Journal of Basrah Researches (Sciences) **51**(2), 245 (2025)
DOI: https://doi.org/10.56714/bjrs.51.2.17

A survey: AI Techniques for Medical Image Analysis

Sarah j. Mohammed*, Rasha I. Ahmed, Rasha M. Mohsin

College of Computer Science, University of Technology, 10066, Baghdad, Iraq.

A R T I C L E I N F O

Received 06 October 2025
Revised 19 November 2025
Accepted 25 November 2025
Published 31 December 2025

K e y w o r d s :

Convolution Neural Networks,
Computer Vision,
Medical Image Analysis.

Citation: S. j. Mohammed et al., J. Basrah Res. (Sci.) **50**(2), 245 (2025). DOI:https://doi.org/10.56714/bjrs.51.2.17

A B S T R A C T

Diagnosing diseases and making the right decision is one of the most controversial topics. Especially the development stage of diseases that affect humanity and the spread of many intractable types that have made it difficult for the doctor to make a decision, as in detecting cancerous diseases and brain tumors. This survey explores recent advancements in artificial intelligence (AI) techniques applied to medical imaging modalities such as MRI, CT, and X-ray. Automated diagnosis can facilitate the process by utilizing all variables and evidence to reach a sound conclusion. In particular, convolutional neural networks (CNNs) and their variants have become the cornerstone of modern medical image analysis, enabling effective segmentation, classification, and detection of regions of interest. Most of these networks are capable of dissecting images and capturing sites of interest when analyzing images. This study presents several types of networks capable of analyzing different types of medical images, such as U-Net and Res-Net. This paper highlights current challenges—such as data imbalance, model interpretability, and generalization—and identifies promising future research directions in the field of AI-driven medical image analysis. We presented a comparison of these works in terms of the level of accuracy, speed, training, and dataset types.

1. Introduction

One of the most important steps that can contribute to the success of medical diagnosis is to accurately find the important areas in the medical image. The first step in analyzing medical images is the segmentation process using intelligent algorithms. It is known that traditional methods used in image segmentation processes do not give accurate results in identifying areas of importance [1,2]. There are applications in which segmentation must be accurate, such as surgical planning, imag-based diagnosis and anatomical structure modeling. Among the factors that lead to the success of analyzing medical images by segmenting them are the quality of the image and the homogeneity of the patients' pathological parameters [3]. To increase efficiency, segmentation interactions with the user must be minimized, The goal of machine learning is to achieve accurate data-driven representations. Examples include Gaussian models and the GrabCut algorithm [4]. To obtain optimal results in image segmentation, deep learning techniques are employed in conjunction with convolutional neural network (CNN) architectures.Most researchers aspire to achieve accurate results by passing through a single processing to save test time. In the recent period, neural networks have improved in two

*Corresponding author email: Sarah.j.mohammed@uotechnology.edu.iq

ISSN: 1817-2695 (Print); 2411-524X (Online)
Online at: https://jou.jobrs.edu.iq

aspects: overcoming the problem of few images in the case of repeated collection of weak samples and large collection. Thus, it depends on taking strong samples that support the accuracy of the work. The other aspect of the networks that deal most with image processing, especially in detecting cells in two-dimensional to three-dimensional images. It is the U-Net network, which has gained wide importance [5, 6]. As mentioned above, all research in this field suffers from gaps. These can be summarized as: a weak (or limited) database, inaccurate selection of segmentation processes, difficulty obtaining accurate results within a standard testing time, and a lack of a suitable neural network selection method for case management.

2. Related works

Several types of convolutional neural networks (CNNs), including VGG, AlexNet, GoogleNet, and ResNet, have been analyzed and employed to perform image classification tasks. Most previous studies have concentrated on processing image pixels and their classification potential, owing to the ability of these models to effectively segment and represent pixel regions. [7]. The accuracy of these networks can be measured and weaknesses identified by testing time. The entire image is fed into neural network and the results are in the form of small, dense parts. One of the problems that appear in neural networks is maximum pooling with weak samples [8]. Some scientists have turned to using groups of deconvolution functions to increase the layers [9]. The U-Net network, which was used in segmenting biomedical images, appeared. Another network called V-Net, was proposed to perform segmentation on prostate MRI images [10]. The quality of the samples, varying from weak to strong, greatly affects the performance of any network, which prompted neural network developers to create networks with a comprehensive convolution of all architectural parameters and available features without ignoring any feature, as in examples of object tracking and segmentation work [11]. To improve segmentation precision, neural networks may be employed to extract multi-scale features. These features can be obtained by repeatedly passing the inputs through the network. The resulting features can be a powerful tool for obtaining classification. Some applications use another type of segmentation called interactive segmentation. User interaction methods vary into types such as bounding box, click-based and contour-based. Figure 1 show U-Net network [12].

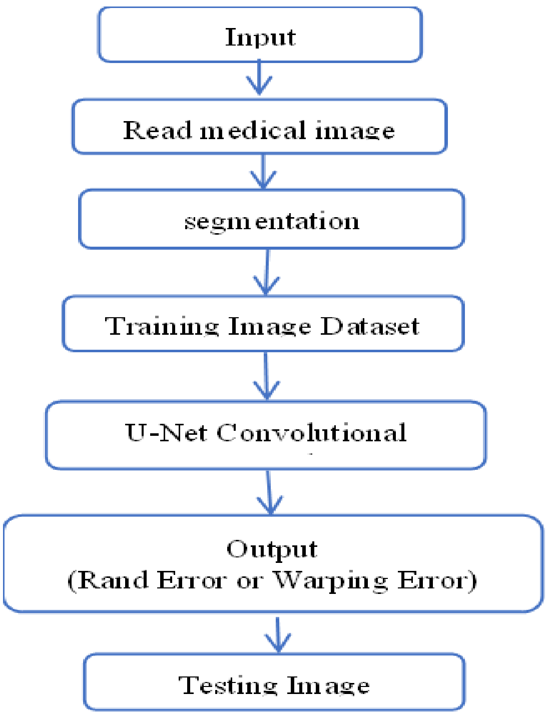

Fig. 1. U-Net network.

It is possible to invest several methods such as Graph-Cuts, GeoS and Random Walks but they use lower features and need to user markers more than usual. These methods succeed in applying them to the internal appearance and the ambiguous boundaries [13]. Machine learning tools have come to clarify User Maker. This network has increased the improvement of interactive segmentation because it has automatic learning of the features available in the images and as a result gives high performance such as the 3D U-Net whose inputs are index images to enter into the semi-automatic segmentation [14]. It is possible to invest the broad lines to define the CNN. Particularly in segmenting MRI images where Deep Cut uses the boundary box technique which provided by user as markers to train the neural network [15].

There are semi-continuous procedures that are not suitable for the testing phase, as they operate within a closed path constrained by annotated data for iterative improvement. When processing deep learning-based models, incoming images are converted into feature maps that serve as input to the convolutional neural network. In contrast, methods based on geodesic distance or spatial dimension encoding that do not employ neural networks tend to yield lower accuracy. Deep learning techniques, however, significantly enhance the precision of image segmentation. [16, 17]. graphical models are invested, for example, the use of strip reference fields. An algorithm is used to reduce the minimum algorithm for the maximum flow of the cut to reduce the energy of the Butch in the case of spatial organization to overcome the problem of the maximum flow. Maps are used to reach formulas like it, as it works on the consistency of collecting pixels with neighbors with similar characteristics [18]. In the case of including long-distance communication in the image segmentation, the CRF is used. CRFs model pairwise probability potentials between spatially adjacent pixels or regions, enabling the network to capture contextual relationships and refine segmentation boundaries. The possibility of adjacent edges is recognized by employing a straight Gaussian mixture. Some approaches utilize feature vectors to integrate the output features [19]. The primary objective is to achieve accurate image segmentation, ensuring the preservation of regional characteristics essential for effective feature extraction from medical images. Many deep learning algorithms are used, the most famous of which is the neural network. Convolutional Most research depends on building a model that can learn from and analyzes the medical characteristics. Obtaining the features can be done by entering the process of classifying the processed images. There are many algorithms used to classify and distinguish images, the most famous of which is the support vector machine, random forests and decision trees, Figure 2 show DeepCut Algorithm [20].

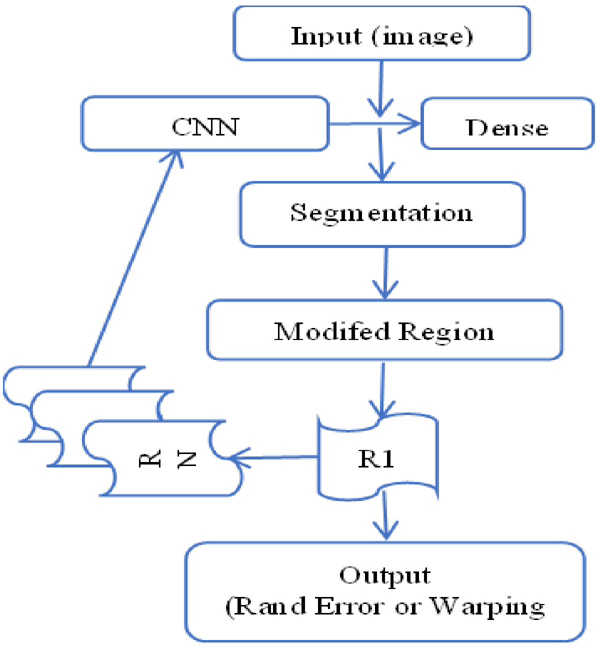

Fig. 2. DeepCut Algorithm

The basics of image processing for medical imaging are three stages: pre-processing, segmentation, feature extraction, and finally classification or discrimination. This paper presents an in-depth review of the techniques applied in medical image analysis, explaining the benefits and aspects of each one. The utilization of neural networks is one of the highly efficient techniques in analyzing and segmenting images. One of the first neural networks used to segment and analyze images is U-Net. It has a U-shaped structure, hence the name [21]. This network introduces a sliding-window model designed to perform pixel-wise classification, assigning a class label to each pixel. It processes 513 × 513 input images on a GPU in under one second, demonstrating high computational efficiency in pixel-level prediction. It has the ability to train a small number of images. His architecture can be memory intensive. The network architecture comprises two primary components: an encoder and a decoder. The encoder is responsible for progressively reducing the spatial dimensions of the input image through a series of convolutional and bottleneck operations, reaching the lowest level of the U-shaped structure. At this stage, feature maps are compressed to capture high-level semantic representations while maintaining computational efficiency. Despite employing a relatively small number of convolutional layers, the network effectively preserves critical contextual information necessary for accurate image reconstruction during the decoding phase. Decoding here returns the original image via deconvolution. It includes a loss function that trains each pixel such as dice loss or cross entropy. To know how well the segmentation is with the real images. U-Net is real time inference [22]. In 2016, Martin Rajchl et al proposed a technique called Deep-Cut. It relies on object segmentation method from the bounding-box annotation utilizing CNN. Aims to develop and extend the work of Grab-Cat to segment images and image deposits. This method solves problems related to low energy in random fields. The proposed algorithm consists of two stages, the first is known as label update and the second is model estimation. It uses GMM while Grab-Cut uses NNM neural network model. The connected graph with the density of images is best used when spreading CNN with the last iteration used. The characteristics of the used CNN are forward that include at least one level. Max Pooling tries to reduce the size of the input and this affects the learning process to represent objects in different environments. This is reflected in reducing the output to limited classes. The network structure includes two types of convolutional layers and pooling levels. These layers are connected in succession to other layers that contain dense neurons and by nature the structure of the network contains a level connected to the foreground and background. As a specific entropy between code distribution can represent the loose function. To enhancement network learning utilize the gaussian distribution intensity. This technique is better than others in segmentation accuracy of brain and lungs images [23].

In 2016, Joseph Redmon et al. presented a paper detailing an object detection model called YOLO (You Only Look Once). It stands out from its peers for its speed and high accuracy. It has undergone several improvements over the years to date. YOLO features a single network capable of analyzing, detecting, and distinguishing entire medical images. Yolo architecture employs a single convolutional neural network, that main task predicates object pounding boxes and their class probabilities within a medical image [24]. This algorithm has gained wide acceptance in medical image applications such as: breast-lesion recognition in mammography [25], lung nodule detection in chest scans [26], also prove its accuracy in brain tumor segmentation [27].

Another technique is Deep Interactive Geodesic Framework which was proposed in 2017 by Guotai wang et .al with the purpose of segmenting and analyzing medical images. The name comes from geospatial (shortest path between two points) and deep learning techniques to solve and simplify complex models by segmenting the image and tracking the boundaries of interest and modifying what is important. In this approach, it is implemented with user interaction as part of the optimization and guidance of the learning process [28]. This proposal uses P-NET as the initial network and R-NET as the next network. Both train the pixels in the image. P-NET usually has a number of input images and as a first step performs an initial automatic segmentation for them. This initial segmentation goes through a user analysis phase by adding scribbles, and is sent as a group to the R-Net. During this process, the P-Net as well as the R-Net maintain the image dimensions as much as possible by employing CRF (conditional-random-field) techniques. The R-Net and P-Net are subsequently learned using backpropagation. The image is divided into two groups: background pixels and shape pixels using scribbles and user interaction. This network uses the receptive fields to obtain high-level features [29]. Guotai wang et .al provided a clear comparison in addition to the best segmentation result between the Geodesic distance between the P-Net and R-Net CNN frameworks. Bellow figure 3 show Refinement Segmentation.

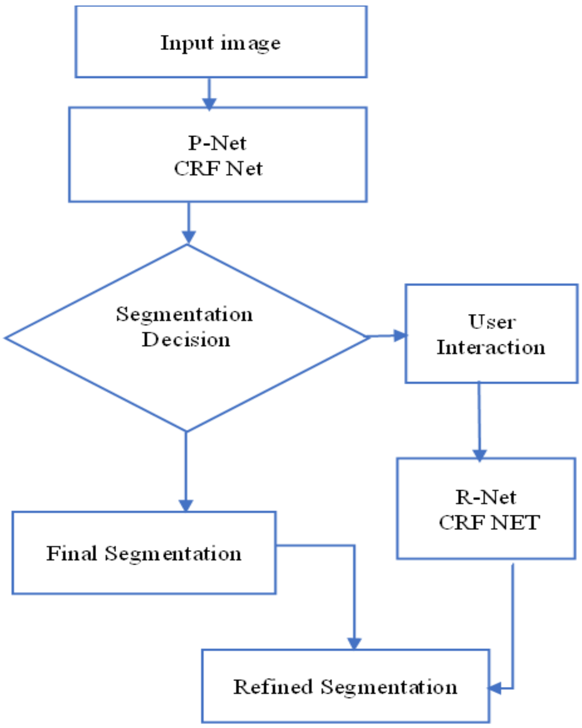

Fig. 3. Refinement Segmentation.

In 2018 guotai Wang et .al proposed a technique called IIS (Interactive Image Segmentation) Investing the Fine-tuning adapted to individual images. This algorithm uses the interactive segmentation method by CNN into the boundary box and scribble related segmentation investor. It is usually used to segment medical images for brain tumors. There has been a significant development in the way of segmenting images and dealing with invisible previous objects.The boundary box of the image used is the input of the CNN and provides a binary segmentation for that input. The boundary box is usually provided by the user at test time. Inside the boundary box, the network is drawn by the BIF segmentation model. The network is trained to interact with different parameters such as contrast and even deal with unseen objects [30].

Another technique as introduced by Deng-Ping Fan et al became known with the emergence of Covid-19. It is (Automatic COVID-19 Lung Infection Segmentation from CT Images). It presents an algorithm that detects pneumonia in cross-sectional images in a simple way. The proposal is called Inf-Net. Its task is to segment images, collect data, and scientifically train the network on the affected areas. The Inf-Net includes an encoding stage that works in parallel to extract high-level features to draw a special structure for them. As a basic step, the boundaries are defined and the representations are improved by the augmented and implicit edges. A semi-supervised scatter model is used to compensate for the lack of available data [31].

The segmentation process is randomly selected, does not necessitate many labeled images, and has the ability to augment unlabeled data. The network includes two convolutional layers that process high-resolution features. As an additional process, it calls low-level features to structure high-level features. In contrast, the decoding stage includes combining all the features to form the Inf-Net structure. It was originally designed from two different parts of the network, one is invested as an accurate season and the other is an approximate indicator. The network uses an adaptive learning method, which eliminates the need to collect features from all levels in parallel. It has undergone many improvements, including increasing the ability to absorb a large number of unlabeled images in order to obtain strong training data. This proposal included 150 axial tomography images of different cases

of Covid-19. When comparing the work with different works, it turned out to be the best at that time [32].
In 2019, Zhijie Zhang et al proposed a new model, "ET-Net: A Generic Edge-aTtention Guidance Network for Medical Image Segmentation". Generally, it utilizes edge detection and relies on it to assist the network. This network performs both encoding and decoding. This model is based on improving the detection of boundaries within medical images, enhancing the features to be used in deep learning processing to obtain accurate segmentation results. The details of the structures present in the image, represented by edges, were captured. This model also has the potential to be applied in multiple medical fields, such as MRI, CT scanners, and ultrasound. After comparing this technique with other techniques, it was found that it achieved an accuracy rate of 91 to 97%, taking into account the differences in image quality and number in the research [33].
Xiangxiang Qin presented research proposing the possibility of a deep neural network featuring multi-level edge attention to achieve greater accuracy in segmenting MRI images. It uses a three-dimensional encoder and decoder that are included in the training process of MRI images. The goal is to extract a high level of accurate information. Data refinement is applied online, reducing the number of image training sessions. It suffers from blurred boundaries due to the low contrast between the prostate and surrounding tissue, making it difficult to segment the prostate in 3D [34].
Justin Ker et al. proposed a method in 2018 that includes multiple convolutional neural networks for medical images. 3D and 2D structures can be analyzed for image recognition. CNNs are capable of handling a wide variety of image-level analysis tasks, such as localization detection, classification, segmentation, and registration [35]. In these tasks, the unique features of the images are preserved, despite the image dimensionality changes. The method focuses on detecting cancer cells. The method's advantages include rapid computation and training, as it reduces the number of calculations required [36]. Bellow figure 4 show classification-based CNN.

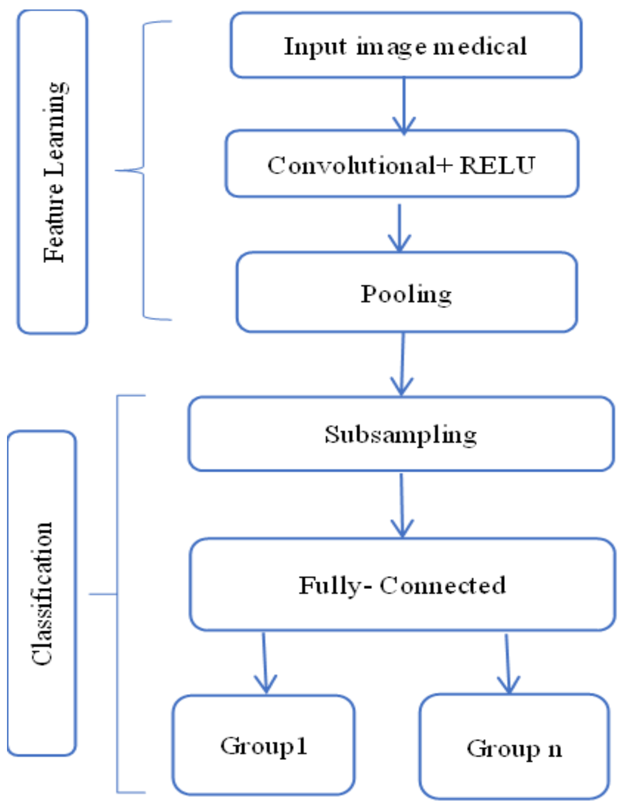

Fig. 4. Classification-based CNN.

Guotai Wang et al. proposed a method in 2020 that involved analyzing pneumonia scrapings from CT scans of COVID-19 patients to detect and diagnose pneumonia in lung images. The training process

includes the image's Dice loss function, along with a self-integration structure. Two-dimensional convolutional neural networks are used for segmentation. Cople-Net now passes through U-Net variables, which contain encoder and decoder structures [37].
In 2020, Trinh Le Ba Khanh et al. presented a study proposing to increase the capacity of the U-Net architecture to increase the segmentation obtained in the spatial channel with minimal computation. They increased efficiency by combining encoder and decoder. This method overcomes the shortcomings of conventional U-Nets. This gateway consists of four main parts as show in figure 5: a prediction module, a fusion module, an encoder module, and a decoder module. Usually, an image is fed into the encoder input, where the spatial channel specification gateway is multiplied by the encoder structure to produce features. This output is fed back into the prediction module, whose task is to identify abnormal tissues. The goal is to detect pathological regions that are subject to segmentation based on the presence and interconnectedness of these regions [38].

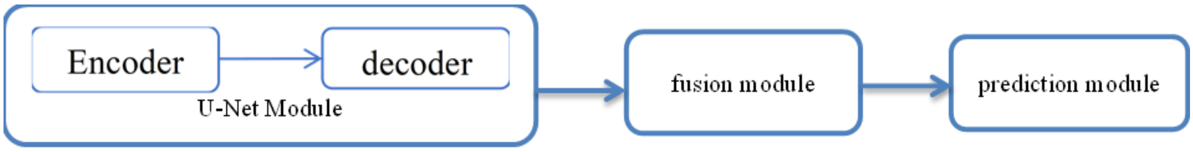

Fig. 5. The proposed Module of [38]

Shumao Pang et al. presented a proposal at the IEEE in 2021. The idea is based on segmenting volumetric MRI images of intervertebral disc sections, and it has proven effective in detecting two major types of spinal diseases and disorders. Initially, the proposal consists of two parts: the input images represent the first part, a two-dimensional U-Net, whereas the second section depicts a three-dimensional graph neural network. A particular spinal unit is represented by each node in the graph. Based on the spinal structure, the vertebral regions are grouped into a single graph, which is represented by a matrix [39]. The self-integrating framework incorporates a noise-specific Dice loss function to enhance the training process. To improve performance when processing noisy images, the model uses an adaptive instructor and an adaptable pupil. For segmentation, a 2D CNN is utilized. The suggested Cople-Net, which has an encoder-decoder structure, is based on variations of U-Net. It incorporates an intermediate layer between the encoder and decoder to align lower-level features and dual pooling to reduce informational loss. This lessens the discrepancy that results from mapping both high- and low-level features [40]. The encoder-decoder structure incorporates an Atrous Spatial Pyramid Pooling (ASPP) module to improve lesion segmentation, enabling better capture of multi-scale segmentation features. Additionally, Each convolutional block incorporates residual connections to make training easier. These connections are refined by a Squeeze-and-Excitation (SE) block, ensuring better channel-wise calibration for improved performance. Furthermore, a 1×1 convolution layer is used for upsampling, followed by bilinear interpolation. The Cople-Net was evaluated against four other networks, with the results presented in a approach. It was demonstrated in which Cople-Net attained the highest performance in segmentation [41].
Xier Chen et al. presented a proposal regarding the segmentation of object images using edge attention networks. This study focuses on instance segmentation via a mask-oriented technique called R-CNN. A fully convolutional box head is employed to detect branches and determine accurate bounding boxes for segmentation. Additionally, a supervised edge-specific module is integrated into the original mask [24]. It is anticipated that the suggested region will help SpinePraseNet generate a coarse segmentation by reflecting the graphical illustration of a semantic picture information. Finally, the segmentation is improved using the two-dimensional ResUNet. A multi-class segmentation task including the intervertebral and vertebral discs in spine pictures is known as "spine parsing." A 3-D graph convolutional segmentation network is used to overcome some of the constraints in the clinical processing of volumetric MR images of the spine. The following elements make up this network [43]:

- 3D Res U-Net
- 3D Graphonomy
- 3D U-Net
- 3D DeepLabv3+ Network

SparseNet is trained using a two-stage training approach. The input consists of 20×18×256×128 poorly populated probability maps and a 19×256×128 MR volume in order to attain 3D coarse segmentation. The output layer is subjected to the Softmax activation function. The Adam optimizer is used to train the 3D Res U-Net, 3D DeepLabv3+ Network, and 3D U-Net with a batch size of 4 for a maximum of 101 iterations. Precision and the Dice Similarity Coefficient (DSC) are utilized to compute the segmentation efficiency. Results indicate that SpineParseNet achieves mean segmentation accuracies of 90.01±6.45% for intervertebral discs (IVDs), 92.44±6.55% for vertebrae, and 88.08±6.20% for all 19 spinal structures combined. The corresponding average precision values are 88.3±4.54%, 87.56±4.82%, and 85.38±3.60%, indicating that false-positive rates in bifurcation scenarios are higher than false-negative rates. Additionally, SpineParseNet attains a mean Dice Similarity Coefficient of 86.33±3.66% for IVDs, 84.79±3.64% for vertebrae, and 88.50±2.81% for all 18 segmented spinal structures [44, 45].

Liming Zhong et al. presented a 2020 paper in the IEEE Journal emphasizing attenuation in MRI images. The method predicts pseudo-CT images from reference T-2-weighted MRI and T-1-weighted MRI images, where the data is fixed and content-specific to calculate the expected matrices. As a first step, the algorithm uses a regression-incentivized approach to localize the pseudo-CT images. The algorithm includes dataset optimization, translation of non-linear features to MRI images, nearest neighbor search, and optimization. MRI images undergo preprocessing, such as removing artifacts such as bias fields, intensity normalization, and spatial normalization. Several methods are used to predict CT images from data such as MRI. Recently, classical methods have been abandoned, and advanced approaches such as INAR have been adopted. These approaches do not require linear mapping between MRI and CT features. For example, if the stability is given, the surrounding training can likely be based on the equation [46].

It is worth noting that Zuhair Al-Amin presented a paper in 2020 at the Sixth International Engineering Conference on Sustainable Technology. The paper addresses statistical methods and statistical analysis of medical images and how to improve the quality of these images. Within the framework of these methods, there is a constant struggle with the quality of medical images due to medical imaging devices. To enhance processing and image quality, attention is turning to the use of statistical methods. The proposed algorithm processes a dataset of poor-quality images using a specific metric and without a reference. After comparing the work with other algorithms, it was found that it performs better using initial image processing alongside statistical methods. The algorithm's core function is to modify the image's contrast at local and global levels. The outputs of this algorithm are then combined using logarithmic image processing. The results are then processed using CST to modify local contrast [47]. The algorithm was compared with four other algorithms in terms of performance improvement, and this algorithm outperformed its peers. As a final result summarizing the most important types of CNN for analyzing medical images and their results.

3. Results and discussion

This study investigated various models developed for the analysis of medical images. A comparative evaluation of these methods was conducted based on key performance criteria, as summarized in Table 1.

Table 1. the most important types of CNN with performance details.

Method Used	Dataset	Dataset Source	Image Type	Accuracy	F1-Score	Dataset Size	Validation Method	Training Time	Ref.
ResNet101	Skin Cancer	ISIC Public Dataset	Clinical Photo	84.09%	0.82	2,000	80/20 Split	4 hrs (GPU)	[48]
2D-CNN	Brain Image	Private	MRI	94%	0.93	3,200	80/20 Split	5 hrs	[49]

Method Used	**Dataset**	**Dataset Source**	**Image Type**	**Accuracy**	**F1-Score**	**Dataset Size**	**Validation Method**	**Training Time**	**Ref.**
3D-CNN	Brain Images	Private Hospital	MRI	96%	0.95	3,500	5-fold CV	10 hrs	[50]
ID3-CNN	Heart Disease	Open HeartDB	CT	91%	0.90	4,200	70/30 Split	6 hrs	[51]
Capsule Network Model	Brain Tumor	BRATS Dataset	MRI	86.5%	0.87	2,800	10-fold CV	8 hrs	[52]
Transfer Learning	Retina	Kaggle Dataset	OCT	96.6%	0.96	6,000	5-fold CV	5 hrs	[53]
Fuzzy C-means	Brain Tumor	BRATS Dataset	MRI	98.6%	0.97	1,800	10-fold CV	7 hrs	[54]
Deep Fusion Feature	Lung Nodules	LIDC-IDRI	CT	96%	0.95	4,000	5-fold CV	9 hrs	[55]
DLM	Brain Tumor	BRATS Dataset	MRI	98%	0.97	2,500	10-fold CV	10 hrs	[56]
Auto Encoder	Lung Disease	LUNA Dataset	CT	93%	0.89	3,600	70/30 Split	6 hrs	[57]
GAN + CNN	Liver Lesion	Internal	CT	92.4%	0.91	3,000	External Validation	12 hrs	[58]

We obtained through this study the AI revolution has proven effective in detecting and treating diseases, particularly in the analysis of x-rays, MRIs, and CT scans of various tumors.

We note that most studies demonstrate their robustness in handling medical image data with high accuracy, especially MRI and CT scans, more than others. This emphasizes the importance of dataset quality and evaluation metrics, there is a relationship between the quality of performance and the amount of data used [48], [48],[49],[57].

In [50], [52] show emerging interest in 3D and capsule-based architectures for richer spatial understanding, though with computational challenges and time. [58] provides evidence that data augmentation can significantly improve CNN model performance, especially when labeled data are scarce. Among the determinants that were observed were high computational and data requirements, like in [50],[52] and explainability in standard CNNs [55], [56].

Based on these findings, the authors conclude that, out of the models they compared, the fuzzy C-means [54] approach performs the best. With just seven hours of training time, it maintained a respectable level of computational efficiency while achieving the highest accuracy (98.6%) and a robust F1-score (0.97). Additionally, its performance and generalizability are statistically validated through evaluation using 10-fold cross-validation. On the other hand, the Deep Learning Model (DLM) [56] needed more training time even though it had a similar F1-score (0.97) and a marginally lower accuracy (98%). Strong performance (96.6% accuracy, 0.96 F1-score) and adaptability across large datasets were also demonstrated by transfer learning techniques, indicating their potential for wider use. However, the Fuzzy C-means method can be regarded as the most efficient and reliable method based on accuracy, F1-score, and validation rigor.

4. Conclusion

This paper brings together several studies of convolutional networks and deep learning methods, demonstrating their use in segmenting and analyzing medical images. Medical image analysis can be thought of as a bridge between computer vision and medical concepts to diagnose diseases and make treatment decisions.

Recent research emphasizes that the effectiveness of medical image analysis is influenced by several critical factors, including the type and quality of datasets, the selection of an appropriate computer vision framework, and the design of a neural network architecture capable of constructing a robust model that balances accuracy and computational efficiency. Nonetheless, certain studies remain constrained by incomplete or insufficient data, which impedes comprehensive model interpretation and consequently reduces the reliability of the reported outcomes. To address these limitations, future work should prioritize the development of explainable deep learning methodologies aimed at creating transparent and interpretable diagnostic models for diverse medical imaging modalities, thereby enhancing clinical decision-making and supporting healthcare professionals.

References

[1] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017, doi: 10.1016/j.media.2017.07.005.

[2] M. Li, Y. Jiang, et al., "Medical image analysis using deep learning algorithms," *Frontiers in Public Health*, vol. 11, 2023, Art. no. 1273253, doi: 10.3389/fpubh.2023.1273253.

[3] Y. Gao, Y. Jiang, Y. Peng, et al., "Medical image segmentation: A comprehensive review of deep learning-based methods," *Tomography*, vol. 11, no. 5, 2025, Art. no. 0052, doi: 10.3390/tomography11050052.

[4] S. Pang, T. Thio, F. Siaw, et al., "Research on improved image segmentation algorithm based on GrabCut," *Electronics*, vol. 13, no. 20, 2024, Art. no. 4068, doi: 10.3390/electronics13204068.

[5] A. Tulbure, et al., "A review on modern defect detection models using DCNNs," *Journal of Advanced Research*, vol. 35, pp. 33–48, 2022, doi: 10.1016/j.jare.2021.03.015.

[6] N. Mathew, "Deep learning for image recognition: Enhancing accuracy and efficiency with big data," M.Sc. thesis, Northumbria Univ., Newcastle upon Tyne, U.K., 2024, doi: 10.13140/RG.2.2.16387.52004.

[7] S. Giri, S. Dutta, et al., "A real-time epidemic alert generation system for rural areas using WBANs and kiosks," in *Proc. Int. Conf. Information Technology (ICIT)*, 2017, doi: 10.1063/5.0113024.

[8] P. Mayerhofer, C. H. Napier, et al., "Neural networks can accurately identify individual runners from their foot kinematics, but fail to predict their running performance," *Journal of Biomechanics*, vol. 185, 2025, Art. no. 112663, doi: 10.1016/j.jbiomech.2025.112663.

[9] G. Montavon, W. Samek, et al., "Methods for interpreting and understanding deep neural networks," *Digital Signal Processing*, vol. 73, pp. 1–15, 2018, doi: 10.1016/j.dsp.2017.10.011.

[10] S. Dutta and M. Roy, "EDOT: Context-aware tracking of similar data patterns of patients for faster diagnoses," in *Proc. 2nd Int. Conf. Electrical, Computer and Communication Technologies (ICECCT)*, 2017, doi: 10.1109/ICECCT.2017.8117871.

[11] M. Taye, "Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions," *Computation*, vol. 11, no. 3, 2023, Art. no. 52, doi: 10.3390/computation11030052.

[12] J. Zhang, H. Wang, et al., “Deep network based on up-and-down blocks using wavelet transform and successive multiscale spatial attention for cloud detection,” *Remote Sensing of Environment*, vol. 261, 2021, Art. no. 112483, doi: 10.1016/j.rse.2021.112483.

[13] W. Gao, Q. Han, et al., “Graph random walk with feature-label space alignment: A multi-label feature selection method,” *arXiv preprint*, arXiv:2505.23228, 2025, doi: 10.48550/arXiv.2505.23228.

[14] A. Pinto, P. Mehta, S. Alle, et al., “DeepEdit: Deep editable learning for interactive segmentation of 3D medical images,” *arXiv preprint*, arXiv:2305.10655, 2023, doi: 10.48550/arXiv.2305.10655.

[15] O. Dorgham, M. Abu Naser, M. H. Ryalat, et al., “U-NetCTS: U-Net deep neural network for fully automatic segmentation of 3D CT DICOM volume,” *Journal of the Franklin Institute*, 2025, doi: 10.1016/j.smhl.2022.100304.

[16] G. Wang, Z. Zuluaga, W. Li, et al., “DeepIGeoS: A deep interactive geodesic framework for medical image segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 7, pp. 1559–1572, 2018, doi: 10.1109/TPAMI.2018.2840695.

[17] Y. Zhang, J. Chen, X. Ma, et al., “Interactive medical image annotation using improved attention U-Net with compound geodesic distance,” *Expert Systems with Applications*, vol. 237, 2024, Art. no. 121282, doi: 10.1016/j.eswa.2023.121282.

[18] S. Pawar, M. Bhushan, et al., “The plant leaf disease diagnosis and spectral data analysis using machine learning: A review,” *Int. J. Advanced Science and Technology*, vol. 29, 2020, doi: 10.2174/9789815049251122010008.

[19] S. H. Chen, Z. Gamechi, F. Dubost, et al., “An end-to-end approach to segmentation in medical images with CNN and posterior CRF,” *Medical Image Analysis*, vol. 76, Feb. 2022, Art. no. 102311, doi: 10.1016/j.media.2021.102311.

[20] A. Salman and W. Al-Jawher, “Image document classification prediction based on SVM and gradient boosting algorithms,” *Journal of Port Science Research*, vol. 6, no. 4, pp. 55–62, Dec. 2023, doi: 10.36371/port.2023.4.5.

[21] L. Liu, G. Cheng, Q. Quan, et al., “A survey on U-shaped networks in medical image segmentation,” *Neurocomputing*, vol. 409, pp. 244–258, Jun. 2020, doi: 10.1016/j.neucom.2020.05.070.

[22] Y. Xin, L. Sun, et al., “U-Net-based medical image segmentation,” *Journal of Healthcare Engineering*, 2022, Art. no. 4189781, doi: 10.1155/2022/4189781.

[23] M. Rajchl, M. Lee, O. Oktay, K. Kamnitsas, et al., “DeepCut: Object segmentation from bounding box annotations using convolutional neural networks,” *IEEE Trans. Med. Imaging*, vol. 36, no. 2, pp. 674–683, 2016, doi: 10.1109/TMI.2016.2621185.

[24] J. Redmon and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proc. IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[25] F. Prinzi, M. Insalaco, A. Orlando, et al., “A YOLO-based model for breast cancer detection in mammograms,” *Cognitive Computation*, vol. 16, pp. 107–120, 2023, doi: 10.1007/s12559-023-10189-6.

[26] K. Gunasekaran, “Leveraging object detection for the identification of lung cancer,” *arXiv preprint*, arXiv:2305.15813, 2023, doi: 10.48550/arXiv.2305.15813.

[27] H. Chegraoui, C. Philippe, V. Dangouloff-Ros, et al., “Object detection improves tumour segmentation in MR images of rare brain tumours,” *Cancers*, vol. 13, no. 23, 2021, Art. no. 6113, doi: 10.3390/cancers13236113.

[28] G. Wang, W. Li, M. Zuluaga, R. Pratt, et al., “Interactive medical image segmentation using deep learning with image-specific fine tuning,” *IEEE Trans. Med. Imaging*, vol. 37, no. 7, pp. 1562–1573, Jul. 2018, doi: 10.1109/TMI.2018.2791721.

[29] A. Plaksyvyi, M. Paszkowska, et al., “A comparative analysis of image segmentation using classical and deep learning approaches,” *Advances in Science and Technology Research Journal*, vol. 17, no. 6, pp. 127–139, 2023, doi: 10.12913/22998624/172771.
[30] G. Wang, W. Li, M. Zuluaga, et al., “Interactive medical image segmentation using deep learning with image-specific fine tuning,” *IEEE Trans. Med. Imaging*, vol. 37, no. 7, pp. 1562–1573, Jul. 2018, doi: 10.1109/TMI.2018.2791721.
[31] D. Fan, T. Zhou, G. Ji, et al., “Inf-Net: Automatic COVID-19 lung infection segmentation from CT images,” *IEEE Trans. Med. Imaging*, vol. 39, no. 8, pp. 2626–2637, Aug. 2020, doi: 10.1109/TMI.2020.2996645.
[32] Q. Dou, H. Chen, Y. Jin, et al., “3D deeply supervised network for automatic liver segmentation from CT volumes,” in *Proc. MICCAI*, Springer, 2016, doi: 10.48550/arXiv.1607.00582.
[33] Z. Zhang, H. Fu, H. Dai, et al., “ET-Net: A generic edge-attention guidance network for medical image segmentation,” *arXiv preprint*, arXiv:1907.10936, 2019, doi: 10.48550/arXiv.1907.10936.
[34] X. Qin, “Transfer learning with edge attention for prostate MRI segmentation,” Ph.D. dissertation, East China Univ. of Science and Technology, 2019, doi: 10.48550/arXiv.1912.09847.
[35] U. R. Ker, L. Wang, J. Rao, and T. Lim, “Deep learning applications in medical image analysis,” *IEEE Access*, vol. 5, pp. 142–151, 2017, doi: 10.1109/ACCESS.2017.2788044.
[36] L. D. Griffin, A. C. F. Colchester, S. A. Röll, and C. S. Studholme, “Hierarchical segmentation satisfying constraints,” in *Proc. British Machine Vision Conf.*, 1994, pp. 135–144, doi: 10.5244/C.8.13.
[37] G. Wang, X. Liu, C. Li, et al., “A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images,” *IEEE Trans. Med. Imaging*, vol. 39, no. 8, pp. 2653–2663, Aug. 2020, doi: 10.1109/TMI.2020.3000314.
[38] T. Khanh, D. Dao, et al., “Enhancing U-Net with spatial-channel attention gate for abnormal tissue segmentation in medical imaging,” *Applied Sciences*, vol. 10, no. 17, 2020, Art. no. 5729, doi: 10.3390/app10175729.
[39] S. Pang, C. Pang, et al., “SpineParseNet: Spine parsing for volumetric MR image by a two-stage segmentation framework with semantic image representation,” *IEEE Trans. Med. Imaging*, vol. 40, no. 1, pp. 262–273, 2021, doi: 10.1109/TMI.2020.3025087.
[40] Y. Ji, H. Zhang, et al., “CNN-based encoder-decoder networks for salient object detection: A comprehensive review and recent advances,” *Information Sciences*, vol. 546, pp. 835–857, Feb. 2021, doi: 10.1016/j.ins.2020.09.003.
[41] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
[42] X. Chen, Y. Lian, et al., “Supervised edge attention network for accurate image instance segmentation,” in *Proc. ECCV Workshops*, 2020, doi: 10.1007/978-3-030-58583-9_37.
[43] E. Hodneland, E. Hanson, A. Munthe-Kaas, et al., “Physical models for simulation and reconstruction of human tissue deformation fields in dynamic MRI,” *IEEE Trans. Biomed. Eng.*, vol. 63, no. 10, pp. 2203–2212, Oct. 2016, doi: 10.1109/TBME.2015.2514262.
[44] P. Chatterjee and S. Dutta, “A survey on techniques used in medical imaging processing,” *J. Physics: Conf. Series*, vol. 2089, no. 1, 2021, Art. no. 012013, doi: 10.1088/1742-6596/2089/1/012013.
[45] D. L. Collins, A. P. Zijdenbos, et al., “Design and construction of a realistic digital brain phantom,” *IEEE Trans. Med. Imaging*, vol. 17, no. 3, pp. 463–468, Jun. 1998, doi: 10.1109/42.712135.

[46] L. Zhong, Y. Chen, X. Zhang, et al., "Flexible prediction of CT images from MRI data through improved neighborhood anchored regression for PET attenuation correction," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 4, pp. 1114–1124, Apr. 2020, doi: 10.1109/JBHI.2019.2927368.

[47] Z. Al-Ameen, "Contrast enhancement of medical images using statistical methods with image processing concepts," in *Proc. 6th Int. Eng. Conf. Sustainable Technology and Development (IEC)*, 2020, pp. 169–173, doi: 10.1109/IEC49899.2020.9122925.

[48] B. Codella, K. Connah, et al., "Analysis of the ISIC image datasets, usage, benchmarks and recommendations," *Medical Image Analysis*, vol. 75, 2022, Art. no. 102305, doi: 10.1016/j.media.2021.102305.

[49] R. Asif, H. Maqsood, and I. Nadeem, "Enhanced MRI brain tumor detection using deep learning with explainable AI SHAP analysis," *Scientific Reports*, vol. 15, 2025, Art. no. 14901, doi: 10.1038/s41598-025-14901-4.

[50] J. Radhika, L. Marta, et al., "Three-dimensional end-to-end deep learning for brain MRI analysis," *arXiv preprint*, arXiv:2506.23916, 2025.

[51] O. Ebenezer and B. Prince, "Computer-aided diagnostics of heart disease risk prediction using boosting SVM," *Journal of Healthcare Engineering*, 2021, Art. no. 3152618, doi: 10.1155/2021/3152618.

[52] A. Parnian, "Capsule networks for brain tumor classification based on MRI images," in *Proc. IEEE ICASSP*, 2019, doi: 10.1109/ICASSP.2019.8683759.

[53] S. Bam, W. Alongbar, et al., "Transfer learning based detection of retina damage using OCT images," in *Computational Methods and Deep Learning for Ophthalmology*, pp. 71–88, 2023, doi: 10.1016/B978-0-323-95415-0.00002-4.

[54] D. Kumar and D. Satyanarayana, "MRI brain tumor detection using optimal possibilistic fuzzy C-means clustering and adaptive KNN," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 2867–2880, 2021, doi: 10.1007/s12652-020-02444-7.

[55] Z. H. Abed Al-Rahman Ali, "Automatic lung cancer classification in CT-scan images using CNN," M.Sc. thesis, Dept. Computer Science, College of Computer Science & Information Technology, University of Kerbala, Kerbala, Iraq, 2023.

[56] M. Mostafa, M. Zakariah, et al., "Brain tumor segmentation using deep learning on MRI images," *Diagnostics*, vol. 13, 2023, Art. no. 1562, doi: 10.3390/diagnostics13091562.

[57] Y. Nasr, R. Fernandez, et al., "Automated lung cancer diagnosis applying Butterworth filtering and sparse CNN to LUNA16 CT images," *Journal of Imaging*, vol. 10, 2024, Art. no. 168, doi: 10.3390/jimaging10070168.

[58] M. Fayad, I. Diamant, et al., "GAN-based synthetic medical image augmentation for CNN performance in liver lesion classification," *Neurocomputing*, vol. 294, pp. 118–127, 2018, doi: 10.1016/j.neucom.2018.09.013.

Journal of Basrah Researches (Sciences) **51**(2), 245 (2025)
DOI: https://doi.org/10.56714/bjrs.51.2.17

استطلاع رأي: تقنيات الذكاء الاصطناعي لتحليل الصور الطبية

سارة جاسم محمد* ، رشا إسماعيل احمد ، رشا محمد محسن

كلية علوم الحاسوب، الجامعة التكنولوجية، بغداد، العراق.

معلومات البحث

الاستلام	06 تشرين الأول 2025
المراجعة	19 تشرين ثاني 2025
القبول	25 تشرين ثاني 2025
النشر	31 كانون أول 2025

الكلمات المفتاحية

الشبكات العصبية التلافيفية، الرؤية الحاسوبية، تحليل الصور الطبية.

الملخص

يُعد تشخيص الأمراض واتخاذ القرار الصائب من أكثر المواضيع إثارة للجدل، لا سيما مرحلة تطور الأمراض التي تصيب البشرية وانتشار العديد من الأنواع المستعصية التي صعّبت على الطبيب اتخاذ القرار، كما هو الحال في الكشف عن الأمراض السرطانية وأورام الدماغ. يستكشف هذا الاستطلاع التطورات الحديثة في تقنيات الذكاء الاصطناعي المطبقة على وسائل التصوير الطبي مثل التصوير بالرنين المغناطيسي والتصوير المقطعي المحوسب والأشعة السينية. يمكن للتشخيص الآلي تسهيل العملية من خلال الاستفادة من جميع المتغيرات والأدلة للوصول إلى استنتاج سليم. على وجه الخصوص، أصبحت الشبكات العصبية التلافيفية (CNNs) ومتغيراتها حجر الزاوية في تحليل الصور الطبية الحديثة، مما يتيح تجزئة وتصنيفًا وكشفًا فعالًا للمناطق ذات الأهمية. معظم هذه الشبكات قادرة على تشريح الصور والتقاط المواقع ذات الأهمية عند تحليل الصور. تقدم هذه الدراسة عدة أنواع من الشبكات القادرة على تحليل أنواع مختلفة من الصور الطبية، مثل U-Net و-Res Net. تُسلّط هذه الورقة البحثية الضوء على التحديات الحالية، مثل اختلال توازن البيانات، وقابلية تفسير النماذج، والتعميم، وتُحدّد اتجاهات بحثية مستقبلية واعدة في مجال تحليل الصور الطبية المُعتمد على الذكاء الاصطناعي. وقد قدّمنا مقارنة بين هذه الأعمال من حيث مستوى الدقة، والسرعة، والتدريب، وأنواع مجموعات البيانات.

Citation: S. j. Mohammed et al., J. Basrah Res. (Sci.) **50**(2), 245 (2025). DOI:https://doi.org/10.56714/bjrs.51.2.17

***Corresponding author email:** Sarah.j.mohammed@uotechnology.edu.iq

ISSN: 1817-2695 (Print); 2411-524X (Online)
Online at: https://jou.jobrs.edu.iq