TAU-PSO: A Transferable Attention U-Net with Particle Swarm Optimization for Optic Cup-Disc Segmentation in Fundus Images

Saba Sheiba[1*], M Neelakantappa[2]
[1] Department of Computer Science and Engineering, BEST Innovation University, Andhra Pradesh, India.
[1]Department of CS&AI, Muffakham Jah College of Engineering and Technology, Hyderabad, India.
[2]Department of Information Technology, Vasavi College of Engineering, Hyderabad, India.
E-mail: saba.sheiba@mjcollege.ac.in 2m.neelakanta@gmail.com
*Corresponding author

Keywords*:* Segmentation, glaucoma, U-Net, optic disc, PSO, retinal fundus images

Received: April 18, 2025

The retina is impacted by diabetic retinopathy, a disorder frequently observed in people diagnosed with diabetes mellitus for an extended period. This condition can ultimately lead to vision loss as a result of pathological retinal microvascular leakage. Early diagnosis of any optic disc abnormalities is crucial for ophthalmologists to treat their patients effectively. Segmenting the optic disc to screen for glaucoma using the optic cup-to-disc ratio is also essential. Several retinal diseases can be significantly helped by automated screening of fundus images utilizing computational approaches. To enhance the efficiency of current annotated samples without necessitating thousands of examples of training for Optic Disc (OD) and Optic Cup (OC) Segmentation, we present the lightweight network known as the Transferable Attention U-Net (TAU-Net) model with Particle Swarm Optimization (PSO) for hyperparameter tuning. The model employs attention modules and two adversarial domain discriminators (feature and attention discriminators) to learn domain-invariant representations across DRISHTI-GS1, RIM-ONE, and REFUGE datasets, while structured dropout blocks reduce overfitting. Experimental results show that our method achieves superior segmentation performance compared to several state-of-the-art approaches. Specifically, for the optic cup, the TAU-PSO model attains 98.2% accuracy, 96.6% precision, 97.2% recall, and 96.4% Dice score, while for the optic disc, it achieves 97.1% accuracy, 96.0% precision, 96.5% recall, and 96.8% Dice score. Compared to baseline methods such as NFN+, SDU-Net, and DU-Net, our approach demonstrates consistent improvements of 2–4% across key metrics, highlighting the effectiveness of combining PSO-based hyperparameter optimization with domain-adversarial learning. The proposed TAU-PSO framework provides a robust and transferable solution for OD–OC segmentation under cross-dataset conditions, offering a reliable tool for early diagnosis of diabetic macular edema.

Povzetek: Predlagani model TAU-PSO izboljša segmentacijo očesnih struktur in omogoča zanesljivejše zgodnje odkrivanje bolezni mrežnice.

1 Introduction

Glaucoma is caused by nerve fiber degeneration, which gradually affects a section of the eye that transforms the retina's information into images and transmits them to the brain. An increase in intraocular pressure (IOP), or the fluid pressure inside the eye, could also harm it. The World Health Organization projects that by 2030, there will be 90 million people with glaucoma, up from 64 million in 2016. [1]. The accumulation of pressure in the eye causes optic nerve fiber injury and thickening of the retinal nerve fiber layer (RNFL) If the eye cannot get rid of excess fluid. As a result, a phenomenon known as cupping occurs, in which the optic cup (OC) grows more significantly than the optic disc (OD). Additionally, it can cause peripapillary atrophy (PPA), which is typically linked to high myopia, which is characterized by the weakening and deformation of the optic nerve's surrounding retinal pigment epithelium [2]. After cataracts, glaucoma is the leading cause of permanent vision loss worldwide [3]. Therefore, to prevent irreversible blindness, glaucoma must be diagnosed early. It is common practice to divide glaucoma into two subtypes: open-angle and closed-angle [4]. Most cases of glaucoma, known as open-angle glaucoma, do not close the cornea or iris drainage angle. There are no warning signs for patients with this form of glaucoma, so they do not know they have it until their vision is severely impaired [5]. The condition known as closed-angle glaucoma develops when the iris partially blocks the drainage angle., preventing the outflow of excess fluid and thus elevating intraocular pressure. Noticeable symptoms experienced by patients with this form of glaucoma include reddening of the eye, abrupt onset of ocular pain, increased intraocular pressure, and abrupt loss of vision [6]. The treatment of clinical glaucoma involves a battery of eye exams performed by ophthalmologists. The elevated intraocular pressure can be measured using tonometry. The range of normal pressure is between 12 and 22 mmHg. High intraocular

pressure (IOP) causes the development of glaucoma [7]. The optic nerve is examined via ophthalmoscopy appearance to diagnose glaucoma. Figure 1 describes the grading of glaucoma diseases.

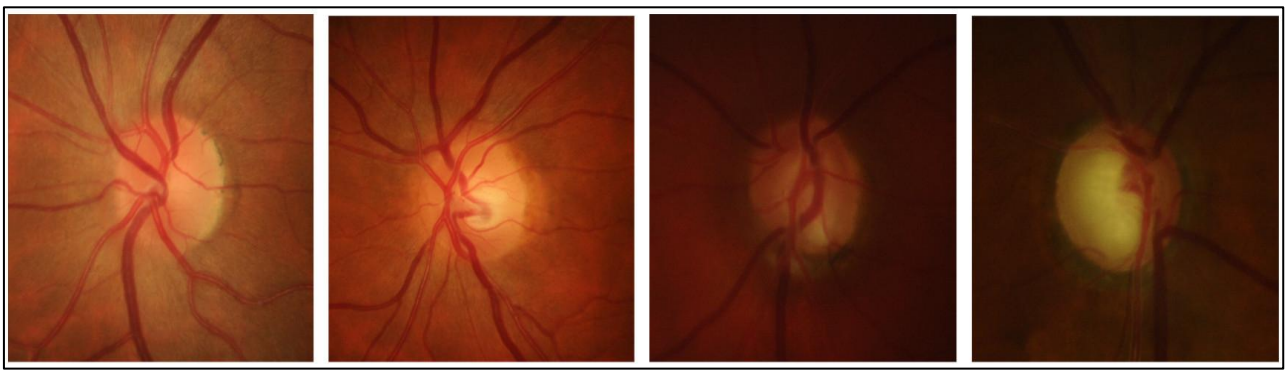

Figure 1. Grading of glaucoma diseases: (a) healthy OD, (b) Mild Glaucoma, (c) Moderate Glaucoma, and (d) severe glaucoma

Patients will be directed to undergo a test called perimetry, a significant change in the shape and color of the optic nerve. Doctors can tell if glaucoma has impacted their patients' eyesight with a perimetry test [8]. The corneal thickness can be examined with a pachymetry test, influencing the eye's pressure readings. Clinicians may use gonioscopy further to confirm an open or restricted fluid outflow angle. Glaucoma is a serious risk for patients with narrow-angle lenses. These methods necessitate the use of qualified specialists for physical examinations of the eye [9]. Inter-observer variability is risky, and these methods are time-consuming and costly [10]. One practical and dependable way to aid in detecting glaucoma is via a computer-aided diagnosis system. An example would be fundus photography's structural analysis of acquired retinal images. As a result, the diagnostic procedure becomes both reliable and economical. Retinal photos are subjective, but computer-aided technologies can assist medical professionals in making more objective assessments. Research has shown a variety of approaches to the early detection of glaucoma [11]. When diagnosing glaucoma, the most reliable technique is to look at the optic nerve head. The OD and OC areas constitute ONH [12]. The cup-to-disc ratio, sometimes known as cupping, is a crucial criterion for diagnosing glaucoma. This can be expressed as the OC's vertical diameter ratio to the OD's vertical diameter [13]. A healthy eye is visible with several optical nerve fibers and a normal-sized cup. However, a rise in intraocular pressure causes nerve fiber degeneration, increasing the cup's surface area and CDR. The CDR value of a healthy eye is below 0.6, while that of a glaucomatous eye is higher than 0.6 [14]. Glaucoma detection using a cup-to-disc ratio alone is inadequate, particularly in those with a hereditary predisposition to a large OC or a history of myopia [15]. Taking note of the neuro-retinal rim's anatomy, a circular area composed of nerve fibers between the disc and cup is another crucial element proposed to detect glaucoma. To screen for glaucoma, doctors follow the ISNT rule, which says that in a healthy eye, The nasal and temporal regions have the thickest rims, followed by the inferior region. The NRR region shrinks as OC size increases in glaucoma; hence, this rule is broken [16].

Although crucial, accurately segmenting OD and OC may provide several obstacles. Segmenting retinal images using fundus cameras has become more arduous because of weak contrast, low resolution, and variable illumination. Aside from reflections and other bright objects, fundus images can also introduce noise in the form of exudates, which look like OD and OC but are very intensely brilliant. Concerned candidates may be mistakenly identified as a result of this. Because there are blood veins in the layer under the retina, it is also difficult to discern the exact boundary of the OC [17]. Additional information that needs to be removed from segmented images is another way noise might be introduced [18]. Previous research has expanded segmentation to include pre-and post-processing steps in response to these issues. Furthermore, before segmentation, several researchers have also proposed localization as a prerequisite, which entails extracting an ROI from images surrounding the OD [19]. Due to its inapplicability to unseen whole-eye fundus images and the need for an OD and OC bounding box, this method cannot be considered generalized. The processing time and total cost are affected by the time required to perform these extra procedures. Joint OC and OD segmentation is challenging, but recent state-of-the-art approaches guarantee competitive performance [19]. Ensemble networks [20] and parameter-heavy generative adversarial networks [16] are how this performance is attained. Using an architecture based on U-Net [2], which is relatively light on parameters, [4] implemented their technique. On the other hand, their approach depends on a heuristic for precisely localizing the OD. Hence, this study's impetus is to slice retinal fundus images into OD and OC regions using a model based on deep learning to remove extra processing blocks that drive up the system's computing cost.

To the best of our knowledge, there is no existing solution that fully addresses the domain shift problem in Optic Disc (OD) and Optic Cup (OC) segmentation. Although segmentation of other retinal lesions is also challenging due to poorly defined boundaries, OD and

OC segmentation under cross-dataset conditions remains a critical task. In this work, we propose a PSO-based Transferable Attention U-Net (TAU) model to address OD and OC segmentation while effectively handling domain shift in fundus image datasets. The proposed framework employs a backbone segmentation network to extract robust and transferable features. Hierarchical attention modules use dynamic learnable coefficients to identify particular OD and OC zones. Two domain discriminators are implemented to encourage the network to learn domain-invariant representations at the feature and picture levels through adversarial training. Be integrated with attention modules. Using two popular datasets of eye fundus images, we assess the suggested technique for adaption tasks. This study overcomes the restrictions using an end-to-end, one-stage deep architecture with image-level inputs of prior approaches' use of multi-stage pipeline processing. It is based on U-Net with PSO for OD segmentation. One of the obstacles to bringing deep learning models to a new domain is the learning rate and momentum, two hyper-parameters of transfer learning that significantly affect network performance. To overcome this, we offer a variation of the PSO algorithm that can automatically choose the best hyper-parameters. The production of numerous unique fine-tuned segments with different optimal training hyperparameters is also made possible by this PSO-based hyperparameter selection procedure. The experimental findings demonstrate that TAU outperforms baseline approaches' segmentation performance on labelled and unlabelled datasets between the destination and source domains. The suggested approach helps identify glaucoma by preserving the medical characteristics of the input eye fundus images, as demonstrated by CDR-based testing.

The main ideas are

- We present TAU-PSO, which combines the adversarial training approach and the attention mechanism to address the domain adaption issues for segmenting OC and O.
- To test the suggested concept, we use three publicly available eye fundus image datasets that have undergone substantial domain shifts. Experimental results show that TAU-PSO performs better on tasks requiring segmentation across datasets than baseline techniques.
- To ensure the proposed strategy is reliable from a medical standpoint, we do experiments based on cup-to-disc ratio (CDR) images. Our findings show that TAU-PSO may be readily adapted to identify glaucoma while retaining the original images' medical characteristics.

The remainder of the paper follows this format: Section 2 reviews the pertinent literature. Section 3 provides a detailed explanation for segmenting OD and OC using the suggested TAU-PSO model. Section Four offers more comprehensive tests and a description of the proposed methodology. Section 5 provides a summary of the paper.

2 Related work

Recently, the segmentation of biomedical images has advanced significantly. A method to extract ROI from an image's red and green channels has been suggested in [21]. The histogram is shown against the red channel image using the Gaussian window to establish the threshold for OD segmentation. The OC's green channel image undergoes a comparable procedure. One study employed an ISNT criteria that considered the neural retinal rim width while screening for glaucoma [22]. The authors of [23] have presented a segmentation method that uses median, mean, and Otsu thresholding. However, these systems are less resilient because patients' colors vary. The intensity-based thresholding method was suggested for discs in [24] to eliminate noisy pixels, and features based on geometry were employed. A method based on vascular tracking bending has been employed for the OC border to monitor and record abrupt changes in blood vessel behavior. With an impressive f-score of 0.9485, this study successfully categorizes fundus images. When dealing with low-contrast images, though, thresholding-based approaches fall short [25]. Because regional information provides robustness against variations in intensity and contrast, edge and region-based segmentation algorithms have been utilized in several publications. One solution to the issues with big datasets was proposed in [26] as a regional information-based approach to getting the cup border using intensity normalization. Using a variety of procedures to achieve a precise disc boundary [27] suggests a way to eradicate peripapillary atrophy. Since non-disc structures are removed, this method's segmentation becomes more accurate. By minimizing the energy function, certain articles have defined the borders of discs and cups using active contours and algorithms based on deformable geometry. In [28], the calculation of CDR and CAR suggested an improved approach to OD and OC segmentation. An adaptive deformable model that can capture shape change and irregularity has been suggested in [29] as a solution to classification problems. The segmentation of optic discs has been suggested [30] utilizing a contour-based active method. However, image noise and pathologies can cause the active contour-based system to become stuck at a local minimum. The effectiveness is heavily reliant on the initialization of the contour model [31]. Regarding segmentation, several researchers have used super pixel-based methods. Segmented disc and cup images were also obtained using an alternative method based on super pixel classification (as described in reference [32]. However, the suggested method has produced inaccurate results for cup identification for both big and small sizes. Preprocessing is crucial to providing high-quality retinal images that can help with segmentation. Noise and lighting problems were resolved in the suggested work [8] by preprocessing fundus images. A classifier was employed to divide the input images into OD, OC, and background regions after obtaining super pixels via the (SLIC) technique based on a super pixel approach. This was done for OD and OC

segmentation. Glaucoma can also be identified by observing the (RNFL) structure [3]. K-nearest neighbour is used to classify images. a framework for glaucoma detection based on statistical features was proposed in [33]. The segmentation results obtained using methods based on machine learning have been satisfactory. The author proposes a technique for effectively and precisely detecting OD in pictures with noise and other abrasions [34]. An automated regression-based approach has been suggested in [35] to determine accurate OC and OD boundaries. The quality of the features extracted by hand from a given dataset is crucial to the success of machine learning algorithms. It takes a lot of time and effort to extract features manually. The extraction of OD contour mathematical morphology and PCA was demonstrated in [36] as an accurate method in the morphology-based approach.

Table 1: Some proposed algorithms to solve retinal blood vessel segmentation

Reference	Year	Methodology	Dataset	Limitations
Karrali et al. [14]	2024	Modified U-Net	DRISHTI	Possible bias due to reliance on a single network architecture and limited cross-dataset validation.
Sule et al. [15]	2023	Deep CNN	RIMONE	Limited comparative evaluation with other state-of-the-art segmentation approaches.
Babu et al. [16]	2023	Deep CNN	REFUGE	Limited robustness and generalizability when applied to larger or heterogeneous datasets.
Zhou et al. [17]	2023	SegR-Net	STARE	Increased architectural complexity and difficulty in handling complex retinal structures.
Yang et al. [10]	2023	Wave-Net	RIMONE	Emphasis on computational efficiency may reduce segmentation accuracy in challenging cases.
Zhou et al. [8]	2023	ResNet	CHASE	Resource-intensive training requirements and increased model complexity.
Alex et al. [20]	2023	CNN	RIMONE	Limited adaptability to colour fundus images and challenges in extending to transformer-based frameworks.
Chen et al. [2]	2024	CNN+Markov	CHASE	Difficulty in effectively integrating deep learning models with probabilistic frameworks.
Dayana et al. [12]	2022	U-Net	REFUGE	High computational cost and extensive training time requirements.
Jiang et al. [13]	2020	Edge-aware U-Net	DRISHTI	Limited discussion on cross-dataset generalization and adaptability of edge-aware features.

Table 1 Recently, there has been a significant increase in the utilization of deep learning-based techniques, which may be trained to learn complex features automatically. Since OD detection and localization are incredibly challenging when working with a large variety of images, and this method will be better suited for future glaucoma identification development, it is recommended that the process begins with localized OD rather than using the clipped image as a PCA grayscale conversion input. According to [37], the OD center can be located using a circular hough transform. Then, a grow-cut method is employed to extract the OD, with the OD center point as the region's initial growth point. Before OD segmentation using polar transform-based adaptive thresholding employing region of interest, preprocessing in [2] includes morphological procedures and OD localization via a circular hough transformation. When compared to the most cutting-edge methods, this strategy performs admirably. The accuracy stability is lower while handling a broad range of images compared to the proposed methodology, even though this method's performance Is good and similar to the datasets used as benchmarks. Intrusion detection and existing IoT security challenges.[38] The first investigates a decision tree-based framework for identifying intrusions in wireless sensor networks [39]. The second looks at IoT evolution and security threats in cyberspace. To improve the network's performance for OD and OC segmentation, the authors of [36] introduced a modified version of the original U-Net CNN. This network uses sampling layers to expand the image dimensions by passing the input image through its contracting and expansive paths. With the shortest prediction time, this approach produces high-quality OD and OC segmentation. For segmenting OD and OC, the authors of [12] have developed a collective learning architecture that draws inspiration from convolution neural networks. A learning architecture for convolutional filters was designed using the entropy sampling approach, which identifies the most informative points. When only a tiny dataset is accessible, the proposed work shines. To determine how effective off-the-shelf CNN designs, such as over feat and VGG-S, are as feature extractors, the authors of [37] conducted an analysis. Fundus images have been preprocessed to assess how well these networks work using vessel inpainting and contrast enhancement methods. Most effective deep

learning models are based on the Fully Convolutional Network (FCN) described by [8]. Instead of creating classification scores, this method generates feature maps utilizing convolution layers extracted from CNN, which is the key component of the technique. Dense pixel-wise output is obtained by up-sampling these feature maps. Here, the CNN is trained end-to-end with input images of similar dimensions, which is a key aspect of the approach for proper object segmentation. Biomedical image segmentation [9] and semantic segmentation [10] are two areas where FCN has found significance. For improved performance, [11] modified the traditional FCN with short and long skip connections. This method, which has been further refined, is now part of the U-Net [12] design. To achieve precise pixel-level localization, the U-Net architecture uses skip connections, which allow for integrating lower-level feature maps with higher-level ones. Multiple image segmentation contests, such as those involving medical images [13,14] and satellite images [15], have demonstrated the efficacy of the encoder-decoder structure. Massive datasets with thousands of annotated samples are necessary for deep network model training. However, this may not be feasible when dealing with biological images.

2.1 Problem Statement

In ophthalmology, precise segmentation of OC and OD is pivotal for diagnosing and managing various eye diseases like glaucoma. However, existing segmentation methods often lack accuracy and efficiency, struggling with variations in vessel width, branching patterns, and noise levels inherent in retinal images. This study develops a deep learning framework based on a modified TAU-PSO approach to address these challenges. By harnessing convolutional neural networks (CNNs) and tailoring modifications to accommodate the intricacies of OC-OD morphology and background noise, the proposed framework aims to achieve precise and robust segmentation of OC and OD. The objective is to enhance diagnostic accuracy, enable early detection of eye diseases, and ultimately improve patient outcomes in ophthalmology.

3 Methodology

OC-OD segmentation has attracted considerable attention from researchers, resulting in the development of various methods tailored for this purpose. To enhance segmentation accuracy, researchers have increasingly turned to deep learning, a subset of artificial intelligence rooted in machine learning principles. Deep learning enables machines to learn autonomously without human intervention in feature calculation and selection, diverging from conventional supervised learning approaches. Deep neural networks, a fundamental aspect of deep learning, analyze input images and automatically discern and internalize their distinctive features, eliminating the necessity for manual feature extraction algorithms. This capability minimizes classification errors during training, facilitating the refinement of classifier parameters and attributes. The progression of deep learning has been propelled by advancements in computational power and the availability of extensive databases, fostering more robust and precise segmentation of OC-OD.

3.1 Dataset description

The suggested approach offers an automated OD-OC segmentation system that can be utilized to identify glaucoma. This automated system will assist ophthalmologists in efficiently analysing retinal images and identifying the disease at an early stage. The method has been tested and evaluated on the DRISHTI-GS1, RIM-ONE, and REFUGE databases. The DRISHTI-GS1 database consists of 101 colour fundus images with 2045 × 1751 pixels resolution to validate OC-OD segmentation. The database comprises notching, average OD and cup boundaries, CDR values, im-age-level judgments, and ground truth for segmentation soft maps annotations. The RIM-ONE database is designed to evaluate glaucoma detection systems, mainly focused on optic nerve head (ONH) segmentation. It consists of retinal fundus images annotated by experts, providing a valuable benchmark for algorithms working on glaucoma diagnosis through OD segmentation. Four hundred retinal fundus images are used for training in the REFUGE dataset. The images are saved in JPEG format and have eight bits per colour channel. Eye doctors take these images as their patients sit up straight using high-resolution 500 images with 2124×2056-pixel. The images are centred on the back to evaluate the ONH, making the macula and OD apparent. It includes 360 images of healthy eyes and 40 images of glaucomatous eyes, together with the corresponding ground facts.

3.2. Preprocessing

To make the auxiliary procedure more straightforward and practical, the retinal colour fundus images are pre-processed to improve the image quality. After the retinal colour fundus images have been pre-processed, our suggested method moves on to picture resizing, a Gaussian filter, erosion, mapping, binary conversion, and masking. When it comes to localization, these sets of procedures are more valuable. Preprocessing a fundus image follows the steps shown in Figure 2. Image resizing across all databases is the initial step in preprocessing to ensure that all images are the same size for automatic processing. The suggested method's efficiency and processing time will both be enhanced by this. This system is drawn from three databases that are available to the public and contain images of varying sizes. For this automated method to work, image resizing is required. Using the initial image resolution for the automation system, this approach performs a basic calculation to resize fundus images. Because every pixel contains vital information essential to detecting the disease at an early stage, resizing images preserves their originality. A threshold-based iterative resizing strategy is adopted. If

either the row size or the column size of the original image exceeds 1000 pixels, both dimensions are scaled down by a factor of 0.5. This resizing operation is applied iteratively until both the row and column dimensions of the image are less than or equal to 1000 pixels. If the original image dimensions are already within this threshold, resizing is not applied. This approach ensures uniformity across datasets while preserving the aspect ratio and anatomical structure of the optic disc and optic cup. The resize dimension of an image can be calculated using Equations (1) and (2).

$$IR_{\text{Rsize}} = \frac{OI_R}{2} \quad (1)$$

$$IR_{\text{Csize}} = \frac{OI_C}{2} \quad (2)$$

In this case, OI_R & OI_C denote the original image's column and row sizes, sub-Rsizeand IR_{Rsize} & IRC_{Size} denote the resized image's row and column sizes. Until an image's combined row and column sizes are fewer than 1000 pixels, this technique will continue to shrink it in an equal and opposite fashion. Due to its ability to decrease image dimensions while preserving the original shape of an OD-OC, this method lends strong support to the circular hough peak approach.

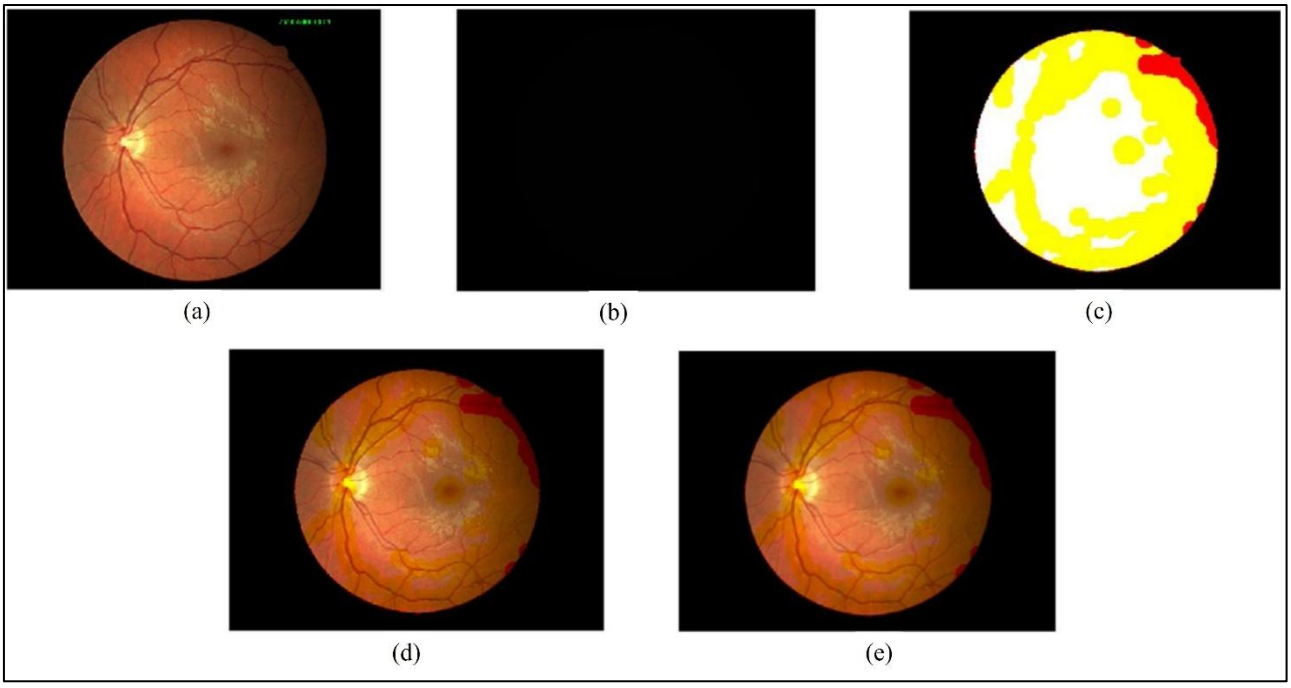

Figure 2: Processing steps for fundus image preprocessing. (a) RGB retinal image (b) Binarized mask (c) Morphological erosion (d) Element-wise multiplication (e) Gaussian filter.

3.3 Transferable attention-based U-Net

We propose TAU to solve the domain adaptation problem for OD and OC segmentation tasks. The suggested TAU is seen in Figure 4. It consists of Four modules for attention: a discriminator in the feature domain, another discriminator for the attention domain, and a segmentation network for the backbone. The backbone segmentation network employs a U-Net architecture and an attention module to determine the OD and OC areas. Using the two domain discriminators at the feature and image levels, the network learns adversarial training to generate domain-invariant representations.

3.3.1 Attention modules

The two main approaches to incorporating the attention mechanism into brain networks that are already in place presented in previous research were multiplicative attention and additive attention. Since the former may be expressed as matrix multiplication, it is theoretically faster and uses less memory. On the other hand, experimental results show that additive attention works well for high-dimensional features [3]. Thus, we incorporate the attention modules into the additive version and employ dynamic learnable coefficients for specifically designed OD and OC segmentation tasks, as seen in Figure 4, to expand on previous studies

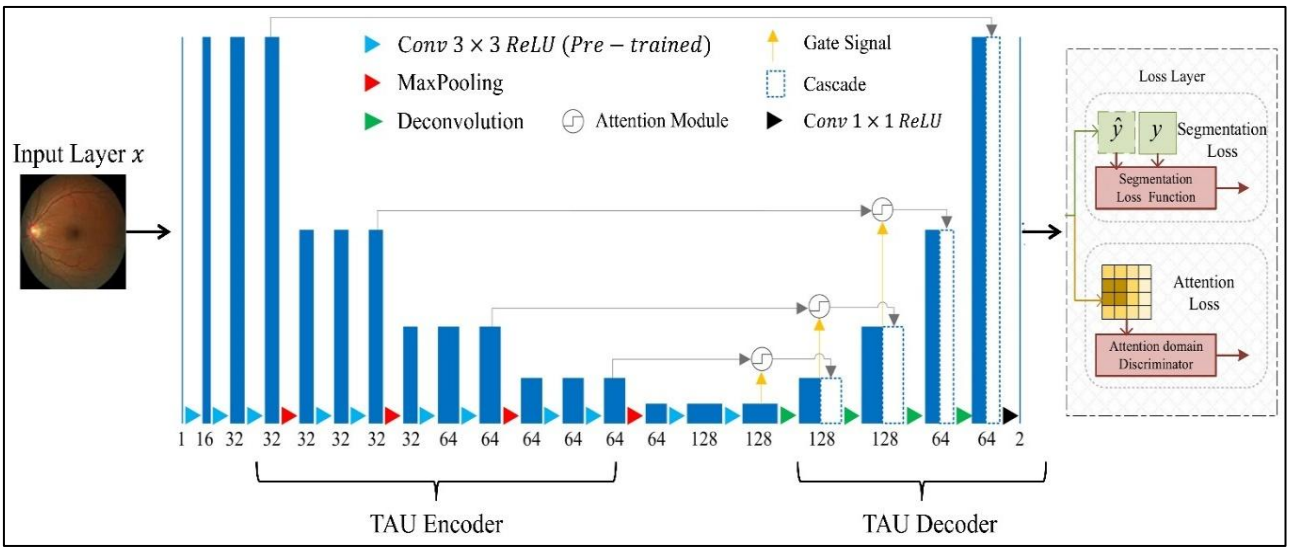

Figure 3: Transfer attention to U-Net architecture

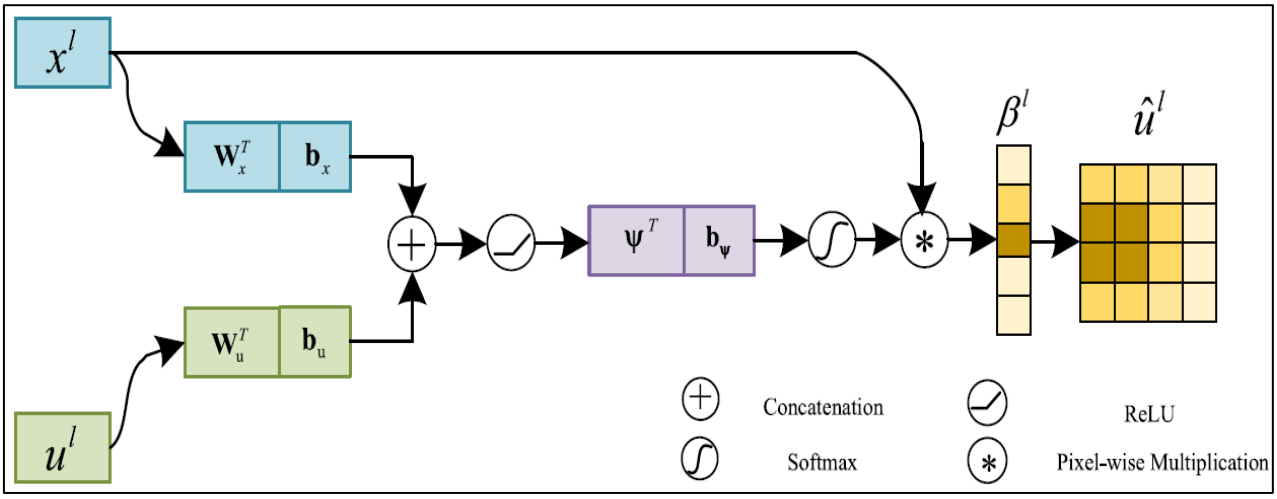

Figure 4: Attention module internal source

Since u^l is the up-sampled feature map that comes from the TAU decoder and x^l is the feature map from the l^{th} layer of the TAU encoder with F_l channels, where $l \in \{1,2,\ldots,L\}$, the attention module α^l is set up as follows:

$$\tilde{\alpha}^l(x^l,u^l;\boldsymbol{\theta}_{\text{attn}}) = \boldsymbol{\Psi}^T\left(\sigma_1([\mathbf{W}_x^T x^l + \mathbf{b}_x] \oplus [\mathbf{W}_u^T u^l + \mathbf{b}_u])\right) + \mathbf{b}_\Psi \quad (3)$$

$$\alpha^l = \beta^l\left(\sigma_2\left(\tilde{\alpha}^l(x^l,u^l;\boldsymbol{\theta}_{\text{attn}})\right)\right) \quad (4)$$

where $\sigma_1(\cdot)$ is a nonlinear activation function at the element level $\text{Re}LU(\cdot)$, and $\oplus$ represents a concatenation of the dimensions of the channel. Characteristics of the attention module $\boldsymbol{\theta}_{\text{attn}}$ consisting of linear transformation matrices $\mathbf{W}_x \in \mathbb{R}^{F_l \times F_{\text{int}}}$, $\mathbf{W}_u \in \mathbb{R}^{F_u \times F_{\text{int}}}$, and $\boldsymbol{\Psi}^T \in \mathbb{R}^{F_{\text{int}} \times 1}$, and bias terms $b_\Psi \in \mathbb{R}$, $\mathbf{b}_x \in \mathbb{R}^{F_{\text{int}}}$, and $\mathbf{b}_u \in \mathbb{R}^{F_{\text{int}}}$. $\sigma_2(\cdot)$ is a normalizing function, like a softmax operation, which guarantees that the attention module's weights add up to one. The matching characteristics in the provided images can be extracted using convolutional layers, which are regarded as high-level feature detectors [60]. However, certain features contribute more significantly to segmentation performance than others, depending on the retinal structure being analysed. A good illustration of this is the relative importance of the boundary feature compared to the texture feature in tasks involving OD and OC segmentation. Clearly defined boundaries make it easier to locate the OC and OD zones. To adjust the attention modules' weights, we create a set of learnable coefficients that may be adjusted over time, represented as β^l, where l is an element of the set $l \in \{1,2,\ldots,L\}$. Using pixel-wise multiplication to build the attended feature map $\hat{u}^l$, and assuming the feature map u^l has the shape $N = h \times w$ in the spatial dimension:

$$\hat{u}^l = \{\alpha_i^l u_i^l\}_{i=1}^N \quad (5)$$

TAU uses attention modules after the skip connections are concatenated before the encoder and decoder. The effectiveness of the attention mechanism is guaranteed in forward and backward transmission. Suppressed during

the backward pass are gradients that originate from the background (irrelevant) regions. Hence, only parameters in the lower levels pertinent to the current task are updated with a high likelihood. It is unnecessary to conduct a Monte Carlo sample with close scrutiny to update the parameters of the attention blocks using ordinary back-propagation.

3.3.2 Domain discriminators

Typically, the input eye fundus pictures utilized for OD and OC segmentation tasks are sourced from various datasets with disparate field-of-view, contrast, and resolution variations. One way to describe this phenomenon, called domain shift, is as the difference in the distributions $P(X_s) \neq P(X_t)$ of the data taken from both the unlabeled target domain datasets and the annotated source domain datasets, where X_s and X_t are the respective variables. Performance drops because of domain shift. Drawing on the work of [12], we develop discriminators for the attention domain and feature domain can assist in obtaining domain-invariant representations of photographs of the eye fundus at the image level and feature level, respectively, across different datasets. The segmentation network's encoded feature map is the input for one discriminator, whereas the input is the relevant segmentation result's attention map. for the other. To achieve an adversarial purpose, the backbone segmentation network uses domain-invariant characteristics to deceive the two domain discriminators and determine if the input image belongs to the dataset for the source or target domain.

3.3.3 Transferable attention U-Net model

As shown in Figure 3, the suggested TAU model incorporates four main sub-networks to address the issue in OD and OC tasks. The first subnetwork generates segmentation results for the provided ocular fundus image, a backbone segmentation network constructed using a modified version of the U-Net. Some attention modules make up the second sub-network; these modules find the OD and OC areas and create an attention map 'out' from the input image. The network's capacity to remain stable across different domains is guaranteed by using two additional modules: the discriminator in both the attention domain discriminator and the feature domain discriminator.

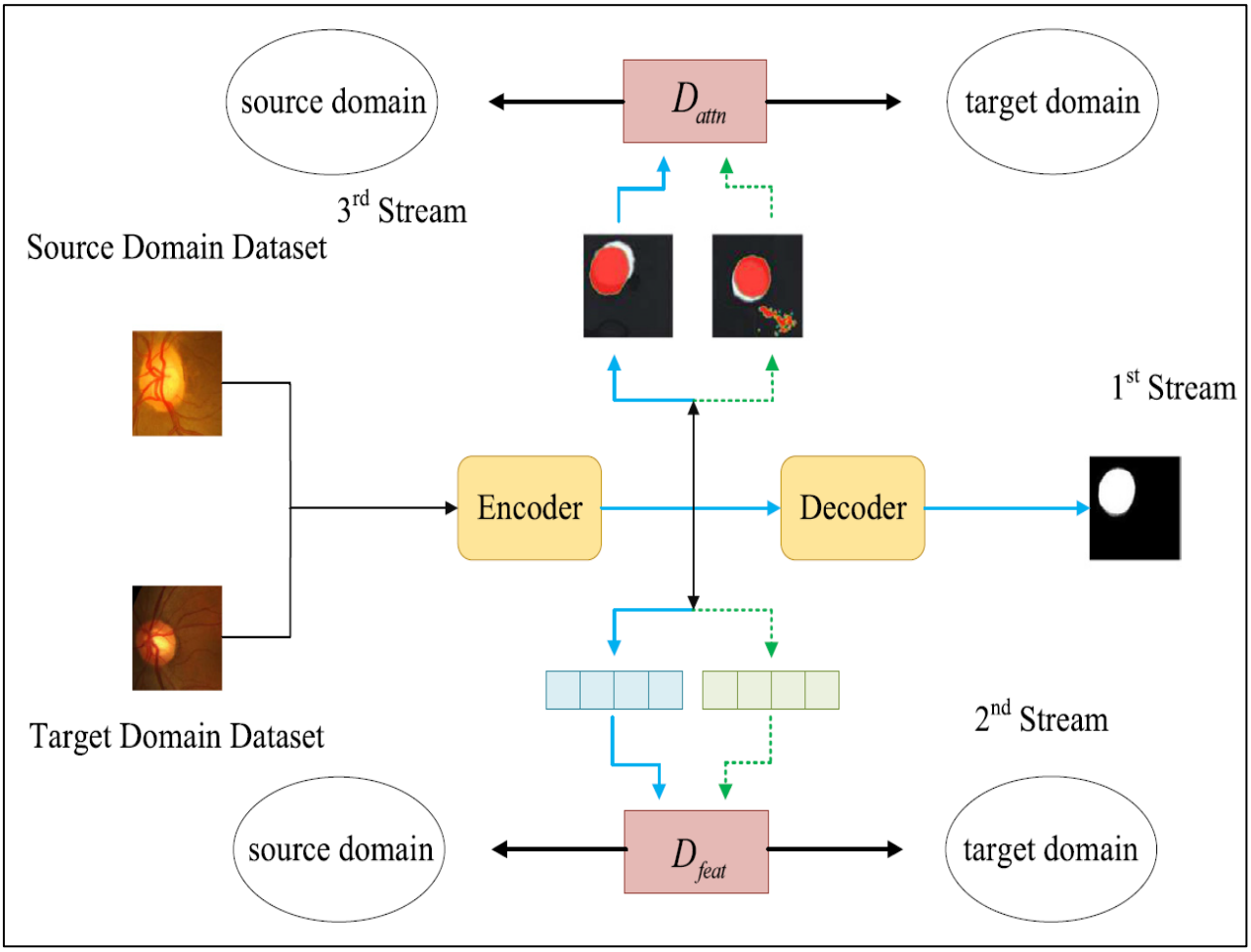

Figure 5: Training of transfer attention U-Net.

These domain discriminators communicate with the attention modules and the backbone segmentation network through adversarial training. In this paper, we formally suggest TAU. There are pixel-wise labelled eye

fundus images in the source domain datasets and unlabelled images in the target domain datasets. The ocular fundus photos with labels from the original domain Y_s, are denoted by X_s, and the photos from the target domain, X_t, are also utilized. By incorporating attention modules into the encoder $E(\cdot)$ and decoder $U(\cdot)$, TAU is defined. The discriminators for the feature and attention are denoted as $D_{\text{feat}}(\cdot)$ and $D_{\text{attn}}(\cdot)$, respectively.

3.3.4 Training

We need data from both the source and target domains to train the TAU model. As shown in Figure 5, the training stage consists of three distinct streams, each serving a unique role. There is a shared input to the TAU model, thus X_s and X_t are input to the model at the same time. We employ an encoder and a decoder to make the backbone segmentation network with attention modules visible. The black solid lines depict the data flow from the source domain to the destination domain. The solid blue lines show the source domain's data flow, whereas the green dashed lines represent the target domain's data flow. In the first stream, we have the training data for the network that will segment the backbone. Fundus images of the eye network take X_s, which are taken from the source domain dataset, as input and produces a segmentation result, Y_s, as well as an additional attention map, α_s^{out}. The formula for $\bar{Y}_S$ is as follows.

$$\bar{Y}_S = U(E(X_S; \boldsymbol{\theta}_E); \boldsymbol{\theta}_U) \tag{6}$$

The segmentation loss Y_s is computed using Eq. (8) and labels Y_s, where $\boldsymbol{\theta}_E$ and $\boldsymbol{\theta}_U$ denote the parameters of the encoder and decoder, respectively.

$$\mathcal{L}_{\text{seg}}(\bar{Y}_S, Y_S; \boldsymbol{\theta}_E, \boldsymbol{\theta}_U) = \text{Dice}(\bar{Y}_S, Y_S) \tag{7}$$

The network gets input images X_s from the target domain dataset and uses them to generate an attention map α_t^{out} that may be used for further training. An adversarial technique is used in the second stream to train adaptive abilities. In TAU, the hidden representations $E(X_s)$ and $E(X_t)$ are encoded when we input X_s and X_t, respectively. The encoded representations lose their domain-invariant properties due to the domain shift. The objective is to maximize the degree of similarity between $E(X_s)$ and $E(X_t)$. The feature domain discriminator is used to train the encoder $E(\cdot)$. The feature maps the encoder produces comprise the first part of the training data set, while the domain labels comprise the second element. We construct pairs $(E(X_S), 0)_{\text{feat}}$ and $(E(X_t), 1)_{\text{feat}}$. Our discriminator with feature domain does a regular supervised classification job, assuming that '0' denotes the source domain and '1' denotes the target domain. Distinguishing between the domains of the hidden representations, $E(X_s)$ and $E(X_t)$, is its responsibility. The feature domain discriminator (D_{feat}) is designed to distinguish between source and target domain feature representations extracted by the encoder. It consists of three convolutional layers with kernel size 3×3, each followed by batch normalization and ReLU activation. The convolutional layers are followed by a fully connected layer and a sigmoid activation function to perform binary domain classification. The discriminator is trained using binary cross-entropy loss. Similarly, the attention domain discriminator (D_{attn}) operates on attention-enhanced feature maps and follows a lightweight architecture comprising two convolutional layers with 3×3 kernels and ReLU activation, followed by a fully connected layer with sigmoid activation. This discriminator enforces domain-invariant attention representations and assists in reducing domain shift between dataset. The cross-entropy most commonly seen is the feature domain discriminator's loss.

$$\mathcal{L}_{\text{feat}}(D_{\text{feat}}(\cdot); \theta_{\text{feat}}) = -\sum_i y_i^d \ln \hat{y}_i^d + (1 - y_i^d) \ln(1 - \hat{y}_i^d) \tag{8}$$

Where $y_i^d \in \{0,1\}$, and $\hat{y}_i^d$ represents the results of the domain label classification using the feature domain discriminator $D(\cdot)$. We use an adversarial training approach for our network, as suggested by [31]. We learn the encoder's parameters $\boldsymbol{\theta}_E$ to maximize the feature domain discriminator's loss and The discriminator for feature domains parameters $\boldsymbol{\theta}_{\text{feat}}$ to minimize its classification loss in order to extract the features that are domain-invariant. The settings of the encoder $E(\cdot)$ and feature domain discriminator $D_{\text{feat}}(\cdot)$ are updated using back-propagation while they engage in a minimax game. Our attention domain discriminator is trained using the adversarial approach in the third stream, just like in the second stream. The forms of the input data are $(\alpha_s^{\text{out}}, 0)_{\text{attn}}$ and $(\alpha_t^{\text{out}}, 1)_{\text{attn}}$, where the domain labels are the second element, and the focus The first component is the attention mappings of the U-Net model. The notion that "1" is the same as the target domain and "0" indicates the source domain one we use. Similarly, the attention domain discriminator uses adversarial training with back-propagation to update its parameters $\boldsymbol{\theta}_{\text{attn}}$, which has a typical classification loss $\mathcal{L}_{\text{attn}}(\alpha^{\text{out}}; \boldsymbol{\theta}_{\text{attn}})$. The suggested TAU model's loss function is formally expressed as:

$$\mathcal{L}(X_s, X_t; \boldsymbol{\theta}_E, \boldsymbol{\theta}_U, \boldsymbol{\theta}_{\text{feat}}, \boldsymbol{\theta}_{\text{attn}}) = \mathcal{L}_{\text{seg}}(\bar{Y}_S, Y_S; \boldsymbol{\theta}_E, \boldsymbol{\theta}_U) + \delta \mathcal{L}_{\text{feat}}(D_{\text{feat}}(\cdot); \boldsymbol{\theta}_{\text{feat}}) + \gamma \mathcal{L}_{\text{attn}}(\alpha^{\text{out}}; \boldsymbol{\theta}_{\text{attn}}) \tag{9}$$

Here, the loss balancing hyperparameters were fixed as δ = 0.5 and γ = 0.5. These values were empirically selected to provide an equal contribution from the feature and attention domain discriminator losses while preserving the dominance of the segmentation loss during training.

3.4 Particle swarm optimization (PSO)

Particle Swarm Optimization is a meta-heuristic search method is based on the intelligence of swarms and is inspired by the way birds find food. In this approach, particles represent individual birds and "move" across a complex search environment, "changing" according to your and your neighbors' past experiences. The piece, like "an individual element in a flock," is one approach that could work. By analyzing local and global data, PSO finds the optimal solution using the particle's velocities and a fitness function. The efficacy of PSO, especially when optimizing parameters in high-dimensional search spaces, was the deciding factor in its selection over other optimization techniques. PSO converges fast, requires fewer parameter settings, and balances exploration and exploitation. It was chosen for its simplicity of implementation and its track record of success in related sectors. Algorithm 1 lays out the steps for the PSO. The algorithm can be defined by the following equations: Eq. (10) to update the position and Eq. (2) to update the velocity.

$$\mathrm{pi}(t+1) = \mathrm{pi}(t) + \mathrm{xi}(t+1) \tag{10}$$

In Equation (10), $\mathrm{Pi}(t)$ is the position of particle i in the search space at iteration t. The velocity can change the particle's position, $xi(t)$.

$$\mathrm{xi}(t+1) = \mathrm{xi}(t)\omega + c_1 r_1[y_i - \mathrm{pi}(t)] + c_2 r_2[\hat{y} - \mathrm{pi}(t) \tag{11}$$

In Equation 11, i is the particle's descriptor and x is the velocity. The cognitive component is described by a parameter c_1, whereas the social component is indicated by the parameter c_2. p_{best}, the optimal particle position, and g_{best}, the optimal global position, is defined by y_i and y, respectively. The inertia weights, r_1 and r_2, stand for the random values between 0 and 1.

$$\mathrm{pbest}_i(t+1) = \begin{cases} \mathrm{pbest}_i(t) & \text{if} \quad f(\mathrm{pbest}_i(t)) \le f(p(t+1)) \\ \mathrm{pi}(t+1) & \text{if} \quad f(\mathrm{pbest}_i(t)) > f(\mathrm{pi}(t+1)) \end{cases} \tag{12}$$

$$\mathrm{gbest}(t+1) = \max\{f(y), f(\mathrm{gbest}(t)) \tag{13}$$

where $y \in \{\mathrm{pbest}_0(t), \mathrm{pbest}_1(t), \ldots, \mathrm{pbest}_n(t)\}$.

Algorithm 1: PSO

Require: Objective function $g: \mathbb{R}^n \to \mathcal{R}$
Require: Hyperparameters n, Fitness function F
Ensure: Optimal solution λ*
Initialization: Initialize the swarm with s particles;
for i=0 to s do
  Initialize position λ_i randomly from $U(bl, bu)$;
  Set best-known position λ_i
  if $g(\lambda_i) > g(\lambda^s)$ then
    Set λ^s as λ_i
  end
  Initialize velocity v_i randomly from U;
end
Evaluation loop:
While termination condition is not met do
  for i=0 to s do
    Generate random numbers rp and rg from $U(0,1)$;
    Update velocity v_i using:

$$v_i \leftarrow wv_i + \emptyset_{prp}\big(g(\lambda_i^*) - g(\lambda_i)\big) + \emptyset_{grg}(g(\lambda^s) - g(\lambda_i);$$

Update position λ_i accordingly
$\lambda_i \leftarrow \lambda_i + v_i$;
if $g(\lambda_i) > g(\lambda^s)$ then
  Set λ^s as λ_i;
  end
 end
end
return λ^s

Assuming are n particles in the swarm, an objective function f is applied to determine particle fitness in a maximization task. For iteration t, the personal best value is updated using equation (12), and the global best value is updated using equation (13). Particle Swarm Optimization was selected due to its fast convergence, low computational overhead, and ability to efficiently explore complex, non-convex search spaces. Compared to grid or random search, PSO requires fewer evaluations, and unlike Bayesian optimization, it does not rely on surrogate models, making it well suited for optimizing deep learning–based segmentation frameworks.

3.4.1 Optimal Selection of Hyper-Parameters via PSO Algorithm

This part details how we optimized the proposed architectural settings using the PSO algorithm. To begin, we need to prioritize which TAU-Net parameters are crucial for optimal performance, and then we may utilize the PSO method to identify those parameters. The performance of a TAU-Net was evaluated in an experimental study with manually modified parameters, and then the optimal parameters were selected. Finding the optimal architectures is the goal, as, as previously stated, different CNN parameter choices yield different possible results for the same task. The study's optimization criteria included the following variables:

- The extent to which convolutional layers use filters.
- The filter sizes used in convolutional layers.
- The maximum pooling layer's pool size.
- The stride size utilized by the max pooling layer.

Here is an example of Algorithm 1 that demonstrates the PSO algorithm and its function in our methodology: $g: \mathbb{R}^n \to \mathbb{R}$ is an objective function defined by the given algorithm, n is the number of hyper-parameters, and F is the fitness function that determines how well the trained CNN model detects. In this study, we want to discover a solution λ such that, for any λ in the set of all hyperparameters χ, the fitness function g(λ) is greater than or equal to g(λ). A swarm of particles, standing in for the hyper-parameter values, undergoes evolution in PSO. The movement of each particle in the swarm is affected by their velocity, $v_i \in \mathbb{R}^n$, and their position in the search space, $\lambda_i \in \mathbb{R}^n$. Assign λ^S to the particle's global position in the swarm and λ_i^* to its local position. The suggested PSO algorithm can readily adjust to any new CNN design since it is not dependent on the optimized CNN. "Swarm Initialization" and "Swarm Evaluation" are the two primary steps of our PSO algorithm implementation. In the end, the output is the best swarm position (λ^S). The number of hyper-parameters used to determine PSO's efficiency is represented as:

$$T_{PSO} = s \cdot g(\lambda_k) \cdot G_{\max} \quad (14)$$

with $G_{\max}$ and s being constants, the main source of time complexity is the evaluation of $f(\lambda_n)$, which grows linearly with the parameter n. The entire framework of our research utilized in this investigation is depicted in Figure 4. To optimize its parameters, the TAU-Net first employs the PSO algorithm. This procedure creates particles by initializing the PSO with the execution settings. Because there is a parameter for the position of each particle that needs to be employed in the proposed TAU-Net architecture, each solution represents a completed M-Net training period. Our TAU-Net architecture is structured in a way that allows it to be compact and versatile. The classification process begins with a block that includes convolutional and max pooling layers, then moves on to a SoftMax activation function. In the proposed TAU-PSO framework, Particle Swarm Optimization (PSO) is employed to optimize the network configuration using fixed architectural settings. The convolutional layers use 64 filters with a kernel size of 3×3, max-pooling layers use a pooling size of 2×2 with a stride of 2, and up sampling layers follow the same configuration. These values are kept constant across all experiments to maintain architectural consistency and computational efficiency. PSO is applied to fine-tune the network parameters during training rather than altering the structural dimensions of the model. The PSO algorithm is configured with 30 particles and 50 iterations. The cognitive and social acceleration coefficients are set to $c_1 = 2.0$ and $c_2 = 2.0$, respectively, and the inertia weight w is linearly decreased from 0.9 to 0.4.

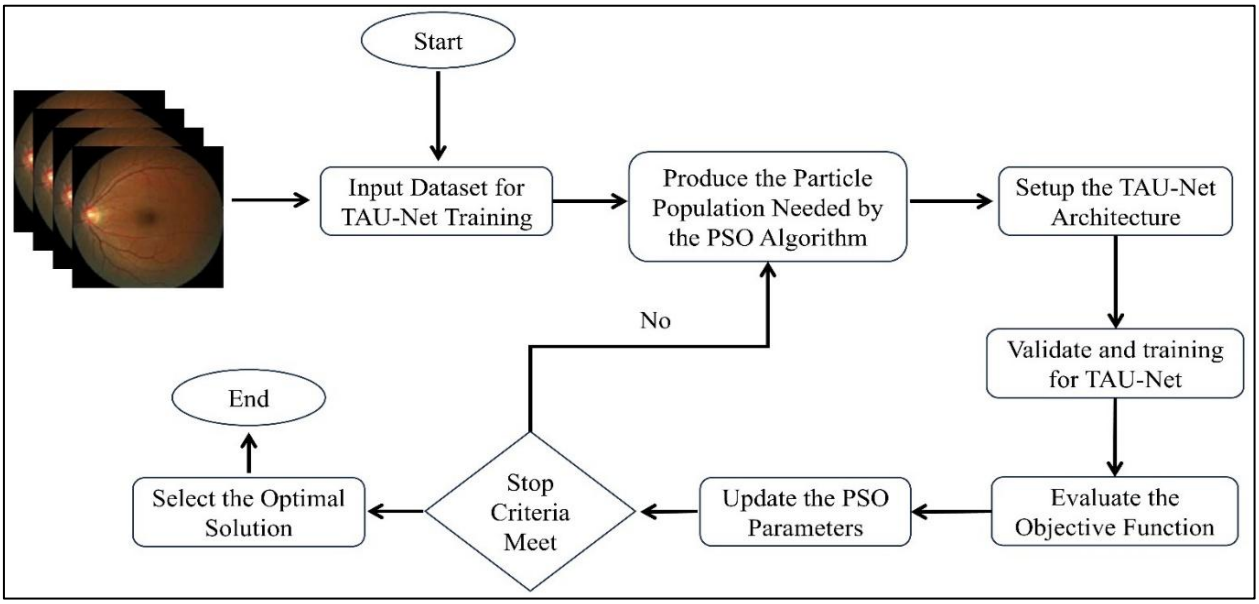

Figure 6: Flowchart of the proposed TAU-Net with PSO

Using the PSO algorithm shown in Figure 6 as a flow diagram, the following are the exact methods for optimizing the TAU-PSO with PSO. The proposed TAU-PSO model has a computational cost comparable to or lower than the standard U-Net. This is achieved by replacing conventional convolutional blocks with structured dropout blocks and lightweight attention modules, which reduce redundant computations and parameter count. Although PSO introduces a small additional overhead during training, it is used only for optimization and does not affect inference time.

4 Results

To determine the effectiveness of the suggested method, it was tested against several performance measures detailed in this section. Also, various bio-inspired optimization techniques and traditional neural network topologies are tested alongside the proposed method. The implementation is executed in MATLAB on a desktop PC with Windows 11, 64-bit OS, Intel i7 core CPU, 8 GB RAM, and an NVIDIA 8 GB GeForce 2050Ti GPU. The experimental training parameters are displayed in Figure 7 and Table 2. The precision, specificity, and sensitivity metrics, as well as DRISHTI-GS1, RIM-ONE, and REFUGE fundus images, are used to assess the effectiveness of the suggested model. The RIM-ONE, REFUGE, and DRISHTI-GS1 databases were also used to compare performance with other neural network classifiers and current optimization methodologies. As a preliminary step in our process, we first pad a color retinal fundus image I to a size of H × W with four predefined boundaries.

Using the feed-forward zero-padded images as input, the OC-OD segmentation network is trained. Random horizontal, vertical, and diagonal flips and rotations were among the data augmentation strategies applied to the dataset training images. This process improves the model's adaptability by developing numerous versions of the original retinal images. The model is trained to be resilient to various retinal circumstances by combining these modifications. Finding a happy medium between adding variety and keeping things realistic requires strategically applying pertinent enhancements and fine-tuning settings. To accurately evaluate performance, it is necessary to maintain consistency in the validation set and not alter it, even while data augmentation improves the model's generalizability. Simply put, our model can precisely partition OC and OD in various clinical situations due to data augmentation, which allows it to learn from many retinal images. We restore the segmented cup and disc images to their original sizes before comparing our findings to the expert annotations from the datasets. If the connected I_o(x) location has a more considerable poise value than the specific threshold T, specifically if T < I_o(x) (we have set T = 0.5), then the I_o(x) location is authenticated.

4.1 Performance evaluation

The proposed framework for partitioning OD and OC based on ground truth has been assessed using the following metrics: Accuracy, Sensitivity, Dice Coefficient, and Intersection-over-Union (IoU).

$$DSC = \frac{2.TP}{2.TP+FP+FN} \quad (15)$$

$$Jaccard\ Index\ (JI) = \frac{TP}{TP+FP+FN} \quad (16)$$

$$Precision = \frac{TP}{TP+FP} \quad (17)$$

$$Recall = \frac{TP}{TP+FN} \quad (18)$$

$$Specificity = \frac{TN}{TN+FP} \quad (19)$$

where TP, FP, TN, and FN are correct, incorrect, and non-correct, respectively. We have also determined the ground truth-segmented region overlapping error (E) and balance accuracy (A). An expression for the parameter A is:

$$A = \frac{1}{2}(\text{Sen} + Sp) \quad (20)$$

where Sen and Sp are specificity and sensitivity. One way to express the overlapping error is as:

$$E = 1 - \frac{\text{Area}\ (S \cup G)}{\text{Area}\ (S \cap G)} \quad (21)$$

where G is the ground truth and S is the segmented area.

Figure 8 and table 3 The results achieved by the suggested algorithm are contrasted with those of state-of-the-art methods once they have been compiled. Here, we put the proposed approach to the test using three open-source datasets: DRISHTI-GS, RIM-ONE, and REFUGE. When comparing this method to others at the field's cutting edge, we look at how well they function through extensive networks or how much processing and localization (ROI extraction) they do before segmentation.

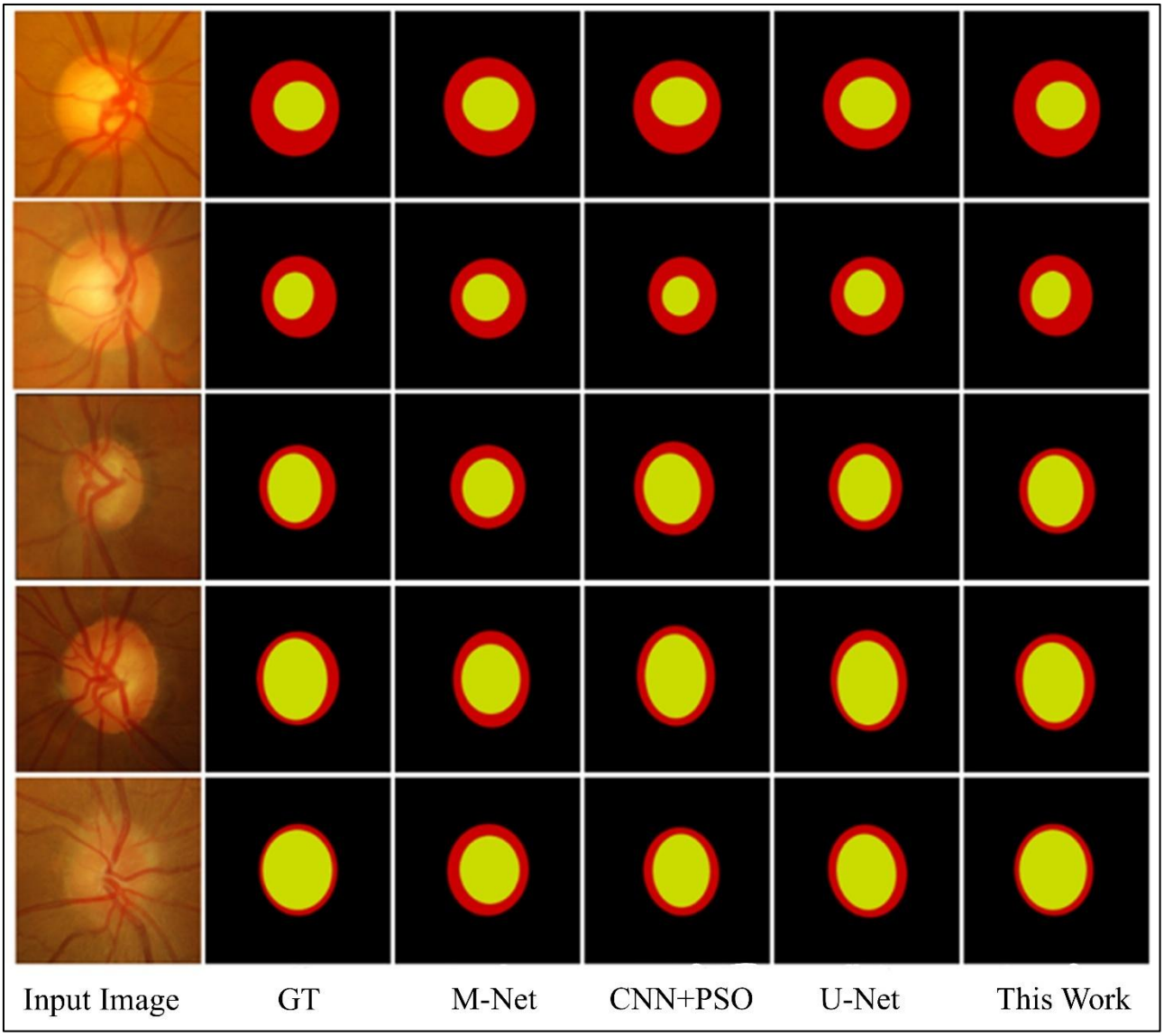

Figure 7: Segmented results of DRISHTI-GS1, RIM-ONE, and REFUGE datasets

Table 2: Functional and training parameter setup

Training Parameters	**Values**
Batch Size	64
Learning Rate	0.001
Epoch	80
Optimizer	PSO
Loss Function	Cross entropy
Range of Zoom	4%
Rotation in Range	$0 - 180^{o}$

Table 3: The suggested approach and conventional neural network techniques are contrasted

Dataset	**Method**	**Evaluation Metrics Cup/Disc**					
		Accuracy (%)	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**IoU (%)**	**Dice Score (%)**
DRISHTI-GS, RIM-ONE, and REFUGE	NFN^{+} [19]	93.8/94.3	82.3/80.2	91.0/92.3	91.3/82.3	×	×
	SDU-Net [31]	95.7/95.3	×	90.3/91.2	90.0/90.1	×	×
	Backbone [14]	96.8/94.2	89.2/91.2	91.3/93.2	×	×	×
	DeepLabv3+ [34]	×	×	×	×	89.6/90.1	90.9/91.2

	DU-Net [35]	94.5/96.2	88.2/80.1	98.0/97.2	91.7/87.2	96.3/93.2	88.6/90.3
	DFU-Net [36]	94.5/94.3	88.6/85.2	97.0/95.2	91.9/89.4	96.0/93.3	90.0/91.2
	This Work	**98.2/97.1**	**96.6/96.0**	**97.2/96.5**	**90.3/94.7**	**96.2/94.3**	**96.4/96.8**

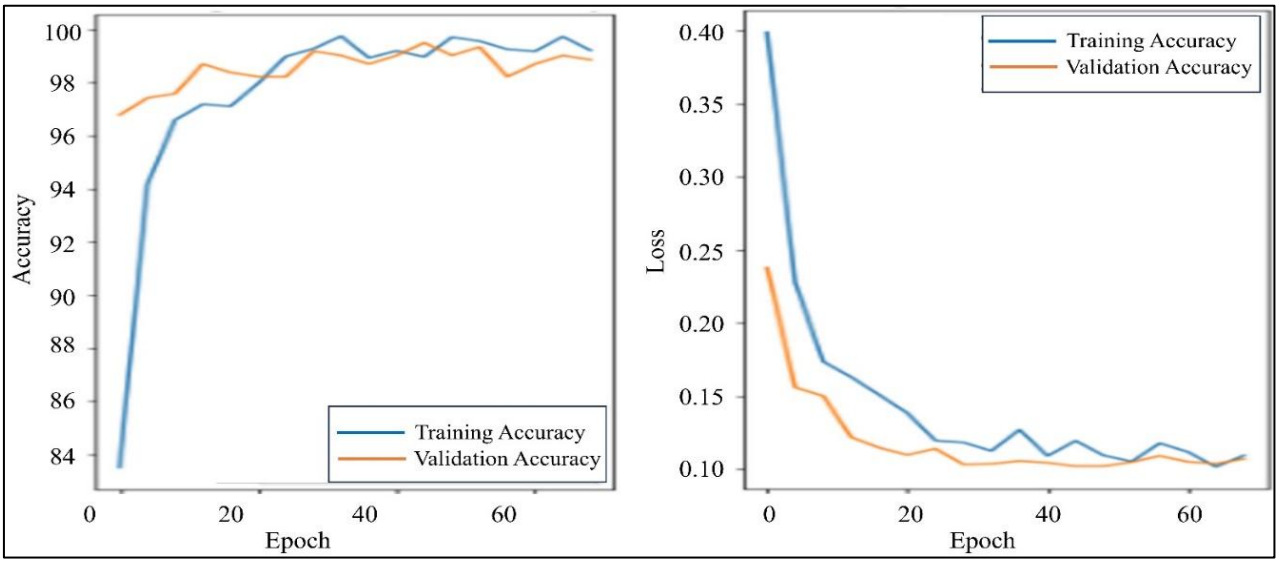

Figure 8: Accuracy vs loss graph of the proposed model

The outcomes of object detection and classification utilizing the RIM-ONE, REFUGE, and DRISHTI-GS datasets are shown in Figure 8, respectively. The evaluation metrics for the cup attained an accuracy of 98.2%, Precision of 96.6%, recall of 97.2%, and dice score of 96.4%. The evaluation metrics for disc attained accuracy of 97.1%, precision of 96.0%, recall of 96.5%, and dice score of 96.8%. The outcomes can be seen in Table 3. On the DRISHTI-GS, RIM-ONE, and REFUGE, our suggested technique hit the most excellent dice value for the cup and disc at 96,43% and the equivalent dice value for the disc at 96.85%. Compared to previous approaches, ours has attained state-of-the-art accuracy values on discs and cups. On three datasets, Table 2 compares these methods to the suggested model. We have also evaluated our model's sensitivity and specificity. The proposed TAU-PSO model shows low sensitivity to PSO hyperparameters. Experimental analysis with varying swarm sizes, inertia weights, and acceleration coefficients demonstrates stable segmentation performance across DRISHTI-GS1, RIM-ONE, and REFUGE datasets. Minor variations mainly affect convergence speed rather than final accuracy, indicating robustness of the proposed framework.

5 Discussion

The proposed TAU-PSO framework consistently outperforms prior methods listed in Table 1, as demonstrated by quantitative results in Table 3. This improvement can be attributed to several architectural and training innovations. First, the attention modules enable the network to focus on salient regions of the optic disc and cup, improving feature representation. Second, the inclusion of feature and attention domain discriminators through adversarial training ensures that the learned representations are transferable across datasets, enhancing cross-domain generalizability—an issue highlighted as a limitation in several SOTA works. Third, Particle Swarm Optimization (PSO) was employed for hyperparameter tuning, enabling optimal selection of filter sizes, dropout rates, and other critical parameters without manual trial-and-error. These modifications collectively contribute to higher Dice and IoU scores for both cup and disc segmentation. Regarding trade-offs, while TAU-PSO incorporates additional attention and adversarial components, the network remains lightweight compared to other deep architectures such as ResNet- or SegR-Net-based models, ensuring competitive inference time suitable for practical clinical applications

6 Conclusions

Automatic glaucoma identification relies on two main modules: OD and OC segmentation. Traditional approaches, however, do not simply transfer to new datasets with high performance due to the substantial differences in the distribution and visual appearance of the various sources of ocular fundus images. An issue with domain adaptation in object detection and object classification was introduced in this work. The TAU model was suggested to solve this issue and achieve satisfactory performance across datasets without costly annotations. We ran OD and OC segmentation tests to ensure the approach worked using the DRISHTIGS, RIM-ONE v3, and REFUGE datasets. According to the experimental results, better segmentation performance and CDR MAE in cross-dataset settings were attained by integrating TAU compared to the baseline approaches. Future work will focus on expanding the suggested model to address the issue with third-party datasets that were

previously noted. The model has the potential to be developed into an automated glaucoma detection system with the inclusion of an additional classifier.

References

[1] Kumar, P. R., Shilpa, B., Jha, R. K., & Chellibouina, V. S. (2024). Spatial attention U-Net model with Harris hawks optimization for retinal blood vessel and optic disc segmentation in fundus images. International Ophthalmology, 44(1), 359.

[2] Chen, N., & Lv, X. (2024). Research on segmentation model of optic disc and optic cup in fundus. BMC ophthalmology, 24(1), 273.

[3] Panahi, A., Askari Moghadam, R., Tarvirdizadeh, B., & Madani, K. (2024). Simplified U-Net as a deep learning intelligent medical assistive tool in glaucoma detection. Evolutionary Intelligence, 17(2), 1023-1034.

[4] Xiao, Y., Zhao, J., Yu, Y., Ding, X., Liu, S., Bao, W., ... & Zhou, X. (2024). SimpleCNN-UNet: An optic disc image segmentation network based on efficient small-kernel convolutions. Expert Systems with Applications, 256, 124935.

[5] Kako, N. A., Abdulazeez, A. M., & Abdulqader, D. N. (2024). Multi-Label Deep Learning for Comprehensive Optic Nerve Head Segmentation through Data of Fundus Images. Heliyon.

[6] Das, S., Chakraborty, S., Mishra, M., & Majumder, S. (2024). Assessment of retinal blood vessel segmentation using U-Net model: A deep learning approach. Franklin Open, 8, 100143.

[7] Badian, Reza A. Schematic sectioning approaches for corneal and retinal surfaces used in ophthalmology and vision-related clinical practice and research. Experimental Eye Research 230: 109442; 2023.

[8] Yu, Dao, Su, Er, Mehnert, Andrew, Yu, Paula K., Cringle, Stephen J., Morgan, William H., and Ian L. McAllister. "Endothelial contraction of retinal veins." Experimental Eye Research 228; 109386: 2023.

[9] Kumar, K.S., Singh, N.P. Analysis of retinal blood vessel segmentation techniques: a systematic survey. Multimed Tools Appl 82, 7679–7733, 2023.

[10] Yang, Y., Wan, W., Huang, S. et al. RADCU-Net: residual attention and dual-supervision cascaded U-Net for retinal blood vessel segmentation. Int. J. Mach. Learn. & Cyber. 14, 1605–1620, 2023

[11] Iqbal, Shahzaib, Naveed, Khuram, Naqvi, Syed S., Naveed, Asim, and Tariq M. Khan. "Robust retinal blood vessel segmentation using a patch-based statistical adaptive multi-scale line detector." Digital Signal Processing 139, (2023): 104075.

[12] Dayana, A.M., and Emmanuel, W.R.S 2022 Deep learning enabled optimized feature selection and classification for grading diabetic retinopathy severity in the fundus image, *Neural Comput & Applic* 34, pp. 18663–18683.

[13] Jiang Y, Wang F, Gao J, and Cao S 2020 Multi-Path Recurrent U-Net Segmentation of Retinal Fundus Image, *Applied Sciences* 10 11 3777.

[14] Karaali, A., Dahyot, R., Sexton, and D.J. 2022 DR-VNet: Retinal Vessel Segmentation via Dense Residual Unet, *Pattern Recognition and Artificial Intelligence* 13363, pp 198–210.

[15] O. O. Sule 2022 A Survey of Deep Learning for Retinal Blood Vessel Segmentation Methods: Taxonomy, Trends, Challenges and Future Directions, *IEEE Access* 10, pp. 38202-38236.

[16] A. A. Babu, V. Jegathesan, D. J. David and K. S. Suriya 2022 Retinal Blood Vessels Segmentation Using Deep Learning Model-A Review, *6th International Conference on Devices, Circuits and Systems (ICDCS)*, pp. 375-379.

[17] Zhou Y, Wang B, and Huang 2021 A benchmark for studying diabetic retinopathy: segmentation, grading, and transferability, *IEEE Trans Med Imaging* 40 3, pp.818–828.

[18] Kar SS, and Maity SP 2018 Automatic detection of retinal lesions for screening of diabetic retinopathy, *IEEE Trans Biomed Eng* 65 3, pp. 608–618.

[19] Yicheng Wu, Yong Xia, Yang Song, Yanning Zhang, and Weidong Cai 2020 NFN+:A novel network followed network for retinal vessel segmentation, *Neural Networks* 126, pp. 153-162.

[20] S. Alex David, C. Mahesh, V. Dhilip Kumar, Kemal Polat, Adi Alhudhaif, and Majid Nour. Retinal Blood Vessels and Optic Disc Segmentation Using U-Net. Mathematical Problems in Engineering, pp.11, 2022.

[21] E. Chakour et al., 2022 Blood vessel segmentation of retinal fundus images using dynamic preprocessing and mathematical morphology, *8th International Conference on Control, Decision and Information Technologies (CoDIT)*, pp. 1473-1478.

[22] Abdel-Basset M, Ding W, and El-Shahat D 2021 A hybrid Harris Hawks optimization algorithm with simulated annealing for feature selection, *Artif Intell Review* 54 1, pp.593–637.

[23] Yang, X., Li, Z., Guo, Y. et al. 2022 DCU-net: a deformable convolutional neural network based on cascade U-net for retinal vessel segmentation, *Multimed Tools Appl* 81, pp. 15593–15607.

[24] L. Jin 2020 3AU-Net: Triple Attention U-Net for Retinal Vessel Segmentation, *2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT, 2020)*, pp. 612-615.

[25] Wang, C., Zhao, Z. and Yu, Y 2021 Fine retinal vessel segmentation by combining Nest U-net and patch-learning, *Soft Comput* 25, pp. 5519–5532.

[26] X. Zuo 2021 Novel Retinal Vessel Segmentation Method Based on U-net and FPN, *International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pp. 686-691.

[27] O. Ramos-Soto et al., 2021 An efficient retinal blood vessel segmentation in eye fundus images by using optimized top-hat and homomorphic filtering, *Comput. Methods Programs Biomed.* 201, pp. 105949.

[28] Y. Yi, C. Guo, Hu Y, Zhou W and Wang W 2022 BCR-UNet: Bi-directional ConvLSTM residual U-Net for retinal blood vessel segmentation, *Front. Public Health* 10, 1056226.

[29] Jiang Y, Liang J, Cheng T, Lin X, Zhang Y, Dong J 2022 MTPA_Unet: Multi-Scale Transformer-Position Attention Retinal Vessel Segmentation Network Joint Transformer and CNN, Sensors 22 12, 4592.

[30] You, A., Kim, J.K., Ryu, I.H. et al. Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey. Eye and vision, vol. 9, Issue 6, pp. 1-19, 2022.

[31] C. Guo, M. Szemenyei, Y. Pei, Y. Yi and W. Zhou 2019 SD-Unet: A Structured Dropout U-Net for Retinal Vessel Segmentation, 19th International Conference on Bioinformatics and Bioengineering (BIBE), pp. 439-444, 2019.

[32] B. Wang, S. Wang, S. Qiu, W. Wei, H. Wang and H. He 2021 CSU-Net: A Context Spatial U-Net for Accurate Blood Vessel Segmentation in Fundus Images, IEEE Journal of Biomedical and Health Informatics 25 4, pp. 1128-1138.

[33] Vaibhav V. Kamble, Rajendra and D. Kokate 2020 Automated diabetic retinopathy detection using radial basis function, Procedia Computer Science 167, pp. 799-808.

[34] Sreng S, Maneerat N, Hamamoto K, Win KY. Deep Learning for Optic Disc Segmentation and Glaucoma Diagnosis on Retinal Images. Applied Sciences. 2020; 10(14):4916.

[35] Jin,Q.,Meng,Z.,Pham,T.D.,Chen,Q.,Wei,L.,Su,R. :Dunet:Adeformable network for retinal vessel segmentation. Knowledge-Based Systems178,149–162(2019).

[36] C. Wu, Y. Cheng, W. Li, Z. Yang and Z. Lu, "DFUNET: A Residual Network for Retinal Vessel," 2021 16th International Conference on Computer Science & Education (ICCSE), 2021, pp. 638-642.

[37] Memari, N., Ramli, A.R., Saripan, M.I.B. et al. 2019 Retinal Blood Vessel Segmentation by Using Matched Filtering and Fuzzy C-means Clustering with Integrated Level Set Method for Diabetic Retinopathy Assessment, J. Med. Biol. Eng. 39, pp. 713–731

[38] Yi, J., Ran, Y., & Yang, G. (2022). Particle swarm optimization-based approach for optic disc segmentation. Entropy, 24(6), 796.

[39] Chen, Y., Bai, Y., & Zhang, Y. (2024). Optic disc and cup segmentation for glaucoma detection using Attention U-Net incorporating residual mechanism. PeerJ Computer Science, 10, e1941.