e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 6, Issue 4, April 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.54

6381 907 438 6381 907 438 ijmrset@gmail.com @ www.ijmrset.com

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604007 |

Detection of Lung Cancer Using Convolutional Neural Network

Ojaswi Banchhor[1], Asst. Prof. Harish Sahu[2]

M. Tech Student, Dept. of Computer Science and Engineering, Columbia Institute of Engineering and Technology, Raipur, Chhattisgarh, India[1]

Assistant Professor, Dept. of Computer Science and Engineering, Columbia Institute of Engineering and Technology, Raipur, Chhattisgarh, India[2]

ABSTRACT: Lung cancer is one of the deadliest forms of the disease; it affects thousands of individuals, and if it is not found in the early stages of the illness, the patient's chances of survival are very low. For the aforementioned reasons as well as to aid in conquering this dreadful condition, prompt identification is crucial and will benefit from artificial intelligence techniques. With the use of a convolutional neural network technique architecture, a computer-aided method was developed through this research to help with the identification of patient cases as normal, benign, or malignant. 550 CT scan images are used as dataset. High accuracy is provided by the suggested model, reaching 99%.

KEYWORDS*:* Lung Cancer, Chest CT scan image, Convolutional Neural Network

I. INTRODUCTION

Cancer is characterized by using the out-of-control division of ordinary cells that may invade and destroy normal human cell. These abnormal tissues are known as a tumor. These cells are having two categories benign and Malignant [1]. Benign is generally now not threatening existence while malignant tumors can threaten someone's lifestyles. Benign type of tumor no longer unfolds in different cells and it is clean and regular in form, opposite to it malignant tumor unfolds to different frame cells to form new cancerous nodule, having uneven form. Cancer is the maximum extreme and massive ailment that is liable for a huge quantity of deaths each 12 months [2].

. According to professionals, lung most cancers is the most general sort of cancer [3].57,795 new instances of lung most cancers had been detected in 2012, with the number expected to be 67,000 new instances in keeping with year by 2020 [4]. Lung most cancers finally ends up killing the lives of more than 7.6 million humans global [5]. As most cancers inside the lungs can be a cause for dying in both ladies and men, so for this reason a timely and unique diagnosis of nodules is crucial for remedy [6].

But one of the maximum hard obligations for radiologists is a most cancers prognosis. A pc-aided diagnostic (CAD) gadget plays a substantial part in helping in challenges, which enables radiologists in successfully identifying, predicting, and diagnosing lung most cancers [7]. Using the machine getting to know algorithms, it will become easy to differentiate among a cancerous and noncancerous part of the image of the lung. In the concern of lung cancer detection, numerous machine learning techniques can be powerful.

As CNN emerge as a famous system getting to know set of rules, it is widely used not simplest in laptop vision but extensively utilized in herbal language processing (NLP), photograph processing, and analysis of clinical images [8]. A

CNN's key strength is its deep architecture, which permits it to extract a hard and fast of different functions at various abstraction degrees [9]. If we take into consideration the hassle of image type in Machine Learning, it is very important to have an intensive knowledge of the area to extract functions manually, that is a prolonged procedure. Feature Extraction is performed mechanically is function of deep learning [10].

In the detection of lung nodules, distinctive varieties of pictures can be used inclusive of chest X-Ray, Computed Tomography (CT) of the lung, and the lung's Magnetic Resonance Imaging (MRI) [11]. Among these the maximum frequent imaging approach for lung cancer diagnosis is volumetric thoracic computed tomography (CT) [2]. As these pix are having noise or can't be used as it is for detection a few pre-processing is carried out on the one's images. Such prepossessed images are then passed thru the exclusive layers of CNN structure in which feature extraction and classification manner are performed, to distinguish among malignant pictures and those which might be benign.

II. LITERATURE REVIEW

Table 1: Comparisons of various techniques and method used in existing system

Ref.	Year	Datasets	Preprocessing	Methods	Results
[26]	2020	Lung Cancer CT Scan images	GLCM features on image	SVM classification algorithm	Technique determines the detection of benign or malignant lung cancer is 83.33%
[27]	2020	Database of Chest X-Ray andLungcancer(LIDC-IDRI)	Morphologicalsegmentation andwatershedsegmentationareused forautomatednodulesegmentation	MAN is used to classifychest X-Rays images andEFTisusedtoclassifythelungCTimages.	DLexactnessis96%forX-Rayimages whilethe exactness is97.27%forCT Images
[28]	2020	Databaseofcancer imagingarchive(CIA)dataset	Multilevel brightness-preservingapproach	improveddeepneuralnetworkandensembleclassifier.	The proposed systemdetect0the tumorwithmaximum accuracy.
[29]	2019	Image was collectedfromCancer imagingArchive(CIA)dataset	Thenoiseisremovedusingweightedmeanhistogramequalizationapproach. Inaddition,improved profuseclusteringmethod(IPCT)isappliedfor segmenting the affectedregion.	Deep learninginstantaneouslytrainedneuralnetwork(DITNN)isutilized.	exactness98.42% & minimumclassificationerrorof0.038.
[30]	2019	DatabaseobtainedfromUCIrepository	Picturesecuring,pre-handling,binarization,thresholding,division, featureextractionareapplied.	The fuzzy neural method isusedtotesttheneuralsystemwithmachinelearningtechnique.	Exactness96.67%.
[31]	2019	Database of Lung Image Dataset Consortium image collection (LIDC-IDRI)	This step involves of segmentation is followed by normalization and zero centering.	A number of classifiers like XGBoost and Random Forest are used.	Exactness84%

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604007 |

[32]	2018	Dataset of Lung Image Dataset Consortium (LIDC)	median filter and Gaussian filter are applied on the CT images.	Watershed segmentation for recognition and SVM for classification of cells as Malignant or benign.	Exactness92%.
[33]	2018	dataset obtained from UCI repository	pre-processing for data cleaning is applied	A number of classifiers including: MLP, Neural Network, Decision Tree, Naïve Bayes, Gradient Boosted Tree, and SVM are assessed	Exactness90%.
[34]	2018	Database of Lung Image Database Consortium (LIDC)	Converting to grayscale image, applying denoising methods such as median blur, Gaussian blur, and bilateral blur, then applying thresholding methods for converting the grayscale image into a binary image.	k-nearest neighbors classifier, support vector machine classifier, decision tree classifier, multinomial naïve Bayes classifier, stochastic gradient descent classifier, random forest classifier, and multi-layer perceptron (MLP) classifier are applied	Exactness88.55%.
[35]	2018	The lung cancer database utilized for training is taken from UCI machine learning dataset	image enhancement and segmentation has been done. Image scaling, color space transformation and contrast enhancement	multi-class SVM classifier	Precision of 97% for cancer recognition and 87% for cancer prediction
[36]	2016	Used Scan CT Image are from VIA and ELCAP dataset	image enhancement using Gabor filter is applied	Region Growing, Marker Controlled Watershed, and Marker Controlled Watershed with Masking are applied.	watershed with masking method has highest exactnessand robustness.

III. PROBLEM INDENTIFICATION

With reference to literature survey, the author has proposed a mechanism for Lung Cancer Detection using SVM and preprocessingsteps. But the accuracy of the present system is very low as well as the implementation model of the existing system is more complicated.

IV. PROPOSED METHODOLOGY

Our system proposed lung cancer detection system which uses CT scan images to detect lung image have Cancerous or non- Cancerous.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604007 |

Modules for the proposed framework are described below in the Fig. 1.

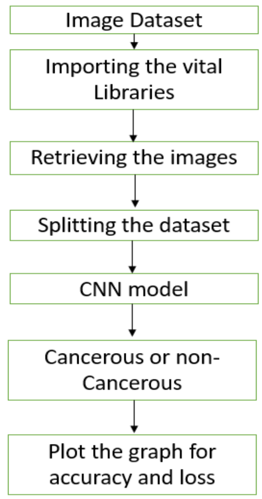


Fig. 1 Proposed system architecture for CT scan image/ dataset for lung cancer recognition

(a) Image Dataset

CT lung cancers prediction using CNN information set given in model folder. The dataset consists of round 550 CT lung cancers pics. In Fig.2 Displays the CT scan image of lung cancer detection.

Fig 2. *Displays the Lung CT scan image as dataset sample*

(b) Importing the necessary libraries

We are using the Python language for the system architecture. Firstly, we import the necessary libraries which include Keras for constructing the main model, sklearn for splitting the training and take a look at information, PIL for changing the images into array of numbers and other libraries consisting of pandas, NumPy, matplotlib and TensorFlow**.**

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604007 |

(c) Retrieving the CT images

We will retrieve the photographs and their labels. Then resize the pix to (224,224) as all images have to have equal size for reputation. Then convert the photographs into NumPy array.

(d) Splitting the dataset

Splits the data into 3 categories

- Training (use for generation of model)

- Private test (use for generation of model)

- Public test (use for evaluating the model)

- Split the dataset into trained and test. Our system finds 80% train data and 20% test facts.

(e) Conventional Neural Network

The targets behind the first module of the path four are:

- To recognize the convolution operation
- To understand the pooling operation
- Remembering the vocabulary utilized in convolutional neural networks (padding, stride, filter, and many others.)
- Building a convolutional neural network for multi-magnificence category in photos.

Convolution Network

In a convolutional network (ConvNet), there are essentially 3 types of layers:

(a) Convolution layer

(b) Pooling layer

(c) Fully connected layer

- *Conventional layer* - It computes the speck object among open fields and the channel. Consequence is an individual entire number of the yield amount. At that factor, channel at some point of following responsive discipline of a comparable statistics picture by means of a Stride and do a similar hobby all over again. It will rehash a comparable interaction and again until it is going via the complete image.

- *Poolinglayer*- Pooling layers are typically used to reduce the size of the inputs and for this reason speed up the computation.

- *Fully connected layer* - Completely associated layer includes masses and inclinations. Fully connected layer connects neurons in an individual layer to next layers. By training, it is used to categories images among distinctive training.

-

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604007 |

(f) Building the model

For building the we will use sequential version from Keras library. Then we will add the layers to make convolutional neural community. In the primary 2 Conv2D layers we've got used 32 filters and the kernel length is (5,5).

In the MaxPool2D layer we've saved pool size (2,2) because of this it will pick the most fee of each 2 x 2 place of the image. By doing this dimension of the photo will reduce through factor of 2. In dropout layer we've got kept dropout charge = 0.25 which means 25% of neurons are removed randomly.

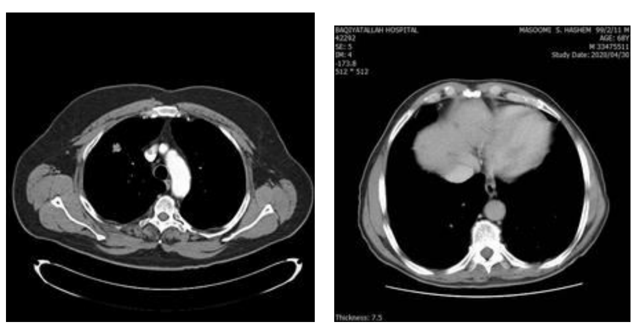
We apply these three layers once more with a few changes in parameters. Then we follow flatten layer to transform 2-D records to one-D vector. This layer is accompanied through dense layer, dropout layer and dense layer again. The ultimate dense layer outputs 2 nodes because the brain tumor or no longer. This layer uses the softmax activation feature which gives opportunity price and predicts which of the 2 options has the best chance.

(g) Apply the model and plot the graphs for accuracy and loss

We will compile the model and observe it using in shape function. The batch size may be 1. Then we will plot the graphs for accuracy and loss. The model creates during the training process comprise of pretrained weights and values, which can be utilized for implementation of a new lung cancer detection problem.

V. RESULT AND DISCUSSION

We have analyzed image of CT scan Lung cancer dataset. We have considered 550 different CT-Scan of lungs as shown in Fig. 3 (a) and (b).



(a) (b)

Fig. 3: Lung CT Scan images (a)Cancerous image and (b) non-Cancerous image

International Journal Of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604007 |

STEP 01: Run program and then open in localhost server

Fig. 4 shows the local server snapshot. A home page is appeared.

Fig. 4 Displays the snapshot of local server

Step 2: Login with username and password

After clicking on login, username and password is required as shown in Fig. 5.

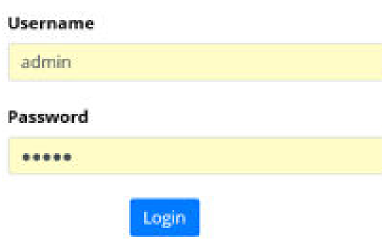


Fig. 5: Displays the login form.

Step 3: Upload an Image

Select the image from the trained database and then click on the submit button as shown in Fig. 6.

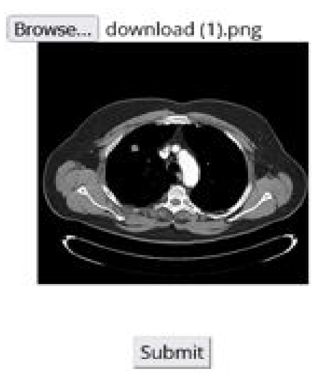


Fig. 6. Displays the browsed trained image from the database

International Journal Of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604007 |

Step 4: Prediction outcome of testing dataset using image processing.

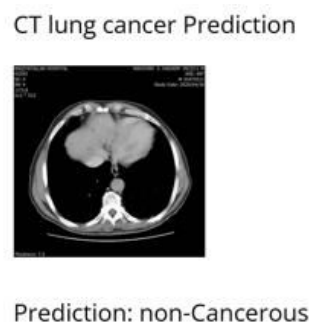


Fig. 5. Shows the of prediction of lung Cancer detection.

This prediction shows the result as the CT image is cancerous or non-cancerous.

Step 5: Performance analysis of the system

PERFORMANCE ANALYSIS

Accuracy:	0.997
Precision:	1.000
Recall:	0.997
F-Measure:	0.997

Fig: 7 displays the performance analysis of the system

Performance analysis gives the output in the form of Accuracy, precision, recall and F-measure. Our system provides 99% accuracy as shown in fig. 7.

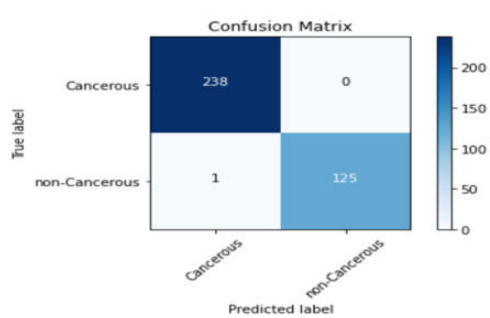


Fig. 8 Displays the confusion matrix

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604007 |

The confusion matrix shows the true label and predicted label of cancerous and non-cancerous images as shown in fig. 8. From the analysis, true label gives the correctly classified lung cancer images and predicted label gives the prediction of lung cancer image.

Step 6: Graphical representation

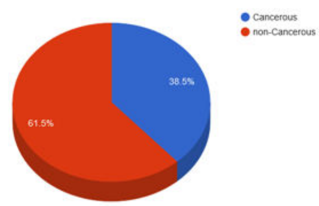


Fig 9: Displays the graph for cancerous and non- cancerous image

Graph shows the graphical representation of cancerous and non-cancerous images present in the database as shown in fig. 9. There are 38.5% images are cancerous and 61.5% images are non-cancerous in our 550 CT scan images dataset.

TABLE II. Accuracy obtained from different frameworks

S.No.	Author	Year	Technique Used	Accuracy
1	Q. Firdaus et al. [26]	2020	SVM Classification Algorithm	83 %
2	Bhandary et al. [27]	2020	DL strategy (Modified AlexNet(MAN)) with SVM	97 %
3	Proposed System	2023	Conventional Neural Network	99%

VI. CONCLUSION

A conventional neural network-based architecture was deployed to recognized the malignancy cell available in the CT scan lung cancer images. CT lung cancer images consists of various shapes and size of the cancerous cells presented at the input training set of the system. Our proposed system is eligible to recognize the presence and absence of cancerous tissue with accuracy about 99%.

In the near future, the system will be trained with huge number of datasets to recognize the type of lung cancer with its size and shape.

REFERENCES

[1] Giang Son Tran, Thi Phuong Nghiem ,VanThi Nguyen ,Chi Mai Luong ,Jean-Christophe Burie, “Improving accuracy of lung nodule classification using deep learning with focal loss,” Journal of Healthcare Engineering 2019 (2019), Volume 2019, Article ID 5156416, 9 pages.

[2] SunyiZheng,Jiapan Guo, Xiaonan Cui, Raymond N. J. Veldhuis, Matthijs Oudkerk, and Peter M.A. van Ooijen, “Automatic pulmonary nodule detection in CT scans using convolutional neural networks based on maximum intensity

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604007 |

projection," IEEE Trans Med Imaging. 2020Mar ; 39(3):797-805.doi10.1109/TMI.2019.2935553. Epub 2019 Aug 15. PMID: 31425026.

[3] FangzhouLiao ,Ming Liang, Zhe Li, Xiaolin Hu, Sen Song," Evaluate the malignancy of pulmonary nodules using the 3-D deep leaky noisy-OR network," in IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 11, pp. 3484-3495, Nov. 2019, doi: 10.1109/TNNLS.2019.2892409.

[4] Dou Q, Chen H, Yu L, Qin J, Pheng-Ann Heng, "Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection," IEEE Trans Biomed Eng. 2017 Jul;64(7):1558-1567. doi: 10.1109/TBME.2016.2613502. Epub 2016 Sep 26. PMID: 28113302.

[5] Gunaydin, Ozge, MelikeGunay, and OznurS¸engel. " Comparison of lung cancer detection algorithms," 2019 Scientific Meeting on Electrical Electronics and Biomedical Engineering and Computer Science (EBBT), IEEE, 2019.

[6] [9] Alam, Janee, Sabrina Alam, and Alamgir Hossan, "Multi-stage lung cancer detection and prediction using multi-class svmclassifie," 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)., IEEE, 2018.

[7] KyameliaRoy,Sheli Sinha Chaudhury,MadhurimaBurman,AhanaGanguly,Chandrima Dutta, et al. "A comparative study of Lung Cancer detection using supervised neural network," 2019 International Conference on Opto-Electronics and Applied Optics (Optronix), IEEE, 2019.

[8] Xinzhuo Zhao, Liyao Liu, Shouliang Qi, YueyangTeng,JianhuaLi,Wei Qian, "Agile convolutional neural network for pulmonary nodule classification using CT images," International journal of computer assisted radiology and surgery, 13.4 (2018): 585-595.

[9] Matko Saric, Mladen Russo, Maja Stella, Marjan Sikora,"CNN-based method for lung cancer detection in whole slide histopathology images,"2019 4th International Conference on Smart and Sustainable Technologies (SpliTech), IEEE, 2019.

[10] Nima Tajbakhsh,Jae Y. Shin, Suryakanth R. Gurudu, R. Todd Hurst, Christopher B. Kendall,Michael B. Gotway, Jianming Liang, "Convolutional neural networks for medical image analysis: fine tuning or full training?," IEEE transactions on medical imaging 35, no. 5 (2016): 1299-1312.

[11] Takiar R, Nadayil D, Nandakumar A,"Projections of Number of Cancer Cases in India (2010-2020) by Cancer Groups,"Asian Pac J Cancer Prev 11.4 (2010): 1045-9.

[12] Q. Firdaus, R. Sigit, T. Harsono and A. Anwar, "Lung Cancer Detection Based On CT-Scan Images With Detection Features Using Gray Level Co-Occurrence Matrix (GLCM) and Support Vector Machine (SVM) Methods," 2020 International Electronics Symposium (IES), IEEE, 2020, pp. 643-648, doi: 10.1109/IES50839.2020.9231663.

[13] Bhandary,A.,Prabhu,G.A.,Rajinikanth,V.,Thanaraj,K.P.,Satapathy,S.C.,Robbins,D.E.,...&Raja,N. S. M. (2020). Deep-learning framework to detect lung abnormality–A study with chest X-Ray andlungCT scanimages. PatternRecognitionLetters,129, 271-278.

[14] Shakeel,P.M.,Burhanuddin,M.A.,&Desa,M.I.(2020).AutomaticlungcancerdetectionfromCTimage using improved deep neural network and ensemble classifier. Neural Computing andApplications,1-14.

[15] Shakeel, P. M., Burhanuddin, M. A., & Desa, M. I. (2019). Lung cancer detection from CT image usingimprovedprofuseclusteringanddeeplearninginstantaneouslytrainedneuralnetworks.Measurement,145, 702-712.

[16] Reddy,U.,Reddy,B.,&Reddy,B.(2019).RecognitionofLungCancerUsingMachineLearningMechanisms with FuzzyNeuralNetworks. Traitementdu Signal,36(1), 87-91.

[17] Bhatia,S.,Sinha,Y.,&Goel,L.(2019).Lungcancerdetection:Adeeplearningapproach.InSoftComputingforProblemSolving(pp.699-705).Springer, Singapore.

[18] Makaju, S., Prasad, P. W. C., Alsadoon, A., Singh, A. K., & Elchouemi, A. (2018). Lung cancer detectionusingCT scanimages. ProcediaComputer Science,125, 107-114.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54|

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604007 |

[19] Faisal, M. I., Bashir, S., Khan, Z. S., & Khan, F. H. (2018, December). An evaluation of machine learningclassifiers and ensembles for early stage prediction of lung cancer. In 2018 3rd InternationalConference on Emerging Trends in Engineering, Sciences and Technology (ICEEST) (pp. 1-4).IEEE.

[20] Singh,G.A.P.,&Gupta,P.K.(2019).Performanceanalysisofvariousmachinelearning-basedapproaches for detection and classification of lung cancer in humans. Neural Computing andApplications,31(10), 6863-6877,

[21] Alam,J.,Alam,S.,&Hossan,A.(2018,February).Multi-stagelungcancerdetectionandpredictionusing multi-class svm classifie. In 2018 International Conference on Computer, Communication,Chemical,MaterialandElectronic Engineering(IC4ME2)(pp. 1-4).IEEE.

[22] Abdillah, B., Bustamam, A., & Sarwinda, D. (2017, October). Image processing-based detection of lungcancer on CT scan images. In Journal of Physics: Conference Series (Vol. 893, No. 1, p. 012063).IOPPublishing.

INNO SPACE

SJIF Scientific Journal Impact Factor

Impact Factor 7.54

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com