**AN INCEPTION-BASED DEEP LEARNING MODEL WITH SPATIAL ATTENTION FOR PAPILLEDEMA CLASSIFICATION**

Subject Area: Medical imaging, ophthalmology, and deep learning

**ABSTRACT**

Papilledema, characterized by optic disc swelling due to elevated intracranial pressure, is a critical neuro-ophthalmic condition that requires prompt detection to prevent vision loss. Traditional diagnostic methods, such as direct ophthalmoscopy, are highly dependent on examiner expertise and often suffer from limited clinical utilization. Deep learning has emerged as a promising solution to automate papilledema detection from retinal fundus images, improving diagnostic accuracy and accessibility. This study proposes a deep learning-based approach utilizing convolutional neural networks (CNNs) with spatial attention mechanisms to enhance feature extraction and classification performance. The model is trained on a balanced dataset of normal, papilledema, and pseudopapilledema fundus images, employing data augmentation techniques to improve generalization. Performance is evaluated using standard classification metrics, demonstrating superior accuracy compared to conventional feature-based machine learning models. By leveraging deep learning for automated papilledema detection, this research aims to assist clinicians in early diagnosis, reducing misdiagnosis rates and improving patient outcomes.

Keywords: Papilledema, Fundus Imaging, Deep Learning, Convolutional Neural Networks (CNN), Spatial Attention

**INTRODUCTION**

Papilledema, defined as optic disc swelling due to elevated intracranial pressure, represents a critical ophthalmic finding that warrants immediate clinical attention [1]. Papilledema can cause increasing visual impairment and possibly total vision loss if it is not identified or treated [1]. The condition may arise from various underlying pathologies including cerebral hemorrhage, spinal cord lesions, head trauma, hydrocephalus, meningitis, impaired cerebral sinus drainage, cranial anomalies, and idiopathic intracranial hypertension (IIH) [1].

Traditionally, papilledema diagnosis relies on direct ophthalmoscopy performed by trained specialists. However, this approach faces significant challenges in practical clinical settings. As noted by Mackay et al. [2], ocular funduscopy is considered to be a "dying art" with physicians demonstrating decreasing confidence in ophthalmoscope usage. The technical difficulty of performing ophthalmoscopy, combined with waning enthusiasm and sometimes even discouragement from medical education preceptors, has further contributed to the decline of this vital clinical skill [2]. Consequently, some clinicians regularly perform ophthalmoscopy, and many struggle to detect ocular fundus disorders [2]. The examination of the ocular fundus, while imperative for diagnosing various medical and neurologic conditions, is frequently underutilized and challenging to perform without pupillary dilation [3].

A further diagnostic challenge lies in distinguishing true papilledema from pseudopapilledema, which refers to optic disc elevation without peripapillary fluid [4]. Pseudopapilledema may result from congenital anomalies such as optic nerve head drusen, tilted disc, hyperopic discs, or myelinated nerve fibers. Misdiagnosis can lead to unnecessary procedures and potentially compromise patient care [4]. Therefore, developing reliable, accessible, and accurate methods for papilledema detection represents an important clinical need.

The emergence of deep learning technologies in medical imaging offers promising opportunities to address these challenges. Deep learning approaches have demonstrated remarkable capabilities in analyzing fundus images for various ophthalmic conditions [5]. These computational methodologies can potentially automate and standardize the detection of papilledema, thereby overcoming the limitations associated with traditional ophthalmoscopy and improving diagnostic accuracy.

This research aims to develop and validate a deep learning-based approach for automated detection of papilledema on fundus images. By leveraging advanced neural network architectures and image processing techniques, we seek to create a reliable tool that can assist clinicians in identifying this critical condition, potentially improving diagnostic efficiency and patient outcomes. The following sections present a comprehensive literature review on existing methodologies for papilledema detection, the application of deep learning in medical image analysis, and the specific challenges and opportunities in this domain.

**Traditional Methods for Papilledema Detection**

The conventional approach to papilledema detection has predominantly relied on direct ophthalmoscopy performed by specialists. However, as highlighted by Bruce et al. [3], this method suffers from significant limitations including dependence on examiner expertise, the requirement for pupillary dilation in many cases, and subjective interpretation of findings. These factors contribute to potential inconsistencies in diagnosis and barriers to widespread implementation, particularly in emergency settings where timely assessment is crucial.Recognizing these challenges, Bruce et al. [3] proposed nonmydriatic fundus photography as an alternative tool to direct ophthalmoscopy. This approach eliminates the need for pupillary dilation, making it more practical in various clinical settings, especially emergency departments. Their study hypothesized that implementing nonmydriatic fundus photography would enhance the detection of relevant ocular fundus abnormalities that might otherwise be missed during routine clinical practice. Despite these advantages, the approach still requires expert interpretation of the images, which can introduce subjectivity and variability in diagnosis.

Another significant challenge in papilledema diagnosis is differentiating it from pseudopapilledema. El-Gendy et al. [4] addressed this diagnostic dilemma, emphasizing that pseudopapilledema patients are occasionally misdiagnosed with papilledema, potentially leading to unnecessary procedures. Their research aimed to examine evidence on various methods for distinguishing between these conditions, highlighting the need for more objective diagnostic approaches. A key limitation identified in their work is the complexity of establishing definitive diagnostic criteria that can be consistently applied across diverse clinical presentations.

**Computational Approaches to Fundus Image Analysis**

Early computational approaches to fundus image analysis focused on traditional image processing techniques. A decision support system for the identification of papilledema using fundus retinal pictures was developed by Akbar et al. [6]. Their method extracted 23 features, including textural features, optic disc margin obscuration features, colour-based features, and vascular features. A Support Vector Machine (SVM) model with a Radial Basis Function (RBF) kernel was utilised for the classification process. While this approach demonstrated promising results, it relied heavily on handcrafted features, limiting its adaptability and generalizability to diverse clinical scenarios. Additionally, the feature extraction process required significant domain expertise and manual intervention, reducing its potential for fully automated implementation.

Medical image analysis has undergone a considerable transformation as a result of the development of deep learning. As outlined by Suzuki [7], the field witnessed a shift from feature-based machine learning, which was dominant before deep learning, to image-based machine learning where image data is learned directly without object segmentation or feature extraction. This paradigm shift has enabled more comprehensive and nuanced analysis of medical images, including fundus photographs. However, Suzuki also noted that the black-box nature of deep learning models presents challenges for clinical interpretation and validation, potentially limiting their acceptance in medical practice.

Fu et al. [8] utilised deep learning for vessel segmentation in fundus images, framing it as a boundary detection challenge with convolutional neural networks (CNNs). They used fully-connected Conditional Random Fields (CRFs) to combine the discriminative vessel probability map with long-range interactions between pixels. While not specifically targeting papilledema, this work demonstrated the potential of deep learning for analysing structural elements in fundus images. A limitation of their approach was the computational complexity introduced by the CRF post-processing step, which could impact real-time application feasibility.

Lee et al. [9] presented a spatio-temporal attention relocation (STARE) method for efficient detection of simultaneously occurring structured activities. Although developed in a different context, their information-theoretic approach for dynamic attention relocation offers valuable insights for medical image analysis, particularly for conditions like papilledema where specific regions of interest need to be prioritized. However, adapting such approaches to medical imaging contexts requires careful consideration of domain-specific constraints and validation requirements.

**Deep Learning in Medical Imaging**

The utilisation of deep learning in medical imaging increased significantly in recent years. Li et al. [5] provided a comprehensive review of deep learning applications in fundus image analysis, systematically categorizing various tasks such as lesion segmentation, biomarker segmentation, disease detection, and image synthesis. Their review highlighted both the advancements and challenges in the field, offering valuable insights for developing new methodologies. A significant limitation identified in their review was the lack of standardized evaluation frameworks, making direct comparisons between different approaches challenging.

Convolutional Neural Networks (CNNs) have emerged as particularly powerful tools for medical image analysis. Li et al. [10] surveyed CNN applications, analysing their architectures, functionalities, and prospects. They highlighted that CNNs have attained cutting-edge performance across multiple fields, including computer vision and natural language processing. making them promising candidates for medical imaging tasks. However, they also noted that CNNs typically require large amounts of labelled data, which can be a significant constraint in medical imaging applications where annotated datasets are often limited.

Implementing deep learning for medical imaging presents unique challenges. Kebaili et al. [11] addressed the issue of insufficient training data in medical imaging. They clarified that data collection is costly and must adhere to privacy regulations. They evaluated deep learning methods such as variational autoencoders, generative adversarial networks, and diffusion models, which can assist in generating additional medical images and addressing data scarcity issues. A limitation of these generative approaches is the potential introduction of artifacts or unrealistic features that may not represent actual pathological conditions, potentially leading to model biases.

Similarly, Hussain et al. [12] investigated differential data augmentation techniques for medical imaging classification tasks. Their research indicated that the degree to which an augmented training set preserves characteristics of the original medical images considerably affects model performance. This finding underscores the importance of appropriate data augmentation strategies for developing robust deep learning models for medical applications. They cautioned that inappropriate augmentation strategies could degrade model performance, highlighting the need for domain-specific considerations in data augmentation design.

Another significant consideration for medical imaging applications is the transfer learning approach. Xie and Richmond [13] investigated pre-training on greyscale ImageNet for the classification of medical images, addressing the mismatch between color images in standard pre-training datasets and single-channel medical images. Interestingly, when they modified the models for disease classification using chest X-ray pictures, they outperformed those pre-trained on colour ImageNet, indicating that colour might not be a crucial characteristic for natural image classification tasks. A limitation of their approach is the potential domain gap between natural images used for pre-training and medical images used for fine-tuning, which may limit the transferability of learned features.

**Attention Mechanisms in Deep Learning**

Attention mechanisms have emerged as powerful components of deep learning architectures. Niu et al. [14] reviewed attention mechanisms in deep learning, providing an integrated model appropriate for the majority of attention frameworks. They categorised current attention models based on four criteria: attention softness, input feature types, input representation, and output representation. This framework offers valuable insights for implementing attention mechanisms in medical imaging applications. However, they noted that the increased model complexity introduced by attention mechanisms could lead to higher computational requirements and potential overfitting on limited data.

Kevin et al. [15] applied attention mechanisms to breast cancer detection using CNN. They combined ResNet-50 with Squeeze-and-Excitation (SE) attention features for better recalibration of feature maps. Their results demonstrated that the inclusion of SE mechanisms improved the model's classification performance, highlighting the potential benefits of attention mechanisms for medical image analysis. A limitation of their approach is the increased computational complexity introduced by the attention mechanisms, which may impact deployment in resource-constrained environments.

**Papilledema Detection Using Deep Learning**

Recent research has begun to explore the application of deep learning specifically for papilledema detection. Saba et al. [16] suggested an automated approach utilising deep learning for the detection and grading of papilledema, employing U-Net and Dense-Net designs. Their approach consisted of two stages: first, localizing and cropping the optic disc and surrounding area for input to Dense-Net for classification; Secondly, the classified papilledema fundus picture is preprocessed using a Gabor filter and subsequently inputted into U-Net to provide a segmented vascular network, from which vessel discontinuity indices are computed for the grading of papilledema. While their approach showed promise, it involved multiple separate models and processing steps, potentially increasing computational complexity and reducing end-to-end learning capabilities. Additionally, the reliance on intermediate feature extraction steps may introduce bottlenecks that limit the full potential of deep learning's feature learning capabilities.

Salaheldin et al. [17] assessed AI-driven techniques for detecting papilledema in retinal fundus pictures, highlighting the challenges involved in diagnosing papilledema in neuro-ophthalmology. They applied novel models using CNNs and recurrent neural networks (RNNs), specifically developing a multi-paths CNN model and a cascaded model which integrates ResNet-50 and LSTM. While their models demonstrated exceptional performance, the approach required a meticulously curated dataset, which may limit generalizability to diverse clinical settings. Furthermore, the complexity of their models may present challenges for clinical deployment and real-time application.

**Evaluation Metrics in Medical Image Segmentation**

Developing reliable evaluation metrics is crucial for assessing the performance of medical image analysis models. Müller et al. [18] elucidated many methodologies for assessing the accuracy of medical image segmentation in both binary and multi-class scenarios. They discussed prevalent difficulties such as imbalanced class distribution and errors in assessment. To enhance research quality and uniformity, they proposed a standardized method for assessing medical image segmentation. A limitation of current evaluation approaches is the lack of consensus on which metrics best reflect clinical utility, potentially leading to optimization for metrics that do not translate to improved patient outcomes.

**MATERIALS AND METHODS**

**Dataset Specification**

This study utilizes the pseudopapilledema dataset curated by Kim (2018), which contains fundus images categorized into three classes: normal, papilledema, and pseudopapilledema [19]. The fundus images in the dataset are displayed in Figure 1. After filtering, the dataset comprises 623 normal fundus images, 236 papilledema images, and 236 pseudopapilledema images. The images exhibit varying resolutions with dimensions specific to the ophthalmologic imaging equipment used during acquisition. A subset analysis revealed an average resolution of approximately 240×240 pixels, which informed our subsequent preprocessing strategy.

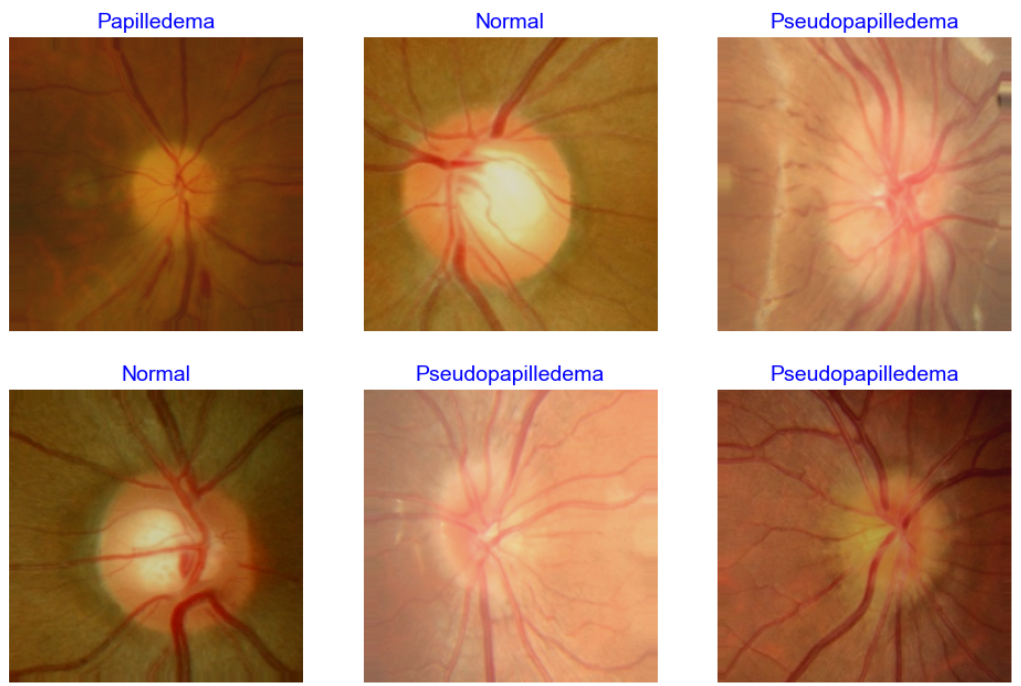


Figure 1: Fundus Images

**Data Preprocessing**

To address the inherent class imbalance in the dataset, we implemented an up-sampling technique where the minority classes (papilledema and pseudopapilledema) were resampled to match the majority class (normal), resulting in 623 images per class. This balanced distribution ensures that the model does not develop biases toward the overrepresented class during training. The dataset was then partitioned using stratified sampling, with 80% allocated for training and 20% for testing. The training set was further divided, with 80% used for model training and 20% for validation. This stratification preserves the class distribution across all subsets.

To enhance model robustness and generalization capabilities, we implemented comprehensive data augmentation techniques exclusively for the training subset. These transformations include:

* Random zoom variations (±10%)
* Brightness adjustments (80%-110%)
* Slight rotations (±1.5°)
* Shear transformations (up to 1.2)
* Channel shifts (up to 50)
* Horizontal and vertical flips

All images were standardized to 240×240 pixels and processed in batches of 30 during training to optimize computational efficiency.

**Model Architecture Design**

Our methodology employs a modified InceptionV3 architecture enhanced with a spatial attention mechanism. The InceptionV3 model, pre-trained on ImageNet, serves as the feature extraction backbone with its top classification layer removed. To balance transfer learning benefits with domain-specific adaptation, we froze all layers except the last 30, allowing these deeper layers to fine-tune to our specific ophthalmological imaging task.

The InceptionV3 architecture utilizes inception modules that perform convolutions at multiple scales simultaneously. The architecture of the InceptionV3 model is illustrated in Figure 2.

A diagram of a block diagram

AI-generated content may be incorrect.

Figure 2: Inception V3 Architecture

Mathematically, an inception module can be represented as:

(1)

Where,

* is the input feature map
* ​ represents convolution operation with filter size
* denotes max pooling operation
* represents the concatenation of feature maps along the channel dimension

This modular structure enables the network to efficiently extract features at various abstraction levels. The network processes 240×240×3 RGB input images through sequential inception modules, producing increasingly abstract feature representations.

On top of this backbone, we implemented a spatial attention mechanism as illustrated in Figure 3, that helps the model focus on the most relevant regions of the fundus images.

A diagram of a process

AI-generated content may be incorrect.

Figure 3: Spatial Attention Architecture

Mathematically, the spatial attention module can be represented as:

(2)

(3)

where,

* represents the feature maps from the InceptionV3 backbone
* ​ denotes a 2D convolution operation with a kernel
* is the sigmoid activation function
* represents element-wise multiplication
* is the attention map
* is the attention-weighted feature map

Following attention-enhanced feature extraction, global average pooling is applied to reduce spatial dimensions while preserving channel information. The resulting features are passed through two fully connected layers with 128 and 64 neurons, each followed by dropout layers (rate = 0.5) to mitigate overfitting.

**Classification**

The proposed approach is designed for multi-class classification of ophthalmological fundus images for papilledema assessment. The input images are resized to 240×240 pixels and normalized to maintain consistent pixel value ranges across the dataset. Each image is represented as a three-dimensional input tensor comprising height, width, and RGB channel dimensions, enabling the model to extract diverse visual characteristics from the fundus photographs.

The model outputs a three-dimensional probability vector through a dense layer employing the softmax activation function. These values represent the probabilities of the input image belonging to the normal, pseudopapilledema, or papilledema classes, with the sum of probabilities equaling 1. Classification is determined by selecting the class with the highest probability value. This multi-class approach allows the model to differentiate between the subtle visual features that distinguish true papilledema from pseudopapilledema and normal optic discs.

**E. System Pipeline**

The complete workflow integrates all components into an automated papilledema detection and classification system:

1. **Dataset Organization**: Images are structured in class-specific directories (normal, papilledema, pseudopapilledema).
2. **Data Balancing**: Class imbalance is addressed via statistical resampling.
3. **Stratified Partitioning**: The dataset is divided into training, validation, and testing subsets while maintaining class distributions.
4. **Image Preprocessing and Augmentation**: Training images undergo resizing, normalization, and augmentation, while validation and test images are only resized and normalized.
5. **Feature Extraction**: The InceptionV3 model processes input images to extract hierarchical features.
6. **Attention Mechanism**: The spatial attention module refines feature representation.
7. **Classification**: The fully connected layers and softmax activation function generate class probabilities.
8. **Evaluation and Interpretation**: The model is assessed using accuracy, precision, recall, and F1-score metrics.

This comprehensive pipeline enables accurate differentiation between normal optic discs, pseudopapilledema, and true papilledema, potentially assisting ophthalmologists in clinical diagnosis and reducing unnecessary medical procedures.

**RESULTS AND DISCUSSION**

The proposed deep learning approach demonstrated significant improvements over the base InceptionV3 model. The baseline InceptionV3 model achieved an accuracy of 78.10% and precision of 76.20%. By implementing the spatial attention mechanism and fine-tuning the model, we substantially enhanced the performance metrics, reaching an accuracy of 94.12% and precision of 94.19%.

**Performance Metrics**

The confusion matrix given in Figure 4, provides a detailed breakdown of the model's classification performance across three diagnostic categories: Normal, Papilledema, and Pseudopapilledema. The matrix demonstrates the model's high classification accuracy, with minimal misclassifications between classes. The diagonal elements (117, 117, 118) represent correctly classified instances, showcasing the model's robust performance in distinguishing between normal fundus images, papilledema, and pseudopapilledema.

A blue and white squares with white text

Description automatically generated

Figure 4: Confusion Matrix

The most notable misclassifications are:

* 7 Normal images classified as Pseudopapilledema
* 4 Papilledema images classified as Normal
* 4 Papilledema images classified as Pseudopapilledema

These minor misclassifications suggest the subtle visual similarities between these diagnostic categories, highlighting the complexity of differentiating between them.

**Class-wise performance**

The per-class performance metrics reveal the model's nuanced diagnostic capabilities:

* High recall values (0.93-0.94) indicate the model's effectiveness in correctly identifying most instances within each class.
* Specificity values (0.96-0.99) demonstrate the model's ability to accurately reject instances not belonging to a specific class.
* The consistent performance across normal, papilledema, and pseudopapilledema classes highlights the model's robust generalization.

**Statistical validation**

The macro-averaged evaluation metrics offer a robust statistical validation of the model's performance. ROC AUC of 0.9922 indicates exceptional discriminative power across all classes. Precision-Recall AUC of 0.9859 demonstrates consistent high-precision predictions across different recall levels. The curves are described in Figure 5. The aggregate performance metrics provide a comprehensive evaluation of the model's classification performance:

* Macro-Averaged ROC AUC: 0.9922
* Macro-Averaged Precision-Recall AUC: 0.9859

The high AUC values indicate the better ability of the model to distinguish between the three classes at various classification thresholds.

A screenshot of a graph

AI-generated content may be incorrect.

Figure 5: ROC curve and Precision-Recall Curve

**Overall performance**

The performance metrics overall provide a comprehensive evaluation of the model's classification capability. The comparison of the metrics with the base model without spatial attention is given in Table I.

* Precision of 0.9419 indicates high accuracy in positive predictions.
* Recall of 0.9412 demonstrates the model's ability to capture most relevant instances.
* The F1-score of 0.9413 represents an optimal balance between precision and recall.
* Overall specificity of 0.9706 underscores the model's exceptional ability to identify negative cases.

Table I: Metrics Comparison Table

| Approach | Metrics | | | |
| --- | --- | --- | --- | --- |
| Accuracy | Precision | Recall | F1 Score |
| Base Model | 78.10% | 76.20% | 82.92% | 77.16% |
| InceptionV3 + Spatial Attention | 87.59% | 83.81% | 88.48% | 85.56 |
| InceptionV3 + Spatial Attention (Fine Tuned) | 94.12% | 94.19% | 94.12% | 94.13 |

The substantial performance improvement from 78.10% to 94.12% can be attributed to the spatial attention mechanism. By dynamically highlighting critical regions in fundus images, the attention module enhances feature extraction, reduces noise, and improves differentiation between normal, pseudopapilledema, and papilledema cases, capturing subtle morphological differences essential for accurate diagnosis

**CONCLUSION**

The high specificity and sensitivity across all classes suggest that this deep learning approach could serve as a valuable diagnostic support tool for ophthalmologists. The model demonstrates remarkable potential in the early detection of papilledema, differentiating between true papilledema and pseudopapilledema, and providing an objective, consistent screening method.

While the findings of this study are promising, further validation using larger and more diverse datasets is essential to ensure the model's robustness and generalizability. Future research could focus on incorporating additional imaging modalities to enhance diagnostic accuracy, investigating model performance across different patient demographics to assess its reliability and fairness, and developing real-time clinical decision support systems for seamless integration into healthcare practice.

**REFERENCES**

1. Rigi M, Almarzouqi SJ, Morgan ML, Lee AG. Papilledema: epidemiology, etiology, and clinical management. Eye Brain. 2015;7:47-57.
2. Mackay DD, Garza PS, Bruce BB, Newman NJ, Biousse V. The demise of direct ophthalmoscopy: A modern clinical challenge. Neurol Clin Pract. 2015;5(2):150-157.
3. Bruce BB, Lamirel C, Wright DW, Ward A, Heilpern KL, Biousse V, et al. Nonmydriatic ocular fundus photography in the emergency department. N Engl J Med. 2011;364(4):387-389.
4. El-Gendy RS, El-Hamid ASA, Galhom AESA, et al. Diagnostic dilemma of papilledema and pseudopapilledema. Int Ophthalmol. 2024;44:272.
5. Li T, Bo W, Hu C, Kang H, Liu H, Wang K, et al. Applications of Deep Learning in Fundus Images: A Review. Med Image Anal. 2021;69.
6. Akbar S, Akram MU, Sharif M, et al. Decision Support System for Detection of Papilledema through Fundus Retinal Images. J Med Syst. 2017;41:66.
7. Suzuki K. Overview of deep learning in medical imaging. Radiol Phys Technol. 2017;10:257-273.
8. Fu H, Xu Y, Wong DWK, Liu J. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI); 2016 Apr; Prague, Czech Republic. p. 698-701.
9. Lee K, Ognibene D, Chang HJ, Kim TK, Demiris Y. STARE: Spatio-Temporal Attention Relocation for Multiple Structured Activities Detection. IEEE Trans Image Process. 2015;24(12):5916-5927.
10. Li Z, Liu F, Yang W, Peng S, Zhou J. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. IEEE Trans Neural Netw Learn Syst. 2022;33(12):6999-7019.
11. Kebaili A, Lapuyade-Lahorgue J, Ruan S. Deep Learning Approaches for Data Augmentation in Medical Imaging: A Review. J Imaging. 2023;9(4):81.
12. Hussain Z, Gimenez F, Yi D, Rubin D. Differential Data Augmentation Techniques for Medical Imaging Classification Tasks. AMIA Annu Symp Proc. 2018:979-984.
13. Xie Y, Richmond D. Pre-training on Grayscale ImageNet Improves Medical Image Classification. In: Proceedings of the European Conference on Computer Vision (ECCV) Workshops; 2018. p. 0-0.
14. Niu Z, Zhong G, Yu H. A review on the attention mechanism of deep learning. Neurocomputing. 2021;452:48-62.
15. Kevin PJ, Asher ER, Bezaleel TP, Prabakeran S. Breast Cancer Detection Using CNN and Attention Mechanism. In: 2024 4th International Conference on Mobile Networks and Wireless Communications (ICMNWC); 2024; Tumkuru, India. p. 1-7.
16. Saba T, Akbar S, Kolivand H, Bahaj SA. Automatic detection of papilledema through fundus retinal images using deep learning. Microsc Res Tech. 2021;84(12):3066-3077.
17. Salaheldin AM, Wahed MA, Talaat M, Saleh N. An evaluation of AI-based methods for papilledema detection in retinal fundus images. Biomed Signal Process Control. 2024;92:106120.
18. Müller D, Soto-Rey I, Kramer F. Towards a guideline for evaluation metrics in medical image segmentation. BMC Res Notes. 2022;15:210.
19. Kim U. Machine learning for Pseudopapilledema [Internet]. 2018 Aug 1 [cited 2025 Apr 4]. Available from: https://doi.org/10.17605/OSF.IO/2W5CE.