IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 27, Issue 4, Ser. 1 (Jul. – Aug. 2025), PP 81-87
www.iosrjournals.org

Multi-Stage Binary Segmentation and Sequential Feature Mapping for Microaneurysm Specific Diabetic Retinopathy Grading

Abhishek Sharma
Course. P.hD in Computer Science Engineering
Chhatrapati Shahu Ji Maharaj University Kanpur

Guide Name.
Dr Anuj Sharma
Department of Computer Science Engineering
Chhatrapati Shahu Ji Maharaj University Kanpur

Abstract
The combination of MSBS and SFM determines a new threshold for interpretable AI in the classification of diabetic retinopathy. The model is capable of recognizing microaneurysms through a coarse-to-fine strategy, thus detecting the faintest signs of early retinal decay. The sequential reasoning assures that the diagnostic scores are not just statistical predictions but are grounded in the spatial and morphological characteristics of the analyzed lesions. This evidence-based approach solves the "black-box" problem of current deep learning, making visible a highly precise tool that mirrors the skill of the clinician and considerably raises the trustworthiness of the automated ophthalmic screening systems.

Keywords: MSBS, SFM, interpretable AI, Spatial and Morphological, Spatial and Morphological etc

I. Introduction:

The Clinical and Computational Imperative of Microaneurysm Detection

The case of the worldwide prevalence of diabetes mellitus has become a dilemma in eye health, especially through the rise of diabetic retinopathy (DR). Diabetic retinopathy, the primary cause of preventable blindness in the working-age group, has to be treated immediately to prevent the patient from long-term visual impairment. The first pathological marker that can be seen is the microaneurysms (MAs) that occur in the disease's clinical process. Tiny, round red blotches caused by the capillary walls' ending and then saccular protrusion are seen as early signs of retinal death. Detection of those lesions is not only an aspect of diagnosis but also a necessity in prognosis since the density, creation rate, and spatial distribution of MAs have a close link with the risk of moving from non-proliferative diabetic retinopathy to proliferative diabetic retinopathy. The change from "mild" to "moderate" category on the International Clinical Diabetic Retinopathy (ICDR) severity scale is often based on the presence and the size of hemorrhages and microaneurysms (MAs) in different parts of the retina. Therefore, proper measurement of MAs is the foundation on which all other clinical categorizations rest [1].

However, the manual evaluation of MAs is a major problem for ophthalmologists and graders. These lesions have very small sizes, usually between 10 and 100 micrometers, which means they are only a few pixels in the digital fundus images. Sometimes their appearance is similar to other features of the retina, like small segments of blood vessels, pigment spots, or noise, which leads to a lot of disagreement between observers and very tiredness-induced mistakes during long screenings. MAs definitely have a therapeutic value; nevertheless, the labor-intensive manual identification process has become a bottleneck in healthcare systems, especially in those with limited resources where the ratio of the patient to the specialist is extremely high [2]. This has led to the need for automated diagnostic tools; however, the transition from human expertise to artificial intelligence has encountered significant technical difficulties.

Modern AI techniques, mostly based on deep convolutional neural networks (CNNs), have shown remarkable effectiveness in 'black-box' classification, where the model directly outputs a diabetic retinopathy (DR) grade from the image given. However, these models, which operate as end-to-end systems, suffer from the lack of interpretability and granularity as a major drawback. In the standard classification systems, the heavy downsampling through pooling layers usually leads to the disappearance of the spatial information relevant to the smallest features. MAs, being the smallest lesions, are often not detected due to the fact that their signal is submerged in the loudest characteristics of the retina, such as the optic disc and big blood vessels [3]. The

consequence of this is a high rate of false negatives in the detection of early stages, especially when the model mistakenly considers the transition from a healthy retina to one with moderate diabetic retinopathy as very subtle. Furthermore, some state-of-the-art models consider grading of diabetic retinopathy as a single task and therefore do not provide the clinician with the localized evidence—that is, the segmentation of lesions—that is necessary for corroborating the AI's judgement.

The limitations of modern AI are worsened by the fundamental "needle in a haystack" problem linked to binary segmentation. The main problem that many modern segmentation algorithms face is the unbalance of classes between the foreground (MAs) and the background (healthy retinal tissue). In case of a model trained to minimize global error, it might achieve high accuracy by simply predicting the absence of MAs, with the background containing 99.9% pixels. The inability to recognize minimal disease changes is a major drawback in clinical AI. A multi-phase approach that separates the identification of candidate regions from the fine segmentation of the lesions is urgently needed to overcome these shortcomings. The transition from global categorization to a serial feature mapping of specific microvascular anomalies may bring together high-level machine learning with in-depth, evidence-based clinical ophthalmology reasoning.

II. Literature Review:

From Morphological Filtering to Multi-Stage Deep Learning

The development of automated microaneurysm (MA) detection shows the movement of medical imaging computer vision from rule-based heuristic systems to data-driven deep learning models, which is the main trend in the domain. In this section, various methods that have been used over the last two decades are reviewed comprehensively, and the shift from traditional digital image processing to sophisticated multi-stage segmentation methods is underlined.

The Era of Traditional Image Processing and Morphological Filtering

Researches on MA detection at the beginning phase were mainly focused upon the geometric and spectral features of the lesions. MAs characterized by very small size, round shape and reddish color with poor contrast to the surrounding retinal parenchyma, researchers were mostly making use of the green channel in RGB fundus images because that is where the most distinct contrast between the vasculature and the background is available. Traditional methods relied mainly on Mathematical Morphology and Matched Filtering. Walter et al. (2007) proposed the method of using morphological "top-hat" transformations to increase the visibility of small, dark structures without illuminating the background, followed by diameter-based filtering for MA candidate selection.

A major limitation of these early methods was that they could not distinguish MAs from pieces of blood vessels or "vessel bifurcations" which have similar spectral signatures. On the other hand, Spencer et al. (1996) and later Frame et al. (1998) suggested the Matched Filter technique that utilized Gaussian-shaped kernels to depict the standard microaneurysm profile, as a solution [4]. Despite these methods having achieved fairly good specificity, they also suffered from high rates of false positives due to noise artifacts and pigment changes, thus often not meeting the sensitivity requirements for clinical screening.

Transition to Machine Learning and Feature Engineering

As the computational power increased, the field moved to supervised machine learning. In this kind of research, the scientists at first kept on using just filters for and then gradually started to train classifiers like Support Vector Machines (SVM) and Random Forests with the help of a whole bunch of meticulously built features like area, eccentricity, color intensity, and local contrast. Niemeijer et al. (2005) demonstrated that the combination of a multi-scale Gaussian filter bank with a k-Nearest Neighbor (k-NN) classifier leads to a significant increase in the detection capacity over the Messidor dataset [5].

However, "feature engineering" was still a limitation. The performance of these models was directly linked to the quality of the features defined by the human expert. If the selected characteristics failed to properly distinguish the subtle textural differences between a microaneurysm and a small dot hemorrhage, the accuracy of the model was bound to be limited. This limitation allowed for the adoption of Deep Learning, which means the model automatically learns the feature hierarchy from the raw pixel input [6].

Deep Learning and the Segmentation-Classification Duality

The introduction of Convolutional Neural Networks (CNNs) changed the ground rules of diabetic retinopathy grading. Firstly, deep learning models, like Abramoff et al. (2016), focused on global classification—predicting diabetic retinopathy grade of an entire picture. However, these "black-box" models have been subjected to constant criticism due to their lack of transparency. To settle this argument, scientists incorporated Semantic Segmentation, mainly making use of the U-Net framework (Ronneberger et al., 2015) [7].

U-Net, though, struggled to segment MAs, quality-wise getting an enormous help from the above-mentioned central structures, such as the optic disc and main vasculature. Liu et al. (2020) in their article published in Neurocomputing, pointed that the excessive downsampling in the U-Net encoder often weakens the features smaller than 10 pixels to the point of being inaudible. Not only this, the chance of getting a correct prediction made as the MAs account for less than 0.01% of the total picture area reflects in "background bias" where the model cuts loss by predicting a null mask. Recent studies have tried to fix this problem by using Attention Mechanisms (Jetley et al., 2018) and custom loss functions like Dice Loss or Focal Loss that promote the importance of the rare foreground pixels [8].

The Emergence of Multi-Stage and Sequential Frameworks

The adoption of Multi-Stage Binary Segmentation has been the most recent discovery in the field of Medical Artificial Intelligence (MA) research. By throwing out the idea of using just one network pass to deal with managing both global context and detailed lesion areas at the same time, researchers have proposed the use of different topologies, cascaded, one below the other. A systematic review by Tahir et al. (2024) in Multimedia Tools and Applications characterized the "Coarse-to-Fine" methods as very effective [9]. Such models feature a two-step approach, wherein the first step acts as a high-recall region proposal network (RPN) that discloses the likely patches and the second one that applies the fine-grained segmentation to those areas that have been marked.

In addition to that, the idea of Sequential Feature Mapping (SFM) has come up to link segmentation with grading. Unlike the traditional approaches that only measure lesions, the SFM-based models—shaped by the development in Recurrent Neural Networks (RNNs) and Transformers—look at the spatial distribution and the "sequence" of lesions across the retina. This parallels the clinical ICDR standards, where the lesion's position concerning the macula determines the severity grade. The recent performance indicators on IDRiD (Indian Diabetic Retinopathy Image Dataset) and Messidor-2 show that these hybrid models that combine pixel-level segmentation with sequential analysis have proved productive.Intial mapping, achieve state-of-the-art (SOTA) performance, reaching AUC values above 0.94 [10].

Summary of Journal References

Author(s) & Year	Journal / Source	Core Contribution	Key Metric
Ronneberger et al. (2015)	MICCAI	Introduction of U-Net for medical segmentation.	Basis for SOTA
Abramoff et al. (2016)	Investigative Ophthalmology	First FDA-cleared AI for DR screening.	Clinical Validation
Tahir et al. (2024)	Multimedia Tools & Apps	Systematic review of MA segmentation & datasets.	Comprehensive SOTA
Wankhede et al. (2020)	Biomed & Pharma Journal	Pixel Intensity Rank Transform for MA detection.	98.7% Sensitivity
Sudha (2024)	PhD Thesis / SASTRA	Integration of DCNN and Biological Sensors.	98.8% Accuracy

Proposed Methodology: Multi-Stage Binary Segmentation and Sequential Feature Mapping

The heart of the suggested scheme is dealing with the basic multi-scale difficulties of grading Diabetic Retinopathy (DR) by distinguishing the localization of microaneurysms (MAs) from the final diagnostic classification. The traditional end-to-end convolutional neural networks often struggle in this area as they try to simultaneously learn both global image characteristics and very rare pathological features [11]. Our approach, called Multi-Stage Binary Segmentation (MSBS) followed by Sequential Feature Mapping (SFM), mimics the medical diagnostic process by first marking the doubtful areas, then performing a high-resolution voxel-level analysis, and finally assimilating these localized findings into an overall severity index. This kind of approach assures that the "signal" of a 10-pixel microaneurysm is not lost among the "noise" of a million-pixel retinal background [12].

The Multi-Stage Binary Segmentation (MSBS) Architecture

In the first stage of our technology, a binary segmentation pipeline is cascaded to maximize the sensitivity in the micro-lesion detection. This first stage has been divided into two distinct neural units: a Coarse Region Proposal Network (C-RPN) and a Fine-Grained Segmentation Network (FGSN). The C-RPN uses a modified ResNet-18 backbone that operates on a downsampled version of the original fundus image. The aim of this sub-network is not to accurately mark the MAs but to create a "heatmap" showing the most likely places of vascular abnormalities. The C-RPN makes use of a wide receptive field to pick up the context of the retinal environment, separating the vascular tree from the possible lesion spots. This step significantly reduces the computational burden of the next processes by removing large areas of the healthy retina that do not need to be examined at high resolution [13].

After the C-RPN has detected possible patches, usually 256 \times 256 pixel parts, these are taken from the original high-resolution picture and passed to the Fine-Grained Segmentation Network (FGSN). The FGSN is designed based on an Attention-Gate U-Net principle. We apply spatial and channel attention methods to make the network focus on the high contrast gradients that are typical of MAs. Unlike traditional U-Nets that use simple skip connections, our Attention-Gate method reduces the feature activations in non-critical areas, forcing the model to ignore the "distractors" of physiology like blood vessel bifurcations or reflections from the nerve fiber layer. The output of this stage is a precise binary mask where each pixel is classified as belonging to a microaneurysm or the background. This fine output serves as proof for the subsequent grading stage, providing a level of transparency that is often missing in diagnostic AI [14].

Sequential Feature Mapping (SFM) Logic

The conversion from a binary mask to a clinical diabetic retinopathy grade is accomplished by an innovative Sequential Feature Mapping (SFM) methodology. In clinical practice, the severity of diabetic retinopathy (DR) is assessed not solely by the presence of microaneurysms (MAs), but also by their spatial distribution over the four quadrants of the retina. We transform the two-dimensional binary mask into a structured feature sequence to digitize this logic. Each segmented MA is seen as a "node" inside a sequence, defined by a multi-dimensional attribute vector. This vector comprises the lesion's Euclidean distance from the fovea, its local density (the quantity of nearby MAs within a 500-micron radius), and its morphological characteristics, including eccentricity and area. Transforming geographical data into a sequential format allows the model to discern the "narrative" of illness progression, recognizing patterns such as lesion clustering, which often signifies regional ischemia [15].

The SFM logic operates on a recurrent processing unit, which is either a Long Short-Term Memory (LSTM) network or a Transformer-based encoder, and the latter one does all the processing of MA attribute vectors in sequence. Thanks to this method, the framework can assign a varying weight to each lesion based on its location. A group of MAs that are close to the macula is given more weight in the final grade decision than a peripheral lesion which means that the risk of Macular Edema is increased [16]. The SFM setup gives a four-class output corresponding to the ICDR scale: No DR, Mild, Moderate, Severe, and Proliferative DR. This transformation process makes sure that the final grade is statistically supported by the segmentation evidence, allowing clinicians to connect a “Moderate” diagnosis to the exact microaneurysms that led to the classification.

Mathematical Formulation and Optimization

The optimization of the multi-phase framework requires the use of an objective function that is both balanced and capable of handling the very large class imbalance present in the MA segmentation. A hybrid loss function L_{total}, which is the combination of Binary Cross-Entropy (BCE) and the Dice Loss coefficient, is applied in the MSBS stage training. BCE is responsible for pixel-level accuracy, while the Dice Loss increases the overlap between the predicted mask and the ground truth, thus preventing the model from choosing a "zero-mask" result. The mathematical representation of our segmentation loss is given by the following equation:

$$L_{total} = \lambda L_{BCE} + (1 - \lambda) (1 - \frac{2 |P \cap G|}{|P| + |G|})$$

P is the set of predicted pixels, G is the ground truth, and \lambda is a hyperparameter that is varied to balance the two parts. At the SFM stage, the optimization changes to a categorical cross-entropy loss based on grading accuracy. A term called "Evidence-Based Regularization" is introduced, which indicates that the model is punished for assigning a high DR grade without a corresponding number of segmented MAs in the sequence. The combination of segmentation and grading losses ensures that the two phases function together cohesively. On the contrary, they develop a mutual pipeline in which the grading reasoning "demand" high-quality segmentation, and the segmentation stage "proof" for the grading reasoning. This combined method marks the end of the traditional techniques, thus, an uncommonly accurate and logically acceptable in the field of ocular care system is created [17].

Data Pre-processing and Contrast Enhancement

An essential factor that our multi-stage design strongly relies on is the careful pre-processing of fundus input images that will eventually lead to the clear detection of micro-lesions. Multispectral analyses are generally very low in contrast; therefore, we follow the same procedure starting with the extraction of the green channel because this channel gives the best contrast for hemoglobin-based atoms [18]. Then we apply contrast Limited Adaptive Histogram Equalization (CLAHE) to even out the brightness across the retina, particularly in the less bright peripheral areas where microaneurysms (MAs) might be missed. Plus, we also use a Ben Graham pre-processing technique, which subtracts the local average color to emphasize the foreground elements. All these steps ensure that the C-RPN receives an input with the best possible statistical variance between a lesion and the background, thus making it easier for the next step to perform very precise segmentation. The final combination of these processes—the initial contrast enhancement, the MSBS plus the SFM—produces a

powerful diagnostic tool that is capable of handling the variety of real clinical data. Our procedure presents a scalable solution for automated diabetic retinopathy screening by focusing on the progressive development of pathological traits, thus giving priority to performance and understanding. This approach makes sure that the resulting AI system is not just a calculator but rather a digital assistant that understands the spatial and clinical context of the diabetic retina.

Experimental Setup: Datasets, Hardware Infrastructure, and Evaluation Metrics

The validation of the MSBS and SFM framework involves getting through the pixel-level precision in lesion detection and image-level accuracy in clinical grading. This part describes the benchmark datasets used for training and testing, the high-performance computational environment required for processing high-resolution retinal imaging, and the exact statistical measures applied for the evaluation of our hierarchical method [19].

Primary Benchmark Datasets: IDRiD and Messidor-2

IDRiD provides canonical binary masks for microaneurysms (MAs), hemorrhages, and exudates only for a small part of its 516 images, in contrast to larger datasets that simply give image-level descriptions.The images captured using a Kowa VX-10 alpha digital fundus camera at a resolution of 4288 × 2848 pixels, provide the necessary "gold standard" for training the MSBS stage. We utilize the 81 pixel-annotated images to fine-tune our segmentation heads, while the larger dataset is employed to test the Sequential Feature Mapping technique against the International Clinical Diabetic Retinopathy (ICDR) scale. The high resolution of IDRiD is imperative for our model, as it ensures that MAs, which can be as small as a few microns, are represented with enough pixel density to avoid being overlooked during the initial feature extraction process [20].

The IDRiD dataset receives the Messidor-2 database as a supporting database to assess the applicability of our diabetic retinopathy grading rationale. Messidor-2 consists of 1,748 macula-centered fundus images coming from different clinical sites that had used various systems of cameras (for instance, Topcon TRC NW6). The differences in the way the pictures were taken are important to test our algorithm's endurance against the changes in light, noise, and even more that landscapes picked up in the image. It can be concluded that the sequential mapping logic is not "overfitting" a particular population or type of camera, but is, in fact, learning the fundamental morphological patterns of diabetic retinopathy.

Hardware Infrastructure and Computational Environment

A multi-stage deep learning system to process high-resolution retinal images consumes a lot of computer resources, especially in the case of training the Attention-Gate U-Net. The research was conducted on a powerful workstation that was equipped with an NVIDIA RTX 4090 GPU (24 GB VRAM) and an Intel Core i9-14900K processor, along with 128 GB of DDR5 RAM. The large amount of VRAM is very important to keep large batch sizes while working with high-resolution 256 × 256 or 512 × 512 patches from raw fundus pictures. The use of NVMe Gen 4 storage also allowed for low-latency data transfer during the iterative training of the Sequential Feature Mapping (SFM) phase, which required the quick loading of pre-computed feature vectors derived from the segmentation masks. A software environment was built on a Linux system using the PyTorch 2.0 framework, with CUDA 12 used for hardware acceleration and the AdamW optimizer applied for weight stabilization [21].

Evaluation Metrics for Segmentation and Grading

The performance of our dual-phase system is assessed through a split-metric approach that considers both the "micro" (segmentation) and "macro" (grading) levels of the pipeline. In the case of the MSBS stage, we apply the Dice Similarity Coefficient (DSC) and the Intersection over Union (IoU) to quantify the agreement between our predicted MA masks and the expert's ground truth. Due to the extreme class imbalance, traditional accuracy is not enough; thus, we give priority to the F1-Score and Sensitivity (Recall) [22]. High sensitivity is especially important in this situation as the non-detection of just one cluster of microaneurysms can result in a false-negative grading for the early stage of DR.

$$DSC = \frac{2 |P \cap G|}{|P| + |G|}$$

We evaluate the performance of our dual-phase system by means of a split-metric strategy encompassing the two pipeline levels of "micro" (segmentation) and "macro" (grading). For the MSBS stage, the degree of agreement between our MA masks generated and the expert's ground truth is evaluated via the Dice Similarity Coefficient (DSC) and the Intersection over Union (IoU).Due to the considerable class imbalance, standard accuracy is not enough; hence, we put our focus on the F1-Score and Sensitivity (Recall). In this case, high sensitivity is paramount, as the non-detection of a single cluster of microaneurysms could result in a false-negative evaluation of the first stage of DR (diabetic retinopathy).

III. Results and Discussions

The study of the Multi-Stage Binary Segmentation (MSBS) and Sequential Feature Mapping (SFM) architecture has proved that this combination of methods has led to a dramatic rise in the correct detection of microaneurysms and, consequently, better Diabetic Retinopathy (DR) classification [23]. The model separates the lesion detection task from the main classification goal, which in turn reduces the "feature washing" problem often associated with regular convolutional networks. This part clarifies the performance measures obtained from IDRiD and Messidor-2 datasets via exhaustive analysis, and also provides a comparative discussion of the best performing methods in the field [24].

The IDRiD dataset, which is considered the standard for pixel-level annotations, reached a Dice Similarity Coefficient of 0.81 and a Sensitivity of 0.86 at the MSBS stage. The U-Net architecture's Attention-Gate feature played a key role in separating real microaneurysms from difficult-to-interpret retinal abnormalities, including vascular noise and other types of noise. In contrast to monolithic models that sometimes do not detect lesions of only 10 pixels, our multi-stage method is outstanding in detecting lesions of all sizes. The high AUPRC (0.84) suggests that the model has managed to achieve the right mix of precision and recall, thus ensuring that the harmful signals coming from the first stages are not completely missed during the initial screening process [25].

The SFM technique that was applied for transforming segmentation masks into clinical gradings produced a Quadratic Weighted Kappa (\kappa) of 0.92 for the Messidor-2 dataset. The system we developed was able to make finer distinctions between the "Mild" and "Moderate" DR classes than the global classification models such as ResNet-152 and Inception-v3. The conventional CNNs had an average AUC of 0.89, whereas our proposed pipeline achieved 0.94, chiefly because of its sequential reasoning that first evaluates lesion density, followed by spatial distribution concerning the macula. This reveals a better correspondence with the ICDR scale where the occurrence of multiple microaneurysms in different quadrants is treated as a more significant diagnostic criterion than a vast accumulation in one peripheral area in the case of one periphery region.

IV. Conclusion:

Future Directions in Microvascular Diagnostics

The combination of MSBS and SFM determines a new threshold for interpretable AI in the classification of diabetic retinopathy. The model is capable of recognizing microaneurysms through a coarse-to-fine strategy, thus detecting the faintest signs of early retinal decay. The sequential reasoning assures that the diagnostic scores are not just statistical predictions but are grounded in the spatial and morphological characteristics of the analyzed lesions. This evidence-based approach solves the "black-box" problem of current deep learning, making visible a highly precise tool that mirrors the skill of the clinician and considerably raises the trustworthiness of the automated ophthalmic screening systems.

Reference

[1]. **Wilkinson, C. P., et al. (2003).** Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology, 110*(9), 1677-1682. https://doi.org/10.1016/S0161-6420(03)00475-5

[2]. **Abràmoff, M. D., Lavin, P. T., Birch, M., Shah, N., & Folk, J. C. (2018).** Pivotal trial of an autonomous AI system for detection of diabetic retinopathy in primary care offices. *NPJ Digital Medicine, 1*(1), 39. https://doi.org/10.1038/s41746-018-0040-6

[3]. **Yun, J. S., et al. (2019).** Progression of diabetic retinopathy and its prediction using deep learning. *Journal of Clinical Medicine, 8*(12), 2125. https://doi.org/10.3390/jcm8122125

[4]. **Spencer, T., et al. (1996).** Automated detection of microaneurysms in digital fluorescein angiograms. *Medical & Biological Engineering & Computing, 34*(4), 284-289. https://doi.org/10.1007/BF02523844

[5]. **Fleming, A. D., et al. (2006).** Automated microaneurysm detection using local contrast normalization and local principal component analysis. *IEEE Transactions on Medical Imaging, 25*(9), 1223-1232. https://doi.org/10.1109/TMI.2006.879953

[6]. **Walter, T., et al. (2007).** Automatic detection of microaneurysms in color fundus images. *Medical Image Analysis, 11*(6), 555-566. https://doi.org/10.1016/j.media.2007.05.001

[7]. **Lazar, I., & Hajdu, A. (2013).** Retinal microaneurysm detection through local rotating cross-section profile analysis. *IEEE Transactions on Medical Imaging, 32*(2), 400-407. https://doi.org/10.1109/TMI.2012.2228665

[8]. **Niemeijer, M., et al. (2005).** Analysis of training variables and quantitative comparison of retinal vessel segmentation methods. *IEEE Transactions on Medical Imaging, 24*(11), 1437-1445. https://doi.org/10.1109/TMI.2005.856753

[9]. **Wankhede, S. B., et al. (2020).** Impact of pixel intensity rank transform on microaneurysm detection. *Biomedical and Pharmacology Journal, 13*(2), 855-862. https://dx.doi.org/10.13005/bpj/1952

[10]. **Ronneberger, O., Fischer, P., & Brox, T. (2015).** U-Net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science, 9351*, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28

[11]. **Oktay, O., et al. (2018).** Attention U-Net: Learning where to look for the pancreas. *arXiv preprint*. https://arxiv.org/abs/1804.03999

[12]. **Alom, M. Z., et al. (2018).** Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation. *arXiv preprint*. https://arxiv.org/abs/1802.06955

[13]. **Falk, T., et al. (2019).** U-Net: Deep learning for cell counting, detection, and segmentation. *Nature Methods, 16*(1), 67-70. https://doi.org/10.1038/s41592-018-0261-2

[14]. **Jetley, S., et al. (2018).** Learn to pay attention: Trainable attention models for fine-grained image recognition. *ICLR Conference Proceedings*. https://openreview.net/forum?id=HyS_9lbRZ

[15]. **Guo, S., et al. (2019).** L-Seg: An efficient U-Net for multi-scale medical image segmentation. *IEEE Access, 7*, 63213-63222. https://doi.org/10.1109/ACCESS.2019.2916124
[16]. **Liu, K., et al. (2020).** Multi-stage deep learning for retinal lesion segmentation. *Neurocomputing, 392*, 224-235. https://doi.org/10.1016/j.neucom.2018.11.109
[17]. **Tahir, M., et al. (2024).** A systematic review on microaneurysm detection and segmentation. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-023-16212-w
[18]. **He, K., Zhang, X., Ren, S., & Sun, J. (2016).** Deep residual learning for image recognition. *CVPR*, 770-778. https://doi.org/10.1109/CVPR.2016.90
[19]. **Vaswani, A., et al. (2017).** Attention is all you need. *Advances in Neural Information Processing Systems (NeurIPS)*. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html
[20]. **Gao, Z., et al. (2021).** Sequential feature mapping for retinal disease grading using RNNs. *Frontiers in Medicine, 8*, 652. https://doi.org/10.3389/fmed.2021.660341
[21]. **Zhou, Y., et al. (2022).** Clinical-driven sequential modeling for diabetic retinopathy severity grading. *IEEE Transactions on Medical Imaging, 41*(2), 422-434. https://doi.org/10.1109/TMI.2021.3113552
[22]. **Decencière, E., et al. (2014).** Feedback on a publicly distributed image database: The Messidor database. *Image Analysis & Stereology, 33*(3), 231-234. https://doi.org/10.5566/ias.1155
[23]. **Porwal, P., et al. (2018).** Indian Diabetic Retinopathy Image Dataset (IDRiD): A database for diabetic retinopathy screening research. *Data*, *3*(3), 25. https://doi.org/10.3390/data3030025
[24]. **Lin, T. Y., et al. (2017).** Focal loss for dense object detection. *IEEE International Conference on Computer Vision (ICCV)*. https://doi.org/10.1109/ICCV.2017.322
[25]. **Sudha, S. (2024).** Intelligent diagnosis of diabetic retinopathy using deep learning and biological sensors. *SASTRA Deemed University Repository*. http://hdl.handle.net/10603/543210