iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: IV Month of publication: April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60150

www.ijraset.com

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

Classification of Diabetic Retinopathy using Deep Learning

Mr.V.Purna Chandra Reddy[1] Sarabu Vignya Sharma[2], M.Reshmi[3], J.Sahithi Reddy[4], D.Sai Priya[5]
Dept. of ECE, GNITS Hyderabad, INDIA

Abstract: Diabetic retinopathy, a potentially fatal retinal disease, occurs in individuals with diabetes, causing vision damage and possible visual impairment traditionally, screening for diabetic retinopathy has been the norm hard by ophthalmologists manually. To facilitate this, we have DR. Our pre-trained model is trained on a large dataset of about 3662 training images, enabling it to automatically detect the DR platform in high-resolution fundus images The dataset used for this is publicly available on kaggle. The DR phases are categorized as 0, 1, 2, 3, and 4. The fundus eye images serve as the input parameters of the model. The pre-trained model then identifies the point of interest in bank eye images, then provides activation function insights By calculating weights, for example patients' visual intensity levels, it helps classify diabetic retinopathy images into their intensity groups appropriately within the possibility of test- methods Provides variation.
Keywords: Convolutional neural network, Deep learning, pre- trained models.

I. INTRODUCTION

Diabetic retinopathy (DR) occurs as a result of structural changes in the micro vascular system of the retina, leading to visual deterioration and potential blindness affecting individuals of all ages worldwide. Early detection of DR is crucial, with research suggesting that 90% of cases can be successfully treated if identified promptly [1]. DR primarily affects various eye structures, including hemorrhages, exudates, and microaneurysms. Exudates, characterized by fluid leakage from retinal blood vessels, are a common early symptom and play a pivotal role in accurate DR screenings shown in figure1. Detecting these exudates not only aids in disease classification but also enhances automated screening methods. Analyzing these structural changes facilitates the identification of DR severity.

Diabetic macular edema (DME), stemming from fluid leakage in the macular region, further complicates the scenario. There exists a complex interplay between DR and DME, necessitating their joint assessment. The risk of DR-DME escalation is heightened when exudates encroach upon the macula [2]. However, conventional methods such as laser photocoagulation for identifying and treating severe DR-DME stages are time-consuming. Early detection of DR-DME, on the other hand, holds promise for effectively managing the conditions and preventing vision loss.

Diabetes mellitus, characterized by elevated blood glucose levels due to insulin deficiency, contributes to various complications, including retinopathy. Both type 1 and type 2 diabetes patients are susceptible to DR, wherein increased glucose levels adversely affect retinal blood vessels. Diabetic retinopathy (DR) presents in two different forms: non- proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR). NPDR is marked by vessel damage and fluid leakage, presenting clinical signs like microaneurysms, hemorrhages [3], and exudates. PDR, conversely, involves abnormal blood vessel growth on the retinal surface. These clinical manifestations are readily observable through fundus imaging, though their severity and progression rates may vary among patients. NPDR is further categorized into four grades, each exhibiting distinct symptoms and implications as in the table 1.

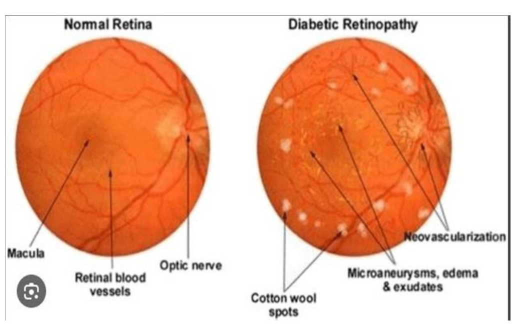


Figure-1: Difference between normal and DR retina.

[4]In the study "Detection of Diabetic Retinopathy and its Classification from Fundus Images" utilizes artificial intelligence for diabetic retinopathy (DR) detection. Employing a binary classification model, the methodology achieves an accuracy rate of 83.12% in identifying DR cases. This approach showcases the potential of AI-based models in accurately diagnosing DR from fundus images, offering a promising avenue for early detection and intervention. By leveraging advanced computational techniques, such as machine learning, the study contributes to improving the efficiency and accuracy of DR screening, ultimately enhancing patient outcomes and reducing the risk of vision-threatening complications associated with diabetes.

[5]The research paper entitled "Applying Supervised Contrastive Learning for Diabetic Retinopathy Detection and Severity Level Assessment from Fundus Images" is authored by Md Robiul Islam, Lway Faisal Abdulrazak, and Md.. Utilizing the "APTOS 2019 Blindness Detection" dataset, they employ supervised contrastive learning (SCL) to identify diabetic retinopathy (DR) and its severity stages from fundus images. Their approach achieves an 84.36% accuracy for multiclass classification, demonstrating SCL's effectiveness in accurate DR detection and severity level classification.

Table 1:severity levels of DR

Stage	Grade	Symptoms
Stage I	Normal	No apparent DR
Stage II	Mild NPDR	Microaneurysms only, Mild Non-Proliferative DR
Stage III	Moderate NPDR	Microaneurysms, **dot and blot hemorrhages**, few exudates
Stage IV	Severe NPDR	All the abnormalities in above two grades in large numbers
Stage V	PDR	Proliferative DR

II. LITERATURE REVIEW

Diabetic retinopathy (DR) is a widespread complication linked to diabetes and remains a major contributor to global blindness. Timely detection and precise diagnosis of DR are imperative for mitigating vision loss among diabetic individuals. Traditional DR diagnostic methods involve manual scrutiny of retinal images by skilled ophthalmologists, a process prone to human error and time-consuming nature.

The advent of deep learning has sparked interest in employing machine learning algorithms, notably Convolutional Neural Networks (CNNs), for automating DR diagnosis. CNN models, such as VGG16 and ResNet50[6], have showcased exceptional performance across diverse image classification tasks, including medical image analysis like retinal image classification.

Transfer learning, a methodology wherein knowledge acquired from training one model is transferred to another, has gained traction for fine-tuning pre-trained CNNs for specialized tasks like DR diagnosis. By leveraging insights gleaned from extensive image datasets like ImageNet, transfer learning empowers models to generalize better to novel tasks with limited training data, thereby enhancing their efficacy and performance.

Cohen's Kappa metric emerges as a statistical tool assessing agreement between two raters beyond random chance. In the context of DR diagnosis, Cohen's Kappa offers a holistic evaluation of model performance by considering not just prediction accuracy but also alignment with human expert judgments. This metric proves invaluable for evaluating the reliability and consistency of automated DR diagnosis systems. In essence, the paper delineates an integrated methodology amalgamating ensemble modelling, transfer learning, and robust performance evaluation utilizing Cohen's Kappa metric for precise and dependable diagnosis of diabetic retinopathy. By harnessing cutting-edge deep learning techniques, this proposed approach holds promise in advancing early detection and management of DR, thus diminishing the risk of vision impairment in diabetic patients.

III. METHODOLOGY

Machines are advancing towards interacting with the world much like humans through the utilization of computer vision. This technology leverages pattern recognition algorithms and vast visual data training to mimic the human brain's ability to perceive visual information. Primarily powered by deep learning, convolutional neural networks (CNNs)[7] have revolutionized the field of computer vision. By employing deep learning methodologies, efforts are underway to address the challenge of developing an automated ensemble model for classifying the severity of diabetic retinopathy.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Convolutional neural networks (CNNs)are deep learning models renowned for their efficacy in computer vision tasks. Unlike traditional neural networks where all layers are fully connected, CNNs typically only have the last layer fully connected. CNNs employ local receptive fields [8] to establish connections between discrete input neuron regions and neurons in hidden layers. These receptive fields traverse the entire image to generate feature maps. By sharing weights and biases, CNNs enable neurons in a feature map to detect related features, enhancing efficiency. In image classification tasks, CNNs often utilize the SoftMax function [9], with fully connected layers linking the final hidden layer to the output layer, while activation functions and pooling operations help in reducing dimensionality. This stands in contrast to artificial neural networks (ANNs) where every layer is fully connected, as depicted in Figure 2.

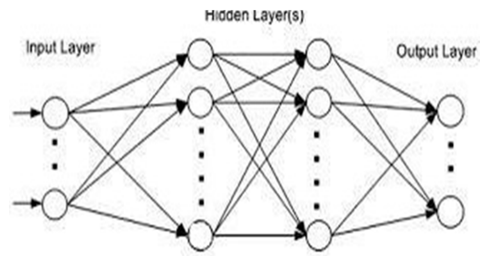


Figure 2: Artificial Neural Network (ANN) Structure.

Artificial neural networks (ANNs) are often deemed unsuitable for image processing tasks due to the potential risk of over fitting arising from the extensive dimensions of images. For instance, consider an image sized [32x32x3]. To input this image into an ANN, it must be flattened into a vector with 32x32x3 = 3072 elements. As a result, the primary layer of the ANN necessitates 3072 weights to handle this input vector, necessitating a more robust processor. Conversely, convolutional neural networks (CNNs) tackle this challenge by selectively connecting a small portion of input layer neurons to neurons within the hidden layer. These connections are localized and termed as local receptive fields, typically small areas like 3x3 regions of the source image, as illustrated in Figure 3. The output of these connections is represented in feature maps, as depicted in Figure 3. This design allows CNNs to effectively process images without the computational burden associated with ANNs.

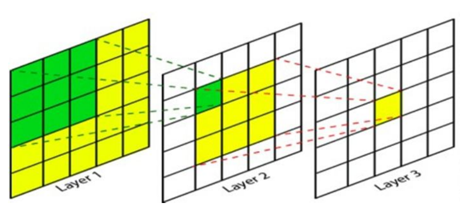


Figure 3: A receptive field of size 3x3 is linked to neurons in the convolutional layer.

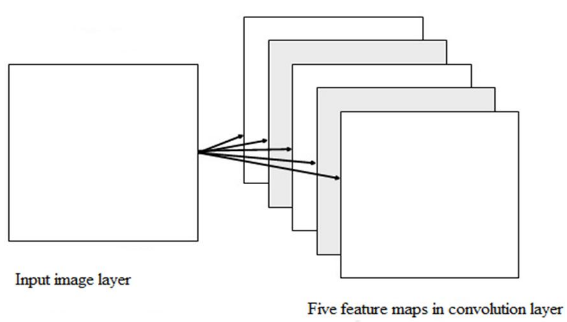


Figure 4: Feature maps generated by the convolutional layer.

Within a convolutional layer, the sliding window method is employed to traverse filters across the entire input image, computing the dot product of filter coefficients and input. Equation (1) outlines the dimensions of the output feature map (OUT), where OUT is equal to 1 plus the result of (N-F) divided by S. Subsequent to the convolutional layer, activation of feature maps occurs through functions such as ReLU, which activate neurons only when weights are equal to or greater than zero, effectively setting all negative weights to zero. The output range of the ReLU function spans from 0 to infinity.

$$OUT = 1 + \frac{N-F}{S} \tag{1}$$

When considering an image size of NxN and a filter size of FxF, convolution is executed utilizing the following Equation-2.

$$[N, N, N_C] * [F, F, N_f] = [[1 + \frac{N-F}{S}], [1 + \frac{N-F}{S}], N_f] \tag{2}$$

Following the ReLU transformation, a pooling stage is introduced to down sample the convolved features, reducing processing time and feature map dimensionality while preserving critical image characteristics. Among the common pooling methods are average pooling and maximum pooling. In maximum pooling, the maximum value within each small neighborhood of the feature map is chosen as the output, discarding other values. Conversely, average pooling computes the mean value within each neighborhood.

The present study adopts the maximum pooling technique. Pooling operations necessitate two parameters: the neighborhood size and the stride. The stride determines the interval between each extracted tile in pixels. Unlike convolution, where a stride of 2 implies overlapping tiles, in pooling, a stride of 2 indicates non- overlapping tiles are extracted from the feature map. Figure 5 depicts both pooling methods, namely max pooling and average pooling.

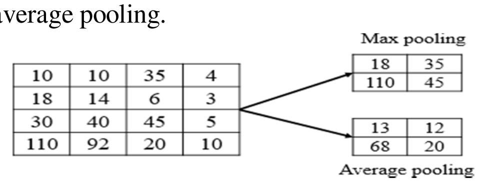


Figure 5: Maximum Pooling and Mean Pooling.

Following the final fully connected layer, the model incorporates the SoftMax function. This activation function produces a vector that reflects the probability distribution across a set of predetermined image classes. A SoftMax layer employs the SoftMax function for this purpose.

The classification layer [10] at the bottom of the model generates the cross-entropy loss for classes which are mutually exclusive. In the classification layer, the model employs the Categorical Cross-Entropy [10] for a 1-of-K coding scheme and the values from the softmax function to categorise each input as belonging to one of the K mutually exclusive classes.

A subset of neural networks, Convolutional Neural Networks (CNNs) exploit the spatial arrangement of input data. These models typically follow a conventional layout, featuring a sequence of convolutional layers interspersed with pooling layers as shown in Figure6.

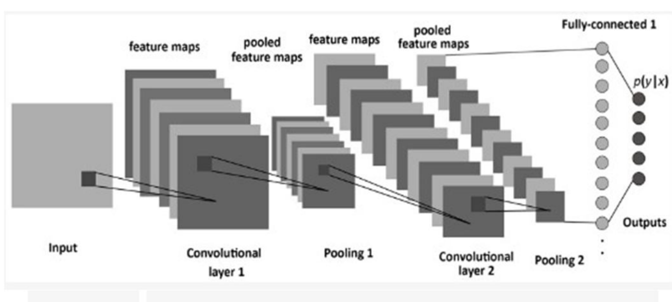


Figure 6: the architecture of a Cnn, Comprising Convolutinal, Pooling, and Fully-Connected layers.

In the proposed workflow, input images undergo pre- processing and data augmentation stages to enhance the quality and diversity of the dataset, respectively. These steps are crucial for preparing the data for subsequent analysis as shown in the figure 7. The pre-processed and augmented data are then fed into pre-trained Convolutional Neural Network (CNN) models, which have been previously trained on large-scale datasets like ImageNet, allowing them to capture intricate features relevant to diabetic retinopathy (DR). The CNN models categorize retinal images into five classes denoting various stages of DR severity: No DR, Mild, Moderate, Severe, and Proliferative DR (PDR). Evaluation of classification outcomes employs two primary metrics: accuracy and Cohen's Kappa.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Accuracy provides a straightforward measure of the proportion of correctly classified instances among all instances evaluated, while Cohen's Kappa offers a more robust assessment by considering the agreement between the model's predictions and human expert judgments, accounting for chance agreement. By employing both metrics, the effectiveness and reliability of the ensemble learning approach utilizing pre-trained CNN models for DR classification can be comprehensively evaluated, ensuring accurate and clinically relevant predictions.
The paper explores ensemble models using pre-trained CNNs like Vgg16 and ResNet50 for diabetic retinopathy severity detection. Transfer learning fine-tunes these models, enhancing accuracy. Cohen's Kappa metric is utilized for robust performance evaluation, providing nuanced insights beyond simple accuracy measures. This integrated approach showcases cutting-edge techniques in deep learning for accurate and reliable diagnosis.

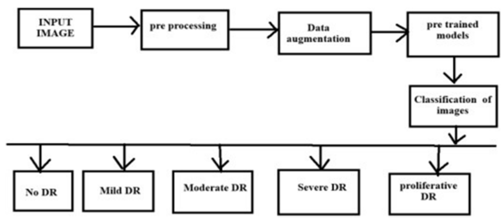


FIGURE 7: Ensemble Learning Utilizing Pre-Trained CNN Models"

The weighted kappa score, a metric used in competitions for evaluation purposes, measures the level of agreement among raters when dealing with categorical data. Unlike simple accuracy, it takes into account the possibility of chance agreement, thus providing a more robust assessment.
Cohen's kappa metric, particularly the Quadratic Weighted Cohen's kappa score, extends this concept by allowing for different weights to be assigned to disagreements, which is particularly useful when dealing with ordered scores. This method involves three matrices: the observed scores, the expected scores, and a weight matrix as shown below.

$$\kappa = \frac{p_0 - p_e}{1 - p_e},$$

Where, Po is observed agreement and Pe is chance of disagreement observed between raters
So, why use a weighted kappa metric instead of other commonly used metrics like confusion matrices or accuracy? Firstly, it's beneficial when the target labels follow a specific order, such as [0,1,2,3,4]. Additionally, it assigns penalties based on the distance between predicted and observed elements in cases of disagreement. This property is advantageous, especially in medical contexts, where misclassifications can have significant consequences. For instance, if a patient is classified as class-0 when they belong to class-1, it could lead to serious implications. Hence, the weighted kappa score provides a nuanced evaluation, prioritizing correct predictions while penalizing misclassifications accordingly.
In this paper, Algorithm used for implementing transfer learning with VGG16:

1) The DR severity condition is classified into five categories: NO DR, Mild, Moderate, Severe, and Proliferative.
2) Exploration data analysis and preprocessing involve utilizing crop functions to eliminate extra dark areas around the images.
3) Analyze the number of training and testing images in the dataset.
4) Load the pre-trained VGG16 model initialized with weights obtained from the ImageNet dataset.
5) Add a Global Average Pooling layer to the output of the 16-layer VGG16 model.
6) Apply a Dropout layer with a dropout rate of on these results.
7) Add a Dense layer with 5 output units and softmax activation.
8) Create a new model (model1) using the modified VGG16 architecture.
9) Compile the model with the Adam optimizer, categorical cross entropy loss, and metrics, including accuracy and a custom Kappa Metric.
10) Utilizing the VGG16 model, achieve an accuracy of 0.817 and a Cohen Kappa metric of 0.91. Refine and enhance based on these results.
11) For same method or the procedure, by utilizing the Resnet model, achieve an accuracy of 0.844 and a Cohen Kappa metric of 0.95.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

IV. RESULTS & DISCUSSIONS

An exploratory data analysis (EDA) is performed on the dataset to gain insights and understand its characteristics from APTOS2019, comprising 3662 training and 1928 test images, encompassing all 5 diabetic retinopathy severity levels as shown in the Figure8.With a relatively small training set, prone to over fitting, pre-processing becomes crucial for performance improvement. Data augmentation is employed to expand the dataset. Prior to augmentation, image conditions are assessed and addressed, including adjustments for darkness, extra black backgrounds, and varying sizes, ensuring uniformity via smoothing techniques, cropping, and resizing.

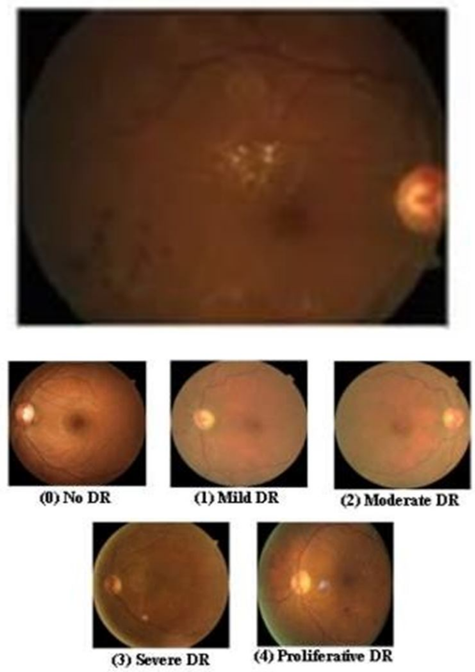
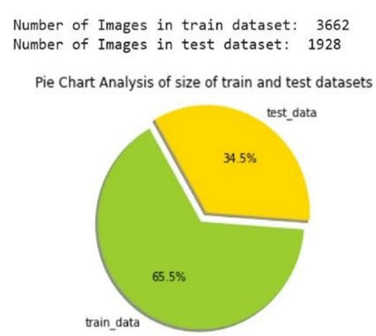
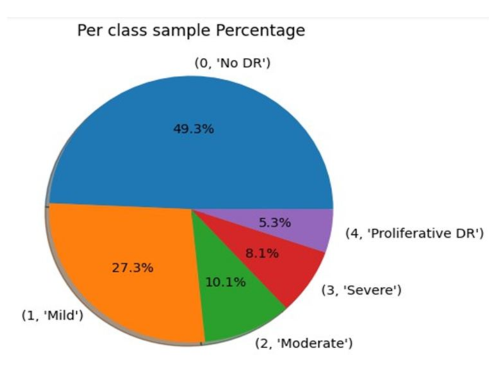


Figure 8: Examples from the dataset





ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

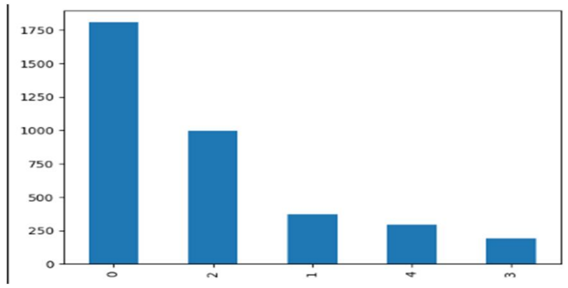


Figure 9: Visual depiction of the dataset.

Transfer learning is a pivotal technique harnessed through pre-trained models, specifically VGG16 and ResNet50, renowned for their efficacy in image classification tasks. The performance metrics, including accuracy values and Cohen's Kappa metric measures, are meticulously documented and presented in Table 2. Leveraging transfer learning allows us to capitalize on the learned features from these pre-trained models, enhancing the efficiency of our classification task. However, to further optimize performance, we introduce an ensemble learning approach. By amalgamating the predictions from multiple models through voting schemes, our proposed ensemble model achieves superior results, particularly evident when applied to previously unseen data. This strategy not only boosts accuracy but also enhances the robustness of the classification, ensuring more reliable predictions for diabetic retinopathy severity assessment.

Table-2: Results on pre-trained models

MODEL	CCURACY SCORE	COHEN'S KAPPA SCORE
VGG16	0.817	0.91
RESNET50	0.844	0.95

V. CONCLUSION

The integration of deep learning in diabetic retinopathy detection shows promise for early and accurate diagnosis. Challenges include model interpretability, bias in datasets, and ethical data use. Validation, regulatory approvals, and model performance comparison underscore the need for rigorous testing. ResNet50 outperforms VGG16, showing higher Cohen Kappa scores, indicating better classification agreement.

REFERENCES

[1] M. T. Hagos and S. Kant, "Transfer learning-based detection of DR from small dataset," arXiv preprint arXiv:1905.07203, 2019.

[2] D.S.W. Ting, L. Carin, M.D. Abramoff, Observations and lessons learned from the artificial intelligence studies for DR screening, JAMA Ophthalmol. 06 (2019).

[3] M.J. Van Grinsven, B. van Ginneken, C.B. Hoyng, T. Theelen, C.I. S´ anchez, Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images, IEEE Trans. Med. Imaging 35 (5) (2016) 1273–1284.

[4] Shelar, M., Gaitonde, S., Senthilkumar, A., Mundra, M., & Sarang, A. (2021). Detection of Diabetic Retinopathy and its Classification from the Fundus Images. In 2021 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-5). IEEE. DOI:10.1109/ICCCI50826.2021.9402347

[5] Islam, M. R., Abdulrazak, L. F., Nahiduzzaman, M., Goni, M. O. F., Anower, M. S., Ahsan, M., Haider, J., & Kowalski, M. (Year). Applying supervised contrastive learning for the detection of diabetic retinopathy and its severity levels from fundus images.

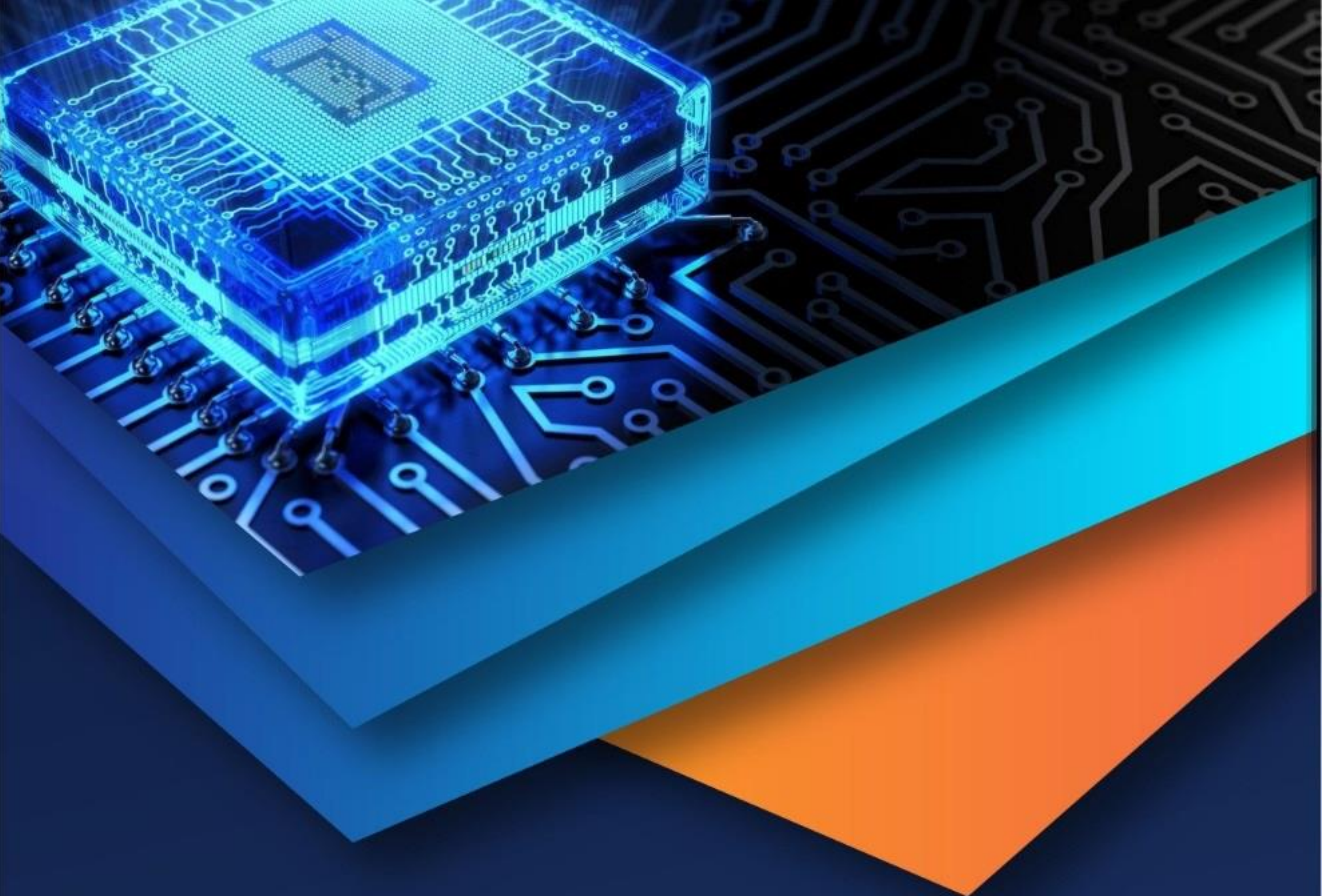
[6] AATILA, M., LACHGAR, M., HRIMECH, H., & KARTIT, A. (Year). Diabetic Retinopathy Classification Using ResNet50 and VGG-16 Pretrained Networks. LTI Laboratory, ENSA, Chouaib Doukkali University, El Jadida 1166, Morocco; University Hassan 1st, ENSA of Berrechid, LAMSAD, B.P 218, Morocco.

[7] M. Jogin, Mohana, M. S. Madhulika, G. D. Divya, R. K. Meghana and S. Apoorva, "Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning," 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2018, pp. 2319-2323, doi: 10.1109/RTEICT42901.2018.9012507.

[8] https://towardsdatascience.com/understand-local-receptive- fields-in-convolutional-neural-networks-f26d700be16c

[9] Alzubaidi, L., Zhang, J., Humaidi, A.J. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8, 53 (2021)

[10] A. Ajit, K. Acharya and A. Samanta, "A Review of Convolutional Neural Networks," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-5, doi: 10.1109/ic-ETITE47903.2020.049.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)