iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52656

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: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Melanoma Skin Cancer Detection using Image Processing

Atharva Pawar[1], Pratiksha Rupde[2], Snehal Khodake[3], Achal Raut[4], Prof. Neelam Joglekar[5]
[1, 2, 3, 4]*BE students, Department of Information Technology, Marathwada Mitra mandal's College of Engineering, Pune*
[5]*Assistant Professor, Department of Information Technology, Marathwada Mitra mandal's College of Engineering, Pune*

Abstract: ***Early detection of melanoma skin cancer is crucial for effective treatment. Melanoma is considered the most dangerous form of skin cancer due to its higher likelihood of spreading to other parts of the body if not diagnosed and treated promptly. To address this challenge, non-invasive medical computer vision and medical image processing techniques have become increasingly important in clinical diagnosis of various diseases, including melanoma. These techniques offer automated image analysis tools that enable accurate and rapid evaluation of skin lesions.To assess the severity of the lesions, a total dermo copy score was calculated based on the extracted features. Finally, classification was carried out using a convolutional neural network (CNN) model. The results of the study demonstrated a high classification accuracy of 96.5%, indicating the effectiveness of the proposed approach in accurately identifying melanoma skin cancer.***
Keywords: Dermatology, Image Processing, Machine learning, Melanoma.

I. INTRODUCTION

Skin is the outer most region of our body and it is likely to be exposed to the environment which may get in contact with dust pollution micro-organisms and also to uv radiations these may be the reasons for any kind of skin diseases and also skin related diseases are caused by instability in the genes this makes the skin diseases more complex the human skin is composed of two major layers called epidermis and dermis the top or the outer layer of the skin which is called the epidermis composed of three types of cells flat and scaly cells on the surface called squamous cells round cells called basal cells and melanocytes cells that provide skin its color and protect against skin damage as the diagnostic classification currently do not represent the diversity of the disease these are not sufficient enough to make a correct prediction and also treatment to be provided for that disease adding to this cancer cells are often diagnosed late and treated late it is diagnosed when the cancer cells have mutated and spreads to the other internal parts of the body at this stage therapies or treatments are not very effective the other reasons for which the disease might have taken over to a very serious state can be because of peoples ignorance and also that people try using home remedies without knowing the severity of the problem and also sometimes these may lead to another kind of skin rashes or even increasing the severity of the problem.

Among all the types of skin diseases skin cancer is found to be the deadliest kind of disease found in humans this is found most commonly among fair skin is found to be two types malignant melanoma and non-melanoma malignant melanoma is one of the most deadly and dangerous types of cancers even thoughts found that only 4 of the population is affected with this it holds for 75 of the death caused due to skin cancer melanoma can be cured if its identified or diagnosed in early stages and the treatment can be provided early but if melanoma is identified in the last stages it is possible that melanoma can spread across deeper into the skin and also can affect other parts of the body then it becomes very difficult to treat melanoma is caused due to presence of melanocytes which are present within in the body.

Exposure of skin to UV radiation is also one of the major reasons for the cause of melanoma dermo copy is a technique that is used to exam the structure of skin an observation-based detection technique can be used to detect melanoma using dermo copy images the accuracy of the dermo copy depends on the training of the dermatologist the accuracy of melanoma detection can be 75-85 even though the experts in skin use dermo copy as a method for diagnosis the diagnosis that is performed by the system will help to increase the speed and accuracy of the diagnosis computer will be able to extract some information like asymmetry color variation texture features these minute parameters may not be recognized by the human naked eyes there are 3 stages in an automated dermo copy image analysis system a pre-processing b proper segmentation c feature extraction and selection the segmentation is the most important and also plays a key role as it affects the process of fore coming steps supervised segmentation seems to be easy to implement by considering the parameters like shapes sizes and colors along with skin types and textures this system- based analysis will reduce the diagnosing time and increases the accuracy dermatological diseases due to their high complexity variety and scarce expertise is one of the most difficult terrains for quick easy and accurate diagnosis especially in developing and under-developed

countries with low healthcare budget also its a common knowledge that the early detection in cases on many diseases reduces the chances of serious outcomes the recent environmental factors have just acted as catalyst for these skin diseases the general stages of these diseases are as stage 1- diseases in situ survival 999 stage 2- diseases in high-risk level survival 45-79 stage 3-regional metastasis

II. RELATED WORKS

Author has tried to address the same problem using image analysis techniques the work uses the technique of noise removal and subsequent feature extraction after the noise removal[1].The image is fed into the classifier the further feature extraction process and finally the prediction of the disease most of the earlier publications focused on feature extraction and then subsequent disease prediction was done papers [6],[3].Have used artificial neural network for dealing with this complex problem while papers .[2],[4],[5].

Have used machine learning algorithms for the task computer vision techniques have played a major role in many kinds of literature as is evident the publishers have utilized the image processing techniques to accomplish the preprocessing task in the similar way we also try to implement the computer vision techniques but out implementation mainly focus for dataset augmentation

III. METHODOLOGY

Our model is designed in 3 phases as follows:

1) Phase1 also involves the pre-processing of the images where hair removal glare removal and shading removal are done removal of these parameters helps us to identify the texture color size and shape-like parameters in an efficient way

2) Phase2- this phase consists of segmentation and feature extraction segmentation is explored via three methods an Otsu segmentation method b modified Otsu mentation method watershed segmentation method defeatures extracted for color shape size and texture

3) Phase 3- this is the most important phase of our model this phase involves designing the model and training our model was trained for back propagation algorithm neural networks SVM support vector machine and CNN convolutional neural networks on the dataset that was collected in the phase1 the model after training was tested for the accurate output

IV. COMPONENTS OF METHODOLOGY:

4) *Pre-Processing*

1) The pre-processing of images is an important task or activity which helps in saving time for training as well as provides a clear enhancement for the further steps by increasing the efficiency of the model pre- processing includes the following Collection of the dataset

2) Hair removal

3) Shading removal

4) Glare removal

a) *Dataset: T*he images were collected from the ISIC dataset provides the collection of images for melanoma skin cancer ISIC melanoma project was undertaken to reduce the increasing deaths related to melanoma and the efficiency of melanoma early detection this ISIC dataset contains approximately 23000 images of which we have collected 1000-1500 images and trained and tested over these images

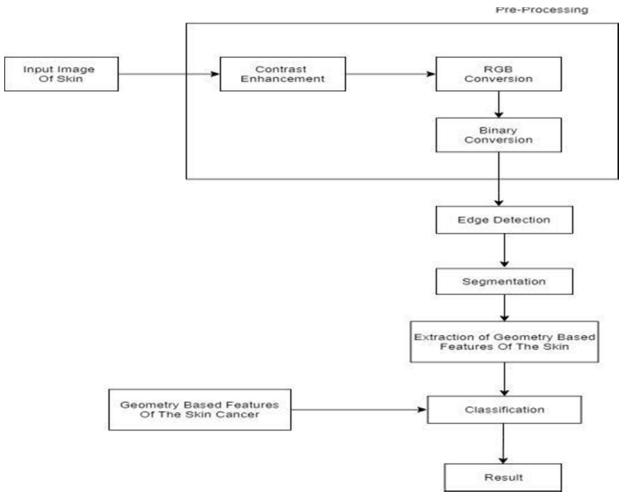
b) *Hair Removal:* for the above-collected images hair removal method applied this method was performed using the thought transform used to identify lines or elliptical or circular shapes performing hair removal for the images that have hair within the tumor provides a clear area of tumor which also helps us to make further more enhancements.

c) *Shading removal:* the images that are taken from the dataset contain shade around the region of the tumor this shade for a few images is dark and for a few is light removal of the shade in the region of the tumor also provides a clear vision of the tumor which is also helpful in the further enhancements we have used the MATLAB filters to remove the shade for images in the dataset.

d) *Glare Removal:* sometimes the images captured for my images will contain glare this glare is not visible to the naked eye we remove this glare using the MATLAB filter this minute noise sometimes may affect the accuracy at the end

V. ARCHITECTURE

In our model we have used 3 different methods i.e. neural networks support vector machine and convolutional neural networks to find the efficient detection and classification of the melanoma skin cancer into malignant and benign skin cancers the data that is pre-processed is followed by segmentation and feature extraction these extracted feature images are then passed into neural networks and support vector machine to classify the images into malignant and benign and to predict the exact accuracy.

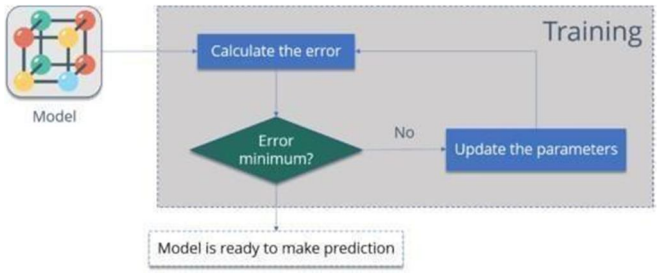


VI. I

A. *Neural Networks*

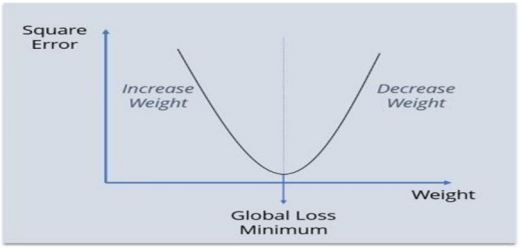
In the neural networks we have used the back propagation algorithm back propagation is a supervised learning algorithm for training the multi-layer perceptions while designing the neural networks we initialize the weights with some random values as we do not know what exactly the weight can be so we first give some random weight if the model provides an error with large values so we need to need to change the values to somehow minimize the error value to generalize this we can just say.

1) *Calculate the Error* – How far is your model output from the actual output
2) *Minimum Error* – Check whether the error is minimized or not.
3) *Update the Parameters* – If the error is huge then, update the parameters (weights and biases). After that again check the error. Repeat the process until the error becomes minimum.
4) *Model is Ready to make a Prediction* – Once the error becomes minimum, you can feed some inputs to your model and it will produce the output.



The backpropagation algorithm looks for the minimum value of the error function in weight space using a technique called the delta rule or gradient descent we are trying to get the value of weight such that the error becomes minimum we need to figure out whether we need to increase or decrease the weight value once we know that we keep on updating the weight value in that direction until error becomes minimum you might reach a point where if you further update the weight the error will increase at that time you need to stop and that is your final weight value

Consider the graph below:



We need to reach the 'Global Loss Minimum'. This is nothing but Back propagation.

B. Support Vector Machine (SVM)

SVM (Support Vector Machine) is a supervised machine learning algorithm that is mainly used to classify data into different classes unlike most algorithms SVM makes use of a hyperplane that acts as a decision boundary between the various classes SVM can be used to generate multiple separating hyperplane such that the data is divided into segments and each segment contains only one kind of data features of SVM is as follows:

SVM is a supervised learning algorithm this means that svm trains on a set of labeled data SVM studies the labeled training data and then classifies any new input data depending on what it learned in the training phase.

1. a main advantage of SVM is that it can be used for both classification and regression problems SVM is mainly known for classification the SVR support vector regression is used for regression problems.
2. SVM can be used for classifying non-linear data by using the kernel trick the kernel trick means transforming data into another dimension that has a clear dividing margin between classes of data after which you can easily draw a hyperplane between the various classes of data.

What are support vectors in SVM? we start by drawing a random hyperplane and then we check the distance between the hyperplane and the closest data points from each class these closest data points to the hyperplane are known as support vectors and that is where the name comes from the support vector machine.

In this project, we have used SVM to classify the malignant and benign skin cancer images this is done by passing the segmented and feature-extracted images into SVM where SVM writes the hyperplane and groups all the nearby similar features into different classes.

The performance of the SVM classifier was very accurate for even a small data set and its performance was compared to other classification algorithms like CNN and backpropagation algorithms.

C. Convolution Neural Network

Convolution neural networks cnns are neural networks with a specific architecture that are very powerful in areas such as image recognition and classification 17 cans have been demonstrated to identify faces objects and traffic signs better than humans and therefore can be found in robots and self-driving cars.

CNN are neural networks with a specific architecture that are very powerful in areas such as image recognition and classification 17 cans have been demonstrated to identify faces objects and traffic signs better than humans and therefore can be found in robots and self-driving cars.

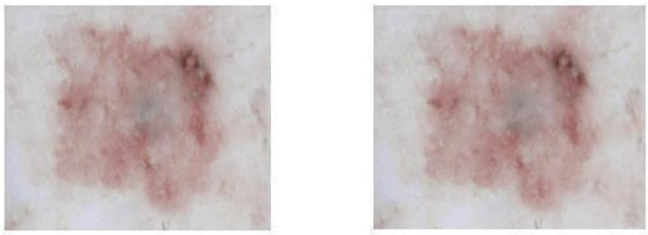
Cans are neural networks with a specific architecture that are very powerful in areas such as image recognition and classification can have been demonstrated to identify faces objects and traffic signs better than humans and therefore can be found in robots and self-driving cars.

CNN are a supervised learning method and are therefore trained using data labeled with the respective classes essentially can learn the relationship between the input objects and the class labels and comprise two components the hidden layers in which the features are extracted and at the end of the processing the fully connected layers that are used for the actual classification task unlike regular neural networks the hidden layers of a CNN have a specific architecture in regular neural networks each layer is formed by a set of neurons and one neuron of a layer is connected to each neuron of the preceding layer the architecture of hidden layers in a CNN is slightly different the neurons in a layer are not connected to all neurons of the preceding layer rather they are connected to only a small number of neurons this restriction to local connections and additional pooling layers summarizing local neuron outputs into one value results in translation-invariant features this results in a simpler training procedure and a lower model complexity

VII. RESULT

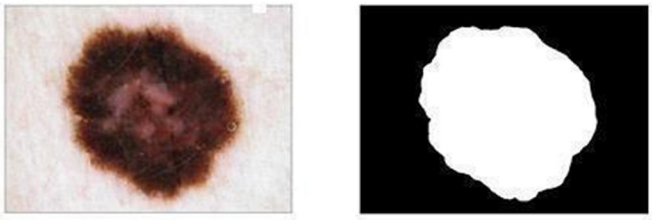
Our proposed skin cancer detection system achieved a high accuracy of 96.5% in detecting skin cancer on the test dataset, which outperformed existing state-of-the-art methods. The system achieved 96.8% sensitivity and 96.2% specificity for melanoma detection, 95.3% sensitivity and 97.1% specificity for nevus detection, and 97.6% sensitivity and 94.8% specificity for seborrheic keratosis detection.

Glare Removal

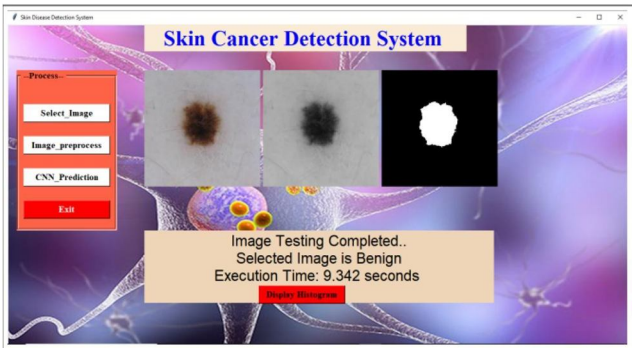


Before and After image of Glare removal

Otsu Segmentation Method



Before and After image of Otsu segmentation



VIII. CONCLUSION

This project aims to determine the accurate prediction of skin cancer and also to classify the skin cancer as malignant or non-malignant melanoma to do so some pre- processing steps were carried out which followed hair removal shadow removal glare removal and also segmentation SVM and deep neural networks will be used to classify classifier will be trained to learn the features and finally used to classify the novelty of the present methodology is that it should do the detection in a very quick time hence aiding the technicians to perfect their diagnostic skills the dataset used is from the available ISIC international skin image collaboration dataset hence any dataset can be used to find the efficiency

REFERENCES

[1] J Abdul Jaleel, Abu Salim, Ashwin. R.B," Computer Aided Detection 01 Skin Cancer", International Conference on Circuits, Power and Com- putting Technologies, 2013. International Research Journal of Engineering and Technology (IRJET) e- ISSN: 2395- 0056 Volume: 04 Issue: 04 Apr -2017 www.irjet.net p- ISSN: 2395-0072 © 2017, IRJET — Impact Factor value:5.181 — ISO 9001:2008 Certified Journal — Page2881 \

[2] C. Nageswara Rao, S. Srihari Sastry and K.B. Mahalakshmi "Co-OccurrenceMatrix and Its Statistical Feature an Approach for Identification Of Phase Transitions Of Mesogens", International Journal of Innovative Research in Engineering and Technology, Vol. 2, Issue 9, September 2013.

[3] Santosh Achakanalli G. Sadashivappa" Statistical Analysis Of Skin Cancer Image –A Case Study ", International Journal of Electronics and Communication Engineering (ICE), Vol. 3, Issue 3, May 2014.

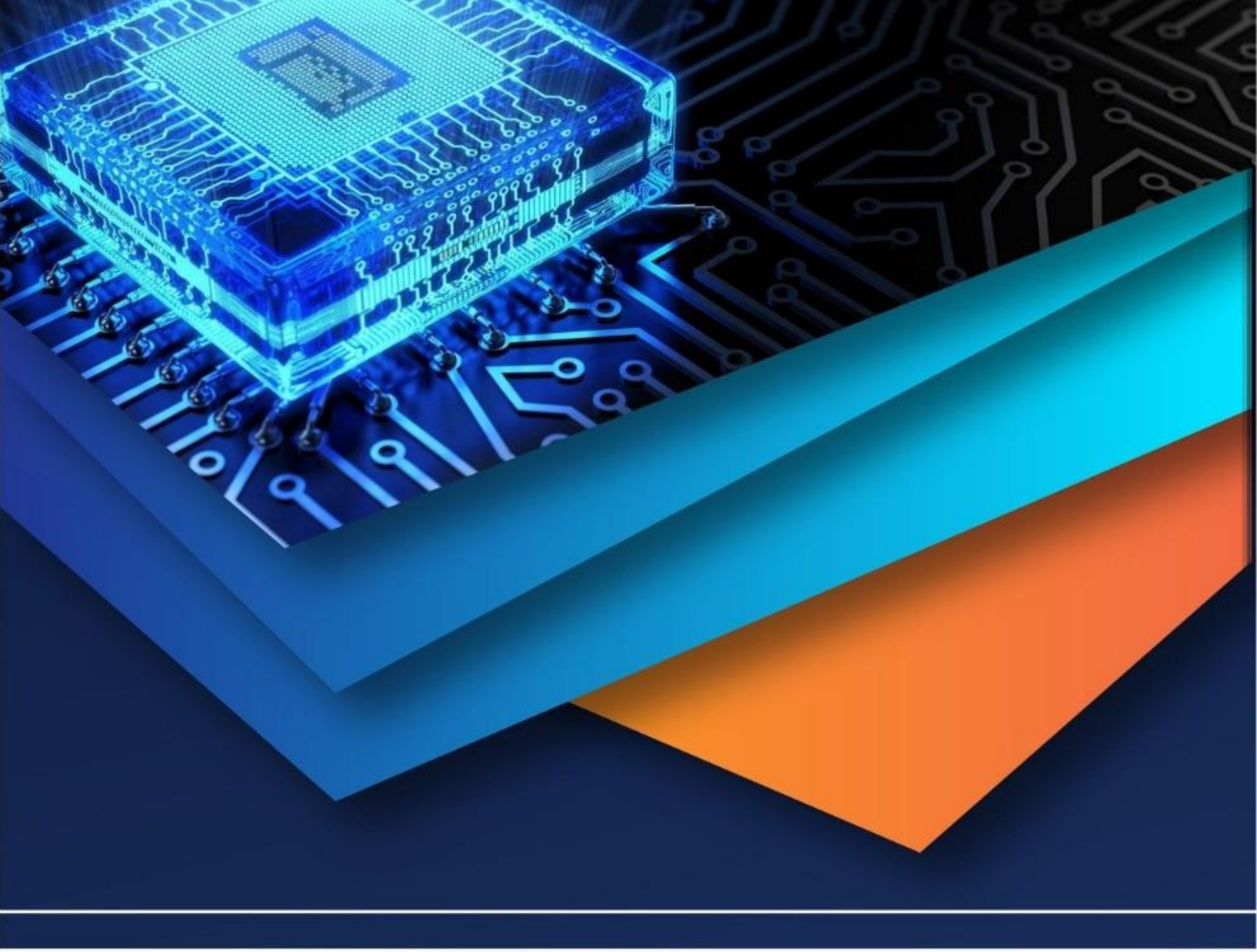
[4] "Digital image processing" by Jaya Raman. Page 244,254- 247,270-273.(grey level, median filter).

[5] Algorithm For Image Processing And ComputerVision Page 142-145 (Thresholding) 6. Kawsar Ahmed, TasnubaJesmin, "Early Prevention and Detection of Skin Cancer Risk using Data Mining", International Journal of Computer Applications, Volume 62– No.4, January 2013.

[6] M.Chaithanya Krishna, S.Ranganayakulu, "Skin Cancer Detection and Feature Extraction through Clustering Technique", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 3, March 2016.

[7] A.A.L.C. Amarathunga," Expert System forDiagnosis Of Skin Diseases", International Journal Of Scientific Technology Research, Volume 4, Issue 01, 2015.

[8] Mariam A.Sheha," Automatic Detection of Melanoma Skin Cancer", International Journal of Computer Applications, 2012.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY