



Deep Learning in Medical Imaging: A Comprehensive Review of Techniques, Challenges, and Future Directions

Marwa Riyadh Ahmed

Technical Engineering College for Computer and AI, Northern Technical University, Mosul, Iraq
Email: Marwa.riyadh@ntu.edu.iq

How to cite this paper: Ahmed, M.R. (2025) Deep Learning in Medical Imaging: A Comprehensive Review of Techniques, Challenges, and Future Directions. *Open Access Library Journal*, 12: e14497. <https://doi.org/10.4236/oalib.1114497>

Received: October 25, 2025
Accepted: November 30, 2025
Published: December 3, 2025

Copyright © 2025 by author(s) and Open Access Library Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Deep learning has come to be one of the most transformative generations in current clinical imaging, imparting advanced answers for diagnosis, segmentation, and sickness class. This complete evaluation explores the most essential deep learning architectures—inclusive of convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep perception networks (DBNs)—and their packages in the course of diverse imaging modalities, such as CT, MRI, PET, and fundus images. The examine highlights how deep studying complements accuracy, overall performance, and automation in scientific photograph interpretation. Moreover, it discusses the cutting-edge annoying conditions handling these structures, at the side of confined categorized information, immoderate computational needs, and the shortage of interpretability. The compare concludes with future commands emphasizing explainable AI, hybrid models, records augmentation, and federated analyzing as pathways to triumph over present boundaries and enhance real-international scientific packages.

Subject Areas

Computer Engineering

Keywords

Deep Learning, Medical Imaging, CNN, Segmentation, Classification, Artificial Intelligence, Diagnostic Accuracy, Future Directions

1. Introduction

The technology of small-scale healthcare databases has long surpassed. With the rapid development of imaging technologies and the growing use of biomedical

records series tools, the healthcare region now produces a good sized quantity of complex and multidimensional facts (CT, MRI, and so on). These datasets are characterised with the aid in their excessive dimensionality, numerous variables, and regularly incompatible sources, which makes scientific picture statistics specially difficult—and fascinating—to analyze [1].

The exponential boom in scientific imaging has created a fantastic workload for scientific experts, whose manual interpretations are time-ingesting, subjective, and liable to inconsistency at some stage in professionals. To deal with the ones boundaries, researchers have grown to turn out to be to machine learning (ML) techniques to automate diagnostic techniques. Yet, traditional ML algorithms have proven constraints while tackling quite complex scientific issues [2]. The integration of immoderate-overall performance computing with advanced mastering algorithms gives an efficient pathway towards analyzing massive medical image datasets with accuracy and velocity. Deep getting to know (DL), in particular, now not simplest assists in function extraction and selection however also allows the automated production of latest, information-driven capabilities. Furthermore, DL models can be expecting disorder effects and generate clinically actionable insights that manual physicians in formulating effective treatment plans (Figure 1) [3].

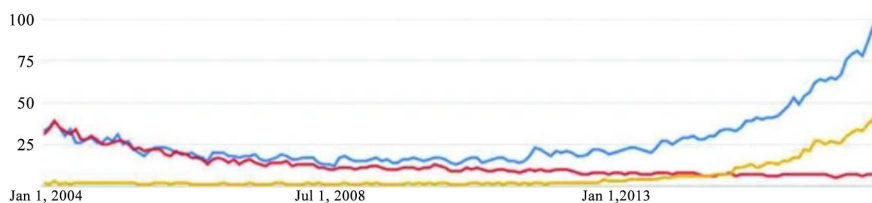


Figure 1. Trends: deep learning vs. machine learning vs. pattern recognition.

Over the beyond decade, both Machine Learning (ML) and Artificial Intelligence (AI) have carried out awesome improvement and function drastically stimulated scientific research and scientific exercising. They have end up important in duties consisting of medical photo processing, computer-aided prognosis, photo registration, segmentation, fusion, and retrieval [4]. ML techniques extract essential data from pictures and arrange it into interpretable formats, enhancing diagnostic precision and assisting faster choice-making [5].

By facilitating in advance and more correct disease detection, AI and ML extend the capabilities of healthcare experts to understand the genetic and natural underpinnings of illnesses. Traditional ML fashions, including Support Vector Machines (SVMs), Neural Networks (NNs), and good enough-Nearest Neighbor (KNN), depend intently handy made functions, requiring domain know-how and good sized preprocessing. In evaluation, deep learning architectures—which include Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Extreme Learning Machines (ELMs), and Generative Adversarial Networks (GANs)—permit the direct

use of uncooked picture records, taking into account fast and automatic function reading [6].

Although classical ML strategies have done astonishing effects in disease detection for many years, modern-day breakthroughs in deep mastering have ushered in a brand new generation of accuracy, speed, and adaptability. DL-based totally techniques have tested effective not handiest in clinical imaging but additionally in domain names which include speech popularity, text assessment, lip analyzing, pc-aided diagnosis, facial reputation, and drug discovery [7].

2. Why Deep Learning over Machine Learning?

The precision of illness evaluation in large part relies upon on two critical additives: picture acquisition and photograph interpretation. In contemporary years, improvements in imaging technology along with X-ray, CT, and MRI have dramatically advanced image high-quality, taking into account radiological visuals of significantly better selection. However, however the rapid progress in imaging devices, the total capacity of automated image interpretation has most effective lately started to be realized [1].

One of the maximum impactful areas of machine learning (ML) is computer vision, but conventional ML algorithms carefully depend on hand made features designed via manner of experts. For example, figuring out lung tumors requires manual extraction of structural characteristics from the records. Because scientific statistics varies extensively across sufferers, such handcrafted strategies regularly lack robustness and consistency [2]. Over time, ML strategies have advanced to cope with large-scale and complex datasets extra effectively, paving the manner for the subsequent technology of reading systems.

Among those advancements, deep gaining knowledge of (DL) has emerged as a innovative pressure—especially within medical photograph evaluation. It is projected that the medical imaging marketplace on my own will surpass \$3 hundred million with the aid of 2021, exceeding the overall investments made in analytical industries in 2016 [3]. DL represents an effective, supervised mastering technique this is predicated on deep neural community (DNN) architectures—complex fashions inspired with the aid of way of the shape and functioning of the human thoughts. Each neuron, the middle computational unit of a DNN, integrates a couple of enter indicators the use of weighted parameters, applies nonlinear changes, and produces outputs that propagate thru the community layers [4].

Neural Networks and Deep Learning Architecture

The foundations of synthetic neural networks (ANNs) hint again to the conceptual have a look at of the human frightened device. One of the earliest neural fashions, the perceptron, emerge as designed with an input layer connected right away to an output layer, capable of classifying linearly separable styles [5]. However, actual-international medical statistics is an extended way more complicated and nonlinear, necessitating deeper architectures comprising enter, hidden, and out-

put layers. In such models, neurons carry out computations on incoming alerts and pass the following outputs to next layers, forming a multilayered network (Figure 2) [6].

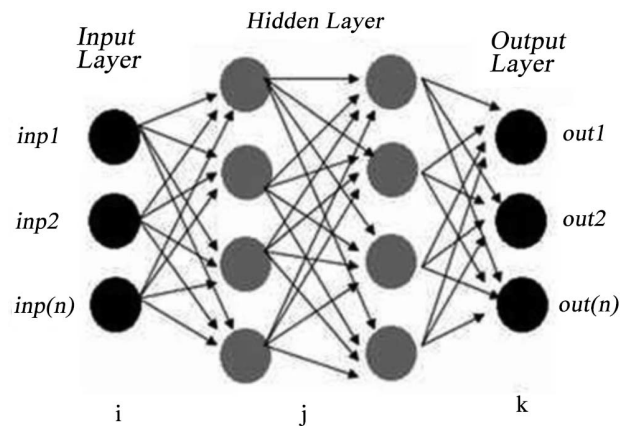


Figure 2. Neural network architecture.

Each neuron aggregates its inputs, applies an activation function, and forwards the end result to the following layer. The inclusion of more than one hidden layers allows the community to learn nonlinear and summary styles. Networks with numerous hidden layers are called Deep Neural Networks (DNNs). Training DNNs turned into traditionally difficult due to computational barriers, but latest enhancements in optimization algorithms and hardware have made deep mastering fee-powerful and efficient (Figure 3) [7].

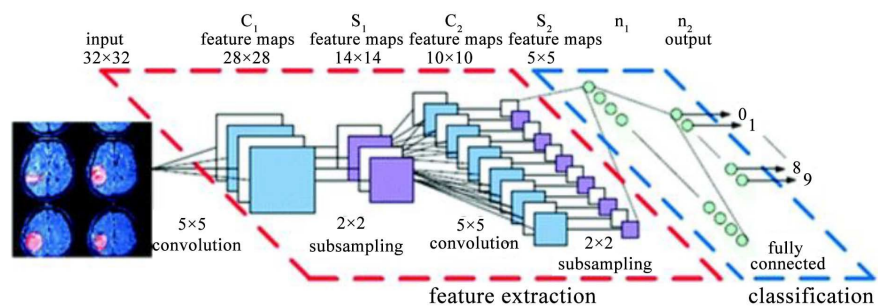


Figure 3. Conventional neural network architecture.

The time period deep learning gained global attention when it turned into listed among the top ten leap forward technologies in 2013. Since then, DL-based systems have confirmed overall performance exceeding human competencies in numerous regions, which includes blood cancer detection, tumor localization in MRI scans, speech popularity, and facial popularity [8]. The better intensity of those neural architectures—comprising numerous hidden layers—has allowed them to perform high-stage abstractions and generate remarkably accurate image analyses.

Modern deep networks might also contain masses or even heaps of layers, al-

lowing them to hierarchically research functions and store great mapping representations between inputs and outputs. With get entry to sufficiently large datasets, those models can generalize to new, unseen scientific cases and carry out clever inferences, which include extrapolations or predictive modeling [9]. Consequently, DL has redefined fields like computer vision, scientific imaging, lip-reading, speech synthesis, and textual content interpretation [10]. All these Comparison of Different are summarized in **Table 1**.

Table 1. Comparison of different architectures of deep learning.

Type of Network	Detail of Networks	Pros	Cons
Deep Neural Network (DNN)	This model includes more than two layers, allowing it to capture complex nonlinear relationships. It is applied in both classification and regression tasks [1].	Highly accurate and widely used in medical image analysis [10].	Training is computationally demanding, as errors are propagated backward through multiple layers, often leading to vanishing gradients and slow learning [3].
Convolutional Neural Network (CNN)	Specially designed for two-dimensional data, CNNs use convolutional filters that transform 2D data into 3D feature representations ([5] [11]).	Demonstrates strong performance and rapid model learning in image classification and segmentation [13].	Requires large amounts of labeled data to achieve accurate classification results [6].
Recurrent Neural Network (RNN)	Capable of learning sequential data, RNNs share weights across all time steps and neurons. Variants include LSTM, BLSTM, MDL-STM, and HLSTM [18].	Excellent at modeling temporal dependencies, providing state-of-the-art results in speech and character recognition, and natural language processing tasks [19].	Suffers from vanishing gradient problems and requires large datasets for effective training [20].
Deep Convolutional Extreme Learning Machine (DC-ELM)	Utilizes Gaussian probability functions to sample local connections, enhancing feature extraction efficiency [8].	Fast training mechanism, computational efficiency, and resilience against random distortions [7].	Initialization can become less effective with simple learning functions or limited labeled data [9].
Deep Boltzmann Machine (DBM)	Based on the Boltzmann family, this network includes unidirectional connections between hidden layers. It incorporates top-down feedback for robust inference [10].	Capable of handling ambiguous data effectively, improving inference robustness [12].	Parameter optimization is challenging for large datasets [4].
Deep Belief Network (DBN)	Consists of multiple layers with unidirectional connections. Each hidden layer serves as the visible layer for the next. It can be used for both supervised and unsupervised learning [2].	Uses a greedy layer-wise strategy for efficient learning, maximizing likelihood and achieving stable inference [5].	Computationally expensive due to complex initialization and training processes [3].
Deep Autoencoder (dA)	Designed primarily for unsupervised learning, it reduces feature dimensionality by reconstructing inputs. The number of inputs equals the number of outputs [11].	Does not require labeled data; variants such as Sparse, Denoising, and Conventional Autoencoders enhance robustness [10].	Requires pre-training, and training stability may degrade without proper optimization [13].

A form of deep learning architectures are presently in use—each tailor-made to unique programs. The maximum prominent encompass Convolutional Neural Networks (CNNs), Deep Neural Networks (DNNs), Deep Belief Networks (DBNs), Autoencoders (dAs), Deep Boltzmann Machines (DBMs), Deep Convolutional Extreme Learning Machines (DC-ELM), and Recurrent Neural Networks (RNNs), along with variants which consist of Bidirectional LSTM (BLSTM) and Multidimensional LSTM (MDLSTM) [11].

Among these, CNNs have received unique interest for their exceptional standard overall performance in virtual picture processing and medical vision responsibilities. Unlike conventional algorithms that require manual feature extraction, CNNs automatically analyze hierarchical features and spatial styles from uncooked pics. Their format is based totally on shared weights, nearby connectivity, and replicated filters, which motive them to computationally efficient and strong [12].

CNN architectures generally include key additives:

- 1) Feature extractors, which exchange among filtering and subsampling layers; and
- 2) Trainable classifiers, which perform prediction based on found out representations.

Several CNN architectures had been widely followed inside the area, including AlexNet (see **Figure 4**), LeNet, Faster R-CNN, GoogleNet, ResNet, VGGNet, and ZFNet, each offering unique blessings in terms of depth, efficiency, and feature generalization (see **Figures 5-8**) [13].

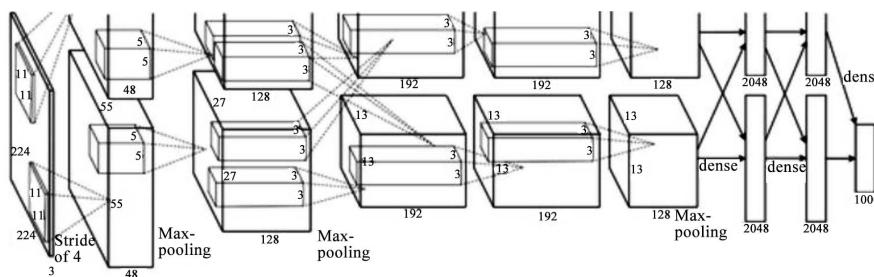


Figure 4. Type of architecture of CNN: AlexNet.

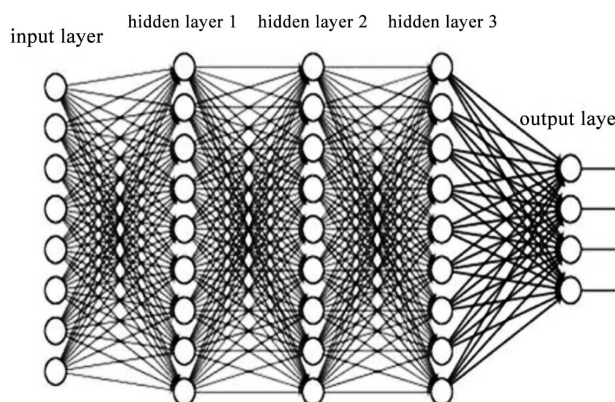


Figure 5. Deep neural network.

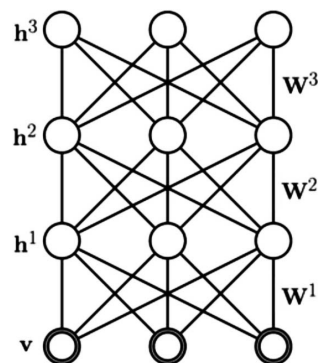


Figure 6. Deep Boltzmann machine (DBM).

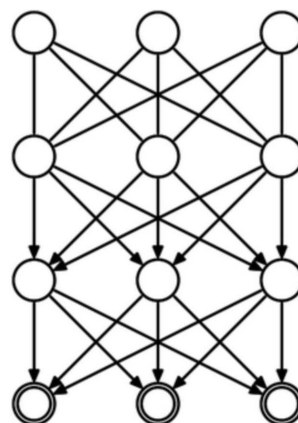


Figure 7. Deep belief network (DBN).

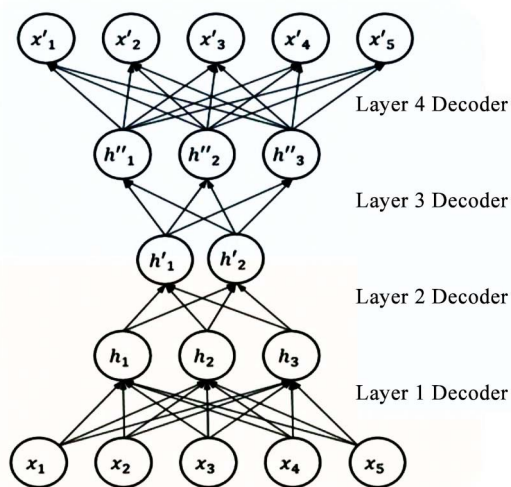


Figure 8. Deep autoencoder (dA).

3. Deep Learning: The Not-So-Near Future in Medical Imaging

The integration of deep learning (DL) into medical imaging stands to turn out to be one of the most transformative technological shifts for the purpose that emergence of digital radiography. Many students count on that in the subsequent

decade and a 1/2 of, DL-powered systems will now not handiest assist human experts however moreover take the lead in executing diagnostic methods, predicting ailments, recommending remedies, and even personalizing scientific prescriptions [1].

Among various clinical disciplines, ophthalmology, pathology, oncology, and radiology are expected to be at the forefront of this modification. Although ophthalmology might witness the earliest revolution, pathology and most cancers diagnostics have already executed brilliant development via DL applications, showing clinically relevant accuracy [2]. For instance, Google DeepMind Health, in partnership with the United Kingdom National Health Service (NHS), signed a 5-yr settlement to manner and study clinical information for as plenty as a million patients—an initiative that demonstrates each the ambition and ability of AI in healthcare [3].

Companies inclusive of IBM Watson Health and Google DeepMind have made good sized investments in clinical imaging. IBM, as an instance, entered the imaging marketplace with a US\$1 billion acquisition of Merge Healthcare, strengthening its AI-pushed diagnostic skills [4]. Despite this large financial momentum, the large-scale adoption of DL in scientific imaging stays hard because of inherent complexities such as information shortage, privateness and felony constraints, and the absence of standardized datasets and algorithms [5].

3.1. Dataset Challenges

The success of deep learning models relies basically on get entry to to large-scale, splendid, and nicely-annotated datasets, which serve as the spine for both version training and validation in the discipline of scientific imaging. However, acquiring such datasets is inherently hard. In scientific and scientific imaging, annotations require now not best expert know-how however also consensus amongst multiple specialists to limit human bias and diagnostic errors. This requirement substantially will increase the complexity and price of dataset practise [6].

Moreover, the availability of information for uncommon conditions is frequently significantly constrained, ensuing in incredibly imbalanced datasets. Certain conditions are underrepresented, which in flip reduces the generalizability of trained models to new or unseen cases. Studies have shown that model overall performance can degrade appreciably while education data lacks enough diversity across extraordinary classes [7]. This imbalance can specially affect deep studying fashions which might be notably facts-based, main to biased predictions favoring overrepresented classes.

The ethical and logistical constraints in medical and medical records series further exacerbate these troubles. Obtaining large volumes of affected person records or uncommon-circumstance instances is often restrained by means of affected person privacy concerns, consent requirements, and resource limitations, making it impractical to acquire sufficiently massive datasets for robust version education. Consequently, dataset imbalance remains one of the primary limitations to ac-

complicating excessive version overall performance [8].

Beyond medical imaging, those challenges are echoed throughout different domain names of deep gaining knowledge of. For example, in actual-time device testing and digital image processing, insufficient facts can lead to fashions that fail underneath real-global conditions, as tested via [14] of their evaluation of tough and tender system responsibilities. The study highlights that artificial datasets, even though useful for preliminary checking out, can't absolutely alternative for complete actual-world statistics due to diffused variations and contextual statistics missing in artificial samples [14].

Furthermore, in domain names like pc imaginative and prescient and wide variety plate popularity, accurate segmentation and detection algorithms closely rely upon well-annotated and balanced datasets. For instance, Mohammad *et al.* (2025) showed that the performance of more than a few plate reputation machine appreciably progressed whilst the dataset included a numerous variety of lights, attitude, and occlusion situations, emphasizing that dataset range is as essential as dataset length [15].

Data preprocessing and enhancement techniques can partially mitigate these demanding situations. Talab *et al.* (2024) [16] explored image evaluation enhancement and equalization strategies, which enhance characteristic extraction for deep mastering models and make amends for a few dataset high-quality obstacles. However, those methods cannot solve the essential difficulty of class imbalance or the shortage of uncommon-situation samples [13].

Hybrid and ensemble fashions, including the ones combining EfficientNet-B0 and ResNet50 with SVM, were proposed to cope with dataset limitations via leveraging complementary strengths of more than one architectures. Qahtan *et al.* (2025) [17] tested that hybrid models can enhance detection accuracy regardless of confined statistics, but they nevertheless remain touchy to imbalances in elegance illustration, underscoring the chronic significance of incredible datasets.

Comprehensive evaluations of deep learning strategies in virtual processing highlight that dataset availability, great, and annotation accuracy continue to be ordinary bottlenecks throughout a couple of domain names, from clinical imaging to engineering programs. Thanoon *et al.* (2022) [18] emphasized that whilst algorithmic innovations continue to enhance, version overall performance is intrinsically tied to the dataset quality, reinforcing the notion that information education is as crucial as version structure design.

In addition, human bias in annotation presents an often-disregarded mission. Annotators may also interpret ambiguous instances in another way, leading to inconsistencies in labels. Multi-professional annotation protocols, although resource-intensive, are vital to lessen labeling mistakes and enhance model reliability. This technique aligns with ethical tips and improves reproducibility in studies settings [6].

Lastly, the interaction between dataset size, high-quality, and augmentation strategies defines the realistic limits of deep getting to know programs. Tech-

niques which include information augmentation, artificial facts generation, and switch learning are increasingly more hired to compensate for limited datasets. Nevertheless, these techniques serve as dietary supplements in preference to replacements for exceptional, properly-annotated datasets, highlighting the essential importance of statistics curation in achieving sturdy model performance [15].

3.2. Privacy and Legal Issues

Sharing scientific imaging facts poses some distance greater complex demanding conditions than sharing regular image information. Legal frameworks, consisting of the Health Insurance Portability and Accountability Act (HIPAA, 1996), strictly regulate how affected person statistics may be stored, accessed, and shared, ensuring that in my opinion identifiable facts remains included [12].

Researchers need to address each technical and sociological factors of records protection. Even after doing away with direct identifiers (e.g., names, social security numbers, or affected person IDs), re-identification risks persist thru bypass-referencing techniques. Emerging techniques which incorporates differential privacy have been proposed to balance records accessibility with confidentiality, but those nevertheless warfare to keep information software program for clinical AI systems [19].

For instance, studies display that combos of easy demographic data—consisting of starting yr, ZIP code, and gender—can uniquely emerge as privy to individuals inside a populace [10]. Consequently, privateness constraints no longer quality lessen dataset period but moreover restriction the variety of records available for model schooling, thereby restricting the effectiveness of DL fashions in real-international healthcare programs [9].

3.3. Data Interoperability and Standards

Data interoperability stays one of the most continual obstacles to DL implementation in healthcare. Imaging statistics varies substantially during incredible hardware gadgets, producers, and acquisition protocols, main to inconsistencies that complicate model training [11].

Effective interoperability calls for that medical photos from diverse gadgets be standardized and prefer minded. Organizations which incorporates HIPAA, HL7, and HITECH have hooked up frameworks and certification requirements—which includes CCHIT and ARRA—to assess and make sure compliance in electronic fitness information [20]. These initiatives purpose to permit seamless information alternate and integration, however huge implementation stays a tremendous challenge in lots of healthcare systems [21].

3.4. The Black-Box Problem in Deep Learning

Despite its revolutionary capacity, deep studying suffers from a big dilemma referred to as the “black-field hassle.” While the mathematical creation of a neural

community can be nicely understood, the reasoning in the back of unique outputs is often opaque and tough to interpret, even for professionals [1].

DL algorithms are capable of identifying subtle styles and relationships in massive datasets that humans may also forget. However, the shortage of transparency in how those fashions obtain their conclusions increases ethical, scientific, and clinical issues, particularly in vital medical selection-making contexts [12]. This problem underscores the want for explainable AI (XAI) systems that can justify their diagnostic predictions and improve take into account in automated healthcare answers [10].

4. Deep Learning in Medical Imaging

Medical image prognosis frequently starts off evolved with figuring out abnormalities and quantifying modifications over the years. Automated image analysis, powered via device reading algorithms, appreciably complements the accuracy and overall performance of analysis and interpretation. Among these techniques, deep getting to know has emerged as an extensively performed approach, reaching modern day performance. It has converted clinical photograph evaluation, providing answers ranging from most cancers detection and disease monitoring to personalized remedy planning. Currently, facts from various assets—which includes radiological imaging (X-ray, CT, MRI), pathology slides, and genomic sequences—offer clinicians huge quantities of information. Yet, device for converting this records into actionable insights continue to be restricted. The following talk highlights superior programs of deep learning in medical imaging, illustrating its large impact on healthcare in recent times.

4.1. Diabetic Retinopathy (DR)

Manual detection of diabetic retinopathy (DR) is tough and time-consuming because of confined get right of entry to specialized equipment and expert clinicians. Early-stage DR frequently affords minimum symptoms, requiring clinicians to take a look at coloured fundus snap shots, that could bring about not on time remedy, miscommunication, and lack of follow-up. Deep studying-primarily based automatic DR detection models have proven advanced accuracy and performance.

Gulshan *et al.* Applied a deep convolutional neural network (DCNN) to the EyePACS-1 and Messidor-2 datasets for identifying mild and extreme DR instances. EyePACS-1 carries approximately 10,000 retinal pictures, even as Messidor-2 has 1, seven-hundred pictures from 874 patients. The authors mentioned 97.5% sensitivity and 93.4% specificity on EyePACS-1, and 96.1% sensitivity and 93.9% specificity on Messidor-2 [1]. Kathirvel skilled a DCNN with dropout layers and tested it on Kaggle fundus, DRIVE, and STARE datasets, accomplishing 94–96% accuracy [2]. Pratt *et al.* Used the NVIDIA CUDA DCNN library on a Kaggle dataset of over 80,000 virtual fundus pictures, resizing them to 512×512 pixels and sprucing the pix. The feature vectors were then labeled the use of Cu-DCNN,

accomplishing as much as 95% specificity, 30% sensitivity, and 75% accuracy [3].

Haloi implemented a 5-layer CNN with dropout for early-degree DR detection on Retinopathy Online Challenge (ROC) and Massidor datasets. Sensitivity reached 97%, specificity 96%, accuracy 96%, and location underneath the curve (AUC) zero.988 on the Massidor dataset, with ROC dataset AUC as much as zero.98 [4]. Alban and Gilligan denoised EyePACS angiograph pictures earlier than CNN evaluation, diagnosing 5 severity classes with 79% AUC and 45% accuracy [5]. Lim *et al.* Extracted vicinity-based totally capabilities and categorised them the use of DCNN on DIARETDB1 and SiDRP datasets [6]. All these outcomes are summarized in **Table 2**.

Table 2. Summary of deep learning (DL) for diabetic retinopathy (DR).

Authors	Model	Dataset	Accuracy / Sensitivity/ Specificity (%)
Gulshan <i>et al.</i> (as cited in [1] [10])	Deep Convolutional Neural Network (CNN)	EyePACS-1, Messidor-2	97.5% sensitivity & 93.4% specificity (EyePACS-1); 96.1% sensitivity & 93.9% specificity (Messidor-2)
Kathirvel (as cited in [5] [11])	CNN with Dropout Layer	Kaggle Fundus, DRIVE, STARE	94% - 96% accuracy
Pratt <i>et al.</i> (as cited in [2])	Cu-DCNN Library	Kaggle	75% accuracy
Haloi <i>et al.</i> (as cited in [12])	Five-Layer CNN	Messidor, ROC	98% AUC (Messidor); 97% AUC (ROC)
Alban <i>et al.</i> (as cited in [3])	Deep Convolutional Neural Network (DCNN)	EyePACS	45% accuracy
Lim <i>et al.</i> (as cited in [4])	Deep Convolutional Neural Network (DCNN)	DIARETDB1, SiDRP	Not reported

4.2. Detection of Histological and Microscopical Elements

Histological evaluation examines cells, tissue businesses, and their business enterprise. Cellular and tissue-stage modifications can be detected the usage of microscopic imaging and marking techniques (coloured chemical substances), regarding steps like fixation, sectioning, staining, and optical imaging. This approach is extensively used for diagnosing skin cancers along with squamous cell carcinoma and melanoma, as well as gastric carcinoma, breast most cancers, malaria, intestinal parasites, and tuberculosis [7]. For example, Plasmodium is the number one parasite responsible for malaria, detected the usage of stained blood smears, whilst Mycobacterium in sputum suggests tuberculosis, commonly recognized through smear microscopy with auramine-rhodamine or Ziehl-Neelsen staining (See **Figure 9**).

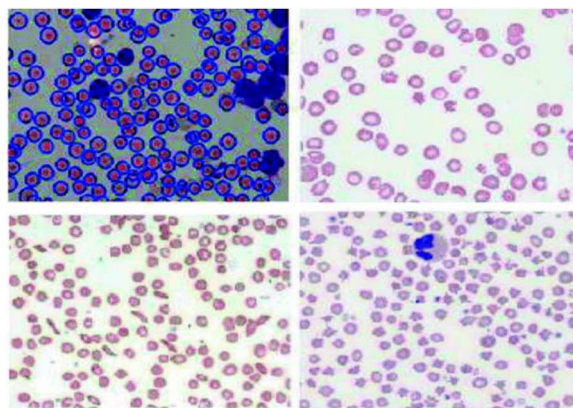


Figure 9. Microscopic blood smear images.

Quinn *et al.* extracted shape and morphological features to detect malaria, tuberculosis, and hookworm from blood, sputum, and stool samples, reporting AUC of 100% for malaria and 99% for tuberculosis and hookworm. Similar DCNN models were applied to malaria detection, intestinal parasites, cell counting, and leukemia detection [9]. Dong and Bryan compared GoogLeNet, LeNet-5, and AlexNet for malaria-infected cell detection, achieving accuracies of 98.13%, 96.18%, and 95.79%, respectively, while SVM-based classification reached 91.66% [10]. **Table 3** summarizes these studies.

Table 3. Summary of deep learning (DL) for histological and microscopical elements detection.

Authors	Model	Data set	Accuracy: acc or sensitivity: sensi or specificity: spec (%)
Bayramogl and Heikkil	Transfer approach with CNN	ImageNet (source for features) HistoPheno—types dataset	—
Quinn <i>et al.</i>	DCNN and shaped features like moment and morphological	Microscopic image	100% for Malaria; 90% for tuberculosis and hookworm
Qiu <i>et al.</i>	DCNN	—	—
Dong <i>et al.</i>	GoogLeNet, LeNet-5, and AlexNet	—	98.66, 96.18 and 95.79% acc

4.3. Detection of Gastrointestinal Diseases (GI)

Gastrointestinal diseases affect the organs involved in digestion, absorption, and waste excretion, including the esophagus, stomach, small intestine, and large intestine. Disorders such as ulcers, polyps, cancer, celiac disease, Crohn's disease, tumors, intestinal obstruction, and vascular abnormalities affect these organs ([11], p. 235). Imaging techniques such as wireless capsule endoscopy (WCE), endoscopy, colonoscopy, X-rays, CT, MRI, and intra-operative enteroscopy are used for diagnosis.

Jia and Meng employed a DCNN on 10,000 WCE images for GI bleeding detection, reporting an F-measure of 99% [12]. Pei *et al.* used FCNs and FCN-LSTM models on cine-MRI sequences to measure bowel contraction frequency and length [15]. Wimmer *et al.* extracted ImageNet features and classified celiac disease from duodenal endoscopic images using CNN-SoftMax [16].

Zhu *et al.* used CNN for feature extraction from endoscopic images, followed by SVM classification, achieving 80% accuracy. Georgakopoulos *et al.* combined CNN and SVM for inflammatory GI disease detection from WCE videos, reporting up to 90% overall accuracy. Ribeiro *et al.* used multiple CNN models to analyze colonoscopy videos with multi-scale image features for polyp detection. These findings are summarized in **Table 4** and **Table 5**.

Table 4. Summary of deep learning (DL) for gastrointestinal (GI) diseases.

Authors	Model	Data set	Accuracy: acc or sensitivity: sensi or specificity: spec (%)
Bayramogl and Heikkil	Transfer approach with CNN	Image Net (source for features) HistoPheno—types dataset	
Qiu <i>et al.</i>	DCNN and shaped features like moment and morphological	Microscopic image	100% for malaria: 99% for tuberculosis and hookworm
Qiu <i>et al.</i>	DCNN	-	-
Dong <i>et al.</i>	GoogLeNet, LeNet-5, and A lex Net		98.66, 96.18 and 95.70% acc

Table 5. Summary of deep learning (DL) for tumor detection.

Authors	Model	Data set	Accuracy: acc or sensitivity: sensi or specificity: spec (%)
Bayramogl and Heikkil	Transfer approach with CNN	Image Net (source for features) HistoPheno—types dataset	-
Quinn <i>et al.</i>	DCNN and shaped features like moment and morphological	Microscopic image	100% for malaria; 99% for tuberculosis and hookworm
Qiu <i>et al.</i>	DCNN	-	-
Dong <i>et al.</i>	GoogLeNet, LeNet-5, and A lex Net		98.66, 96.18 and 95.79% acc

4.4. Cardiac Imaging

Deep getting to know has proven good-sized promise in cardiac imaging, mainly for quantifying calcium scores. Numerous programs have been developed, with CT and MRI scans being the maximum usually used modalities. The primary goal

for image segmentation is regularly the left ventricle. Manual identity of Coronary Artery Calcium (CAC) in cardiac CT pics calls for huge expert involvement, making it each time-consuming and impractical for big-scale or epidemiological research. To deal with these demanding situations, semi-automatic and automatic calcium scoring methods were brought for cardiac CT. Recent studies has centered on CT angiographic picture-based CAC calculation using deep convolutional neural networks (DCNNs), as illustrated in **Figure 10** [1].

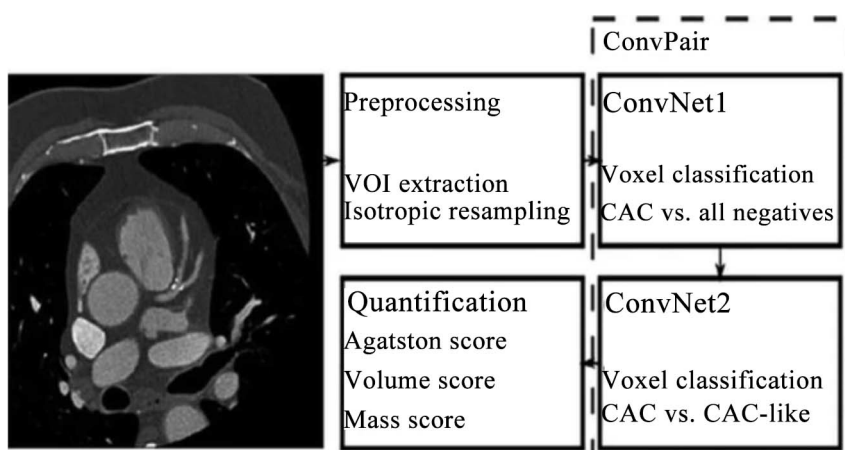


Figure 10. Calcium score classification. Source: Wolterink *et al.*

4.5. Tumor Detection

Tumors, or neoplasms, arise whilst cells develop abnormally to shape a mass. They may be categorized as benign (non-cancerous) or malignant (cancerous). Benign tumors continue to be localized and typically pose minimal danger, whilst malignant tumors spread to other parts of the frame, complicating remedy and worsening diagnosis (**Table 6**) [2].

Table 6. Summary of deep learning (DL) for tumor detection.

Authors	Model	Data set	Accuracy: acc or sensitivity: sensi or specificity: spec (%)
Bayramogl and Heikkil	Transfer approach with CNN	Image Net (source for features) HistoPheno—types dataset	–
Quinn <i>et al.</i>	DCNN and shaped features like moment and morphological	Microscopic image	100% for malaria; 99% for tuberculosis and hookworm
Qiu <i>et al.</i>	DCNN	–	–
Dong <i>et al.</i>	GoogLeNet-, LeNet-5, and A lexNet	–	98.66, 96.18 and 95.79% acc

Wang and Qu applied a pipeline on 482 mammographic images from women

aged 32 - 70, of which 246 contained tumors. Images were denoised using a median filter, and breast tumors were segmented using region growing, morphological operations, and modified wavelet transformation. Morphological and textural features were then classified using extreme learning machines (ELM) and support vector machines (SVM), achieving total error rates of 84% and 96%, respectively [3]. Kooi and den Heeten worked with limited datasets of malignant masses and benign cysts, employing CNN models augmented with image variations to achieve an AUC of up to 87% (Figure 11) [4].

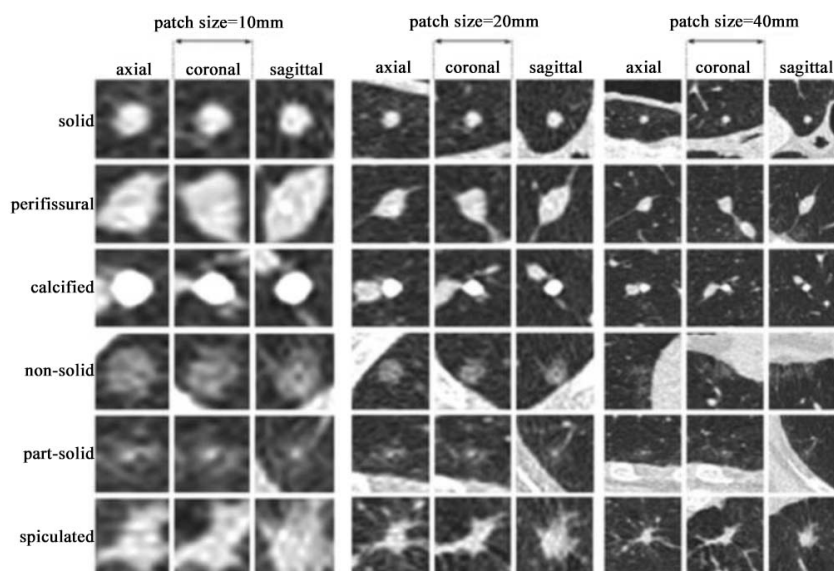


Figure 11. Lung nodule segmentation. Source: Cui *et al.*

4.6. Detection of Alzheimer's Disease and Parkinson's Disease

Arevalo *et al.* Performed a study on 736 mediolateral indirect and craniocaudal mammographic images from 344 sufferers, manually segmenting 310 malignant and 426 benign lesions. Images had been improved and fed into a CNN for classification, attaining an AUC of 82.6% [5]. Huynh and Giger carried out CNNs to breast ultrasound photos with 2,393 areas of hobby from 1, a hundred twenty-five sufferers, acting two experiments: category of CNN-extracted features and handmade functions via SVM. The CNN features reached 88% AUC, at the same time as handmade capabilities reached eighty-five% AUC. Further, the usage of a CNN for function extraction and SVM for category, they executed 86% AUC on 219 lesions in 607 breast pics [6].

Antropova and Giger explored transfer gaining knowledge of from ImageNet datasets, using four,096 functions labeled by means of SVM for 551 MRI breast pictures, achieving up to 85% AUC. Samala *et al.* Employed a frozen switch mastering DCNN on 2282 mammograms, reporting 99% AUC, which confirmed at 90% on tool tree blob photographs. Shin *et al.* Nice-tuned CNN fashions skilled on ImageNet to discover thoraco-stomach lymph node and interstitial lung lesions, reaching 83% - 85% sensitivity and as much as 95% AUC [7].

Parkinson's sickness is a neurodegenerative sickness characterized by way of innovative motor decline because of basal ganglia disorder, related to dopaminergic neuron loss [8]. Alzheimer's diagnosis commonly employs neurological testing (e.g., MMSE) and mind imaging. Sarraf *et al.* Carried out LeNet-5 CNN on 4D fMRI, education on 270,900 photographs and checking out on 90,300, accomplishing 96.86% accuracy for Alzheimer's detection. Additional tactics include deep belief networks (DBM), 3D-CNN, and sparse autoencoders, applied on MRI and PET datasets, reporting accuracies ranging from 87.76% to 100% across various studies [9]. All these outcomes are summarized in **Table 7**.

Table 7. Summary of deep learning (DL) for Alzheimer's and Parkinson's detection.

Application Area	Input Data	Deep Learning Method
Cardiac CAC	CT	CNN (Wolterink <i>et al.</i> , 2017; Lessmann <i>et al.</i> , 2018)
Lung Cancer	MRI	CNN (Sakamoto & Nakano, 2020)
Lung Cancer	CT	DNN (Ciompi <i>et al.</i> , 2017; Paul <i>et al.</i> , 2019)
Diabetic Retinopathy	Fundus Image	CNN (Pratt <i>et al.</i> , 2016; Gulshan <i>et al.</i> , 2016)
Blood Analysis	Microscopic	CNN ([13]; Xie <i>et al.</i> , 2019)
Blood Analysis	Microscopic	DBN (Duggal <i>et al.</i> , 2019)
Blood Vessel Detection	Fundus	DNN (Liskowski & Krawiec, 2016)
Blood Vessel Detection	Fundus	CNN (Ngo & Han, 2019)
Brain Lesion Segmentation	MRI	CNN (Kamnitsas <i>et al.</i> , 2017; Cui <i>et al.</i> , 2019; Kleesiek <i>et al.</i> , 2016)
Polyp Recognition	Endoscopy	CNN (Yuan & Meng, 2017; Segu <i>et al.</i> , 2020; Wimmer <i>et al.</i> , 2016)
Alzheimer's Disease	PET	CNN (Sarraf <i>et al.</i> , 2016)

5. Open Research Issues and Future Directions

The speedy development of deep learning has been driven with the aid of 3 primary factors: the availability of huge datasets, superior algorithms stimulated through the human brain, and growing computational electricity. Although the ability of deep gaining knowledge of is substantial, reaching effective effects involves good sized effort and funding. Leading generation agencies which includes Google DeepMind and IBM Watson, in collaboration with hospitals and medical device vendors, are actively developing premier answers for clinical imaging. Major providers including Siemens, Philips, Hitachi, and GE Healthcare have made good sized investments. Similarly, research labs like Google and IBM are operating to create practical imaging programs. For instance, IBM Watson collaborates with extra than 15 healthcare carriers to explore real-international deep gaining knowledge of packages. Google DeepMind partners with the NHS within the UK to analyze anonymized eye scans of one.6 million patients, aiming to stumble on early signs and symptoms of diseases which can purpose blindness. GE Healthcare,

in collaboration with Boston Children's Hospital, is developing smart imaging technologies to pick out pediatric brain problems. Additionally, GE Healthcare and the University of California, San Francisco, announced a three-12 months undertaking to create algorithms that differentiate everyday results from the ones requiring professional attention [1].

5.1. Requires Extensive Inter-Organization Collaboration

Despite the efforts of fundamental stakeholders and predictions of deep learning's growth in scientific imaging, there is ongoing debate about changing human knowledge with machines. Nevertheless, deep learning holds tremendous promise for disease prognosis and remedy. Achieving this ability, but, calls for great collaboration amongst hospitals, carriers, and device studying researchers. Such collaboration can deal with the shortage of to be had annotated clinical data, which is one of the key demanding situations. Another difficulty is developing state-of-the-art techniques to deal with the growing extent of healthcare facts, especially because the industry movements in the direction of body sensor network-based totally tracking [2].

5.2. Need to Capitalize on Big Image Data

Deep getting to know programs require massive, annotated datasets, yet scientific photograph annotation stays luxurious, arduous, and time-consuming. Unlike actual-international photo annotation, scientific statistics annotation demands expert time and is complex by the subjective nature of professional opinions. Annotation for uncommon sicknesses is specifically scarce. Sharing datasets among healthcare providers can assist triumph over those limitations, enhancing the availability of statistics for gadget learning research [3].

5.3. Advancement in Deep Learning Methods

Most deep learning techniques in healthcare are supervised, but acquiring annotated clinical records is often impractical, specially for uncommon illnesses. To conquer facts scarcity, there's a need to shift from supervised to unsupervised or semi-supervised approaches. The effectiveness of these strategies in healthcare stays below investigation. Additionally, making sure that unsupervised or switch learning approaches keep the accuracy required for scientific packages is a tremendous undertaking. Despite advances, many questions continue to be unresolved, and deep mastering theories haven't begun to offer complete answers [4].

5.4. Black-Boxes and Their Acceptance by Health Professionals

Healthcare experts remain careful because deep getting to know fashions are frequently "black containers," providing little transparency on how selections are made. While device gaining knowledge of researchers argue that interoperability can be much less of an issue, acceptance in healthcare is predicated heavily on human accept as true with. Professionals searching for evidence of deep studying

fulfillment in vital actual-international applications, inclusive of autonomous cars and robotics. Legal implications of black-field systems, consisting of accountability for wrong diagnoses, remain a barrier. Unlocking the inner choice-making of these fashions is a key studies consciousness (Figure 12) [5].

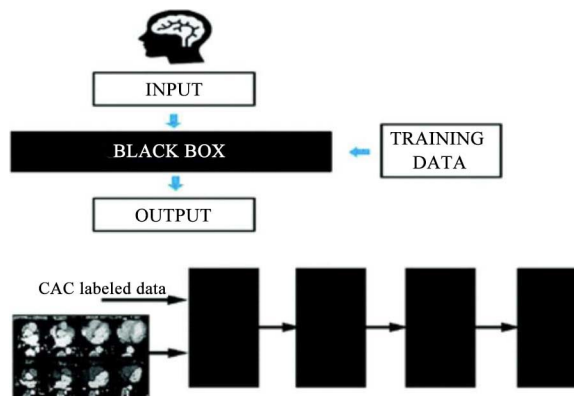


Figure 12. Deep learning: a black box.

5.5. Privacy and Legal Issues

Privacy and felony concerns in healthcare information are each technical and sociological. Regulations like HIPAA provide sufferers with rights over their identifiable records and obligate providers to guard and restriction its use. However, anonymizing large, dynamic healthcare datasets gives full-size challenges. Limited facts access can reduce valuable facts content material, and the steady increase and variation of records make present techniques inadequate for correct coping with [6].

6. Conclusions

In recent years, deep mastering has end up critical to automating numerous duties and has outperformed traditional gadget gaining knowledge of in lots of regions. Researchers count on that within the next 15 years, self-sustaining systems will manage increasingly more day by day activities, which includes clinical duties. However, deep gaining knowledge of adoption in healthcare, mainly scientific imaging, remains slower than in other domains. This chapter highlighted key demanding situations hindering deep gaining knowledge of boom in healthcare and supplied cutting-edge packages in clinical image analysis.

While huge research businesses continue to expand deep learning solutions for medical imaging, challenges inclusive of the shortage of annotated datasets continue to be. The performance of deep learning fashions relies upon heavily on the availability of amazing records. Recent trends endorse that large datasets improve outcomes, but the volume to which big facts can be leveraged in healthcare remains unsure. Despite fine results, the complexity and sensitivity of healthcare records necessitate extra sophisticated deep studying methods. The potential to enhance healthcare thru deep studying is significant, with numerous possibilities

for advancing analysis and remedy [7].

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] Vishraj, R., Gupta, S. and Singh, S. (2022) A Comprehensive Review of Content-Based Image Retrieval Systems Using Deep Learning and Hand-Crafted Features in Medical Imaging: Research Challenges and Future Directions. *Computers and Electrical Engineering*, **104**, Article 108450. <https://doi.org/10.1016/j.compeleceng.2022.108450>
- [2] Altaf, F., Islam, S.M.S., Akhtar, N. and Janjua, N.K. (2019) Going Deep in Medical Image Analysis: Concepts, Methods, Challenges, and Future Directions. *IEEE Access*, **7**, 99540-99572. <https://doi.org/10.1109/access.2019.2929365>
- [3] Birjais, R. (2025) Challenges and Future Directions for Segmentation of Medical Images Using Deep Learning Models. In: *Deep Learning Applications in Medical Image Segmentation: Overview, Approaches, and Challenges*, Wiley-IEEE Press, 243-264.
- [4] Kumar, R.R. and Priyadarshi, R. (2024) Denoising and Segmentation in Medical Image Analysis: A Comprehensive Review on Machine Learning and Deep Learning Approaches. *Multimedia Tools and Applications*, **84**, 10817-10875. <https://doi.org/10.1007/s11042-024-19313-6>
- [5] Xu, Y., Quan, R., Xu, W., Huang, Y., Chen, X. and Liu, F. (2024) Advances in Medical Image Segmentation: A Comprehensive Review of Traditional, Deep Learning and Hybrid Approaches. *Bioengineering*, **11**, Article 1034. <https://doi.org/10.3390/bioengineering11101034>
- [6] Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M. and Henkel, J. (2024) A Comprehensive Survey of Convolutions in Deep Learning: Applications, Challenges, and Future Trends. *IEEE Access*, **12**, 41180-41218. <https://doi.org/10.1109/access.2024.3376441>
- [7] Ghorbian, M., Ghorbian, S. and Ghobaei-Arani, M. (2025) Advancements in Machine Learning for Brain Tumor Classification and Diagnosis: A Comprehensive Review of Challenges and Future Directions. *Archives of Computational Methods in Engineering*, 1-36. <https://doi.org/10.1007/s11831-025-10340-6>
- [8] Abdalredha Kadhim, D. and Abed Mohammed, M. (2024) A Comprehensive Review of Artificial Intelligence Approaches in Kidney Cancer Medical Images Diagnosis, Datasets, Challenges and Issues and Future Directions. *International Journal of Mathematics, Statistics, and Computer Science*, **2**, 199-243. <https://doi.org/10.59543/ijmscs.v2i.9747>
- [9] Mahmood, T., Rehman, A., Saba, T., Nadeem, L. and Bahaj, S.A.O. (2023) Recent Advancements and Future Prospects in Active Deep Learning for Medical Image Segmentation and Classification. *IEEE Access*, **11**, 113623-113652. <https://doi.org/10.1109/access.2023.3313977>
- [10] Zhou, S.K., Greenspan, H., Davatzikos, C., Duncan, J.S., Van Ginneken, B., Madabhushi, A., et al. (2021) A Review of Deep Learning in Medical Imaging: Imaging Traits, Technology Trends, Case Studies with Progress Highlights, and Future Promises. *Proceedings of the IEEE*, **109**, 820-838. <https://doi.org/10.1109/jproc.2021.3054390>
- [11] Li, X., Zhang, L., Yang, J. and Teng, F. (2024) Role of Artificial Intelligence in Medical

- Image Analysis: A Review of Current Trends and Future Directions. *Journal of Medical and Biological Engineering*, **44**, 231-243. <https://doi.org/10.1007/s40846-024-00863-x>
- [12] Rasool, N. and Iqbal Bhat, J. (2023) Unveiling the Complexity of Medical Imaging through Deep Learning Approaches. *Chaos Theory and Applications*, **5**, 267-280. <https://doi.org/10.51537/chaos.1326790>
- [13] Razzak, M.I., Naz, S. and Zaib, A. (2017) Deep Learning for Medical Image Processing: Overview, Challenges and the Future. In: Dey, N., Ashour, A. and Borra, S. Eds., *Lecture Notes in Computational Vision and Biomechanics*, Springer International Publishing, 323-350. https://doi.org/10.1007/978-3-319-65981-7_12
- [14] Younis, N.K., Ahmed, M.R., Talab, A.W. and Ali, R.R. (2025) Testing Real-Time System Algorithms Performance on Synthetic Data: Analytical Study of Hard and Soft System Tasks. *Edelweiss Applied Science and Technology*, **9**, 607-614. <https://doi.org/10.55214/25768484.v9i5.6956>
- [15] Mohammad, S.H., Alrikabi, I.Z.C. and Al-Fayyadh, H.R.D. (2025) Number Plate Recognition System Based on an Improved Segmentation Method. *International Journal of Computational and Electronic Aspects in Engineering*, **6**, 42-50. <https://doi.org/10.26706/ijceae.6.1.20250207>
- [16] Talab, A.W., Younis, N.K. and Ahmed, M.R. (2024) Analysis Equalization Images Contrast Enhancement and Performance Measurement. *OALib*, **11**, 1-11. <https://doi.org/10.4236/oalib.1111388>
- [17] Qahtan, M.H., Asaad, A.M. and Younis, A.K. (2025) Bone Fracture Detection Using Hybrid Efficientnet-B0 and Resnet50 with SVM: A Comparative Performance Analysis. *Ingénierie des Systèmes d'Information*, **30**, 1775-1782. <https://doi.org/10.18280/isi.300710>
- [18] Thanoon, K.H., Shareef, A.F. and Alsaif, O.A. (2022) Digital Processing and Deep Learning Techniques: A Review of the Literature. *NTU Journal of Engineering and Technology*, **1**, 15-24. <https://doi.org/10.56286/ntujet.v1i3.223>
- [19] Verma, S., Goel, T., Tanveer, M., Ding, W., Sharma, R. and Murugan, R. (2023) Machine Learning Techniques for the Schizophrenia Diagnosis: A Comprehensive Review and Future Research Directions. *Journal of Ambient Intelligence and Humanized Computing*, **14**, 4795-4807. <https://doi.org/10.1007/s12652-023-04536-6>
- [20] Sakib, M., Mustajab, S. and Alam, M. (2024) Ensemble Deep Learning Techniques for Time Series Analysis: A Comprehensive Review, Applications, Open Issues, Challenges, and Future Directions. *Cluster Computing*, **28**, Article No. 73. <https://doi.org/10.1007/s10586-024-04684-0>
- [21] Muzammul, M. and Li, X. (2025) Comprehensive Review of Deep Learning-Based Tiny Object Detection: Challenges, Strategies, and Future Directions. *Knowledge and Information Systems*, **67**, 3825-3913. <https://doi.org/10.1007/s10115-025-02375-9>