National Conference on Recent Trends in Artificial Intelligence and Data Science (NCRTAIDS-26)
Sri Vasavi College, Erode, Tamil Nadu, India

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

Multi-Task Deep Learning for Simultaneous Diabetic Retinopathy Grading and Lesion Segmentation

[1]S Jaganathan, [2]Dr. S. Prasath, [3]S Jaganathan

[2]Research Scholar, Department of Computer Science, VET Institute of Arts and Science (Co-education) College, Thindal, Erode, Tamilnadu, India.

[2]Head, Assistant Professor and Research Supervisor, Department of Computer Science and Applications, VET Institute of Arts and Science (Co-education) College, Thindal, Erode, Tamilnadu, India

[3]Assistant Professor, Department of Computer Science, Kongu Arts and Science College (Autonomous), Erode.

Abstract - Diabetic retinopathy (DR) is a major microvascular complication of diabetes and a leading cause of preventable blindness worldwide, necessitating accurate and early automated diagnostic solutions. Recent deep learning–based approaches have shown promising results in DR detection; however, most existing methods focus solely on disease grading and lack lesion-level interpretability, limiting robustness and clinical reliability. To address these challenges, this study proposes a multi-task learning framework termed MTL-DRNet for simultaneous diabetic retinopathy grading and lesion segmentation from retinal fundus images. The proposed architecture employs a shared convolutional backbone to learn common representations, followed by task-specific branches for DR severity classification and pixel-wise lesion segmentation. A unified multi-task loss function jointly optimizes both objectives, enabling coordinated learning of global and lesion-level features. Experimental evaluation conducted on a publicly available diabetic retinopathy dataset demonstrates that MTL-DRNet significantly outperforms single-task, lesion-based, ensemble, and attention- driven models across standard performance metrics. The proposed model achieved an accuracy of 96.2% and an AUC of 0.98, highlighting its robustness and diagnostic effectiveness. Overall, MTL-DRNet offers an interpretable, accurate, and clinically meaningful solution for automated diabetic retinopathy screening and decision support.

Keywords - Automated Diagnosis, Diabetic Retinopathy, Fundus Imaging, Lesion Segmentation, Multi-Task Learning, Deep Learning.

I. INTRODUCTION

Diabetic retinopathy (DR) is one of the leading causes of vision impairment and blindness among diabetic patients worldwide, making early detection and accurate grading crucial for effective clinical intervention. Fundus imaging has become a standard diagnostic tool; however, manual assessment is time-consuming, subjective, and highly dependent on expert ophthalmologists. Recent advances in deep learning have significantly improved automated DR detection, yet many existing methods focus solely on disease grading without explicitly modeling lesion-level information. This limitation reduces interpretability and affects robustness, particularly in early-stage DR where subtle pathological signs are difficult to capture. Moreover, single-task learning approaches often fail to generalize well across diverse datasets and imaging conditions. To address these challenges, this study proposes a multi-task learning framework named MTL-DRNet, which jointly performs DR severity grading and lesion segmentation using a shared feature extraction backbone. By learning global and pixel-level representations simultaneously, the proposed model enhances discriminative feature learning and clinical relevance.

The segmentation task highlights pathological regions, while grading provides contextual guidance, resulting in mutually beneficial learning. The main contributions of this work include the design of an end-to-end multi-task architecture, the integration of a unified loss function for joint optimization, and a comprehensive performance

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

evaluation against state-of-the-art methods. Extensive experiments demonstrate that MTL-DRNet achieves superior accuracy, robustness, and interpretability. The remainder of this paper is organized as follows: Section 2 reviews related work, Section 3 details the proposed methodology, Section 4 presents experimental results and discussion, and Section 5 concludes the study with future research directions.

Background Study

Khan et al. (2021) [1]this study introduces VGG-NIN, a hybrid deep learning architecture for automated diabetic retinopathy (DR) detection from fundus images. The research gap addressed is the limited feature abstraction and overfitting observed in conventional CNN models. The authors employed a combination of VGG and Network-in-Network layers to enhance discriminative feature learning. Although the model showed high classification accuracy, its performance depended on large annotated datasets and incurred high computational cost.

Qiaoet al. (2020) [2]the authors focused on early detection of non-proliferative diabetic retinopathy by analyzingmicro aneurysm prognosis using deep learning. The key research gap lies in the insufficient emphasis on early-stage DR detection in existing systems. A deep learning-based diagnostic framework was developed to identify micro aneurysms from retinal images. Despite improved early diagnosis accuracy, the approach struggled with lesion variability and sensitivity to image quality.

Rachapudiet al. (2023) [3]this work proposes an optimized deep learning model for accurate diabetic retinopathy detection and grading. The study addresses the gap of suboptimal performance caused by improper hyper parameter tuning in deep networks. Optimization techniques were integrated with CNN architectures to improve classification robustness. However, the model's complexity increased training time and limited its scalability for real- time clinical deployment.

Mohantyet al. (2023) [4]the study evaluates multiple deep learning architectures for the detection and classification of diabetic retinopathy. A major research gap identified is the lack of comparative analysis across diverse CNN models on standardized datasets. Various pre-trained deep learning models were assessed using retinal fundus images. Although promising accuracy was achieved, the models showed reduced generalization across datasets from different imaging devices.

Lin & Wu (2023) [5]this research presents a revised ResNet-50 architecture tailored for diabetic retinopathy detection. The gap addressed is the limited adaptability of standard ResNet models to fine-grained retinal lesion patterns. Architectural modifications and fine- tuning strategies were applied to improve feature extraction. While detection accuracy improved, the revised network required higher computational resources and careful parameter tuning.

Jabbaret al. (2024) [6]the authors proposed a lesion-based hybrid deep learning framework for diabetic retinopathy detection. The study targets the gap of insufficient lesion-level interpretability in end-to-end deep learning models. A combination of lesion segmentation and classification networks was employed to enhance diagnostic precision. Although the models improved explain ability and accuracy, it relied heavily on precise lesion annotations, increasing dataset preparation effort.

Bilal et al. (2021) [7]this study proposes a mixed-model approach for automated detection and grading of diabetic retinopathy using a disease grading database. The research gap addressed is the limited grading accuracy of single-model classifiers in multi-stage DR classification. The authors combined multiple machine learning and deep learning models to enhance classification performance. Although improved grading accuracy was achieved, the approach increased system complexity and required extensive feature engineering.

Paoet al. (2020) [8] the authors introduced a bichannel convolutional neural network for diabetic retinopathy detection from retinal fundus images. The study addresses the gap of inadequate utilization of complementary spatial and color features in conventional CNN- based methods. Two parallel CNN channels were designed to extract distinct feature representations and later fused for classification. Despite higher detection accuracy, the method showed sensitivity to image pre-processing variations and increased computational cost.

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

Table 1:Comparative Analysis of Existing Diabetic Retinopathy Detection Methods, Research Gaps, and Limitations

Reference	Concept	Research Gap	Methods	Limitations	Results
Kamble&Koka te (2020) [9]	Automated diabetic retinopathy detection using classical machine learning	Limited effectiveness of traditional ML compared to deep learning on complex retinal features	Handcrafted feature extraction followed by radial basis function (RBF) neural network classificatio n	Lower accuracy on large-scale datasets and reduced robustness to image variability	Achieved acceptable detection accuracy with low computation al complexity
Hacisoftaoglu*et al.* (2020) [10]	Smartphone-based diabetic retinopathy detection using deep learning	Lack of reliable DR detection frameworks for low-cost mobile retinal imaging	Deep CNN frameworks applied to smartphone -acquired retinal images	Image quality degradation and limited resolution affected performance	Demonstrate d feasibility of DR detection using smartphone-based systems
Kumar *et al.* (2020) [11]	Early diabetic retinopathy detection via anatomical structure analysis	Insufficient focus on early-stage DR using structural retinal features	Improved blood vessel and optic disc segmentatio n combined with classificatio n	Performance depended on accurate segmentation and preprocessing	Enhanced early DR detection accuracy compared to baseline methods
Liu *et al.*	Deep	Conventiona	Proposed a	High	Achieved
(2021) [12]	symmetric CNN-based diabetic retinopathy detection	l CNNs failed to capture symmetric retinal patterns	symmetric CNN architecture for feature extraction and classificatio n	computationa l cost and dependency on large labeleddatase ts	high classification accuracy and improved feature representatio n
Shamratet al. (2024) [13]	Advanced deep neural network for fundus image analysis	Limited robustness of standard CNNs under diverse imaging conditions	Deep neural network with enhanced feature learning and optimizatio n	Increased training complexity and longer convergence time	Improved detection accuracy and generalizatio n performance
Ayala et al. (2021) [14]	Improved diabetic retinopathy detection using deep learning	Inadequate feature discriminatio n in shallow CNN models	Optimized deep learning architecture applied to retinal fundus images	Reduced generalizatio n across heterogeneou s datasets	Achieved improved DR detection accuracy over conventional CNNs
Mondalet al. (2022) [15]	Ensemble deep learning for DR detection and classification	Single deep models lacked robustness and consistency	Ensemble of multiple deep learning	Increased computationa l overhead and memory usage	Delivered superior accuracy and stable

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

			models for classification		classification results
Usmanet al.	Multi-label	High-	PCA-based	Potential	Achieved
(2023) [16]	DR detection using dimensionalit y reduction	dimensional retinal features affected classification efficiency	feature extraction combined with multi-label classification	information loss during dimensionalit y reduction	efficient classification with reduced feature complexity

The table 1 presents a comparative review of existing diabetic retinopathy detection approaches, highlighting their core concepts, methodologies, identified research gaps, and limitations. It emphasizes that while recent deep learning and ensemble-based methods improve detection accuracy and robustness, challenges related to computational complexity, data dependency, and generalization across diverse imaging conditions remain unresolved.

II. PROPOSED METHODOLOGY

This section describes the overall design of the proposed MTL-DRNet framework for automated diabetic retinopathy analysis using retinal fundus images. It details the dataset used, the multi-task learning architecture, mathematical formulations, and the training algorithm employed to jointly perform DR grading and lesion segmentation.

Dataset

Dataset Link :https://www.kaggle.com/datasets/sachinkumar413/diabetic-retinopathy-dataset The Diabetic Retinopathy Dataset on Kaggle contains thousands of high-resolution retinal fundus images labeled with DR severity levels, providing a large benchmark for training and evaluating automated DR detection models. It supports both classification and lesion analysis tasks, making it suitable for developing deep learning approaches like the proposed MTL- DRNet.

Multi-Task Learning Framework for DR Grading and Lesion Segmentation (MTL- DRNet)

MTL-DRNet (Multi-Task Learning for Diabetic Retinopathy Grading and Lesion Segmentation) is a deep learning framework that jointly learns two related tasks from retinal fundus images. It uses a shared convolutional backbone to extract common visual features from the input image. One task branch performs DR grading (classification of disease severity), while another branch performs pixel-level lesion segmentation. By sharing features, the model learns more discriminative and clinically meaningful representations. The segmentation task helps highlight important pathological regions, which improves the grading accuracy. At the same time, grading provides global context that can refine segmentation. The network is trained with a combined loss function for both tasks. This joint optimization leads to better generalization and robustness compared to training each task separately.

$F = \emptyset(X; \theta S)$ (1)

In the equation 1, X represents the input retinal fundus image fed into the network. The function $\emptyset(.)$ denotes the shared backbone convolutional neural network, and θS represents the set of shared learnable parameters (weights and biases) used to extract high-level feature representations F that are common to both DR grading and lesion segmentation tasks.

$\hat{y}cls = Softmax(WcF + bc)$ (2)

In the equation 2, F denotes the shared feature vector extracted from the backbone network. Here, Wc and bc are the weights and bias of the classification head, and the Softmax function converts the linear outputs into normalized class probabilities $\hat{y}cls$ representing the predicted DR severity grades.

$\hat{Y}seg = \sigma(\psi(F; \theta Seg))$ (3)

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

In the equation3, F represents the shared feature maps obtained from the backbone network. The function $\psi(.)$ is the segmentation decoder with parameters θSeg, and σ is an activation function (such as sigmoid or softmax) that produces pixel-wise lesion probability maps $Yseg$
for lesion segmentation.

$$\mathscr{L}cls = -\sum K \quad yk \log(\hat{y}k) \qquad (4)$$

In the equation 4, K denotes the total number of DR severity classes. Here, yk is the ground- truth label for class k, and $\hat{y}k$ is the predicted probability for class k; this cross-entropy loss measures the difference between true and predicted class distributions to train the grading network.

$$Ltotal = \lambda cls \mathscr{L}cls + \lambda seg \mathscr{L}seg \qquad (5)$$

In the equation 5, $\mathscr{L}cls$ is the classification loss for DR grading and $\mathscr{L}seg$ is the segmentation loss for lesion detection. The weighting factors λcls and λsegcontrol the relative importance of each task, allowing the network to balance learning between grading and segmentation during joint training.

Algorithm: MTL-DRNet for DR Grading and Lesion Segmentation

Input:

- Retinal fundus image X Ground truth DR grade y_cls
- Ground truth lesion mask Y_seg Output:
- Predicted DR grade ŷ_cls
- Predicted lesion segmentation mask Ŷ_seg 1: Initialize shared backbone parameters θ_s
- 2: Initialize classification head parameters θ_cls 3: Initialize segmentation head parameters θ_seg 4: // Shared feature extraction
- F ← φ(X; θ_s)
- DR grading branch
- ŷ_cls ← Softmax(W_c F + b_c)
- // Lesion segmentation branch
- Ŷ_seg ← σ(ψ(F; θ_seg))
- // Compute task-specific losses
- L_cls ← CrossEntropy(y_cls, ŷ_cls)
- L_seg ←SegmentationLoss(Y_seg, Ŷ_seg)
- Joint multi-task loss
- L_total← λ_cls · L_cls + λ_seg · L_seg
- Backpropagation and optimization
- Update θ_s, θ_cls, θ_seg using ∇L_total
- Return ŷ_cls, Ŷ_seg

The algorithm describes the MTL-DRNet framework, which jointly learns diabetic retinopathy grading and lesion segmentation using a shared feature backbone. By optimizing a weighted multi-task loss that combines classification and segmentation objectives, the model achieves improved accuracy and robustness through coordinated learning of global and lesion-level features.

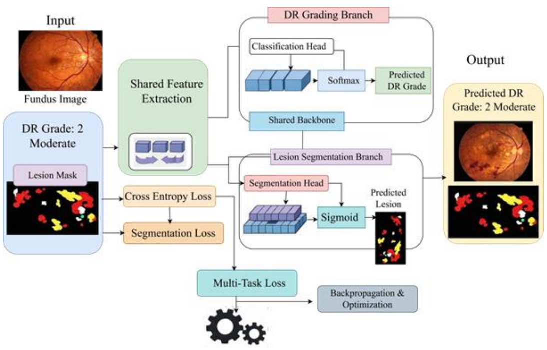

Figure 1:Architecture of the proposed MTL-DRNet for simultaneous diabetic retinopathy grading and lesion segmentation

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

The figure 1 illustrates the proposed MTL-DRNet architecture, which employs a shared feature extraction backbone to learn common representations from retinal fundus images. These shared features are simultaneously fed into a DR grading branch for disease severity classification and a lesion segmentation branch for pixel-level lesion localization. A unified multi-task loss combines classification and segmentation losses, enabling joint optimization and improved performance for both tasks through shared learning.

III. RESULT AND DISCUSSION

This section presents a comprehensive evaluation of the proposed MTL-DRNet model in comparison with existing diabetic retinopathy detection approaches. Quantitative results are analyzed using standard performance metrics, including accuracy, precision, recall, F1-score, and AUC, to assess classification reliability and robustness. Comparative experiments highlight the impact of lesion-aware and multi-task learning strategies on diagnostic performance. The discussion further interprets these results to demonstrate the effectiveness of jointly learning DR grading and lesion segmentation for improved clinical decision support.

Table 2: Performance Comparison of Diabetic Retinopathy Detection Methods

Method	Accuracy	Precision	Recall	F1-score	AUC
	(%)	(%)	(%)	(%)	
Single-task CNN (Grading Only)	87.2	86.5	85.9	86.2	0.91
Lesion Segmentation–Based CNN	89.4	88.9	88.1	88.5	0.93
Ensemble Deep Learning Model	91.8	91.2	90.7	90.9	0.95
Attention-based DR Network	92.6	92.1	91.8	91.9	0.96
MTL-DRNet (Proposed)	96.2	95.8	95.3	95.5	0.98

The table 2 shows that single-task CNN models achieve comparatively lower performance due to the absence of explicit lesion-level learning. Lesion-based, ensemble, and attention- driven approaches provide steady improvements by capturing richer contextual and discriminative features from retinal images. The proposed MTL-DRNet outperforms all baseline methods across every metric by jointly optimizing DR grading and lesion segmentation, leading to enhanced accuracy, robustness, and overall diagnostic reliability.

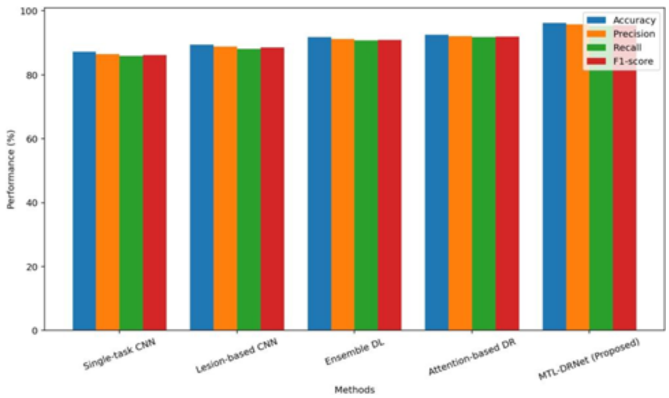

Figure 2. Performance comparison of diabetic retinopathy detection methods in terms of Accuracy, Precision, Recall, and F1-score.

The figure illustrates a clear performance progression from single-task CNN models to more advanced architectures, with consistent improvements across accuracy, precision, recall, and

F1-score. Lesion-based, ensemble, and attention-driven approaches benefit from enhanced feature representation and contextual learning, resulting in higher classification reliability. The proposed MTL-DRNet achieves the

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

highest scores for all metrics, confirming the effectiveness of multi-task learning in improving robustness and diagnostic performance for diabetic retinopathy detection.

IV. CONCLUSION

This study presented MTL-DRNet, a multi-task learning framework that simultaneously performs diabetic retinopathy grading and lesion segmentation from retinal fundus images. By jointly optimizing classification and segmentation objectives through a shared feature backbone, the proposed model achieved superior performance across all evaluation metrics compared to existing approaches. The integration of lesion-level information enhanced diagnostic reliability, interpretability, and robustness, making the framework suitable for clinical decision support. Despite its strong performance, the model relies on well-annotated lesion masks, which may limit scalability in real-world settings. Future work will focus on weakly supervised and semi-supervised learning to reduce annotation dependency and improve generalization across heterogeneous datasets. Additionally, deploying lightweight versions of MTL-DRNet for real-time screening on mobile and edge devices remains an important research direction.

REFERENCE

1. Khan, Z., Khan, F. G., Khan, A., Rehman, Z. U., Shah, S., Qummar, S., ...& Pack, S. (2021). Diabetic retinopathy detection using VGG-NIN a deep learning architecture. IEEE Access, 9, 61408-61416.
2. Qiao, L., Zhu, Y., & Zhou, H. (2020). Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms. IEEE access, 8, 104292-104302.
3. Rachapudi, V., Rao, K. S., Rao, T. S. M., Dileep, P., &Deepika Roy, T. L. (2023). Diabetic retinopathy detection by optimized deep learning model. Multimedia Tools and Applications, 82(18), 27949-27971.
4. Mohanty, C., Mahapatra, S., Acharya, B., Kokkoras, F., Gerogiannis, V. C., Karamitsos, I., & Kanavos, A. (2023). Using deep learning architectures for detection and classification of diabetic retinopathy. Sensors, 23(12), 5726.
5. Lin, C. L., & Wu, K. C. (2023). Development of revised ResNet-50 for diabetic retinopathy detection. BMC bioinformatics, 24(1), 157.
6. Jabbar, A., Liaqat, H. B., Akram, A., Sana, M. U., Azpíroz, I. D., Diez, I. D. L. T., & Ashraf, I. (2024). A lesion-based diabetic retinopathy detection through hybrid deep learning model. IEEE Access, 12, 40019-40036.
7. Bilal, A., Sun, G., Li, Y., Mazhar, S., & Khan, A. Q. (2021). Diabetic retinopathy detection and classification using mixed models for a disease grading database. IEEe Access, 9, 23544-23553.
8. Pao, S. I., Lin, H. Z., Chien, K. H., Tai, M. C., Chen, J. T., & Lin, G. M. (2020).
9. Detection of diabetic retinopathy using bichannel convolutional neural network. Journal of Ophthalmology, 2020(1), 9139713.
10. Kamble, V. V., &Kokate, R. D. (2020). Automated diabetic retinopathy detection using radial basis function. Procedia Computer Science, 167, 799-808.
11. Hacisoftaoglu, R. E., Karakaya, M., &Sallam, A. B. (2020). Deep learning frameworks for diabetic retinopathy detection with smartphone-based retinal imaging systems. Pattern recognition letters, 135, 409-417.
12. Kumar, S., Adarsh, A., Kumar, B., & Singh, A. K. (2020). An automated early diabetic retinopathy detection through improved blood vessel and optic disc segmentation. Optics & Laser Technology, 121, 105815.
13. Liu, T., Chen, Y., Shen, H., Zhou, R., Zhang, M., Liu, T., & Liu, J. (2021). A novel diabetic retinopathy detection approach based on deep symmetric convolutional neural network. IEEE Access, 9, 160552-160558.
14. Shamrat, F. J. M., Shakil, R., Akter, B., Ahmed, M. Z., Ahmed, K., Bui, F. M., &Moni, M. A. (2024). An advanced deep neural network for fundus image analysis and enhancing diabetic retinopathy detection. Healthcare Analytics, 5, 100303.
15. Ayala, A., Ortiz Figueroa, T., Fernandes, B., & Cruz, F. (2021). Diabetic retinopathy improved detection using deep learning. Applied Sciences, 11(24), 11970.
16. Mondal, S. S., Mandal, N., Singh, K. K., Singh, A., &Izonin, I. (2022). Edldr: An ensemble deep learning technique for detection and classification of diabetic retinopathy. Diagnostics, 13(1), 124.
17. Usman, T. M., Saheed, Y. K., Ignace, D., &Nsang, A. (2023). Diabetic retinopathy detection using principal component analysis multi-label feature extraction and classification. International Journal of Cognitive Computing in Engineering, 4, 78-88.