iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 14 **Issue:** V **Month of publication:** May 2026

**DOI:** https://doi.org/10.22214/ijraset.2026.82248

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 14 Issue V May 2026- Available at www.ijraset.com*

# Enhancing Brain Tumor MRI Classification Using Transfer Learning and Diffusion-Based Augmentation

Rohan Shinde[1], Janhavi Manjarekar[2], Vishal Akale[3], Pramod Jadhavo[4], Pritish Bisne[5]
*Department of MCA, Trinity Academy of Engineering, Pune, India*

***Abstract:*** ***Brain tumor classification using magnetic resonance imaging (MRI) is essential for early diagnosis, but its performance is often limited by small and imbalanced datasets. This study uses 196 MRI images from Kaggle for binary classification. To address data scarcity, multiple approaches are evaluated, including traditional augmentation, patch-based learning, diffusion-based augmentation, and transfer learning. A baseline CNN achieves 87% accuracy but suffers from overfitting. Conventional augmentation provides limited improvement, while patch-based and diffusion-based methods negatively affect performance due to label noise and loss of structural details. In contrast, transfer learning with MobileNetV2 significantly improves generalization. The results show that MobileNetV2 achieves the best performance, highlighting that preserving data quality and leveraging transfer learning are more effective than simply increasing dataset size.***

***Keywords: Brain Tumor Detection, MRI Imaging, Data Scarcity, Image Preprocessing, Patch-Based Learning, Computer-aided Diagnosis(CAD), Deep Learning, Convolutional Neural Networks (CNN), Diffusion Models, Data Augmentation, MobileNetV2***

## I. INTRODUCTION

Medical image classification now plays a huge part in early diagnosis and clinical decision-making. Take brain tumor classification with MRI it's especially important, since getting this right can mean better treatment plans and fewer diagnostic errors. But there's a real problem: unavailability of enough labeled medical images. Most datasets are small because of privacy rules and, honestly, because it takes real experts to label this stuff.

In this research, we look at classifying brain tumors from just 196 MRI images. Since the dataset is small, the model easily overfits, memorizing training samples instead of learning general patterns. We first trained a basic CNN, which performed reasonably well but showed clear overfitting. Conventional data augmentation techniques, such as image rotation and flipping, were subsequently applied to improve model generalization; however, these transformations resulted in only marginal performance improvements and did not significantly mitigate the issue of overfitting.

In order to increase the size of the training dataset, image patches were created; thus, more data became available to the model. While such a procedure allowed for an expansion of the available data, it inevitably led to a lot of noise because not all patches included parts of tumors, negatively impacting the model's training process. Transfer learning was then employed, using MobileNetV2, which allowed the model to benefit from pre-trained features based on large amounts of image data. This method improved the results and reduced the risk of overfitting. Diffusion augmentation was also tried out to create additional image data, but the resulting artificial images tended to lack structure compared to actual MRI scans and negatively impacted the results.

All of this demonstrates that having a higher amount of training data does not necessarily lead to better results. More attention needs to be paid to the quality of data provided to the algorithm. The research presented here provides a comparative overview of various methods for obtaining training data in cases when there is a shortage of MRI scan images.

## II. LITERATURE SURVEY

AI and deep learning have changed the landscape of medical image analysis in a big way. Now, doctors can rely on algorithms to spot diseases, classify them, and even isolate problem areas in scans more quickly and accurately. These tools help clinicians act faster and stay consistent with their decisions. However, a key challenge arises: deep learning models are only as good as the data they're trained on. They need large, varied, and well-labeled datasets to work well. Getting that kind of data for medical imaging is tough. Privacy rules make sharing data complicated, annotating images is expensive, and there aren't enough specialized experts around. On top of that, the data itself is often unbalanced, with some conditions represented much more than others.

Because of all this, tools like data augmentation and synthetic data generation have become essential. Generative Adversarial Networks (GANs) are everywhere in medical imaging research right now. Islam et al. [2] did a deep dive into how GANs are being used to create, enhance, and segment medical images. They found that images generated by GANs help expand datasets and make models more robust. But nothing's perfect—training these networks can be unstable, and they sometimes go off the rails and produce repetitive results, which limits their clinical reliability. Kumar et al. [4] also explored GANs in healthcare and found they're pretty good at keeping the important visual and statistical details while helping with privacy. Still, making GAN-generated images look clinically real enough for true medical use is a hurdle the field hasn't fully cleared. Recently, diffusion-based generative models have gotten some attention as an alternative to GANs. Yang et al. [1] came up with a diffusion-model-based system for grading skin disease images, and it outperformed traditional GANs, producing higher-quality and more stable images. These models avoid common GAN problems like mode collapse and offer better control during generation. The trade-off? Diffusion models use a lot more computing power, which makes them less practical in places with limited resources. Some researchers have merged generative models with clever data augmentation to boost results. For example, Villanueva and Kumar [11] combined GANs with a customized ResNet-50 network for recognizing multiple skin diseases. They brought in Explainable AI tools like LIME and SHAP to help make sense of what the model was focusing on—which is important in medicine. Their setup reached 92.5% accuracy and a 98.82% AUC score. Impressive, but again, everything hinges on how realistic and useful the synthetic images are. If the fake data isn't high-quality, the whole system stumbles. On the more traditional side, people are still coming up with smart ways to tweak and expand datasets. Faryna et al. [6] automated data augmentation for histopathology, adjusting for staining differences to fight issues with inconsistent data. It helped reduce the effects of domain shift, so models stayed more reliable. Sun et al. [7] introduced LCAMix—an augmentation method that preserves important structural and edge details, fixing the shortcomings of simple techniques like rotation or flipping. Later, they came up with HSMix, which blends hard and soft mixing strategies to make new samples that retain their meaning. Both methods focus on keeping the critical features intact for accurate diagnosis. Another area getting more attention is fairness and bias. Juwara et al. [9] showed that generating synthetic data can help make underrepresented classes more visible, cutting down on bias and boosting the accuracy and fairness of the models. Balancing datasets this way is promising, but creating synthetic data that's totally unbiased and stays clinically valid is still a tricky task. People aren't just generating new images, either recent studies are using Large Language Models (LLMs) for data augmentation. Ding et al. [10] looked at ways to use LLMs for prompt based data generation and transformation, especially useful when there isn't a lot of labeled data. These models can generate diverse, context-rich data, but they also bring headaches like "hallucination" (making up details), inconsistency, and evaluation difficulties. Not quite ready for mission-critical medical use, but the potential's there. GAN-based methods have even been extended to challenging work like brain tumor segmentation. Kalejahi et al. [5] developed a framework based on 3D GANs that captures spatial relationships in MRI scans, giving better segmentation results. Chatterjee and Byun [3] found that GANs can help fix imbalanced datasets, a common problem both in medicine and industry. But, once again, the real-world impact depends on generating synthetic samples that are up to the job. In spite of big progress, several obstacles remain. Step one: synthetic data has to make medical sense, anything subpar drags down performance. Second: Accuracy is great, but many current approaches don't explain their predictions, which is a serious problem for clinical acceptance. Finally, getting models to perform consistently across different imaging protocols and datasets is still a major challenge. Clearly, the next step is building a smarter, broader framework that brings together advanced generative methods, efficient deep learning, and real transparency through explainable AI. There's untapped ground in blending GANs, diffusion models, and interpretability tools— especially for brain tumor analysis using limited MRI sets. Cracking this will set the stage for AI systems that doctors can trust, bringing reliable, generalizable solutions to medical image analysis.

## III. METHODOLOGY

### A. *Dataset Description*

The study uses a brain tumor MRI dataset from Kaggle comprising 196 images classified into tumor and non-tumor categories. Due to the limited dataset size, there is a high risk of overfitting. To address this, stratified sampling was applied to split the data into 70% training, 15% validation, and 15% testing, ensuring balanced class distribution across all sets.

### B. *Data Preprocessing*

Preprocessing plays a crucial role in transforming raw data into a structured and usable format.

The following steps are performed:

*1)* Resize images to 224x224 pixels — this fits with pre- trained deep learning models.

2) Normalize pixel values to [0,1] for steadier training.
3) Reduce noise just enough for clearer images.
4) Shuffle everything to avoid bias.

### C. Problem Formulation

The dataset is structured as follows:
D = { (x_i, y_i) | i = 1, 2, ..., N }
Where x_i's the MRI scan, and y_i □ {0,1} gives you the class. For the mapping:
f(x; θ)→y (1), where θrepresents the model parameters optimized to minimize classification loss.

### D. Baseline CNN Model

We developed a baseline Convolutional Neural Network (CNN) to establish initial performance. The model extracts spatial features through convolutional layers, and is subsequently processed through fully connected layers for classification. The network is trained using the cross- entropy loss function:
L = −∑ y_i log (ŷ_i) (2)
Here's the setup: - Optimizer: Adam - Learning rate: 0.001 - Batch size: 32 - Epochs: 25
The baseline model achieved a peak accuracy of approximately 87%; however, it exhibited clear signs of overfitting, a common issue when training with limited data.

### E. Traditional Data Augmentation

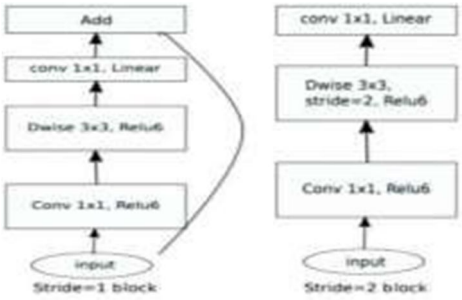
To enhance data variability, standard augmentation techniques such as rotation, flipping, scaling, and brightness adjustment were applied:
x' = T(x) (3), where ( T ) denotes a transformation applied to the input image. Although these methods increased data diversity, they did not introduce substantial new structural information, resulting in only limited performance improvement.

### F. Patch-Based Learning Approach

To further expand the dataset, each MRI image was partitioned into smaller patches, thereby increasing the number of training samples.
x → {p1, p2, ..., pk} (4)
This approach increased the dataset size from 196 images to more than 3000 patches. However, an important limitation arises, as patches that do not contain tumor regions are still assigned the same labels. This introduces label noise, which can interfere with effective feature learning and ultimately reduces classification accuracy.

### G. Transfer Learning Using MobileNetV2

To address the challenge of limited data, transfer learning was employed using MobileNetV2. The model utilizes weights pre-trained on ImageNet for feature extraction and subsequently fine-tunes the final layers for the classification task. In model terms:
f(x) = g(h(x)) (5) with h(x) extracting features and g(.) handling the class prediction.
Fine-tuning significantly improves model performance.
This method adapts to MRI data, boosts generalization, and reduces overfitting.




MobileNetV2 Architecture

### H. Diffusion-Based Data Augmentation

Diffusion models were also explored to generate synthetic MRI samples. The forward process introduces noise into the data: $x_t = \sqrt{\alpha_t}x + \sqrt{1 - \alpha_t}\varepsilon$ (6)
and the reverse builds images back:
$x_{t-1} = (1/\sqrt{\alpha_t})(x_t - \sqrt{1 - \alpha_t}\varepsilon\theta(x_t, t))$ (7)
This looked promising, but in practice, it blurs out key features like tumor edges and results in a decline in classification accuracy.

### I. Experimental Setup

All models were implemented using TensorFlow or Keras, with GPU acceleration to improve computational efficiency. Model performance was evaluated using standard metrics, including accuracy, precision, recall, and F1-score, providing a comprehensive assessment, particularly in the presence of class imbalance.

### J. Comparative Analysis

The experimental results can be summarized as follows: the baseline CNN exhibits rapid overfitting; conventional data augmentation yields only marginal improvements; patch- based learning introduces label inconsistencies; transfer learning provides the most effective performance; and diffusion-based augmentation leads to a decline in accuracy.

### K. Key Findings

These findings indicate that increasing data volume alone is insufficient; data quality plays a more critical role than quantity.
Preserving the structural integrity and diagnostic relevance of medical images is essential. In scenarios with limited data, transfer learning proves to be the most effective approach.

### L. Summary of Proposed Framework

The methodology provides a clear comparative analysis of different approaches to brain tumor classification under limited data conditions.
In summary, emphasis should be placed on high-quality data augmentation and transfer learning rather than merely increasing the number of samples. These findings offer a useful reference for future work in medical imaging.

## IV. RESULTS AND DISCUSSION

### A. Experimental Results

To examine the impact of different approaches on brain tumor classification under limited data conditions, five experiments were conducted using a dataset of 196 MRI images. Each experiment focused on a different strategy to address overfitting and improve model generalization. The performance of all approaches is summarized in Table I.

| Experiment | Method | Accuracy (%) |
|---|---|---|
| Exp 1 | Baseline CNN | 87 |
| Exp 2 | Data Augmentation | 86 |
| Exp 3 | Patch-Based Learning | 75 |
| Exp 4 | MobileNetV2 (Transfer Learning) | 79 |
| Exp 5 | Diffusion-Based Augmentation | 75 |

Table 1 – Experimental Results of Proposed Approaches

*1) Exp 1: Baseline CNN Performance*

A standard Convolutional Neural Network (CNN) was trained on the original dataset and achieved the highest accuracy of 87%. However, further analysis revealed clear overfitting, as the model performed well on training data but poorly on unseen validation samples. This indicates that the model memorized training patterns rather than learning generalized features, which is common in small datasets.

*2) Exp 2: Traditional Data Augmentation*

Data augmentation techniques such as rotation, flipping, and brightness adjustment were applied to increase data diversity. The model achieved 86% accuracy, which is very close to the baseline. While augmentation slightly reduced overfitting, the improvement was minimal, suggesting that simple transformations are insufficient for capturing complex variations in MRI data.

*3) Exp 3: Patch-Based Learning Analysis*

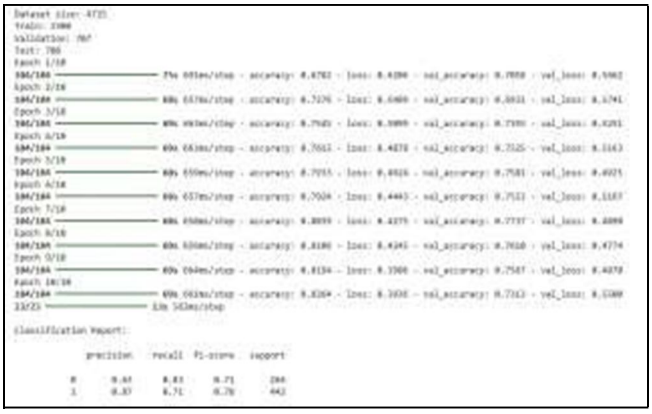
To increase dataset size, images were divided into smaller patches, expanding the dataset significantly. However, accuracy dropped to 75%. This decline is due to loss of contextual information, as many patches lacked tumor regions but were still labeled similarly, introducing noise and misleading the model.

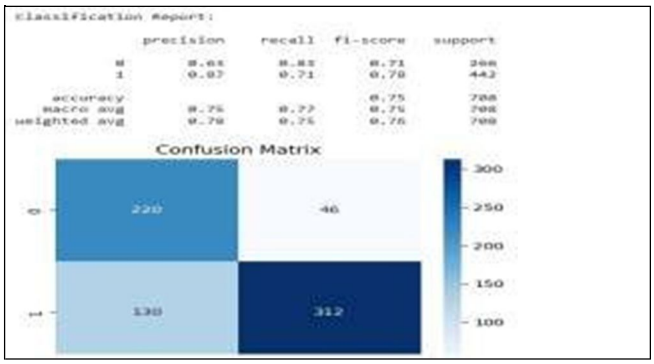
*4) Exp 4: Transfer Learning using MobileNetV2*

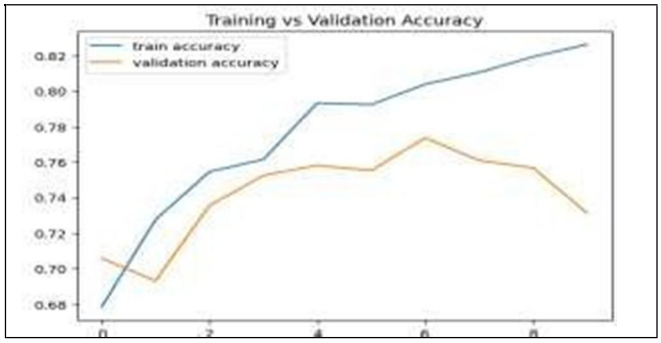
Transfer learning was implemented using a pretrained MobileNetV2 model. The model achieved 79% accuracy, which is lower than the baseline CNN. However, this model demonstrated better generalization and reduced overfitting. Unlike the CNN, MobileNetV2 learned more robust and transferable features from pretrained weights, making it more reliable for unseen data despite lower accuracy.

*5) Exp 5: Diffusion-Based Augmentation*

Synthetic data generation using diffusion techniques increased dataset size to approximately 4715 samples. However, accuracy remained at 75%. The generated images lacked realistic anatomical structure, leading the model to learn incorrect patterns.








Exp 5 Results

Key Insights:
Increasing dataset size alone does not guarantee better performance.
CNN models trained from scratch tend to overfit on small datasets.
Traditional augmentation provides only marginal improvements.
Patch-based methods can degrade performance due to loss of spatial context.
Transfer learning is more robust and reliable for small medical datasets.
Synthetic data must be highly realistic to be effective.

## V. CONCLUSION AND FUTURE WORK

This study took on the problem of classifying brain tumors from MRI scans when there aren't many images to work with. With only 196 MRIs, the main priority was to stop the models from overfitting while still keeping their predictions reliable. I tried several strategies—starting with a basic CNN, then adding data augmentation, patch-based learning, transfer learning with MobileNetV2, and finally, generating synthetic images using diffusion models.
The experiments showed that these approaches don't all have the same effect, especially when you don't have much data.

### *A. Key Findings*

The basic CNN actually reached high accuracy—on the training data—but clearly overfit and didn't do well on new images.
Old-school augmentation, like rotating or flipping images, didn't really help much. Those tricks don't add real variety, no new structures, just the same info shuffled around.
Patch-based learning gave the model a lot more data, but accuracy suffered because patches often miss the bigger picture or even introduce noise instead of useful features.
Out of everything tried, transfer learning with MobileNetV2 worked best. It stayed stable and generalized better, handling small datasets far more reliably than the other methods.
Diffusion-based synthetic data just didn't cut it—those generated images didn't look realistic or contain the right structure for clinical relevance, so models didn't benefit.
So, based on all this, it's clear: In medical image classification, having high-quality, relevant data matters much more than just having more data.

### *B. Conclusion*

Just adding more training samples wasn't the answer. When new data doesn't capture the real features of brain tumors, it actually makes things worse. Of all the approaches tested, transfer learning was hands-down the best way to deal with not having much data.
The bigger lesson here is that you can't ignore quality or clinical meaning when building AI for healthcare. If your method doesn't maintain anatomical or diagnostic information, it won't be trustworthy.

### *C. Future Work*

This project lays some good groundwork, but there are several ways to push further:

Work on better generative models—like more advanced GANs or better diffusion techniques—that can actually keep the medical details intact.

Try focusing on regions or tumor-specific areas for learning, instead of just chopping images into patches.

Incorporate bigger and more varied datasets; that always helps with making the models less biased and more robust. Bring in explainable AI tools, like Grad-CAM, so researchers and doctors can really see why the model made a certain call.

And finally, explore more hybrid architectures, maybe combining transformers or attention mechanisms, to see if those boost performance.

## REFERENCES

[1] T.-T. Yang, C.-W. Ma, C.-H. Lee, S.-X. Qiu, and C.-C. E. Lan, "AI grading of lateral canthal lines: Novel models for unseen synthetic image generation and data augmentation," Clinical, Cosmetic and Investigational Dermatology, 2024.

[2] S. Islam, M. T. Aziz, H. R. Nabil, J. Rahman, M. F. Mridha, M. M. Kabir, N. Asai, and J. Shin, "Generative adversarial networks (GANs) in medical imaging: Advancements, applications, and challenges," IEEE Access, 2024.

[3] S. Chatterjee and Y.-C. Byun, "Leveraging generative adversarial networks for data augmentation to improve fault detection in wind turbines with imbalanced data," Results in Engineering, 2024.

[4] A. Kumar, P. P. Kumar, and O. Prakash, "Exploring generative adversarial networks for synthetic data generation in healthcare applications," NeuroQuantology, vol. 19, no. 10, pp. 523–532, 2021.

[5] B. K. Kalejahi, S. Meshgini, and S. Danishvar, "Segmentation of brain tumor using a 3D generative adversarial network," Diagnostics, vol. 13, 2023.

[6] K. Faryna, J. van der Laak, and G. Litjens, "Automatic data augmentation to improve generalization of deep learning in H&E stained histopathology," Computers in Biology and Medicine, vol. 170, 2024.

[7] D. Sun, F. Dornaika, and J. Charafeddine, "LCAMix: Local-and-contour aware grid mixing based data augmentation for medical image segmentation," Information Fusion, vol. 110, 2024.

[8] D. Sun, F. Dornaika, and N. Barrena, "HSMix: Hard and soft mixing data augmentation for medical image segmentation," Information Fusion, vol. 115, 2025.

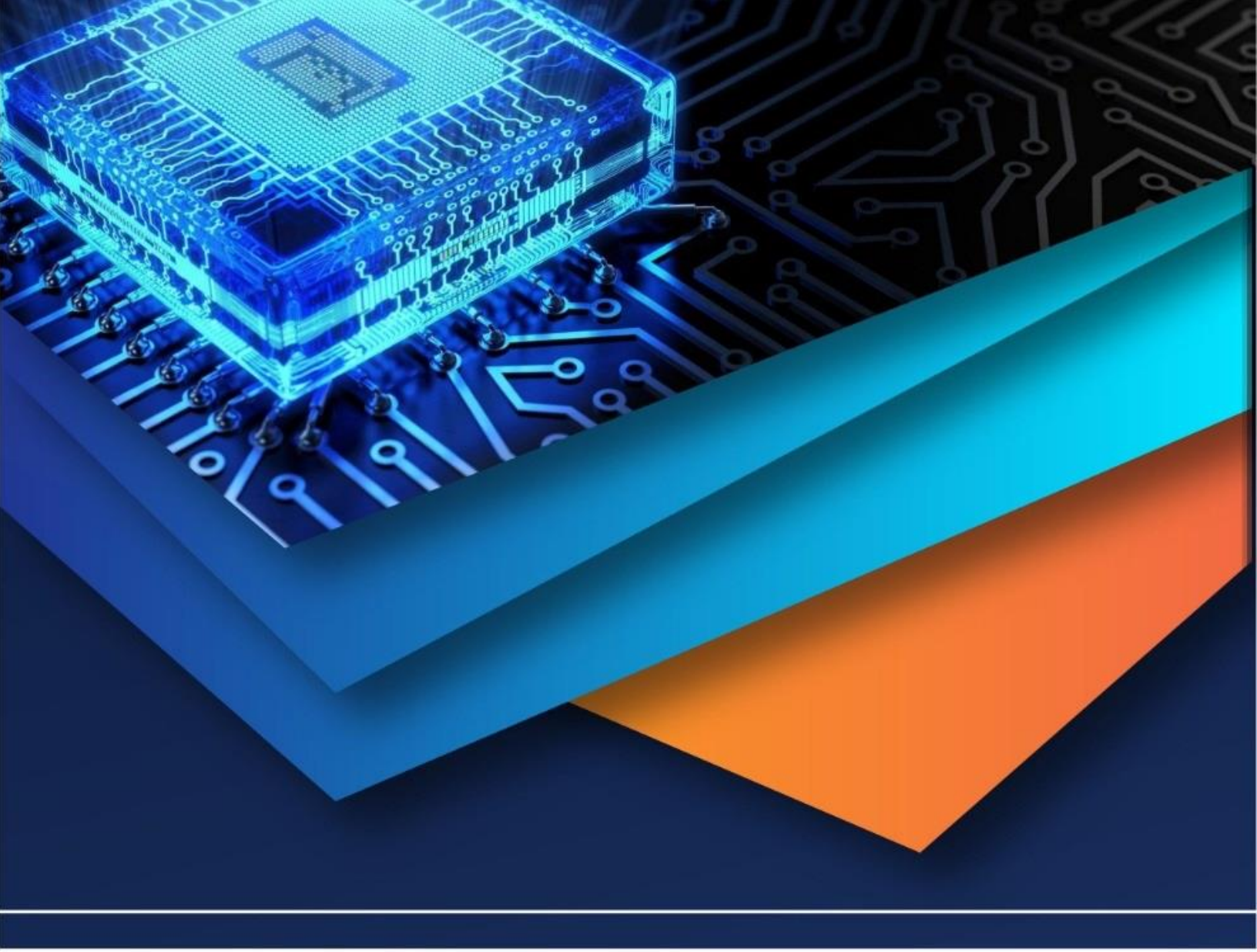
[9] L. Juwara, A. El-Hussuna, and K. El Emam, "An evaluation of synthetic data augmentation for mitigating covariate bias in health data," Patterns, vol. 5, 2024.

[10] B. Ding, C. Qin, R. Zhao, T. Luo, X. Li, G. Chen, W. Xia, J. Hu, A. T. Luu, and S. Joty, "Data augmentation using LLMs: Data perspectives, learning paradigms and challenges," Findings of ACL, 2024.

[11] K. G. A. Villanueva and P. Kumar, "XAI-driven skin disease classification: Leveraging GANs to augment ResNet-50 performance," arXiv preprint, 2025.

[12] "Data augmentation strategies for GAN-based medical image analysis: An empirical review," 2023.

[13] "Diffusion-based generative model for medical image augmentation," 2024.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY