STATE-OF-THE-ART CNN APPLICATIONS IN EYE CANCER DETECTION: A REVIEW OF RECENT DEVELOPMENTS (2021-2024)

Aparna Paliya and Amit Singhal

Department of Computer Science,
Institute of Engineering and Technology, Dr. Bhimrao Ambedkar University, Khandari, Agra, UP, India

ABSTRACT

Eye cancers such as retinoblastoma, ocular melanoma, and eyelid tumors present diagnostic difficulties due to their rarity, diversity, and the increased risk of vision loss or death if detected late. Recent advancements in deep learning, particularly through convolutional neural networks (CNNs), have significantly improved the precision and automation in the detection and segmentation of ocular tumors. This review article examines the latest applications of Convolutional Neural Networks (CNNs) in detecting eye cancer, drawing on findings from ten recent studies conducted between 2021 and 2024. It highlights the importance of CNNs in improving diagnostic accuracy and efficiency in ophthalmology, suggesting they could significantly change clinical practices. Results reveal developments in CNN architectures that lead to better detection rates on the retinoblastoma and melanoma among different eye cancer types by analysing retinal images (e.g. fundus imaging, OCT, MRI) and histological records. It also introduces the clinical impact of these improvements referring to embedding the CNNs in routine tests. Recent advancements in deep learning, particularly through convolutional neural networks (CNNs), have significantly improved the precision and automation in the detection and segmentation of ocular tumors. The models analyzed include traditional architectures like LeNet and VGG16, as well as more advanced methods such as variations of U-Net, multi-view CNNs, ConvNeXt, and hybrid systems like FedCNN paired with XGBoost. Nonetheless, issues remain concerning dataset diversity, clinical validation, interpretability, and the practical use of these technologies. This review highlights both the notable advancements achieved and the gaps that still need to be filled to enable the incorporation of CNN-based systems into standard ocular cancer treatment.

KEYWORDS

Eye cancer detection, Convolutional Neural Networks, deep learning, retinoblastoma, ocular melanoma, early diagnosis, medical imaging.

1. INTRODUCTION

1.1. Background

Eye cancers include a wide range of intraocular and periocular malignancies, the most famous of which are retinoblastoma, ocular melanoma, and eyelid tumors. Retinoblastoma, a malignancy that mostly affects children and is associated to abnormalities in the RB1 gene, is the most prevalent intraocular tumor in this age group, with a median diagnosis of two years (Priya et al., 2021). Symptoms such as leukocoria and strabismus can help with early detection; nevertheless, misdiagnosis is common due to parallels with other ocular disorders (Rahdar et al., 2023). Every year, around 8,000 new cases are reported worldwide, with a disproportionately high frequency in low-resource settings (Priya et al., 2021).

The most common type of intraocular malignancy in adults is ocular melanoma, namely uveal melanoma. It is caused by melanocytes in the uveal tract and is highly likely to metastasize, particularly to the liver (Strijbis et al., 2021). Other malignancies, such as eyelid tumors, including basal cell carcinoma and sebaceous carcinoma, have a significant impact on ocular health and appearance (Wang et al., 2024).

The early detection of these tumors is critical; nevertheless, clinical diagnosis remains challenging. Traditional procedures such as ophthalmoscopy, OCT, MRI, histology, and angiography have limitations due to expense, subjectivity, and the availability of specialists (Sinha et al., 2021; Goswami, 2021). Furthermore, emerging methods such as ion beam-based Raman spectroscopy have demonstrated potential for detecting changes in tumor microstructure, but their clinical application is still in its early stages (Sur et al., 2023). These barriers underscore the importance of automated, precise, and easily accessible diagnostic solutions.

1.2. Role of CNNS in Medical Imaging

By making it easier to automatically extract distinctive features from complex datasets, deep learning—in particular, Convolutional Neural Networks, or CNNs—has revolutionized the field of medical image analysis. As opposed to traditional machine learning techniques that rely on manually constructed features, CNNs may learn hierarchical representations directly from imaging data, which makes them incredibly effective for medical imaging tasks.

Ocular oncology applications include a range of methods. LeNet and VGG16 architectures are two fundus-based methods that have been used to detect retinoblastoma and ocular melanoma with 92–95% accuracy (Priya et al., 2021; Sinha et al., 2021). According to Rahdar et al. (2023), the use of sophisticated semi-supervised CNNs has greatly increased segmentation accuracy; for retinoblastoma tumors in fundus pictures, Dice coefficients are higher than 0.90. Multi-view CNNs in MRI analysis have made it easier to segment retinoblastoma volumetrically, which helps determine how well a tumor responds to treatment (Strijbis et al., 2021). Ocular tumor segmentation accuracy has also improved thanks to OCT-based models that use U-Net and Inception backbones (Goswami, 2021).

In recent years, hybrid and federated learning approaches have become more prevalent. For example, the combination of FedCNN and XGBoost has demonstrated significant effectiveness in strabismus identification, highlighting the potential of privacy-preserving training in a variety of educational settings (Jabbar et al., 2024). Parallel to this, ConvNeXt designs demonstrated an AUC of 0.95 across 30 categories of ocular illnesses, including malignancies, when trained using ultra-widefield fundus angiography (Wang et al., 2024b). The significance of CNN-inspired approaches in multi-modal tumor research is further demonstrated by the use of ensemble deep learning and Raman spectroscopy for the characterization of retinoblastoma tissue, which goes beyond imaging techniques (Sur et al., 2023).

Together, these approaches show that CNNs can identify minor imaging characteristics that doctors might miss, reducing diagnostic subjectivity and enabling scalable screening for ocular malignancies.

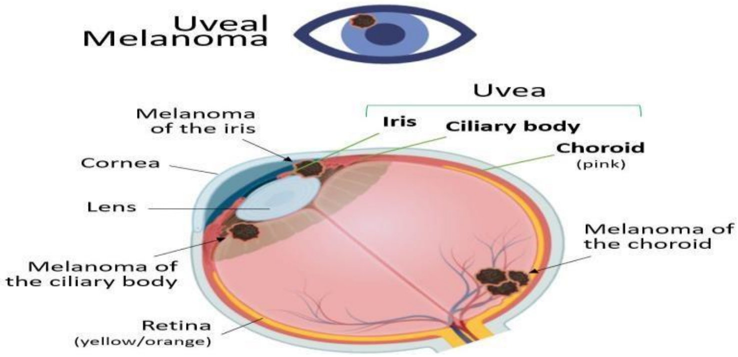

Fig. 1 Uveal Melanoma (from google)

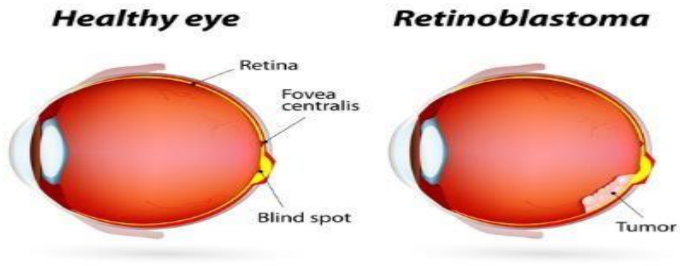

Fig. 2 Healthy eye and Retinoblastoma (from google)

2. Literature Review

2.1. Retinoblastoma Detection and Segmentation

A rare intraocular cancer in children, retinoblastoma has attracted a lot of interest in CNN-based eye cancer research. Several imaging modalities have been studied in numerous studies to improve treatment oversight and diagnosis accuracy.

One of the first CNN designs, LeNet, was used by Sinha et al. (2021) to identify cancers in general fundus images. With accuracies of over 92%, their model showed that deep learning might be used to detect ocular cancers in routine imaging procedures. Similarly, in order to detect retinoblastoma in fundus pictures and evaluate tumor regression after therapy, Priya et al. (2021) presented a VGG16-based system that integrated CNN classification with image pre-processing. The multi-view CNN (MV-CNN) presented by Strijbis et al. (2021) was intended to automatically segment retinoblastoma in MRI images in an effort to achieve segmentation, and it achieved a Dice coefficient of 0.91. This development made exact volumetric analysis possible, which is crucial for monitoring treatment outcomes. Additionally, Rahdar et al. (2023) developed a semi-supervised CNN for fundus image segmentation using Gaussian Mixture pre-labelling, which produced Dice scores higher than 0.90. Their study demonstrated how well-suited semi-supervised approaches are for dealing with the problem of sparsely annotated data.

Additionally, U-Net versions combined with VGG16, Inception, and ResNet backbones were used by Goswami (2021) to segment ocular malignancies using OCT and OCT angiography. These topologies demonstrated the efficacy of U-Net derivatives for high-resolution retinal imaging by achieving Dice scores ranging from 0.87 to 0.97. Furthermore, Sur et al. (2023) achieved classification accuracies of over 90% by applying ensemble machine learning to Raman spectroscopic data obtained from retinoblastoma tissues, going beyond imaging approaches. The increasing importance of CNN-inspired techniques in multimodal tumor diagnosis is highlighted by this methodology.

These studies collectively demonstrate the effectiveness of CNN-based models in detecting and classifying retinoblastoma in a variety of modalities, with results routinely exceeding 90% accuracy or similar measures.

2.2. Ocular Melanoma and Ocular Surface Tumors

Deep learning has also been applied to the treatment of ocular surface squamous neoplasia (OSSN) and ocular melanoma. AI-driven classifier for OSSN detection was developed by Sinha et al. (2024) in a recent work. While their solution does not solely rely on CNNs, it does highlight the versatility of machine learning models in non-invasive ocular surface cancer screening methods.

MRI is still an important imaging method for intraocular malignancies. As previously noted, Strijbis et al. (2021) further demonstrated the utility of CNNs beyond pediatric cancers by showing that they may successfully define tumor margins in cases of ocular melanoma.

2.3. Eyelid Tumors

Eyelid tumors are among the most common periocular cancers, and recent developments in CNN applications have expanded to include these cancers. A proteomics-driven AI model was developed by Wang et al. (2024a) with the goal of categorizing eyelid cancers. Despite not being a conventional CNN, the study used high-dimensional proteomic data in conjunction with machine learning approaches, yielding an accuracy of 84.8% (AUC 0.80–1.0). This demonstrates the potential of combining omics data with AI techniques in ocular oncology, particularly in situations where imaging features by themselves are insufficient.

2.4. Hybrid Models and Multi-Disease Detection

Hybrid frameworks and systems for multi-disease detection have been studied recently, indicating a shift from single-disease CNN applications to a wider range of clinical applications. A federated CNN (FedCNN) combined with XGBoost was presented by Jabbar et al. (2024) for the purpose of detecting strabismus using OCT and eye-tracking data. Their findings demonstrate the adaptability of CNN-based methods in ocular diagnostics and underscore the promise of federated learning for cross-institutional collaboration while protecting patient privacy, despite their primary focus being on a non-oncological ailment.

Similarly, Wang et al. (2024b) created a ConvNeXt CNN backbone for ultra-widefield fundus fluorescein angiography (UWFFA) analysis. Their model demonstrated the scalability of complex CNN architectures for large multi-class datasets by achieving AUC values ranging from 0.94 to 0.95 across 30 ocular illnesses, including malignancies.

3. KEY FINDINGS

- Even LeNet, one of the original CNN designs, was shown by Sinha et al. (2021) to be able to accurately identify eye cancers in fundus pictures with an accuracy of above 92%.
- This was further upon by Priya et al. (2021) using VGG16, which produced AUC values of roughly 0.92 for the detection of retinoblastoma.

- More sophisticated designs were demonstrated by Goswami (2021), where U-Net variants combined with VGG16, Inception, and ResNet backbones produced Dice coefficients as high as 0.97 in OCT/OCTA-based segmentation.
- The work of Strijbis et al. (2021) advanced the field of MRI applications by using a multi-view CNN to achieve a Dice score of 0.91 for volumetric segmentation of retinoblastoma.
- Accuracy: Most models achieved accuracies above 90%. For example, Jabbar et al. (2024) reported a 95.2% accuracy using FedCNN in conjunction with XGBoost for the identification of eye disorders.
- Dice Scores: Segmentation tasks using semi-supervised CNNs (Rahdar et al., 2023) and U-Net versions (Goswami, 2021) yielded Dice coefficients between 0.90 and 0.97.
- AUC Values: Wang et al. (2024b) used ConvNeXt to perform a thorough analysis of fundus angiography, and the results showed AUC values ranging from 0.94 to 0.95 for 30 ocular disorders, including malignancies.
- In order to reduce reliance on large labelled datasets, Rahdar et al. (2023) introduced a semi-supervised convolutional neural network that pre-labels using Gaussian mixture models.
- In order to demonstrate how multimodal data might improve limited imaging datasets, Sur et al. (2023) combined ensemble classifiers with Raman spectroscopy.
- Federated learning, as demonstrated by Jabbar et al. (2024), improves data diversity while maintaining privacy by enabling collaboration across several institutions without requiring data exchange.
- Wang et al. (2024a) classified eyelid cancers with an accuracy rate of 84.8% using AI models using proteomic data.
- This development highlights CNN-inspired models' potential for integrating multi-omics and improving morphological imaging with molecular insights for more precise tumor characterisation.
- The case for integrating CNNs into diagnostic procedures is strengthened by the excellent performance shown across multiple modalities.
- Scalable and privacy-preserving methods appropriate for practical application are provided by federated learning and semi-supervised approaches (Jabbar et al., 2024; Rahdar et al., 2023).
- Multiple diseases CNN frameworks, as demonstrated by Wang et al. (2024b), provide a major step forward in the direction of clinical integration and demonstrate the possibility for comprehensive ocular diagnostic systems.

Author(s)	Year	Problem (Eye Cancer Type / Task)	Model Used	Accuracy / Dice / AUC
Sinha et al.	2021	Eye tumour detection (general, fundus images)	LeNet CNN	~92–94% Acc
Priya et al.	2021	Retinoblastoma detection (fundus images)	VGG16 CNN	AUC ~0.92
Goswami	2021	Ocular tumor segmentation (OCT/OCTA)	U-Net variants (VGG16, Inception, ResNet hybrids)	Dice 0.87–0.97
Strijbis et al.	2021	Retinoblastoma segmentation (MRI)	Multi-view CNN (MV-CNN)	Dice 0.91
Rahdar et al.	2023	Retinoblastoma segmentation (fundus images)	Semi-supervised CNN (with GMM pre-labeling)	~93% Acc (Dice)
Sur et al.	2023	Retinoblastoma tissue imaging (Raman spectroscopy + ML)	Ensemble learning with CNN-like classifiers	>90% Acc
Wang et al. (2024a)	2024	Eyelid tumor classification (proteomics)	AI/ML classifier (non-CNN)	84.8% Acc (AUC 0.80–1.0)
Jabbar et al	2024	Strabismus / ocular disease detection (eye-tracking, OCT)	FedCNN + XGBoost	95.2% Acc
Wang et al. (2024b)	2024	Multi-disease detection incl. ocular tumors (UWFFA)	ConvNeXt CNN backbone	AUC 0.94–0.95
Sniha et al.	2024	Ocular surface squamous neoplasia (OSSN) detection	AI/ML classifier (non-CNN)	~90% Acc

Table 1. Comparative Table of CNN Applications in Eye Cancer Detection (2021–2024)

4. CHALLENGES AND LIMITATIONS

One of the major obstacles is the scarcity of annotated datasets for ocular cancers, especially the rare diseases retinoblastoma and ocular melanoma. The generalizability of CNN models is limited by the availability of small, single-center datasets (Strijbis et al., 2021; Sinha et al. 2021). The issue of dataset diversity is related to this; many models have not been validated in different populations, imaging systems, or clinical settings, which calls into question how reliable they will be in practical settings (Priya et al. 2021).

Another drawback is the possibility of overfitting when models are created with limited data. Goswami (2021), for example, used U-Net versions to generate impressive segmentation results; however, these models might not perform as well outside of controlled datasets. Similarly, Sur et al. (2023) found great accuracy using Raman spectroscopy data; nevertheless, its clinical usefulness is limited due to the lack of external validation.

Additionally, there are few attempts to integrate various modalities; most research is still focused on a single modality, either fundus, OCT, or MRI data. Although Wang et al. (2024a) highlighted the potential of proteomics, there isn't yet a unified framework that combines omics and imaging data into a comprehensive diagnostic tool.

In the end, there are still barriers to clinical translation. Many CNNs operate as "black boxes," offering little interpretability, which erodes physician confidence (Wang et al., 2024b). Complex designs such as ConvNeXt have computational requirements that make it difficult to apply them in resource-constrained environments. Logistically, even privacy-preserving strategies like federated learning (Jabbar et al. 2024) face challenges such high communication costs and the need for uniformity throughout several centers.

5. FUTURE DIRECTIONS

A comprehensive approach is required to address these issues. First and foremost, creating large, varied, and multi-institutional datasets is essential. International collaborations can produce libraries that reflect a range of imaging modalities and demographics. The creation of synthetic data using GANs and semi-supervised learning (Rahdar et al., 2023) offer complementary strategies for addressing the problem of annotation scarcity.

Second, architectural innovation should continue to be emphasized. To strike a compromise between accuracy and computational efficiency, future studies should give special emphasis to transformer-enhanced models, lightweight CNNs, and attention processes (Wang et al., 2024b). These models can be used in low-resource settings and may be able to detect subtle tumor features more accurately.

Third, there is potential for further research in the field of multi-modal integration. Combining omics data like proteomics (Wang et al., 2024a) or spectroscopy (Sur et al., 2023) with imaging modalities (fundus, OCT, MRI) may provide a more thorough understanding of tumor biology and enhance precision oncology in ocular tumors.

Fourth, the use of explainable AI (XAI) methods, including Grad-CAM and saliency maps, can improve transparency and make CNN predictions easier for regulators and doctors to understand (Wang et al., 2024b).

Fifth, while safeguarding patient data, the spread of federated learning frameworks among international centers would enable cooperative model training. Interoperability and reducing communication overhead should be the main priorities (Jabbar et al., 2024).

In order to evaluate CNN systems in practical workflows, prospective clinical studies are desperately needed. The impact of models on patient outcomes, diagnostic efficiency, and accuracy should be assessed. Given the high prevalence of retinoblastoma in low-resource locations, mobile-based CNN screening techniques designed for underdeveloped infrastructure should also be given priority in future studies.

6. CONCLUSION

Recent advances in deep learning, particularly convolutional neural networks (CNNs), have had a significant impact on the diagnosis and detection of ocular cancer. CNN-based models demonstrated remarkable accuracy in a variety of modalities, including fundus photography, OCT, MRI, Raman spectroscopy, and proteomics, between 2021 and 2024. While more complex architectures, such as U-Net variations, multi-view CNNs, ConvNeXt, and hybrid federated techniques, consistently obtained accuracies, Dice scores, and AUC values surpassing 90%, more conventional models, such as LeNet and VGG16, validated the usefulness of CNNs in ocular cancer. These findings demonstrate that CNNs are state-of-the-art instruments that can identify, segment, and classify ocular cancers with performance comparable to that of clinicians.

Even with these developments, there are still problems. Clinical translation is hampered by the scarcity of annotated datasets, the lack of demographic variety, the limited integration of several modalities, and interpretability problems. Many research have not yet received extensive, prospective validation and are still in the proof-of-concept stage. Prior to practical deployment, the computational needs and ethical issues must also be carefully addressed, particularly with regard to pediatric diseases like retinoblastoma.

Future developments in the field include the creation of lightweight and attention-based architectures, the multimodal fusion of imaging and omics, the development of explainable AI, the sharing of multi-institutional datasets, and the creation of federated learning frameworks. By overcoming these obstacles, CNNs can evolve from experimental prototypes to clinically integrated, internationally available diagnostic instruments. When properly implemented, they have the potential to revolutionize ocular oncology by promoting early identification, improving treatment planning, and enhancing patient outcomes globally.

REFERENCES

[1] Sinha, A., Aneesh, R. P., & Nazneen, N. S. (2021). Eye tumour detection using deep learning. In 2021 7th International Conference on Bio Signals, Images, and Instrumentation (ICBSII) (pp. 1–5). IEEE. https://doi.org/10.1109/ICBSII51839.2021.9445172

[2] Priya, R. L., Morogiri, O., Bauskar, P., Sagane, G., & Brahmapurkar, P. (2021). Proposed idea on detection of retinoblastoma and its response to treatment. In Proceedings of the 2nd International Conference for Emerging Technology (INCET) (pp. 1–6). IEEE. https://doi.org/10.1109/INCET51464.2021.9456339

[3] Goswami, M. (2021). Deep learning-based U-Net variants for ocular tumor segmentation using OCT and OCTA. Computers in Biology and Medicine, 134, 104462. https://doi.org/10.1016/j.compbiomed.2021.104462

[4] Strijbis, V. I., Hooijmans, J. F., Koopman, T., Beenakker, J. W. M., & de Graaf, P. (2021). Automatic segmentation of retinoblastoma on MRI using deep learning. Scientific Reports, 11(1), 16707. https://doi.org/10.1038/s41598-021-93905-2

[5] Rahdar, A., Ahmadi, M. J., Naseripour, M., Akhtari, A., Sedaghat, A., Hosseinabadi, V. Z., Yarmohamadi, P., Hajihasani, S., & Mirshahi, R. (2023). Semi-supervised segmentation of retinoblastoma tumors in fundus images. Scientific Reports, 13(1), 13010. https://doi.org/10.1038/s41598-023-39909-6

[6] Sur, S., Reddy, M. S., Paikaray, D., Barve, A., Faujdar, M., Sharma, S., & Awasthi, A. (2023). Nanostructure analysis in polymeric materials with ion beam–based Raman spectroscopy for retinoblastoma tumor imaging using ensemble machine learning. Optical and Quantum Electronics, 55(12), 940. https://doi.org/10.1007/s11082-023-05167-z

[7] Wang, H., et al. (2024a). Proteomic classification of eyelid tumors using artificial intelligence. Oncogenesis, 13(1), 767. https://doi.org/10.1038/s41698-024-00767-8

[8] Jabbar, A., Ahmad, J., Khan, W. A., Rho, S., & Mehmood, I. (2024). Federated deep learning-based ocular disease detection using eye-tracking and OCT. Scientific Reports, 14, 72919. https://doi.org/10.1038/s41598-024-72919-6

[9] Wang, J., et al. (2024b). Development of a deep learning system for ultra-widefield fundus angiography analysis. Eye and Vision, 11(1), 900. https://doi.org/10.1038/s40123-024-00900-7

[10] Sinha, S., Ramesh, P. V., Nishant, P., Morya, A. K., & Prasad, R. (2024). Novel automated non-invasive detection of ocular surface squamous neoplasia using artificial intelligence. World Journal of Methodology, 14(2), Article 92267. https://doi.org/10.5662/wjm.v14.i2.92267

AUTHORS

Aparna Paliya holds an M.Sc. in Computer Science. She is currently pursuing her Ph.D. at Dr. Bhimrao Ambedkar University, Agra. Her research interests include Machine learning, Deep learning and emerging areas of computer science and information technology. She has presented a paper at international conference and possesses over 6 months of teaching experience in higher education, where she has actively contributed to academic instruction, student development, and various departmental activities.

Dr. Amit Singhal is a Assistant Professor in the Department of Computer Science at Dr. Bhimrao Ambedkar University Agra. He holds a Ph.D. in Computer Science from the same University. He has nearly 24 years of teaching experience to PG classes. His research interest includes Wireless Networks and Mobile Computing. Dr. Singhal has published twelve research papers in the well reputed journals (API & IEEE) and attended two international conferences abroad and many national and international conferences in India.