IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Glaucoma Detection using Self-Attention GAN (SAGAN)

Mrs.R.Anitha, (Author1)
Department of Computer Science and Applications,
SRM Institute of Science and /technology,
Chennai, India – 600089

Mrs.K.Pandimeena (Author2)
Department of Computer Science and Applications,
SRM Institute of Science and /technology,
Chennai, India – 600089

Abstract: Glaucoma is a group of eye conditions that damage the optic nerve. The optic nerve sends visual information from eye to brain and is vital for good vision.As this nerve gradually deteriorates, blind spots develop in vision and the nerve damage is usually related to increased pressure in the eye.

Elevated eye pressure happens as the result of a buildup of fluid that flows throughout the inside of the eye. This fluid is also known as the aqueous humour. It usually drains through a tissue located at the angle where the iris and cornea meet. This tissue also is called the trabecular meshwork. The cornea is important to vision because it lets light into the eye. When the eye makes too much fluid or the drainage system doesn't work properly, eye pressure may increase.

Self-Attention Generative Adversarial Networks (SAGAN) is used in Glaucoma Detection that involves the capabilities of self-attention mechanisms to enhance feature extraction in medical imaging.

Keywords— SAGAN, Glaucoma.

I. Introduction

Self-Attention Generative Adversarial Networks (SAGAN) is primarily a specific architecture that incorporates self-attention mechanisms into the traditional GAN framework. However, variations and extensions of SAGAN can be categorized based on their applications and modifications. Here are some notable types and variations:

1. Standard SAGAN-The original model that integrates self-attention layers to improve feature learning and focus on relevant parts of the image, enhancing both the generator and discriminator.

2. Conditional SAGAN-This variation conditions the GAN on additional information (e.g., class labels, textual descriptions) to generate images that align with specific categories, useful in applications like image-to-image translation.

3. Progressive SAGAN-Incorporates a progressive training approach where images are generated at low resolutions initially and gradually increase in resolution. This can lead to higher-quality outputs and stability during training.

4. Semi-Supervised SAGAN-Utilizes both labelled and unlabelled data for training. It enhances learning by leveraging the structure of the data while still benefiting from the self-attention mechanism.

5. Multi-Scale SAGAN-Implements self-attention across multiple scales, allowing the model to capture features at different resolutions, which can be particularly beneficial in complex image datasets.

6. Style-Based SAGAN-Combines ideas from style transfer with SAGAN, allowing for the manipulation of style attributes in generated images while maintaining content consistency.

7.Dual SAGAN-Involves using two GANs in a dual setup where one focuses on generating images and the other on extracting features, enhancing the training dynamics and potentially improving performance.

8.Hybrid SAGAN-Combines SAGAN with other architectures, such as Variational Autoencoders (VAEs) or existing convolutional neural networks (CNNs), to leverage their strengths for specific tasks.

Applications of SAGAN Variants

These variations of SAGAN can be applied in various domains, including

Medical Imaging: For detecting conditions like glaucoma, retinal diseases, etc.

Image Synthesis: Generating high-resolution images or creating artistic styles.

Data Augmentation: Enhancing datasets with synthetic examples for better model training.

These adaptations and extensions enable SAGAN to be a flexible and powerful tool in many fields, particularly where complex image generation and feature extraction are crucial.

SAGAN in Glaucoma Detection: It incorporates self-attention mechanisms in its architecture, allowing the model to focus on relevant features in the input images. This is particularly useful in medical imaging, where subtle patterns can indicate conditions like glaucoma.

Steps for Glaucoma Detection

1. Data Acquisition

Collect a diverse dataset of eye images (e.g., OCT, fundus) labelled as healthy or glaucomatous. Public datasets or clinical databases can be valuable sources.

2. Data Preprocessing

Normalization: Scale pixel values to a range (e.g., 0 to 1).

Resizing: Standardize image sizes for consistent input.

Augmentation: Apply techniques like rotation, flipping, and contrast adjustments to increase the dataset's variability.

3. Model Architecture

Generator: Generates synthetic images that mimic the characteristics of glaucomatous eyes.

Discriminator: Evaluates the authenticity of images (real vs. synthetic) while employing self-attention layers to capture global dependencies.

Self-Attention Mechanism: Allows the model to focus on important areas of the image, improving its ability to detect features related to glaucoma, such as changes in the optic nerve head.

4. Training the Model

Train the SAGAN model using adversarial loss to simultaneously improve the generator and discriminator.

Use a combination of standard loss functions (e.g., binary cross-entropy) and any specific loss metrics relevant to medical imaging.

5. Evaluation Metrics

Assess model performance using metric such as accuracy, sensitivity, specificity, F1-score, and AUC-ROC. This helps gauge how well the model distinguishes between healthy and glaucomatous cases.

6. Visualization and Interpretability

Implement techniques like Grad-CAM or attention maps to visualize which features or regions of the images the model focuses on. This can provide insights into the decision-making process of the model and enhance trust in its predictions.

7. Validation and Testing

Validate the model on a separate test set to ensure it generalizes well to unseen data. Cross-validation can also be employed to assess stability and robustness.

8. Deployment

Deploy the trained model in clinical settings or as part of diagnostic tools. Ensure compliance with healthcare regulations and guidelines.

9. Continuous Improvement

Continuously update the model with new data and feedback to improve accuracy and adapt to changes in population demographics or imaging technology.

II. ARCHITECTURE

The SAGAN Self-Attention Module is a self-attention module used in the Self-Attention GAN architecture for image synthesis.

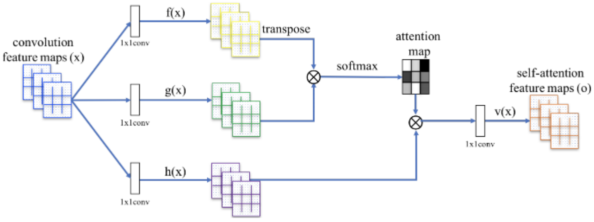

Fig .1. SAGAN Architecture

In the module, image features from the previous hidden layer $x \in R^{C \times N}$ are first transformed into two feature spaces f, g to calculate the attention, where $f(x) = W_f x$, $g(x)=W_g x$.

We then calculate:

$$\beta_{j,i}=\exp(s_{ij})/\sum_{i=1}^{N}\exp(s_{ij})$$

where $s_{ij}=f(x_i)^T g(x_i)$ and $\beta_{j,i}$ indicates the extent to which the model attends to the ith location when synthesizing the jth region. Here, C is the number of channels and N is the number of feature locations of features from the previous hidden layer. The output of the attention layer is $o=(o_1,o_2,\ldots,o_j,\ldots,o_N) \in R^{C \times N}$, where,

$$o_j=v\left(\sum_{i=1}^{N}\beta_{j,i}h(x_i)\right)$$

$$h(x_i)=W_h x_i$$

$$v(x_i)=W_v x_i$$

In the above formulation, $W_g \in R^{\bar{C} \times C}$, $W_f \in R^{\bar{C} \times C}$, $W_h \in R^{\bar{C} \times C}$ and $W_v \in R^{C \times \bar{C}}$ are the learned weight matrices, which are implemented as 1×1 convolutions. Here we choose $\bar{C}=C/8$.
In addition, the module further multiplies the output of the attention layer by a scale parameter and adds back the input feature map.
Therefore, the final output is given by,

$$y_i=\gamma o_i+x_i$$

where γ is a learnable scalar and it is initialized as 0. Introducing γ allows the network to first rely on the cues in the local neighbourhood. Since this is easier and then gradually learn to assign more weight to the non-local evidence.

RELATED WORK

In this model, Self-attention mechanisms is used in BaMSGAN, to significantly enhance the generation of anime faces. Self-attention introduces a critical element to this architecture, enabling the model to capture long-range dependencies and intricate spatial relationships within the input images. By implementing self-attention, BaMSGAN can simultaneously focus on different regions of the image, effectively capturing the global contextual information. This capability empowers both the generator and discriminator to discern complex structures and connections within anime faces, a task often challenging for traditional convolutional architectures.

The integration of self-attention is pivotal for improving the quality of our generated anime faces. It ensures that the model comprehends and synthesizes anime features more cohesively, resulting in sharper and more realistic images. Furthermore, it enhances the discriminator's ability to distinguish between real and generated images, thereby facilitating a more robust and stable training process. The inclusion of self-attention in BaMSGAN underscores our commitment to overcoming the limitations of earlier GAN architectures. It plays a central role in elevating the capabilities of our model, ultimately leading to superior anime face generation and the establishment of new benchmarks in the field.

As for our training approach, we first extract a portion of the samples from the original dataset for edge-blur processing to create a blur dataset. Subsequently, the images from the original dataset serve as positive examples, while the generated images and images from the blur dataset are used as negative examples during training. Once the training converges, the images generated from the sampled portion are preserved as a memory dataset.

During the training process, we continually sample and randomly remove some memory data, which then serves as negative examples to further enhance the training of the BaMSGAN. The whole training process and model architecture are shown in the figure

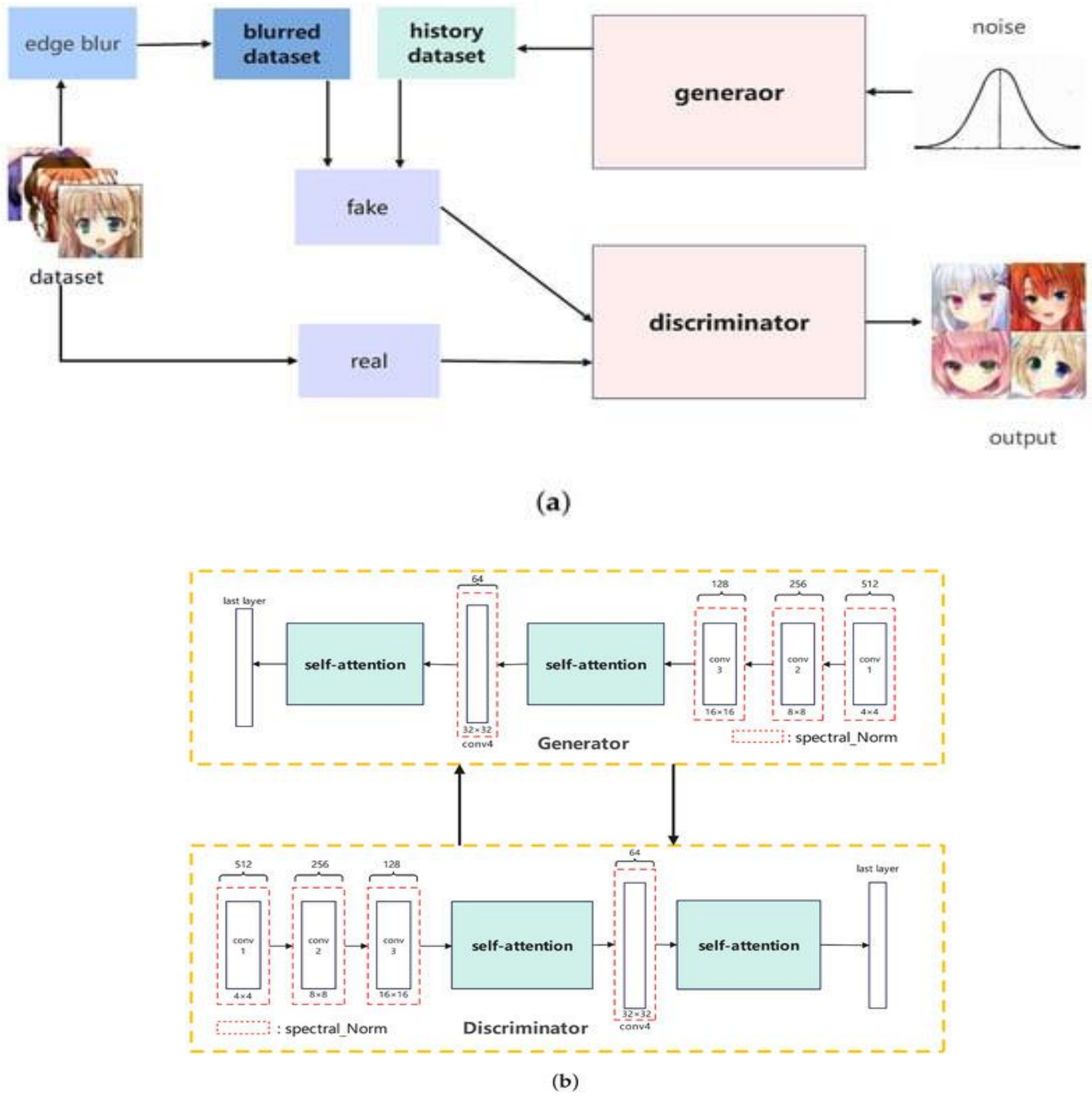

Fig .2. An overall schematic of BaMSGAN.
(a) The overall architecture of our BaMSGAN.
(b) Our BaMSGAN generator and discriminator hierarchy.

The experimental results clearly demonstrate that the BaMSGAN consistently achieves superior performance, both in the short-term and long-term training, producing anime avatars that closely resemble real characters. In the edge-only group experiment, it's evident that the image quality in the short-term is significantly lower than the BaMSGAN, and there's a noticeable deterioration in image quality over prolonged training. In the memory-only group experiment, it becomes apparent that without the additional training facilitated by the edge-blur loss, the BaMSGAN struggles to generate higher-quality images, with some images appearing missing or blurred.

These ablation experiments highlight the indispensable role played by each component in our loss function, emphasizing their collective contribution to the BaMSGAN's remarkable performance in anime face generation.

SAGAN is used for enhancing the historical information and edge blurring for anime face generation. Our model introduces regularization operations to the original SAGAN to enhance the stability of SAGAN training. We selected a subset of images from the source dataset, applied edge detection using the Canny algorithm, and used Gaussian blurring to create a blurred dataset.

Additionally, we saved a portion of the images generated during the training process to construct a historical dataset. These two datasets are then employed as fake data to bolster training, resulting in improved performance in anime face generation. Experimental results demonstrate that our model outperforms DCGAN, WGAN, and SAGAN, generating clearer, higher-quality images and converging more rapidly.

In our future work, we aim to explore additional methods to enhance GAN performance in specific tasks through modifications to the loss function. We will also test our model on a wider range of datasets to enhance its generalization capabilities. Furthermore, we intend to investigate the potential impact of the blurred dataset on the overall image generation style.

III.EXPERIMENTS AND RESULTS

A. Evaluating the Proposed Stabilization Techniques

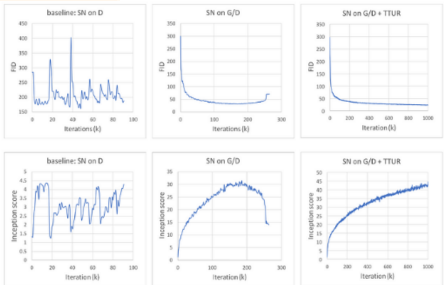

Fig.3. Training curves for the baseline model and models with the proposed stabilization techniques

In the baseline model, SN is only utilized in the discriminator. When we train it with 1:1 balanced update for the discriminator (D) and the generator (G), the training becomes very unstable. It exhibits mode collapse very early in training.As shown in the middle sub-figures, adding SN to both the generator and the discriminator greatly stabilized the model "SN on G/D". For example, the image quality as measured by Fréchet Inception distance (FID) and Inception score (IS) is starting to drop at the 260k-th iteration. When applying the imbalanced learning rates to train the discriminator and the generator, the quality of images generated by the model "SN on G/D+TTUR" improves monotonically during the whole training process.

B. Self-Attention Mechanism

Model	no attention	SAGAN				Residual			
		$feat_8$	$feat_{16}$	$feat_{32}$	$feat_{64}$	$feat_8$	$feat_{16}$	$feat_{32}$	$feat_{64}$
FID	22.96	22.98	22.14	**18.28**	18.65	42.13	22.40	27.33	28.82
IS	42.87	43.15	45.94	51.43	**52.52**	23.17	44.49	38.50	38.96

Fig.4. Comparison of Self-Attention and Residual block

Several SAGAN models are built by adding the self-attention mechanism to different stages of the generator and discriminator. SAGAN models with the self-attention mechanism at the middle-to-high level feature maps (e.g., feat32 and feat64) achieve better performance than the models with the self-attention mechanism at the low level feature maps (e.g., fe at8 and feat16).

For example, the FID of the model "SAGAN, feat8" is improved from 22.98 to 18.28 by "SAGAN, feat32".The attention mechanism gives more power to both generator and discriminator to directly model the long-range dependencies in the feature maps. Compared with residual blocks with the same number of parameters, the self-attention blocks also achieve better results. For example, the training is not stable when we replace the self-attention block with the residual block in 8×8 feature maps, which leads to a significant decrease in performance (e.g.: FID increases from 22.98 to 42.13).This comparison demonstrates that the performance improvement given by using SAGAN is not simply due to an increase in model depth and capacity.

Fig.5. Visualization of attention maps

In each cell, the first image shows three representative query locations with colour coded dots.

For example, in the top-left cell, the red point attends mostly to the body of the bird around it, however, the green point learns to attend to other side of the image. In this way, the image has a consistent background.

Similarly, the blue point allocates the attention to the whole tail of the bird to make the generated part coherent. Those long-range dependencies could not be captured by convolutions with local receptive fields.

As shown in the top-right cell, SAGAN is able to draw dogs with clearly separated legs. The blue query point shows that attention helps to get the structure of the joint area correct.

C. SOTA Comparison

Model	Inception Score	FID
AC-GAN [31]	28.5	/
SNGAN-projection [17]	36.8	27.62*
SAGAN	**52.52**	**18.65**

Fig.6.Comparison of the proposed SAGAN with state of-the-art GAN models [19, 17] for class conditional image generation on ImageNet

As shown in the above table, the proposed SAGAN achieves the best Inception score and FID.

SAGAN can better approximate the original image distribution by using the self-attention module to model the global dependencies between image regions.

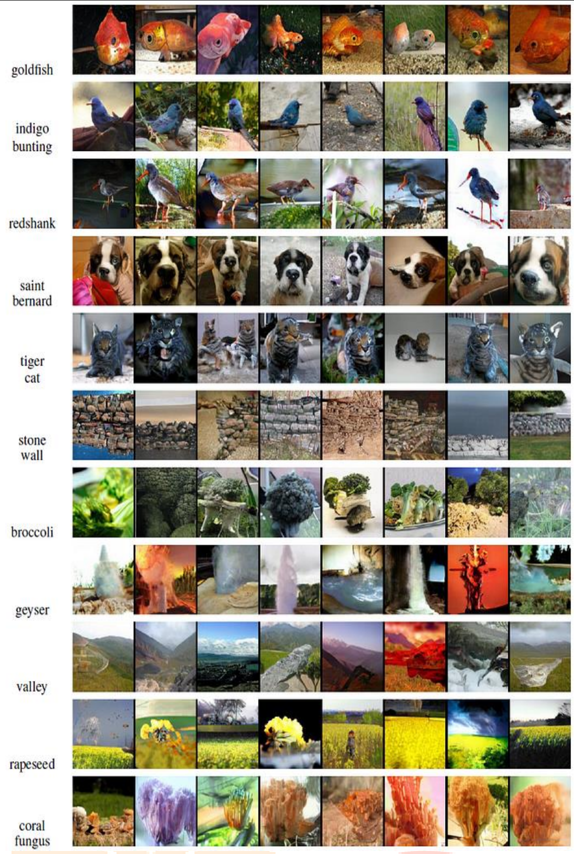

Fig .7.128×128 example images generated by SAGAN for different classes. Each row shows samples from one class

IV.CONCLUSION

SAGAN for glaucoma detection presents a promising advancement in the field of medical imaging and artificial intelligence. This approach enhances the accuracy and efficiency of identifying early signs of glaucoma by utilizing the power of generative adversarial networks and self-attention mechanisms

The ability to analyse complex retinal images and focus on critical features can significantly aid ophthalmologists in making timely diagnoses, ultimately contributing to better patient outcomes. However, successful implementation requires high-quality datasets, rigorous validation, and careful consideration of clinical integration challenges.

Overall, the combination of cutting-edge technology and medical expertise has the potential to transform glaucoma detection, leading to improved vision preservation and enhanced healthcare delivery.

V REFERENCES

[1] Glaucoma Detection using GAN-based Domain Adaptation Approach,S.Kalaiselvi; S. Poorani;A. Amirtharaj; M. Mythreyan; G. Vignesh Kumar 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS) Year: 2023 | Conference Paper | Publisher: IEEE

[2] Domain Adaptation Using DCGAN For Glaucoma Diagnosis, Siddhi Chourasia; Rhugved Bhojane; Rushikesh Patil; Deepali M. Kotambkar,2023 IEEE 8th International Conference for Convergence in Technology (I2CT) Year: 2023 | Conference Paper | Publisher: IEEE

[3] Glaucoma Disease Detection Using Hybrid Deep Learning Model, Manikandan J; Sidharth Raj M; Yogesh Kumar R; Sasi Varma C,
2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM),Year: 2023 | Conference Paper | Publisher: IEEE

[4] Synthesis of Optical Nerve Head Region of Fundus Image ,Anurag Deshmukh; Jayanthi Sivaswamy,2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)Year: 2019 | Conference Paper | Publisher: IEEE

[5] Generating Twin VFs of Unlabeled and Unpaired VF through Generative Model, Zhenyu Zhang; Haogang Zhu,2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)
Year: 2023 | Conference Paper | Publisher: IEEE

[6] Unsupervised Domain Adaptation Based Image Synthesis and Feature Alignment for Joint Optic Disc and Cup Segmentation, Haijun Lei; Weixin Liu; Hai Xie; Benjian Zhao; Guanghui Yue; Baiying Lei,

IEEE Journal of Biomedical and Health Informatics, Year: 2022 | Volume: 26, Issue: 1 | Journal Article | Publisher: IEEE

[7] Utilizing Advanced Deep Learning for Accurate Multilabel Classification of Ocular Diseases,M Kiran Mayee;M Humera Khanam,
2024 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU),Year: 2024 | Conference Paper | Publisher: IEEE

[8] Glaucoma Prediction Using Deep Learning Techniques - A Comparative Study, P. Malin Bruntha;Iwin Thanakumar Joseph S;Saravanan S;D. Narmadha;Suresh Subramanian; G. Naveen Sundar, 2024 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU)

[9] Automated cup-to-disc ratio evaluation for generalised glaucoma diagnosis using deeplabv3+, Alice Divya Nelson;J Raghul;A Devika Gireesh;Kalpana George;S S Anup , 2024 IEEE 12th International Symposium on Signal, Image, Video and Communications (ISIVC)

[10] Convolutional Neural Network Based Glaucoma Detection and Classification using Optic Disc Localization, Mahesh B Elgart. C. Manjunath;Raj Kumar Patra;Sripelli Jagadish;Debarshi Mazumder ,2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)

[11] Glaucoma Detection Using Ensemble Techniques with InceptionV3, VGG16 and ResNet50, Puttaraj Kajagar;Ajeet Shyanavad;Anupama P Bidargaddi;Vinod Soddi;Prashant Ibrahimpur;Meena S M, 2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE)

[12] Performance Analysis of Deep Learning Models for Accurate Glaucoma Detection, Sambhav Dixit;Bhupendra Singh Kirar; Dheraj Kumar Agrawal ,2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)

[13] Multi-Modal AI/ML Integration for Precision Glaucoma Detection: A Comprehensive Analysis using Optical Coherence Tomography, Fundus Imaging, RNFL, and Vessel DensityT Raghunathan;Aditya Mishra;Akhilesh Kumar Mahur;Balaji B 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)

[14] Automated Glaucoma Detection from Fundus Images using Deep Learning, J.Nageswara Rao;Gurijala. Kusuma Priya;Shaik. Sadiya Tabassum;Shaik. Khaleel, 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)

[15] Early Glaucoma Detection through ANSAN-Infused Retinal Vessel Segmentation, Keerthivasan E;Senthil Kumar Thangavel;Madhusudana Rao Nalluri;Somasundaram K;Sathyan Parthasaradhi;Meenakshi Y Dhar;Avadhani Bindu, 2024 International Conference on Inventive Computation Technologies (ICICT)

[16] Classification of Glaucoma Severity Stages Using Deep Learning, Passawut Uttakit;Panawit Hanpinitsak;Sukhumal Thanapaisal;Pattarawit Polpinit 2024 21st International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)

[17] Glaucoma Detection through Synergistic Combination of Image Processing and Deep CNN using MATLAB B, Subbarayudu; Harshitha Yadalam; Jeevana Seepiri;Jayalakshmi Bhogathi, 2024 5th International Conference for Emerging Technology (INCET)

[18] Glaucoma Identification Using Convolutional Neural Networks Ensemble for Optic Disc and Cup Segmentation,Sandra Virbukaitė;Jolita Bernata vičienė;Daiva Imbrasienė,IEEE Access Year: 2024 | Volume: 12 | Journal Article | Publisher: IEEE

[19] CA-Net: A Novel Cascaded Attention-Based Network for Multistage Glaucoma Classification Using Fundus Images,,Dipankar Das;Deepak Ranjan Nayak;Ram Bilas Pachori,IEEE Transactions on Instrumentation and Measurement Year: 2023 | Volume: 72 | Journal Article | Publisher: IEEE

[20] Automated Tool Support for Glaucoma Identification with Explainability Using Fundus Images,Thisara Shyamalee;Dulani Meedeniya;Gilbert Lim;Mihipali Karunarathne,IEEE Access Year: 2024 | Volume: 12 | Journal Article | Publisher: IEEE

[21] Selective Routing-Based Optic Disc Segmentation for Glaucoma Diagnosis,Goutam Balla;Mohammad Farukh Hashmi;Deepali M. Kotambkar,IEEE Sensors Letters,Year: 2024 | Volume: 8, Issue: 5 | Journal Article | Publisher: IEEE

[22] Glaucoma Grading Using Multimodal Imaging and Multilevel CNN,Marcos Melo Ferreira;Geraldo Braz Junior;João Dallyson Sousa de Almeida;Anselmo Cardoso Paiva, IEEE Latin America Transactions Year: 2023 | Volume: 21, Issue: 10 | Journal Article | Publisher: IEEE

[23] Predicting Clinician Fixations on Glaucoma OCT Reports via CNN-Based Saliency Prediction Methods,,Mingyang Zang;Pooja Mukund;Britney Forsyth;Andrew F. Laine;Kaveri A. Thakoor, IEEE Open Journal of Engineering in Medicine and Biology Year: 2024 | Volume: 5 | Journal Article | Publisher: IEEE

[24] Performance Evaluation of Retinal OCT Fluid Segmentation, Detection, and Generalization Over Variations of Data Sources, Nchongmaje Ndipenoch;Alina Miron;Yongmin Li, IEEE Access Year: 2024 | Volume: 12 | Journal Article | Publisher: IEEE

[25] Reconstruction-Driven Dynamic Refinement Based Unsupervised Domain Adaptation for Joint Optic Disc and Cup Segmentation, Ziyang Chen;Yongsheng Pan; Yong Xia ,IEEE Journal of Biomedical and Health Informatics, Year: 2023 | Volume: 27, Issue: 7 | Journal Article | Publisher: IEEE

[26] Geometric Correspondence-Based Multimodal Learning for Ophthalmic Image Analysis, Yan Wang;Liangli Zhen;Tien-En Tan;Huazhu Fu; Yangqin Feng;Zizhou Wang;Xinxing Xu;Rick Siow Mong Goh;Yipin Ng;Claire Calhoun;Gavin Siew Wei Tan; Jennifer K. Sun; Yong Liu; Daniel Shu Wei Ting , IEEE Transactions on Medical Imaging, Year: 2024 | Volume: 43, Issue: 5 | Journal Article | Publisher: IEEE

[27] GLIM-Net: Chronic Glaucoma Forecast Transformer for Irregularly Sampled Sequential Fundus Images, Xiaoyan Hu;Ling-Xiao Zhang; Lin Gao; Weiwei Dai.

[28] Xiaoguang Han; Yu-Kun Lai;Yiqiang Chen, IEEE Transactions on Medical Imaging,Year: 2023 | Volume: 42, Issue: 6 | Journal Article | Publisher: IEEE