Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD30225

Deep Learning in Medical Image Analysis

Minxi Mo

Sydney Smart Technology College, Northeastern University at Qinhuangdao, Qinhuangdao, China
190152mo@gmail.com

Abstract. Deep learning has assisted disease diagnosis with the increasing availability of medical imaging data. This article investigates the current applications of deep learning in medical image analysis. The study evaluates their strengths, limitations, and clinical relevance by analyzing the architectures and performances of CNN, U-Net, and DNN model across various imaging modalities and disease types. In addition to model comparison, the review addresses significant data challenges in medical imaging, including limited annotated data, privacy constraints, and class imbalance. The findings show that CNNs excel in image classification for diseases like pneumonia and skin lesions. At the same time, U-Net demonstrates superior performance in tasks requiring precise spatial localization such as brain tumor segmentation. This paper contributes to the growing field by providing an integrated perspective on model performance and data limitations and suggests future directions for improving deep learning adoption in clinical practice.

Keywords: Medical image analysis, Deep learning, CNN, U-Net, Classification

1. Introduction

With the development of medical technology, more and more medical imaging data is used for disease diagnosis. However, due to the complexity of the pathological conditions and the shortage of human experts, accurately and promptly judging these medical images has become a significant problem. Therefore, clinicians and researchers increasingly use computer-assisted intervention to enhance diagnostic support [1]. Deep learning has become a powerful solution for interpreting medical images in recent years.

Deep neural networks (DNNs) and convolutional neural networks (CNNs) can automatically extract hierarchical features from raw images and have demonstrated exceptional performance in image classification and segmentation tasks. These techniques improve diagnostic accuracy and reduce human workload, making them highly promising for clinical deployment. This paper adopts a literature analysis approach to systematically review the application of deep learning in medical image analysis, focusing on classification and segmentation. We examine models such as CNN, U-Net, and DNN, showing their performance and limitations in medical images.

The primary purpose of this study is to analyze the current application of deep learning technology in medical diagnostic images. This study analyses the role of deep learning in medical imaging by comparing the performances of models such as Convolutional Neural Network (CNN), DNN and U-Net on images of different diseases in medicine. This review also explores the main

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD30225

obstacles in clinical application: data issues. Problems include limited training datasets, unbalanced categories, and inconsistent imaging sources. This article aims to bridge the gap between theoretical progress and clinical practice, and support the development of more accurate, user-friendly and reliable artificial intelligence-assisted tools for medical image analysis.

2. Deep learning

2.1. The significance and function of deep learning

Machine learning is how computer systems learn from data and apply algorithms to perform tasks without explicit programming. Machine learning models can make predictions about new data. In contrast, deep learning is inspired by the structure and function of the human brain. It employs hierarchical algorithm networks, enabling machines to handle complex data types such as images, texts, and documents [2]. Deep learning demonstrates exceptional efficacy across various domains, including speech recognition, natural language processing, image analysis, and autonomous system development.

Deep learning can automatically extract features from raw data, eliminating the need for manual feature engineering. This improves the model efficiency and enables the models of different data types and domains to have greater scalability. It can extract high-level features from the front-end image data through architectures such as convolutional neural networks. This speeds up the diagnostic process and reduces human errors. This has dramatically improved the diagnostic efficiency

2.2. Deep learning models

Deep Neural Networks (DNNs) consist of multiple hidden layers of interconnected neurons. Each layer processes its input via weighted connections and nonlinear activation functions, facilitating the network's capacity to capture intricate and abstract representations. Deep neural networks excel in modeling high-dimensional feature spaces, making them highly effective for applications involving structured prediction and hierarchical decision-making processes. This model is widely adopted due to its high classification accuracy in various tasks [3].

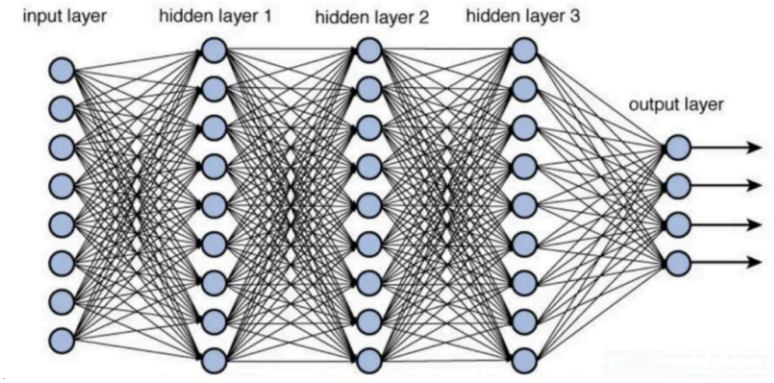

Figure 1. The structure of DNN

Figure 1 shows a standard deep neural network (DNN) with an input layer, three hidden layers, and an output layer. Each layer is fully connected to the next, forming a feedforward structure. The input layer takes raw features, while hidden layers apply weighted transformations and nonlinear

DOI: 10.54254/2755-2721/2025.LD30225

activations to extract higher-level features. The network's depth allows it to learn complex patterns. Depending on the task, the output layer produces predictions using a softmax or linear function. This multilayer perceptron design is standard in classification and regression and forms the basis for more advanced models.

Convolutional Neural Networks (CNNs) are explicitly designed to handle grid-like data structures, such as images or time series. Convolutional neural networks (CNNs) are effective in image analysis tasks because they retain the spatial relationships within input images during processing [4]. In contrast to DNNs, which establish full connectivity between neurons and the preceding layer, CNNs employ localized filters (kernels) that traverse the input, effectively capturing spatial hierarchies and maintaining local feature dependencies. This results in significant parameter reduction and improved efficiency for pattern recognition tasks. CNNs are particularly effective at recognizing shapes, textures, and hierarchical patterns in data.

The U-Net architecture is characterized by a unique U-shaped configuration, comprising a contracting encoder pathway and a symmetrical expanding decoder pathway, facilitating efficient feature extraction and localization within biomedical image datasets [5]. Unlike general CNNs optimized for classification, U-Net is tailored for pixel-wise prediction tasks and has become a standard model in medical image segmentation, particularly for functions such as tumor boundary delineation or organ contour extraction in MRI and CT images.

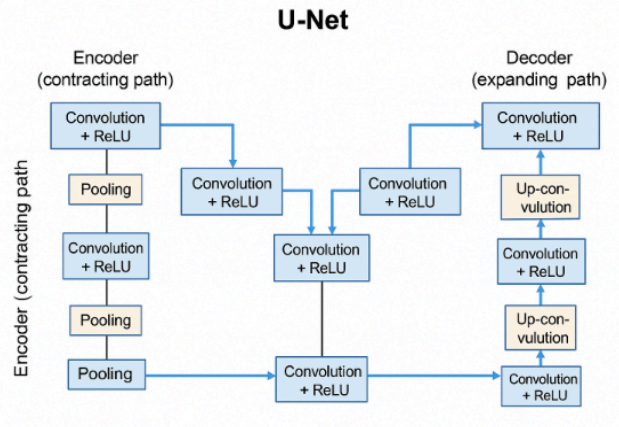

Figure 2. The structure of U-Net

Figure 2 shows the architecture of the U-Net model, which consists of two main components: the encoder on the left and the decoder on the right. The encoder applies repetitive convolution and ReLU activation operations, then pools layers to gradually reduce the spatial dimension and capture advanced features. The decoder uses convolutional layers and encoder connections, including skip connections, to reflect this structure, thereby allowing the network to restore spatial resolution and precisely locate features.

2.3. The role of the model in medical imaging

DNN has powerful feature learning and nonlinear modeling capabilities, can handle complex and high-dimensional medical image data, and is suitable for disease classification and risk prediction tasks. Decision trees, characterized by their inherent structure and transparent logic, offer robust interpretability in medical diagnostics. They are well-suited for preliminary screening, feature selection, or integration within ensemble methodologies, such as random forests. CNN is one of the core models in medical image processing. Because it is good at extracting spatial structure features,

DOI: 10.54254/2755-2721/2025.LD30225

it is widely used in tasks such as lung X-ray detection, brain MRI segmentation, and breast cancer recognition. It can achieve precise positioning and classification of the lesion area.

3. Data challenges in medical imaging

3.1. Data privacy and limited data availability

One of the most significant barriers to developing robust deep learning models in medical imaging is the restricted access to high-quality, labeled datasets. Medical images, such as MRI, CT, and X-rays, often contain sensitive patient information and are subject to strict privacy regulations, including the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in Europe. While essential for protecting patient rights, these regulations pose challenges for data sharing across institutions and hinder the creation of large-scale, diverse datasets required for effective model training.

Moreover, annotating medical images requires domain expertise and considerable time, especially for complex tasks like tumor segmentation or multi-class diagnosis. In contrast to general image datasets, which can be annotated via crowd-sourcing, medical image annotation necessitates expert input from radiologists or clinicians, resulting in protracted and costly data acquisition. This shortage of well-annotated data severely limits the training capacity of deep neural networks, which typically require thousands of examples to generalize effectively.

3.2. Class imbalance in medical image datasets

Class imbalance is also a major data-related challenge in medical imaging. Diagnostic label imbalance often adversely affects a model's sensitivity or specificity, as the model may internalize and replicate biases inherent in the training data distribution [6]. In many clinical datasets, abnormal cases such as specific diseases or rare pathologies occur far less frequently than usual or healthy cases. For instance, in lung nodule detection or breast cancer classification tasks, the proportion of positive cases is typically much smaller than that of negative ones. As a result, machine learning models trained on such imbalanced datasets may become biased toward the majority class, leading to poor sensitivity and high false-negative rates, an unacceptable outcome in clinical practice.

Class imbalance can hinder the model's performance, especially in tasks requiring a high recall rate. Deep learning models tend to support most classes during optimization, usually minimizing losses by completely ignoring underrepresented classes. This undermines the model's reliability and may lead to misleading accuracy metrics, where high overall accuracy masks the poor performance of key minority classes.

DOI: 10.54254/2755-2721/2025.LD30225

4. Compare the usage of different models for images of different diseases

4.1. Medical image analysis process

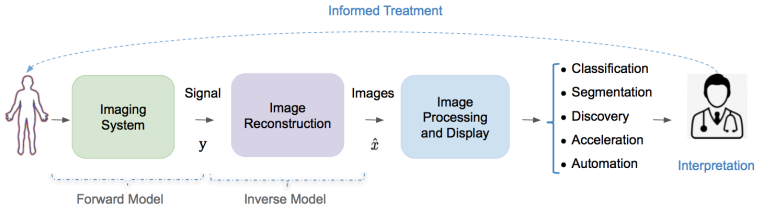

Figure 3. The workflow of the medical image analysis [7]

The workflow of medical image acquisition and analysis can be divided into two main components: the forward and inverse models. The process begins with the imaging system, where physical signals from the human body are acquired. These signals y are then transformed into visual data through the image reconstruction stage, forming readable medical images $\hat{x}$.

When generated, images are passed through image processing and display stages, which can enhance visibility, normalize formats, and prepare the data for interpretation. The emergence of machine learning has introduced data-driven approaches that offer notable advantages.Specifically, deep learning frameworks have demonstrated considerable promise in improving reconstruction performance, facilitating more adaptable and task-oriented modeling approaches [8]. Convolutional neural networks (CNNs), U-Nets, and other architectures are applied to perform advanced tasks such as classification, segmentation, discovery, acceleration, and automation. These tasks contribute directly to clinical interpretation and enable informed treatment planning.

By learning complex features from large volumes of image data, deep models enhance each stage of the image analysis pipeline, making the process more accurate, efficient, and scalable across clinical settings.

4.2. Compare the modeling of images of various diseases in medicine

4.2.1. Pneumonia

Pneumonia is commonly diagnosed through the analysis of chest X-ray images, which are low-resolution grayscale images suitable for binary classification. CNN-based models have demonstrated exceptional performance in this domain. For example, a study employing an Inception-V3-based CNN achieved an accuracy of 99.29% and a recall of 99.73% in pneumonia detection, significantly outperforming traditional DNN approaches [9].

4.2.2. Brain tumors

Brain tumor diagnosis and treatment planning commonly rely on MRI scans, which are complex three-dimensional images requiring precise segmentation. With its symmetric encoder-decoder structure and skip connections, U-Net is particularly well-suited for this task. A recent study reported that a modified U-Net achieved a Dice coefficient of 87.47% in brain tumor segmentation, surpassing the performance of the original U-Net model [10]. This illustrates U-Net's strength in tasks that demand accurate spatial localization within high-dimensional medical data.

DOI: 10.54254/2755-2721/2025.LD30225

4.2.3. Skin lesions

The diagnosis of skin lesions, including melanoma, often relies on dermoscopy images, which are high-resolution color images rich in texture and detail. CNN models have proven highly effective in this domain as well. In a recent investigation, a bespoke convolutional neural network (CNN) demonstrated a classification accuracy of 95.18% for diverse cutaneous lesions, surpassing the performance of conventional machine learning methodologies [11]. This demonstrates the model's robustness in handling fine-grained visual information, essential for dermatological diagnosis.

Table 1. Model comparison

Disease	Imaging Modality	Task Type	Best Model	Performance Metrics
Pneumonia	Chest X-ray	Classification	CNN	Accuracy: 99.29%, Recall: 99.73%
Brain Tumor	MRI	Segmentation	Modified U-Net	Dice coefficient: 87.47%
Skin Lesions	Dermoscopy	Classification	Custom CNN	Accuracy: 95.18%

5. Conclusion

This paper reviews the application of deep learning in medical image analysis. Comparing DNNs, CNNs, and U-Net models, the findings show that each model plays a different role in different medical images. For example, CNNs outperform other models in tasks involving relatively simple, low-dimensional or richly colored images, such as pneumonia detection or classifying skin lesions. Furthermore, U-Net demonstrates significant efficacy in high-resolution, detail-sensitive segmentation applications, such as precise delineation and localization of intracranial neoplasms via magnetic resonance imaging The paper also indicates data-related challenges, such as the scarcity of accurately annotated datasets, privacy restrictions on patient data, and persistent class imbalance issues. These constraints significantly complicate model training and practical deployment in healthcare contexts. Deep learning technology offers substantial potential in medical imaging. Although this article provide a comprehensive overview of deep learning in medical image, there are still some limitations. Because of based on published passage, it does not include original experimental validation for which it can't draw a deep conclusion. Future research efforts should focus on improving model generalizability, exploring synthetic data generation, employing transfer learning techniques, and establishing adequate, privacy-compliant data-sharing infrastructures. Encouraging closer collaboration between artificial intelligence specialists and medical practitioners will also be vital for bridging the gap between theoretical advances and clinical application.

References

[1] Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. Annual Review of Biomedical Engineering, 19, 221–248. https: //doi.org/10.1146/annurev-bioeng-071516-044442

[2] Rana, M., & Bhushan, M. (2023). Machine learning and deep learning approach for medical image analysis: Diagnosis to detection. Multimedia Tools and Applications, 82, 26731–26769. https: //doi.org/10.1007/s11042-022-14305-w

[3] Tran PV (2016) A fully convolutional neural network for cardiac segmentation in short axis MRI. arxiv: 1604.00494. abs/1604.00494

[4] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (pp. 234–241). Springer. https: //doi.org/10.1007/978-3-319-24574-4_28

DOI: 10.54254/2755-2721/2025.LD30225

[5] Ker, J., Wang, L., Rao, J., & Lim, T. (2018). Deep learning applications in medical image analysis. IEEE Access, 6, 9375–9389. https: //doi.org/10.1109/ACCESS.2017.2788044
[6] Osuala, R., et al. (2022). Data synthesis and adversarial networks: A review and meta-analysis in cancer imaging. Medical Image Analysis, 78, 102396. https: //doi.org/10.1016/j.media.2022.102396
[7] Yedder, H. B., Cardoen, B., & Hamarneh, G. (2020). Deep learning for biomedical image reconstruction: A survey. Artificial Intelligence Review, 55, 215–251. https: //doi.org/10.1007/s10462-020-09861-2
[8] Cahyawijaya, S. (2023). Biomedical Image Reconstruction: A Survey. arXiv preprint arXiv: 2301.11813. https: //arxiv.org/abs/2301.11813
[9] Ibrahim, A. U., et al. (2022). Pneumonia classification using deep learning from chest X-ray images during COVID-19. Computers in Biology and Medicine, 139, 105002. https: //doi.org/10.1016/j.compbiomed.2021.105002
[10] Alfaouri, M., et al. (2024). Enhanced brain tumor segmentation using a novel hybrid deep learning model. Mathematical Biosciences and Engineering, 21(3), 5437–5455. https: //doi.org/10.3934/mbe.2024033
[11] Saba, T., et al. (2022). A comprehensive method for skin lesion classification using deep learning and dermoscopy images. Scientific Reports, 12, 18296. https: //doi.org/10.1038/s41598-022-22644-9