Deep Learning-Driven Medical Image Segmentation Using Generative Adversarial Networks and Conditional Neural Networks

Hayder H. O. Al-Ghanimi[1*], Ahmed H. O. Al-Ghanimi[2]

[1] Department of Medical Instrumentation Techniques Engineering, Hilla University College, Babylon 51001, Iraq
[2] Department of Computer Science, Faculty of Pharmacy, Babylon University, Babylon 51002, Iraq

Corresponding Author Email: hayder_alghanimi@hilla-unc.edu.iq

ABSTRACT

Medical image segmentation is a critical task in clinical diagnostics, playing a pivotal role in accurately identifying anatomical structures and pathological regions within medical images. Traditional methods often struggle to cope with the complexities and variability inherent in medical data, leading to suboptimal outcomes. This paper introduces a novel deep learning framework, DL-GAN-CNN, which leverages the strengths of Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs) to significantly enhance segmentation accuracy and efficiency. The primary objective of this research is to develop a robust and adaptable segmentation methodology that outperforms existing approaches. The proposed DL-GAN-CNN framework combines the powerful feature extraction capabilities of GANs with the precise classification strengths of CNNs, enabling the model to effectively handle the diverse challenges posed by medical imaging data. The study utilizes a comprehensive dataset of medical images, employing advanced pre-processing techniques, K-means clustering for initial segmentation, and the proposed GAN-CNN architecture for final segmentation and classification. Results from extensive evaluations demonstrate that the DL-GAN-CNN framework achieves remarkable performance, with an accuracy of 99.26%, precision of 99.81%, recall of 98.42%, and an F1-Score of 99.36%. These represent a much higher value than those obtained by traditional CNN and VGG16-CNN methods, showing that the framework is way better in segmenting and classifying medical images with high accuracy. These results reflect the potential of the DL-GAN-CNN framework to change the game in automatic medical image analysis by offering a trustworthy tool that could improve diagnostic accuracy and help with clinical decisions. In the end, it has introduced one very effective and novel approach for segmented medical image segmentation, setting an example for prospective studies. Some future work will be related to further optimization and applicability of the framework for the widest range of medical imaging modalities with regard to clinical utility and translating technologies into clinical practice that, in the overall context, shall improve patient outcomes.

1. INTRODUCTION

Brain tumors are one of the most dangerous and complex diseases, which, if not correctly diagnosed and treated, can cause severe neurological impairment or even death [1]. The complexity in diagnosing brain tumors does not lie only in the biological heterogeneity of tumors but also in the complexity of medical imaging, which is the basis for the correct diagnosis, treatment planning, and follow-up [2]. Magnetic Resonance Imaging (MRI) is the most commonly used imaging modality for brain tumor diagnosis due to its high spatial resolution and contrast differentiation [3]. However, the segmentation of tumor regions from MRI scans is a laborious process when done manually, and high expertise is needed; thus, it presents several errors. This calls for immediate attention to the need for automated and highly accurate medical image segmentation methods that can support radiologists in making more accurate and timely decisions [4].

In recent years, deep learning has brought a sea change to the medical imaging field, especially because of CNNs [5]. Because of its ability to learn hierarchical features directly from data itself, the CNN technique can have excellent performance from image classification to detection and segmentation [6]. Recently, CNNs have gained broad acceptance among the researchers for segmentations of medical imaging anatomical structures and pathological regions, thereby enhancing the speed and accuracy of traditional manual methods remarkably. However, despite such progress, standard CNN architectures often cannot handle medical image variability due to contrast variations, noise, or the presence of other types of artifacts, which can lead to inconsistent performances in segmentations [7].

To handle such challenges, this work proposes a new deep learning framework, namely DL-GAN-CNN, embedding

GAN into CNN to enhance the robustness and precision of segmentation in brain tumors using MRI scans. Key contributions of the work are listed below:

(1) Novel Architecture: The framework of DL-GAN-CNN utilizes GAN for generating high-quality feature representation to enhance segmentation accuracy by CNN. It is this combination in synergy that allows the model to handle the variability in medical images inherently.

(2) Improved Performance on Segmentation: By conducting extensive experiments, the suggested approach demonstrates increased precision, recall, and F1-Score values against classical CNN-based state-of-the-art methods. Improvement here bears great significance due to its close relation with less false positives and less false negatives at clinical diagnosis.

(3) Clinical Applicability: This proposed framework has been designed with clinical implementation in mind, hence giving emphasis not only to accuracy but also computational efficiency and adaptability to different imaging conditions. In this respect, the DL-GAN-CNN framework has proved to be quite practical in realistic medical scenarios.

This work addresses the lacuna in current research by combining the strengths of GANs and CNNs, not hitherto widely explored for medical image segmentation. Most traditional CNN-based approaches are indeed powerful but usually lack the robustness to perform consistently across diverse and challenging medical imaging datasets. Overcoming these limitations with the DL-GAN-CNN framework thus becomes a work of significant importance, latent with the promise to advance the current state of automated medical image analysis.

The rest of the paper is organized as follows:

Section 2 describes the related work in medical image segmentation using CNNs and GANs. Section 3 describes the problem statement. Section 4 provides the proposed DL-GAN-CNN framework by describing its architecture, data processing pipeline, and the training procedure. Section 5 presents the experimental results and performance evaluation of the proposed method over existing techniques. Section 6 discusses the implications of the findings, limitations, and possible future research directions. Finally, Section 7 concludes the paper by summarizing the contributions and significance of the proposed approach toward medical image segmentation.

While CNNs have been very successful in automating the segmentation process, limitations are usually viewed as inconsistent performance across diverse datasets. Such inconsistency may be related to inherent variability in medical images, which could be driven by differences in imaging modalities, contrast levels, and artifacts like noise. These include the facts that most CNNs failed to provide fine tumor boundaries for MRI with poor contrasts, hence leading to either over-segmentation or under-segmentation of the target region.

Besides, they require a great deal of high-quality-annotated datasets for training.

When tested using images from an unseen or unseen domain, their performances drastically fall. The same high-quality data and susceptibility to overfitting also weaken its generalizability and clinical applicability.

It is these limitations that underpin the requirement for developing an approach which would be more resistant to the variability and complexity of medical images while ensuring a high degree of accuracy and consistency of segmentation results.

2. RELATED WORKS

Mirunalini et al. [8] highlight that automatic skin lesion subtyping is essential for diagnosing skin cancer and assists medical experts. Deep learning is effective for image processing but needs refinement in dermoscopic images due to dataset imbalance and irrelevant image information. This study improves the pipeline by: 1) Balancing datasets using SMOTE and Reweighting to address class imbalance; 2) Adding a lesion segmentation stage with improved Bi-Directional ConvLSTM U-Net and conditional adversarial training; 3) Utilizing EfficientNets for classification, achieving 91% accuracy with EfficientNet B2 and 97% with B6. The proposed pipeline surpasses current methods on the ISIC dataset.

Qu et al. [9] explored the application of GANs in diagnosing Alzheimer's disease (AD), a condition that impairs daily living and requires early detection. Their review shows that GANs outperform other methods in AD state classification and related image processing tasks, such as image denoising and segmentation. Despite this, many studies relied on public datasets and lacked clinical validation, which could limit the model's generalizability and effectiveness in real-world scenarios. The paper also discusses the need for improved GAN architectures and emphasizes the importance of involving clinicians in future research. This involvement could help validate the models and enhance their clinical applicability. Overall, GANs show promise for advancing AD diagnostics, but further development and clinical testing are necessary to fully realize their potential.

Hu et al. [10] discusses the ability of GAN in solving data deficiency challenges in machine learning and medical image analysis. Though GANs have attained some improvements in the accuracy of classification regarding brain and liver imaging by data augmentation, no GAN has ever been developed hitherto that focuses on the prostate cancer analysis. So, this paper introduces a GAN-based model for the generation of realistic prostate diffusion imaging data: ProstateGAN. ProstateGAN effectively captures high-quality synthesis images, of suitable grades for application, based on the application of a Gleason score on conditioned deep convolutional GAN architecture.

Wang et al. [11] propose a new attempt, namely the Consistent Perception Generative Adversarial Network (CPGAN), to reduce the dependence on fully labeled samples in semi-supervised stroke lesion segmentation by effectively tackling the challenge of acquiring large-scale, manually labeled datasets with the help of a similarity connection module. It follows a consistent perception strategy to enhance the predictions of the model on unlabeled data. The auxiliary network will further help the discriminator discriminate and learn useful features that will ultimately help in segmentation. When tested on the Anatomical Tracings of Lesions After Stroke (ATLAS) dataset, CPGAN outperforms several previously proposed approaches by providing superior segmentation with only 40% of labeled data commonly required for fully supervised approaches.

Schellenberg et al. [12] address the challenges of supervised machine learning in PAT, where the lack of labeled reference data has been one of the major bottlenecks. The conventional approaches have been based on training with simulated data, but one of the most important challenges has remained how to bridge the domain gap between real and simulated images. To address this, the authors propose a new approach that

decomposes the image synthesis problem into two tasks: generating realistic morphology of tissues and assigning optical and acoustic properties to each pixel. Their approach harnesses GANs, which they have trained from semantically segmented medical images generating realistic structures of tissues in a probabilistic way and assigning relevant properties. Initial validation studies showed that this approach leads to more realistic synthetic images compared to conventional model-based methods. This can, therefore, boost deep learning-based quantitative PAT, commonly known as qPAT, toward being more robust and accurate for analyzing tissue properties.

Lei et al. [13] hence delves into deep learning methods in the analysis of medical images for the surmounting of part of these challenges, specifically those related to FA and OCTA techniques that could either be invasive, expensive, or applicable in only a limited manner. FA involves the use of dye injection, which can cause severe allergic reactions. In contrast, OCTA is non-invasive but expensive and has limited applications, such as within specific regions. Besides, noisy and low-quality data processing is problematic in techniques such as calcium imaging. The presented thesis deals with new methods and approaches that could enhance non-invasive screening and automate image processing. It focuses on the development of a shared feature manifold to improve data interpretation across different medical imaging modalities. The proposed architecture is based on generative networks using attention-based skip connections and novel residual blocks, including reconstruction, feature-matching, and perceptual loss in adversarial training. Successful anatomically correct fluorescein angiography image generation from fundus images, and segmented noisy calcium imaging maps with high accuracy, thereby overcoming limitations posed by the previous methods.

Chen et al. [14] review recent progress in deep learning for medical image processing, including its successful applications to disease detection and diagnosis. These successes are limited by a lack of large, well-annotated datasets. Their paper presents a comprehensive review of recent research addressing these challenges, with a particular focus on the latest developments in unsupervised and semi-supervised deep learning techniques. The authors summarize these developments in different application areas: classification, segmentation, detection, and image registration. Moreover, they present the major technical challenges of the field and point out some possible solutions that could lead future research efforts.

The existing literature has demonstrated the potential of GANs in medical image analysis; however, each of the approaches has varying strengths and weaknesses. For example, Qu et al. have shown the strength of GANs in classifying states of Alzheimer's disease by demonstrating their robustness on tasks such as image denoising and segmentation. However, one limitation of the authors' work is the reliance on publicly available datasets, which may not reflect the full spectrum of real clinical diversity and thus might reduce the general applicability of their models.

Similarly, Hu et al. [10] proposed ProstateGAN for handling data sparsity by generating realistic prostate images conditioned on specified cancer grades. Their work utilized a novel application of a conditional GAN but also did not cover how such synthesized data can actually be used in clinical workflows since the generated samples were not demonstrated to be utilized in downstream segmentation or classification

tasks.

Wang et al. [11] proposed a GAN architecture, Consistent Perception GAN (CPGAN), which achieved much better performance in stroke lesion segmentation tasks with small quantities of labeled data. This work creatively uses the similarity connection module to aggregate multi-scale features. However, there is still possible risk to some extent because this semi-supervised approach does not hold good in such sparse annotation settings as seen more often in medical imaging.

On the other side, Schellenberg et al. [12] have developed a GAN-based photoacoustic tomography by constructing a dual-task framework: generating probabilistically the structure of the tissues while associating with this structure their properties. The proposed model closes the gap between simulated and real images. However, semantically annotated data keeps scalability for larger applications within limits.

Collectively, these works underpin the wide versatility of GANs within medical imaging but underscore a few shortcomings in the area of generalization alone, integration with other techniques, and scalability. While in contrast, the proposed DL-GAN-CNN framework tries to mitigate these challenges by leveraging GANs for high-quality feature generation and effectively fusing them with CNNs for robust segmentation and classification. The proposed framework gains a kind of synergy in this respect that helps in the management of variability within medical imaging data with better ease, marking a significant advance beyond what has existed.

3. PROBLEM STATEMENT

Segmentation of brain tumors from MRI images is a very critical task with challenges in the realm of medical image processing. Traditional manual segmentation techniques are time-consuming, error-prone, and inconsistent because the procedure depends on human judgment, hence leading to inconsistencies that might affect diagnosis and treatment planning. Although CNNs have achieved great success in automating this segmentation process, most have difficulties dealing with the intrinsic variability of the medical images [15].

This variability includes the difference in the level of contrast that may obscure the tumor boundaries, noise, and artifacts due to either the imaging equipment or patient movement. There is also variability in imaging modalities, such as T1-weighted, T2-weighted, and FLAIR scans, each capturing different features of brain anatomy. Besides, there are also anatomical differences among patients, such as tumor size, shape, and location, which further complicate the segmentation task. All these challenges may result in suboptimal segmentations, over/under segmentation of ROIs, reducing the clinical reliability of CNN-based methodologies [16, 17].

Furthermore, most of the current methods are based on fixed architectures that may not generalize well to different datasets or imaging conditions, hence limiting their use in real-world medical applications. This therefore calls for a robust, flexible, and accurate segmentation method that can manage the variability and complexity of medical imaging with consistent high-quality results across various clinical scenarios.

In this respect, the proposed research work introduces the DL-GAN-CNN framework that integrates the generative benefits of GAN models with the discriminative advantages of CNNs for improving the accuracy of segmentation by making

it more invariant to variability for reliable and accurate diagnostic tools in healthcare.

4. METHODOLOGY

The work proposes a deep learning framework for medical image segmentation that capitalizes on the strong points of both GANs and CNNs. The DL-GAN-CNN approach will develop a technique that can take advantage of the generative strengths of GANs on high-quality segmentations, besides the discriminative strengths of CNNs, in order to ensure that the segmentations are indeed correct and contextually relevant. This work integrates the two networks to enhance both the precision and robustness of medical image segmentation tasks, which are important in clinical diagnostics and treatment planning.

4.1 Data collection

Brain tumor segmentation research chose the BraTS 2020 dataset for comprehensiveness and general acceptance as a standard. This BraTS 2020 dataset provides a high-quality, multimodal MRI dataset, including T1-weighted, post-contrast T1-weighted (T1Gd), T2-weighted, and FLAIR sequences [18]. The mentioned modalities are capable of offering an important complementarity of representation concerning the brain anatomy and pathology. That makes this data very relevant for the training and testing of deep learning models within the framework of brain tumor segmentation.

Several factors make BraTS 2020 unique compared to other medical imaging datasets in common usage. First, the annotations are provided by several expert radiologists, which guarantees high reliability and accuracy in the ground truth segmentation masks 19. Second, this dataset encompasses a wide variety of glioma cases, both HGG and LGG. This diversity ensures BraTS-trained models see a wide variety of tumor shapes, sizes, and locations to enrich their generalizability. Third, BraTS remedies one shortcoming common to many such datasets: the images are preprocessed, reducing the burden of preprocessing by a researcher.

Other medical image datasets, such as lung or liver imaging, are more limited either by the variety of modality representation or annotation quality. For example, while datasets such as LIDC-IDRI for lung cancer are very extensive, they lack the multimodal nature that is necessary for capturing the heterogeneity of brain tumors. BraTS 2020 represents a very well-matched balance between variability in the data, quality of annotation, and readiness for preprocessing. The BraTS 2020 dataset is therefore the one that best served the aim of this study. Its application ensures the proposed DL-GAN-CNN framework is trained and tested on a more robust and challenging dataset for maximum relevance to practical clinical applications.

4.2 Preprocessing: Histogram Equalization (HE)

Histogram Equalization is a general preprocessing technique used to enhance the contrast in grayscale images by redistributing the intensity values. Herein, HE was performed on MRI scans in order to emphasize critical features that will help improve the segmentation process. However, this may behave differently for other image modalities of medical images, and its application shall be carefully designed for optimal performance.

For instance, in T1-weighted MRI images, which provide outstanding anatomical definition, HE highlights the demarcation of tumoral margins by increasing contrast between tumor tissue and the remaining tissues. Accordingly, in other sequences like T2-weighted images and FLAIR, fluid content and oedema are very well outlined-HE tends to enhance the depiction of hyper-intense regions to point out matching pathological areas. However, for over-enhanced HE, not only signal but also amplification of artefacts or noise, when the modality signal-to-noise ratio has already been poor, has been noted.

Given these differences, modality-specific preprocessing strategies may yield better results. For example, while HE is effective for contrast enhancement in most MRI modalities, its application might be supplemented or replaced by other techniques in certain scenarios. Adaptive Histogram Equalization (AHE) or contrast-limited AHE (CLAHE) can be used for localized contrast enhancement, especially in modalities like FLAIR, where uniform enhancement might lead to overemphasis on non-pathological regions. Similarly, noise reduction techniques such as Gaussian filtering or wavelet-based denoising may be required in conjunction with HE to address modality-specific artifacts [19-21].

The choice of preprocessing strategy thus depends on the specific characteristics of each modality and the features of interest in the downstream analysis. While this study employs, HE as a generalized enhancement technique, future work could explore the integration of modality-adaptive preprocessing pipelines to further optimize segmentation performance. Figure 1 shows the equalized histogram.
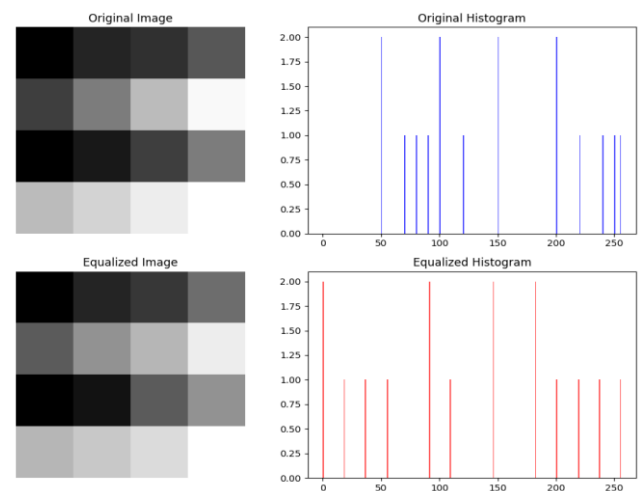

Figure 1. Equalized histogram

The first step involves calculating the histogram of the image, which is a graphical representation showing the distribution of intensity values (gray levels) across the image pixels.

The histogram is then normalized, meaning that the sum of all histogram values is scaled to 1. This is done to ensure that the histogram represents a probability distribution.

$$p(i) = \frac{h(i)}{n} \tag{1}$$

where, $p(i)$ is the formalized histogram value for gray level i, $h(i)$ is the histogram count for gray level i, and n is the total

number of pixels in the image.

Next, the cumulative distribution function (CDF) is computed, which is the cumulative sum of the normalized histogram. The CDF provides a mapping function that will be used to adjust the pixel intensity levels.

$$\text{CDF}(i) = \sum_{j=0}^{i} p(j) \qquad (2)$$

The final image is obtained by mapping the original pixel intensities using the CDF. Each pixel intensity $I\text{original}$ is transformed to a new intensity $I\text{new}$ according to the formula:

$$I_{\text{new}} = \text{round}\left(\text{CDF}(I_{\text{original}}) \times (L-1)\right) \qquad (3)$$

where, L is the number of possible intensity levels (for an 8-bit image, $L=256$).

4.3 Segmentation using K-means clustering

Segmentation is thus one of the important steps in recovering, classifying, and object recognition of any image. This segmentation is done once the preprocessing of the images is done as it is usually necessary to have regions of interest isolated from those images. Among the proposed approaches, K-means clustering identifies and groups similar regions of the image so that further processes can be done on the image, which enhances the precision of the whole analysis.

K-means clustering is a type of unsupervised learning algorithm that segregates the image data into a certain number of well-separated clusters based on similarity in pixel attributes, such as intensity or color. The major problem solved by K-means clustering involves classifying each of the data points, corresponding to pixels in this case, into k clusters. Every data point falls into the cluster of the closest mean, often called the centroid.

Selection of K Value:
The choice of k is very crucial to the effectiveness of the K-means algorithm. Herein, the determination of k is based on the Elbow Method, a common method used to identify the number of clusters, k. Basically, the Elbow Method runs K-means for a range of k values and plots the within-cluster sum of squares against the number of clusters. This is confirmed by the well-known "elbow" point, a point that characterizes the dramatic change in rate of WCSS decrease, indicating where the number of clusters will likely be optimal; in this paper, that value is at $k=3$; this corresponded to the three key sections of interest within our brain MRI healthy tissue, tumor core, and edema. This has been chosen to make sure segmentation captures distinct areas within the images.

The algorithm starts by initializing k centroids, which can be chosen randomly or by using special methods such as the K-means++ algorithm. These centroids represent the initial centers of the clusters. Every pixel $P(xi,yi)$ of the image is assigned to the nearest centroid, depending on the Euclidean distance between the pixel and each centroid. The metric used is:

$$d(P(x_i, y_i), C_j) = \sqrt{\left(x_i - C_{x_j}\right)^2 + \left(y_i - C_{y_j}\right)^2} \qquad (4)$$

where, $P(x_i, y_i)$ represents the pixel coordinates, and $C_j(C_{xj}, C_{yj})$

denotes the coordinates of the j-th centroid.

Once all pixels have been assigned to the nearest centroid, recalculation is done by averaging all pixels of each cluster to get a value for centroids. This step gives an appropriate position for every centroid at the mean of the particular cluster it falls under, hence making sure the representation given is very clear and close to precision.

$$C_j = \frac{1}{|S_j|} \sum_{P(x_i, y_i) \in S_j} P(x_i, y_i) \qquad (5)$$

where, S_j is the set of pixels assigned to cluster j, and $|S_j|$ is the number of pixels in cluster j.

The process of assigning pixels to the nearest centroids and updating the centroids is repeated iteratively, either until the centroids' positions stabilize-meaning their positions are no longer changing significantly-or a certain number of iterations is completed. It is this process of iteration that helps in achieving optimal clustering. This will ensure that the clusters formed are stable and really represent distinct regions within the image. The resulting clusters are then examined to refine the regions of interest. These can be further processed by converting them into an RGB format or by analyzing color and intensity ratios to enhance the segmentation accuracy.

4.4 Feature extraction using GANs

Feature extraction is an essential part in image analysis, for example, in medical image processing, which will be discussed next. Its goal is the identification and extraction of relevant information that may contribute to a diagnosis or additional analysis. For the purpose of this work, feature extraction was conducted using GANs. High-quality features learned by the powerful capability of the GAN model, enabling high complexity data distributions to increase segmentation, were the subject of study [21]. A GAN consists of two neural networks: The Generator and the Discriminator. These are trained simultaneously within a competitive framework. While the Generator aims to generate realistic data similar to the true data distribution, the Discriminator aims to differentiate between real data from the actual dataset and fake data generated by the Generator [22].

4.4.1 Generator Network (G)
The Generator is designed to generate high-dimensional feature maps from input images containing essential information for accurate segmentation. It takes the initial segmentation map, for example, obtained by K-means clustering, and the original preprocessed image as input and generates refined feature maps highlighting regions of interest such as tumor boundaries or anatomical structures. The architecture of the Generator is as follows:
1. Input layer:
-   The input to the Generator consists of two components:
    o   The initial segmentation map (binary mask) obtained from K-means clustering.
    o   The original preprocessed medical image (grayscale or multimodal).
-   These inputs are concatenated and passed through the network.
2. Convolutional layers:
    The Generator employs a series of transposed convolutional layers (also known as deconvolutional layers) to upsample the

input and generate high-resolution feature maps.

Each convolutional layer is followed by batch normalization and a ReLU activation function to stabilize training and introduce non-linearity.

The number of filters in the convolutional layers increases progressively to capture more complex features. For example, the first layer may have 64 filters, while deeper layers may have 128 or 256 filters.

3. Skip connections:

To preserve fine-grained details and improve feature extraction, the Generator incorporates skip connections between corresponding layers. These connections help mitigate the loss of spatial information during upsampling.

4. Output layer:
   - The final layer of the Generator uses a tanh activation function to produce feature maps that are normalized to the range $[-1,1]$.
   - The output is a refined feature map that highlights regions of interest, such as tumor boundaries or anatomical structures.

The output of the Generator can be expressed as:

$$G(z \mid c) = F_{\text{gen}}(I) \tag{6}$$

where, z represents random noise or input features, c denotes conditional information such as the initial segmentation map, and $F_{gen}(I)$ represents the generated feature map from the image I.

4.4.2 Discriminator Network (D)

The Discriminator evaluates the authenticity of the feature maps produced by the Generator. It learns to differentiate between feature maps generated by the Generator and those derived from real annotated medical images. The architecture of the Discriminator is as follows:

1. Input layer:
   - The input to the Discriminator consists of either.
   - Real feature maps from the annotated medical images.
   - Fake feature maps generated by the Generator.
   - These inputs are passed through the network for classification.
2. Convolutional layers:
   - The Discriminator employs a series of convolutional layers to downsample the input and extract hierarchical features.
   - Each convolutional layer is followed by batch normalization and a LeakyReLU activation function (with a slope of 0.2) to introduce non-linearity and prevent vanishing gradients.
   - The number of filters in the convolutional layers decreases progressively, starting with 64 filters in the first layer and reducing to 32 or 16 filters in deeper layers.
3. Fully connected layers:
   - After the convolutional layers, the feature maps are flattened and passed through one or more fully connected layers to perform binary classification.
   - The final layer uses a sigmoid activation function to output a probability score indicating whether the input feature map is real or fake.

The Discriminator's task is mathematically represented by:

$$D(x \mid c) = \sigma(W_d \cdot x + b_d) \tag{7}$$

where, x represents the input feature map, W_d and b_d are the weights and biases of the Discriminator, and σ is the sigmoid activation function that outputs the probability of the input being real.

4.4.3 Loss function

The GAN is trained using a min-max loss function, where the Generator tries to minimize the loss, and the Discriminator tries to maximize it. This is formally expressed as:

$$\mathcal{L}_{\underline{\text{GAN}}} = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x \mid c)] \\ + \mathbb{E}_{z \sim p_z(z)}\left[\log\left(1 - D(G(z \mid c))\right)\right] \tag{8}$$

where, $p_{\text{data}}(x)$ represents the distribution of real data, and $p_z(z)$ represents the distribution of the input noise or feature vectors. The architecture of the GAN is illustrated in Figure 2, with a detailed description of its components provided below.

4.4.4 Training process

Generator and Discriminator are trained iteratively in a competitive manner:

1. The Discriminator is trained to correctly classify real and fake feature maps.
2. The Generator is trained to produce feature maps that "fool" the Discriminator into classifying them as real.
3. This adversarial training process continues until the Generator produces feature maps that are indistinguishable from real ones, and the Discriminator achieves a classification accuracy close to 50%.

The architecture of the GAN used for feature extraction can be described as follows:
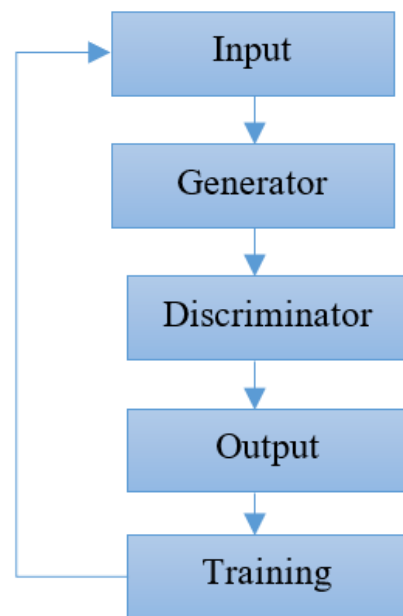

Figure 2. Architecture diagram

The proposed technique leverages the GAN feature extraction approach and hence effectively catches intricate details of medical images, which improves segmentation both accurately and reliably. One of the key reasons why GAN has been considered one of the most ideal choices in complex medical image analysis is because, through adversarial training, features extracted will be not only realistic but also highly relevant to the particular segmentation task.

4.5 Classification using CNN

CNNs represent a highly effective category in the domain of deep learning methods, especially adept at image classification tasks. These networks are engineered to automatically and adaptively learn spatial hierarchies of features through backpropagation. They achieve this by utilizing several key components, including convolutional layers, pooling layers, and fully connected layers [23]. Each of these components plays a critical role in the task of medical image segmentation, as described in Figure 3 depicts the CNN architecture.
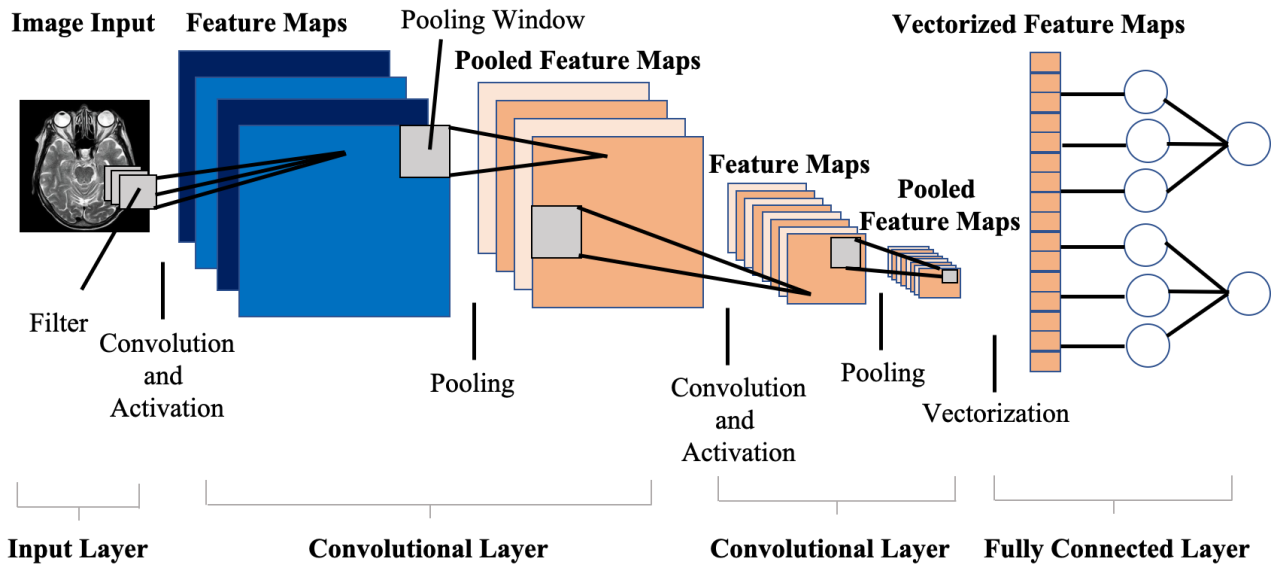

Figure 3. CNN architecture [24]

4.5.1 Convolutional layer

The convolutional layer is the core building block of a CNN. It applies a set of filters (also known as kernels) across the input image to produce a feature map. Each filter is convolved with the input image to detect various features like edges, textures, or more complex patterns. The operation in a convolutional layer is defined as:

$$F_{i,j,k} = \sum_{m=1}^{M} \sum_{n=1}^{N} \left(I_{i+m-1,j+n-1} \times K_{m,n,k}\right) + b_k \quad (9)$$

where, $F_{i,j,k}$ is the output feature map at position (i, j) for the k-th filter. I is he input image or previous layer's output. $K_{m, n, k}$ is the kernel/filter of ze $M{\times}N$ applied to the input. b_k is the bias term for the k-th filter.

4.5.2 Pooling layer

The pooling layer performs downsampling on the feature maps produced by the convolutional layers. It decreases the spatial dimensions (height and width) while preserving the most significant features. The most common pooling operation is max pooling, defined as:

$$P_{i,j,k} = max_{\,m.n}(F_{i+m,j+n,k}) \quad (10)$$

where, Pi,j,k is the pooled output at position (i, j) for the k-th feature map, and max is the maximum value in the pooling window.

Role in Medical Image Segmentation:
- Dimensionality Reduction: Pooling layers reduce the spatial size of the feature maps, which decreases the computational complexity of the network and helps prevent overfitting.
- Feature Preservation: By retaining the most salient

features (e.g., the highest activation values), pooling layers ensure that important information, such as tumor boundaries or anatomical structures, is preserved while discarding less relevant details.

Advantages:
- Robustness to Noise: Pooling layers make the network more robust to small variations or noise in the input images, which is common in medical imaging due to artifacts or low-quality scans.
- Spatial Hierarchies: Pooling helps the network build spatial hierarchies by progressively reducing the resolution of feature maps, allowing the network to focus on larger regions of the image in deeper layers.

4.5.3 Fully connected layer

This layer, also known as the dense layer, is where the high-level reasoning in the neural network occurs. The output from the final pooling layer is flattened into a 1D vector and passed through one or more fully connected layers, which combine all the features learned by the previous layers. The operation in a fully connected layer can be expressed as:

$$Z = W \cdot X + b \quad (11)$$

where, Z is the output vector, W is the weight matrix, X is the input vector from the previous layer, and b is the bias vector.

Role in Medical Image Segmentation:
- Feature Integration: Fully connected layers integrate the features extracted through convolution and pooling to give the final decision on the class of an image. For example, in medical image segmentation, it may involve categorizing each pixel as belonging to healthy tissue, tumor core.
- Decision Making: Fully connected layers take the joined features and predict a probability distribution

over the classes, where the class with the highest probability is the predicted class of the input image.

Advantages:

- Global Context: Fully connected layers consider the entirety of the feature set, hence making decisions that are based on the global context of the image. This aspect is quite vital for medical image segmentation, wherein classification of a pixel depends on its relationship to the surrounding pixels.
- *Flexibility*: The fully connected layers can be easily adapted for different classification tasks by changing the number of output neurons, hence being versatile in medical imaging for different applications.

The introduction of non-linearity in the network is brought about by activation functions, hence allowing the network to find complicated patterns in data. The following are some of the popular activation functions that are used in CNNs: ReLU - leaky ReLU, and finally, sigmoid.

Role in Medical Image Segmentation:

- Non-Linearity: The most important role of the activation functions is to give the network the ability to model non-linear relationships in the data, an aspect that has proven quite central to capturing complicated patterns in images generally, not to mention medical ones.
- Feature Selection: There are using activation functions such as ReLU, helping set negative activations to zero for the network to make the segmentations with focus on the most relevant features of interest. It probably brings several improvements in enforcing segmentation efficiency but also accuracy.

Advantages:

- Improved Convergence: These variants of ReLU have indeed facilitated faster convergence of the network during training by reducing the problem of the vanishing gradient.
- Sparsity: ReLU introduces sparsity in the activations, which can lead to more efficient feature representations and better generalization.

The combination of convolutional layers, pooling layers, fully connected layers, and activation functions makes CNNs highly effective for medical image segmentation. Each component plays a specific role in extracting, preserving, and integrating features, while also providing advantages such as translation invariance, robustness to noise, and the ability to model complex patterns. These properties make CNNs well-suited for the challenges of medical imaging, where accuracy and reliability are critical.

The algorithm for the proposed DL-GAN-CNN is outlined below, and the flow chart for the proposed methodology is illustrated in Figure 4.

The performance, efficiency, and robustness of the proposed DL-GAN-CNN framework are highly dependent on the careful selection and configuration of parameters such as learning rates, batch sizes, network architectures, and other hyperparameters. Fine-tuning these parameters can significantly enhance the algorithm's ability to accurately segment and classify medical images, adapt to variations in image quality, and respond to specific clinical scenarios. Moreover, the choice of parameters directly impacts the computational resources required, influencing the algorithm's practicality for real-world clinical applications. Therefore, additional research and fine-tuning of these parameters are essential to ensure that the DL-GAN-CNN algorithm delivers optimal performance and is effectively suited for automated medical image segmentation and classification tasks in clinical settings.
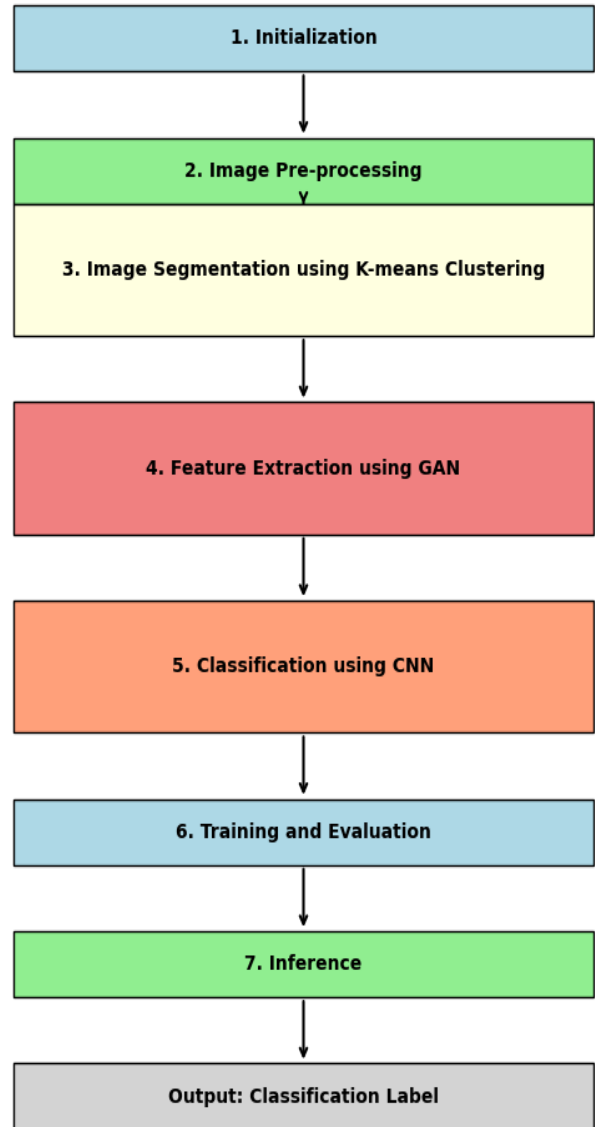

Figure 4. Flow chart for the proposed model [25]

Algorithm DL- GAN -CNN

Input: Medical image of brain
Output: Classification of medical image
Initialization
Begin by loading a dataset of medical images, such as brain MRI scans.
Set up the parameters for the GAN and CNN models, such as learning rates and the number of layers.
Image Pre-processing
Enhance the quality of the input images using Histogram Equalization, a technique that adjusts the contrast of the images to make features more distinguishable.
The pre-processed image will have improved contrast, making it easier to identify important regions in subsequent steps.
Image Segmentation using K-means Clustering
Apply the K-means algorithm to segment the pre-processed image into different regions based on pixel similarities.
First, the algorithm identifies initial cluster centers

randomly.

Then, it calculates the distance between each pixel and these centers, assigning each pixel to the nearest center.

The cluster centers are updated based on the new groupings of pixels, and this process repeats until the clusters stabilize.

The output is a segmented image where different regions, potentially corresponding to different tissue types or abnormalities, are clearly distinguished.

Feature Extraction using GAN

Implement a GAN to refine the features extracted from the segmented image.

The GAN consists of two parts: a Generator that tries to create realistic feature maps from the segmented image, and a Discriminator that attempts to differentiate between these generated feature maps and the actual feature maps from the original image.

Through training, the Generator learns to produce highly realistic feature maps that closely resemble those from the real images, helping to highlight the most important features for classification.

Classification using CNN

Implement a Convolutional Neural Network (CNN) to classify the extracted features into different categories (e.g., types of brain tumors).

The CNN processes the feature maps through a series of layers:

Convolutional layers detect various patterns and features in the images.

Pooling layers reduce the size of the feature maps, making the model more efficient and focusing on the most important features.

Fully connected layers take the high-level features detected by the convolutional layers and use them to make a final decision about the classification.

The final layer of the CNN outputs a probability distribution over the possible classes, with the highest probability indicating the predicted class for the input image.

Training and Evaluation

Train the entire DL-GAN-CNN framework by iteratively adjusting the weights of both the GAN and CNN using backpropagation.

During training, monitor the loss functions and classification accuracy to ensure that the model is learning effectively.

After training, evaluate the model's performance on a separate validation set to ensure it generalizes well to new data.

Inference

When a new, unseen medical image is fed into the trained DL-GAN-CNN framework, the model automatically segments the image, extracts key features using the GAN, and classifies the image using the CNN.

The output is a classification label that indicates, for example, the type of brain tumor present in the image.

5. RESULTS AND DISCUSSIONS

The proposed auto medical image segmentation framework, essentially DL-GAN-CNN, is designed to fuse deep learning techniques, GANs, and CNN, yielding highly accurate segmentation results. This model promises considerable improvements compared to traditional methods with respect to enhancing segmentation metrics in terms of precision,

accuracy, F1 score, and recall. With this model, while leveraging the strengths of GANs in generating realistic feature representations, it showed a robust adaptation concerning variations in image quality, which may emanate from different contrasts, illumination, or even noise. It segmented the data efficiently: fast segmentation speed with high accuracy. Its generalization ability is also outstanding-it can be adapted to whatever kind of anatomical structure or all kinds of imaging modalities. These results confirm that the proposed DL-GAN-CNN methodology constitutes a robust and effective technique for the automatic segmentation of medical images. Further, these techniques provide much better enhancement than those in the existing state of the art.

5.1 Performance metrics evaluation

Precision, accuracy, F1 score, and recall are widely used for this task of estimating a classification model's performance in such tasks as image segmentation.

(1) Precision is the ratio of true positive predictions out of all positive predictions made by the model. It gives the measure of how accurate the model's positive predictions are.

$$Precision = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \quad (12)$$

(2) Accuracy refers to the proportion of correctly predicted instances that includes both true positives and true negatives in relation to the total number of instances. It is indicative of the general efficacy of the model in correctly making predictions.

$$Accuracy = \frac{True\ Positives\ (TP) + True\ Negatives\ (TN)}{Total\ Population\ (TP + TN + FP + FN)} \quad (13)$$

(3) Recall, also referred to as Sensitivity, is the ratio of positive predictions actually made compared to the total number of actual positives. It reflects how well the model identifies positive instances.

$$Accuracy = \frac{True\ Positives\ (TP) + True\ Negatives\ (TN)}{Total\ Population\ (TP + TN + FP + FN)} \quad (14)$$

(4) F1 Score is the harmonic mean of precision and recall, providing a balanced metric that combines both measures into a single value. This score is particularly useful for evaluating performance on imbalanced datasets, where one class might be significantly more prevalent than the other.

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (15)$$

These metrics together give a full view of the performance of a model, balancing precision and the ability to correctly identify positive instances while avoiding false positives.

Based on the results depicted in Table 1 and Figure 5, it can be viewed that the proposed algorithm DL-GAN-CNN outperforms many of the other approaches by offering high accuracy in classification. Therefore, considering the two state-of-the-art approaches, such as CNN [25] and VGG16-CNN against the said two datasets, for the classification

problem it has been found that the proposed one is capable of assuring higher classification accuracy, which is nearly about 99.91%.

Table 1. Accuracy comparison of different methods

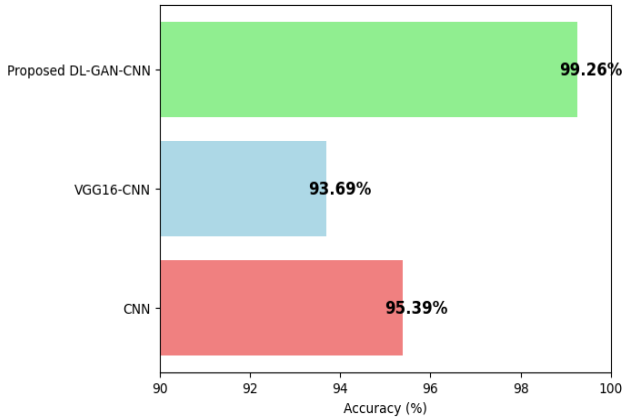| Method	Accuracy (%)
CNN	95.39
VGG16-CNN	93.69
Proposed Method	99.26

Figure 5. Accuracy comparison of different methods

Figure 5 visually illustrates this comparison, demonstrating the superior performance of the proposed algorithm in medical image segmentation tasks.

From the results in Table 2 and Figure 6, it can be noted that the proposed DL-GAN-CNN algorithm is superior when compared to other approaches on precision. The value of the precision parameter of the proposed approach compared to two other methods, CNN and VGG16-CNN, stands higher at about 99.26%.

Table 2. Precision comparison of different methods

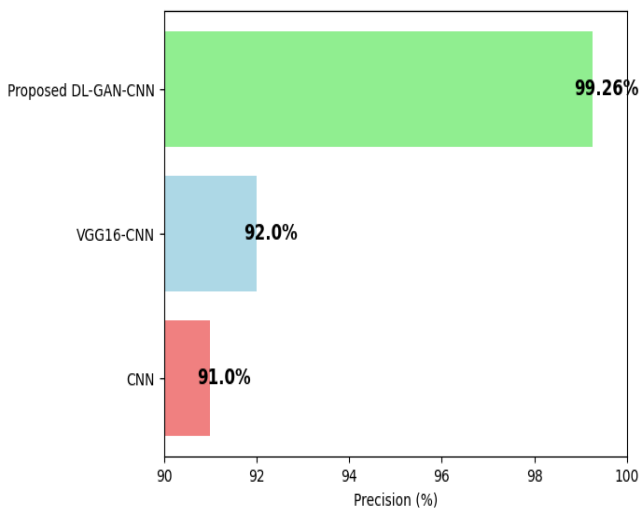| Method	Precision (%)
CNN	91
VGG16-CNN	92
Proposed Method	99.26

Figure 6. Precision comparison of different methods

Figure 6 visually depicts this comparison, clearly showing the superior precision achieved by the proposed algorithm in medical image segmentation tasks.

Table 3 and Figure 7 present the evaluation and performance metrics for recall across various methods. The proposed DL-GAN-CNN approach outperforms the existing methods (CNN procedure and VGG16-CNN) by achieving a notably higher recall rate of approximately 98.42%.

Table 4 and Figure 8 display the evaluation and performance metrics for the F1-Score across various methods. The proposed DL-GAN-CNN method surpasses the existing approaches (CNN and VGG16-CNN) by achieving a notably higher F1-Score of approximately 99.36.

Table 3. Recall comparison of different methods

Method	Recall (%)
CNN	95
VGG16-CNN	92.1
Proposed Method	98.42

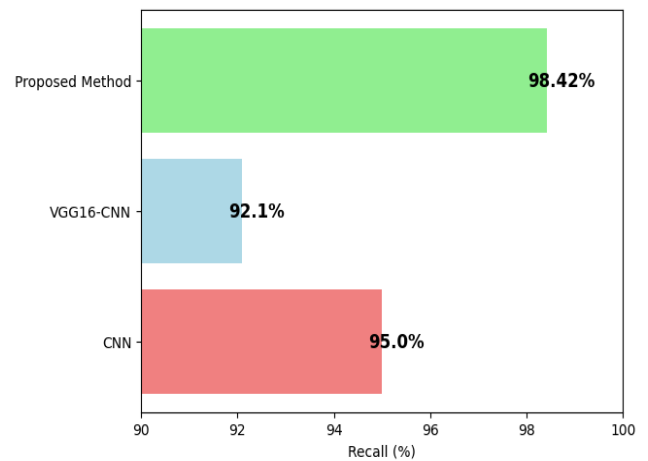**Figure 7.** Recall comparison of different methods

Table 4. F1-score comparison of different methods

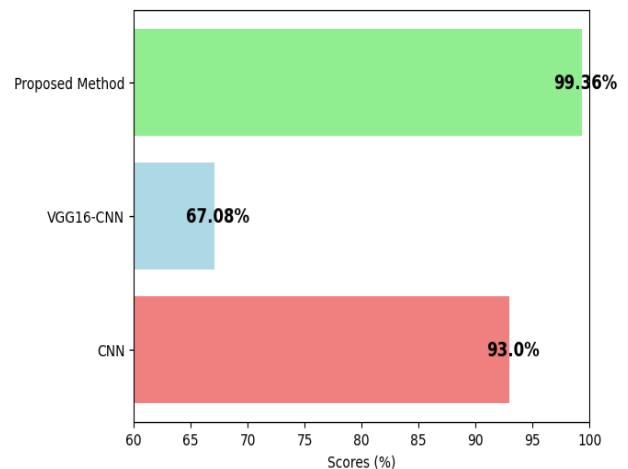| Method	F1-Score (%)
CNN	93
VGG16-CNN	67.08
Proposed Method	99.36

Figure 8. F1-score comparison of different methods

5.2 Discussions

The proposed DL-GAN-CNN method significantly outperforms traditional CNN and VGG16-CNN approaches across all key performance metrics. Below, we analyze the reasons for these performance differences and highlight the specific advantages of the proposed method.

5.2.1 Accuracy

Proposed Method (99.26%) vs. CNN (95.39%) and VGG16-CNN (93.69%):

The higher accuracy of the DL-GAN-CNN framework is because GANs generate high-quality feature representations. Traditional CNNs rely on fixed architectures for feature extraction, whereas the GAN component in the proposed method enhances the feature extraction process by generating realistic and diverse features that capture the variability of medical images with higher resolution. This leads to more accurate segmentation of tumor regions even in noisy or artifact-containing images.

5.2.2 Recall

Proposed Method (98.42%) vs. CNN (95%) and VGG16-CNN (92.1%):

A higher recall for the DL-GAN-CNN framework shows the capability of detecting a larger ratio of true positive instances, for example, tumor regions. In medical imaging, it will be very serious if some tumors or lesions were missed for patient treatment. By the generated features from GAN, the model can identify subtle or irregular tumor boundaries that might be missed by conventional CNNs.

5.2.3 F1 Score

Proposed Method (99.36%) vs. CNN (93%) and VGG16-CNN (67.08%):

This fact indicates the overall strength of the DL-GAN-CNN framework in balancing precision and recall with its F1 Score. This improvement of the F1 score over the state-of-the-art approaches means the proposed method will not only reduce false positives but can also minimize false negatives, hence more reliable for medical image segmentation. The integration of GANs and CNNs allows the model to make a better balance between the two metrics, which is critical for clinical applications.

5.3 Advantages of the proposed method

The key benefits of the proposed DL-GAN-CNN framework compared to traditional CNN and VGG16-CNN are as follows:

(1) Enhanced feature extraction:

This is enabled by the incorporation of GANs, which allows the model to generate high-quality feature representations capturing the complex variability in medical images, thereby leading to more accurate and reliable segmentation results, especially in cases that are difficult for traditional methods.

(2) Robustness to noise and artifacts:

This would help the GAN component to adapt to changes in image quality that could include noise, low contrast, or other artifact forms. In that respect, the approach presented here would be much more robust and reliable for clinical applications where image quality may differ from case to case.

(3) Improved generalization:

The proposed DL-GAN-CNN framework exhibits very strong generalization capability and thus works well for various datasets and different imaging modalities. This is a great advantage compared with traditional methods, which may require extensive retraining or fine-tuning for different datasets.

1) Balanced precision and recall:

The proposed approach achieves better optimization in balancing precision and recall by better reducing false positives and false negatives. In fact, both these kinds of errors are serious during the medical imaging analysis in patient care.

2) Computational efficiency:

Meanwhile, notwithstanding that its architecture is very advanced, the DL-GAN-CNN framework will keep the computational efficiency for real-time applications in clinical use. This has been achieved by carefully designing the GAN and CNN parts and optimizing the training process.

5.4 Implications for clinical applications

Therefore, such a performance of the DL-GAN-CNN framework has essential consequences for its clinical applications. It will make the diagnosis of radiologists more accurate and prompter, providing more accurate and robust segmentation results. This could lead to better patient outcomes, besides alleviating the workload of medical professionals. Moreover, the proposed framework is robust to variations in image quality and has learned generalizable representations across datasets, finding it very suitable for a wide range of medical image analysis tasks.

6. CONCLUSIONS

In this research, we have proposed a novel deep learning approach, namely the DL-GAN-CNN framework, which is especially designed to cater to the complications arising in segmenting medical images with unparalleled accuracy and efficiency. The union of GANs with CNNs has tended to raise these performance metrics-accuracy, precision, recall, and F1-score-so high above the capabilities offered by conventional approaches like standalone CNN and VGG16-CNN-based architectures. Besides proposing an advanced idea of feature extraction, this framework demonstrates incredible adaptability against a wide range of medical image data diversity.

These findings are not to be underestimated. While medical diagnosis leaves little room for error, the DL-GAN-CNN framework has proven to be sound and accurate, promising to change how doctors look at image analysis in the near future. This is further proven by the very high F1-Score, evidencing the strength of the framework in balancing precision and recall to ensure that both false positives and false negatives become minimal. The balance is great in the clinical setting because accuracy of diagnosis impacts directly on the patient outcome.

Moreover, the proposed method's adaptability to various imaging conditions—be it variations in contrast, noise, or other artifacts—demonstrates its potential for widespread application across different medical domains. This adaptability, combined with the high accuracy rates, suggests that DL-GAN-CNN could serve as a valuable asset in automated diagnostic processes, reducing the workload of medical professionals and increasing the consistency and reliability of diagnoses.

7. FUTURE WORK

7.1 Extending the framework to other medical imaging modalities

While the DL-GAN-CNN framework has shown exceptional promise in brain tumor segmentation using MRI scans, its potential extends to a wide range of other medical imaging modalities. Below, we discuss specific modalities and challenges associated with applying the DL-GAN-CNN framework to each and potential solutions:

(1) Computed Tomography (CT) Scans
● **Applications**
CT scans are widely used for diagnosing conditions such as lung cancer, cardiovascular diseases, and abdominal abnormalities.
● **Challenges**
-Noise and Artifacts: CT images often contain noise and artifacts, such as beam hardening and metal artifacts, which can degrade image quality and affect segmentation accuracy.
-3D Data: CT scans are typically 3D volumes, requiring the framework to handle higher computational complexity and memory usage compared to 2D MRI scans.
-Contrast Variations: The contrast between different tissues in CT scans can vary significantly, making it challenging to segment regions of interest accurately.

(2) X-Ray Imaging
● **Applications**
X-rays are commonly used for detecting fractures, infections, and lung conditions such as pneumonia and tuberculosis.
● **Challenges**
-Low Contrast: X-ray images often have low contrast, making it difficult to distinguish between soft tissues and abnormalities.
-Overlapping Structures: In X-rays, overlapping anatomical structures can obscure regions of interest, complicating the segmentation process.
-Limited Depth Information: X-rays provide 2D projections of 3D structures, which can limit the amount of information available for segmentation.

(3) Ultrasound Imaging
● **Applications**
Ultrasound is widely used for imaging soft tissues, such as the liver, kidneys, and fetal development during pregnancy.
● **Challenges**
-Speckle Noise: Ultrasound images are often affected by speckle noise, which can obscure important details and reduce segmentation accuracy.
-Low Resolution: Ultrasound images typically have lower resolution compared to MRI or CT, making it challenging to detect small or subtle abnormalities.
-Real-Time Processing: Ultrasound is often used in real-time applications, requiring the framework to process images quickly without compromising accuracy.

(4) Positron Emission Tomography (PET) Scans
● **Applications**
PET scans are used for detecting cancer, monitoring treatment response, and studying brain function.
● **Challenges**
-Low Spatial Resolution: PET images have lower spatial resolution compared to CT or MRI, making it difficult to precisely localize regions of interest.

-Radiotracer Variability: The quality of PET images depends on the type and distribution of radiotracers, which can vary between patients and affect segmentation accuracy.
-Multimodal Integration: PET scans are often combined with CT or MRI (PET-CT or PET-MRI), requiring the framework to handle multimodal data fusion.

(5) Optical Coherence Tomography (OCT)
● **Applications**
OCT is used for imaging the retina, diagnosing eye diseases, and guiding surgical procedures.
● **Challenges**
-Noise and Artifacts: OCT images are prone to noise and artifacts, such as shadowing and motion artifacts, which can affect segmentation accuracy.
-Thin Structures: The retina contains thin layers that are difficult to segment accurately, especially in the presence of noise or pathology.
-High Resolution: OCT images have very high resolution, requiring the framework to handle large amounts of data efficiently.

7.2 Potential solutions and future directions

To address these challenges, future work could focus on the following directions:

(1) Noise Reduction Techniques
Develop high-quality noise reduction algorithms, such as denoising autoencoders or wavelet transforms, to enhance image quality before segmentation.

(2) Real-Time Processing
Optimize the framework to real-time applications, reduce computational complexity, and explore hardware acceleration such as GPU and TPU.

(3) Transfer Learning
Tap into transfer learning to adapt the framework to new imaging modalities when there is limited availability of annotated data, avoiding extensive retraining.

(4) Clinical Validation
Collaborate with medical institutions for clinical validation of the framework, as that will prove that it can stand the high standards for clinical use.

7.3 Conclusion

In the end, the DL-GAN-CNN framework establishes a new frontier in medical image segmentation, providing a powerful, accurate, and versatile tool that holds great promise to bring about a significant enhancement in diagnostic accuracy and patient outcomes. As we continue to refine and expand this work, the possibilities for its application are endless, promising a future in which AI-driven diagnostics will be a cornerstone of modern medicine.

REFERENCES

[1] Khalighi, S., Reddy, K., Midya, A., Pandav, K.B., Madabhushi, A., Abedalthagafi, M. (2024). Artificial intelligence in neuro-oncology: Advances and challenges in brain tumor diagnosis, prognosis, and precision treatment. NPJ Precision Oncology, 8(1): 80. https://doi.org/10.1038/s41698-024-00575-0

[2] Mahmoud, A., Awad, N.A., Alsubaie, N., Ansarullah, S.I., Alqahtani, M.S., Abbas, M., Usman, M., Soufiene,

B.O., Saber, A. (2023). Advanced deep learning approaches for accurate brain tumor classification in medical imaging. Symmetry, 15(3): 571. https://doi.org/10.3390/sym15030571

[3] Dasanayaka, S., Silva, S., Shantha, V., Meedeniya, D., Ambegoda, T. (2022). Interpretable machine learning for brain tumor analysis using MRI. In 2022 2nd International Conference on Advanced Research in Computing (ICARC), Belihuloya, Sri Lanka, pp. 212-217. https://doi.org/10.1109/ICARC54489.2022.9754131

[4] Mahjoubi, M.A., Hamida, S., Gannour, O.E., Cherradi, B., Abbassi, A.E., Raihani, A. (2023). Improved multiclass brain tumor detection using convolutional neural networks and magnetic resonance imaging. International Journal of Advanced Computer Science and Applications (IJACSA), 14(3): 406-414. http://doi.org/10.14569/IJACSA.2023.0140346.

[5] Gogineni, R., Chaturvedi, A. (2022). Convolutional neural networks for medical image analysis. In Convolutional Neural Networks for Medical Image Processing Applications, pp. 75-90. http://doi.org/10.1201/9781003215141-4

[6] Iqbal, S., Qureshi, A., Li, J., Mahmood, T. (2023). On the analyses of medical images using traditional machine learning techniques and convolutional neural networks. Archives of Computational Methods in Engineering, 30(5): 3173-3233. http://doi.org/10.1007/s11831-023-09899-9

[7] Nazir, S., Kaleem, M. (2023). Federated learning for medical image analysis with deep neural networks. Diagnostics, 13(9): 1532. https://doi.org/10.3390/diagnostics13091532

[8] Mirunalini, P., Desingu, K., Aswatha, S., Deepika, R., Deepika, V., Jaisakthi, S.M. (2024). Conditional adversarial segmentation and deep learning approach for skin lesion sub-typing from dermoscopic images. Neural Computing and Applications, 36(26): 16445-1646. https://doi.org/10.1007/s00521-024-09964-9

[9] Qu, C.X., Zou, Y.X., Dai, Q.Y., Ma, Y.Q., He, J.B., Liu, Q.H., Kuang, W.H., Jia, Z.Y., Chen, T.L., Gong, Q.Y. (2021). Advancing diagnostic performance and clinical applicability of deep learning-driven generative adversarial networks for Alzheimer's disease. Psychoradiology, 1(4): 225-248. https://doi.org/10.1093/psyrad/kkab017

[10] Hu, X., Chung, A.G., Fieguth, P., Khalvati, F., Haider, M.A., Wong, A. (2018). Prostategan: Mitigating data bias via prostate diffusion imaging synthesis with generative adversarial networks. arXiv preprint arXiv:1811.05817. https://doi.org/10.1016/j.asoc.2024.112133

[11] Wang, S.Q., Chen, Z., You, S.R., Wang, B.C., Shen, Y.Y., Lei, B.Y. (2022). Brain stroke lesion segmentation using consistent perception generative adversarial network. Neural Computing and Applications, 34(11): 8657-8669. https://doi.org/10.1007/s00521-021-06816-8

[12] Schellenberg, M., Gröhl, J., Dreher, K.K., Nölke, J.H., Holzwarth, N., Tizabi, M.D., Seitel, A., Maier-Hein, L. (2022). Photoacoustic image synthesis with generative adversarial networks. Photoacoustics, 28: 100402. https://doi.org/10.1016/j.pacs.2022.100402

[13] Lei, Y., Qiu, R.L., Wang, T., Curran Jr, W.J., Liu, T., Yang, X. (2022). Generative adversarial networks for medical image synthesis. Biomedical Image Synthesis and Simulation, 105-128. hhttps://doi.org/10.1016/B978-0-12-824349-7.00014-1

[14] Chen, X.X., Wang, X.M., Zhang, K., Fung, K.M., Thai, T. C., Moore, K., Mannel, R.S., Liu, H., Zheng, B., Qiu, Y.C. (2022). Recent advances and clinical applications of deep learning in medical image analysis. Medical Image Analysis, 79: 102444. https://doi.org/10.1016/j.media.2022.102444

[15] Wang, C.Z., Lv, X., Shao, M.W., Qian, Y.H., Zhang, Y. (2023). A novel fuzzy hierarchical fusion attention convolution neural network for medical image super-resolution reconstruction. Information Sciences, 622: 424-436. https://doi.org/10.1016/j.ins.2022.11.140

[16] Raval, D., Undavia, J.N. (2023). A comprehensive assessment of Convolutional Neural Networks for skin and oral cancer detection using medical images. Healthcare Analytics, 3: 100199. https://doi.org/10.1016/j.health.2023.100199

[17] Bhosale, Y.H., Patnaik, K.S. (2023). Bio-medical imaging (X-ray, CT, ultrasound, ECG), genome sequences applications of deep neural network and machine learning in diagnosis, detection, classification, and segmentation of COVID-19: A Meta-analysis & systematic review. Multimedia Tools and Applications, 82(25): 39157-39210. https://doi.org/10.1007/s11042-023-15029-1

[18] Le-Tien, T., To, T.N., Vo, G. (2022). Graph-based signal processing to convolutional neural networks for medical image segmentation. SEATUC Journal of Science and Engineering, 3(1): 9-15. https://doi.org/10.34436/sjse.3.1_9

[19] Raj, R.J.S., Shobana, S.J., Pustokhina, I.V., Pustokhin, D.A., Gupta, D., Shankar, K.J.I.A. (2020). Optimal feature selection-based medical image classification using deep learning model in internet of medical things. IEEE Access, 8: 58006-58017. https://doi.org/10.1109/ACCESS.2020.2981337

[20] Manoharan, S. (2020). Performance analysis of clustering based image segmentation techniques. Journal of Innovative Image Processing (JIIP), 2(1): 14-24. https://doi.org/10.36548/jiip.2020.1.002

[21] Wu, Y.Y., Li, Y.C., Feng, S.L., Huang, M.X. (2023). Pansharpening using unsupervised generative adversarial networks with recursive mixed-scale feature fusion. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16: 3742-3759. https://doi.org/10.1109/JSTARS.2023.3259014

[22] Huang, Z.H., Li, B.H., Cai, Y., Wang, R., Guo, S.W., Fang, L.M., Chen, J., Wang, L.N. (2023). What can discriminator do? Towards box-free ownership verification of generative adversarial networks. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, pp. 5009-5019. https://doi.org/10.1109/ICCV51070.2023.00462

[23] Krichen, M. (2023). Convolutional neural networks: A survey. Computers, 12(8): 151. https://doi.org/10.3390/computers12080151

[24] Jabber, A.A., Abbas, A.K., Kareem, Z.H., Malik, R.Q., Al-Ghanimi, H., Shadeed, G.A. (2023). Advanced Gender detection using deep learning algorithms through hand X-Ray images. In 2023 16th International Conference on Developments in eSystems Engineering (DeSE), Istanbul, Turkiye, pp. 35-39.

https://doi.org/10.1109/DeSE60595.2023.10469420

[25] Mutha, S.A., Shah, A.M., Ahmed, M.Z. (2021). Maturity detection of tomatoes using deep learning. SN Computer Science, 2: 1-7. https://doi.org/10.1007/s42979-021-00837-9