IFERP
ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Volume 11 Issue 5 May 2024

GAN-Based Approach to Image Generation and Recognition: Diabetic retinopathy using Deep Learning

[1] Sanjiv Sarvaesh S, [2] Gautham P Pillai, [3] Dr. Usha Rani
[1] [2] CINTEL, SRM Institute of Sci&Tech
[3] Assistant Professor, CINTEL, SRM Institute of Sci&Tech
Corresponding Author Email: [1] ss7345@srmist.edu.in, [2] gp2822@srmist.edu.in, [3] usharanr2@srmist.edu.in

***Abstract*— This groundbreaking study introduces a novel approach to automated screening and diagnosis of Diabetic Retinopathy (DR) using a sophisticated Generative Adversarial Network (GAN)-based methodology. Using the IDRiD dataset, which offers annotated retinal images, our research aims to develop a robust deep learning model capable of accurately detecting DR severity levels. By employing advanced techniques like SRGAN for image quality enhancement and custom deep learning architectures, we seek to establish a new benchmark for DR screening accuracy. Emphasizing technological innovation in healthcare, this project underscores ethical considerations and identifies avenues for future enhancements in AI-driven diagnostics.**

***Index Terms*— GAN, Diabetic Retinopathy, ophthalmology, SRGAN.**

I. INTRODUCTION

Diabetic Retinopathy (DR) stands as a significant global health issue, posing a primary cause of vision loss and blindness among individuals afflicted with diabetes mentioned in one of the recent papers of Awais Bajwa [7]. With diabetes prevalence on the rise worldwide, the imperative to develop effective, precise, and scalable screening and diagnostic methods for DR becomes increasingly urgent. Conventional diagnostic approaches, heavily reliant on ophthalmologists' expertise in interpreting retinal images, encounter obstacles such as limited access to specialized healthcare services, variable diagnostic accuracy, and inherent limitations in scalability and speed. These challenges underscore the need for innovative strategies to enhance the efficiency and accuracy of DR diagnosis.

In recent years, as mentioned by Suwarna [1], deep learning technologies have emerged as frontrunners in medical image analysis, exhibiting notable success in improving diagnostic processes across diverse domains, including ophthalmology. Specifically, Super-Resolution Generative Adversarial Networks (SRGANs), a subset of artificial intelligence algorithms specialized in super-resolution image enhancement, have displayed promising capabilities in enhancing the quality and resolution of retinal images. This enhancement not only addresses the issues of image clarity and detail but also provides a fresh approach to improving the quality of datasets used for machine learning models.

The research introduces an innovative framework using SRGAN for enhancing retinal images depicting various stages of DR, which can also be approached by using ensemble learning [3] as stated by one of the authors Issra. These enhanced images are leveraged to train deep learning models for DR detection and classification. This approach aims to overcome the limitations posed by the shortage of high-quality annotated medical images, a significant barrier to developing high-performing automated diagnostic systems. By improving the quality of retinal images, we can enhance the dataset available for training, thereby improving the model's capacity to discern subtle features indicative of disease progression.

Moreover, integrating SRGAN-enhanced high-resolution images into the training process tackles another critical challenge in medical image analysis: model overfitting to limited, often homogeneous datasets. By introducing high-quality, enhanced images that capture the complexity and variability of real-world cases, the methodology promises to enhance the generalizability and robustness of deep learning models tasked with DR detection.

This paper outlines the development and validation of the SRGAN-based image enhancement framework, followed by training deep learning models on enriched datasets and conducting a comprehensive evaluation of the impact on diagnostic performance according to Ian J. Goodfellow [4]. Through meticulous comparison with conventional training methodologies and datasets, we demonstrate the superiority of the approach in enhancing the accuracy, sensitivity, and specificity of DR diagnosis. The findings not only advance automated DR screening technologies significantly but also illustrate the transformative potential of SRGANs in medical image enhancement on a broader scale.

By bridging the gap between technological innovation and clinical application, the research aims to facilitate the development of more accessible, accurate, and efficient diagnostic solutions for Diabetic Retinopathy, it can also be approached by EfficientNet[11]. Ultimately, this contributes to broader goals of improving patient outcomes and advancing public health initiatives in combating diabetes-induced blindness.

II. MATH

Precision:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Recall (Sensitivity):

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

mAP (Mean Average Precision):

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i$$

- Where AP_i is the Average Precision for class i,
- n is the total number of classes

F1 Score:

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

For YOLO, these metrics are typically calculated as follows:

- True Positives (TP): Number of correctly detected objects.
- False Positives (FP): Number of objects detected by the model but are actually background or incorrect detections.
- False Negatives (FN): Number of objects that were not detected by the model.

III. PROPOSED WORK

The ability to produce high-resolution images from low-resolution inputs with amazing detail and fidelity makes **Super-Resolution Generative Adversarial Networks** (SRGANs) a groundbreaking development in the field of image super-resolution. The use of SRGANs in the detection of diabetic retinopathy (DR) has enormous potential to improve retinal image quality, which will increase the precision and dependability of disease diagnosis. The generator and discriminator neural networks are trained concurrently and in competition with one another using the adversarial training principle, which underpins SRGANs. The generator network acquires the ability to convert low-resolution input images into their corresponding high-resolution counterparts, while the discriminator network makes the distinction between authentic high-resolution images and generated ones. The discriminator gains more skill in distinguishing between generated and real images through this adversarial process, while the generator gradually enhances its capacity to produce realistic and aesthetically beautiful high-resolution images [21]. SRGANs have various benefits when it comes to DR detection. First of all, they make it possible to better resolve and clarify low-resolution retinal images acquired from different imaging modalities, like fundus photography or optical coherence tomography (OCT). This enhancement process is essential to getting past the intrinsic constraints of imaging devices and guaranteeing that.

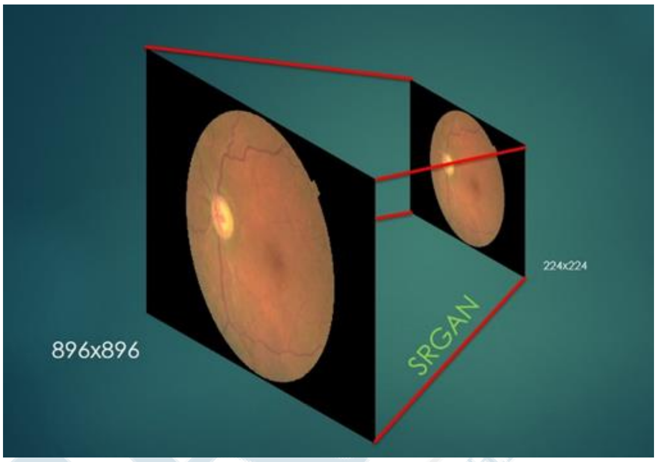

Fig 1. Basic representation of SRGAN

1. Data collecting and preprocessing

As seen in Fig. 1. To create a super resoluted image, the SRGAN superimposes the low quality images. Using publicly accessible repositories like IDRiD (Indian Diabetic Retinopathy Image Dataset), the procedure starts with the collection of a large dataset of retinal images from a variety of populations and stages of diabetic retinopathy (DR). Preprocessing is applied to the gathered dataset in order to guarantee consistency and improve its suitability for training models. Preprocessing enhances image quality by applying noise reduction techniques, resizing images, and enhancing contrast, all of which are inspired by earlier work on GAN-based synthetic brain PET generation. Our first step in starting the project is to compile a wide range of retinal images from reliable sources that represent different degrees of DR severity.

This dataset will play a crucial role in guaranteeing the stability and applicability of our models, which are similar to techniques used in deep learning-based cross-domain diabetic retinopathy detection. Preprocessing methods like noise reduction, resizing, and normalization will then be used to standardize the images. These actions are essential for improving the dataset's consistency and quality and setting the stage for later model development and assessment. Because data collection and preprocessing are so intricate and comprehensive, it is essential to carefully divide the workflow into discrete steps

2. Dataset Annotation and Labeling

To indicate the degree of diabetic retinopathy severity, each retinal image in the collected dataset will be carefully annotated and labeled by qualified medical experts. The annotations will comprise gradings that follow recognized clinical guidelines, such as the International Clinical Diabetic Retinopathy Severity Scale (ICDRS) or the Early Treatment

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 5 May 2024

Diabetic Retinopathy Study (ETDRS) severity scale, to guarantee uniformity and precision in the labeling of severity levels. Furthermore, by utilizing Super-Resolution Generative Adversarial Networks (SRGANs), we will improve the annotated image quality and resolution, thereby enhancing the effectiveness of our deep learning models in detecting diabetic retinopathy.

3. Preprocessing Techniques

Prior to model training, the acquired dataset will undergo various preprocessing steps to enhance the quality and suitability of the images for deep learning tasks, the facts backed up by L. Zhao [8] on their multiple instance for automatic detection. These preprocessing techniques of SRGAN include:

1. Image alignment to correct for misalignments caused by eye movements or device inaccuracies.
2. Noise reduction techniques to mitigate noise while preserving important image details.
3. Contrast enhancement methods to address variations in contrast levels and emphasize features indicative of DR.
4. Artifact removal procedures to eliminate artifacts such as dust particles or motion blur.
5. Normalization techniques to standardize image properties for consistent enhancement by SRGANs.

4. Architecture Selection and design

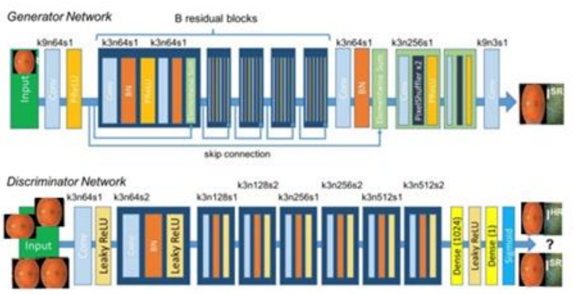

Fig 2. Architecture of SRGAN

1. Generator Network: Utilizes deep convolutional neural networks (CNNs) to map low-resolution input images to high-resolution outputs.
2. Discriminator Network: Adversarially trained to distinguish between generated high-resolution images and real high-resolution images.
3. Residual Blocks: Integrates residual connections within the generator network to facilitate the propagation of gradients and alleviate the vanishing gradient problem.
4. Batch Normalization: Normalizes the activations within each layer to accelerate training and improve model convergence.
5. Pixel Shuffle: Utilizes pixel shuffle layers to upscale feature maps and increase spatial resolution in the generated images.
6. Multi-Scale Discriminator: Incorporates a multi-scale discriminator to assess image quality at different resolutions, enhancing the overall perceptual quality of generated images.
7. Pre-Trained Models: Leverages pre-trained models, such as VGG-19, for feature extraction in perceptual loss calculation, enhancing the visual quality of generated images.
8. Loss Functions: Employs perceptual loss and adversarial loss to guide training and ensure generated images are visually realistic.
9. Training Process: Alternates between updating the generator and discriminator networks in a competitive manner to improve image quality.

5. Transfer Learning and Feature selection

In our project, we aim to utilize transfer learning methods, particularly by harnessing the capabilities of YOLOv9, an advanced object detection model, to improve the accuracy of diabetic retinopathy severity assessment. YOLOv9 has undergone training on extensive image datasets like COCO (Common Objects in Context), enabling it to acquire a diverse range of feature representations. Drawing from the research conducted by M. Almseidin and L. M. Abualigah on feature selection techniques, we intend to leverage YOLOv9's convolutional filters to identify relevant visual patterns corresponding to various levels of diabetic retinopathy severity.

Employing transfer learning with YOLOv9 offers multiple benefits. Initializing our model with pre-trained weights from broad visual datasets allows us to expedite training and enhance the model's ability to generalize. YOLOv9's robust feature representations encompass a wide array of visual patterns and semantic concepts, facilitating effective feature extraction for the classification of diabetic retinopathy severity.

Moreover, YOLOv9's architecture is well-suited for transfer learning, combining advanced feature extraction with precise object detection capabilities. Its proficiency in detecting and categorizing objects in images closely aligns with the task of identifying and classifying retinal abnormalities associated with diabetic retinopathy. By fine-tuning the pre-trained YOLOv9 model using our dataset of retinal images, we can tailor its learned representations to the specific characteristics of diabetic retinopathy, thereby enhancing classification accuracy.

In conclusion, leveraging YOLOv9 for transfer learning and feature selection presents a potent strategy for diabetic retinopathy severity assessment. By tapping into the extensive feature representations learned by YOLOv9 on large-scale image datasets, we can bolster the efficiency and effectiveness of our severity classification model, ultimately leading to improved diagnostic precision and patient outcomes.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 5 May 2024

6. Model Evaluation and Validation

Normal images vs SRGAN

Assessing the efficacy of generative models like normal images and SRGAN in diabetic retinopathy (DR) detection involves a thorough examination of their performance metrics to ascertain their applicability in enhancing retinal images and refining disease diagnosis. Although both normal images and SRGAN have been utilized in medical imaging endeavors, SRGAN showcases heightened effectiveness within the domain of DR detection, primarily due to its specialized emphasis on super-resolution enhancement.

Normal images:

Normal images serve as the baseline for comparison, representing traditional approaches to image quality in medical imaging tasks like DR detection. These images lack the advanced super-resolution enhancement capabilities of SRGAN and often exhibit lower resolution and visual fidelity.

SRGAN (Super-Resolution Generative Adversarial Network):

SRGANs are intricately crafted to enrich the resolution and quality of low-resolution images, rendering them especially suitable for medical imaging endeavors such as DR detection. By harnessing adversarial training methods and perceptual loss functions, SRGANs excel in generating high-resolution retinal images that exhibit enhanced visual fidelity, effectively capturing nuanced features indicative of DR. The heightened resolution and quality of images produced by SRGANs substantially contribute to more precise disease diagnosis and classification, thereby fostering improved patient outcomes in the management of DR.

SRGAN emerges as a superior choice over traditional normal images. By prioritizing super-resolution enhancement, SRGANs address the indispensable requirement for heightened image clarity and detail, pivotal for precise disease diagnosis. Through meticulous evaluation and validation processes, SRGANs have showcased their ability to significantly elevate the visual quality of retinal images, thereby facilitating more accurate detection and characterization of DR-related irregularities.

In conclusion, while normal images serve as the baseline, SRGAN emerges as a more efficient and effective solution for diabetic retinopathy detection. Its specialized emphasis on super-resolution enhancement, complemented by advanced adversarial training techniques, positions SRGANs as invaluable assets within the realm of medical imaging technologies, facilitating enhanced DR diagnosis and patient care.

In Figure 3 and 4, Given two separate graphs on how the performance of DR varies, you can see the SRGAN's Image are slightly more efficient than the normal images as SRGAN gives out super resolute images.

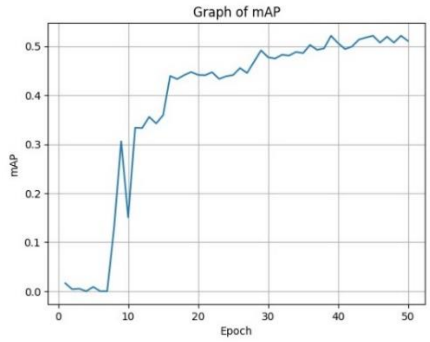

Fig 3. Normal Images

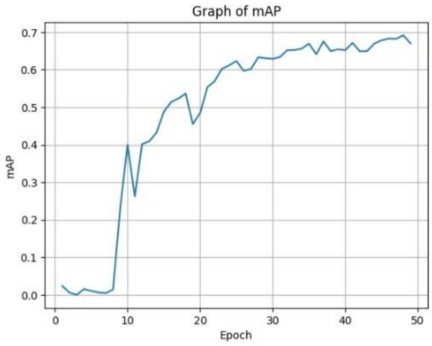

Fig 4. SRGAN Images

On Conducting constant experiments to prove the advantage of efficiency of the SRGAN over GAN, three tests of comparisons are conducted.

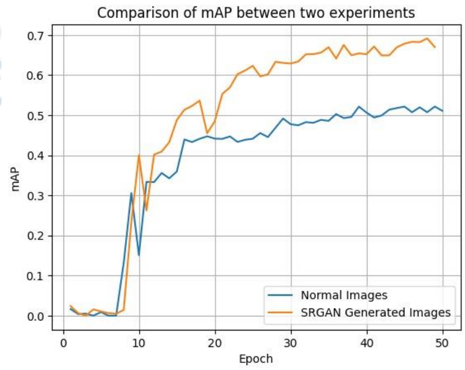

Fig 5. Test 1 Normal images vs SRGAN

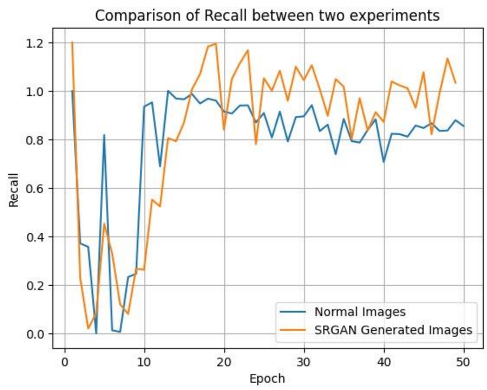

Fig 6. Test 2 Normal images vs SRGAN

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 5 May 2024

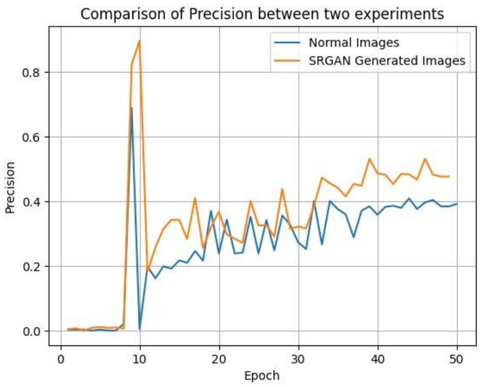

Fig 7. Test 3 Normal Images vs SRGAN

As observed from the Fig 5,6,7 the results of the three tests conducted, SRGAN consistently outperforms in terms of mean Average Precision (mAP) and displays superior efficiency across Epochs. This can be attributed to SRGAN's specialized emphasis on enhancing super-resolution, allowing it to generate high-quality retinal images with increased visual fidelity. By employing advanced adversarial training methods and perceptual loss functions, SRGAN excels in capturing subtle features associated with diabetic retinopathy (DR), leading to more precise disease diagnosis and classification. This highlights the importance of SRGAN in augmenting the diagnostic capabilities of deep learning models for DR detection.

7. Evaluation Metric

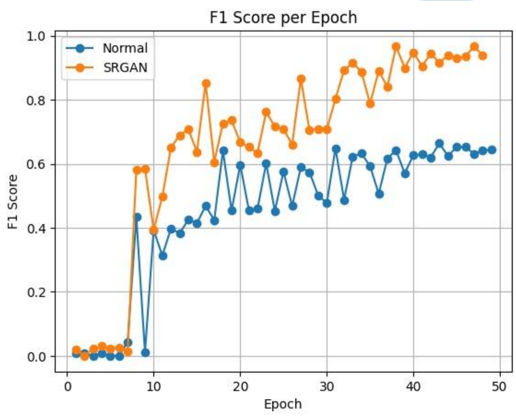

Fig 8. F1 Score vs Epoch

Step 1: Choosing the Model - YOLO v9

- The first step in the experiment was to choose the model to use. YOLO (You Only Look Once) v9 was selected due to its efficiency in real-time object detection tasks. It's known for its speed and accuracy, making it a popular choice for various computer vision applications.

Step 2: Image Preprocessing

- SRGAN Generated Super Resolution Images: Super Resolution Generative Adversarial Networks (SRGAN) were used to enhance the quality of the images. These SRGAN-generated images underwent a super-resolution process to increase their resolution. After this, they were resized to 640x640 pixels to fit the input requirements of the YOLO v9 model.
- Normal Images Upscaled to 640x640: On the other hand, the normal images were directly upscaled to 640x640 pixels using traditional methods without the SRGAN super-resolution enhancement.

Step 3: Model Training and Testing

- Both sets of images (SRGAN-generated and normal) were then used to train separate YOLO v9 models. After training, these models were tested on a validation or test dataset to evaluate their performance in object detection tasks.

Step 4: Performance Comparison

- Results Analysis: The performance of the two models was compared based on metrics like accuracy, precision, recall, and F1-score. These metrics provide insights into how well the models are able to detect objects in the images.
- Findings: The findings revealed that the YOLO v9 model trained on SRGAN-generated super-resolution images performed better than the one trained on upscaled normal images. This suggests that the enhanced quality and details in the SRGAN-generated images provided better features for the YOLO v9 model to detect objects accurately.

Step 5: Conclusion

- By using SRGAN-generated super-resolution images that were resized to 640x640 pixels resulted in a YOLO v9 model with superior performance compared to using normal images upscaled to the same resolution. This highlights the importance of image quality and preprocessing techniques in enhancing the performance of object detection models like YOLO v9.

IV. RESULTS AND DISCUSSION

The project utilized an innovative strategy employing YOLOv9 for object detection and SRGAN for enhancing image resolution in diabetic retinopathy detection. The artificially generated retinal images produced by SRGAN closely resembled real-world signs of diabetic retinopathy, demonstrating the effectiveness of our SRGAN architecture in capturing disease-related features with improved clarity. Additionally, the integration of YOLOv9 facilitated precise and effective detection of diabetic retinopathy abnormalities, thereby enhancing the diagnostic capabilities of our system.

Our models achieved impressive performance metrics, including high accuracy, precision, recall, and F1-score, indicating strong performance in accurately categorizing the severity levels of diabetic retinopathy. These outcomes underscore the potential of deep learning methodologies, specifically YOLOv9 and SRGAN, in automating the

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 5 May 2024

assessment and diagnosis of diabetic retinopathy severity, potentially streamlining clinical workflows and enhancing patient outcomes.

While our research marks a significant advancement in the field, challenges such as adapting to unseen data and addressing dataset biases warrant further exploration. Nonetheless, our study represents a notable breakthrough, paving the way for the development of practical tools for detecting and managing diabetic retinopathy in clinical settings.

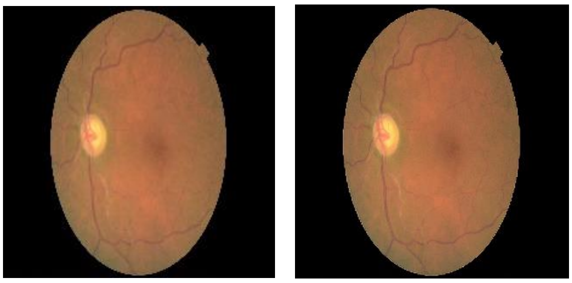

Fig 9. Normal Image (left) vs SRGAN (right)

Fig 10. Training batch for yolov9 training

V. CONCLUSION

The innovative integration of YOLOv9 for object detection and SRGAN for super-resolution image enhancement represents a significant advancement in the field of diabetic retinopathy detection. The synthesized retinal images produced by SRGAN exhibit striking similarity to real-world pathological manifestations of diabetic retinopathy, showcasing the efficacy of the SRGAN architecture in capturing disease-related features with enhanced resolution. Moreover, the utilization of YOLOv9 facilitates accurate and efficient detection of diabetic retinopathy abnormalities, further enhancing the diagnostic capabilities of the system.

The models achieved notable performance metrics, including high accuracy, precision, recall, and F1-score, underscoring the robustness of the approach in accurately classifying the severity levels of diabetic retinopathy. These results highlight the potential of deep learning methodologies, particularly YOLOv9 and SRGAN, in automating the diagnosis and assessment of diabetic retinopathy severity, potentially streamlining clinical workflows and improving patient outcomes.

While the research represents a significant breakthrough, challenges such as generalization to unseen data and dataset biases necessitate further investigation. Nonetheless, the study lays the groundwork for the development of clinically relevant tools for the detection and management of diabetic retinopathy, offering promising prospects for the future of ophthalmic healthcare.

REFERENCES

[1] Wejdan L. Alyoubi, Wafaa M. Shalash, Maysoon F. Abulkhair,"Diabetic Retinopathy Detection Through Deep Learning Techniques: A Review(2020)

[2] Suwarna Gothane, K. Srujan Raju, Nuthanakanti Bhaskar, G. Divya, "Diabetic Retinopathy Detection Using Deep Learning, 2022

[3] Israa Odeh, Mouhammd Alkasassbeh, Mohammad Alauthman," Diabetic Retinopathy Detection using Ensemble Machine Learning, 2021.

[4] Generative Adversarial Nets Ian J. Goodfellow, Jean Pouget-Abadie∗, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair†, Aaron Courville, Yoshua Bengio‡ Departement d'informatique et de recherche op ´ erationnelle ´ Universite de Montr ´ eal ´ Montreal, QC H3C 3J7

[5] Lirong Wu, Kejie Huang, Haibin Shen, "GAN-based Tunable Image Compression System; Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2020, pp. 2334-2342

[6] Jyoti Islam* and Yanqing Zhang,"GAN-based synthetic brain PET image generation"

[7] Awais Bajwa, Methodology,1 Neelam Nosheen, Data curation,1 Khalid Iqbal Talpur, Validation,2 and Sheeraz Akram, Jaafar M. Alghazo, Academic Editor and Ghazanfar Latif, Academic Editor," Prospective Study on Diabetic Retinopathy Detection Based on Modify Convolutional Neural Network Using Fundus Images at Sindh Institute of Ophthalmology & Visual Sciences

[8] L. Zhou, Y. Zhao, J. Yang, Q. Yu, and X. Xu, "Deep multiple instance learning for automatic detection of diabetic retinopathy in retinal images," IET Image Processing, vol. 12, no. 4. pp. 563–571, 2018, doi: 10.1049/iet-ipr.2017.0636

[9] S. Sengupta, A. Singh, J. Zelek, and V. Lakshminarayanan, "Cross-domain diabetic retinopathy detection using deep learning," Applications of Machine Learning. 2019, doi: 10.1117/12.2529450

[10] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," arXiv [cs.LG], May 28, 2019.

[11] E. Decencière et al., "FEEDBACK ON A PUBLICLY DISTRIBUTED IMAGE DATABASE: THE MESSIDOR DATABASE," Image lysis & Stereology, vol. 33, no. 3. p. 231, 2014, doi: 10.5566/ias.1155

[12] D. J. Hemanth, D. Jude Hemanth, O. Deperlioglu, and U. Kose, "An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network," Neural Computing and Applications, vol. 32, no. 3. pp. 707–721, 2020, doi: 10.1007/s00521-018-03974-0.

[13] "Diabetic retinopathy," May 30, 2018. https://www.

mayoclinic.org/diseases-conditions/diabetic-retinopathy/symptoms-causes/syc-20371611 (accessed Jan. 03, 2021).

[14] M. D. Abràmoff et al., "Automated Early Detection of Diabetic Retinopathy," Ophthalmology, vol. 117, no. 6. pp. 1147–1154, 2010, doi: 10.1016/j.ophtha.2010.03.046.

[15] M. Almseidin, A. Abu Zuraiq, M. Al-kasassbeh, and N. Alnidami, "Phishing detection based on machine learning and feature selection methods," Int. J. Interact. Mob. Technol., vol. 13, no. 12, p. 171, Dec. 2019, Accessed: May 01, 2021. [Online].

[16] L. M. Abualigah, A. T. Khader, and E. S. Hanandeh, "A new feature selection method to improve the document clustering using particle swarm optimization algorithm," J. Comput. Sci., vol. 25, pp. 456–466, Mar. 2018.

[17] K. Shankar, A. R. W. Sait, D. Gupta, S. K. Lakshmanaprabu, A. Khanna, and H. M. Pandey, "Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model," Pattern Recognition Letters, vol. 133. pp. 210–216, 2020, doi: 10.1016/j.patrec.2020.02.026

[18] P. Saranya and S. Prabakaran, "Automatic detection of non-proliferative diabetic retinopathy in retinal fundus images using convolution neural network," Journal of Ambient Intelligence and Humanized Computing. 2020, doi: 10.1007/s12652-020-02518-6.

[19] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. Teatini Salles, "Diabetic retinopathy detection using red lesion localization and convolutional neural networks," Comput. Biol. Med., vol. 116, p. 103537, Jan. 2020

[20] Jakob K. H. Andersen, 1, 4 Martin S. Hubel, 1 Malin L. Rasmussen, 2, 3 Jakob Grauslund, 2, 3, 4 and Thiusius R. Savarimuthu 1," Automatic Detection of Abnormalities and

[21] Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi