International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 05 | May 2025 www.irjet.net p-ISSN: 2395-0072

CNN-based Architecture with Attention Mechanisms for Enhanced Diabetic Retinopathy Detection and Classification

Meenal Katole[1], Prof. Pramila M Chawan[2]

[1]MTech student, Dept of Computer Engineering and IT, VJTI college Mumbai, Maharashtra, India
[2]Associate Professor, Dept of Computer Engineering and IT, VJTI college Mumbai, Maharashtra, India

---***---

Abstract - *Diabetic retinopathy is one of the leading causes of visual loss, and early detection is critical to effective treatment. This research discusses in detail the current deep learning practices, particularly convolutional neural networks and attention mechanisms that are applied to improve the detection and grading of DR. The study considers different algorithms and technologies employed in this area, assesses their performance, and examines the way attention mechanisms increase CNN performance. We conclude by highlighting key developments and identifying future research directions.*

Key Words: Diabetic Retinopathy, Convolutional Neural Networks (CNN), Attention Mechanisms, Medical Image Classification, Retinal Images.

1.INTRODUCTION

1.1 Diabetic Retinopathy

Diabetic Retinopathy is a severe complication of diabetes, characterized by damage to blood vessels of the retina, progressive visual impairment, and potentially leading to blindness. Early detection of this complication is very important to prevent the DR from progressing to an advanced stage, such as NPDR and PDR. Traditionally, DR detection has been performed manually by analyzing the retinal images. It is a very time-consuming process and prone to human errors; therefore, it requires an automated detection system.

1.2 Convolutional Neural Networks (CNNs)

Convolutional neural networks have become powerful tools for image classification, including medical imaging of diabetic retinopathy. They inherently learn the spatial hierarchies of features from images without requiring manual feature extraction. Thus, they are quite efficient in the diagnosis of DR from fundus images. However, CNNs sometimes fail to capture the subtle yet critical features of fundus images, which are very vital in the diagnosis of DR in its early stages limitation to their effectiveness.

1.3 Attention Mechanisms

To such CNNs, attention mechanisms are utilized to guide the network on where in the image it should focus its attention. Spatial attention guides the model on significant regions of attention, usually with hemorrhages and exudates, whereas channel attention models support the model to emphasize feature maps relevant to further processes. Integrating an attention mechanism will improve the focus of the important areas of the CNN models, classifying with high precision, and better interpreting the results.

1.4 Algorithms for DR Detection

Various approaches have, therefore, been developed for DR detection:

These techniques depend on the features extracted manually and then perform the classification, typically using algorithms such as support vector machines or random forest.

• CNN-based methods: DR detection performance has improved significantly due to architectures such as ResNet and Inception, which learn features directly from retinal images using CNNs.

• Hybrid approaches: Some techniques involve multi-model architectures, such as combining CNNs with other models for data augmentation or attention mechanisms to enhance feature focus.purposes.

2. LITERATURE REVIEW

[1] A CNN model that was trained on the Kaggle dataset was presented by Kokane et al. (2021), who were able to obtain 74.8% accuracy for DR detection. The authors hypothesized that by enabling the model to focus more on the salient characteristics of lesions, the attention processes may improve performance.

[2] An enhanced ResNet architecture optimization for DR classification was presented by Jiao and Tao (2023). Their network's ability to capture intricate retinal features was further enhanced by their hierarchical residual-like connections, which led to an improvement in performance compared to conventional CNNs.

[3] To capture the subtle changes in retinal lesions, Pan and Yang (2023) created an encoder-decoder model with a dual branch structure that uses LGE and attention mechanisms. For applied, it demonstrated notable improvements in accuracy for identifying early-stage DR using the Messidor dataset.

[4] To overcome the problem of data scarcity in DR detection, Wietse ten Dam et al. (2023) suggested a method using Generative Adversarial Networks (GANs) to generate synthetic retinal pictures. Even though CNNs were not directly evaluated in this work, it is anticipated that using the generated data for model training may enhance CNN performance.

[5] Hybrid CNN Model: To enhance feature extraction for the categorization of DR in 2022, this model merged the ResNet and DenseNet architectures. With the Messidor dataset, the hybrid model achieved a high accuracy of 96.22% demonstrating the benefits of combining several network topologies.

[6] The study by Muhammad Mohsin Butt and colleagues (2022) focuses on improving the automatic detection of Diabetic Retinopathy (DR) using a hybrid deep learning approach. In their review of past work, the authors discuss two main categories of DR detection methods: traditional machine learning techniques and modern deep learning models.

[7] For autonomous DR classification, the Hybrid CNN Approach (2020) integrated CNN layers with conventional machine learning classifiers. This model shows the potential of hybrid techniques by achieving good performance on the APTOS and Messidor datasets.

[8] GANs were used in 2023 Deep Learning Models for DR Detection in conjunction with CNNs to create artificial retinal images. This method seeks to improve CNNs capacity to identify DR, particularly in light of the problem of the scarcity of medical data.

[9] P. K. Darabi focuses on the early diagnosis of diabetic retinopathy (DR) to prevent vision loss. Reviews traditional diagnostic tools such as fundus photography and OCT. The study emphasizes the growing role of AI, especially deep learning models, in automating DR detection. Highlights the benefits of these tools in improving accessibility and accuracy of screening.

[10] The 2024 study by Meenal Katole and Prof. Pramila M. Chawan explores using transfer learning with ensemble learning for early detection of diabetic retinopathy. It utilizes pre-trained CNNs to extract features from retinal images, enhancing classification accuracy. The approach addresses limited data challenges and aims to support early diagnosis. unavoidable.

3. PROPOSED SYSTEM

3.1 Problem Statement

To develop a Convolutional Neural Network (CNN) and Attention Mechanisms architecture to improve the detection and classification of Diabetic Retinopathy in retinal images.

The ultimate goal will be to develop an architecture that will integrate CNNs with attention mechanisms, yielding better detection and classification of diabetic retinopathy on retinal images. CNN-based Architecture with Attention Mechanisms for Enhanced Diabetic Retinopathy Detection and Classification system.

3.2 Problem Elaboration

Diabetic retinopathy is one of the most common complications associated with diabetes, affecting millions globally. The condition progresses through several stages, from mild non-proliferative to severe proliferative DR, and can eventually lead to blindness. Early detection of DR can prevent vision loss, but manual detection by ophthalmologists is time consuming and subjective. Automated systems using deep learning have shown promise in detecting DR from retinal images; however, standard CNN architectures cannot consistently focus on important areas within an image, leading to suboptimal detection accuracy, especially in early-stage DR. In this project, we propose using CNNs enhanced with attention mechanisms, which can guide the model to focus on regions that are most relevant for DR detection, improving sensitivity to small lesions and enhancing overall classification accuracy.

3.3 Dataset Description

The Kaggle Diabetic Retinopathy Detection Dataset is used for this research. This dataset contains retinal fundus images labeled according to the severity of diabetic retinopathy. The images are categorized into two classes:

Class 0: Diabetic Retinopathy

Class 1: No Diabetic Retinopathy

The dataset consists of :- Training Set : 35,126 images

Each image in the dataset is captured under various conditions, including varying image resolution and lighting. Preprocessing techniques such as resizing, normalization, and data augmentation (e.g., rotation, flipping, and scaling) are applied to enhance the quality and diversity of the training data.

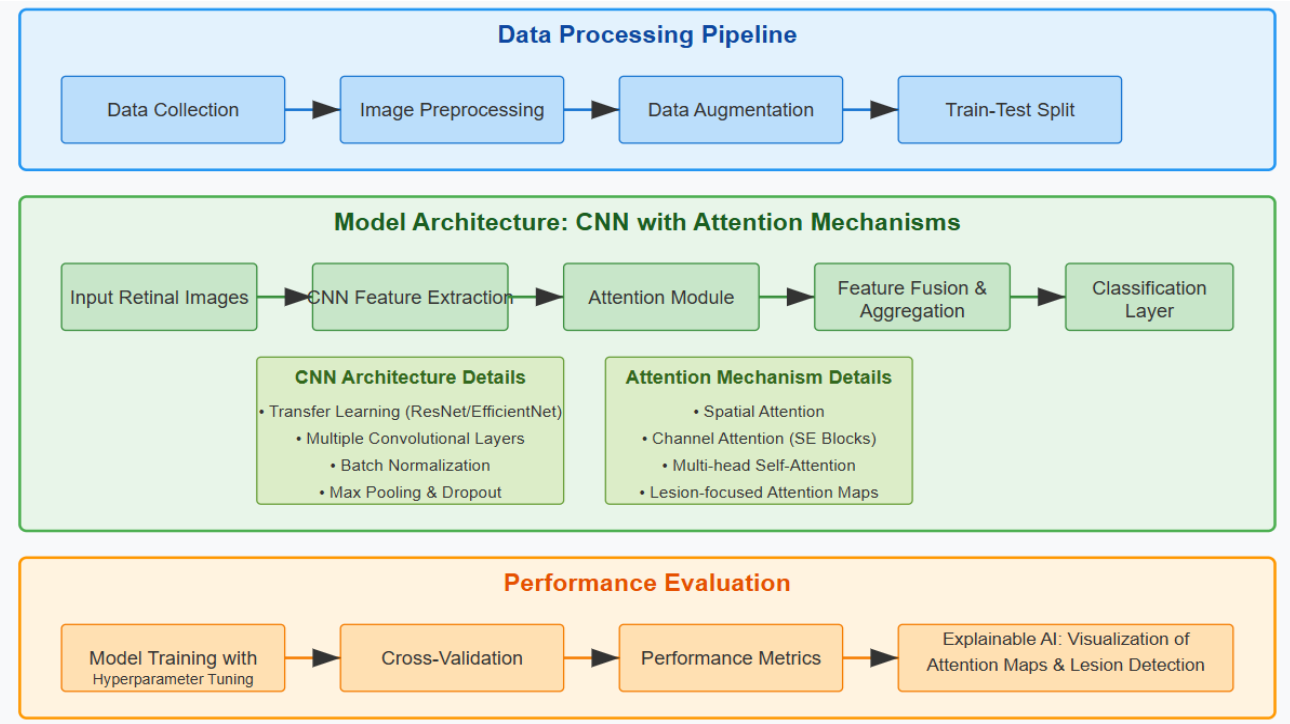

Fig -1: DR System Architecture

3.4 Data Preprocessing

The dataset comprises retinal images from individuals across diverse age groups, captured under various lighting conditions, leading to inconsistent pixel intensity distributions. To mitigate these variations, color normalization techniques are applied to standardize image appearance. Given the high resolution of the original images and their substantial memory requirements, the RGB images are first transformed into grayscale format to reduce computational load. The images are then uniformly resized to 224x224 pixels, which helps retain the relevant retinal features while optimizing memory usage and training efficiency. As the architecture is based on a Convolutional Neural Network, extensive preprocessing is not necessary beyond these essential steps.

3.5 Model Description

To effectively detect and classify diabetic retinopathy (DR) from retinal fundus images, we propose a deep learning architecture that integrates Convolutional Neural Networks (CNNs) with attention mechanisms. This hybrid approach enables the model to not only learn hierarchical features from the images but also focus on the most diagnostically relevant regions, such as microaneurysms, hemorrhages, and exudates.

The CNN backbone is responsible for feature extraction. It consists of multiple convolutional layers with ReLU activation, interleaved with max pooling layers to reduce spatial dimensions while retaining key features. These layers capture both low-level textures and high-level patterns critical for distinguishing between DR severity levels.

To enhance interpretability and performance, an attention module is introduced after the convolutional layers. This mechanism dynamically weighs feature maps, allowing the model to attend to regions most likely to contribute to the final classification. This mimics a clinician's behavior of focusing on abnormal retinal zones during diagnosis.

The network concludes with fully connected layers followed by a Softmax classifier, predicting one of the two standard DR classes (Diabetic Retinopathy, No Diabetic Retinopathy). Dropout layers are used to prevent overfitting, and the model is trained using cross-entropy loss and an adaptive optimizer such as Adam.

The use of attention mechanisms alongside CNN not only boosts classification accuracy but also helps localize critical lesions, making the model more robust and clinically relevant. Key Architectural Parameters:

Input size: 224x224 pixels (preprocessed grayscale image)
Convolutional filters: 32, 64, 128 (increasing with depth)
Kernel size: 3x3
Pooling: Max pooling (2x2)
Fully connected layers: 1–2 layers with ReLU
Output classes: 2
Activation function (final layer): Softmax
Batch size: 64
Optimizer: Adam
Attention mechanism: Channel and/or spatial attention (ex. CBAM or SE block).

4. IMPLEMENTATION

The proposed system is implemented on a Windows 11 machine equipped with an Intel® Core™ i7 processor and NVIDIA RTX 3060 GPU to facilitate efficient training of deep learning models. Python is used as the programming language, and the Jupyter Notebook environment is chosen for ease of visualization and iterative development.

```
# Define Architecture For Retinopathy Model
class CNN_Retino(nn.Module):

    def __init__(self, params):

        super(CNN_Retino, self).__init__()

        Cin,Hin,Win = params["shape_in"]
        init_f = params["initial_filters"]
        num_fc1 = params["num_fc1"]
        num_classes = params["num_classes"]
        self.dropout_rate = params["dropout_rate"]

        # CNN Layers
        self.conv1 = nn.Conv2d(Cin, init_f, kernel_size=3)
        h,w=findConv2dOutShape(Hin,Win,self.conv1)
        self.conv2 = nn.Conv2d(init_f, 2*init_f, kernel_size=3)
        h,w=findConv2dOutShape(h,w,self.conv2)
        self.conv3 = nn.Conv2d(2*init_f, 4*init_f, kernel_size=3)
        h,w=findConv2dOutShape(h,w,self.conv3)
        self.conv4 = nn.Conv2d(4*init_f, 8*init_f, kernel_size=3)
        h,w=findConv2dOutShape(h,w,self.conv4)
```

```
    # compute the flatten size
    self.num_flatten=h*w*8*init_f
    self.fc1 = nn.Linear(self.num_flatten, num_fc1)
    self.fc2 = nn.Linear(num_fc1, num_classes)

def forward(self,X):

    X = F.relu(self.conv1(X));
    X = F.max_pool2d(X, 2, 2)
    X = F.relu(self.conv2(X))
    X = F.max_pool2d(X, 2, 2)
    X = F.relu(self.conv3(X))
    X = F.max_pool2d(X, 2, 2)
    X = F.relu(self.conv4(X))
    X = F.max_pool2d(X, 2, 2)
    X = X.view(-1, self.num_flatten)
    X = F.relu(self.fc1(X))
    X = F.dropout(X, self.dropout_rate)
    X = self.fc2(X)
    return F.log_softmax(X, dim=1)
```

Fig -2: Defined Architecture for Retinopathy Model

The following Python libraries are employed:

NumPy: For numerical computations and image data manipulation.
Pandas: To manage and preprocess metadata and label information.
PIL(Python Imaging Library): For image reading, resizing, normalization, and augmentation.
TensorFlow and Keras: Initially used for model prototyping and experimentation.
PyTorch: Used for final implementation of the CNN with attention mechanism due to its dynamic computation graph and ease of customization.

5. RESULTS

The performance of the proposed deep learning model, built using PyTorch, was evaluated on a dataset of retinal fundus images labeled for diabetic retinopathy severity. The architecture combined convolutional layers with an attention mechanism to improve the network's focus on disease-relevant features.

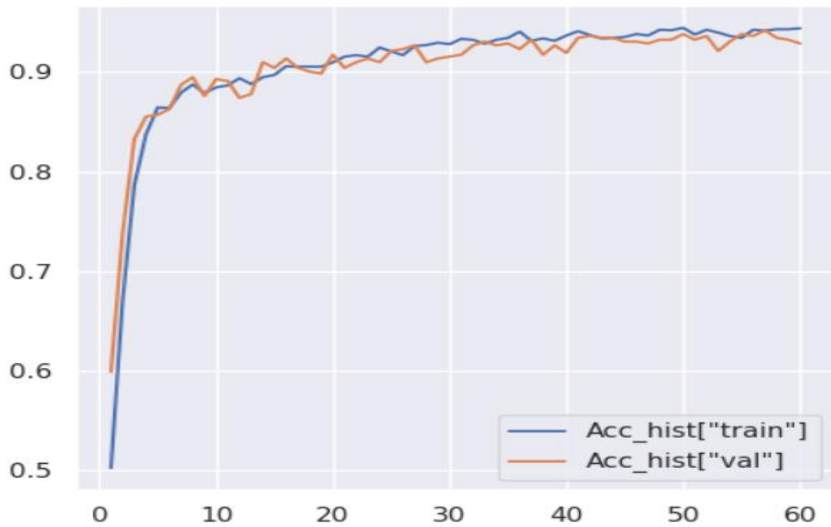

Chart -1: Accuracy of Model

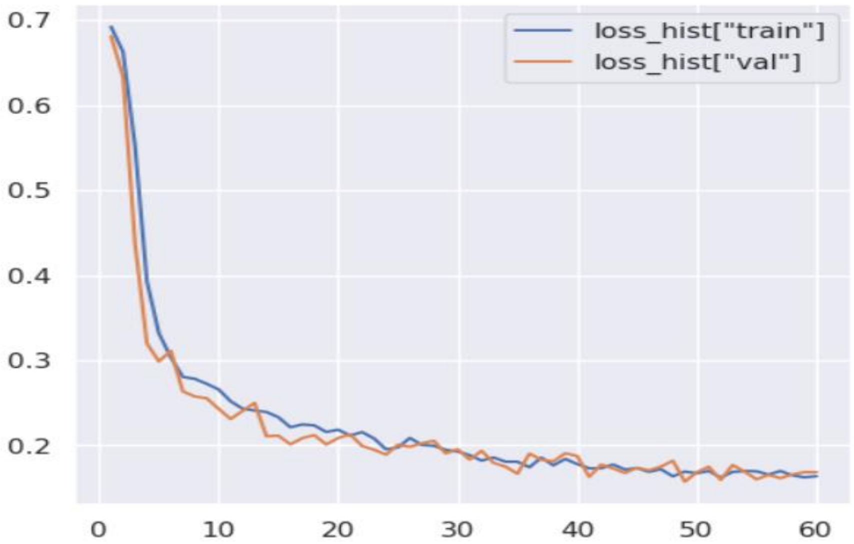

Chart -2: Loss of Model

During training, data augmentation techniques such as rotation, flipping, and brightness adjustment were applied to improve the model's generalization and reduce overfitting. An accuracy of 94% is achieved.

	precision	recall	f1-score	support
0.0	0.96	0.92	0.94	113
1.0	0.93	0.97	0.95	118
accuracy			0.94	231
macro avg	0.94	0.94	0.94	231
weighted avg	0.94	0.94	0.94	231

Fig -3: Classification Report for Retinopathy Classification Model based on Test Set

6. CONCLUSION

The proposed system utilizes the strengths of CNNs and attention mechanisms to enhance diabetic retinopathy detection and classification. Guiding the model toward focusing on the critical areas in retinal images ensures that the architecture overcomes the limitation of traditional CNNs in the detection of early DR. Increased ability by the model for distinction between the stages of DR will lead to a timelier diagnosis with greater accuracy, enabling scalability for clinical use. Furthermore, the inclusion of attention mechanisms will make the model interpretable and not only accurate but also focus on those features relevant to the medical professional analyzing them by hand.

REFERENCES

[1] Aryan Kokane, Gourhari Sharma, Akash Raina, Shubham Narole, and Prof. P M Chawan, "Detection of Diabetic Retinopathy using Neural Networks," International Research Journal of Engineering and Technology (IRJET), Vol 8, Issue 3, 2021.

[2] Jiao X, Tao*, "Classification of Diabetic Retinopathy based on Improved ResNet Model," In Proceedings of the ACM Conference on Information Technology, 2023.

[3] Anning Pan, Jingzong Yang, "Research on Deep Learning-based Detection of changes in Diabetic Retinopathy Lesions," In Proceedings of the ACM Conference on Medical Imaging Applications, 2023.

[4] Wietse ten Dam, Meike Grol, Zelda Zeegers, Alireza Dehghani, and Huib Aldewereld, "Representative Data Generation of Diabetic Retinopathy Synthetic Retinal Images," In Human-Centered AI Education and Practice Conference 2023 (HCAIep '23), ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3633083.3633175

[5] Yasashvini R., Vergin Raja Sarobin M., Rukmani Panjanathan, Graceline Jasmine S. and Jani Anbarasi L, "Diabetic Retinopathy Classification using CNN and Hybrid Deep Convolutional Neural Networks," Symmetry, MDPI, 2022, https://doi.org/10.3390/sym14091932

[6] Muhammad Mohsin Butt , D N F Awang Iskandar , Sherif E Abdelhamid , Ghazanfar Latif , Runna Alghazo, "Diabetic Retinopathy Detection from Fundus Images of the Eye Using Hybrid Deep Learning Features," MDPI, National Library of Medicine, 2022.

[7] Wietse ten Dam, Meike Grol, Zelda Zeegers, Alireza Dehghani, and Huib Aldewereld, "Deep Learning Models for Diabetic Retinopathy Detection Using GANs and CNNs," ACM Transactions on Medical Applications, 2023.

[8] Ali Ghulam, Aqsa Dastgir, Muhammad Waseem Iqbal, Muhammad Anwar, "A Hybrid Convolutional Neural Network Model for Automatic Diabetic Retinopathy Classification from Fundus Images," IEEE Journal of Translational Engineering in Health and Medicine, 2023.

[9] P. K. Darabi, "Diagnosis of Diabetic Retinopathy," ResearchGate, 2024, [Online], Available: https://doi.org/10.13140/RG.2.2.13037.19688

[10] Meenal Katole, Prof. Pramila M Chawan, "Transfer Learning model with Ensemble Learning to detect Diabetic Retinopathy from retinal images, enhancing early diagnosis: A Survey," International Research Journal of Engineering and Technology (IRJET), Vol 11, Issue 11, 2024.

BIOGRAPHIES

Meenal Katole is presently pursuing her MTech in Software Engineering from Veermata Jijabai Technological Institute, Mumbai. She has done her B.E in Computer Science and Engineering.

Prof. Pramila M. Chawan is an Associate Professor at Veermata Jijabai Technological Institute (VJTI), Mumbai. She has done her B.E. (Computer Engineering) & M.E. (Computer Engineering) from VJTI College, Mumbai, with over 31 years of teaching experience. Prof. Chawan has published an outstanding amount of 191 papers across international journals, conferences, and symposiums. She has been actively involved in guiding students in their research, with over 100 M.E./MTech. projects and 150 B.E./B.Tech. projects under her supervision. She has also worked on the editorial boards of many international scientific journals. She has been awarded with the 'Innovative and Dedicated Educationalist Award Specialization: Computer Engineering and I.T.' by The Society of Innovative Educationalist and Scientific Research Professional, Chennai (SIESRP).