



# A COMPREHENSIVE REVIEW OF DEEP LEARNING BASED TECHNIQUES FOR THE RETINITIS PIGMENTOSA DETECTION: A COMMON INHERITED RETINAL DISORDER

<sup>1</sup>Mr Manjunatha C Parameshwarappa, <sup>2</sup>Dr Jithendra P R Nayak

<sup>1</sup>Assistant Professor and Research Scholar, <sup>2</sup>Associate Professor

<sup>1</sup>Department of Computer Science, Sir M. V. Govt. Science College, Bhadravathi, India and Srinivas University, Mangalore, India. <sup>2</sup>Department of CSE, Srinivas Institute of Technology, Valachil, Mangalore, India

**Abstract:** Retinitis pigmentosa (RP) is a progressive inherited retinal disorder primarily due to the degeneration of rod and cone cells, leading to loss of vision as the time passes and may lead to legal blindness. As early diagnosis plays a crucial role in patient care and treatment planning, there has been growing interest in leveraging deep learning based techniques to improve the accuracy and efficiency of RP detection. This paper presents a comprehensive review of the current deep learning-based approaches developed for identifying RP using ophthalmic imaging modalities, such as fundus photographs and optical coherence tomography. We review frequently used neural network architectures, including convolutional neural networks (CNNs), as well as emerging techniques like transfer learning. In addition to discussing the datasets employed and performance benchmarks reported, this article also addresses key challenges such as data scarcity, model interpretability. This paper aims to provide researchers and clinicians with a clear understanding of the current landscape and future prospects of deep learning in the automated detection of retinitis pigmentosa.

**Index Terms:** Deep learning, Machine learning, Retinitis pigmentosa, Image processing

## 1. Introduction

Retinitis pigmentosa (RP) belongs to a group of inherited retinal diseases that occur due to defects in genes responsible for the normal functioning of the retinal layers. The occurrence of RP ranges between one patient out of 1000 to one patient out of 4000 individuals checked for RP. RP can present as early as infancy or in adulthood, leading to progressive visual deterioration due to the predominant degeneration of photoreceptors and the retinal pigment epithelium (RPE). The most common symptoms of RP include night blindness, loss of visual field (tunnel vision), sensitivity to bright light, and reduced central vision. RP is diagnosed by an ophthalmologist (retina specialist) through a detailed medical history and a fundus (retinal) examination. Signs of RP on fundus examination include attenuation of retinal blood vessels, bony spicule-like pigmentation, optic disc pallor, RPE degeneration, and chorioretinal atrophy involving the central macula or the rest of the retina [1].

Various therapies are under research to slow the degeneration of retinal layers and prevent further vision loss. Treatment approaches include gene therapy, cell therapy, optogenetics, and neuroprotection[2, 3]. At present the best approach in managing the RP focuses on early diagnosis, low vision aids, and providing guidance on future plans. RP can impose a significant social and economic burden, as it often leads to legal blindness at a young age[4]. Therefore, early and accurate diagnosis of RP is crucial for better patient management, family guidance, and future planning. At the same time, with promising newer therapies for RP, large-scale qualitative image analysis of the retina is becoming increasingly important as a clinical trial outcome measure.

Ancillary imaging modalities like images of color fundus, optical coherence tomography (OCT) and fundus autofluorescence (FAF) play a very significant contribution in the early detection of RP, prior to clinical symptoms are evident during an ophthalmologic examination. These imaging methods are non-invasive, fast, and effective in identifying subtle retinal changes. Deep learning (DL) has transformed various fields of medicine by providing automated, precise, and efficient diagnostic solutions. The application of deep learning-based image analysis in RP diagnosis and disease progression monitoring is emerging as a powerful tool[5].

This paper investigates the applications of DL techniques for detecting RP using OCT and color fundus imaging. It examines various image preprocessing procedures, neural network architectures, and performance metrics employed in RP diagnosis. Additionally, the study addresses current challenges and limitations, explores future directions, and considers the impact of DL on clinical practice in ophthalmology.

## 2. Imaging Modalities

### 2.1 Fundus image and retinitis pigmentosa

Color fundus photography employs the principles of indirect ophthalmoscopy, utilizing multiple lenses and a camera to capture high-resolution retinal images. Fundus imaging has the capability of capturing retinal views ranging from 50° up to a wide field of 200°. Advances in wide-field and ultra-wide-field fundus imaging have significantly transformed ophthalmic practice by enabling rapid image acquisition within a few seconds capturing details from the central to peripheral retina. Among these modalities, Optos ultra-wide-field imaging (UWFI) currently offers the most extensive retinal coverage (200 degrees) in clinical practice[6]. These imaging techniques also overcome the challenges of imaging a pediatric patient and in eyes with hazy ocular media, where conventional examinations may be challenging. Additionally, fundus imaging serves as an effective educational tool for patients and their families.

The findings of RP in color retinal fundus photographs vary according to the patient's age at examination and the specific gene responsible for RP. Early retinal signs include arterial narrowing, loss of the macular reflex, and subtle pigmentary mottling. As RP progresses, more distinct abnormalities emerge, such as attenuation of both arteries and veins, waxy pallor of the optic disc, prominent black bony spicule-like pigmentation, and chorioretinal atrophic patches, with or without macular involvement[1]. While ophthalmologists can identify these findings through examination, ensuring consistent and precise descriptions can be challenging due to interobserver variability. Features such as pigmentation and atrophy can differ in size, shape, location, and severity between eyes and among patients, introducing subjective errors. This variability highlights the importance of image-based analysis like deep learning, which enables automated and highly accurate RP detection. Furthermore, deep learning reduces the burden of manual grading, which often leads to significant time investment and may result in variability [7].

Deep learning can assess several key retinal features in RP, including the extent of optic disc pallor, blood vessel caliber, quadrant-wise distribution of background retinal changes, pigment characteristics (size, shape, and color density), chorioretinal atrophy extent, and macular abnormalities (e.g., pigmentation, thinning, or loss of normal reflex). Analyzing these features not only aids in early RP detection but also facilitates disease monitoring and visual prognosis [5, 8].

### 2.2 Optical coherence tomography (OCT) image and retinitis pigmentosa

Optical coherence tomography (OCT) utilizes low-coherence interferometry to generate high-resolution, cross-sectional images of the retina, capturing its structure layer by layer. It is one of the most frequently used imaging approaches for retinal diseases, evolving from time-domain OCT to spectral-domain and swept-source OCT, which provide enhanced image resolution. Like fundus imaging, OCT also provides wide-field as well as ultra-wide-field imaging technique options. However, OCT scans most commonly focus on the centre of the retina, with a primary target on the macula. OCT provides detailed visualization of all ten retinal layers, which includes the inner limiting membrane, retinal nerve fiber layer, ganglion cell layer, inner plexiform layer, inner nuclear layer, outer plexiform layer, outer nuclear layer, external limiting membrane, photoreceptor layer (ellipsoid zone), and the retinal pigment epithelium (RPE)[9]. Additionally, it allows for the examination of the preretinal/vitreous phase and deeper choroidal layers, offering comprehensive structural analysis.

In RP, OCT reveals distinct retinal changes. Characteristic OCT findings include signs of vitreous degeneration, thinning of the outer retinal layers, parafoveal tapering (from the outer nuclear layer to the RPE), reduced overall retinal thickness, residual ellipsoid zone and outer nuclear layer width at the macula, choroidal thinning, and hyperreflective dots[10]. OCT is also instrumental in detecting RP-associated complications such as cystoid macular edema, full-thickness macular holes, vitreomacular traction, and lamellar macular holes. Given its ability to provide detailed structural insights, OCT is far superior to direct clinical examination in assessing disease severity.

Deep learning applications in OCT imaging for RP offer significant advantages, as manual layer-by-layer analysis is time-consuming, labor-intensive, and prone to human error. DL algorithms can analyze key OCT features, including total retinal thickness, segmented layer analysis, residual ellipsoid zone and outer nuclear layer width, RPE thinning, and choroidal layer segmentation. Furthermore, deep learning-based OCT analysis has shown strong correlations with fundus autofluorescence and electroretinogram findings, improving disease severity assessment, identifying patient-specific patterns, predicting visual outcomes, and determining eligibility for clinical trials[7, 11].

## 3. Machine Learning(ML) And Deep Learning(DL)

Machine learning is the principal subset of artificial intelligence (AI). In the year 1959, Arthur Samuel, a prominent influencer in the area of AI called the machine learning term[12]. In the 17<sup>th</sup> century, the derivatives of least-squares methods by Laplace and Bayesian, along with the advent of Markov chains, laid the foundation step for numerous techniques that are used prominently in machine learning today [13]. From then, ML has evolved at fast pace and is now extensively used in many fields of technology, science and also in everyday life. It plays a vital role in areas such as data mining [14], computer vision [15], natural language processing(NLP) [16], human biometric recognition [17], diagnosis of medical images [18], fraud detection in banking transactions [19], prediction and analysis stock market[20], speech recognition [21], real time gaming systems, and in the design of robots [22].

Machine learning and deep learning are two key subfields of artificial intelligence that have gained significant attention in recent years, driven by the availability of big data [23] and advancements in computing power [24]. Machine learning is a data analysis method that automates model building, allowing computers to learn and improve from experience without explicit programming. On the other hand deep learning, a subset of machine learning, uses neural networks inspired by the human brain to solve complex problems and process unstructured or unlabeled data such as images [25], speech [26] and text [27].

Several deep learning frameworks exist for these tasks, including deep neural networks (DNNs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs) [28]. These technologies have transformed industries such as healthcare [29, 30], finance [31, 32], and retail industry [33, 34], impacting society by enabling automation, predictive analytics, and more sophisticated decision-making systems.

Computer vision is one such field that has experienced remarkable progress in today's world. By leveraging convolutional neural networks (CNNs), researchers have developed highly accurate models for object recognition and image classification [35]. These advancements have the potential to enhance medical imaging analysis [36], optimize quality control in manufacturing industries [37] and support driverless vehicles [38].

DL techniques provide promising solutions for automating the classification of retinal images [8, 39-43] and identifying genetic variations associated with RP [44, 45]. These models can analyze vast amounts of imaging data, detecting subtle features linked to specific genes. By leveraging DL, diagnostic accuracy and accessibility can be significantly improved [46].

## 4. Image Processing Approaches

Image processing techniques are essential for image acquisition, pre-processing, clustering, segmentation, and classification. They are extensively used to enhance analysis and interpretation across various image types such as vegetables and fruits, vehicles, text as well as in medical imaging [47].

### 4.1 Contrast Enhancement

Key features such as contrast and color, especially around the blood vessels and optic disc are captured in color retinal images. However, these images often suffer from noise caused by uneven illumination during the acquisition process. To reduce this noise and enhance image quality, image enhancement methods are applied [48].

A commonly used method is histogram equalization, known for its simplicity and low computational cost. Among various enhancement techniques, Contrast Limited Adaptive Histogram Equalization (CLAHE) has gained significant attention for its effectiveness in enhancing still images [49]. As proposed and summarized in various studies [50, 51], CLAHE is especially useful in improving contrast while limiting noise amplification, making it well-suited for retinal image enhancement [48].

### 4.2 Noise reduction:

Noise is an unwanted signal that can degrade the quality of an image, often arising during the capturing and transmission process. It can occur from various sources such as low light conditions, slow shutter speeds, sensor heat, and sensor fill factor [52]. Gaussian noise, balanced noise, and impulse noise are the general classifications of image noises [53]. Among these, impulse noise appears as randomly distributed light and dark pixels that not only distort the true information within the image but also significantly impair its visual quality [52].

The median filter is a nonlinear filtering technique commonly used in digital image processing due to its strong ability to preserve edges while effectively reducing impulse noise [53, 54]. The weighted average filter improves the functionality of the average filter by incorporating weighted contributions from neighboring pixels [55, 56].

### 4.3 Data Augmentation

Traditionally Geometric distortions, such as affine transformations, are widely used for data augmentation to increase training samples in deep neural networks [57], in order to balance the size of datasets [58]. These affine transformations also improve efficiency of the dataset [59]. Popular methods include histogram equalization, contrast and brightness enhancement, white balancing, sharpening, and blurring [60].

Synthetic image generation techniques such as Generative Adversarial Networks (GAN) are also used for data augmentation which are effective when used with small, significantly imbalanced datasets, but has limited impact for large datasets [61].

Other data augmentation approaches, such as image style transformation, generate high-quality images by blending the content of one image with the visual style of another. These newly created images can be used for pre-training a neural network, enhancing its overall training performance [62].

## 5. Deep Learning Approaches For RP Detection

### 5.1 Convolutional Neural Networks (CNNs)

CNNs are frequently used in the task of classification of images because of their ability to learn spatial hierarchies of features [63]. For the detection of RP, CNNs extract features related to retinal abnormalities and classify images as RP-positive or negative.

Some of the popular architectures include

ResNet: A deep residual network for capturing complex medical image patterns. One of the popular deep CNN architectures is ResNet, which played very important role in DL by effectively training the deep neural networks. Traditional deep networks often suffer from the vanishing gradient problem, where gradients become too small, leading to inefficient training and poor model performance. To address this issue, ResNet introduces skip connections, which allow certain layers to be bypassed. These connections enable the effective training of deeper neural networks by enhancing the flow of gradients. Though the ResNet enhances accuracy, it does so at the cost of higher memory consumption and greater computational complexity [64, 65].

**EfficientNet:** This architecture achieves a balance between computational efficiency and high accuracy. It is a CNN architecture designed for optimal performance in DL applications and in image processing. Introduced by Google in 2019, this architecture revolutionized model scaling by efficiently adjusting depth, width, and resolution using a compound scaling method. This approach enables high performance while reducing computational costs compared to traditional CNN architectures [66].

**VGGNet:** VGGNet (Visual Geometry Group) is a widely used convolutional neural network (CNN) architecture known for its effectiveness in image classification, feature extraction, and transfer learning. A key feature of VGGNet is its use of small 3×3 convolutional filters stacked together, which increases network depth without significantly impacting computational cost. Additionally, 2×2 max pooling layers help reduce spatial dimensions while enhancing feature extraction. The final layers consist of fully connected (FC) layers followed by a Softmax activation function for classification [67]. The 19 and 16-layer variants are called as VGG-19 and VGG-16 respectively. In particular, VGG-16, is having approximately 138 million parameters [67].

**Inception v3:** Inception v3 is another image classification DL model, which is the improved version of GoogLeNet(Inception v1) developed by Google. It efficiently scales deep networks by utilizing factorized convolutions, which break down large convolutional filters into smaller operations, reducing computational cost and memory usage. To enhance training stability and mitigate overfitting, the model incorporates regularization techniques such as batch normalization, label smoothing, and auxiliary classifiers. These optimizations contribute to its superior performance, allowing Inception v3 to achieve state-of-the-art results in the ILSVRC 2012 classification challenge, demonstrating its effectiveness in real-world applications [68].

## 5.2 Transfer Learning

Since medical datasets are often limited, transfer learning leverages pre-trained CNN models such as ResNet, Inception models and VGG trained on large image datasets and fine-tunes them on retinal images. This approach improves performance, especially when data availability is a challenge [42, 69, 70].

Inception-V3, VGG19, U-net, Xception and ResNet-50 are commonly used as base models due to their relatively small size, which ensures computational efficiency while maintaining strong classification performance [8, 41-43]. These architectures have shown promising results in handling complex image data, including retinal images, and have been successfully used in previous ophthalmology studies [71].

## 6. Dataset

The availability of both public and private datasets plays a very important role in advancing AI-driven research on retinitis pigmentosa (RP). However, there are very few datasets available, and only a limited number are publicly accessible for training deep learning models. Public datasets from Kaggle and contributions from various ophthalmology research institutions, provide retinal fundus images and optical coherence tomography (OCT) scans that enable model training and validation.

On the other hand, private datasets from hospitals and specialized research centers often contain higher-quality, well-annotated images. However, access to these datasets is restricted due to privacy concerns and ethical regulations. The combination of public and private datasets enhances the generalizability and robustness of deep learning models for RP detection and prognosis.

The challenge of diagnosing RP is further compounded by the limited availability of high-quality datasets, making it difficult for researchers to develop accurate and cost-effective diagnostic solutions. This issue is addressed by EDORP (Early Detection of Retinitis Pigmentosa) dataset of Kaggle which integrates three distinct medical imaging modalities—Fundus Autofluorescence (FAF), Ultra-Wide Field Pseudocolor (UWPC), and Colored Fundus Imaging. This multi-modal approach provides a more comprehensive view of the retina, significantly enhancing diagnostic accuracy. EDORP is a curated combination of several publicly available datasets, including RIPS, RFMD, Multiclass Retinal Disease, Masumoto FAF, and Masumoto UWPC. It consists of 932 images, with 427 labeled as RP and 505 as normal eyes [<https://www.kaggle.com/datasets/karmayasser/edorp-alltypes/data>].

Deep learning models are data hungry models. The prediction of a deep learning model is strongly influenced by the size and quality of the dataset used for training [24]. When working with small datasets, it can be difficult to allocate enough data for both training and validation without compromising the amount of information available for effective model learning [8]. Cross-validation addresses this challenge by rotating the roles of training and validation data across multiple iterations, ensuring that all samples are utilized for both purposes and minimizing data wastage [72].

## 7. Evaluation Metrics And Performance Assessment

In deep learning, particularly in classification tasks, TP, TN, FP, and FN are key performance metrics used to evaluate model performance. The above 4 terms are based on the confusion matrix, which evaluates predictions by comparing them to the actual labels.

**True Positive (TP)** – The DL model accurately predicts a positive case. **True Negative (TN)** – The model accurately predicts a negative case. **False Positive (FP) (Type I Error)** – The model wrongly predicts a positive case when it is negative. **False Negative (FN)(Type II Error)** – The model wrongly predicts a negative case when it is positive.

TP = Model correctly identifies an eye with RP.

TN = Model correctly detect a healthy eye.

FP = Model wrongly classifies a healthy eye as RP.

FN = Model fails to detect RP in an affected eye.

A high FN rate (missed cases) is dangerous in medical applications, as it means diseased individuals go undiagnosed. Hence, models are often optimized to maximize sensitivity (recall) to reduce FN cases. High recall is critical because missing an RP diagnosis (false negatives) could delay treatment and lead to disease progression. High recall may come at the cost of lower precision, meaning more false positives. Balancing recall and precision are crucial, often evaluated using the F1-score. AUC-ROC (Area Under Curve - Receiver Operating Characteristic) evaluates the overall diagnostic performance of the model. A higher AUC-ROC value signifies a strong DL model which is capable of effectively differentiating between RP and non-RP cases.

Table 1. Performance metrics[73, 74]

Performance Metric	Formula	Description
Accuracy	$= \frac{TP + TN}{(TP + TN + FP + FN)}$	Measures overall correctness of classification. (Number of correct classifications to the Total number of classification)
Sensitivity	$= \frac{TP}{(TP + FN)}$	Measures the capability to correctly detect RP-positive cases.
Specificity	$= \frac{TN}{(TN + FP)}$	Measures the capability to correctly identify non-RP cases.
Precision	$= \frac{TP}{(TP + FP)}$	Measures the proportion of correctly identified RP cases out of all cases the model labeled as RP.
Recall	$= \frac{TP}{(TP + FN)}$	Measures the model's ability to correctly identify patients who actually have RP
F1 score	$= \frac{2 \times Precision \times Recall}{Precision + Recall}$	F1-score, which is the harmonic mean of precision and recall, provides a balanced assessment of performance, especially when there is an imbalance between positive and negative cases in the dataset [8].

Table 2. State of the art methodologies and their performance measures in detection of RP using RIPS dataset

Methods	Accuracy	Sensitivity	Specificity	F Measure
Random Forest [46]	99.14	58.26	99.46	47.93
AdaBoost [46]	99.01	64.29	99.30	46.76
U-Net 48 × 48 [41]	99.00	55.70	99.40	50.60
U-Net 72 × 72 [41]	99.00	62.60	99.30	52.80
U-Net 96 × 96[41]	99.20	55.20	99.60	55.10
RPS-Net [39]	99.52	80.54	99.60	61.54
P3S-Net[75]	99.67	83.17	99.73	63.24
SE-ResNet [43]	99.16	97.70	96.93	90.47

## 8. Challenges And Limitations

Developing deep learning models for the detection of retinitis pigmentosa (RP) faces several critical challenges. A primary concern is the limited availability of large, labeled datasets containing RP-positive cases. This scarcity not only hampers the development of well-generalized models but also results in pronounced class imbalance, potentially skewing predictions in favor of the majority class. Another big challenge in DL systems is the inherent lack of interpretability. Often described as black boxes, these models can be difficult to trust in clinical environments due to their opaque decision-making processes [76]. To address this, the integration of explainable AI (XAI) techniques has become increasingly important. The trust among the clinicians can be increased with the help of XAI methods by developing more understandable and justifiable models which inturn improves model transparency [77, 78]. Furthermore, discrepancies in imaging modalities caused by differences in hardware, acquisition protocols, and settings introduce variability in image quality and appearance [52]. These inconsistencies can impair the model's ability to perform reliably across diverse clinical environments, limiting its practical applicability and generalization capability.

## 9. Future Scope

Combining multiple data modalities such as fundus imaging and Optical Coherence Tomography (OCT) can greatly improve the accuracy of retinitis pigmentosa (RP) diagnosis by providing a richer, more comprehensive view of both retinal structure and function. Simultaneously, the integration of explainable artificial intelligence techniques, like Grad-CAM based visualizations, enhances the interpretability of deep learning models, allowing clinicians to gain insight into how diagnostic decisions are made and fostering greater trust in AI systems [79]. Moreover, embedding these advanced AI models into practical clinical applications such as handheld diagnostic tools or telemedicine platforms—has the potential to broaden access to RP screening and care, particularly in underserved or remote communities.

## 10. Conclusion

Deep learning has emerged as a powerful and promising tool in the early diagnosis and identification of retinitis pigmentosa, offering the potential to automate and enhance traditional ophthalmic assessment methods. Through the use of advanced neural network architectures, particularly convolutional neural networks, researchers have demonstrated encouraging results in identifying RP from a variety of imaging modalities, including fundus photography, OCT and FAF images. These models have shown the ability to detect subtle pathological features that may be missed during routine clinical evaluation.

However, despite notable progress, several challenges remain. The limited availability of large, annotated datasets hinders the development of highly generalizable models, while the black-box nature of deep learning raises concerns about interpretability and clinical trust.

Moreover, variations in imaging quality, population diversity, and disease presentation further complicate model robustness across real-world settings. To address these issues, future research should prioritize the creation of diverse, high-quality datasets, explore techniques such as transfer learning and self-supervised learning to reduce reliance on labeled data, and develop explainable AI frameworks that enhance clinician confidence in model predictions. In conclusion, while deep learning for RP detection is still evolving, its integration into clinical practice holds significant promise. With continued interdisciplinary collaboration between clinicians, engineers, and data scientists, deep learning can play a transformative role in the early identification and management of retinitis pigmentosa.

## REFERENCES

- [1]. Hamel, C., Retinitis pigmentosa. *Orphanet journal of rare diseases*, 2006. 1: p. 1-12.
- [2]. Zhang, Q., Retinitis pigmentosa: progress and perspective. *Asia-Pacific journal of ophthalmology*, 2016. 5(4): p. 65-271.
- [3]. Cross, N., et al., Current and future treatment of retinitis pigmentosa. *Clinical Ophthalmology (Auckland, NZ)*, 2022. [4]. Cross, N., et al., Retinitis pigmentosa: burden of disease and current unmet needs. *Clinical Ophthalmology*, 2022:
- [5]. Liu, T.Y.A., et al., Prediction of visual impairment in retinitis pigmentosa using deep learning and multimodal fundus images. *British Journal of Ophthalmology*, 2023. 107(10):
- [6]. Mishra, C. and K. Tripathy, Fundus camera. 2022.
- [7]. Musleh, A.M., et al., Diagnostic accuracy of artificial intelligence in detecting retinitis pigmentosa: A systematic review and meta-analysis. *Survey of Ophthalmology*, 2024. 69(3): p. 411-417.
- [8]. Ferreira, H., et al., Retinitis Pigmentosa Classification with Deep Learning and Integrated Gradients Analysis. *Applied*
- [9]. Enaholo, E.S., M.J. Musa, and M. Zeppieri, Optical Coherence Tomography, in *StatPearls [Internet]*. 2024, StatPearls
- [10]. Kim, H.M., J.K. Oh, and S.H. Tsang, The Use of Optical Coherence Tomography in Evaluation of Retinitis Pigmentosa,
- [11]. Ran, A. and C.Y. Cheung, Deep learning-based optical coherence tomography and optical coherence tomography angiography image analysis: an updated summary. *Asia-Pacific Journal of Ophthalmology*, 2021. 10(3): p. 253-260.
- [12]. Samuel, A.L., Some studies in machine learning using the game of checkers. *IBM Journal of research and development*,
- [13]. Andrieu, C., N. De Freitas, A. Doucet, and M.I. Jordan, An introduction to MCMC for machine learning. [14]. Mitchell, T.M., *Machine learning and data mining*. Communications of the ACM, 1999. 42(11): p. 30-36.
- [15]. Voulodimos, A., N. Doulamis, A. Doulamis, and E. Protopapadakis, Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience*, 2018. 2018(1): p. 7068349.
- [16]. Cambria, E. and B. White, Jumping NLP curves: A review of natural language processing research. *IEEE Computational intelligence magazine*, 2014. 9(2): p. 48-57.
- [17]. Chaki, J., N. Dey, F. Shi, and R.S. Sherratt, Pattern mining approaches used in sensor-based biometric recognition: a review. *IEEE Sensors Journal*, 2019. 19(10): p. 3569-3580.
- [18]. Bakator, M. and D. Radosav, Deep learning and medical diagnosis: A review of literature. *Multimodal Technologies and Interaction*, 2018. 2(3): p. 47.
- [19]. Modi, K. and R. Dayma. Review on fraud detection methods in credit card transactions. in *2017 International Conference on Intelligent Computing and Control (I2C2)*. 2017. IEEE.
- [20]. Chong, E., C. Han, and F.C. Park, Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, 2017. 83: p. 187-205.
- [21]. Nassif, A.B., et al., Speech recognition using deep neural networks: A systematic review. *IEEE access*, 2019. 7: p.
- [22]. Pierson, H.A. and M.S. Gashler, Deep learning in robotics: a review of recent research. *Advanced Robotics*, 2017.
- [23]. Zhou, T., Z. Song, and K. Sundmacher, Big data creates new opportunities for materials research: a review on methods 7
- [24]. Whang, S.E., Y. Roh, H. Song, and J.-G. Lee, Data collection and quality challenges in deep learning: A data-centric ai perspective. *The VLDB Journal*, 2023. 32(4): p. 791-813.
- [25]. Wang, J., et al. Research on image recognition technology based on multimodal deep learning. in *2024 IEEE 2nd*

- International Conference on Image Processing and Computer Applications (ICIPCA). 2024. IEEE.
- [26]. Kheddar, H., M. Hemis, and Y. Himeur, Automatic speech recognition using advanced deep learning approaches: A comprehensive survey. *Journal of Imaging*, 2024. 10(1): p. 18.
- [27]. AlKendi, W., F. Gechter, L. Heyberger, and C. Guyeux, Advancements and challenges in handwritten text recognition: A comprehensive survey. *Journal of Imaging*, 2024. 10(1): p. 18.
- [29]. Kolasa, K., et al., Systematic reviews of machine learning in healthcare: a literature review. *Expert Review of Pharmacoeconomics & Outcomes Research*, 2024. 24(1): p. 63-115.
- [30]. Chakraborty, C., M. Bhattacharya, S. Pal, and S.-S. Lee, From machine learning to deep learning: Advances of the recent data-driven paradigm shift in medicine and healthcare. *Current Research in Biotechnology*, 2024. 7: p. 100164.
- [31]. Kanaparathi, V., Transformational application of Artificial Intelligence and Machine learning in Financial Technologies and Financial services: A bibliometric review. *arXiv preprint arXiv:2401.15710*, 2024.
- [32]. Paramesha, M., N.L. Rane, and J. Rane, Artificial intelligence, machine learning, deep learning, and blockchain in financial and banking services: A comprehensive review. *Partners Universal Multidisciplinary Research Journal*, 2024.
- [33]. Hassan, D.O. and B.A. Hassan, A comprehensive systematic review of machine learning in the retail industry: Unveiling consumer preferences: A two-stage deep learning approach to enhance accuracy in multi-channel retail sales forecasting. *Expert Systems with Applications*, 2024. 257: p. 125066.
- [34]. Wu, J., H. Liu, X. Yao, and L. Zhang, Unveiling consumer preferences: A two-stage deep learning approach to enhance accuracy in multi-channel retail sales forecasting. *Expert Systems with Applications*, 2024. 257: p. 125066.
- [35]. Zhao, X., et al., A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, 2024.
- [36]. Takahashi, S., et al., Comparison of vision transformers and convolutional neural networks in medical image analysis: a systematic review. *Journal of Medical Systems*, 2024. 48(1): p. 84.
- [37]. Ghelani, H., Enhancing PCB Quality Control through AI-Driven Inspection: Leveraging Convolutional Neural Networks for Automated Defect Detection in Electronic Manufacturing Environments. Available at SSRN 5160737, 2024.
- [38]. Kishor, S., et al. Lane detection for autonomous vehicles with canny edge detection and general filter convolutional neural network. in 2024 11th International Conference on Computing for Sustainable Global Development (INDIACom). 2024. IEEE.
- [39]. Arsalan, M., et al., Deep learning-based detection of pigment signs for analysis and diagnosis of retinitis pigmentosa. *Sensors*, 2020. 20(12): p. 3454.
- [40]. Shah, A., et al., Comparative analysis of median filter and its variants for removal of impulse noise from gray scale images. *Journal of King Saud University-Computer and Information Sciences*, 2022. 34(3): p. 505-519.
- [41]. Zhu, Y. and C. Huang, An improved median filtering algorithm for image noise reduction. *Physics Procedia*, 2012. 25: p. 505-519.
- [42]. Lin, P.-H., B.-H. Chen, F.-C. Cheng, and S.-C. Huang, A morphological mean filter for impulse noise removal. *Optical Engineering*, 2008. 47(3): p. 345-352.
- [43]. Jin, L., C. Xiong, and D. Li, Selective adaptive weighted median filter. *Optical Engineering*, 2008. 47(3): p. 345-352.
- [44]. Galdran, A., et al., Data-driven color augmentation techniques for deep skin image analysis. *arXiv preprint arXiv:2107.02970*, 2021.
- [45]. Sundaram, S. and N. Hulkund, Gan-based data augmentation for chest x-ray classification. *arXiv preprint arXiv:2107.02970*, 2021.
- [46]. Mikołajczyk, A. and M. Grochowski. Data augmentation for improving deep learning in image classification problem. in 2018 international interdisciplinary PhD workshop (IIPhDW). 2018. IEEE.
- [47]. Wang, W., et al., HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection. *IEEE access*, 2017. 6: p. 1792-1806.
- [48]. He, K., X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [49]. Xu, W., Y.-L. Fu, and D. Zhu, ResNet and its application to medical image processing: Research progress and challenges. *Computer Methods and Programs in Biomedicine*, 2023. 240: p. 107660.
- [50]. Tan, M. and Q. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. in *International conference on machine learning*. 2019. PMLR.
- [51]. Simonyan, K. and A. Zisserman, Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [52]. Szegedy, C., et al. Rethinking the inception architecture for computer vision. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [53]. Kora, P., et al., Transfer learning techniques for medical image analysis: A review. *Biocybernetics and biomedical engineering*, 2022. 42(1): p. 79-107.
- [54]. Kim, H.E., et al., Transfer learning for medical image classification: a literature review. *BMC medical imaging*, 2022. 22(1): p. 69.
- [55]. Raschka, S., Model evaluation, model selection, and algorithm selection in machine learning. *arXiv preprint arXiv:1811.12808*, 2018.
- [56]. Shekar, S., N. Satpute, and A. Gupta, Review on diabetic retinopathy with deep learning methods. *Journal of Medical Systems*, 2023. 48(1): p. 104263.
- [57]. Choudhary, A., et al. A deep learning-based framework for retinal disease classification. in *Healthcare*. 2023. MDPI.
- [58]. Imran, S.M.A., A. Hussain, N. Salem, and M. Arsalan, Artificially Intelligent detection of retinal pigment sign using P3S-Net for retinitis pigmentosa analysis. *Results in Engineering*, 2025: p. 104263.
- [59]. Van der Velden, B.H., H.J. Kuijff, K.G. Gilhuijs, and M.A. Viergeever, Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis*, 2022. 79: p. 102470.
- [60]. Paudel, D., et al., Interpretability of deep learning models for crop yield forecasting. *Computers and Electronics in Agriculture*, 2023. 206: p. 107663.
- [61]. Teng, Q., et al., A survey on the interpretability of deep learning in medical diagnosis. *Multimedia Systems*, 2022. 28(6): p. 2335-2355.
- [62]. Papandrianos, N.I., et al., An explainable classification method of SPECT myocardial perfusion images in nuclear cardiology using deep learning and grad-CAM. *Applied Sciences*, 2022. 12(15): p. 7592.