BIBLIOGRAPHY

Aatila, M., Lachgar, M., Hrimech, H., and Kartit, A. (2021), "Diabetic retinopathy classification using ResNet50 and VGG-16 pretrained networks", *International Journal of Computer Engineering and Data Science (IJCEDS),* Volume 1, Issue 1, Pp. 1-7.

Abbood, S. H., Hamed, H. N. A., Rahim, M. S. M., Rehman, A., Saba, T., and Bahaj, S. A. (2022), "Hybrid retinal image enhancement algorithm for diabetic retinopathy diagnostic using deep learning model", *IEEE Access,* Volume 10, Pp. 73079-73086.

Abdel Maksoud, E., Barakat, S., and Elmogy, M. (2020), "Diabetic retinopathy grading system based on transfer learning", arXiv preprint arXiv:2012.12515.

AbdelMaksoud, E., El-Sappagh, S., Barakat, S.I., Abuhmed, T., and Elmogy, M. M. (2021), "Automatic Diabetic Retinopathy Grading System Based on Detecting Multiple Retinal Lesions", *IEEE Access,* Volume 9, Pp. 15939-15960.

Aharon, M., Elad, M. and Bruckstein, A. (2006), "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation", *IEEE Transactions on Signal Processing,* Volume 54, Pp. 4311-4322.

Ahmed, H., Zhang, Q., Donnan, R., & Alomainy, A. (2024), "Denoising of Optical Coherence Tomography Images in Ophthalmology Using Deep Learning", *A Systematic Review. Journal of Imaging,* 10(4), 86.

Al-Antary, M. T., and Arafa, Y. (2021), "Multi-scale attention network for diabetic retinopathy classification", *IEEE Access,* Volume 9, Pp. 54190-54200.

Alimanov, A., and Islam, M. B. (2022), "Retinal Image Restoration and Vessel Segmentation using Modified Cycle-CBAM and CBAM-UNet", *IEEE Innovations in Intelligent Systems and Applications Conference (ASYU),* Pp. 1-6.

Aljehane, N. O. (2022), "An Intelligent Moth Flame Optimization with Inception Network for Diabetic Retinopathy Detection and Grading", *IEEE International Conference on Computing and Information Technology (ICCIT),* Pp. 370-373.

Alyoubi, W. L., Abulkhair, M. F., and Shalash, W. M. (2021), "Diabetic retinopathy fundus image classification and lesions localization system using deep learning", *Sensors*, Volume 21, Issue 11, Pp. 3704.

Andreini, P., Ciano, G., Bonechi, S., Graziani, C., Lachi, V., Mecocci, A., Sodi, A., Scarselli, F., and Bianchini, M. A. (2021), "Two-stage GAN for high-resolution retinal image generation and segmentation", *Electronics*, Volume 11, Issue 1, Pp. 1-17.

Anilet B., Chiranjeeb H. and Punith C. (2014), "Image Denoising Method using curvelet Transform and Wiener Filter", *International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering*, Volume 3, Issue 1, Pp. 6943-6950.

Arcadu, F., Benmansour, F., Maunz, A., Willis, J., Haskova, Z., and Prunotto, M. (2019), "Deep learning algorithm predicts diabetic retinopathy progression in individual patients", *NPJ digital medicine*, Volume 2, Issue 1, Pp. 92.

Azad, R., Asadi Aghbolaghi, M., Fathy, M., and Escalera, S. (2019), "Bi-Directional ConvLSTM U-Net with Densley Connected Convolutions", *IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, Pp. 406-415.

Aziz, T., Charoenlarpnopparut, C., and Mahapakulchai, S. (2023), "Deep learning-based hemorrhage detection for diabetic retinopathy screening", *Scientific reports*, Volume 13, Issue 1, Pp. 1479.

Bandyopadhyay, S., Saha, S., Maulik, U., and Deb, K. (2008), "A Simulated Annealing-Based Multiobjective Optimization Algorithm: AMOSA", *IEEE Transactions on Evolutionary Computation*, Volume 12, Pp. 269-283.

Beck, A., and Teboulle, M. (2009), "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems", *IEEE transactions on image processing*, Volume 18, Issue 11, Pp. 2419-2434.

Beevi, S. Z. (2023), "Multi-Level severity classification for diabetic retinopathy based on hybrid optimization enabled deep learning", *Biomedical Signal Processing and Control*, Volume 84, Pp. 104736.

Bilal, A., Sun, G., Li, Y., Mazhar, S., and Khan, A. Q. (2021), "Diabetic retinopathy detection and classification using mixed models for a disease grading database", *IEEE Access,* Volume 9, Pp. 23544-23553.

Bodapati, J. D., Shaik, N. S., and Naralasetti, V. (2021), "Composite deep neural network with gated-attention mechanism for diabetic retinopathy severity classification", *Journal of Ambient Intelligence and Humanized Computing,* Volume 12, Issue 10, Pp. 9825-9839.

Buades, A., Coll, B., and Morel, J. M. (2005), "A non-local algorithm for image denoising", *IEEE computer society conference on computer vision and pattern recognition (CVPR'05),* Volume 2, Pp. 60-65.

Butt, M. M., Iskandar, D. A., Abdelhamid, S. E., Latif, G., and Alghazo, R. (2022), "Diabetic Retinopathy Detection from Fundus Images of the Eye Using Hybrid Deep Learning Features. Diagnostics", Volume 12 Issue 7, Pp. 1607.

Chakraborti, T., Jha, D. K., Chowdhury, A. S., and Jiang, X. (2015), "A self-adaptive matched filter for retinal blood vessel detection", *Machine Vision and Applications,* Volume 26, Issue 1, Pp. 55–68.

Chang, S. G., Yu, B., and Vetterli, M. (2000), "Adaptive wavelet thresholding for image denoising and compression", *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*, Volume 9, Issue 9, Pp. 1532–1546.

Chen, P. N., Lee, C. C., Liang, C. M., Pao, S. I., Huang, K. H., and Lin, K. F. (2021), "General deep learning model for detecting diabetic retinopathy", *BMC Bioinformatics,* Volume 22, Issue 5, Pp. 1-15.

Chen, Y., Li, J., Xiao, H., Jin, X., Yan, S., and Feng, J. (2017), "Dual path networks", *In Proceedings of Advanced Neural Information Processing Systems*, Pp. 1-11.

Chen, Y., Long, J., and Guo, J. (2021), "RF-GANs: a method to synthesize retinal fundus images based on generative adversarial network", *Computational Intelligence and Neuroscience*, Pp. 1-17.

Church, J. C., Chen, D. Y. and Rice, D. S. (2008), "A Spatial Median Filter for Noise Removal in Digital Images", *IEEE SoutheastCon,* Pp. 618- 623.

Church, J. C., Chen, Y., and Rice, S. V. (2008), "A spatial median filter for noise removal in digital images", *IEEE South east Con*., Pp. 618-623.

Cordero Martinez, R., Sanchez, D., and Melin, P. (2022), "Hierarchical genetic optimization of convolutional neural models for diabetic retinopathy classification", *International Journal of Hybrid Intelligent Systems*, (Preprint), Pp. 1-13.

Costa, P. and Campilho, A. (2017), "Convolutional bag of words for diabetic retinopathy detection from eye fundus images", *IPSJ Transactions on Computer Vision and Applications*, Volume 9, Pp. 1-6.

Das, D., Biswas, S. K., and Bandyopadhyay, S. (2022), "A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning", *Multimedia Tools and Applications*, Volume 81, Issue 18, Pp. 25613-25655.

Dayana, A. M., and Emmanuel, W. S. (2022), "An enhanced swarm optimization-based deep neural network for diabetic retinopathy classification in fundus images", *Multimedia Tools and Applications*, Volume 81, Issue 15, Pp. 20611-20642.

Dekhil, O., Naglah, A., Shaban, M., Ghazal, M., Taher, F., and Elbaz, A. (2019), "Deep Learning Based Method for Computer Aided Diagnosis of Diabetic Retinopathy", *IEEE International Conference on Imaging Systems and Techniques (IST)*, Pp. 1-4.

Donoho, D. L. (1995), "Denoising by Soft-Thresholding", *IEEE Transactions on Information Theory*, Volume 41, Pp. 613-627.

Donoho, D. L., and Johnstone, I. M. (1995), "Adapting to Unknown Smoothness via Wavelet Shrinkage", *Journal of the American Statistical Association*, Volume 90, Pp. 1200-1224.

El Houby, E. M. (2021), "Using transfer learning for diabetic retinopathy stage classification", *Applied Computing and Informatics*, Pp. 1-11.

Eszes, D. J., Szabo, D. J., Russell, G., Lengyel, C., Varkonyi, T., Paulik, E., and Petrovski, B. É. (2021), "Diabetic retinopathy screening in patients with diabetes using a handheld fundus camera: The experience from the South-Eastern region in Hungary", *Journal of Diabetes Research*, Pp. 6646645.

Fukushima, K., (1980), "A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", *Biol. Cybern.,* Volume 36, Pp.193-202.

Gadekallu, T. R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P. K. R., Ra, I. H., and Alazab, M. (2020), "Early detection of diabetic retinopathy using PCA-firefly based deep learning model", *Electronics*, Volume 9, Issue 2, Pp. 274.

Gangwar, A. K., and Ravi, V. (2021), "Diabetic retinopathy detection using transfer learning and deep learning", *In Evolution in Computational Intelligence: Frontiers in Intelligent Computing: Theory and Applications (FICTA 2020),* Springer Singapore, Volume 1, Pp. 679-689.

Gao, Z., Li, J., Guo, J., Chen, Y., Yi, Z., and Zhong, J. (2018), "Diagnosis of diabetic retinopathy using deep neural networks", *IEEE Access,* Volume 7, Pp. 3360-3370.

Gayathri, S., Gopi, V. P., and Palanisamy, P. (2020), "A lightweight CNN for Diabetic Retinopathy classification from fundus images", *Biomedical Signal Processing and Control*, Volume 62, Pp. 102115.

Goodfellow, I., Bengio, Y., and Courville, A. (2016), "Deep learning". *MIT press*.

Gu, S., Zhang, L., Zuo, W., and Feng, X. (2014), "Weighted nuclear norm minimization with application to image denoising", *In Proceedings of the IEEE conference on computer vision and pattern recognition*, Pp. 2862-2869.

Gundluru, N., Rajput, D. S., Lakshmanna, K., Kaluri, R., Shorfuzzaman, M., Uddin, M., and Rahman Khan, M. A. (2022), "Enhancement of detection of diabetic retinopathy using Harris hawks optimization with deep learning model", *Computational Intelligence and Neuroscience*, 8512469, Pp. 1-13.

HaoQi, G., and Ogawara, K. (2020), "CGAN-based synthetic medical image augmentation between retinal fundus images and vessel segmented images", *IEEE International Conference on Control and Robotics Engineering (ICCRE),* Pp. 218-223.

Harangi, B., Toth, J., Baran, A., and Hajdu, A. (2019), "Automatic screening of fundus images using a combination of convolutional neural network and hand-crafted features", *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* Pp. 2699-2702.

Hemanth, D. J., Deperlioglu, O. and Kose, U. (2020), "An Enhanced Diabetic Retinopathy Detection and Classification Approach using Deep Convolutional Neural Network", *Neural Computing and Applications*, Volume 32, Issue 3, Pp. 707-721.

Hilgert, G. R., Trevizan, E. and Souza, J. M. D. (2019), "Use of a handheld fundus camera as a screening tool for diabetic retinopathy", *Revista Brasileira de Oftalmologia*, Volume 78, Pp. 321-326.

Irsch, K. and Guyton, D. L. (2009), "Anatomy of Eyes", In: Li, S.Z., Jain, A. (eds) *Encyclopedia of Biometrics*, Springer, Pp. 11-16.

Jayanthi, J., Jayasankar, T., Krishnaraj, N., Prakash, N. B., Sagai Francis Britto, A., and Vinoth Kumar, K. (2021), "An Intelligent Particle Swarm Optimization with Convolutional Neural Network for Diabetic Retinopathy Classification Model", *Journal of Medical Imaging and Health Informatics*, Volume 11, Issue 3, Pp. 803-809.

Jena, P. K., Khuntia, B., Palai, C., Nayak, M., Mishra, T. K., and Mohanty, S. N. (2023), "A novel approach for diabetic retinopathy screening using asymmetric deep learning features", *Big Data and Cognitive Computing*, Volume 7, Issue 1, Pp. 25.

Jintasuttisak, T. and Intajag, S. (2014), "Color Retinal Image Enhancement by Rayleigh Contrast Limited Histogram Equalization", *International Conference on Control, Automation and Systems*, Pp. 692-697.

Kassani, S. H., Kassani, P. H., Khazaeinezhad, R., Wesolowski, M. J., Schneider, K. A., and Deters, R. (2019), "Diabetic Retinopathy Classification Using a Modified Xception Architecture", *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Pp. 1-6.

Kawakami, T., Murahira, K., and Taguchi, A. (2009), "Modified histogram equalization with variable enhancement degree for image contrast enhancement", IEEE *International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, Pp. 570-573.

Khan, M. F., Khan, E. and Abbasi, Z. A. (2014), "Segment dependent dynamic multi-histogram equalization for image contrast enhancement", *Digital Signal Processing*, Volume 25, Pp. 198-223.

Khan, M. F., Khan, E., and Abbasi, Z. A. (2014), "Segment selective dynamic histogram equalization for brightness preserving contrast enhancement of images", *Optik*, Volume 125, Issue 3, Pp. 1385-1389.

Khan, Z., Khan, F. G., Khan, A., Rehman, Z. U., Shah, S., Qummar, S., Ali, F. and Pack, S. (2021), "Diabetic retinopathy detection using VGG-NIN a deep learning architecture", *IEEE Access*, Volume 9, Pp. 61408-61416.

Kim, Y. T. (1997), "Contrast enhancement using brightness preserving bi-histogram equalization", *IEEE transactions on Consumer Electronics*, Volume 43, Issue 1, Pp. 1-8.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983), "Optimization by Simulated Annealing", *Science*, Volume 220, Pp. 671 - 680.

Kornblith, S., Shlens, J. and Le, Q. V. (2019), "Do Better Imagenet Models Transfer Better?", *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, Pp. 2661-2671.

Krizhevsky, A., Sutskever, I. and Hinton, G. E., (2012), "Imagenet Classification with Deep Convolutional Neural Networks", *Advances in neural information processing systems*, Pp. 1097-1105.

Lam, C., Yi, D., Guo, M. and Lindsey, T. (2018), "Automated Detection of Diabetic Retinopathy using Deep Learning", *In AMIA Summits on Translational Science Proceedings*, Pp. 147-155.

Le Cun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. and Jackel, L.D. (1989), "Backpropagation Applied to Handwritten Zip Code Recognition", *Neural computation*, Volume 1, Issue 4, Pp.541-551.

Li, X., Hu, X., Yu, L., Zhu, L., Fu, C. W. and Heng, P. A. (2020), "CANet: Cross-Disease Attention Network for Joint Diabetic Retinopathy and Diabetic Macular Edema

Grading", *IEEE transactions on medical imaging,* Volume 39, Issue 5, Pp. 1483–1493.

Liu, H., Yue, K., Cheng, S., Pan, C., Sun, J., & Li, W. (2020), "Hybrid Model Structure for Diabetic Retinopathy Classification", *Journal of Healthcare Engineering*, Pp. 1-9.

Maaliw, R. R., Mabunga, Z. P., De Veluz, M. R. D., Alon, A. S., Lagman, A. C., Garcia, M. B., Lacatan, L and Dellosa, R. M. (2023), "An Enhanced Segmentation and Deep Learning Architecture for Early Diabetic Retinopathy Detection", *IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC),* Pp. 0168-0175.

Magister, L. C., and Arandjelovi, O. (2021), "Generative image inpainting for retinal images using generative adversarial networks", *IEEE 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society*, Pp. 2835-2838.

Majumder, S., and Kehtarnavaz, N. (2021), "Multitasking Deep Learning Model for Detection of Five Stages of Diabetic Retinopathy", *IEEE Access,* Volume 9, Pp. 123220-123230.

Mall, P. K., Singh, P. K., Srivastav, S., Narayan, V., Paprzycki, M., Jaworska, T., & Ganzha, M. (2023), "A comprehensive review of deep neural networks for medical image processing", *Recent developments and future opportunities. Healthcare Analytics,* 100216.

Mallat, S. G. (1989), "A Theory for Multi Resolution Signal Decomposition: The Wavelet Representation", *IEEE transactions on pattern analysis and machine intelligence,* Volume 11, Issue 7, Pp. 674-693.

Mallat, S.G. (2008), "A Wavelet Tour of Signal Processing: The Sparse Way", *Academic Press, Elsevier*.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953), "Equation of state calculations by fast computing machines", *The journal of chemical physics,* Volume 21, Issue 6, Pp. 1087-1092.

Mohideen, S., Arumuga, Perumal and Sathik, Mohamed (2008), "Image Denoising using Discrete Wavelet transform", *International Journal of Computer Science and Network Security*, Volume 8, Issue 1, Pp. 213-216.

Mushtaq, G., and Siddiqui, F. (2021), "Detection of diabetic retinopathy using deep learning methodology", *In IOP conference series: materials science and engineering*, Volume 1070, Issue 1, Pp. 12049.

Nakayama, L. F., Ribeiro, L. Z., Gonçalves, M. B., Ferraz, D. A., Dos Santos, H. N. V., Malerbi, F. K., Morales, P. H., Maia, M., Regatieri, C. V. S. and Mattos, R. B. (2022), "Diabetic retinopathy classification for supervised machine learning algorithms", *International Journal of Retina and Vitreous*, Volume 8, Issue 1, Pp. 1-5.

Odstrcilik, J., Kolar, R., Tornow, R. P., Jan, J., Budai, A., Mayer, M., Vodakova, M., Laemmer, R., Lamos, M., Kuna, Z., Gazarek, J., Kubena, T., Cernosek, P., and Ronzhina, M. (2014), "Thickness related textural properties of retinal nerve fiber layer in color fundus images", *Computerized medical imaging and graphics: the official journal of the Computerized Medical Imaging Society*, Volume 38, Issue 6, Pp. 508–516.

Ozbay, E. (2023), "An active deep learning method for diabetic retinopathy detection in segmented fundus images using artificial bee colony algorithm", *Artificial Intelligence Review*, Volume 56, Issue 4, Pp. 3291-3318.

Paletta, Q., Terrén-Serrano, G., Nie, Y., Li, B., Bieker, J., Zhang, W., ... & Feng, C. (2023), "Advances in solar forecasting: Computer vision with deep learning", *Advances in Applied Energy*, 100150.

Pires, R., Rocha, A., and Wainer, J. (2019), "Image analytics techniques for diabetic retinopathy detection", *In Anais do XXXII Concurso de Teses e Dissertações*, SBC, Pp. 1-6.

Pisano, E. D., Zong, S., Hemminger, B. M., DeLuca, M., Johnston, R. E., Muller, K., Braeuning, M. P. and Pizer, S. M. (1998), "Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms", *Journal of Digital imaging*, Volume 11, Pp. 193-200.

Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., and Meriaudeau, F. (2018), "Indian Diabetic Retinopathy Image Dataset (IDRiD): A Database for Diabetic Retinopathy Screening Research", *Data*, Volume 3, Issue 3, Pp. 25.

Priya, R.P., Sivarani, T.S., and Saravanan, A.G. (2021), "Deep Long and Short Term Memory based Red Fox Optimization Algorithm for Diabetic Retinopathy Detection and Classification*", International Journal for Numerical Methods in Biomedical Engineering*, Pp. 38.

Qummar, S., Khan, F. G., Shah, S., Khan, A., Shamshirband, S., Rehman, Z. U., Ahmed Khan, I., and Jadoon, W. (2019), "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection", *IEEE Access*, Volume 7, Pp. 150530-150539.

Qureshi, I., Ma, J., and Abbas, Q. (2019), "Recent Development on Detection Methods for the Diagnosis of Diabetic Retinopathy", *Symmetry*, Volume 11, Issue 6, Pp. 749.

Radhamani, V., Venkataramanan, V., Diwakaran, S., Subramanian, M. and Rajasekaran, A. S. (2021), "Wavelet thresholding techniques implementation in retinal images for suppressing noises", *Materials Today: Proceedings*, Volume 57, Part 5, Pp. 2124-2133.

Sadollah, A., Bahreininejad, A., Eskandar, H., and Hamdi, M. (2013), "Mine Blast Algorithm: A New Population based Algorithm for Solving Constrained Engineering Optimization Problems", *Applied Soft Computing*, Volume 13, Issue 5, Pp. 2592-2612.

Sadollah, A., Eskandar, H., and Kim, J. H. (2014), "Geometry Optimization of a Cylindrical Fin Heat Sink using Mine Blast Algorithm, *The International Journal of Advanced Manufacturing Technology*, Volume 73, Pp. 795-804.

Sadollah, A., Yoo, D. G., and Kim, J. H. (2015), "Improved mine blast algorithm for optimal cost design of water distribution systems", *Engineering Optimization*, Volume 47, Issue 12, Pp. 1602-1618.

Saeed, F., Hussain, M., and Aboalsamh, H. A. (2021), "Automatic diabetic retinopathy diagnosis using adaptive fine-tuned convolutional neural network", *IEEE Access,* Volume 9, Pp. 41344-41359.

Sakaguchi, A., Wu, R., and Kamata, S. I. (2019), "Fundus image classification for diabetic retinopathy using disease severity grading", *In Proceedings of the 2019 9th International Conference on Biomedical Engineering and Technology,* Pp. 190-196.

Salman, H., Ilyas, A., Engstrom, L., Kapoor, A., and Madry, A. (2020), "Do Adversarially Robust Imagenet Models Transfer Better?, *Advances in Neural Information Processing Systems*, Volume 33, Pp. 3533-3545.

Santos, J., Carrijo, G., Santos, F., Ferreira, J., Sousa, P. and Patrocinio, A. (2020), "Fundus image quality enhancement for blood vessel detection via a neural network using CLAHE and Wiener filter", *Research on Biomedical Engineering*, Pp. 36.

Saranya, P., Pranati, R., and Patro, S. S. (2023), "Detection and classification of red lesions from retinal images for diabetic retinopathy detection using deep learning models", *Multimedia Tools and Applications*, Pp. 1-21.

Sebastian, A., Elharrouss, O., Al-Maadeed, S., and Almaadeed, N. (2023), "A Survey on Deep Learning-based Diabetic Retinopathy Classification", *Diagnostics*, Volume 13, Issue 3, Pp. 345.

Shaik, F., Giri Prasad, M. N., Rao, J., Abdul Rahim, B. and Somasekhar, A. (2010), "Medical image analysis of electron micrographs in diabetic patients using contrast enhancement", *International Conference on Mechanical and Electrical Technology,* Pp. 482-485.

Sheikh, S. and Qidwai, U. (2021), "Smartphone-based Diabetic Retinopathy Severity Classification using Convolution Neural Networks", *In Proc. of the Advances in Intelligent Systems and Computing,* Springer, Volume 1252, Pp. 469–481.

Sikder, N., Chowdhury, M., Arif, A. and Nahid, A. (2020), "Early Blindness Detection Based on Retinal Images Using Ensemble Learning", *22nd International Conference on Computer and Information Technology (ICCIT),* Pp. 1-6.

Sikder, N., Masud, M., Bairagi, A. K., Arif, A. S. M., Nahid, A. A., and Alhumyani, H. A. (2021), "Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images", *Symmetry*, Volume 13, Issue 4, Pp. 1-26.

Sim, K. S., Tso, C. P., and Tan, Y. Y. (2007), "Recursive sub-image histogram equalization applied to gray scale images", *Pattern Recognition Letters*, Volume 28, Issue 10, Pp. 1209-1221.

Solomon, S. D., and Goldberg, M. F. (2019), "ETDRS grading of diabetic retinopathy: still the gold standard?", *Ophthalmic research*, Volume 62, Issue 4, Pp. 190-195.

Sreejini, K. S., and Govindan, V. K. (2019), "Retrieval of pathological retina images using Bag of Visual Words and pLSA model", *Engineering Science and Technology, an International Journal*, Volume 22, Issue 3, Pp. 777-785.

Srinivasan, V. and Rajagopal, V. (2022), "Multi-Scale Attention-Based Mechanism in Gradient Boosting Convolutional Neural Network for Diabetic Retinopathy Grade Classification", *International Journal of Intelligent Engineering and Systems*, Volume 15, Issue 4, Pp. 489-498.

Srinivasan, V. and Rajagopal, V. (2023), "An Optimization Method for Tuning Hyper-Parameters of SGAN with Ensemble Classification Regression Model", *International Journal of Electrical and Electronics Engineering*, Volume 10, Pp. 24-36.

Srinivasan, V., and Rajagopal, V. (2022), "Computer-Aided Diabetic Retinopathy Diagnosis Using Conventional and Deep Learning Techniques-A Comparison", In: Kumar, N., Shahnaz, C., Kumar, K., Abed Mohammed, M., Raw, R.S. (eds) *Advance Concepts of Image Processing and Pattern Recognition. Transactions on Computer Systems and Networks*, Springer, Singapore, Pp. 131-153.

Sutour, C., Deledalle, C. A., and Aujol, J. F. (2014), "Adaptive regularization of the NL-means: Application to image and video denoising", *IEEE Transactions on image processing*, Volume 23, Issue 8, Pp. 3506-3521.

Tamim, N., Elshrkawey, M., and Nassar, H. (2021), "Accurate diagnosis of diabetic retinopathy and glaucoma using retinal fundus images based on hybrid features and genetic algorithm", *Applied Sciences*, Volume 11, Issue 13, Pp. 6178.

Tariq, H., Rashid, M., Javed, A., Zafar, E., Alotaibi, S. S., and Zia, M. Y. I. (2021), "Performance Analysis of Deep-Neural-Network-Based Automatic Diagnosis of Diabetic Retinopathy", *Sensors*, Volume 22, Issue 1, Pp. 205.

Tian, F., Li, Y., Wang, J. and Chen, W. (2021), "Blood Vessel Segmentation of Fundus Retinal Images Based on Improved Frangi and Mathematical Morphology", *Computational and Mathematical Methods in Medicine*, Pp. 1-11.

Ullah, N., Mohmand, M. I., Ullah, K., Gismalla, M. S., Ali, L., Khan, S. U., and Ullah, N. (2022), "Diabetic Retinopathy Detection Using Genetic Algorithm-based CNN Features and Error Correction Output Code SVM Framework Classification Model", *Wireless Communications and Mobile Computing*, Pp. 1-13.

Uppamma, P., and Bhattacharya, S. (2023), "Diabetic Retinopathy Detection: A Blockchain and African Vulture Optimization Algorithm-Based Deep Learning Framework", *Electronics*, Volume 12, Issue 3, Pp. 742.

Vikramathithan, A. C., Pooja, P., Bhaskar, M. S., Navya, S., and Rakshith, K. B. (2022), "Identification of diabetic retinopathy (DR) using Image Processing", *Journal of Physics: Conference Series*, Volume 2161, Pp. 1-9.

Waibel, A., Sawai, H. and Shikano, K., (1989), "Modularity and scaling in large phonemic neural networks", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Volume 37, Issue 12, Pp.1888-1898.

Wang, L. and Schaefer, A. (2020), "Diagnosing diabetic retinopathy from images of the eye fundus", *Cs230*, Stanford Edu, Pp. 1-6.

Xu, Z., Liu, X. and Ji, N. (2009), "Fog Removal from Color Images using Contrast Limited Adaptive Histogram Equalization", *International Congress on Image and Signal Processing*, Volume 10, Pp. 1-5.

Yang, T., Wu, T., Li, L., and Zhu, C. (2020), "SUD-GAN: deep convolution generative adversarial network combined with short connection and dense block for retinal vessel segmentation", *Journal of Digital Imaging*, Volume 33, Issue 4, Pp. 946-957.

Yao, Z., Yuan, Y., Shi, Z., Mao, W., Zhu, G., Zhang, G., and Wang, Z. (2022), "FunSwin: A deep learning method to analysis diabetic retinopathy grade and macular edema risk based on fundus images", *Frontiers in physiology*, Pp. 1462.

Yavuz, Zafer and Köse, Cemal. (2017), "Blood Vessel Extraction in Color Retinal Fundus Images with Enhancement Filtering and Unsupervised Classification", *Journal of Healthcare Engineering*, Pp. 1-12.

Yoo, T. K., Choi, J. Y., and Kim, H. K. (2020), "Cycle GAN-based deep learning technique for artifact reduction in fundus photography", *Graefe's Archive for Clinical and Experimental Ophthalmology*, Volume 258, Issue 8, Pp. 1631-1637.

Yu, Z., Xiang, Q., Meng, J., Kou, C., Ren, Q., and Lu, Y. (2019), "Retinal image synthesis from multiple landmarks input with generative adversarial networks", *Biomedical Engineering Online*, Volume 18, Issue 1, Pp. 1-15.

Zhao, L. and Liu, Y. (2017), "A New Generalized Orthogonal Matching Pursuit Method", J*ournal of Electrical and Computer Engineering*, Pp. 1-7.

Zheng, Y., Vanderbeek, B. L., Xiao, R., Daniel, E., Stambolian, D., Maguire, M. G., O'Brien, J., and Gee, J. C. (2012), "Retrospective illumination correction of retinal fundus images from gradient distribution sparsity", *IEEE International Symposium on Biomedical Imaging (ISBI)*, Pp. 972-975.

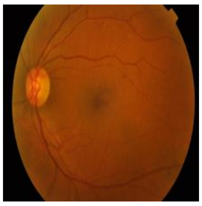
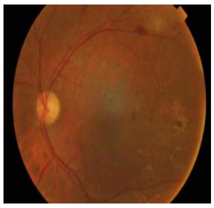
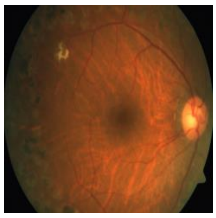
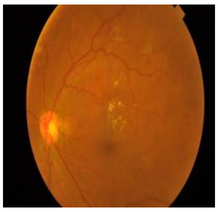
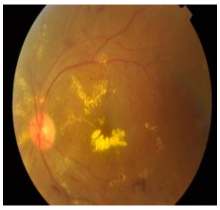
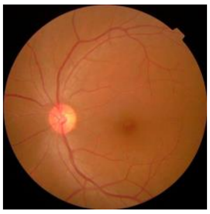
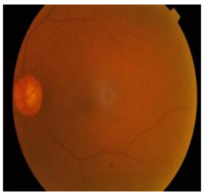
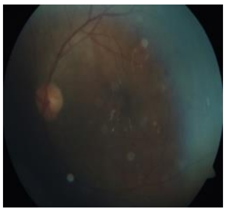
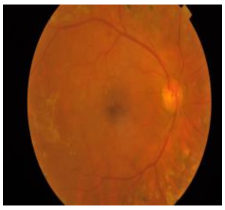
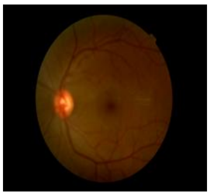
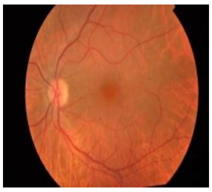
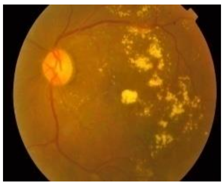
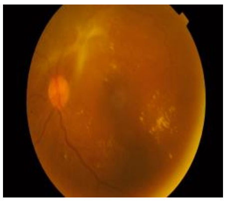
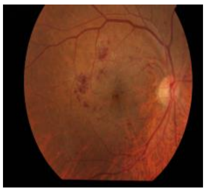
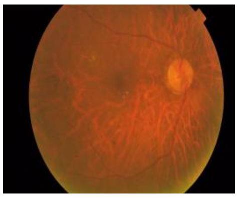
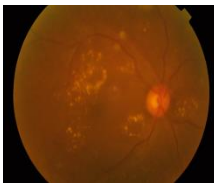
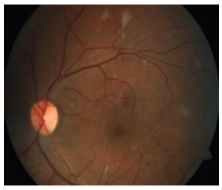
Website References

1. https://www.kaggle.com/c/aptos2019-blindness-detection/
2. https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid
3. https://www.who.int/
4. https://www.nih.gov/
5. https://www.cdc.gov/health-topics.html
6. https://www.niddk.nih.gov/health-information/diabetes

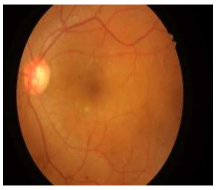
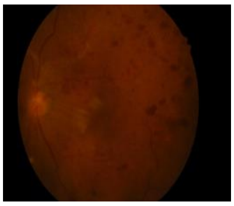
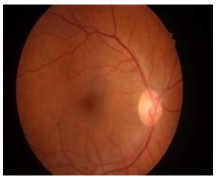
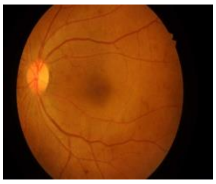
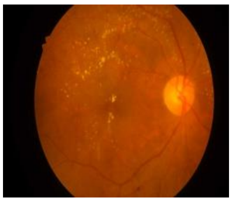
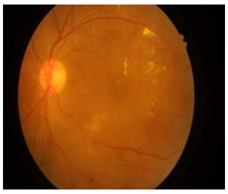
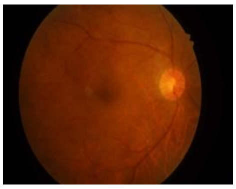
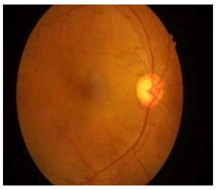
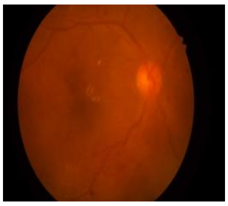
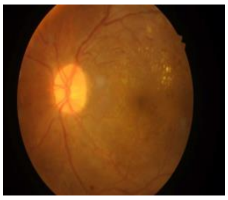
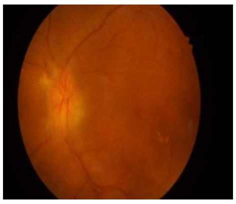
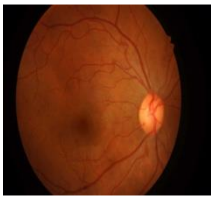
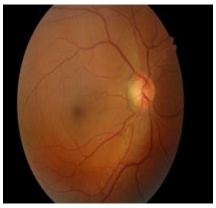
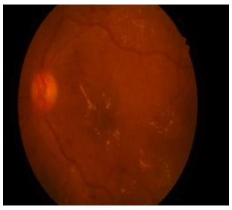
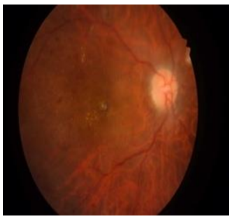
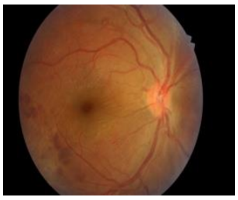
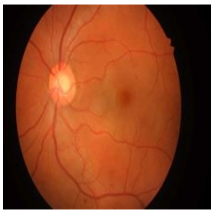
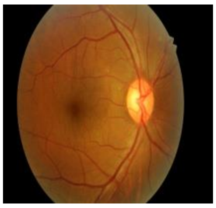
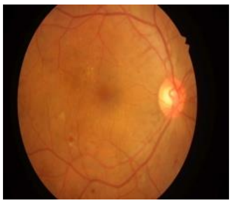
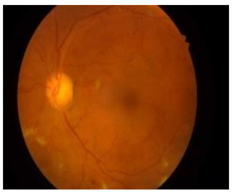
ANNEXURES

ANNEXURE I

<u>APTOS 2019 BLINDNESS DETECTION DATASET SAMPLES</u>

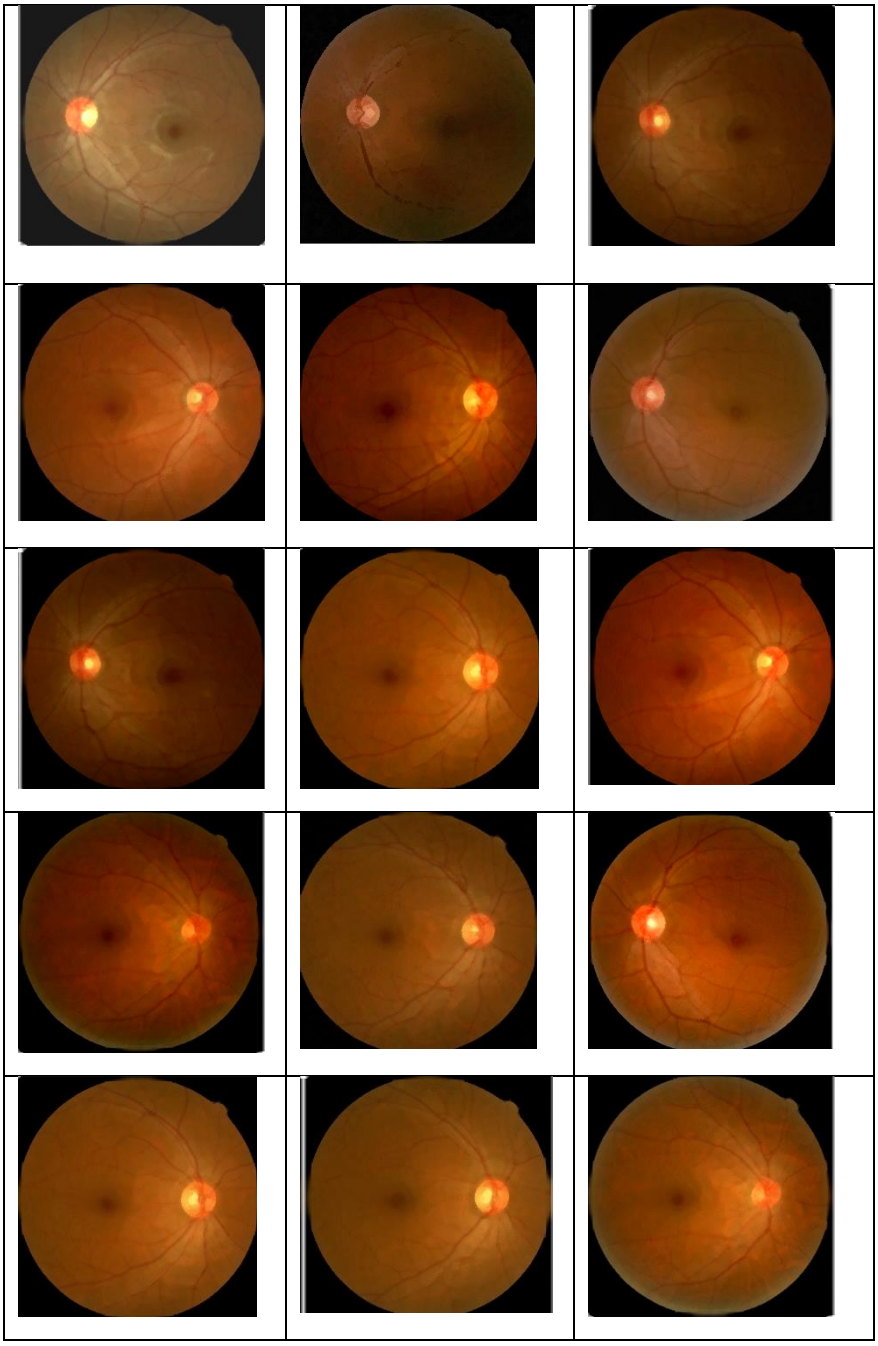
Class 0 **No DR**	**Class 1** **Mild**	**Class 2** **Moderate**	**Class 3** **Severe**	**Class 4** **Proliferative DR**
				
				
				
				
				

INDIAN DIABETIC RETINOPATHY IMAGE DATASET (IDRiD) SAMPLES

Class 0 No DR	Class 1 Mild	Class 2 Moderate	Class 3 Severe	Class 4 Proliferative DR
				
				
				
				
				

ANNEXURE II

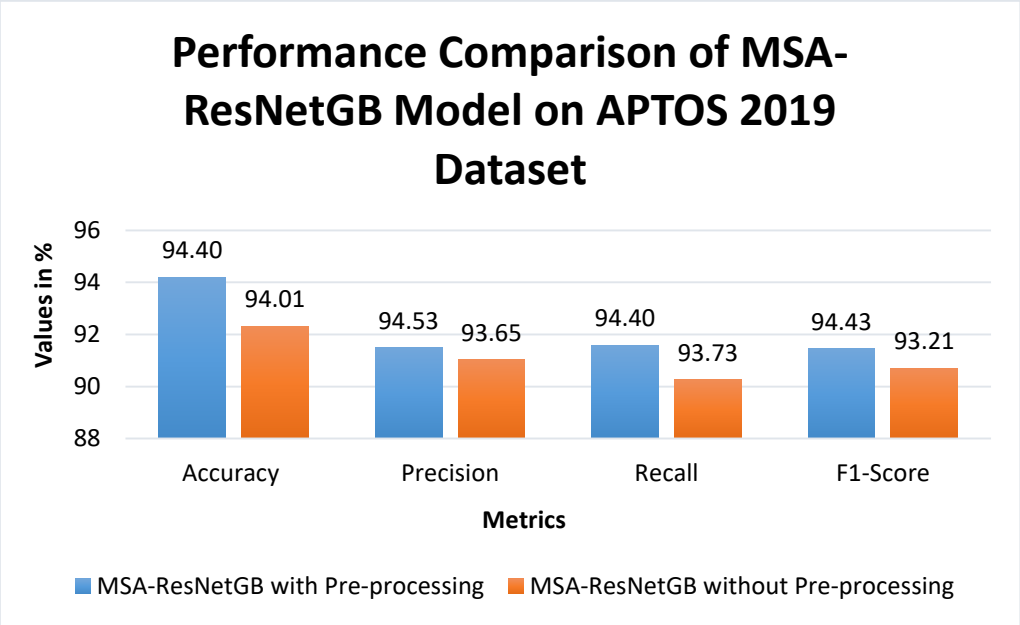
SPECIAL GENERATIVE ADVERSARIAL NETWORK (SGAN) GENERATED IMAGES

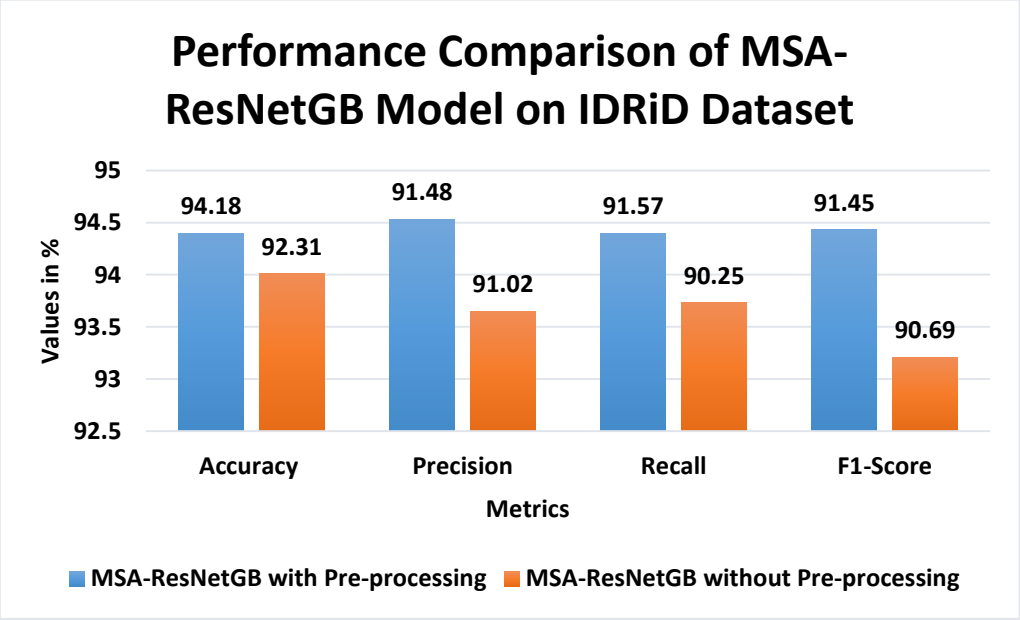


ANNEXURE III

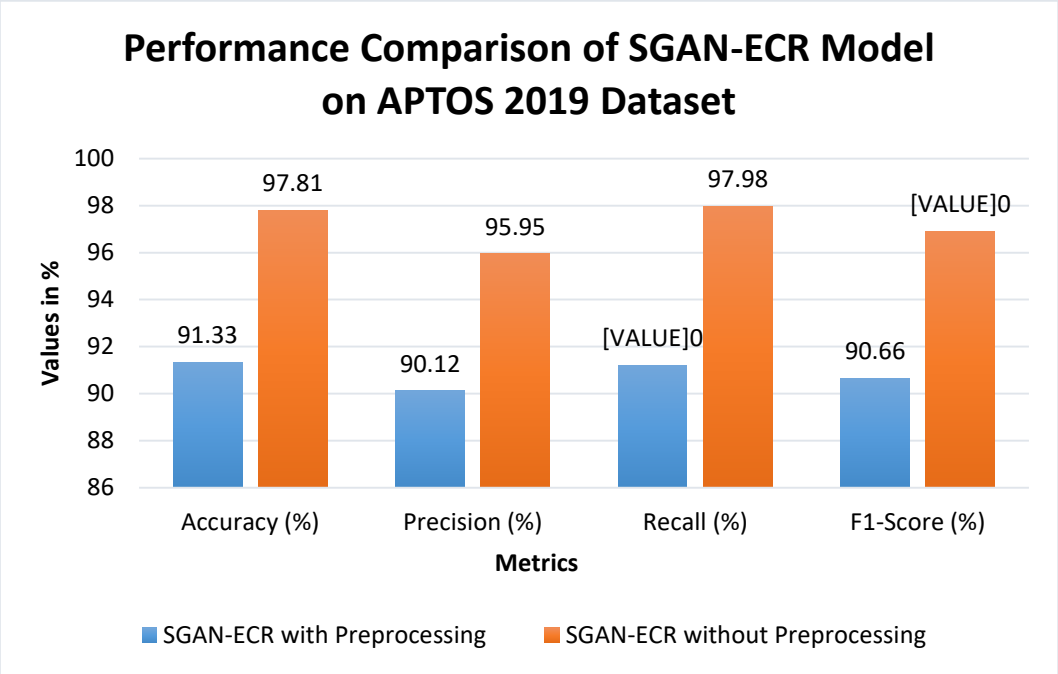
PERFORMANCE COMPARISON OF MODELS ON APTOS AND IDRiD DATASET WITH AND WITHOUT PREPROCESSING

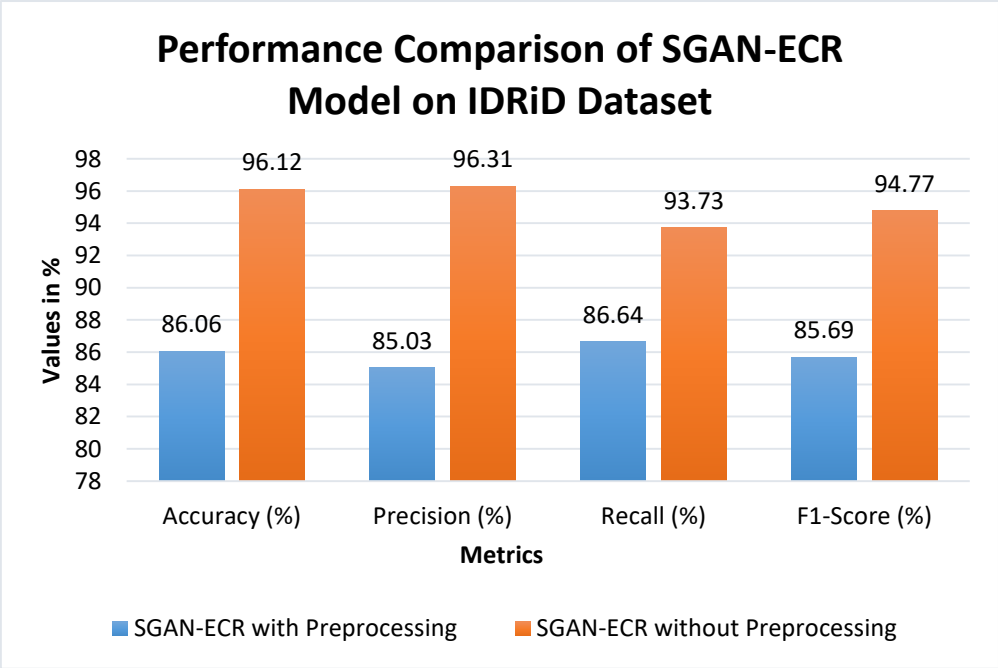
1. PERFORMANCE COMPARISON OF MSA-RESNETGB MODEL



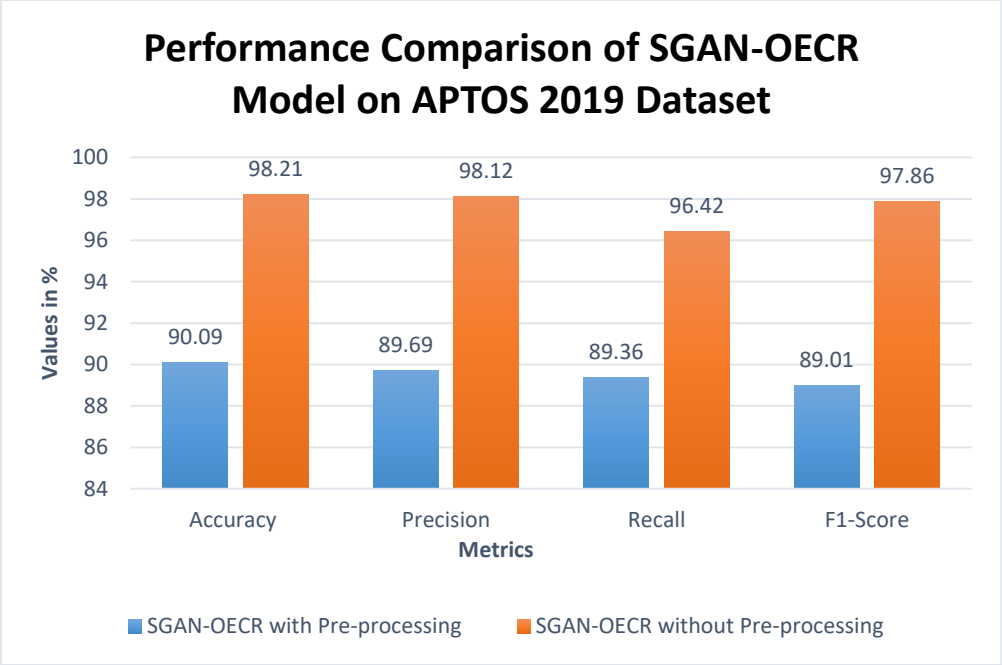


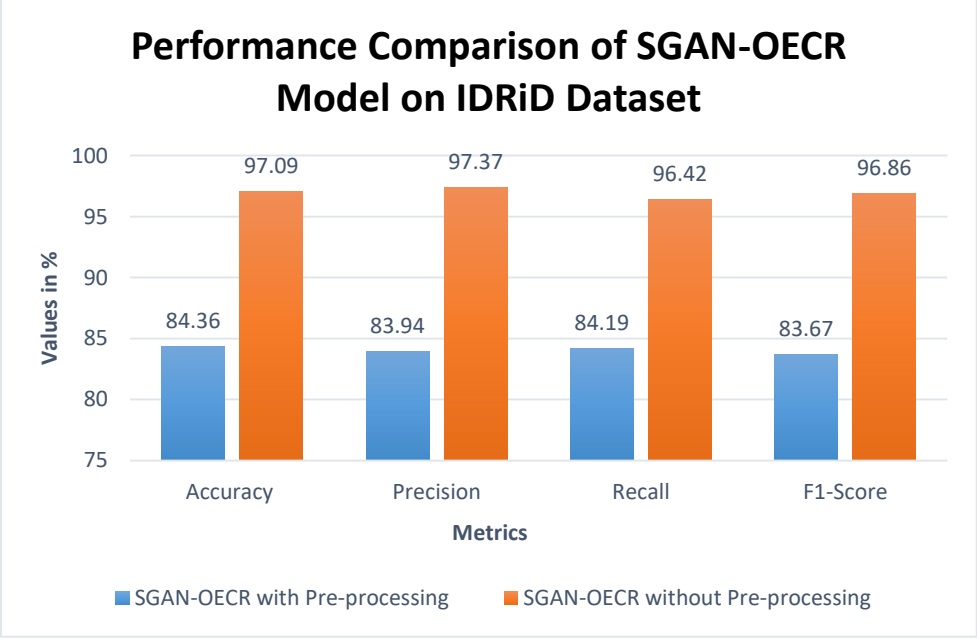
2. PERFORMANCE COMPARISON OF SGAN-ECR MODEL





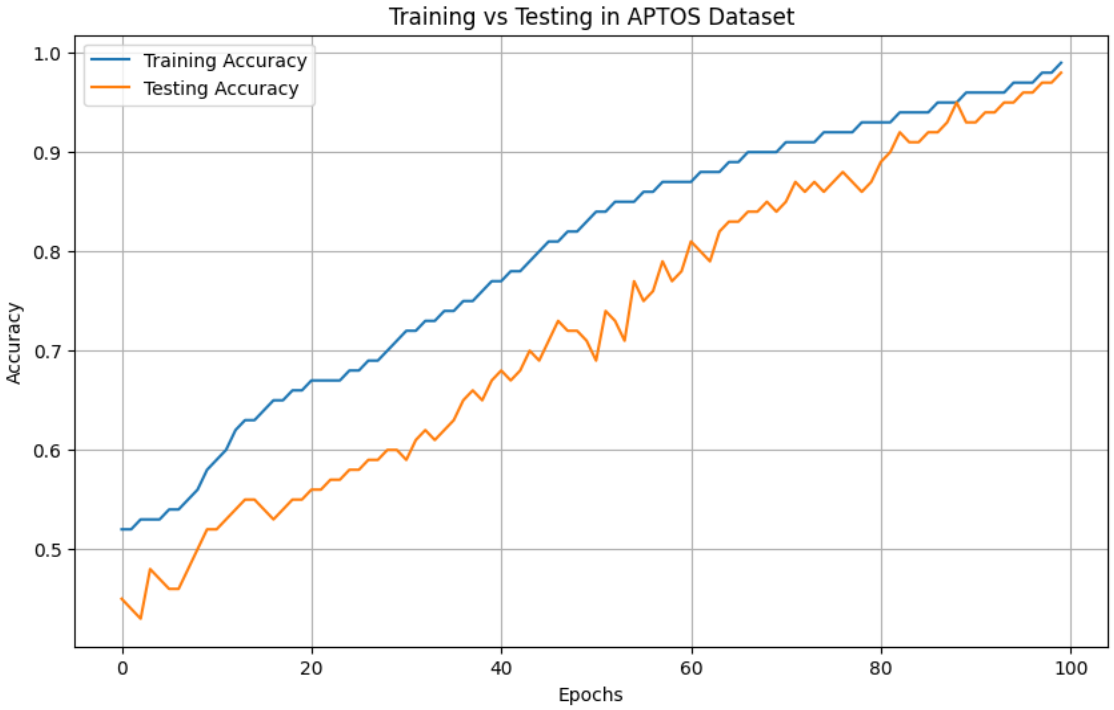
3. PERFORMANCE COMPARISON OF SGAN-OECR MODEL

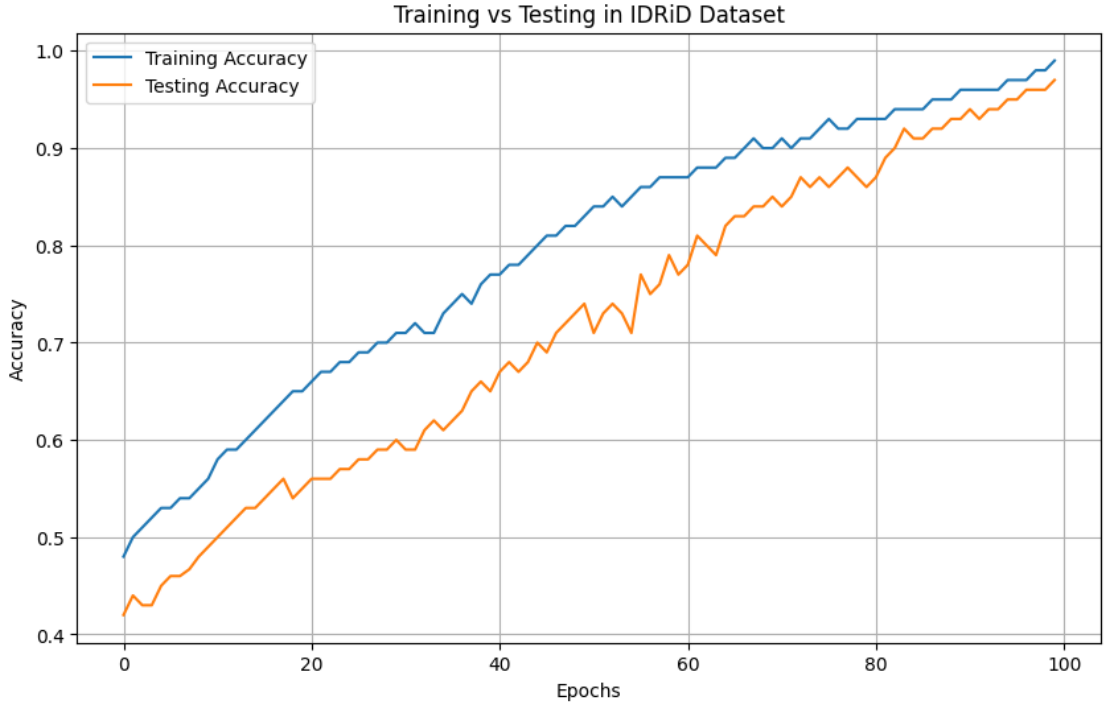




ANNEXURE IV

TRAINING & TESTING PERFORMANCE ON APTOS & IDRID DATASET





PUBLICATIONS

INTERNATIONAL JOURNALS:

1. Valarmathi S and Dr.R.Vijayabhanu, "A Review on Diabetic Retinopathy Disease Detection and Classification using Image Processing Techniques", International Research Journal of Engineering and Technology, Volume 07, Issue 09, 2020. e-ISSN: 2395-0056. **(Peer-reviewed)**

2. S Valarmathi and R Vijayabhanu, "A Comparative Analysis of Contrast Enhancement of Retinal Fundus Images using Histogram-Based Methods"**,** Indian Journal of Natural Sciences, Volume 13, Issue 71, 2022. **(Web of Science, Zoological Records)**

3. Valarmathi Srinivasan and Vijayabhanu Rajagopal, "Multi-Scale Attention-Based Mechanism in Gradient Boosting Convolutional Neural Network for Diabetic Retinopathy Grade Classification", International Journal of Intelligent Engineering & Systems, Volume 15, Issue 04, 2022. DOI: 10.22266/ijies2022.0831.44 **(Scopus-Indexed)**

4. Valarmathi Srinivasan and Vijayabhanu Rajagopal, "An Optimization Method for Tuning Hyper-Parameters of SGAN with Ensemble Classification Regression Model" in SSRG International Journal of Electrical and Electronics Engineering, Volume 10, Issue 4, 2023. DOI: 10.14445/23488379/IJEEE-V10I4P103 **(Scopus-Indexed)**

5. Valarmathi Srinivasan and Vijayabhanu Rajagopal, "Generative Adversarial Network generated images for Diabetic Retinopathy Disease Classification", Bulletin of Electrical Engineering and Informatics**. (Scopus-Indexed, Accepted)**

INTERNATIONAL CONFERENCES:

1. S.Valarmathi and R.Vijayabhanu, "An Efficient Wavelet-based Image Denoising Technique for Retinal Fundus Images", International Conference on Innovation & Sustainability, Intelligent Systems, Algorithms for Intelligent Systems Book series, Springer-Singapore, 2021. **(Springer, Scopus-Indexed)**
   Online ISBN: 978-981-16-2248-9.
   DOI: https://doi.org/10.1007/978-981-16-2248-9_36

2. Valarmathi S and Vijayabhanu R, "A Survey on Diabetic Retinopathy Disease Detection and Classification using Deep Learning Techniques", Seventh International (virtual) Conference on Biosignals, Images, and Instrumentation (ICBSII-IEEE), 2021.
   Electronic ISBN**:** 978-1-6654-4126-1. **(IEEE, Scopus-Indexed)**
   DOI: 10.1109/ICBSII51839.2021.9445163

3. S.Valarmathi and R.Vijayabhanu, "Computer-Aided Diabetic Retinopathy Diagnosis Using Conventional and Deep Learning Techniques—A Comparison", Advance Concepts of Image Processing and Pattern Recognition, Transactions on Computer Systems and Networks Book series, Springer – Singapore, 2022. **(Springer, Scopus-Indexed)**
   Online ISBN: 978-981-16-9324-3.
   DOI: https://doi.org/10.1007/978-981-16-9324-3_8

Avinashilingam Institute for HomeScience and Higher Education for Women
(Deemed to be University Estd. u/s 3 of UGC Act 1956, Category 'A' by MHRD
Re-accredited with A++ Grade by NAAC. CGPA 3.65/4, Category 1 by UGC
Coimbatore - 641 043, Tamil Nadu, India

Appendix L2

(Item No 5 of Check List) Details of Research Publications

S.No	Article	Journal	Other Details Vol/No/Page No/ Year	Published in UGC-CARE / Scopus Indexed/ Web of Science
1	Multi-Scale Attention-based Mechanism in Gradient Boosting Convolutional Neural Network for Diabetic Retinopathy Grade classification.	International Journal of Intelligent Engineering & Systems	vol. 15, No. 4, 2022 pg. no: 489-498	Scopus-Indexed
2	An optimization Method for Tuning Hyper-parameters of SGAN with Ensemble classification Regression Model.	SSRG International Journal of Electrical & Electronics Engineering	vol. 10 Issue 4 24-36 April 2023	Scopus-Indexed

*Proof of list of Journals from Internet to be attached along with copies of reprints.

Scholar : S. C 8/6/23

Supervisor : Vijayabhaud 8/6/23

Verified 27/9/23

Checked By: 27/9/2023

HoD/Dean of Respective School

The scholar Miss. Valarmathi, S has published her article in "International Journal of Intelligent Engineering and Systems" (Vol. 15, No. 4, 2022) which is indexed and active in the Scopus database from

2008 to present and

2. SSRG International Journal of Electrical and Electronics Engineering is also active & indexed in Scopus database from 2019 to present. Her paper was published in Vol. 10, No. 4, April 2023 issue.

08.06.23

The above two journals are indexed and active in Scopus database as of today (27.09.2023).

27.09.23.

The above two journals are indexed and active in Scopus database as of today (02.11.2023). This may be considered.

02.11.2023

Received: April 24, 2022. Revised: June 10, 2022.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Multi-Scale Attention-Based Mechanism in Gradient Boosting Convolutional Neural Network for Diabetic Retinopathy Grade Classification

Valarmathi Srinivasan[1]* **Vijayabhanu Rajagopal[1]**

[1]*Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India*
* Corresponding author's Email: valarmathis313@gmail.com

Abstract: Diabetic Retinopathy (DR) is a common complication of Diabetes Mellitus (DM) that produces retinal abnormalities and can lead to blindness if not diagnosed and treated on time. To address this concern, an adaptive Convolutional Neural Network (CNN) model with Gradient Boosting (GB) called ResNetGB has been used from the literature where a Principal Component Analysis (PCA) based Fully Connected (FC) layer is used to capture the discriminative characteristics from the Retinal Fundus (RF) samples. It is essential to extract more effective features to categorize the DR grades. Hence, in this article, the Multi-Scale Attention (MSA) strategy is incorporated into the ResNetGB model for effective DR grade classification. First, the encoder network is used to embed the RF image in a high-level interpretational space in which the mixture of mid and high-level characteristics is considered to enhance the representation. Then, a Multi-Scale Feature Pyramid (MSFP) is added to define the retinal pattern in various localities and the MSA strategy is applied to the high-level interpretation. Moreover, the entire MSA-ResNetGB framework is trained by the cross-entropy loss to categorize the patients with respective DR grades. Finally, the experimental analysis exhibit that the MSA-ResNetGB model achieves the 94.40% and 94.17% accuracy on two benchmark datasets: Kaggle-APTOS and IDRiD, respectively compared to the cutting-edge models.

Keywords: Diabetic retinopathy, CNN, Gradient boosting, Multi-scale feature pyramid, Attention strategy.

1. Introduction

Diabetes Mellitus is one of the most prevalent causes of DR, which produces vision impairment. It involves varying degrees of severity and is treated when identified ahead of time [1]. DR affects the retina, which is in charge of turning light into an electronic signal that can be processed to produce an image. A network of blood vessels surrounds the retina, supplying it with oxygen and nutrients. Diabetes damages blood vessels, resulting in a shortage of blood supply to the retina. This deteriorates the health of the retina and, as a result, affects visual acuity. The most severe kind of DR is Non-proliferative retinopathy. DM does not influence vision at this level, although it does alter blood vessels. The arteries can expand moderately (microaneurysms – MAs, exudates - EXs, or retinal hemorrhages - HEs) [2].

In the final phase, DR progresses to proliferative retinopathy, which causes greater damage to the retina than non-proliferative retinopathy. When the majority of the retina is deprived of sufficient blood flow, it raises the possibility of visual impairment and poses a danger of vision loss. Manual detection of DR by ophthalmologists or skilled graders is costly. Diagnosis of a huge proportion of diabetic patients for probable DR incidence places a significant strain on ophthalmologists or evaluators, reducing their reliability and avoiding DR diagnoses. The screening is highly subjective, with various graders emerging with many interpretations [3].

To tackle this concern, modified and pre-learned CNN structures are used for the DR classification [4]. The most advanced pre-learned CNN structures, such as VGG and Residual Network (ResNet) [5] and Dual Path Network (DPN) [6] are often learned on ImageNet [7] and convert the low-level

characteristics to high-level characteristics. As DR grading is dependent on the existence of lesions in RF images such as MAs, EXs, and HMs, Saeed et al. [8] developed a 2-phase transfer learning employing a pre-learned CNN structure such as VGG19, ResNet152 and DPN107. The approach encompasses a RF scan as input, analyzes it using the modified framework, and then classifies it as healthy or DR. The CNN structure trains the domain-specific structure of low and high-level characteristics. First, the foremost layer of the pre-trained CNN structure called ResNetGB is reinitialized by the Regions-of-Interest (ROI) of tumors retrieved from the labeled RF samples. For lesion ROI extraction, E-optha is considered since it involves pixel-level tumor labeling. Then, the structure is optimized, wherein the low-level units train the neighborhood patterns of the tumor and healthy areas. Because the FC units convert high-level characteristics, these are substituted by the novel FC unit depending on the PCA and apply it in an unsupervised scheme to capture discriminative characteristics from the RF samples. Also, the GB-based classifier unit is added to estimate the DR scores of RF scans. Although it extracts high-level features to identify and classify the retina lesions into DR classes, highly effective features are essential to categorize the DR severity stage. The following are the proposed work's significant contributions:

1. The ResNetGB with MSA approach is presented to increase the accuracy of DR grade categorization. To begin, the encoder network is utilized to place the RF image in a high-level interpretational space, with a blend of mid and high-level data added to reinforce the interpretation. The MSFP is also built to specify the RF pattern in various locations.
2. The MSA strategy is employed in the high-level interpretation to improve the discriminative power of the feature interpretation.
3. Using the cross-entropy loss, the MSA-ResNetGB model is trained to detect DR patients based on their DR grades. As a result, distinguishing between normal and DR patients is simplified.
4. The proposed MSA-ResNetGB model is tested on two public benchmark datasets: Kaggle-APTOS 2019 dataset and Indian Diabetic Retinopathy Image Dataset (IDRiD).
5. The MSA-ResNetGB model is validated using four performance metrics: Accuracy, Precision, Recall, and F1-Score.

The rest of this paper is prepared as follows: The recent work linked with the DR detection and classification model is discussed in Section 2. Section 3 explains the ResNetGB-MSA model, while Section 4 demonstrates its efficacy. Section 5summarizes this paper and suggests its possible improvement.

2. Related work

A novel modified Xception-based feature mining and Multi-Layer Perceptron (PLP) classification model [9] have been developed to categorize DR criticality and diagnose them efficiently. An ensemble model [10] has been designed, which combines CNN and classical hand-crafted features into a single structure to classify RF images. A framework [11] was developed by considering a 3-channel RF image as input and resulting in the criticality of DR. Also, transfer learning was applied to the pre-trained MobileNetV2 and a weighted loss function was utilized.

A new Cross-disease Attention Network (CANet) [12] was designed to jointly grade DR and diabetic macular edema by finding the internal correlation among the diseases with only image-level supervision. A modified EfficientNet structure [13] was suggested to classify the early and advanced grades of the DR disease.

An improved cross-entropy loss function and three hybrid CNN models [14] have been developed to classify DR. A new CNN structure [15] was introduced to capture characteristics from RF scans. A new technique [16] was presented for DR prognosis depending on the gray-level pixels and fine details from the RF scans by the decision tree-based ensemble training. A composite Deep Neural Network (DNN) structure [17] was integrated with a gated-attention strategy for automated prognosis of DR. Deep transfer learning [18] was investigated based on the AlexNet, GoogleNet, InceptionV4, Inception ResNetV2 and ResNext50 to automatically diagnose DR.

CNN-based computer-aided diagnosis system [19] was developed to categorize RF images into different grades of DR. Two deep learning-based models [20] were developed: a CNN512 was utilized to categorize the RF image into different grades of DR, as well as, an adopted YOLOv3 was utilized to identify and localize the DR lesions.

A multi-task deep learning system [21] was developed using a modified Squeeze Excitation densely connected DNN and Xecption network as a multitasking scheme. Also, the MLP was used as a classification to categorize the DR grades. A new deep learning hybrid model [22] was developed, in which transfer learning was applied on pre-trained Inception-resNetV2 and a custom module of CNN layers was added on top of Inception-resNetV2 to recognize DR disorders.

A new early blind recognition technique [23] was designed using the color information obtained from RF images based on the ensemble learning scheme such as Extra tree model. Different deep learning-based classifiers [24] have been analyzed such as VGG16, ResNet50, InceptionV3 and DenseNet121 to partition the retina areas and categorize DR severity grades. The MSA Network (MSA-Net) [25] based on the encoder and decoder structure was designed for DR categorization. A Graph Neural Network (GNN) model [26] was designed to categorize DR severity.

2.1 Problem definition

From the above-studied related works, the issues in classifying the DR severity grades are:

- The number of training and testing images was limited, which impacts the accuracy of classifying the DR grades.
- Also, the accuracy was degraded because of imbalanced classes in the RF image databases.
- The imbalanced classes affect the significance of the feature maps during training and classification.
- Some CNN structures trained by only image-level supervision, which makes the model very difficult to recognize the exact abnormal signs like soft EXs, hard EXs, MAs and HEs.
- The classification efficiency mainly relies on the RF image resolution and the optimal hyperparameters of CNN structures such as learning rate, number of epochs, number of layers and batch size.
- Though high-level features were extracted, additional features are essential to enhance the accuracy of classifying the different DR grades.

2.2 Research contribution

This research focuses on enhancing the accuracy of classifying the different grades of DR severity.

Table 1. List of notations

Notations	Description
G	Encoder structure
θ_1	Encoding variable
I	Retinal fundus image
F_{ec}	Interpretation tensor
$\mathcal{A}$	Attention tensor
$\mathbb{K}_{pw}$	Point-wise convolution kernel
h	Height of the feature map
w	Width of the feature map
c	Number of channels
σ	Sigmoid activation
$\odot$	Point-wise multiplication
F	Absolute feature interpretation vector
θ_2	Learning variables for the multi-scale and attention units
$\mathcal{L}(\theta, \varphi)$	Loss factor

To achieve this task, a MSA-based ResNetGB model is proposed. First, the RF image database is obtained and fed to the encoder network to create the interpretation tensor. Then, the MSFP is applied to represent high-level features at various scales. Also, an attention strategy is performed to get the attention maps and multiply them with the high-level feature map representations to obtain the final feature interpretation. Further, the global interpretation is created and classified into various classes of DR grades.

3. Materials and methods

In this research work, the ResNetGB model [8] is used as a baseline model under different system settings. This section explains the MSA-ResNetGB model for DR severity categorization briefly. Table 1 presents the notations used in this study.

Pseudo code for the proposed MSA-ResNetGB model:

Input: ROI lesions from the E-Optha dataset and RF images $I_1, \ldots, I_n$ (from Kaggle-APTOS and IDRiD)

Output: DR grade classification (0-No DR, 1-Mild DR, 2-Moderate DR, 3-Severe DR and 4-Proliferative DR)

***Step 1*:** Collect the annotated RF images and split the images into training and test set.

Step 2: In the training set, embed the collected images into the high-level representational space using the ResNetGB encoder; generate the interpretation tensor F_{ec} as:

$$F_{ec} = G(\theta_1; I) \tag{1}$$

Step 3: Add the attention strategy to the multi-scale feature interpretation;

Step 4: Improve the discriminative ability of high-level feature interpretation; Use the point-wise convolution among the pyramid representations.

Step 5: Generate the attention tensor $\mathcal{A}$ as:

$$\mathcal{A}^{h\times w\times 64} = \left(F_{all}^{h\times w\times 4c} * \mathbb{K}_{pw}\right) \quad (2)$$

Step 6: Obtain the global interpretation by the GMP and create the final feature interpretation vector F as:

$$F^{1\times 1024} = \frac{GMP\left(\sigma\left(F^{h\times w\times c} \odot \mathcal{A}^{h\times w\times 1}\right)\right)}{GMP\left(\mathcal{A}^{h\times w\times 1}\right)} \quad (3)$$

Step 7: Map the feature vectors using the PCA layer to the required outcomes;

Step 8: Train the classification layer using GB classifier and calculate the loss factor $\mathcal{L}\ (\theta,\ \varphi)$ and adjust the training variables.

Step 9: Classify the test images based on DR severity grades using the trained MSA-ResNetGB model.

3.1 ResNetGB as encoder

The initial unit of the model is the encoder module. As shown in Fig. 1, the ResNetGB structure is utilized as the encoder in this model to insert the RF images in the high-level interpretational space. The model efficiency is enhanced with the help of the number of units in the deep learner. On the other hand, there is a challenge in this network called vanishing gradients. So, this challenge is resolved by the short links in the ResNetGB model, i.e. direct paths among the outcome of all layers including the input of the nearby unit. The ResNetGB trains the residuals.

As the ResNetGB is comparatively simple to adopt, the efficiency is improved by incorporating additional units. The initial unit is $7 * 7$ convolutional layer followed by the four different units which contain 3, 8, 12 and 3 residual blocks, correspondingly. The last unit comprises a mean pooling layer. It is observed that the ResNetGB architecture is utilized without an FC layer, i.e. PCA layer as the encoder of the MSA model.

The encoder structure G having an encoder variable θ_1 in this ResNetGB-MSA model considers the RF scan I and creates the interpretation tensor (F_{ec}) given in Eq. (1).

3.2 Multi-scale feature extraction and interpretation

The characteristics are obtained from the series of residual blocks in the encoder module. Such obtained characteristics near the input scans include greater resolution and so contain more data regarding the neighborhood characteristics when the characteristics near the final unit have multiple semantic data. Then, multi-level features such as mid and high-level characteristics are concatenated to use both types of data in the consecutive phases. As such characteristics contain varying spatial resolutions, a scaling method is applied to generate them all of the same size. Next, a set of various features is passed through an atrous convolution to retrieve features with varied scales. Convolutional filters with varying view dimensions are used in the atrous convolution. The network can encode more

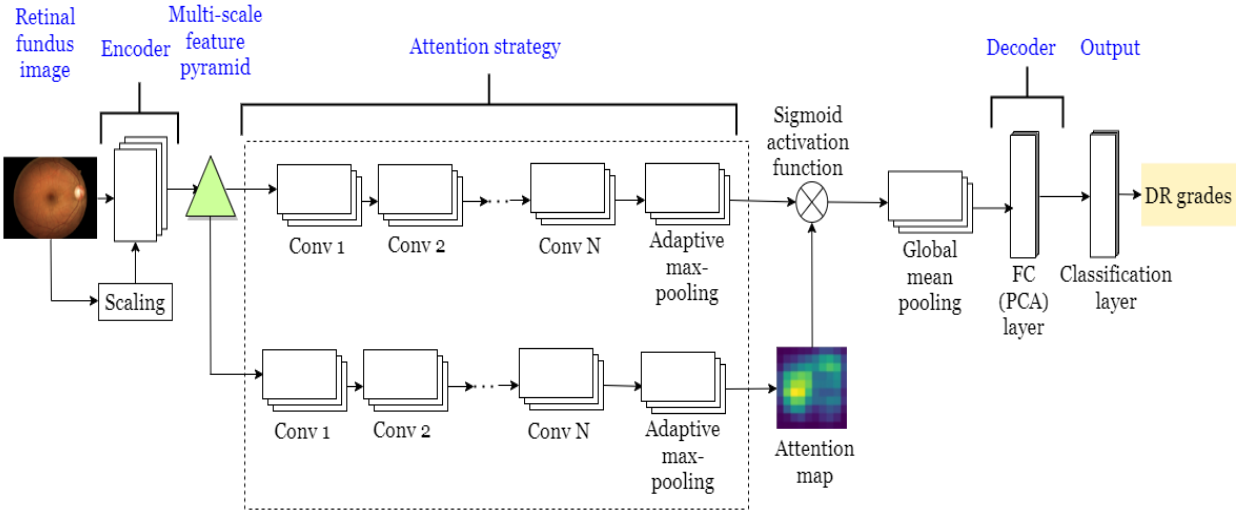


Figure. 1 Multi-scale attention-based mechanism for diabetic retinopathy grade categorization

local details by using a limited view for feature extraction. But, more global details are considered for a larger image context. The resulting multi-scale feature interpretation encapsulates the input scan in a small subspace, allowing it to train DR symptoms of varying sizes, localization and severity.

3.3 MSA strategy

Based on the DR criticality grade, the pattern of the retina is altered. The alteration can induce specific impairment to the RF. To detect such impairments, the high-level feature interpretation is considered which differentiates various classes. However, its efficiency is not satisfactory because of the lack of diabetic patterns. Therefore, the attention strategy is added to the multi-scale feature interpretation to improve the discriminative ability of high-level feature interpretation. The main goal of this attention strategy is to learn where to search for visual loss and adjust the feature space. Particularly, the attention strategy in this MSA-ResNetGB structure focuses on the abnormal regions in the RF images while ignoring the healthy areas.

In this MSA-ResNetGB model, the attention strategy is a sequence of convolution units used for the multi-scale feature interpretation. Initially, the point-wise convolution is used among the pyramid representations to generate a small interpretation as given in Eq. (2). In Eq. (2), $\mathcal{A}$ indicates the attention tensor, F is the interpretation tensor created using the multi-scale unit, $\mathbb{K}_{pw}$ is the point-wise convolution kernel, h, w are the height and width of the feature maps and c is the number of channels. Subsequently, the created small interpretation is applied to the sequence of convolution to produce the attention map $\mathcal{A}^{h\times w\times 1}$ which is multiplied by the high-level interpretation $F^{h\times w\times c}$ to limit the interpretation. To normalize the outcome range from 0 to 1, sigmoid activation (σ) is used. Moreover, the global interpretation is retrieved by the Global Mean Pooling (GMP) of the final interpretation, which is normalized by the GMP data of $\mathcal{A}^{h\times w\times 1}$. So, the created absolute feature interpretation vector F as given in Eq. (3).

In Eq. (3), $\odot$ defines point-wise multiplication. Significantly, the learning variables for the multi-scale and attention units are defined as θ_2. The MSA strategy enhances the training and helps to increase the precision of RF image classification depending on DR criticality grade because it uses the outcomes of the preceding units with different significance. It differentiates this presented framework from the ResNetGB or other CNN structures that do not examine discrepancy.

3.4 Decoder and classification units

The decoder module has the FC layers called PCA layers to map the feature vectors to the required outcomes. The FC units of the pre-learned structure are trained from the ImageNet database to convert the global high-level characteristics, which are not appropriate to the usual and tumor characteristics in the RF scans. Also, the FC units include a massive amount of trainable variables. To solve this problem, all FC units are discarded from the modified structure and included the PCA unit has 153 neurons to minimize the training difficulty. Such 153 neurons are chosen by the greedy method depending on ROIs and the modified ResNet model. Thus, the PCA unit retrieved the characteristics associated with the common and tumor patterns in the RF scans. Those obtained characteristics are then classified by the categorization unit that applies the GB classifier to predict whether the given RF scans belong to healthy people or the DR patients with their severity grades.

The key intention of these units is to categorize the RF images into healthy and DR with severity grades. So, the loss factor $\left(\mathcal{L}(\theta,\varphi)\right)$ is represented as the categorization error for the MSA-ResNetGB structure with encoder and attention variables $(\theta = \theta_1 + \theta_2)$ along with the categorization branch variable φ. The cross-entropy error is implemented in both the estimated and the actual labels. Moreover, the non-learnable weight is added in $\left(\mathcal{L}(\theta,\varphi)\right)$ to reduce the significance of all class errors on the absolute error rates. The error aims to limit the impact of imbalanced scans during the learning phase.

Because the vital goal of automated DR identification is to support the ophthalmologist and decrease the screening complexity, this framework is developed to assist the physician in detecting DR. Also, labeling the normal and abnormal without accurate DR grades is much simpler than proper grading for the physicians. So, this weak labeling is handled conveniently. The secondary process is learned with variables $\left(\mathcal{L}(\theta,\varphi)\right)$ by the cross-entropy error factor. As well, the image augmentation schemes including resizing and flipping are involved in the training phase to prevent overfitting. For this reason, the ROIs around the lesions with various dimensions (16×16, 32×32 and 64×64) are initially extracted, which comprise various context data. After that, those are resized to the equal dimension and all ROIs are rotated in multiple directions $\left[40^{\circ}, 120^{\circ}, 180^{\circ}, 275^{\circ}\right]$ as well as flipped horizontally.

4. Experimental analysis

The MSA-ResNetGB model is evaluated using four distinct measures, including Accuracy, Precision, Recall, and F1-Score, on two different benchmark datasets, Kaggle-APTOS and IDRiD. The system is learned for 120 epochs employing adam optimization with a batch of 3 and a training rate of 10^{-4}.

4.1 Dataset description

Two benchmark datasets Kaggle-APTOS and IDRiD were used in this experimental analysis section to evaluate the proposed model. This subsection further describes the dataset in detail, including the number of collected samples, annotated samples, DR grades and ground truths. The sample RF images from both datasets are shown.

4.1.1. APTOS 2019 blindness detection dataset:

The RF images from the Asia Pacific Tele-Ophthalmology Society (APTOS) 2019 Blindness Detection dataset were used in this study [27]. This Kaggle image collection comprises 3662 samples obtained from a diverse range of rural Indian people. Aravind Eye Hospital in India collected and organized the data. However, a panel of medical experts analyzed and classified the collected samples using the International Clinical Diabetic Retinopathy Disease Severity Scale (ICDRSS). The Kaggle-APTOS dataset samples are classified into five groups on the scale of 0-4, No DR, Mild DR, Moderate DR, Severe DR, and Proliferative DR. The first classification comprises healthy RF samples that do not have DR. Each of the subsequent classifications contains more defective retinas than the previous class. The last classification, proliferative DR, includes samples with vitreous or pre-retinal HEs. Fig. 2 shows the RF samples from each class in Kaggle-APTOS.

4.1.2. Indian diabetic retinopathy image dataset (IDRiD):

IDRiD sub-challenge 2 from the IEEE ISBI - 2018 has been employed in this study [28]. It has 516 images with a range of clinical states of DR and DME, including 413 and 103 training and test images, respectively under Disease grading. IDRiD is the first dataset to represent an Indian population. Fig. 3 shows the RF samples from each class in the IDRiD dataset. Each sample in the IDRiD collection is annotated with Diabetic Retinopathy and Diabetic Macular Edema severity grades at a pixel level. Based on the severity scale, the DR grade is labeled into five classes on the scale of 0 – 4, categories similar to the Kaggle-APTOS dataset.

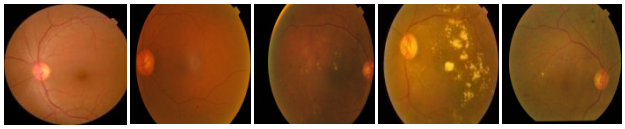


Figure. 2 APTOS 2019 blindness detection dataset samples from each class (0-4)

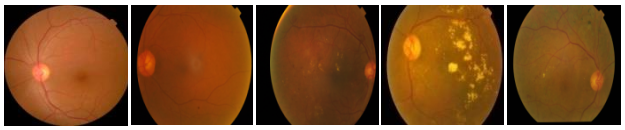


Figure. 3 Indian diabetic retinopathy image dataset (IDRiD) samples from each class (0-4)

4.2 Performance metrics

The performance of the MSA-ResNetGB model is examined with the cutting-edge models using the following measures: Accuracy in Eq. (4), Precision in Eq. (5), Recall in Eq. (6) and F1-Score in Eq. (7). The following are the mathematical formulae utilized to compute the metrics:

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

Whereas True Positive (TP) is the accurate categorization of the number of samples as positive, True Negative (TN) is the correct classification of the number of samples as negative. Further, False Positive (FP) denotes the ratio of negative class samples categorized as a positive class, while False Negative (FN) denotes the ratio of positive class samples classified as negative class.

4.3 Evaluation

In this section, the proposed MSA-ResNetGB model is validated using the performance measures: Accuracy, Precision, Recall and F1-Score on two public-benchmark datasets. The MSA-ResNetGB model was trained on two different datasets to classify the DR images based on the severity grades.

Table 2. Performance analysis of the proposed model and the cutting-edge models for DR classification on the Kaggle-APTOS 2019 blindness detection dataset

Reference	Year	Classification method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
[19]	2019	CNN model	77.00	-	-	-
[9]	2019	Modified Xception model	83.09	-	88.24	-
[17]	2021	Composite gated attention DNN	82.54	82.00	83.00	82.00
[20]	2021	CNN512 model	84.10	-	-	-
[21]	2021	Xception Multitask model	86.00	77.00	70.00	73.00
[22]	2021	Hybrid Inception ResNet-v2	82.18	-	-	-
[23]	2019	ExtraTree model	91.07	90.40	89.54	89.97
[11]	2020	MobileNetV2 model	78.47	68.66	60.01	64.04
[24]	2021	DenseNet121 model	90.50	93.00	90.00	88.47
[16]	2021	Tuned XGBoost model	94.20	94.34	92.68	93.51
[25]	2021	MSA-Net model	84.60	-	91.00	-
Proposed	**2022**	**MSA-ResNetGB model**	**94.40**	**94.52**	**94.40**	**94.42**

In both datasets, multi-class classification is performed as the dataset comprises five classes (0 – 4). Further, APTOS and IDRiD experimental results are discussed in subsequent subsections, which include the performance comparison table and confusion matrices. The class-wise performance evaluation on APTOS and IDRiD are discussed.

4.3.1. APTOS results:

In this subsection, the proposed model has trained on Kaggle-APTOS 2019 Blindness Detection dataset and is evaluated to obtain DR classification results. The Kaggle-APTOS dataset consists of 3662 training samples in which 10% of labeled samples are taken as test data. The multi-class model is validated by comparing it with the latest literature work. Table 2 shows the performance analysis of the proposed model and the cutting-edge models for DR classification on the Kaggle-APTOS 2019 Blindness Detection dataset. The most commonly used metric in the literature is accuracy and some of the other metrics which are not reported in the literature are denoted as '-'.

Table 2 aggregates and compares the recent research work on the Kaggle-APTOS dataset. It is found from the literature that CNN is the most popularly used deep learning method for medical image analysis. In Table 2, eleven distinct classification models from recent literature were utilized to compare the performance of the MSA-ResNetGB model on the Kaggle-APTOS dataset using four metrics: Accuracy, Precision, Recall, and F1-Score. According to Table 2, the proposed method surpassed the cutting-edge models on the Kaggle-APTOS. CNN models such as CNN [19] and CNN512 [20] obtained 77% and 84% accuracy, respectively, which is 17.4% and 10.3% less than the proposed model. The CNN variant models mentioned in the literature [9, 11, 17] and [21-25] performed better but not greater than the proposed model.

Moreover, the proposed model surpassed the Hybrid Inception ResNet-v2 [22] by 12.22%. This proves the proposed model's robustness. The confusion matrix for the proposed MSA-ResNetGB model on the APTOS dataset is shown in Fig. 4. It provides the distribution of samples by class, the ratio of accurately classified and the misclassified samples. In class 0, 70 samples were correctly identified as having no DR, 70 samples were correctly classified in class 1 as having mild DR, 71 samples were correctly classed in class 2 as having moderate DR and 69 samples were correctly classified in class 3 as having severe DR, as well as, 70 samples were correctly classified in class 4 as having PDR.

4.3.2. IDRiD results:

In this subsection, the proposed model is trained on IDRiD and is evaluated to obtain DR classification results. The IDRiD consists of 516 samples in which 20% of labeled samples are taken as test data. The multi-class model is validated by comparing it with the latest literature work.

Table 3 shows the performance analysis of the proposed model and the cutting-edge models for DR classification on the IDRiD. The most commonly used metric in the literature is accuracy and some of the other metrics which are not reported in the literature are denoted as '-'.

Table 3 aggregates and compares the recent research work on the IDRiD. In Table 3, four

Table 3. Performance analysis of the proposed model and the cutting-edge classification models for DR classification on the IDRiD dataset

Reference	Year	Classification method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
[10]	2019	AlexNet model	90.07	-	-	-
[12]	2020	CANet model	65.10	-	-	-
[13]	2021	EfficientNet B0 model	86.00	-	-	-
[26]	2019	GNN-based model	79.30	-	-	-
Proposed	**2022**	**MSA-ResNetGB model**	**94.17**	**91.48**	**91.57**	**91.45**

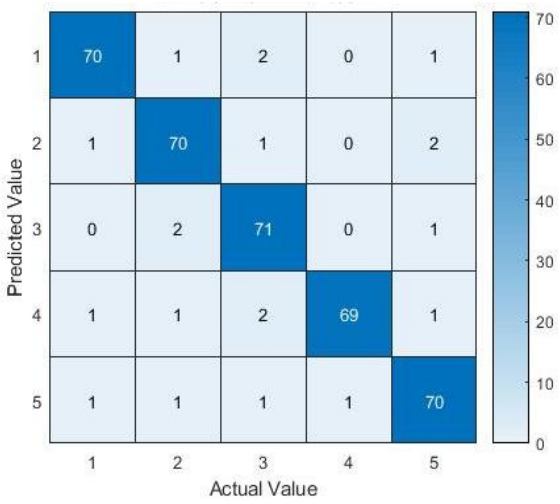


Figure. 4 Confusion matrix for MSA-ResNetGBmodel on the Kaggle-APTOS 2019blindness detection dataset

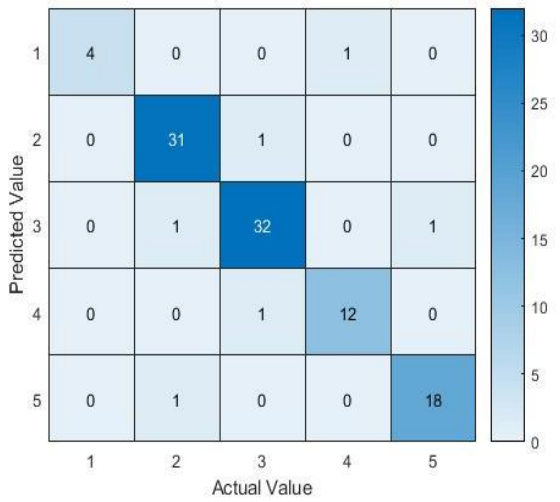


Figure. 5 Confusion matrix for MSA-ResNetGB model on the IDRiD

distinct classification models from recent literature were used to analyze the performance of the MSA-ResNetGB model on the IDRiD using four metrics: Accuracy, Precision, Recall, and F1-Score. According to Table 3, CNN and the pre-trained models [10, 12, 13, 26] surpassed the proposed model on the IDRiD. The proposed model obtained 4.1% greater than AlexNet model [10], 29.07% greater than CANet model [12], 8.17% greater than EffiecientNet B0 model [13] and 14.87% greater than GNN-based model [26]. The confusion matrix for the MSA-ResNetGB model on the IDRiD is shown in Fig. 5.

The matrix shown in Fig. 5 provides the distribution of image samples by class, as well as the ratio of accurately classified and misclassified samples. In class 0, 4 samples were correctly identified as having no DR, 31 samples were correctly classified in class 1 as having mild DR, 32 samples were correctly classed in class 2 as having moderate DR and 12 samples were correctly classified in class 3 as having severe DR, as well as, 18 samples were correctly classified in class 4 as having PDR.

4.4 Discussion

The empirical values of the proposed MSA ResNetGB model on the Kaggle-APTOS and IDRiD datasets are discussed and validated in this section. The model performance on the two datasets was analyzed and evaluated individually with the cutting-edge models in the previous sections. According to the previous section, it is discovered that both benchmark datasets outperformed the literature models. The model learns the DR structure with varied feature selection and localization due to the combination of local and global feature representations, resulting in improved performance. The multi-scale attention strategy, which is implemented in the high-level interpretation space, focuses on the key region in identifying the DR severity levels. Table 4 compares the performance of the MSA-ResNetGB model on the Kaggle-APTOS and IDRiD datasets.

The MSA-ResNetGB model on the Kaggle-APTOS dataset achieved 94.40% accuracy, 94.53% precision, 94.40% recall, and 94.43% F1-Score, which is higher than the performance on the IDRiD dataset, which achieved 94.18% accuracy, 91.48% precision, 91.57% recall, and 91.45% F1-Score.

When compared to the Kaggle-APTOS dataset, the performance loss in IDRiD is mostly due to a lack of image samples. For computer vision and

Table 4. Performance of MSA-ResNetGB model on the Kaggle-APTOS and IDRiD dataset

Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
APTOS	94.40	94.53	94.40	94.43
IDRiD	94.18	91.48	91.57	91.45

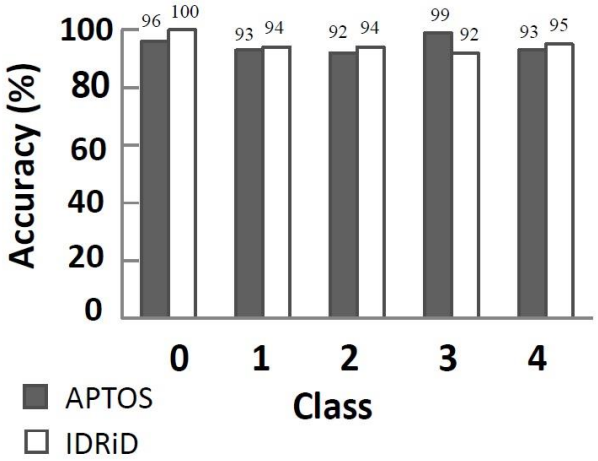


Figure. 6 Class-wise performance evaluation of MSA-ResNetGB model on Kaggle-APTOS and IDRiD dataset

images, data augmentation strategies produce effective outcomes. Fig. 6 represents the class-wise performance evaluation of the MSA-ResNetGB model on the Kaggle-APTOS and IDRiD datasets.

As previously stated, these datasets have five distinct classes labeled from 0 to 4. According to Fig. 6, on the Kaggle-APTOS dataset, 96%, 93%, 92%, 99% and 93% of accuracy were attained in classes 0, 1, 2, 3, and 4, respectively. On the IDRiD dataset, accuracy rates of 100%, 94%, 94%, 92%, and 95% were reached for classes 0, 1, 2, 3, and 4, respectively. The Kaggle-APTOS dataset shows that class 3 performs better while class 2 performs less. In the IDRiD dataset, however, class 0 generates better results while class 3 produces lower results.

5. Conclusion

In this paper, the MSA-ResNetGB model was proposed to detect and classify RF samples based on DR severity grades. Initially, the encoder network was used to integrate the RF picture into the high-level feature interpretation space. The MSFP model was then used to describe the RF pattern at various scales. In addition, the MSA strategy was used on the high-level interpretation to increase the feature interpretation's differentiation efficiency. Further, the whole MSA-ResNetGB structure was trained on the cross-entropy error to correctly characterize the DR criticality grades. Finally, when validated with cutting-edge models to classify DR severity classes, the MSA-ResNetGB test results on Kaggle-APTOS and IDRiD revealed 94.40% and 94.17% accuracy, respectively.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, methodology, software, validation, Valarmathi Srinivasan; formal analysis, investigation, Vijayabhanu Rajagopal; resources, data curation, writing—original draft preparation, Valarmathi Srinivasan; writing—review and editing, Vijayabhanu Rajagopal; visualization; supervision, Vijayabhanu Rajagopal;

References

[1] D. J. Eszes, D. J. Szabo, G. Russell, C. Lengyel, T. Varkonyi, E. Paulik, and B. E. Petrovski, "Diabetic retinopathy screening in patients with diabetes using a handheld fundus camera: the experience from the south-eastern region in Hungary", *Journal of Diabetes Research*, pp. 1-9, 2021.

[2] S. D. Solomon and M. F. Goldberg, "ETDRS grading of diabetic retinopathy: still the gold standard?", *Ophthalmic Research*, Vol. 62, No. 4, pp. 190-195, 2019.

[3] K. S. Sreejini and V. K. Govindan, "Retrieval of pathological retina images using bag of visual words and pLSA model", *Engineering Science and Technology, an International Journal*, Vol. 22, No. 3, pp. 777-785, 2019.

[4] Z. Gao, J. Li, J. Guo, Y. Chen, Z. Yi, and J. Zhong, "Diagnosis of diabetic retinopathy using deep neural networks", *IEEE Access*, Vol. 7, pp. 3360-3370, 2018.

[5] M. Aatila, M. Lachgar, H. Hrimech, and A. Kartit, "Diabetic retinopathy classification using ResNet50 and VGG-16 pretrained networks", *International Journal of Computer Engineering and Data Science*, Vol. 1, No. 1, pp. 1-7, 2021.

[6] Y. Chen, J. Li, H. Xiao, X. Jin, S. Yan, and J. Feng, "Dual path networks", In: *Proc. of Advanced Neural Information Processing Systems*, pp. 1-11, 2017.

[7] S. Kornblith, J. Shlens, and Q. V. Le, "Do better imagenet models transfer better?", In: *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, pp. 2661-2671, 2019.

[8] F. Saeed, M. Hussain, and H. A. Aboalsamh, "Automatic diabetic retinopathy diagnosis using adaptive fine-tuned convolutional neural network", *IEEE Access*, Vol. 9, pp. 41344-41359, 2021.

[9] S. H. Kassani, P. H. Kassani, R. Khazaeinezhad, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Diabetic retinopathy classification using a modified xception architecture", In: *Proc. of the IEEE International Symposium on Signal Processing and Information Technology*, pp. 1-6, 2019.

[10] B. Harangi, J. Toth, A. Baran, and A. Hajdu, "Automatic screening of fundus images using a combination of convolutional neural network and hand-crafted features", In: *Proc. of the 41st Annual International Conf. of the IEEE Engineering in Medicine and Biology Society*, pp. 2699-2702, 2019.

[11] L. Wang and A. Schaefer, "Diagnosing diabetic retinopathy from images of the eye fundus", *Cs230. Stanford. Edu*, 2020.

[12] X. Li, X. Hu, L. Yu, L. Zhu, C. W. Fu, and P. A. Heng, "CANet: cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading", *IEEE Transactions on Medical Imaging*, Vol. 39, pp. 1483-1493, 2020.

[13] E. Abdelmaksoud, S. Barakat, and M. Elmogy, "Diabetic retinopathy grading system based on transfer learning", *International Journal of Advanced Computer Research*, Vol. 11, No. 52, pp. 1-12, 2021.

[14] H. Liu, K. Yue, S. Cheng, C. Pan, J. Sun, and W. Li, "Hybrid model structure for diabetic retinopathy classification", *Journal of Healthcare Engineering*, Vol. 2020, pp. 1-9, 2020.

[15] S. Gayathri, V. P. Gopi, and P. Palanisamy, "A lightweight CNN for diabetic retinopathy classification from fundus images", *Biomedical Signal Processing and Control*, Vol. 62, pp. 1-11, 2020.

[16] N. Sikder, M. Masud, A. K. Bairagi, A. S. M. Arif, A. A. Nahid, and H. A. Alhumyani, "Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images", *Symmetry*, Vol. 13, No. 4, pp. 1-26, 2021.

[17] J. D. Bodapati, N. S Shaik, and V. Naralasetti, "Composite deep neural network with gated-attention mechanism for diabetic retinopathy severity classification", *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-15, 2021.

[18] H. Tariq, M. Rashid, A. Javed, E. Zafar, S. S. Alotaibi, and M. Y. I. Zia, "Performance analysis of deep-neural-network-based automatic diagnosis of diabetic retinopathy", *Sensors*, Vol. 22, No. 1, pp. 1-15, 2022.

[19] O. Dekhil, A. Naglah, M. Shaban, M. Ghazal, F. Taher, and A. Elbaz, "Deep learning based method for computer aided diagnosis of diabetic retinopathy", In: *Proc. of the IST 2019-IEEE International Conf. on Imaging Systems and Techniques*, pp. 1-4, 2019.

[20] W. L. Alyoubi, M. F. Abulkhair, W. M. Shalash, "Diabetic retinopathy fundus image classification and lesions localization system using deep learning", *Sensors*, Vol. 21, No. 11, pp. 1-22, 2021.

[21] S. Majumder and N. Kehtarnavaz, "Multitasking deep learning model for detection of five stages of diabetic retinopathy", *IEEE Access*, Vol. 9, pp. 123220-123230, 2021.

[22] A. K. Gangwar and V. Ravi, "Diabetic retinopathy detection using transfer learning and deep learning", *Evolution in Computational Intelligence*, pp. 679-689, 2021.

[23] N. Sikder, M. S. Chowdhury, A. S. M. Arif, and A. A. Nahid, "Early blindness detection based on retinal images using ensemble learning", In: *Proc. of the 22nd International Conf. on Computer and Information Technology*, pp. 1-6, 2019.

[24] S. Sheikh and U. Qidwai, "Smartphone-based diabetic retinopathy severity classification using convolution neural networks", In: *Proc. of the Advances in Intelligent Systems and Computing*, Vol. 1252, pp. 469-481, 2021.

[25] M. T. A. Antary and Y. Arafa, "Multi-Scale attention network for diabetic retinopathy classification", *IEEE Access*, Vol. 9, pp. 54190-54200, 2021.

[26] A. Sakaguchi, R. Wu, and S. I. Kamata, "Fundus image classification for diabetic retinopathy using disease severity grading", In: *Proc. of the 9th International Conf. on Biomedical Engineering and Technology*, pp. 190-196, 2019.

[27] APTOS 2019 Blindness Detection. [Online] Av ailable: https://www.kaggle.com/c/aptos2019-bl indness-detection/

[28] P. Prasanna, P. Samiksha, K. Ravi, K. Manesh, D. Girish, S. Vivek, and M. Fabrice, "Indian diabetic retinopathy image dataset (IDRiD)", *IEEE Dataport*, 2018.

SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V10I4P103

Volume 10 Issue 4, 24-36, April 2023

Original Article

An Optimization Method for Tuning Hyper-Parameters of SGAN with Ensemble Classification Regression Model

Valarmathi Srinivasan[1], Vijayabhanu Rajagopal[2]

[1, 2]*Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Tamilnadu, India*

[1]*Corresponding Author : valarmathis313@gmail.com*

Received: 18 January 2023 Revised: 15 March 2023 Accepted: 03 April 2023 Published: 27 April 2023

Abstract *- Diabetic Retinopathy (DR) is a potential condition of Diabetes Mellitus (DM) that causes lesions on the retina, reducing vision and perhaps leading to disability if not properly treated. A Special Generative Adversarial Network with Ensemble Classification Regression (SGAN-ECR) model was introduced for the effective categorisation of various DR grades with the generation of high-contrast and low-saturated RF images. However, the Multi-Scale Attention Residual Network with Gradient Boosting (MSA-ResNet-GB) model used in SGAN-ECR requires manually assigning values for many hyper-parameters. An inappropriate value of the hyper-parameter leads to an increased error rate. This paper presents a method using the Enhanced Mine Blast Algorithm (EMBA) for selecting optimal hyper-parameters of MSA-ResNet-GB, such as the number of convolution layers, number of filters, filter size, number of Fully Connected (FC) layers and the hidden units in the FC layer to improve the architecture. Essential principles of EMBA are derived from the mine bomb explosion in real-time applications. The initial population of shrapnel fragments represents the initial hyper-parameter, and the computation of their subsequent locations represents the search for the best hyper-parameter. SGAN-ECR with parameter-optimised MSA-ResNet is named SGAN-OECR. In SGAN-OECR, optimised MSA-ResNet is trained by training images, and then the trained model is used to recognise the abnormalities in test images. MLP classifiers in the last FC layers of MSA-ResNet classify the severity of DR lesions. Finally, the experimental results of SGAN-OECR on Kaggle-APTOS and IDRiD datasets have 99.03% and 98.63% of accuracy, respectively, which is higher than the existing DR classification models.*

Keywords *- Diabetic retinopathy, Generative adversarial network, Hyper-parameters, Mine blast algorithm, Optimisation, Retinal fundus images.*

1. Introduction

Diabetes is one of the most prominent causes of DR, which greatly impairs eyesight. It fluctuates in severity and is medicated when detected early [1]. The retina transforms light into an electrical signal that can be altered to generate an image. The retina is surrounded by a network of blood arteries supplying oxygen and nutrients. Diabetes alters blood vessels, which reduces blood supply to the retina. This impairs the retina's health and reduces visual acuity. The most chronic kind of DR is non-proliferative retinopathy. At this stage, diabetes has minimal effects on eyesight, despite altering blood vessels. DR lesions include arteries that expand considerably (microaneurysms), leak fluid and proteins (exudates) or bleed blood (retinal haemorrhages) [2].

It is expensive to detect DR manually by ophthalmologists or professional graders. Various computeraided DR detection and diagnosis strategies have been developed in the past year. Computer-Aided Detection (CADe) methods identify and discriminate lesions at the pixel level, while Computer-Aided Diagnosis (CADx) models distinguish DR at the image level [3,5]. Evaluators can use the CADe and CADx systems to help them take patients to an ophthalmologist. Machine Learning (ML) and Deep Learning (DL) algorithms have been used to execute alternative strategies for CADx systems. Scale-invariant features extracted from DR lesion images are used to identify visual interest points. Probabilistic latent semantic analysis reduces feature dimensions. Support vector classifiers identify images as normal or pathological using visual terms [6].

ML does not dynamically produce features from the dataset that uses only the given features. The given features are either inaccurate representations of tumor forms in RF images or are not fitted to them [7]. Recently, DL has been used to estimate DR from RF pictures, and the results are

promising [8,9] and address the difficulties mentioned above. Convolutional Neural Network (CNN) is a popular DL model comprising millions of base learners and requires a high number of RF scans to train in the case of DR, resulting in generalisation error. Some advanced and pre-learned CNN architectures like VGG [11], ResNet [11] and Dual Path Network (DPN) [12] are used for DR lesion detection. These models efficiently transform low-level features into high-level features. The pre-learned model produced the best results for even fewer training samples. However, these models failed to properly categorise some pixels at the edges of tumours and healthy regions.

ResNetGB [13] is a pre-trained CNN structure constructed to train the Regions-of-Interest (ROI) in tumours from pixel-level tumor labelling. The layout was designed to train the tumor and normal neighbourhood patterns. A GB-based classifier was also included in determining the DR values of RF images. Multiple layers of features are required to explain the DR severity stage, despite the fact that the highest range of features are retrieved to identify and label retinal lesions as DR. To solve this, MSA-ResNetGB [15] model was developed to extract a mix of mid and high-level features to reinforce the interpretation. The Multi-Scale Feature Pyramid (MSFP) was also constructed to define the RF pattern in various positions GB classifier was used to classify DR grades. Although this model performs well for RF images, the accuracy is ineffective for low-contrast and saturated RF images. A Special Generative Adversarial Network with Ensemble Classification Regression (SGAN-ECR) model was developed to generate high-contrast and low-saturated RF images from low-contrast and saturated RF images for efficient training. In SGAN, pixel-2-pixel GAN and UNet++ are used in the generator and patch-GAN is used in the discriminator. MSA-ResNet was used to retrieve features. Multi-Layer Perceptron (MLP) algorithm instead of GB is used to categorise the different phases of DR severity. However, a classifier's accuracy mainly depends on assigning values for hyper-parameters used in MSA-ResNet.

Hence, an EMBA optimisation method is developed in this article to optimise the hyper-parameters of the MSA-ResNet structure in the SGAN-ECR model. The optimal hyper-parameters, such as the number of convolution layers, No. of filters, filter size, No. of FC layers, and the hidden units in the FC layer are selected by using the EMBA scheme. The initial population of shrapnel pieces of EMBA is analogous to the initial hyper-parameters. Their updated location computation is analogous to searching for the best hyperparameters. The updated locations' fitness values (classification error) are improved iteratively until all individuals have identical fitness values. Since SGAN-ECR is reconstituted with parameter-optimised MSA-ResNet is referred to as SGAN-OECR. The presented model improves the classification of DR severity from RF images.

This manuscript is arranged as Section II reviews the previous research associated with DR severity. Section III explains the SCAN-OECR algorithm for categorising the DR severity from RF images by optimising the hyper-parameter in MSA-ResNet using EMBA, and Section IV displays its efficacy. Section V summarises the whole study and suggests further development.

2. Literature Survey

A DL interpretable classifier was presented [16] to classify the DR grades. This model determined categorisation results by applying a score to every hidden and input data. Such scores represent the importance of each pixel to the objective categorisation. These results were obtained by utilising a new framework for the pixel-wise transmission of results across all neurons. However, efficiency was diminished due to an absence of DR images with subjectively designated lesions.

A new hybrid model was introduced [17] based on image processing and CNN layer to identify and classify DR from retinal fundus images. At first, pre-processing was performed using histogram equalisation and the CLAHE to enhance the input images for better classification. Then, the CNN was applied to classify those images to detect the DR. But, it needs additional databases and new pre-processing methods to increase the efficiency.

The decision tree-based ensemble learning (DTEL) method was introduced as a novel technique [18] for DR identification using gray-level intensity and texture data collected from fundus pictures. First, images containing noise were filtered out using pre-processing. The histogram and gray-level co-incident matrix of every image was computed to get the statistical features. In addition, the genetic technique was used to identify the most relevant features, which were then taught by the XGBoost algorithm to estimate the intensity of DR. However, the model parameters must be fine-tuned to improve efficiency.

A new hybrid method was presented [20] for prior DR identification and categorisation. First, the collected images were pre-analysed using the adaptive histogram equalisation, contrast stretching and median filtering schemes. After, enhanced images were given to the bright and red lesion recognition schemes to extract the features on the recognised areas. Moreover, those features were classified by the SVM, KNN and binary trees along with the combined voting scheme. On the other hand, the pre-processing stage was time-consuming because of engaging all image-processing phases.

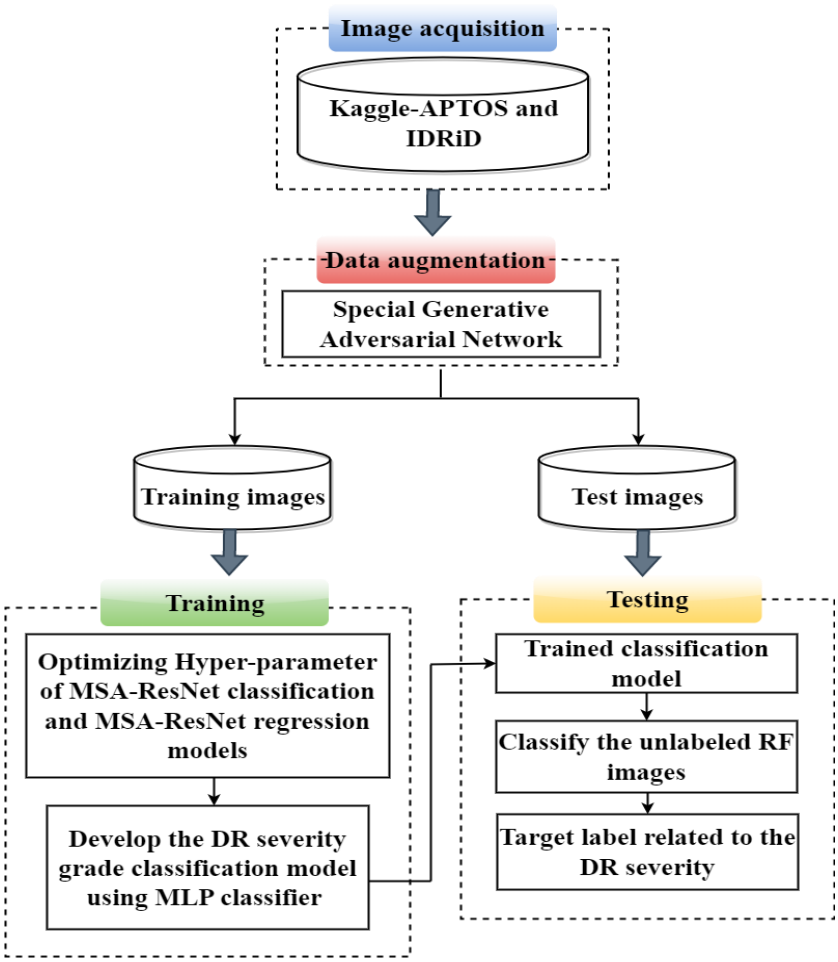


Fig. 1 Flow diagram of the SGAN-OECR model for DR disease classification

A VGG-NIN method was developed [21] to classify the different DR stages with the minimum feasible trainable variables to accelerate learning and model convergence. This model integrated the VGG16, spatial pyramid pooling layer, and network-in-network to produce a very indiscriminate, persistent structure for deep identification of DR images. However, the accuracy was less since the pre-processing steps were not effective.

An automated DR grading model was developed [22] depending on identifying multiple retinal lesions. Initially, the retinal images were pre-processed, and the grey-level run-length matrix average was determined to distinguish normal and DR images. Then, the U-Net was used to extract the exudates, microaneurysms, haemorrhages and blood vessels. Also, different characteristics were retrieved and classified by the multi-mode SVM to classify the DR grades. But, it needs to consider the different disease features concurrently to increase the accuracy.

Deep Transfer Learning (DTL) was introduced [23] based on the different pre-trained CNN models such as "AlexNet, GoogleNet, InceptionV4, Inception ResNetV2 and ResNext50" to detect Retinopathy automatically. First, the DR images were acquired and labelled with a suitable diagnosis method. Then, the different CNN models were trained to classify the severity of DR images. However, a more in-depth analysis was required, which needed a massive amount of patient information.

A Model was developed [25] for recovering retinal images using a cycle-consistent generative adversarial network (CycleGAN) with a convolution block attention component. Using a modified UNet, the retinal vessels in the regenerated retinal images were segmented. This model has two generators and two discriminators. Generators convert low-quality images from one classification to another and vice versa. Discriminators classify real and manufactured images. In order to get segmented images, the model for retinal vascular segmentation incorporates downsampling, bottlenecking, and up-sampling layers. CBAM was utilised to enhance the feature extraction of these models. This model demands a considerable amount of data training time.

A DL protocol Funswim was created [26] to determine the rate of DR and the risk of retinal edema based on FIR. The

swim transformer works remarkably well in this technique for categorising medical images. The ImageNet-based TL model was applied as a training model to collect sufficient low-level features for learning. Therefore, when the model was combined with high-dimensional features, the systems could be analysed more accurately (particularly potential disease classification bases) and improve classification efficiency. However, this model generated a slower convergence rate.

A Deep CNN (DCNN) model for autonomous identification and categorising DR from color FIR was described [27]. Each image was standardised using Min-Max standardisation to minimise the CNN from training the image's underlying background noise. The CNN was then utilised to categorise and identify the DR. Finally, dynamically, the quadratic kappa metric was employed to assess the accuracy of the predictions. However, sample sizes for all cohorts were rather small.

A Hybrid DL (HDL) framework was devised [28] to recognise and categorise DR in the FIR of the eye. CNN models were subjected to TL to obtain features from ResNet-18 and GoogleNet models, which were then connected to generate a composite feature vector. Using these attributes as input, multiple classifiers performed binary and multiclass categorisation of DR images from the acquired dataset. Integrating GoogleNet and ResNet-18 features increases the aggregate multiclass recognition efficiency of the system. However, this system has a slower processing speed.

3. Proposed Methodology

In this part, the SGAN-OECR model is described briefly. Figure 1 depicts the flowchart of the provided hyper-parameter maximising system for classifying the various degrees of DR severity from RF images. Primarily, different RF image databases are acquired and then the SGAN-ECR is applied to get the number of high-quality RF images. Such images are further trained by the classification and regression models using MSA-ResNet structure [15] with MLP classifier. In addition, the generated model is deployed to the test samples to categorise the various severity levels of DR. The following sections describe these different processes.

3.1. MSA-ResNet Structure Optimization

This framework intends to propose an efficient method for autonomously optimising the hyper-parameters and developing an MSA-ResNet for DR disease classification. The efficiency of the MSA-ResNet architecture is extremely dependent on the hyper-parameter values. Since optimising all MSA-ResNet hyper-parameters would require considerable time. The number of convolution layers, the number of filters, the filter size, the number of FC layers, and the hidden units in the FC layer are thereby optimised thoroughly in this suggested framework**.**

Table 1. Configuration of hyper-parameters for MSA-ResNet structure

Model	Hyper-parameter	Value
MSA-ResNet	Conv_1 layer	7*7
	Conv_2 layer	3*3
	Batch Size	4
	Epochs	100
	Learning Rate	10^{-4}
	Optimizer	Adam Optimizer
	Activation Function	Sigmoid
	Pooling	Adaptive and Global Mean Pooling
	Residual Blocks	3, 8, 12, and 3

The optimal MSA-ResNet structure is expressed as the following list of hyper-parameters:

$$H = H_{conv}, H_{fc} \quad (1)$$

From above Equation (1), H_{conv}stands for the parameters of the convolutional layer and H_{fc}stands for the parameters of the FC layer. The convolutional layer's parameters are defined as follows:

$$H_{conv} = \{c_0, \dots, c_{N-1}\} \quad (2)$$

Where c_N represents the number of convolutional layers in Equation (2). Every Convolutional layer in the a^{th}layer consists of two tuples,

$$c_a = (d_{count}, d_{size}) \quad (3)$$

Where, the d_{count} and d_{size} are the No. of filters and size of the filters of the a^{th} layer, respectively, in equation (3). Equation (4) formulates the set of the fully-connected layers,

$$H_{FC} = t_0, t_{k-1} \quad (4)$$

Where, t represents the hidden division number in the FC layer of the a^{th} layer, and k is the No. of FC layers.

The possible set of MSA-ResNet configurations is denoted by $\mathcal{Q}$. The objective is to identify a configuration $q \in \mathcal{Q}$ which provides a minimum classification error rate. In this proposed model, the Classification error rate (Cer) is considered the objective function to enhance the classification accuracy. EMBA selects the possible hyper-parameter values in this research work. The selection of optimal parameters reduces the utilisation of computer resources. Table 1 depicts the Hyper-parameter and its range of MSA-ResNet.

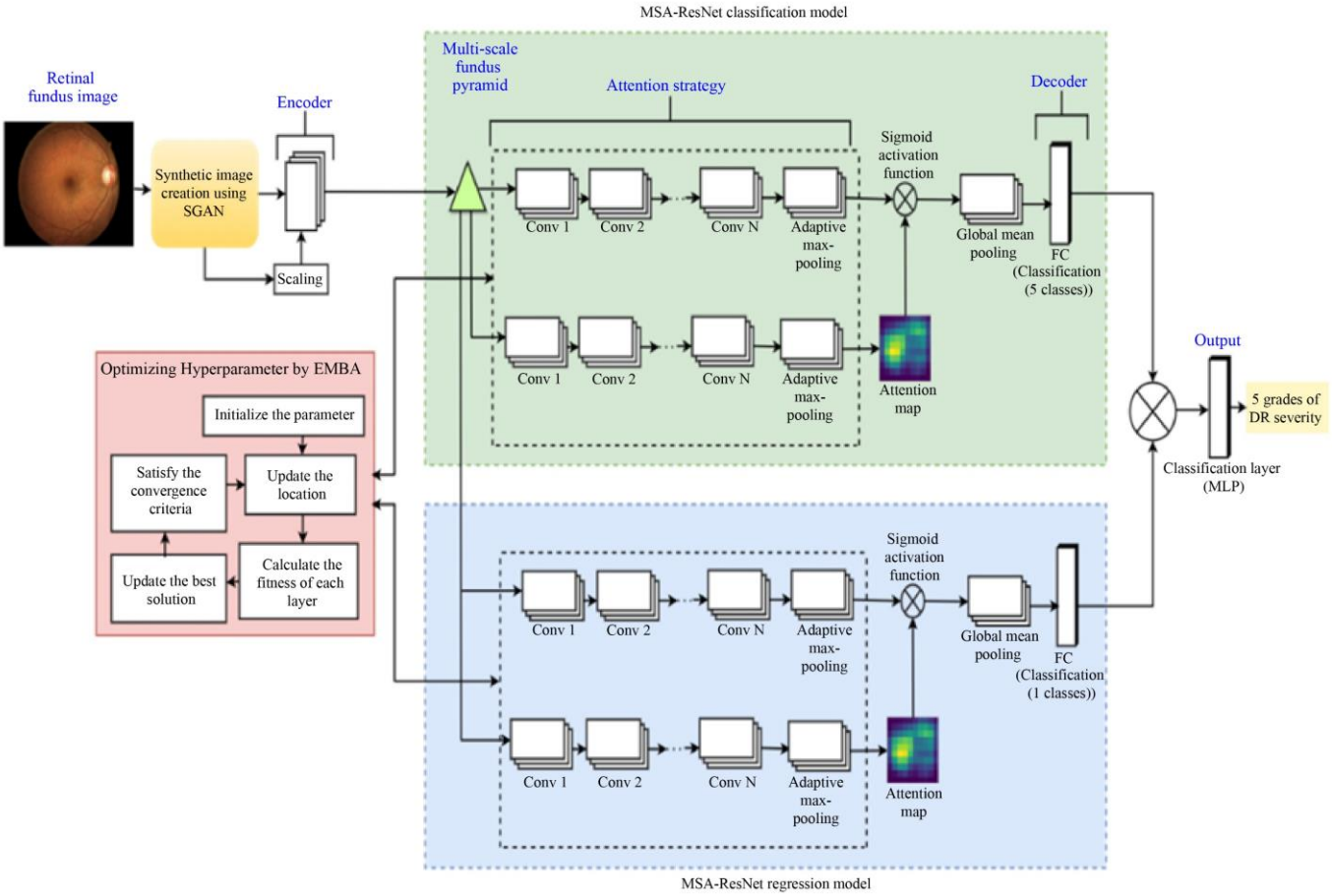


Fig. 2 Schematic representation of SGAN-OECR model

3.2. EMBA for Optimising Hyper-Parameters in MSA-ResNet Structure

The structure of the SGAN-OECR model is provided in Figure 2. This is the optimised version of SGAN-ECR. SGAN-OECR comprises of SGAN and ECR. SGAN has GAN with UNet++ as a generator and patch-GAN as a discriminator for generating quality images to get a highly relevant training model from the classifier. The ECR part has MSA-ResNet regression and MSA-ResNet classification structures for feature extraction, training and classifying for DR severity grades. The EMBA algorithm in the proposed work optimises the MSA-ResNet model's hyperparameters. An appropriate selection of parameters provided high accuracy in less number of iterations.

MBA is one of the most recently developed metaheuristic algorithms, and EMBA is an upgraded version of MBA. This algorithm has given superior outcomes when compared to other metaheuristic optimisation techniques. The explosion of landmines influenced this MBA method. In summary, the MBA claims that when a landmine explodes, the resulting shrapnel disperses and collides with other landmines, destroying them. In order to achieve the optimal solution, it is essential to identify the initial blasting point at which sufficient collisions are generated to remove the region of mines completely. Consequently, when an explosion occurs at an initial location(X_0). Shrapnel fragments are produced and disseminated over the (search) zone for the first time. As a result, more explosions occur when these shrapnel particles interface with surrounding landmines. The position of every explosion differs from the position of the preceding explosion, and the number of casualties is generated and represented by $f(X)$ for each point (X). This method is a population-based technique, which necessitates the generation of a starting population. This population (which is the ideal alternative at the time) is generated by the initial blast (first shot), and its size is determined by the number of shrapnel fragments generated by the original shot. The two parts of this MBA procedure are exploration (global search) and exploitation. The proliferation of shrapnel prompted the examination stage to numerous additional regions, whereas the exploitation stage comprises exploiting the existent resources in these designated sites. Figure 3 displays a diagrammatic description of the MBA in terms of exploration (color lines) and exploitation (black color lines) processes, where $x_0, \ldots, x_5$ represent the beginning positions and the subsequent positions.

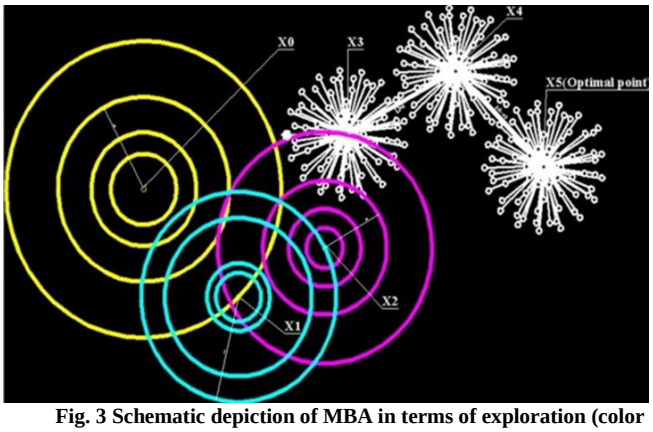


Fig. 3 Schematic depiction of MBA in terms of exploration (color lines) and exploitation (white color lines) operations

A population generation is here represented as the analogy of generating the set of MSA-ResNet hyper-parameters configurations Q, and the solutions or fitness of each population is represented as,

$$f(Q) = f(X) = \frac{1}{1+|Cer|} \quad (5)$$

In Equation (5), $f(Q)$ represent the fitness or objective value for the configuration set Q. $f(X)$ $represents$ the fitness of exploitation at shot point X. Here X is the analogy for the configuration set Q. So, in the following sections, X represent the configuration set for hyperparameters of MSA-ResNet. Cer is the classification error of MSA-ResNet for a selected configuration parameter.

3.2.1. Steps Involved in MBA

The initial shot point (hyperparameter set Qs for MSA-ResNet) determines the response space to be searched, and this initial shot point is chosen arbitrarily. This EMBA uses the lower limit parameters of a given problem as though the location of the original shot point is irrelevant and performs a fresh first shot. EMBA employs a changing randomly completed value using Equation (6),

$$X_0^{current} = \ell b + rand * (ub - \ell b) \quad (6)$$

Suppose, $X_0^{current}$is the point of the initial explosion point,$\ell b$and ub depicts lower bound upper bound search space, respectively. This means the minimum and maximum range of values for the hyperparameter set Qs for MSA-ResNet. The next explosion location (the updated configuration sets for MSA-ResNet) will occur later, which will be represented as,

$$X = X_T, T = 1, 2, 3, .., n_d \quad (7)$$

Where,n_d is a search space for positions X in Equation (7). Assume further locations of the detonated landmine, as identified by the fragments (shrapnel), using Equation (8),

$$X_{n+1}^b = X_{z(n+1)}^b + \exp\left(-\sqrt{\frac{T_{n+1}^b}{d_{n+1}^b}}\right) * X_n^b \quad (8)$$

$$n = 0,1,2, \ldots d$$

$X_{z(n+1)}^b$represents the location of the exploding mine bomb, whereas T (n+1)b and d (n+1)b are the orientation and range of the shrapnel pieces accomplished in every iteration. The subscript b indicates the total amount of initial shot points.

Determine the exploding landmine location by Equation (9) as flows,

$$X_{z(n+1)}^b = d_{n+1}^b \times r \times \cos(\theta) \quad (9)$$

Where, r is the random integer from 0 to 1 and is the shrapnel angle. This angle θ corresponds to $360/Nb$.

The distance and the directions of shrapnel pieces are computed by equations (10) and (11), and the position of the exploded landmine is calculated by utilising equation (9).

$$d_{n+1}^b = \sqrt{(X_{n+1}^b - X_n^b)^2 + (F_{n+1}^b - F_n^b)^2} \quad (10)$$

$$T_{n+1}^b = \frac{F_{n+1}^b - F_n^b}{X_{n+1}^b - X_n^b} \quad (11)$$

F is the value of the fitness function at the location of X. In the exploitation phase, the EMBA uses a fitness value of the current solution. In order to attain a balance between these two random operations, a fitness function $(F)$from Eq. (5) is used in Eqs. (10) and (11). The fitness function represents the classification error of MSA-ResNet for the parameters selected at location X.

Two techniques are employed to enhance the exploration and exploitation of EMBA.

The Exploration Process:

The user specifies an initial distance, which is utilised to search for the optimal solution within a range determined by the ratio of the original length and an arbitrarily determined number. In addition, exploration is incorporated to perform exploration of the layout space at lesser and greater distances. This factor is utilised in the early iterations of the method, compared to an iteration number. If it is more than the iterative number, the resulting space is explored as follows:

$$d_{n+1}^b = X_n^b * r^2 \quad (12)$$

$$X_{z(n+1)}^b = d_{n+1}^b * \cos(\theta) \quad (13)$$

The Exploitation Process:

This procedure is determined by the exploration factor (γ) if it is lower than the number of iterations(b). By minimising the constant value, the distance traversed by fragments in an explosion is decreased in this instance. Candidates will obtain the fixed value. This elimination in the distance is calculated by Eq. (13) as follows:

$$X_n^b = \frac{\gamma_{n-1}^b}{\exp\left(\frac{b}{\varepsilon}\right)} \tag{14}$$

Convergence Criteria:

In the majority of metaheuristic algorithms, the optimal result is obtained when the terminating criteria are considered to be the greatest iterative values, the query distance, or a smaller value set as the permitted deviation among the previous two outcomes. This is the primary distinction between EMBA and MBA. The explosion positions are updated until the below convergence criteria are satisfied.

The difference in the length among the current exploded point is $X_{z(n+1)}^b$ and the greatest solution will be X_{best}. The enhancement is considered as follows

$$X_{n+1}^b = X_{z(n+1)}^b + exp\left(-\sqrt{\frac{1}{ED}}\right) \times r \otimes \left\{X_{best} - X_{z(n+1)}^b\right\}, \quad j = 1,2,\ldots n \tag{15}$$

$$ED = \left[\sum_{i=1}^{T}\left(X_{best} - X_{z(n+1)}^b\right)^2\right]^{1/2} \tag{16}$$

Where ED indicates the Euclidean distance between the optimal solution location X best and the current explosion point $X_{z(n+1)}^b$in T dimensions. The current method did not utilise preceding optimal location information, hence accelerating the convergence of the algorithm. The general pseudo code of SGAN-OECR is given below.

3.3. Pseudo Code for SGAN-OECR Model

Input : DR lesions and healthy images for 5 different DR stages.
//Images from Kaggle-APTOS and IDRiD for 5 different DR grades; Get the RF images Perform SGAN model to create the synthetic RF images and add them to the real RF image databases; Divide the newly obtained RF image databases into the training set and test set;

Output : Different grades of DR severities

Begin

Step 1 : Choose the preliminary parameters: N_d, γ, $\varepsilon, \mathcal{ub}, \mathcal{lb}$ and maximum number of iterations p_{max} for EMBA.

Step 2 : $while(p < P_{max})$

Step 3 : $P \leftarrow p + 1$;

Step 4 : Initialize the initial shot points using Equation (6) //this may cause the algorithm to explore alternative regions in the resulting space. For an analogy, this point has hyperparameter configuration set Q for MSA-ResNet, i.e. No. of convolution layers, No. of filters, filter size, No. of. FC layers, hidden units in the FC layer.

Step 5 : Generate other shrapnel pieces using Equation (7), Equation (8) and Equation (9) and update their locations.

Step 6 : Compute the distance and directions of shrapnel pieces direction by equations (10) and (11).

Step 7 : Calculate the fitness of all location and direction updated shrapnel pieces using MSA-ResNet classification and regression, saving the highest fitness as the best temporal shrapnel piece.

Step 8 : Improve the shrapnel pieces' position by exploration-exploitation from Equation (12) to Equation (14)

Step 9 : Examine the convergence criteria. If the stopping requirement is met, the algorithm will be terminated. Otherwise, proceed to step 5.

Step 10 : If true, save the improved shrapnel piece as the best temporal response

Step 11 : If not, replace the shrapnel's location with the optimal temporal response

Step 12 : Examine the convergence criteria. If the stopping rule is satisfied, the algorithm will be terminated. Else, proceed to step 5.

Step 13 : End while

Step 14 : Select highest fitness shrapnel pieces and obtain the trained model from the MSA-ResNet classification model;

Step 15 : Classify the test images into different grades of DR severity using the trained model;

End

Table 2 depicts the MBA parameters and their value for performing MSA-ResNet and GAN.

Table 2. MBA parameter settings

Parameter	Value
Upper bound	1
Lower bound	0
Maximum distance	2
Number of iterations	200
Population size	20
KNN Classifier	5

4. Experimental Results

The SGAN-OECR model is assessed using four different metrics like Accuracy, Precision, Recall, and F1-Score using benchmark datasets, Kaggle-APTOS and IDRiD.

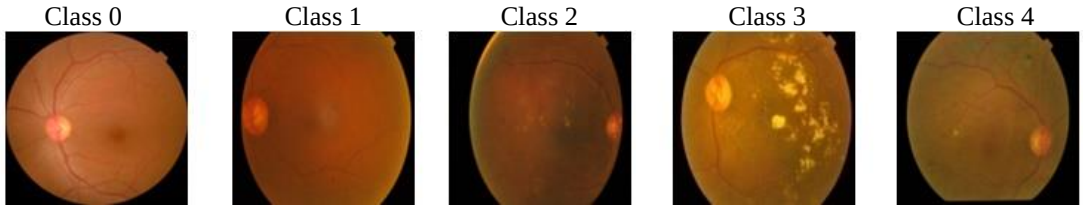


Fig. 4 Sample RF images taken from the Kaggle-APTOS corpus

Fig. 5 Augmented RF images taken from the Kaggle-APTOS corpus

Fig. 6 Sample RF images taken from the IDRiD

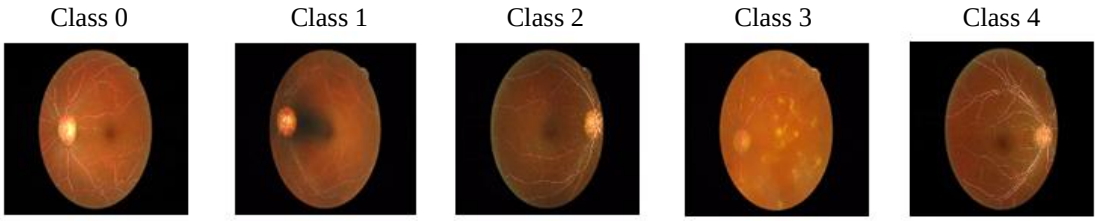


Fig. 7 Augmented RF images taken from the IDRiD

4.1. Dataset Description

The proposed model was assessed using benchmark datasets, Kaggle-APTOS and IDRiD. A number of samples that were gathered, the samples that were annotated, the DR grades, and the ground truths are all detailed in this subsection. From both datasets, sample RF pictures are displayed.

4.1.1. APTOS 2019 Blindness Detection Database

This study used RF images from the "Asia Pacific Tele-Ophthalmology Society (APTOS) 2019 Blindness Detection dataset" [29]. This Kaggle image collection has 3662 images obtained from various rural Indians. "Aravind Eye Hospital" in India gathered and organised the data. However, a group of medical specialists used the International Clinical Diabetic Retinopathy Disease Severity Scale to examine and classify the obtained samples.

APTOS dataset images are divided into five classes on a scale of 0 to 4: "no DR, mild DR, moderate DR, severe DR, and proliferative DR". The initial class includes healthy RF samples that do not include DR. The preceding classes have more injured retinas than the previous ones. The final categorisation was proliferative DR, which includes instances with vitreous or pre-retinal HEs. Figure 4 illustrates the RF samples from every Kaggle-APTOS class, whereas Figure 5 depicts the enhanced RF samples from each Kaggle-APTOS class.

4.1.2. Indian Diabetic Retinopathy Image Database (IDRiD)

Images from the IEEE ISBI-2018 are used in the "Indian Diabetic Retinopathy Image Database (IDRiD)" [30]. It includes 516 images with medical indications for DR and DME, containing 413 training images and 103 test images under the heading "Disease grading" as appropriate. IDRiD was the first dataset to represent the Indian samples. Figure 6 shows the RF images from every category in the IDRiD, whereas Figure 7 shows the enhanced RF samples from every class in the IDRiD.

Table 3. Performance analysis of the proposed model and the cutting-edge models for DR classification on "APTOS 2019 blindness detection dataset"

Ref.	Year	Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
[18]	2021	DT-EL	86.81	87.31	87.74	87.51
[13]	2021	ResNetGB	88.40	88.72	88.40	88.47
[21]	2021	VGG-NIN	87.65	92.86	92.86	92.86
[22]	2021	U-Net	87.37	91.76	93.98	92.86
[26]	2022	Funswim	88.89	92.59	94.94	93.75
[27]	2022	DCNN	89.47	94.12	94.12	94.12
[15]	2022	MSA-ResNetGB model	94.40	94.52	94.40	94.42
-	-	SGAN-ECR	97.81	95.95	97.98	96.90
Proposed Model	2023	SGAN-OECR (Conventional MBA)	97.92	96.97	98.12	97.85
		SGAN-OECR (Enhanced MBA)	98.63	98.83	98.28	98.55

Table 4. Performance analysis of the proposed model and the existing models for DR classification on the "IDRiD Database"

Ref.	Year	Classification method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
[17]	2020	CNN model	86.02	89.67	90.24	91.93
[20]	2021	Mixed Model	87.50	91.76	93.38	92.86
[13]	2021	ResNetGB	88.89	93.51	93.51	93.51
[23]	2022	DTL model	89.01	93.83	93.83	93.83
[25]	2022	CBAM-UNet	89.47	94.94	92.59	93.75
[28]	2022	HDL Model	90.11	93.83	95.00	94.41
[15]	2022	MSA-ResNetGB model	94.17	91.48	91.57	91.45
-	-	SGAN-ECR	96.12	96.31	93.73	94.77
Proposed Model	2023	SGAN-OECR (Conventional MBA)	98.61	98.26	95.92	96.89
		SGAN-OECR (Improved MBA)	99.03	99.39	96.00	97.47

The prevalence of DR and DME is reported at the pixel level for every instance in the IDRiD collection. Based on the intensity scale, the DR grade was categorised into five groups from 0 to 4, equivalent to the Kaggle-APTOS dataset.

4.2. Performance Metrics

Using the following metrics, the efficiency of the SGAN-OECR model is compared to that of cutting-edge models: "Accuracy, Precision, Recall and F1-Score" in equations 17 – 20, respectively. The quantitative formulas used to calculate the metrics are as follows:

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \quad (17)$$

$$Precision = \frac{TP}{TP+FP} \quad (18)$$

$$Recall = \frac{TP}{TP+FN} \quad (19)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (20)$$

From the above equations (17-20), True Positive (TP) appropriately identifies the proportion of positive samples. In contrast, True Negative (TN) correctly categorises the number of negative samples. False Positive (FP) refers to the proportion of negative class samples that are identified as positive class samples, whereas False Negative (FN) refers to the proportion of positive class samples that are classed as negative class samples.

4.3. Performance Evaluation

This section validates the suggested SGAN-OECR model using the "Accuracy, Precision, Recall, and F1-score" from the available benchmark databases. The SGAN-OECR model was trained on two distinct databases for classifying RF pictures into different degrees of DR severity. Both databases use multiclass classification since the database has five classes. In addition, the conclusions of the APTOS and IDRiD tests are described in the parts that follow, including the performance comparison table and confusion matrices. APTOS and IDRiD class-wise performance evaluations are offered. Table 3 displays the comparative performance analysis of the presented model and existing methods for DR categorisation using the Kaggle-APTOS 2019 Blindness Detection dataset.

Table 3 combines and evaluates the modern framework of the Kaggle-APTOS database. According to the research, CNN is the most extensively used deep learning model for clinical image assessment. Based on the "Accuracy, Precision, Recall, and F1-Score", Table 3 evaluates the efficacy of the SGAN-OECR model on the Kaggle-APTOS database using four categorisation models from current research. According to Table 3, the suggested model surpasses cutting-edge Kaggle-APTOS models.

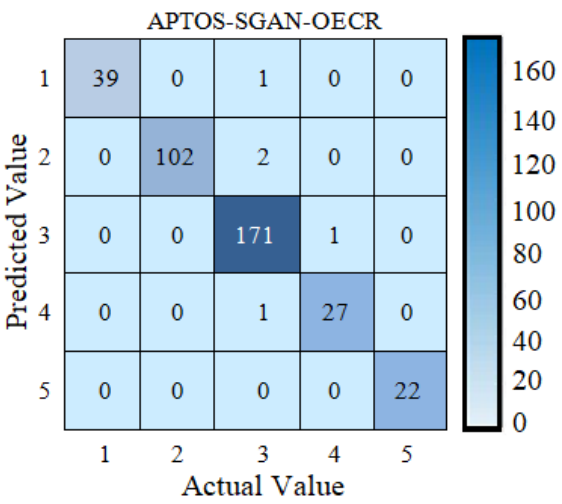


Fig. 8 Confusion matrix for SGAN-OECR model on the Kaggle-APTOS 2019 blindness detection database

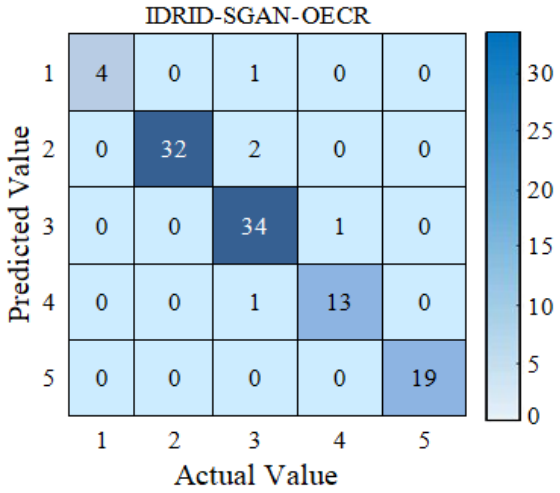


Fig. 9 Confusion matrix for SGAN-OECR model on the IDRiD

Existing models like ResNetGB [13], DT-EL [18], VGG-NIN [21], U-Net [22], Funswim [26], DCNN [27], MSA-ResNetGB [15], and SGAN-ECR achieved an accuracy of 88.40%, 86.81%, 87.65%, 87.37%, 88.89%, 89.47%, 94.40%, 97.81%. Thus, it is noticed that these models' performance is better but not superior to that of the suggested model. This validates the reliability of the suggested model.

Figure 8 depicts the confusion matrix for the suggested SGAN-OECR algorithm on the APTOS database. It provides the prevalence of images by category and the fraction of accurately and incorrectly classified images. In class 0, 39 images were accurately diagnosed as having no DR, 102 images were accurately categorised in class 1 as having mild DR, 171 images were accurately categorised in class 2 as having considerable DR, 27 images were accurately defined in class 3 as having extreme DR, and 22 images were accurately classed in class 4 as having PDR.

While using the IDRiD dataset, the suggested model is trained and evaluated to generate DR categorisation results. The IDRiD consists of 516 photos, of which 20% of annotated images are evaluated for testing. The multiclass model is validated by comparing it to contemporary creative works. The efficiency study of the proposed and current algorithms for DR categorisation on the IDRiD is presented in Table 4.

In Table 4, relevant categorisation methods such as ResNetGB [13], CBAM-UNet [25], MSA-ResNetGB [15], and SGAN-ECR were used to assess the results of the SGAN-OECR model on the IDRiD using the evaluation metrics like "Accuracy, Precision, Recall, and F1-Score".

The existing models such as the CNN model [14], Mixed Model [20], ResNetGB [13], DTL model [23], CBAM-UNet [25], HDL Model [28], MSA-ResNetGB [15], and SGAN-ECR [16] obtained 86.02%, 87.50%, 88.89%, 89.01%, 89.47%, 90.11%, 94.17%, 96.12%, 99.03% accuracy, respectively, which is 15.12%, 13.17%, 11.41%, 11.25%, 10.69%, 9.89%, 5.17% and 3.03% less than the proposed model.

Figure 9 depicts the confusion matrix for the SGAN-OECR model on the IDRiD. Figure 9 depicts the distribution of photos by class and the fraction of correctly and incorrectly classified photographs. In class 0, 4 photos were appropriately recognised as having no DR; in class 1, 32 images were precisely categorised as having mild DR; in class 2, 34 images were accurately classified as having considerable DR; in class 3, 13 images were accurately defined as having extreme DR; and in class 4, 19 images were accurately categorised as having PDR.

The proper parameter selection for the classifier changes the model's architecture towards less error rate, as depicted in Figure 10. Figure 10 provides the average error between the restored and original images. The lower the Mean Square Error (MSE) value, the less degraded the restored image. This precaution prevents overfitting and, as a result, improves training effectiveness by drastically reducing the number of parameters generated during the training process. An optimal variable selection accelerates model training and convergence.

On the basis of attempting specific hyper-parameters, the produced model is used to estimate the value of all hyper-parameters space. Then, a specific number of high-quality hyper-parameters are picked to train the model repeatedly.

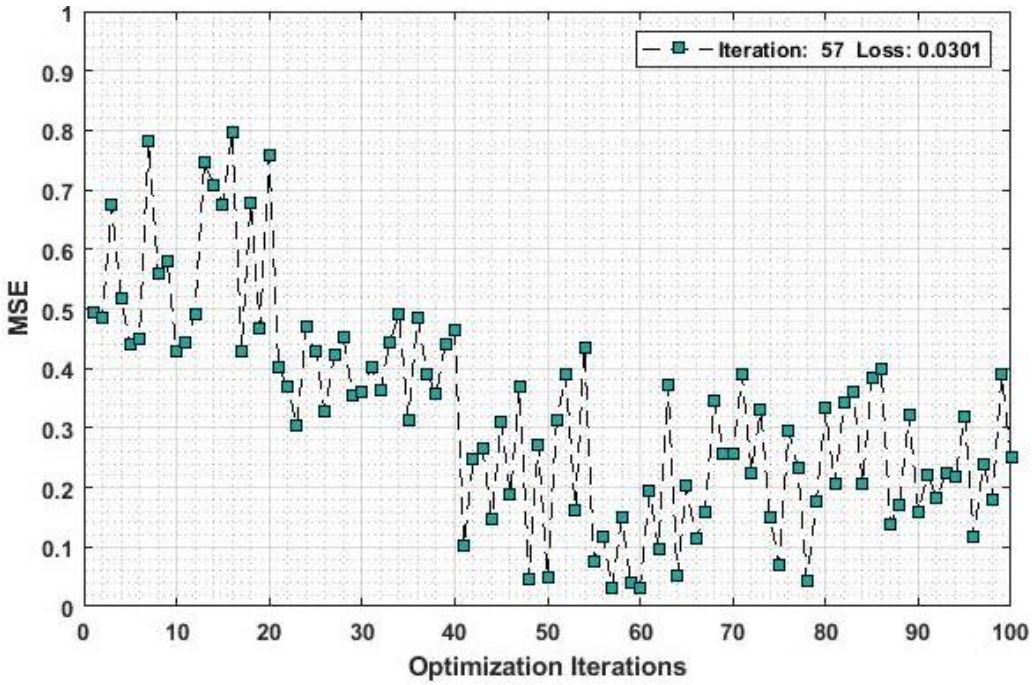


Fig. 10 MSE vs optimisation iterations

Table 5. Configuration of hyper-parameters for MSA-ResNet and GAN structure

Model	Hyper-parameter	Range	Optimal Value
MSA-ResNet	Batch Size	[5, 10, 20]	20
	Learning Rate	[0.001, 0.01, 0.1, 0.2]	0.01
	Filter Size	[1-10]	5
	Number of Feature Map	[32-256]	128
	Number of Filters	50-500	400
	Drop out	0.1~0.2	0.2
	Channel Size	[18-30]	24
	Optimizer	[Adam, SGD, Adagrad, Adadelta]	Adam
	Activation Function	[“ReLU”, “tanh”, “Linear”, “Sigmoid”]	Sigmoid
	Pooling	[Adaptive, Global]	Adaptive
	Residual Blocks	[3, 8, 12, 3]	8
	Scaling	3-5	5
GAN	Generator (Gen_layersize)	[16-40]	32
	Discriminator [Disc_layersize]	[6-20]	18
	Batch Size	[32-128]	64
	Gen_ learning rate	[1-50]	24
	Disc_learning rate	[1-50]	24

Table 6. Performance of SGAN-OECR model on the "Kaggle-APTOS and IDRiD database"

Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
APTOS	99.03	99.39	96.00	97.47
IDRiD	98.63	98.83	98.28	98.55

The amount of training data diminishes steadily with each iteration, but the model's accuracy increases fast, and eventually, the suitable hyper-parameter is acquired. Table 5 depicts the Hyper-parameter and its range of MSA-ResNet and GAN. The suggested SGAN-OECR model is discussed, and its experiential values on the "Kaggle-APTOS and IDRiD datasets" are evaluated. Using the classical models discussed above, the performance of the models on the two databases was evaluated separately. Both reference databases are proposed models that provide better results than the models in the literature. Table 6 shows results from tests on the "Kaggle-APTOS and IDRiD databases" using the SGAN-OECR

model. On the Kaggle-APTOS database, the SGAN-OECR model obtained 99.03% accuracy, 99.39% precision, 96.00% recall, and 97.47% F1-Score, outperforming the IDRiD database, which obtained 98.63% accuracy, 98.83% precision, 98.28% recall, and 98.55% F1-Score.

5. Conclusion

This article presented an automated method called SGAN-OECR model to enhance the efficiency of categorising the DR severity grades. At first, the binary encoding of MSA-ResNet architectures was computed. Each layer in the CNN module has its own set of hyper-parameters encoded, with all layers having the same constant length dependent on layer count and input image size. Then, the metaheuristic algorithm (EMBA) was used to select the best hyper-parameters and improves the MSA-ResNet structure. Finally, the fitness function was calculated to improve the model's detection of DR lesion abnormalities in FIR. This model efficiently resolves the problem of designing an automated hyper-parameter of MSA-ResNet architectures for optimising DR lesion grading. At last, the experimental findings of SGAN-OECR on Kaggle-APTOS and IDRiD proved that it has 99.03% and 98.63% accuracy, correspondingly, compared to the existing DR grade classification models.

References

[1] Dóra Júlia Eszes et al., "Diabetic Retinopathy Screening in Patients with Diabetes Using a Handheld Fundus Camera: The Experience from the South-Eastern Region in Hungary," *Journal of Diabetes Research*, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] Solomon and Goldberg, "ETDRS Grading of Diabetic Retinopathy: Still The Gold Standard?," *Ophthalmic Research*, vol. 62, no. 4, pp. 190-195, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[3] Wei Zhou, Chengdong Wu, and Xiaosheng Yu, "Computer-Aided Diagnosis for Diabetic Retinopathy Based on Fundus Image," *2018 37th Chinese Control Conference (CCC), IEEE*, pp. 9214-9219, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[4] Meteb Altaf et al., "Monitoring Diabetic Foot by Designing a New Smart Sole," *SSRG International Journal of Pharmacy and Biomedical Engineering*, vol. 7, no. 3, pp. 15-20, 2020. [CrossRef] [Publisher Link]

[5] Nabila Eladawi et al., "Computer-Aided Diagnosis System Based on a Comprehensive Local Features Analysis for Early Diabetic Retinopathy Detection using OCTA," *Diabetes and Fundus OCT*, pp. 1-23, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] K.S. Sreejini, and V.K. Govindan, "Retrieval of Pathological Retina Images Using Bag of Visual Words and Plsa Model," *Engineering Science and Technology, An International Journal*, vol. 22, no. 3, pp. 777-785, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[7] Gwenolé Quellec et al., "Automatic Detection of Referral Patients due to Retinal Pathologies through Data Mining," *Medical Image Analysis*, 29, pp 47-64, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[8] Ping-Nan Chen et al., "General Deep Learning Model for Detecting Diabetic Retinopathy," *BMC Bioinformatics*, vol. 22, no. 5, pp. 1-15, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Zhentao Gao et al., "Diagnosis of Diabetic Retinopathy Using Deep Neural Networks," *IEEE Access*, vol. 7, pp. 3360-3370, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[10] Ms.S.Deepa, and Mr.S.Vijayprasath, "Certain Investigation of the Retinal Hemorrhage Detection in Fundus Images," *SSRG International Journal of Electronics and Communication Engineering*, vol. 2, no. 2, pp. 24-34, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[11] Mustapha Aatila et al., "Diabetic Retinopathy Classification Using Resnet50 and VGG-16 Pretrained Networks," *International Journal of Computer Engineering and Data Science*, vol. 1, no. 1, pp. 1-7, 2021. [Google Scholar] [Publisher Link]

[12] Yunpeng Chen et al., "Dual Path Networks," *Advances in Neural Information Processing Systems*, vol. 30, 2017. [Google Scholar]

[13] Fahman Saeed, Muhammad Hussain, and Hatim A. Aboalsamh, "Automatic Diabetic Retinopathy Diagnosis using Adaptive Fine-Tuned Convolutional Neural Network," *IEEE Access*, vol. 9, pp. 41344-41359, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] S.Bibiana Vincy, and Dr. C. Chitra, "Detection of Exudates in Color Fundus Image," *SSRG International Journal of Electronics and Communication Engineering*, vol. 2, no. 11, pp. 1-10, 2015. [CrossRef] [Publisher Link]

[15] Valarmathi Srinivasan, and Vijayabhanu Rajagopal, "Multi-Scale Attention-Based Mechanism in Gradient Boosting Convolutional Neural Network for Diabetic Retinopathy Grade Classification," *International Journal of Intelligent Engineering and System*, vol. 15, no. 4, pp. 489-498, 2022. [CrossRef] [Google Scholar]

[16] Jordi de la Torre, Aida Valls, and Domenec Puig, "A Deep Learning Interpretable Classifier for Diabetic Retinopathy Disease Grading," *Neurocomputing*, vol. 396, pp. 465-476, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] D. Jude Hemanth, Omer Deperlioglu, and Utku Kose, "An Enhanced Diabetic Retinopathy Detection and Classification Approach Using Deep Convolutional Neural Network," *Neural Computing and Applications*, vol. 32, no. 3, pp. 707-721, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Niloy Sikder et al., "Severity Classification of Diabetic Retinopathy Using an Ensemble Learning Algorithm through Analysing Retinal Images," *Symmetry*, vol. 13, no. 4, pp. 1-26, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Mohammed Al-Mijalli, "Hemodynamic in Human Diabetic Abdominal Aneurysmal Aorta Using Computational Fluid Dynamics," *SSRG International Journal of Pharmacy and Biomedical Engineering*, vol. 8, no. 2, pp. 1-5, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] Anas Bilal et al., "Diabetic Retinopathy Detection and Classification Using Mixed Models for a Disease Grading Database," *IEEE Access*, vol. 9, pp. 23544-23553, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[21] Zubair Khan et al., "Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture," *IEEE Access*, vol. 9, pp. 61408-61416, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Eman Abdelmaksoud et al., "Automatic Diabetic Retinopathy Grading System Based on Detecting Multiple Retinal Lesions," *IEEE Access*, vol. 9, pp. 15939-15960, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[23] Hassan Tariq et al., "Performance Analysis of Deep-Neural-Network-Based Automatic Diagnosis of Diabetic Retinopathy," *Sensors*, vol. 22, no. 1, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] B. Venkaiahppalaswamy, PVGD Prasad Reddy, and Suresh Batha, "An Effective Diagnosis of Diabetic Retinopathy Based on 3d Hybrid Squeezenet Architecture," *International Journal of Engineering Trends and Technology*, vol. 70, no. 12, pp. 147-159, 2022. [CrossRef] [Publisher Link]

[25] Alnur Alimanov, and Md Baharul Islam, "Retinal Image Restoration and Vessel Segmentation using Modified Cycle-CBAM and CBAM-Unet," *2022 Innovations in Intelligent Systems and Applications Conference (ASYU), IEEE*, pp. 1-6, 2022. [Google Scholar]

[26] Zhaomin Yao et al., "Funswin: A Deep Learning Method to Analysis Diabetic Retinopathy Grade and Macular Edema Risk Based on Fundus Images," *Frontiers in physiology*, 1462, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[27] Najib Ullah et al., "Diabetic Retinopathy Detection Using Genetic Algorithm-Based CNN Features and Error Correction Output Code SVM Framework Classification Model," *Wireless Communications and Mobile Computing*, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] Muhammad Mohsin Butt et al., "Diabetic Retinopathy Detection from Fundus Images of the Eye Using Hybrid Deep Learning Features," *Diagnostics*, vol. 12, no. 7, p. 1607, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[29] APTOS 2019 Blindness Detection Dataset. [Online]. Available: https://www.kaggle.com/c/aptos2019-blindness-detection/

[30] Prasanna Porwal et al., "Indian Diabetic Retinopathy Image Dataset (Idrid): A Database for Diabetic Retinopathy Screening Research," *Data*, vol. 3, no. 3, p. 25, 2018. [CrossRef] [Google Scholar] [Publisher Link]

Avinashilingam Institute for Home Science and Higher Education for Women
(Deemed to be University Estd. u/s 3 of UGC Act 1956, Category A by MHRD)
Re-accredited with 'A++' Grade by NAAC.CGPA 3.65/4, Category I by UGC
Coimbatore - 641 043, Tamil Nadu, India

PLAGIARISM CHECK REPORT (THESES)

1.	Name of the Research Scholar	S. Valarmathi
2.	Roll No. and Year of Registration	19PHCSF006, 2019
3.	Department	Computer Science
4.	Name of the Research Guide	Dr. R. Vijayabhanu
5.	Title of the Thesis / Dissertation	An Optimized Convolutional Neural Network-based Ensemble Classification and Regression Framework for Classifying the Stages of Diabetic Retinopathy
6.	Similarity Content (%) Identified	**8%**
7.	Software Used	Turnitin
8.	Date of Verification	13-05-2024

Note : The report is excluding 14 Consecutive words, Review of Literature and Quoted Materials.

Checked by :

13/5/24
Information Scientist

13/05/2024
Research Scholar

13.05.24
Assistant Librarian

Research Guide

Date: 13-05-2024

An Optimized Convolutional Neural Network-based Ensemble Classification and Regression Framework for Classifying the Stages of Diabetic Retinopathy

by Central Library Avinashilingam

Submission date: 13-May-2024 04:22PM (UTC+0530)
Submission ID: 2378179029
File name: Plag_4_Final_Plag_check-VS_Th.docx (18.42M)
Word count: 32689
Character count: 183345

An Optimized Convolutional Neural Network-based Ensemble Classification and Regression Framework for Classifying the Stages of Diabetic Retinopathy

ORIGINALITY REPORT

8%	5%	6%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	oaji.net Internet Source	1%
2	www.internationaljournalssrg.org Internet Source	1%
3	"Intelligent Systems", Springer Science and Business Media LLC, 2021 Publication	<1%
4	dokumen.pub Internet Source	<1%
5	ebin.pub Internet Source	<1%
6	www.researchgate.net Internet Source	<1%
7	Mohammad T. Al-Antary, Yasmine Arafa. "Multi-scale Attention Network for Diabetic Retinopathy Classification", IEEE Access, 2021 Publication	<1%