

Using Convolutional Neural Networks (CNN) for Medical Image Classification

Name: Areej Abdulaziz Alotaibi

Abstract

This paper discusses the way to create a more effective Convolutional Neural Network (CNN) library that will be able to enhance the medical image classification. CNNs have demonstrated exceptional capabilities to learn spatial hierarchies and subliminal information within medical images as digital imaging becomes an increasingly significant tool of clinical diagnosis. However, their effective use in clinics is hindered by the problem of small datasets, poor generalization, and high computing costs. Based on the experience of the hybrid feature extraction and ensemble learning methods, the proposed research presents a better classification model by combining CNN feature with hand-designed descriptors, including Local Binary Patterns (LBP), and using ensemble mechanisms, including stacking and prediction averaging. The study under review is an example of experimental quantitative research that includes the selection of datasets, preprocessing, the creation of the baseline model, the creation of the hybrid model, and the evaluation of its performance. The suggested model will reach more accuracy, better robustness in a variety of datasets and will be more efficient in computation as opposed to standard CNNs. The paper will serve to put forth a more trustworthy and flexible classification pipeline to be utilized in real-life medical imaging and offer evidence-based suggestions on how the model could be optimized in the future.

Key Words

Convolutional Neural Networks; Medical Image Classification; Hybrid Feature Extraction; Ensemble Learning

Chapter 1 Introduction

Introduction

One of the most transformative uses of artificial intelligence (AI) in the contemporary healthcare is medical image classification. Following the massive advancement of digital imaging technology including CT, MRI, X-ray, PET, dermoscopy and histopathology scanners, the need to have automated and dependable diagnostic support mechanism has become more urgent. A deep learning architecture called Convolutional Neural Networks (CNNs) is used to detect spatial hierarchies of visual patterns and has become a popular solution to the analysis of medical imaging as it offers strong feature extractors and offers an impressive accuracy in classification (Shrivastava et al., 2022). Over the past years, CNN-based systems have attained performance similar to or even better than the human experts in some diagnostic tasks, which points to their potential to minimize diagnostic errors, speed up clinical processes, and increase early detection of diseases.

Despite their success, CNNs still face several challenges that do not allow their full incorporation into clinical settings. Medical imaging is usually complicated, heterogeneous, and scarce in terms of privacy and the expensiveness of data acquisition. Consequently, CNNs can be prone to overfitting, poor generalization, or susceptible to noises and differences between imaging devices. Additionally, the typical CNN models even though effective might be insufficient to encode the rich textual data and local variations inherent in medical images, particularly when the dataset is small or imbalanced. This has prompted the need to explore hybrid, optimized and ensemble-based CNN designs that will incorporate two or more feature extraction methods so as to improve on robustness and classification. An example is that using deep CNN features together with handcrafted features, e.g., Local Binary Patterns (LBP), has helped to achieve significant improvement in more particular areas of medical imaging (Rocha et al., 2023).

The other method under development is ensemble learning methods such as stacking, bagging and prediction augmentation, to stack multiple CNN architecture and decrease model variance. Ensemble learning has been demonstrated to provide more robust and better predictions, especially in the complex medical imaging setting where only a single CNN model might not identify all of the discriminative patterns. Research has also established that when a series of CNN models are

stacked there can be significant improvements in F1-score, precision, and robustness, making ensemble learning a major future trend in medical image classification pipeline (Müller et al., 2023).

With this in mind, research is still ongoing to enhance CNN-based classification models, which are accurate and generalizable, computationally efficient, as well as explainable. The research proposal will make a contribution towards this emerging field by exploring and devising an optimized CNN-based medical image classification model which incorporates hybrid feature extraction, ensemble learning and architecture optimization. Finally, it is aimed at improving the accuracy of the diagnosis and overcoming the constraints that still do not allow CNNs to become a commonly used tool in clinical practice.

Background

The classification of medical images is a central aspect in clinical decision-making, allowing to identify any disease at an early stage, stage it, plan its treatment, and monitor the health of patients. Conventional image classification was based on hand-generated features, which were either texture descriptors, edge detectors or shape-based features, which had to be tuned to domain knowledge and were often manual. These techniques frequently lacked capabilities to deal with the large variation in medical images and could not learn more semantic patterns. In 2012, with the success of AlexNet in the ImageNet competition, deep learning came into a revolution, introducing CNNs as automated feature extractors and capable of learning hierarchical patterns directly on the data (Shrivastava et al., 2022).

The CNNs have proven to be very successful in many different biomedical imaging tasks such as tumor detection, organ segmentation, diabetic retinopathy classification, pneumonia detection, breast cancer screening, and histopathological diagnosis. Their capacity to form intricate spatial representations renders them especially appropriate in the analysis of delicate patterns and subtle abnormalities of medical images. CNN models, however, have their weaknesses, although they are good in their own capacity. The size of medical datasets is usually limited as opposed to natural image datasets, and CNNs need huge volumes of data to generalize successfully. Moreover, class imbalance, image noise, scanner variation, and small labels are often a problem to be encountered in real world performance.

Due to these difficulties, the hybrid CNN architectures have been introduced to combine deep features and handcrafted features. One important instance is the Hybrid CNN Ensemble (HCNNE) suggested by Rocha et al. (2023) that involves integrating CNN-derived features with LBP texture descriptors. In their approach, weights are allocated according to the confidence of the features and a multilayer perceptron and a combination of classifiers are used to give the final predictions. The approach led to a 10 % higher accuracy of odontogenic cyst histology classification, which proves that hybrid methods can overcome the limitations of pure CNN models with the low cost of the computation.

Similar to hybrid feature engineering, ensemble learning has started to emerge as a technique to enhance CNN performance in medical imaging. Ensemble methods combine several CNN predictions, which use their variations to lessen variance, prevent overfitting, and enhance generalization. Muller et al. (2023) did an extensive analysis of stacking, bagging, and augmentation-based ensembles with 9 CNN architectures on four medical imaging datasets. Their findings revealed that stacking has a F1-score increment of up to 13% which is better than any other ensemble method. Interestingly, they also found that simple statistical pooling functions, like averaging, frequently performed as well as more complex learning-based pooling models which supports the usefulness of simple and reliable ensemble in medical practice.

The use of ensemble techniques and fusion of features from both groups is a significant advancement in the architecture of medical image classification systems. Increasingly modern CNN-based systems are focusing on robustness, interpretability and adaptability as opposed to accuracy per se. In clinical use, intermodality reliability, noise resistance, and reduced computation needs are important factors of use. The biomedical imaging has already changed as Shrivastava et al. (2022) point out that the shift in the field was moved not only to deep learning but also to the handcrafted machine learning. However, to reach clinically feasible solutions, current limitations should be surmounted by means of architectural inventions, optimized pooling techniques, and multi-model integration.

Research Problem

The role of medical image classification in contemporary healthcare is critical since it allows making an early diagnosis and facilitating clinical decision-making. Despite the fact that Convolutional Neural Networks (CNNs) have made tremendous breakthrough in this field through

better image parting, feature incorporation and enhanced image classification, the system still faces a number of limitations that negatively impact the system in practical use. Available literature indicates that CNNs are capable of effectively learning spatial hierarchies and intricate patterns of medical images, but most of the studies lack empirical support, work with small datasets, or do not focus on issues regarding generalization and interpretability (Kshatri and Singh, 2023). Also, a majority of CNN-based models do not perform well when used in other imaging modalities or with heterogeneous sources of data, and their clinical effectiveness is diminished (Sarvamangala and Kulkarni, 2021).

Hybrids and ensemble systems, e.g., CNN-LBP feature fusion or stacking-based systems, have been shown to improve accuracy and robustness and obtain significant gains in specialized datasets (Rocha et al., 2023; Müller et al., 2024). The methods are however usually tested on small datasets and it is hard to make generalization on the large medical setting. Computational complexity is also a limiting factor, especially in situations that require CNNs in a healthcare setting characterized by time sensitivity (Nirthika et al., 2022). Besides, the majority of models do not incorporate optimization pooling strategy and domain adaptation mechanisms and therefore lack a lot of adaptability.

It is therefore clear that studies are needed to devise the best CNN-based classification techniques that integrate hybrid feature extraction, learning via ensembles and superior architecture design. It is possible to close such gaps to enhance the accuracy of diagnosis and strength of the models and enable the practical application of CNN models in diverse medical imaging scenarios (Al-Zubaidi and Mijwil, 2021).

Research Objectives

Primary Goal:

To build the optimal CNN-based medical image classification network that enhances the diagnosis accuracy, strength, and performance.

Specific Objectives:

- To examine the shortcomings of the currently employed CNN based methods of medical image classification in terms of accuracy, generalizability, and interpretability.

- To investigate how optimizations of architecture like hybrid feature extraction and improved pooling can affect CNN performance.
- To assess ensemble learning strategies (e.g. stacking, bagging) to enhance reliability of CNN classification in separate medical imaging data.
- To design and propose the most efficient CNN-based classification model and incorporates the efficient hybrid, pooling, and ensemble methods.
- To prove the proposed model on the basis of several medical image datasets and to compare its results with the existing state-of-the-art CNN-based solutions.

Research Scope

This research will aim at the design, development and testing of an optimized medical image classification Convolutional Neural Network (CNN)-based model to overcome the shortcomings noted by previous literature with respect to the generalizability, feature extraction and the model robustness. The proposed research is focused on the combination of hybrid feature extraction methods (corresponding to deep CNN) and ensemble learning methods to achieve a better classification result and a more stable performance across the existing medical imaging datasets. This area is based on the evidence provided by Rocha et al. (2023), who have shown that hybrid models can greatly enhance the level of texture detection, and Müller et al. (2023), whose article indicates the possibility of significant performance improvements with the implementation of ensemble models.

The study will be restricted to actor-supervised learning to identify images on classification tasks that use the publicly accessible medical image datasets including CT, MRI, X-ray, or histopathology collections of images. These datasets are going to be chosen on the basis of class variety and suitability to modern diagnostic issues. The research is not going to be conducted as a clinical trial or intervention of a real patient; only computational experiments and the evaluation of algorithms will be involved in the study. The proposed model will be evaluated according to its performance by comparison with state-of-the-art CNN architecture and existing reported hybrid and ensemble techniques. The accuracy, precision, recall, F1-score, and the efficiency of the computations will be considered as evaluation metrics, which are based on the findings reported by Shrivastava et al. (2022).

There is also the scope of exploring the optimized pooling strategies and architectural changes to enhance feature representation. The study, however does not consider the development of transformer based, 3D CNN models or multimodal learning systems since they require a wider scope of computational complexity than the resources and time of the project could have. Also, the ethical aspects like the preservation of privacy, explainability of the models, and federated learning are recognized but are not the primary focus of the experiment.

Chapter 2 Literature review

Literature Review

Kshatri & Singh (2023) gives a literature on convolutional neural networks (CNNs) used in medical image processing, specifically in preprocessing, segmentation, and diagnosis of MRI. The authors underline the fact that CNNs with the help of convolution and pooling and backpropagation layers successfully learn the spatial hierarchy of medical images and make efficient classification and feature retrieval. Their review also classifies CNN-based research by their usage as pattern recognition and classifies medical images processing with a taxonomy. The review concludes that CNNs have made new breakthroughs in the field of medical imaging because segmentation and noise reduction processes have been automated, which significantly enhances diagnostic accuracy. Nevertheless, even though the paper is written in a highly technical way, it is not quantitatively evaluated and experimentally attested. The second weakness that limits the generalization of the findings is the lack of analysis of the transfer learning and domain adaptation methods.

Rocha et al (2023). suggests a hybrid CNN-based model named as Hybrid CNN Ensemble (HCNNE) to classify medical images. This model combines deep CNN features with Local Binary Pattern (LBP) features which are handcrafted in order to enhance medical image texture based discrimination. The features extracted are trained with multilayer perceptrons and ensemble classifiers yielding 10% accuracy improvement compared to the earlier models of the state-of-the-art modeling odontogenic cyst histological images. The study shows that the high-level and low-level features used together are beneficial to increase the performance of classification with low cost in terms of computation. However, the primary weakness of the study is that it was only validated on one dataset and therefore it cannot be applied widely to other medical imaging fields. Its strength, however, is that hybrid feature fusion is effective in balancing accuracy and efficiency in CNN-based classification systems.

Muller et al (2024). discusses how ensemble learning techniques may be used to achieve optimal CNN performance in medical image classification pipelines. This paper explores various ensemble methods, such as stacking, bagging, and augmenting, and it was used with nine CNNs and four medical imaging datasets of different complexity. By empirically assessing the methods, the

authors show that the stacking method provides the best improvement and the F1-score can be improved by up to 13% and then bagging and augmenting, which are more effective at generalization and robustness. Surprisingly, the study also concludes that simple statistical pooling functionality tends to be more efficient than the complex learning-based pooling functionality, such as a support vector machine. The advantages of this work are the thoroughness of the experimental design and reproducibility, which provide viable insights on CNN ensemble optimization. Nevertheless, it does not discuss the issue of computational trade-offs and interpretability, which is vital in clinical acceptance.

Nirthika et al (2022). provide a comprehensive survey and empirical results of the pooling methods in CNNs in terms of their effects on medical image classification problems. The authors survey several pooling mechanisms such as max, average, stochastic and spatial pyramid pooling and provide a comparison experiment of HEP-2 cell and diabetic retinopathy data. Their results show that pooling plays an important role in CNN performance because the type of pooling to use should be related to the magnitude and distribution of class-specific features relative to image size. The experiment also shows that the application of optimized pooling strategies based on the ResNet architecture can bring new state-of-the-art results. Although the paper presents excellent theoretical foundations and experimental support, its weakness is that it has limited number of datasets, which limits the generalizability. However, it gives useful recommendations in the choice of pooling activities in future CNN-based medical applications.

Sarvamangala & Kulkarni (2021) provides a comprehensive review of how CNNs can be used to understand medical image, which includes classification, segmentation, localization, and detection of different types of diseases involving the brain, breast, and the lungs. The authors claim that CNNs outperform the conventional methods of diagnosis, and even humans, in a number of recognition tasks because they are able to learn complex spatial features automatically. The paper includes a brief description of CNN architectures and their award-winning structures, which also furnishes evidence that CNNs enhance diagnostic performance and lower the workload of clinical personnel. Nevertheless, issues like interpretability, data imbalance, and computational costs are also revealed by it, and they are still not resolved in clinical practice. Although conceptually deep and all-encompassing, the research lacks the latest developments such as attention-based networks or transformers, which restricts its ability to capture the latest trends in CNN.

Al-Zubaidi & Mijwil (2021) use a practical CNN model to classify COVID-19 medical images with the help of the GoogleNet architecture. The model was trained using 425 CT-coronavirus images to help in early detection of an infection. The CNN had a validation accuracy of 82.14% using a total training time of about 74 minutes which shows the possibility of the model to be used in automated pandemic diagnostics. The study also emphasizes the fact that CNNs can exploit features that are directly extracted out of medical data and classify them without much preprocessing. The main merit of the work is the fact that it can be applied in real life and that CNN is efficient in a time-sensitive medical environment. Nevertheless, this makes the dataset rather small and presents a risk of overfitting, and the absence of an external validation restricts its clinical applicability. Even with these constraints, the study has a good evidence of the effectiveness of CNNs in the rapid detection of diseases in emergency situations.

Kim et al. (2022) present a thorough transfer learning review on medical image classification. Their analysis of 121 qualified studies proves that transfer learning is extremely powerful to address a lack of data which is one of the significant challenges in medical imaging. Inception and ResNet are deep models that are common with feature extractors, and they save greatly on training time and computation costs without compromising accuracy. The review also discusses various transfer learning setup (feature extractors, feature extractors, fine-tuning, hybrid feature extractors) and points out inconsistencies in the way researchers are choosing models, initial pre-trained weights, and fine-tuning approaches. The authors suggest that systematic benchmarking is the way to enhance the reproducibility and provide optimal model configuration.

Ajagbe et al. (2021), created a deep convolutional neural network (DCNN) model based on multi-classification of the stage of Alzheimer disease using MRI scans. They combine CNN structures with transfer learning models, including VGG16 and VGG19. The research uses various metrics of evaluation, such as F1-score, AUC, recall, and precision. Findings indicate that VGG19 was superior to other models, indicating that pretrained deep architectures are useful in medical imaging processes. Nonetheless, the analysis is confined to the data of MRI of Alzheimer disease and does not compare the generalizability of the models to other imaging applications.

Aljuaid et al. (2022) also introduce a computer-aided diagnosis model to classify breast cancer by a transfer learning artificial neural network (Deep neural networks ResNet18, ShuffleNet, and Inception-V3). Their approach had a very high precision up to 99.7 percent on binary classification

and 97.81 percent on multi-class classification, which shows the effectiveness of deep CNNs in conjunction with transfer learning with large datasets as BreakHis. The strength of the study is based on the multi-network comparison and strong performance. It however, is only specific to breast cancer histopathology and may not be directly applicable to other types of medical imaging.

Lastly, Takahashi et al. (2024) authoritatively compare Vision Transformers (ViTs) and CNNs. Their analysis of 36 articles suggests that ViTs can tend to be better than CNNs in the activity with complex medical data because they possess better global attention mechanisms and can scale. They highlight the fact that ViTs need pretraining and mention that ViTs can outperform CNNs in feature extraction when large datasets are present. However, ViTs have large computational requirements and need large data, whereas CNNs are practical with small medical datasets.

Taken together, these six papers indicate the increased maturity of CNN research in medical image classification. They show a steady tendency to combine hybrid architecture, ensemble learning and optimization methods to achieve better diagnostic precision. Nevertheless, general difficulties remain throughout the literature, such as the lack of data diversity, interpretability, and computational requirements, implying the possibility of future research.

Synthesis of Findings

In all the explored researchers, CNNs are repeatedly determined as an effective method of classifying medical images automatically. Kshatri & Singh (2023) and Sarvamangala & Kulkarni (2021) presented the background theoretical viewpoints regarding CNNs are repeatedly identified as a useful way of automatically classify medical images in all the researched researchers. The background theoretical perspectives on CNN architecture and applications are introduced by Kshatri and Singh (2023) and Sarvamangala and Kulkarni (2021), and Rocha et al. (2023) and Muller et al. (2024) are the ones that provide the initial break and take the hybrid and ensemble developments into account. Meanwhile, Nirthika et al. (2022) bring certain architectural concepts in the picture, and Al-Zubaidi and Mijwil (2021) demonstrate CNNs in an actual healthcare setting. The papers considered collectively reveal the strength of CNNs in terms of feature selection and pattern recognition but exhibit a difference in their priorities: some of them are more concerned with accuracy and robustness through complex architectures, others are more inclined to the face of the value of computational efficiency and flexibility. The approaches can be different, but the requirement of CNNs in the development of diagnostic imaging is unanimous, and the issues of

explainability, cross-domain generalization, and real-time deployment should become the research needs of the future.

The literature reviewed proves the outstanding performance of CNNs in medical image classification both in theory and practical clinical environment implementation. Such innovations as hybridization and ensemble learning increased the robustness and accuracy of the models considerably. Nevertheless, there still remain challenges such as incomplete cross-domain adaptability, lack of explainability and data scarcity that prevent full clinical deployment. The way forward must focus on explainable AI, federated learning to maintain data privacy, and the integration of multimodal CNN to make sure that medical image classification systems are both diagnostic and ethically transparent.

Chapter 3 Research methodology

Research Methodology

The proposed research follows an experimental and quantitative approach that should be used to design, test, and optimize a Convolutional Neural Network (CNN)-based medical image classifier model. The methodology is planned to be organized in five stages, namely, dataset selection, preprocessing, model development, training and evaluation, and comparative analysis. These phases will make sure that the research goals (associated with the enhancement of accuracy, robustness, and computational efficiency) are achieved in a well-organized manner.

Dataset Selection

In order to have reproducibility and reliable evaluation, the study will utilize publicly available medical image datasets, including:

- BreakHis - histopathology images of breast cancer (benign vs malignant).
- Chest X-Ray (Kaggle: Pneumonia Dataset) pneumonia vs normal.
- Alzheimer MRI Dataset (Kaggle or ADNI subset) – multi-stage classification.

A single dataset will be chosen to make use of its availability, classes, balance, and appropriateness to train CNN. The exclusion of ethical limitations by utilizing publicly available datasets and the possibility to compare the outcomes with previous work is standardized.

Data Preprocessing

To obtain constant input into the CNN model, the following preprocessing mechanisms will be used:

- Scaling of images to a given size (e.g., 224x224).
- Normalize pixel values of 0 to 1.

- Enhancement of data diversity (data rotation, flipping, zooming) to help decrease the overfitting.
- Training/validation/test split (e.g., 70% / 15% / 15%).

The preprocessing operations enhance the generalization capability of the model to solve known data scarcity problems in medical imaging.

Model Development

The development of the model entails the building of:

A. Baseline CNN Model

The implementation will be of a single CNN in Python (TensorFlow or PyTorch).

Example components:

- Convolution layers
- Max/Average pooling
- Flatten + Dense layers
- Multi-class classification softmax output.
- level of performance.

B. Refined/Recommended CNN Model.

The enhanced model will incorporate:

- Hybrid CNN + LBP features (extracted features)
- Ensemble learning (e.g. stacking or averaging multiple CNN outputs)
- Increased pooling (e.g., average, max or pyramid pooling)

Such improvements are based on the results of the research like Rocha et al. (2023) and Muller et al. (2024).

Training and Evaluation

Models will be trained on the chosen data with the acceleration of the graphical processors when it is possible.

Examples of important training settings are:

- Batch size (e.g., 32)
- Learning rate optimization
- Loss: categorical cross-entropy.
- Optimizer: Adam or SGD

Once trained, the baseline and improved models are going to be compared using the conventional performance metrics:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion matrix

These measures enable an in-depth comparison of diagnostic performance, strength and pattern of errors.

Comparative Analysis

In order to achieve the research objectives, it will be a comprehensive comparison of:

- Baseline CNN model
- Hybrid/ensemble CNN model.
- The latest models described in the literature.

The comparison will assess:

- Increased accuracy in classification.
- Generalization and strength improvement.
- Training time, model size (computational efficiency)
- Whether hybrid and ensemble improvements are of great value.

This comparison process makes the offered model to be assessed in equal opportunities to the current methods.

Tools and Implementation

The following will be used in the experimental work:

- Python
- TensorFlow or PyTorch
- opencv / scikit-image to extract LBP.
- Acceleration of training on Google Colab / Kaggle with GPUs.

An implementation of the clean code will be made simpler and clean to make it understandable and repeatable.

Ethical Considerations

None of the research involves any personal or sensitive information about patients since the data is publicly available and anonymized. Therefore, there is no need of ethical approval.

Expected outcomes.

- Better Classification Accuracy.

The CNN model is optimized to include hybrid features extraction and ensemble; therefore, the optimized CNN model should have a higher accuracy rate than the baseline CNN and will show a better capacity to identify faint patterns in medical images and increase the reliability of the analysis.

- Improved Model Generalization and Robustness.

The proposed model will achieve enhanced generalization in diverse variations of the image, optimized pooling, and ensemble learning, which are set to reduce overfitting and enhance the performance of the model on unseen test data.

- Fewer Computational Complexities.

Through combination of effective pooling techniques and feature fusion algorithms, the model will be useful in providing dependable performance of classification with reduced computing cost, therefore, more applicable in real-time or resource constrained state of medical applications.

- The Hybrid Feature Fusion has been demonstrated to be effective.

It is predicted that the research will demonstrate that deep CNN features would be more successful with regard to handcrafted features (LBP), which is the case in the literature.

- Comparison to the State-of-the-Art Maquettes.

The research will likely present an unambiguous experimental data on the effectiveness of the proposed optimized CNN model in comparison with the current state-of-the-art methods, which will illuminate potential limitations and opportunities of the research and point to the weakness, strength, and the ways of future research.

Gantt Chart

Task / Activity	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
1. Literature Review Expansion								
2. Dataset Selection & Preprocessing								
3. Baseline CNN Model Development								
4. Optimized CNN Model (Hybrid/Ensemble)								
5. Model Training & Evaluation								
6. Comparative Analysis								

(Baseline vs Proposed)								
7. Writing Chapters 1–3 & Final Editing								
8. Final Proofreading & Submission								

References

- Al-Zubaidi, E. A., & Mijwil, M. M. (2021). *Medical image classification for coronavirus disease (COVID-19) using convolutional neural networks*. *Iraqi Journal of Science*, 62(8), Article 27. <https://doi.org/10.24996/ij.s.2021.62.8.27>
- Kshatri, S. S., & Singh, D. (2023). *Convolutional neural network in medical image analysis: A review*. *Archives of Computational Methods in Engineering*, 30, 2793–2810.
- Müller, D., Soto-Rey, I., & Kramer, F. (2024). *An analysis on ensemble learning optimized medical image classification with deep convolutional neural networks*. *IEEE Access*.
- Nirthika, R., Manivannan, S., Ramanan, A., & Wang, R. (2022). *Pooling in convolutional neural networks for medical image analysis: A survey and an empirical study*. *Neural Computing and Applications*, 34, 5321–5347.
- Rocha, M. M. M., Landini, G., & Florindo, J. B. (2023). *Medical image classification using a combination of features from convolutional neural networks*. *Multimedia Tools and Applications*, 82, 19299–19322.
- Sarvamangala, D. R., & Kulkarni, R. V. (2021). *Convolutional neural networks in medical image understanding: A survey*. *Evolutionary Intelligence*, 15, 1–22.
- Shrivastava, A., Chakkaravathy, M., & Shah, M. A. (2022, December). A comprehensive analysis of machine learning techniques in biomedical image processing using convolutional neural network. In *2022 5th international conference on contemporary computing and informatics (IC3I)* (pp. 1363-1369). IEEE.
- Ajagbe, S. A., Amuda, K. A., Oladipupo, M. A., AFE, O. F., & Okesola, K. I. (2021). Multi-classification of Alzheimer disease on magnetic resonance images (MRI) using deep convolutional neural network (DCNN) approaches. *International Journal of Advanced Computer Research*, 11(53), 51.

Aljuaid, H., Alturki, N., Alsubaie, N., Cavallaro, L., & Liotta, A. (2022). Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning. *Computer methods and programs in biomedicine*, 223, 106951. <https://doi.org/10.1016/j.cmpb.2022.106951>

Kim, H. E., Cosa-Linan, A., Santhanam, N., Jannesari, M., Maros, M. E., & Ganslandt, T. (2022). Transfer learning for medical image classification: a literature review. *BMC medical imaging*, 22(1), 69. <https://doi.org/10.1186/s12880-022-00793-7>

Takahashi, S., Sakaguchi, Y., Kouno, N., Takasawa, K., Ishizu, K., Akagi, Y., Aoyama, R., Teraya, N., Bolatkan, A., Shinkai, N., Machino, H., Kobayashi, K., Asada, K., Komatsu, M., Kaneko, S., Sugiyama, M., & Hamamoto, R. (2024). Comparison of Vision Transformers and Convolutional Neural Networks in Medical Image Analysis: A Systematic Review. *Journal of medical systems*, 48(1), 84. <https://doi.org/10.1007/s10916-024-02105-8>