iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78697

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Web-Based Platform for Multi-Modal Medical Image Analysis Using X-Rays, MRI, and CT Data

Ms. Lakshmi Deepthi N[1], Nithish Yemma[2], Ramya Swarna[3], Nikhitha M[4]
Dept. of CSE (Data Science), Institute of Aeronautical Engineering, Dundigal, Hyderabad, India

Abstract: In this application medical images of different modalities are used for image analysis and fusion. For final fusion MRI (Magnetic Resonance Imaging) and CT (Computed Tomography) images are taken as input and which gives fused image as output. This fused image has more information than individual images. Proposed method uses VMD (Variational Mode Decomposition) decomposition method and IMF method for fusion. Final fused image is compared with individual MRI and CT image using different metrics such as entropy, mutual information, edge intensity, PSNR, SSIM, MSE etc. Proposed method used CNN (Convolutional Neural Network) for analysis of fused image and to predict the final results as a normal scan or scan with tumor. Proposed CNN model finds results very perfectly than state of art machine learning models.
Keywords: Medical Image Analysis, Image Fusion , VMD , PSNR , Entropy , SSIM, CNN.

I. INTRODUCTION

Medical imaging is an essential part of healthcare in the contemporary world since it aids physicians to identify and diagnose illnesses at an early stage. Nonetheless, single imaging can result in delayed or incorrect diagnosis, particularly in such severe diseases as lung cancer, pneumonia, or Brain tumor. In order to overcome this issue, scientists are working on the integration of various imaging technologies, including X-rays, MRI, CT scans, and even audio messages, including cough or lung sounds, and artificial intelligence. The purpose of these systems is to enhance the accuracy of diagnoses, decrease the anthropogenic factor, and make decisions quicker. Our primary research problem is to create a safe and user-friendly web platform to enable medical practitioners to post and analyze multi-modal medical images using AI methods to identify diseases effectively. A variety of multimodal strategies have demonstrated good outcomes. Ghori [1] suggested to combine the results of chest X-ray with the results of lung-sound signals to detect the lung cancer at an initial stage and it was found that the combination of both the data has much better sensitivity than the single-input result with the traditional analysis. Their system introduces the role of audio-visual fusion and further demonstrates how the web based access can be used to help with real time medical help. In the same line of reasoning, Malik and Anees [2] came up with an AI model that combines chest X-rays, CT images as well as cough sound images to classify a number of respiratory diseases. Their results show that multimodal data improve learning features and increase accuracy but they point out that validation has to be undertaken in actual clinical settings in which the quality of the data differs greatly.

Improvements in healthcare automation have more recently existed in the vision-language models. Al-Hamadani [3] proposed a system, which can automatically process medical images and produce reports on diagnosis and enhance the efficiency of the working process. This method suggests the opportunities in the future where physicians will have visual and textual information that is supported by AI provided directly through digital channels. Besides model architecture, training strategies are also being enhanced by researchers. Kingsley and Izuchukwu [4] demonstrated that multidisease and multidomain multidimensional data augmentation and regularization methods can assist AI models to work more effectively. Moreover, vision transformer (as suggested by Kaliappan et al. [5]) is a type of unified deep learning models that can be used to classify various categories of medical images, which is why such solutions may be deployed by hospitals with limited resources. Based on these developments, our paper aims at developing a multi-modal medical imaging web-based platform and uploading X-ray, MRI and CT images to enable automated analysis with deep learning. The system shall offer a reduction in noise, the detection of abnormalities and the provision of helpful visual information to facilitate medical decision making. Another point that we focus on is secure data management and convenient design because doctors should be able to assess various types of images in a single location effectively. In this study, we will make a contribution towards more rapid, more precise, and more accessible healthcare diagnostics.

II. LITERATURE SURVEY

Ghori introduces a multimodal model of lung cancer detection integrating the use of both the chest X-ray images and lung-sound recordings in making the diagnosis of the disease at an earlier stage. The system is based on CNNs to analyze images and GRU-based models to interpret the sound and provides the results in an accessible medical support via the web platform.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

The research indicates high sensitivity with combination of the two modalities, which validates the fact that audio-visual fusion has the ability to detect anomalies at an earlier stage than the conventional single-input systems. Nevertheless, the methodology remains to be widely tested in clinics and better explainable, so it could be adopted by the medical community. [1]

Malik and Anees introduce a deep learning system that distinguishes nine significant chest conditions through a combination of chest X-ray, CT, and cough-sound images transformed into scalogram images. They use the CNN-based design with better pooling, normalization, and balancing with SMOTE to address the limitations of a dataset to have very high diagnostic accuracy relative to conventional models. This study shows that images and cough sounds come together as a good way of learning features of respiratory diseases. However, the reliance on publicly available data makes it questionable how the technology can be used in practical settings under various clinical scenarios.[2]

This paper presents a vision-language model (VLM) system of automated medical image analysis and clinical reporting of CT, MRI, X-ray, and ultrasound scans. The system combines natural language processing and anomaly detection, providing color overlays and formatted reports to aid radiologists. Its zero-shot means that it will show more dependency on massive labeled datasets. Despite its potential in efficient workflow, the model needs to be strictly validated to avoid the problem of errors with automatic text generation in the context of sensitive medical facilities.[3]

The authors are concerned with the enhancement of diagnostic AI models by means of sophisticated data augmentation and regularization means within CNN models. Their approach gives them stronger resistance to changes in medical images and their accuracy and segmentation performance is superior on diseases such as pneumonia, breast cancer and diabetic retinopathy. The results indicate that appropriately designed augmentation can help greatly to increase the generalization under limited datasets. Nonetheless, the issues of deployment are still present, such as the costs of computing and the necessity to ensure that augmented data corresponds to actual-patient situations. [4]

Kaliappan et al. introduce a single Vision Transformer capable of classifying medical images of various modalities, including X-ray, MRI, and dermatology photos, in the same model. The attention-based design enhances feature attention and can be explained by Grad-CAM visualization, making clinicians have confidence in the decision-making procedure. The model can be used with great accuracy, which is reported to be high, in resource-restrained regions, where numerous individual diagnostic systems are hard to maintain. Nevertheless, transformer models are computationally-intensive and need domain-specific fine-tuning. [5]

The present review paper will examine the current advances in AI-based medical imaging and define both technical successes and significant obstacles to clinical implementation. It outlines that there is a necessity to have a secure and fast data exchange, effective model deployment, standard imaging processes, and robust interpretability to sustain real healthcare decisions. The authors stress that it is not sufficient to use deep learning but rather effective systems need to fit within clinical settings. Their observations form a guiding principle on how to come up with safe and reliable diagnostic instruments.[6]

III. PROPOSED METHOD

Proposed method block diagram is explained below,

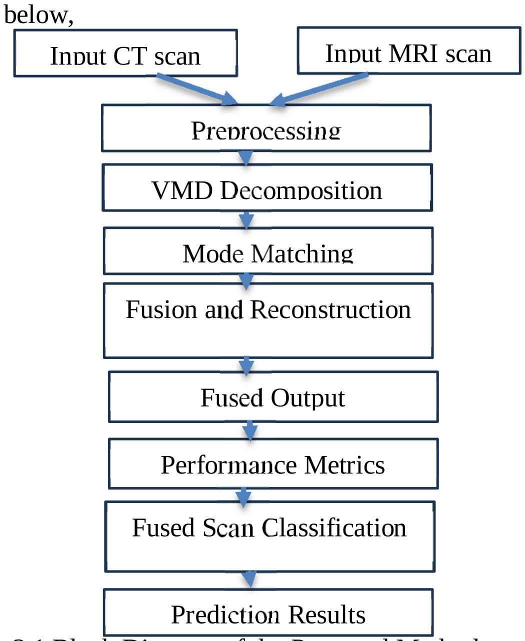


Fig.3.1 Block Diagram of the Proposed Method

Proposed method all blocks are explained step by step below in detail,

A. Overall System Purpose

The system is structured in such a way that the medical professionals can post the X-ray, MRI, and CT images by a secure web interface and get automated diagnostic assistance. All image modalities present independent clinical information. X-rays give a clear image of the skeletal structure and lung structure, MRI gives emphasis on the soft-tissue structure and CT scans give anatomical accuracy in three-dimensions. The system increases the reliability of the diagnostic process, leads to reduced decision-making time, and increases the availability of highly sophisticated tools in hospitals and remote clinical settings by using deep learning to analyze images and deliver multimodal interpretation through a single web platform.

B. Input

The system receives uploaded medical images of various modalities which include the X- rays, MRI scan and CT scan. These images could be used to depict different parts of the body depending on the need in medicine. The uploaded files present the raw diagnostic input which can have a diverse format, image resolution, brightness range, and noise level as a result of variation in imaging equipment and imaging acquisition conditions.

C. Pre-Processing and Standardization

In order to deliver consistent analysis on a wide range of inputs, the system will carry out pre-processing, which normalizes the quality and format of images. All images are brought to an equal resolution, the intensity is normalized, and noise caused by scanning devices or other factors in the environment is minimized. Should it be necessary, contrast enhancement techniques can be used to enhance relevant anatomical structures and segmentation can be used to isolate important features like lungs or brain tissues so that they can be interpreted more easily. It is a unified pre-processing pipeline which guarantees that the deep learning model is fed with clean, comparable, and structurally consistent input samples.

D. Deep Feature Extraction

After being standardized, images can be fed through pre-trained deep neural network models like ResNet, EfficientNet, or Vision Transformer. These nets are automatic feature extractors which are trained to learn important visual patterns such as textures, shapes, lesion patterns, tissue densities, as well as structural abnormalities. The deep model learns to draw the line between normal tissue structure and pathological patterns based on modalities instead of manually engineering features. This is achieved by an automated representation learning that identifies modality-specific and cross-modality medical cues that are important in the correct diagnosis

E. Multimodal Fusion

The system then applies multimodal fusion after extracting meaningful features in each of the input modalities that deep features are then merged to create one representative of diagnostic features. In feature-level fusion, the system is able to use complementary clinical data: bony data of X-rays, soft-tissue data of MRI, and volume or cross-sectional anatomical clarity of CT scans. The combination of these various medical viewpoints by the fusion process results in a more valuable and valid diagnostic interpretation than in single-modality analysis and has a much higher degree of robustness and clinical usefulness.

F. Abnormality Detection and Classification

The fused feature representation is sent to a classification head that either finds out whether there is disease or not and classifies abnormalities identified. The model gives prediction labels and probability scores in terms of confidence levels. Besides, the system employs class-activation mapping techniques to highlight suspicious areas hence explainable. This diagnostic prediction layer is used as the decision core and it is an indication of the learned wisdom of the system on the basis of extensive annotated medical data.

G. Visualization of Heatmap and Explainability

To improve clinical trust and support medical practitioners during the interpretation, the system produces attention-based heatmaps that represent visualization of abnormal tissue regions. The explainability methods based on grad-cam identify the precise spatial regions that influenced the decision made by the model by which transparency and validation are attained.

This will make sure that the tool does not act as a black-box system but as an aiding diagnostic engine that clearly reveals pathological areas to help radiologists and doctors.

H. Web-Based Output Interface

The ultimate deliverable is presented in an easy to use web interface, in which the results are presented with prediction labels, confidence scores, and heat maps. The interface also enables user to log in securely, manage and download diagnostic summaries, as well as manage patients and images. The web platform also makes the solution scalable, accessible and friendly to hospitals as it allows clinicians to use advanced AI-based medical imaging tools without installing software on their computers or high computational resources on their computers and makes the solution easily accessible.

I. Performance Measures and Review

Performance of the model is measured against clinical image databases and reported in terms of the important medical imaging measures of accuracy, precision, recall, specificity, sensitivity and F1-score. Prediction consistency is determined by receiver-operating characteristic curves and the correctness of predictions on a per-class scale is determined by confusion matrices. The ongoing assessment will maintain a high level of medical-grade reliability of the system, reduce false alarms, and make the system clinically reliable prior to implementation.

IV. RESULT ANALYSIS

Proposed image analysis complete detail study and results are shown below,

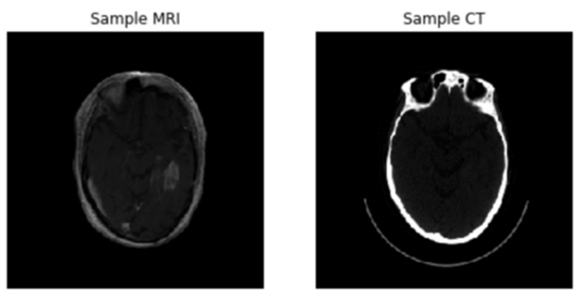


Fig. 4.1 Sample MRI and CT scan for the same person

Here the image of MRI and CT scan of the same person is selected which are used for further image fusion and analysis.

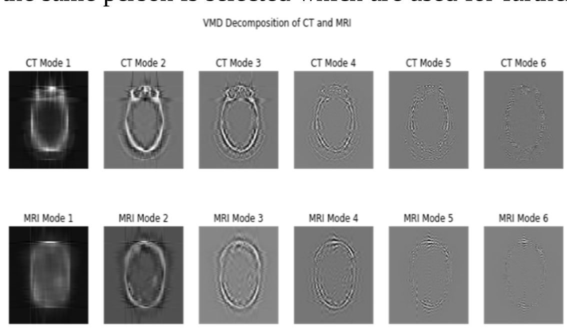


Fig.4.2 VMD decomposition of CT and MRI scans

Before applying the image fusion of the MRI CT scans both are using VMD decomposition for decomposing images to different modes and these different modes are used for fusion of images.

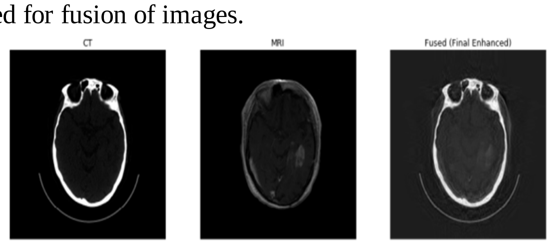


Fig.4.3 input individual CT, MRI and Fused Final Image

In above results, first image is individual CT scan of a patient and second image is the MRI scan of same person while third image shows the fused image using both MRI and CT scans of the same person.

	Metric	CT	MRI	Fused
0	Entropy	1.74702	1.78513	3.06403
1	Standard Deviation	0.182517	0.061112	0.144396
2	Edge Intensity	0.020709	0.00702254	0.0211916
3	Mutual Information (CT)	0.296422	0.258318	0.621202
4	Mutual Information (MRI)	0.0612756	0.735509	0.431125
5	Cross-Correlation (CT)	1	0.344283	0.983461
6	Cross-Correlation (MRI)	0.344283	1	0.420515
7	PSNR (CT)	100	15.2173	18.328
8	PSNR (MRI)	15.2173	100	14.4145
9	SSIM (CT)	1	0.747767	0.208809
10	SSIM (MRI)	0.747767	1	0.111727
11	MSE (CT)	0	0.0300798	0.014696
12	MSE (MRI)	0.0300798	0	0.0361864

Fig. 4.4 Performance Metrics

Different parameters are calculated to show the superior performance of proposed image fusion compared to individual images.

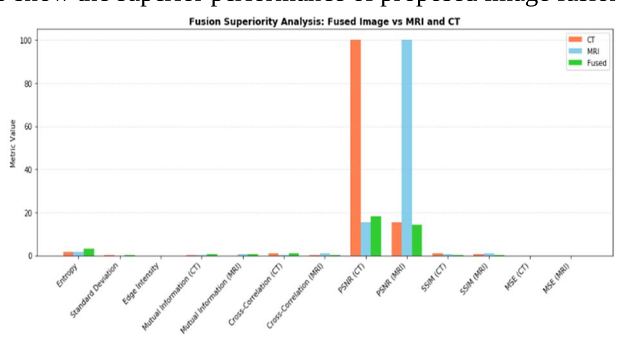


Fig. 4.5 Plot of all the metrics for individual and fused image

Different metrics shown in the above tabular format are plotted. Using matplotlib we plotted bar graph to represent the performance of proposed fusion-based image analysis.

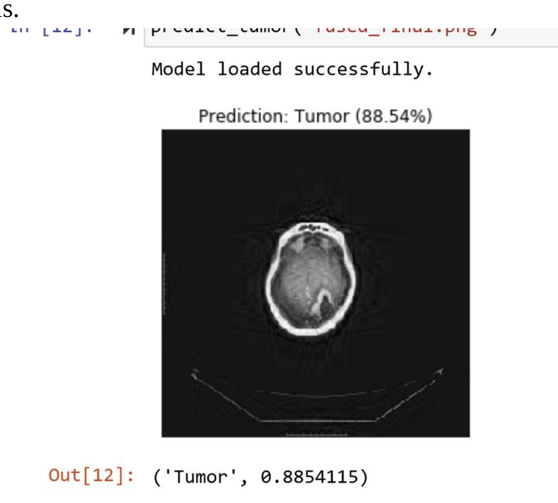


Fig. 4.6 Predicted Results by CNN

The above fused scan is predicted as having Tumor. The confidence score for tumor is shown as 0.8854. fused image is more likely to get correct output as they have more features than the individual images.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

V. CONCLUSION

Proposed method successfully designed using python and relevant dataset. The dataset contains MRI and CT scan of the same person with the similar part. VMD decomposition and IMF are used for fusion from input MRI scan and CT scan. The performance of proposed fusion is calculated using different metrics such as PSNR, SSIM, Entropy, MSE etc. Further fused image is passed to CNN algorithm for analysis. CNN image is first trained with normal images and tumor images. The new fused test image is fed to CNN algorithm for testing and CNN will predict the fused image is either normal or having tumor, if tumor then how much percentage confidence.

Future this application can be integrated with android application to make this application available to everyone.

REFERENCES

[1] Ghori, Khawaja Waqas Ur Rehman. "Multimodal Deep Learning for Lungs Cancer Detection: Integrating Audio and Image Analysis with Web-Based Accessibility." PhD diss., Dublin, National College of Ireland, 2025.

[2] Malik, Hassaan, and Tayyaba Anees. "Multi-modal deep learning methods for classification of chest diseases using different medical imaging and cough sounds." Plos one 19, no. 3 (2024): e0296352.

[3] Al-Hamadani, Samer. "Intelligent Healthcare Imaging Platform An VLM-Based Framework for Automated Medical Image Analysis and Clinical Report Generation." arXiv preprint arXiv:2509.13590 (2025).

[4] Kingsley, Nwizua Felix, and Amannah Constance Izuchukwu. "Optimization of Medical Image Analysis Models for Effective Disease Diagnosis through Data Augmentation Techniques." Journal of Infectious Diseases and Patient Care (2025).

[5] Kaliappan, M., E. Mariappan, V. Manimaran, and B. Revathi. "Unified Vision Transformer for Multimodal Medical Image Classification." Available at SSRN 5281786.

[6] Panayides, Andreas S., Amir Amini, Nenad D. Filipovic, Ashish Sharma, Sotirios A. Tsaftaris, Alistair Young, David Foran et al. "AI in medical imaging informatics: current challenges and future directions." IEEE journal of biomedical and health informatics 24, no. 7 (2020): 1837-1857.

[7] Singh, M., & Azam, M. A Review of Multimodal Medical Image Fusion Techniques. Journal of Medical Systems, 2020. Wiley Online Library

[8] Trägårdh, E., et al. RECOMIA — a cloud-based platform for artificial intelligence research in medical imaging. EJNMMI Physics, 2020. SpringerOpen

[9] Egger, J., et al. Studierfenster: an open science cloud-based medical imaging platform for visualization and analysis (CT, MRI). Journal of Digital Imaging, 2022. SpringerLink

[10] Saleh, M. A. A Brief Analysis of Multimodal Medical Image Fusion: techniques and assessment. Electronics, 2022. MDPI

[11] Liu, S., et al. Two-Scale Multimodal Medical Image Fusion Based on Spatial-Frequency Methods. Frontiers in Computational Neuroscience, 2022. Frontiers

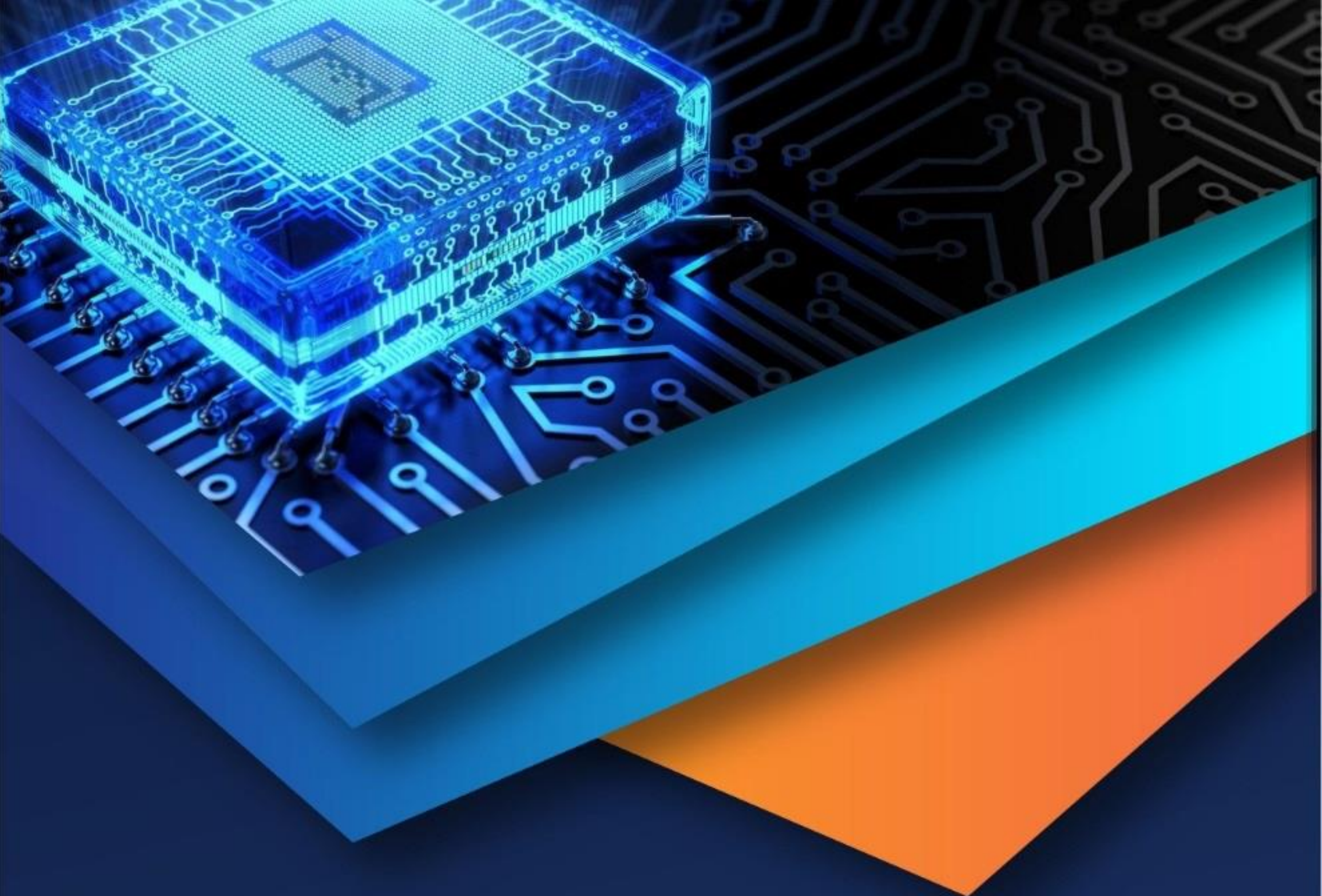
[12] Mei, X., et al. RadImageNet: an open radiologic deep learning pretraining dataset and models. Radiology: Artificial Intelligence / RSNA, 2022. Radiological Society of North America

[13] [13] Li, M., et al. Medical image analysis using deep learning algorithms: a review (Frontiers). 2023. Frontiers

[14] [14] Jozić, K., et al. DICOM SIVR: a web architecture and platform for seamless image and volume rendering of DICOM images. Journal/Conference, 2022. ScienceDirect

[15] [15] Min, Q., et al. Web-Based Technology for Remote Viewing of Radiological Images. Journal of Medical Internet Research, 2020. JMIR Publications

[16] [16] AboArab, M. A. Advancing Progressive Web Applications to Leverage Medical Imaging: a modular PWA design for DICOM and MPR visualization. JMIR Medical Informatics, 2024.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY