in the wake of smoking for a long time, you can essentially diminish your odds of creating lung cancer.Early identification of this malignancy is generally significant. Indeed, even the manual conclusion of the infection and the stage ID are excessively monotonous and dangerous. The picture segmentation like mean shift, machine learning like Naïve Bayes and SVM classification, deep learning like ANN and CNN are utilized to play out the early location of lung malignant growth utilizing advanced CT pictures.

B. Convolutional Neural Networks

A Convolutional Neural Network (CNN/ConvNet) is a Deep Learning calculation which can take in an information picture, appoint significance (learnable weights and biases) to different viewpoints/questions in the picture and have the option to separate one from the other. The pre-processing required in a ConvNet is a lot of lower when contrasted with other order calculations. While in crude techniques channels are hand-designed, with enough training, ConvNets can become familiar with these channels/qualities.

The engineering of a ConvNet is practically equivalent to that of the availability example of Neurons in the Human Brain and was enlivened by the association of the Visual Cortex. Singular neurons react to upgrades just in a confined locale of the visual field known as the Receptive Field. An assortment of such fields cover the whole visual region. This can be accomplished through the three distinct layers. They are Convolution Layer, Pooling/Sub-Sampling Layer, Fully Connected Layer and Soft-Max Layer. This can be accomplished through the two unique advances. At first the component learning can be acted in the convolution and pooling layers and the most significant characterization can be acted in the completely associated and delicate max layer.

C. Related works

In existing, F Thaher and R Sammouda have proposed the lung malignancy identification through the division based system, for example, Artificial Neural Networks and Fluffy clustering strategies. It has been performed through the Computer Aided Diagnosis (CAD). Indeed, even the machine learning techniques like SVM classifier, the lung cancer can be performed dependent on the classification on SVM and clustering methods. As like as Werghi N et al have proposed the lung cancer recognition and division utilizing the Bayesian Classification and mean move shift segmentation algorithms utilized for plainly distinguish the sputum cells.

Indeed, even a novel algorithms have likewise proposed dependent on the machine learning techniques like, the creators have planned and built up their own novel algorithms for the usage of lung cancer detection. In [1], the authors Tefti et al and Iyer .An et al have planned and built up the deep learning techniques to improve the accuracy of the lung cancer detection rather than the existing.

In [2], Cengil E and Cinar A have proposed the lung cancer distinguishing proof through deep learning for grouping and furthermore it very well may be actualized utilizing the tensorFlow and 3D CNN architecture of deep learning classification and expectation of stages. Vas M and Dessai A played out the picture handling for lung cancer detection utilizing mathematical morphological activities for segmentation of lung area of intrigue. Potghan et al [3] have been proposed an approach to distinguish lung cancer with tumorous or non tumorous. The enhancements should be possible through the creative executions to the segmentation by threshold and k means clustering. Indeed, even the segmentation can be acted right now in two unique ways that is for lung volumes and lung nodules. At that point the distinguished highlights can be extricated through the Gray Level Co event grid. Along these lines the grouping can be performed to recognize the lung nodules through the K-Nearest Neighbor and MultiLayer Perception algorithm with 98.30 and 98.31% of precision. Alam. J [4] et al have proposed a productive lung cancer detection and expectation methodology with the changed multi-class SVM classification method. As like this proposed instrument, in each phase of characterization, picture improvement and division have been done independently.

III.PROPOSED SYSTEM MODEL

The proposed model design has been delineated in the figure 4.1 with this layer subtleties of Convolutional Neural; Networks. Indeed, even it just speaks to the piece of execution of CNN itself. The input image ought to be preprocessed to acquire exact outcomes. The input image can be scaled to the 500 x 500 pixels in size and the picture ought to be applied to Gaussian or Gradient channel to make de-noised picture for the usage.

The convolution neural network can be applied to the input image as layer by layer. The image can be at first applied to the convolution layer and sub sampling layer to acquire the component learning. After the recognizable proof of highlights, the arrangement can be acted in the indiscretion associated and soft max layer of this CNN architecture. The yield names can be gotten from the usage of soft max layer. The input CT lung image can be conveyed to the underlying preprocessing stage to get the preprocessed image. Here the image can be scaled up to the scope of pixels. Therefore the image needs to distinguish and evacuate the clamors utilizing the Gaussian or Gradient Filter. This can be picked under the pre-checking of the nature of the CT images.

The de-noised picture can be applied to the element extraction period of CNN. Here the highlights can be separated through the convolution and pooling layer of this mechanism. The emphasis goes up three times to get the smooth degree of highlights and afterward it is sent to the completely connected and soft max layer to get the yield output labels. This output label shows the cancer identification with rate and furthermore not perceived output labels. The pseudo code for this proposed model can be clarified in detail for the execution and simple comprehension.

Pseudo code for this proposed model:

Step 1: Obtain the input CT lung image from the user

Step 2: Preprocess the image to get the appropriate scale 500 x 500 pixels

Step 3: Apply Gaussian/ Gradient filter to obtain denoised image for the preprocessed image

Step 4: Apply Convolution Neural Network to obtain the resultsStart initial iteration of Convolution and Pooling

a. Perform convolution using conv2D with pixels, padding and input_shape parameters

b. Activate the relu layers

c. Perform sub sampling using MaxPooling2D with poolSize and Strides

Start second iteration of Convolution and Pooling

d. Perform convolution using conv2D with pixels and padding parameters

e. Activate the relu layers

f. Perform sub sampling using MaxPooling2D with poolSize and Strides

Start third iteration of Convolution and Pooling

g. Perform convolution using conv2D with pixels and padding parameters

h. Activate the relu layers

i. Perform sub sampling using MaxPooling2D with poolSize and Strides

Step 5: Perform first set of fully connected layer using Flatten and Dense functions

Step 6: Perform softmax classification using dense and activation of softmax function

Step 7: Display the resultant image with the stage classificationThis proposed model can be graphically designed to describe the implementation flow of the work as in figure 4.2.

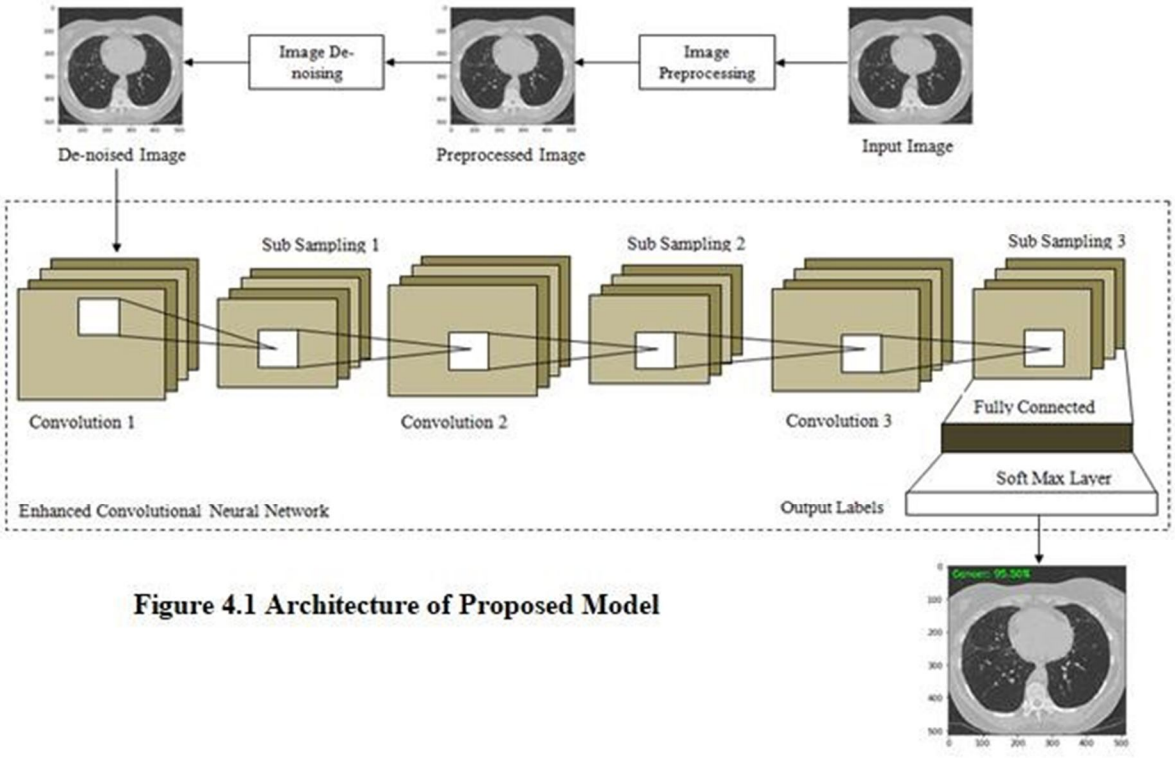


Figure 4.1 Architecture of Proposed Model

Fig. 4.1 Architecture of Proposed Model

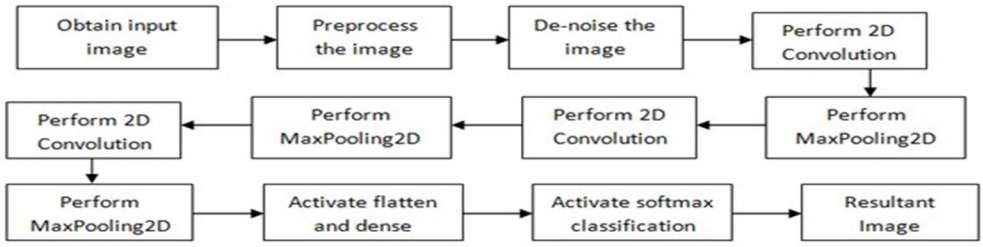


Figure 4.2 Flow of proposed model

Fig. 4.2 Flow of proposed model

This model can be experimentally verified in the Python environment to get the detection results.

IV. IMPLEMENTATION RESULTS

Here the deep learning can be acted in the Python TensorFlow condition. It very well may be acquired utilizing the keras.models classes from the TensorFlow condition.

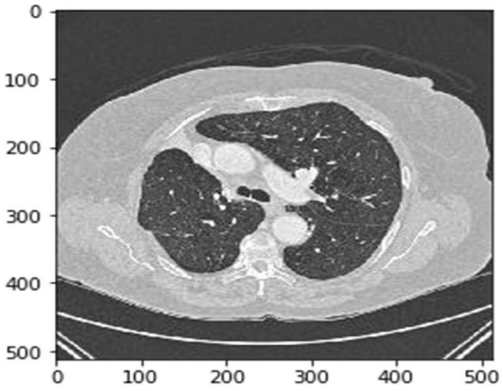


Fig. 5.1 Input image_1

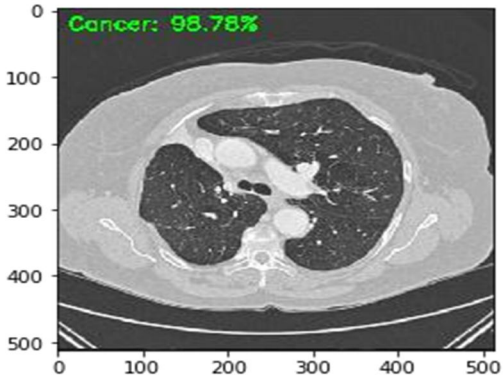


Fig. 5.2 Obtained output_1

The Kaggle dataset can be gathered and used for the execution of the proposed detection model. Python Linguistic structure: Using TensorFlow backend;The over two outcomes that show the positive outcomes of the malignant growth recognizable proof and it likewise shows the how far the lung malignancy cells can be spread over the picture in percentage. In the figure 5.5 and figure 5.6, shows the input and the resultant pictures with the negative indication of the lung malignancy what's more, it shows the outcome as not perceived.

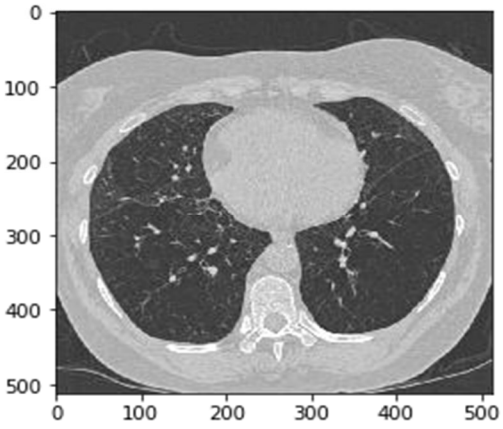


Fig. 5.3 Input image_2

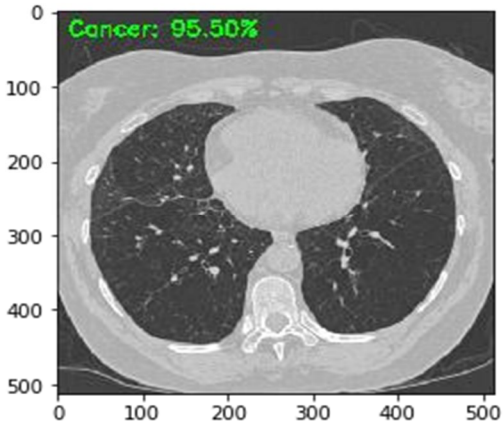


Fig. 5.4 Obtained output_2

The precision can be acquired with the correlation of preparing tests.

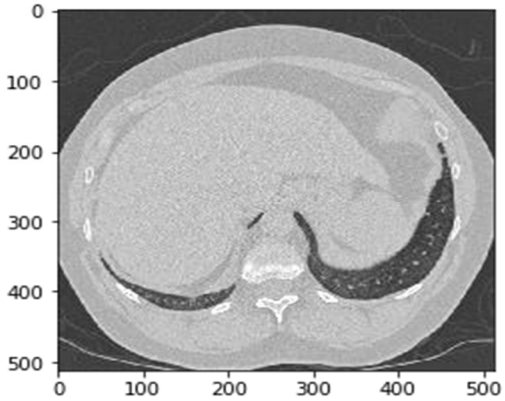


Fig. 5.5 Input image_3

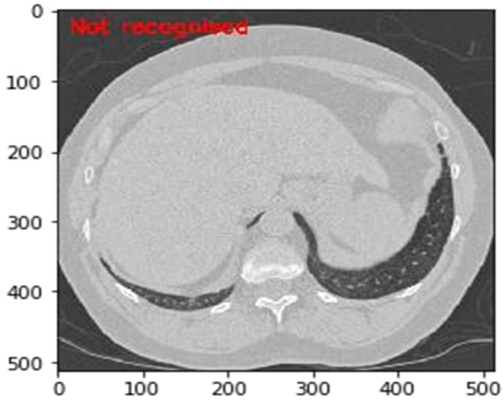


Fig. 5.6 Obtained output_3

The confusion matrix can likewise be gotten through the outcomes as from anticipated labels and genuine names of cancer cells and it has been outlined in the figure 6. In light of the testing samples, the quantity of test samples were tested to get this confusion matrix with true positives (TP), true negatives (TN), false positives (FP) and false negative (FN). Let us figure the precision, recall, true positive rate and false positive rate at the limit of 0.5. At that point we make the confusion matrix has shown in figure 7.

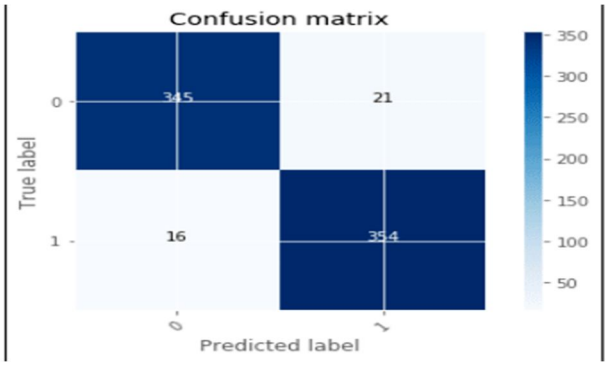


Fig. 5.7 Confusion Matrix

The performance metrics such as precision, recall and f1-score can be obtained through the following notations based on the results obtained through the confusion matrix. Precision = TP / (TP+FN) = 345 / (354+21) = 0.92

Recall = TP / (TP+FP) = 345 / (354+16) = 0.93

$$\text{F-Score} = 2x\,\frac{\text{precision}-\text{recall}}{\text{precision}+\text{recall}} = 0.01$$

The accuracy for this proposed mechanism can be obtained through the below formula using the values through the confusion matrix.

Accuracy = (TP+TN) / (TN+TP+FP+FN)

= (345+354) / (345+354+21+16)

= 699 / 736

= 94.97

Accuracy of this proposed mechanism is 94.97% obtained through the identified matrix has illustrated in the figure 5.8. Thus the accuracy can be varied from the usual CNN model used in the existing works.

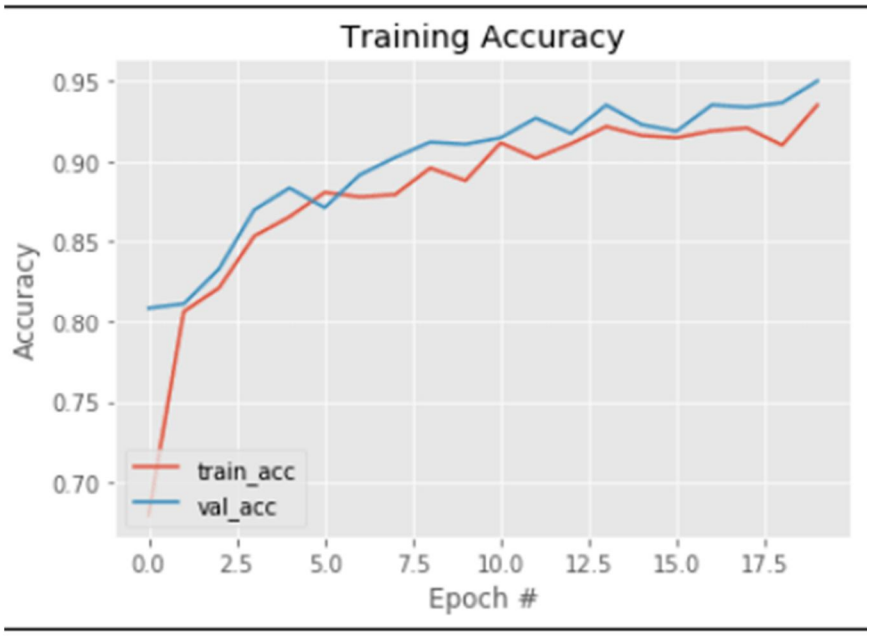


Fig. 5.8 Performance graph based on accuracy

From these outcomes, the acquired outcomes can be improved through the alteration in the element extraction acquired from the convolution and pooling layer of the CNN model.

V. CONCLUSIONS

In this way the stage identification of the Lung cancer can be performed through the viable execution of the proposed component utilizing Convolution Neural Networks.

Right now, four layers of improvements may assists with playing out the deep learning based order. Along these lines the layers are convolution layer, sub sampling layer/pooling layer, completely connected layer and soft max layer. Beginning two layers are iterated three times to get the exact outcomes from the model. The anticipated outcomes can acquired in normal of 95% exactness and it should nearly great as opposed to the current works. In future, it is to improve through the upgrades in the security of the preparation testing samples to acquire the more exact answers for the classification and prediction.

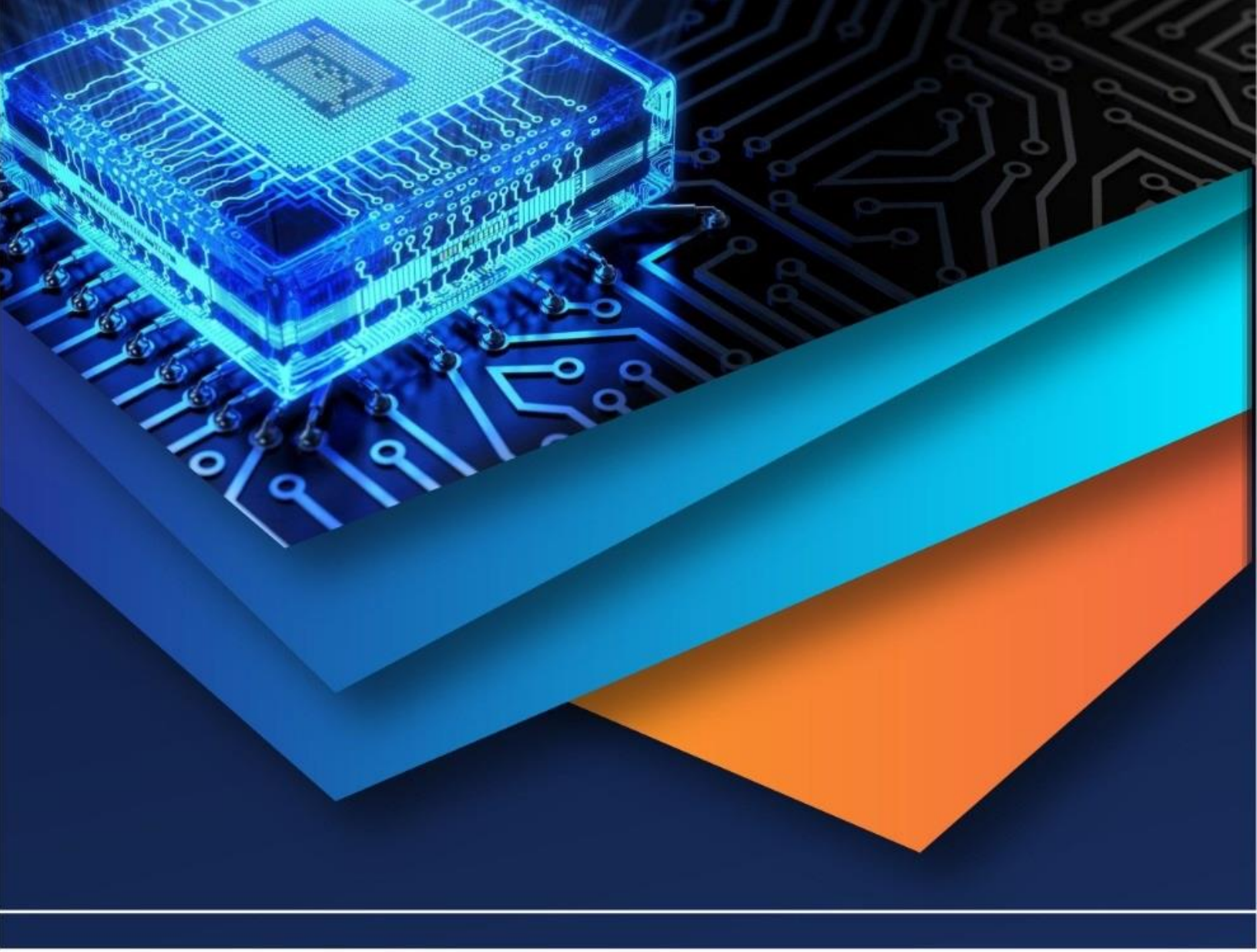
REFERENCES

[1] Tafti A. P., Bashiri F. S., La Rose E., & Peissig P., "Diagnostic Classification of Lung CT Images Using Deep 3D Multi- Scale Convolutional Neural Network", 2018 IEEE International Conference on Healthcare Informatics (ICHI), 2018, USA, pp.412-418.

[2] CENGIL E., & CINAR A., "A Deep Learning Based Approach to Lung Cancer Identification", 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), 2018, Turkey.

[3] Potghan S., Rajameenakshi R., & Bhise A., "Multi-Layer Perception Based Lung Tumor Classification", 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018, India, pp.492-502.

[4] Alam J., Alam S., & Hossan A., "Multi-Stage Lung Cancer Detection and Prediction Using Multi-class SVM Classification", 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), 2018, Bangladesh.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY