



Medical image segmentation using deep semantic-based methods: A review of techniques, applications and emerging trends

Imran Qureshi ^{a,b}, Junhua Yan ^{a,*}, Qaisar Abbas ^{c,*}, Kashif Shaheed ^d, Awais Bin Riaz ^e, Abdul Wahid ^e, Muhammad Waseem Jan Khan ^f, Piotr Szczuko ^d

^a Key Laboratory of Space Photoelectric Detection and Perception, Ministry of Industry and Information Technology, College of Astronautic, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 211106, China

^b Department of Computer Software Engineering, Military College of Signals, National University of Sciences and Technology (MCS-NUST), Islamabad 44000, Pakistan

^c College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), P.O. Box 5701, Riyadh 11432, Saudi Arabia

^d Department of Multimedia Systems, Faculty of Electronics, Telecommunication, and Informatics, Gdańsk University of Technology, 80-233 Gdańsk, Poland

^e School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad 44000, Pakistan

^f Department of Management Sciences, University College of Zhob, BUITEMS, Balochistan 85200, Pakistan

ARTICLE INFO

Keywords:

Deep learning
Medical imaging
Optimization techniques
Transfer learning
Semantic segmentation

ABSTRACT

Semantic-based segmentation (Semseg) methods play an essential part in medical imaging analysis to improve the diagnostic process. In Semseg technique, every pixel of an image is classified into an instance, where each class is corresponded by an instance. In particular, the semantic segmentation can be used by many medical experts in the domain of radiology, ophthalmologists, dermatologist, and image-guided radiotherapy. The authors present perspectives on the development of an architectural, and operational mechanism of each machine learning-based semantic segmentation approach with merits and demerits. In this regard, researchers have proposed different Semseg methods and examined their performance in a variety of applications such as medical image analysis (e.g., medical image classification and segmentation). A review of recent advances in Semseg techniques are presented in this paper by applying computational image processing and machine learning methods. This article is further presented a comprehensive investigation on how different architectures are helpful for medical image segmentation. Finally, advantages, open challenges, and possible future directions are elaborated in the discussion part, beneficial to the research community to understand the significance of the available medical imaging segmentation technology based on Semseg and thus deliver robust segmentation solutions.

1. Introduction

Image segmentation is useful for medical imaging as it is crucial for analyzing and diagnosing a particular disease. The precise identification of region-of-interest (ROI) within the sample is a critical step to perform any feature segmentation. Semantic-based Image segmentation is a classification task at each pixel level as it allows us to detect biological structures and quantify the morphology. Further, it is useful because it will enable us to capture objects' shapes quantitatively and provide the means of doing high-resolution spatial statistics [1–3]. Most authors divide deep learning-based semantic segmentation into three different important categories: fully convolutional network (FCN), region-based, and weakly supervised segregation methods [5], as depicted in Fig. 1.

Region-based methods require domain experts to identify the region of interest (ROI) and precisely draw the boundaries for image segmentation tasks. The next step in these methods is correctly annotating every pixel according to the ground truth classes. So, these methods are time-consuming and tedious as they need domain experts and resources to select the region of interest (ROI). But one thing that shows the importance of these methods is that fully convolutional networks (FCN) based methods and weekly supervised methods use the knowledge generated by region-based methods. As region-based approaches involve manual processes, these methods are feasible for small datasets. Secondly, if the images are very high resolution, it is difficult to correctly identify the boundary values through region-based approaches [6]. These methods are helpful for medical imaging because, for clinical

* Corresponding author.

E-mail addresses: yjh9758@126.com (J. Yan), qaabbas@imamu.edu.sa (Q. Abbas).

image analysis, we need the help of domain experts like physicians to segment specific images for the prediction of some diseases.

Weakly supervised segmentation, also known as semi-supervised segmentation methods, used automated algorithms for segmentation tasks with little interaction of the domain experts with the systems to accurately identify the results produced by these methods [7]. Domain experts may be required when selecting ROI, which will be generalized to the whole image using weakly supervised algorithms. Some examples of these methods are seeded region growing (SRG) algorithms that use the knowledge provided at the initial seed point and generalize it to a neighboring pixel to merge them and segment them accordingly [8]. Level set-based active contour models iteratively change the shape of the boundary by expanding or shrinking using automated algorithms. The advantage of using these methods is that they do not need prior knowledge of ROI [9]. Another method is the localized region-based active contour technique in which information about foreground and background is collected using region parameters. The benefit of using a small, localized region is that it handles the heterogeneous texture accurately [10].

Fully convolutional networks (FCN) are fully automatic segmentation techniques because domain experts do not need to select ROI. But most of these methods involve supervised learning and require training data, e.g., deep neural networks, shape models, and random forests. However, if an unsupervised learning approach is used, we need some labeled images as validation and training data. Some challenges with FCN-based methods are the variation in image size, texture, shape, and ROI selection in low-resolution images [11]. Automatic segmentation methods usually contain fully convolutional models; since these are fully convolutional, we can change the input shape and the number of classes, which makes it efficient for semantic segmentation tasks [14, 15]. But sometimes, if we have noisy data for image analysis, image acquisition variation in different features [57] can affect the output of automatic segmentation models. Methods like watershed algorithms, clustering techniques, and some Machine learning (ML) based models face global applicability challenges. More methods involving human features engineering with machine learning-based support vector machine (SVM) and some other neural network (NN) is time-consuming, fails to process raw features, and cannot adapt to new features. Deep Learning (DL) based methods solve all these problems of other ML-based arts. These methods can easily process raw data, generalize it on unseen datasets, and do not need domain experts for handcrafted features [12]. These approaches have shown effective results for semantic segmentation of natural scene images and biomedical imaging [13]. With upgraded CPUs

and GPUs, training and test time are reduced; that is why deep learning methods are widely used in image segmentation, image classification, and other image analysis tasks. Multiclass segmentation using DL has many attractive qualities; first, it does not influence human factors and reduces segmentation time. In addition, it provides an opinion for spatial annotations to eliminate the manual efforts of experts for labeling samples and to perform semantic segmentation for various medical imaging modalities [63–68].

In this review article, we have reviewed available data sources and applications of semantic-based segmentation (Semseg) methods in the domain of medical imaging modalities. Fig. 2 shows the organization of imaging modalities covered in this review paper about semantic-based segmentation. In the Semseg field, all papers are classified into subsections according to the applications concerning deep-learning (DL) algorithms and imaging modalities. In this article, we have explained the importance of image modalities in diagnosing human diseases using DL algorithms. To further highlight the research gap in this field, the new DL algorithms are also described that can be used in the future to diagnose image modalities.

1.1. Paper organization

Section 1 introduces the problem and describes various segmentation methods used for different kinds of medical image datasets. These semantic segmentation methods for medical imaging can be categorized into three different classes, namely: (1) Region-Based Segmentation, (2) Fully Convolutional Neural Network (FCN), and (3) Weakly supervised segmentation. **Section 2** describes the research protocol of our review paper. It also covers the research questions with the motivation to address them. A comparison of our paper to the existing published paper has been presented in this **section 2**. It also includes the research paper quality evaluation to define the protocols to select the best paper for our review. Finally, the detail about the database used to evaluate the performance of existing segmentation methods was demonstrated. **Section 3** illustrates the literature on developed segmentation methods. These include deep learning-based and multimodal-based segmentation schemes for biomedical image analysis problems. **Section 4** presents a study about a deep learning model used for medical image segmentation, such as Model Compression Based, Attention Based, Encoder-Decoder Based, and Sequential Model-based Segmentation methods.

Section 5 demonstrates the optimization functions for image segmentation which include: (1) Cross-Entropy, (2) Focal loss, (3) Dice loss, (4) Tversky loss, and (5) Exponential logarithmic loss. **Section 6**

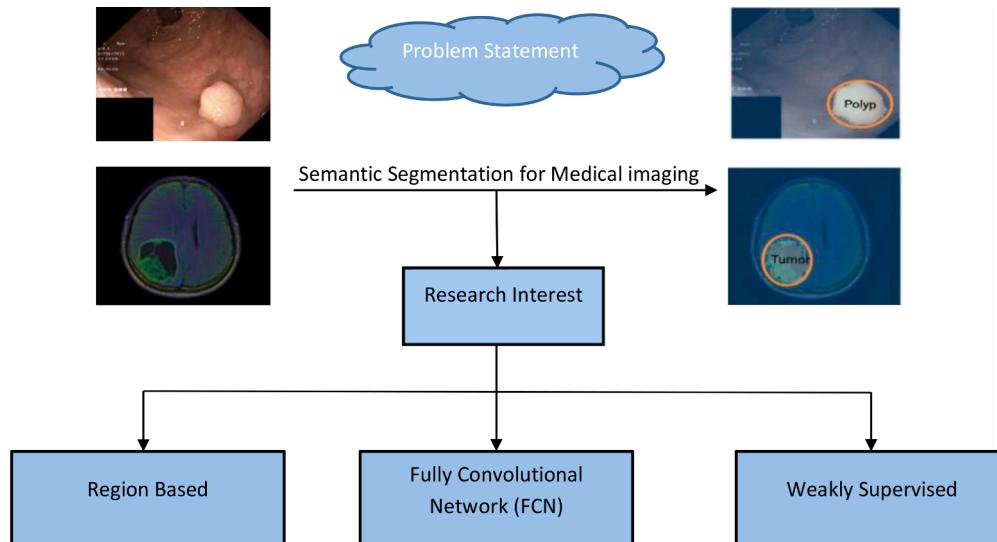


Fig 1. Classification of major semantic segmentation methods available for different kinds of medical image datasets.

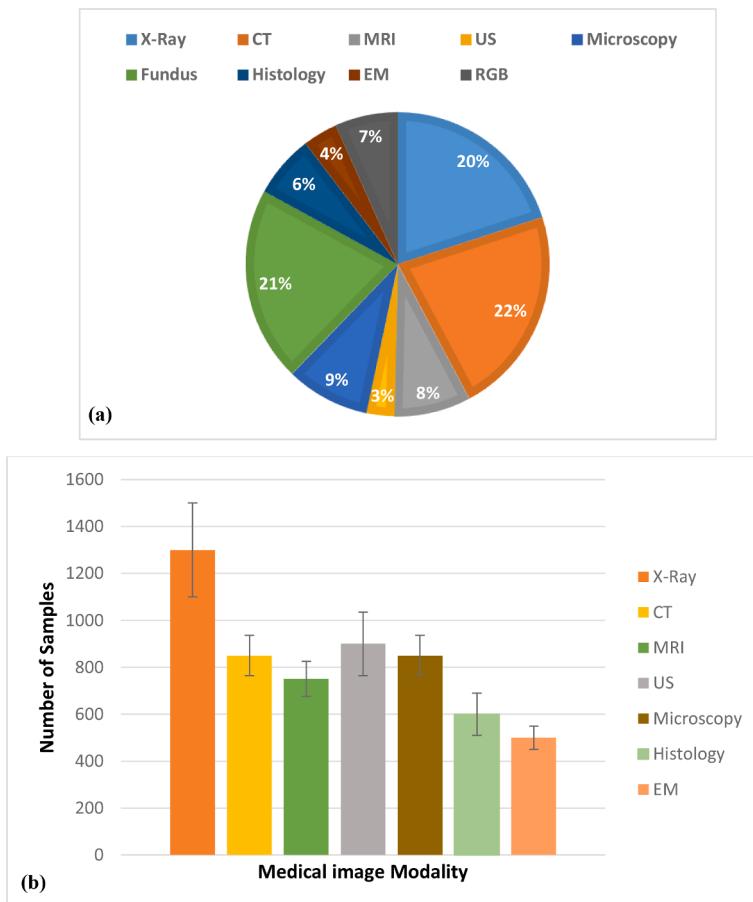


Fig 2. Organization of imaging modalities covered in this review paper about semantic-based segmentation of medical images, where figure (a) shows the selected research papers published for each modality, and figure (b) shows the number of samples used in each modality.

discusses thorough literature on the semantic segmentation approaches used to diagnose various potential diseases. The current deep learning architecture, optimization function, and performance are discussed in section 7. This section also debated possible future research directions and presented future research prospects in medical image segmentation. Finally, our paper is concluded in section 8.

1.2. Research highlights

The main contributions of this review article are as follows:

- Semantic-based segmentation (Semseg) methods are presented in deep learning architecture.
- A comparative study is presented for reviewed paper based on performance.
- Existing challenges and problems in DL-based semantic segmentation approaches are discussed.
- The challenges of Semseg applications are mentioned for other researchers.
- Prospect for future work in this area for regular medical image segmentation.

2. Research protocol

The research protocol described the complete layout of this review article, defining the methods used in this study. It also shows the state-of-the-art survey of recently conducted reviews about semantic segmentation methods for medical images. It will be helpful for beginners to learn about recent advancements in the semantic segregation of

biomedical imaging. Accordingly, the evolution of semantic segmentation arts started from 2011 to 2022 for the medical imaging analysis is presented in Fig. 3. This review describes s performance comparisons of different deep learning-based techniques and methods available for semantic segmentation tasks and explains the optimization functions used for semantic image segmentation to improve the performance measures. This review will also highlight the limitations of currently available methods with experiments and results comparisons.

2.1. Research questions

The research questions are the fundamental step of any study, either a literature survey or research project. This is the starting point of any kind of scientific research. Following are some important research questions highlighted in Table 1 that will be addressed in this review paper.

2.3. Analysis of articles

The current trend of semantic-based segmentation (Semseg) methods for the diagnosis of different human diseases are described visually by Fig.3. This information is collected from different platforms such as PubMed and top journals from the years 2011 to 2022. In the year 2022, it is noticed that the latest trend is to use deep-learning architectures in multimodal features are utilized by many researchers. To limit the scope of the review article, we have reviewed only Semseg methods used in the medical domain.

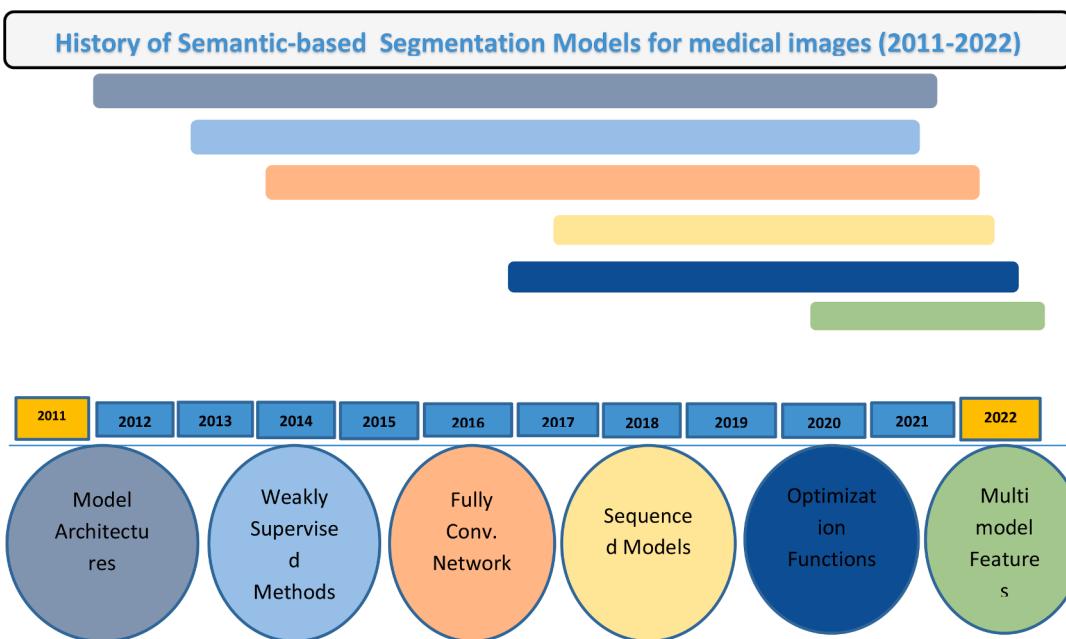


Fig. 3. An overview of the latest trend to show the methodological evolution in the development of semantic-based segmentation models for medical images

Table 1
Summary of Research Questions and Motivations Involves in this Study.

S. No.	Research Questions	Motivations
RQ1	What are the currently available semantic segmentation methods for biomedical images?	To know more about recent studies on semantic segmentation methods for biomedical images.
RQ2	What are the Deep Learning based improvements applied to medical image segmentation?	To know about a recent famous research study about medical image segmentation.
RQ3	What optimization functions used for semantic image segmentation to improve the performance measure?	To know about semantic image segmentation approaches which used optimization functions to improve performance.
RQ4	How are Encoder-Decoder Decoder-based methods helpful for medical image segmentation?	To know the significant factor of Encoder-Decoder-based methods for the image.
RQ5	What statistical measures can be used for performance evaluation of available methods?	To understand the statistical measure used for performance evaluation.
RQ6	Why are weakly supervised or unsupervised methods frequently used in medical image segmentation compared to supervised or fully automatic semantic segmentation methods?	To know about the methods of machine learning for image segregation pertaining to medical.
RQ7	What are the current limitations and future guidance for semantic segmentation of medical images?	To study and know about the limitations and future directions for semantic segmentation.
RQ8	Which optimization technique performs well in segmentation tasks?	To know and evaluate the performance of various optimization techniques.
RQ9	Would the transfer learning technique improve the performance of segmentation tasks?	To analyze the performance of several pre-trained models.
RQ10	What kinds of datasets are used for medical image segmentation?	To study and investigate various image benchmarks suitable for various medical image segmentation tasks.

2.2. Comparison with existing survey articles

We studied the literature of recently published review articles and avant-garde methods for semantic segregation of images pertaining to medical. We have shown an extensive comparison of our review article

with some recently published review articles in well-reputed journals. **Table 2** shows the difference between our review article and other available surveys and review papers.

2.4. Papers quality evaluation

This review study focuses solely on the previous research categories listed below: the paper written in English, which focuses on the medical image segmentation task. That work focuses on semantic segregation based on deep learning for medical image analysis. However, instance-based segmentation was disregarded from the exploration. The keywords "Semantic Segmentation and Deep Learning-based Semantic Segmentation" were searched for related papers. The four top digital databases were chosen to explore for a target research paper: (1) IEEE Xplore, (2) Science Direct, (3) ACM, and (4) Springer, as shown in **Table 3**. We also include others as a new column in the table below, which includes paper from other digital databases such as SPIE, MDPI, etc. These resources provide many research publications on semantic segmentation in medical image analysis. The research under consideration was published from 2015 through 2021.

Three filtering iterations and screening methods were used to choose and search for a relevant research paper from these databases. The duplicate research paper was eliminated during the initial screening procedure. The second screening phase identifies and eliminates all irrelevant manuscripts by examining the title and abstract. All research papers that passed the second screening procedure were subjected to a full-text review in the third screening step. All iteration phases were subjected to the same eligibility criterion. The last group of research articles covers deep learning-based semantic segmentation and semantic segmentation. To consider the newly published work, the inquiry was undertaken in September 2021. **Table 3** shows the total number of papers found in each database using the mentioned keywords for medical image segmentation. After applying the filtering and screening method, a total of 167 papers were selected for this study.

2.5. Segmentation datasets for performance evaluation

Collecting adequate data into the data set for any deep learning model segmentation is critical. The quality of the segmentation algorithm is determined by the high-quality image data provided by the

Table 2

Summary of existing survey papers on medical image segregation based on deep learning methods.

S.No.	Related Review Articles	Reviewed up to	Research Questions									
			RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	RQ7	RQ8	RQ9	RQ10
1	İşın, Ali et al. [5]	2016	✓				✓					
2	Anwar, Syed Muhammad, et al. [34]	2018		✓			✓		✓			
3	Tongxue Zhou et al. [69]	2019		✓	✓	✓	✓			✓	✓	
4	Taghanaki, Saeid Asgari, et al. [2]	2020	✓	✓	✓	✓	✓		✓	✓		✓
5	Tajbakhsh, Nima, et al. [70]	2020	✓	✓			✓	✓	✓		✓	
6	Wang, Risheng et.al. [156]	2022			✓			✓	✓	✓	✓	
7	Malhotra, Priyanka et.al. [157]	2022	✓	✓			✓		✓			✓
8	Our Review	2021,2022	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 3

Number of Studies for Initial Search for Each Keyword for All Databases.

Keyword Search	Springer	IEEE Xplore	Science Wiley	Science Direct	Conferences	Others
Semantic Segmentation	20	15	5	24	34	11
Semantic Segmentation+ Deep learning	28	18	10	77	30	18

experts and the associated label-standardized data set, which allows for fair comparison amongst systems. In this subsection, we have listed some of the datasets used in the recent existing literature. Table 4 demonstrates different datasets used for medical image segmentation. It shows the name of the dataset, modalities, total number of samples, and available URL of these datasets.

3. Overview of semantic segmentation (Semseg) methods

Semantic image segregation, categorized at the pixel level, is the process of grouping components of the image together that correspond to the object of the same class. Image segregation, like prediction at the

pixel level, which categorizes every pixel. Convolutional neural networks have strong feature extraction capabilities; human image feature extraction or unnecessary image preprocessing are not required. As a result, CNN has recently been employed in medical image segregation. Therefore, in this section, we focus on those methods which involve semantic segmentation for medical image segregation based on deep learning.

3.1. Deep learning-based segmentation

The Convolutional Neural Networks (CNNs) is a primarily used algorithm in medical images and computer vision tasks. For instance, some studies shows the significance of deep learning CNN models such as Covid-19 detection [153], Fetal ultrasound standard plane recognition [154], Biometric recognition system [158], pattern recognition system [159,160], brain function detection, and cardiovascular disease screening [152]. For example, A comprehensive review study [153] was proposed that focuses on the scope and contribution of AI to combat Covid-19. Pu et al. [154] developed an automatic fetal ultrasound standard plane recognition model based on CNN and recurrent neural network (RNN) in the industrial internet of things (IIOT) environment. The proposed model learn spatial and temporal feature of the ultrasound video stream. The CNN component uses every video frame to identify the four potential fetal standard planes and other critical anatomical components of the fetus. To precisely localize and monitor fetal organs across frames, the RNN component collects the temporal information spanning adjacent frames. The proposed model is more accurate than the competing baselines. Chen et al. [152] proposed a configured DL framework, which comprises a library of DL model components and a library of MIA task elements to precisely describe different possible MIA activities for many diseases. The framework allows creation of a customized DL model by defining MIA tasks and setting DL model components in response to a specific MIA need for various diseases.

Additionally, the structure of the DL model for each configuration can be further modified to enhance performance by comparison with the current machine learning (ML)/DL models. The key benefit of this approach is its ability to create flexible, individualized DL models and to continuously tweak the model structure to enhance model performance. But the generally Convolutional Neural Networks (CNNs) approaches apply to 2D images instead of 3D images. In contrast, more medical data is available in 3D images. In [17], the authors proposed a fully Convolutional Neural Networks (CNNs) approach for 3D image segmentation based on volumetric. The author introduced the Dice overlap coefficient between the background and predicted segmentation as a novel objective function to optimize training data. Various acquisition procedures and various types of equipment are used to collect the

Table 4
Demonstrate the dataset used in Medical Image Segmentation.

Dataset Name	Modalities	Total Number of Samples	Available URL
MSD	MRI, CT	2633	http://medicaldecathlon.com/
BRATS	MRI	Not Reported	https://www.med.upenn.edu/sbia/brats2018/data.html
DDSM	Mammography	5000	http://www.eng.usf.edu/cvpr/Mammography/Database.html
ISLES	MRI	166	http://www.isles-challenge.org/
LiTS	CT	200	https://competitions.codalab.org/competitions/17094
PROMISE12	MRI	Not Reported	https://promise12.grand-challenge.org/
LIDC-IDRI	CT	1018	https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI
OASIS	MRI, PET	Not Reported	https://www.oasis-brains.org/
DRIVE	FUNDUSCOPY	40	https://drive.grand-challenge.org/
STARE	FUNDUSCOPY	400	http://homes.esat.kuleuven.be/~mblaschk/projects/retina/
CHASEDB1	FUNDUSCOPY	84	https://blogs.kingston.ac.uk/retinal/chasedb1/
MIAS	X-Ray	322	https://www.repository.cam.ac.uk/handle/1810/250394?show=full
SCD	MRI	45	http://www.cardiacatlas.org/studies/
SK110	MRI	Not Reported	http://www.ski10.org/
HVSMDR2018	CMR	Not Reported	http://segchd.csail.mit.edu/

medical dataset from different hospitals. The proposed method is evaluated in terms of two parameters such as the Hausdorff and Dice coefficient. The authors concluded that an improvement in experimental evaluation shows that the proposed method shows excellent performance on challenging datasets whereas utilized a bit more time than the previous method [17].

The author in [16] highlights the central problem of segmentation in breast ultrasound images to diagnose breast cancer. This study overcame the problem of manual segmentation in breast ultrasound images by using the automated CNN classifier to segment the tissues separately. The convolutional neural networks (CNNs) classify breast ultrasound images into four functional tissues: mass, skin, fibro glandular, and fatty tissue. The proposed method is evaluated based on four quantities performance parameters: recall, Precision, $F1_{measure}$, and accuracy. The author concluded that the value of all performance parameters reached over 80%, which shows that the proposed method could accurately segment the ultrasound images into functional tissues. The value of another evaluation metric, the Jaccard similarity index (JSI), also increases from 74.54% to 85.1% compared to existing studies. The author concluded that the proposed method could improve the other medical images and breast cancer clinical diagnosis. Similarly, some other authors have also proposed breast cancer segmentation methods using different image modalities, including Mammography, MRI, US, CT, and Optical imaging [80,81,82,83,84].

In [18], the author says that the real reason for benign situations of some skin cancers is the late diagnosis of cancer. The early detection of skin cancer, especially melanoma, is essential to prevent benign situations. It's a challenging process to detect the early skin cancer stages, even by specialists. The author addressed this problem by optimizing the Convolutional neural network (CNN) with a whale optimization algorithm. The whale optimization algorithm inspires the process of bubble-net hunting in humpback whales to trap the prey. For performance analysis, the proposed optimized Convolutional neural network (CNN) method was compared with existing ten methods, including Inception-v3, semi-supervised method, ResNet, Spot-mole tool, Ordinary CNN, AlexNet, and VGG-16, in terms of five performance parameters such as sensitivity specificity, NPV, PPV, and accuracy. The author concluded that the proposed method shows virtuous performance for diagnosing skin cancer [18]. Similarly, other authors have also proposed skin cancer segmentation methods using different image modalities, including Surface Microscopy, US, MRI, and Confocal Microscopy [93, 94,95,96,97].

The author in [161] proposed a new deep medical image segmentation framework named dual swin transformer U-Net (DS-TransUNet), to integrate the hierarchical swin transformer into both the encoder and the decoder of the conventional U-shaped architecture. Their model takes advantage of the self-attention computation in the swin transformer and the specially developed dual-scale encoding, which can accurately simulate the non-local relationships and multiscale contexts for improving the semantic segmentation quality of different medical images. They also proposed a well-designed transformer interactive fusion (TIF) module to combine multiscale data using the self-attention method successfully. To better investigate the long-range contextual information during the up-sampling process, we also incorporate the swin transformer block into the decoder. Extensive tests on four typical medical image segmentation tasks show that DS-TransUNet works well, and their method performs noticeably better than state-of-the-art techniques. Recently, in another work, two deep learning networks, USegTransformer-P and USegTransformer-S were proposed for medical image segmentation tasks. The presented models combine transformer-based and convolution-based encoders [166] to segment medical images with high precision while taking advantage of local and global features. Both proposed models show more promising results than the previous state-of-the-art models in a variety of segmentation tasks, including segmentation of brain tumors, lung nodules, skin lesions, and nuclei.

In a previous study [19], the author compares deep learning network capabilities with other classifiers such as SVM to evaluate computational efficiency and performance. The author considers two networks named the network in network (NIN) and LeNet to compare the performance and computational efficiency in the detection and classification problems. The experiments performed on the art images, burn wounds images, and facial databases show differences between deeply learned network architectures in terms of computational efficiency and performance. By comparing both deep learned networks, the author concluded that the architectures of NIN and LeNet are more efficient for low complex classification than more complex classification. The more input size may not assure better results. However, this problem can be limited by adding more layers to a network.

The semantic segmentation of the medical images in the context of 3D multimodal is a challenging task compared to 2D single modal. In this study, the author focuses on the problem that segmentation of 3D multimodal medical images with semi-supervised learning is such a difficult task. To address this problem, the author proposed an exciting novel approach named Generative Adversarial Networks (GANs) to differentiate in segmentation networks that the output image either belongs to labeled or unlabeled images. The method supports unlabeled data to categorize between generated fake and true patches appropriately. The MRBrain-2013 and Iseg-2017 datasets of Brain MRI images are considered to evaluate the proposed method for the segmentation problem. The proposed segmentation method is compared with the state-of-the-art segmentation networks trained on fully supervised datasets. The results show a significant improvement in the proposed method's DCS and ASD parameters. It is concluded that the proposed method can reduce the burden of acquiring annotated medical images.

In a previous study [21], the author proposed a method to overcome the accuracy problems in several segmentation results. The accuracy issues occur due to dependency on reliable appearance models and a better initial guess. An approach proposed based on distance regularized level set named fully automated non-rigid segmentation. The Deep Belief Networks (DBNs) were initially utilized for training the datasets based on the results extracted from the structure inference. The major problem of result accuracy was stunned by DBNs that formed reliable primary guess and appearance models. To check the accuracy of the proposed innovative method, the experiments are performed on the demanding MICCAI-2009 left ventricle (LV) segmentation datasets. The results conclude that the proposed method is modest compared to the existing methods in terms of accuracy results on the MR images in left ventricle segmentation. The proposed method might encompass the 3D geometric model for the segmentation process.

Diagnosing chest X-ray (CXR) digital images is a demanding problem due to the inconsistency shape of the lungs of individuals and the existence of robust edges on the clavicle and rib. The automated segmentation of lung approaches is applied to diagnose the chest X-ray (CXR) digital images. The author proposed an innovative hybrid method for lung segmentation to conquer the deficiency of the existing lung segmentation method. The new method is based on the fusion of distance regularized level set and deep structured inference. Deep learning produces intense training with minor annotated datasets, down-segmentation, and structured inference. The advantages of the level set approach are the appearance prior, optimization techniques, and use of shape. The author evaluated the proposed method using the Japanese Society of Radiological Technology (JSRT) dataset, which is widely available. In this paper, the author concluded that the new lung segmentation method produces more accurate results than others based on the initial guess in DSC and ACD. The average accuracy of lung segmentation results significantly increases from 94.8% to 98.5%. The quantitative results show that the initial guess used in the method proposed compared to other methods gives the best results. [22]. Similarly, other authors have also proposed segmentation methods to examine the chest and lungs using different image modalities, including X-ray, CT, and MRI [76,77,78,79]. Fig. 4 helps the readers to swiftly learn the ratio

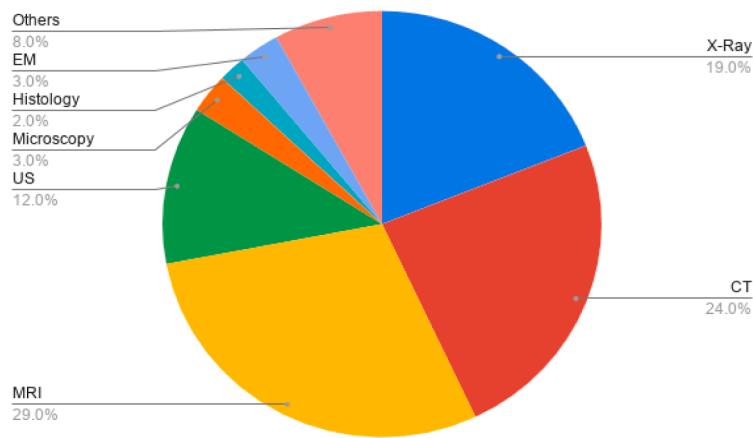


Fig 4. Distribution of Different image modalities used in cited papers.

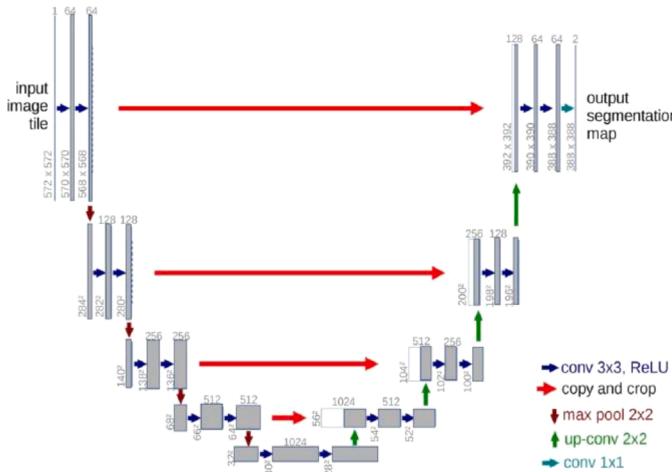


Fig 5. U-Net architecture proposed by Ronneberger et al. 2015. [45]

of a diverse range of medical imaging modalities for segmentation.

In the research article [23], the author proposed a new method based on the combination of the level set and deep learning designed for the automated segmentation of the heart's left ventricle from cardiac cine magnetic resonance (MR) data. Combining deep learning and level set aims to defeat the segmentation problems as interesting objects visually appear in huge shapes and variations in appearance. The level set method uses a small, annotated training dataset, but there are still limitations of large shapes and variations in appearance. Deep learning arts can overcome the drawbacks of an interesting visual object's oversized shape and appearance variations. The deep learning method must utilize small, annotated training to report the variations and regularization to create an excellent overview. So, combining both methods gives precise segmentation results with small, annotated training datasets. The proposed method is tested on the challenge segmentation database as MICCAI 2009 left ventricle, which holds 15 sequences for testing, 15 for training, and 15 for validation. The experiment determined more accurate results for fully automated segmentation comparatively semi-automated problems.

Similarly, authors in [27] suggested a lung segregation method using deep learning U-Net architecture. Their approach eliminated unnecessary information from lung CT images and obtained excellent results. Segmenting lesions or organs from medical images has long been a problematic global problem because of the diversity of parameters such as size, shape, location, and severity. To address this issue, Liu et al. [151] propose a novel region-to-boundary deep learning model. To begin, they employ a U-shaped network with two branches behind the

final layer, one of which generates the target probability map and the other the corresponding signed distance map. Second, using the signed distance map and the acquired multi-scale characteristics, they concentrate on the boundary of the target lesions or organs to be segmented. To obtain the results, they combine the region and border features. The study demonstrates that the suggested model outperforms the comparable approaches for most evaluation criteria, particularly boundary tracking.

Few studies have been found to segment the bacteria from microscopic images for anthrax disease detection. For instance, the study in [149] proposed a deep learning framework based on two networks, UNet and UNet++ to automatically segment and detect the anthracis bacteria in microscopic images. This method offers the same level of accuracy as the human diagnostic specialist, and in some cases, it outperforms. According to research, *B. anthracis* bacteria segmentation of microscopic pictures taken under various circumstances can be automated using these deep architectures, especially UNet++. This work achieved an accuracy of 97% on both patch and whole raw image datasets. However, in UNet architecture, skip connection plays a vital role in these segmentation tasks. Therefore, Hoorali et al. [150] proposed the improved version of UNet architecture for *Bacillus anthracis* bacteria and immune cell segmentation and detection using microscopic images. The proposed approach, known as IRUNet, integrates multi-scale characteristics and benefits from inception and residual blocks in skip connections to extract excellent segmentation features. This proposed study outperformed state-of-the-art techniques and produced better segmentation results, with a precision of 92.8%, recall of

Table 5

Cited papers using deep learning techniques for biomedical image segmentation.

Cited Paper	Modality	Method	Description	Performance metrics & Results
Badea et al. [19]	Medical images	CNN (LeNet and NiN)	CNN-based models LeNet and Network in Network (NiN) are used to classify and detect skin, burn, light burn, and severe burn medical images.	LeNet Accuracy 75.91% and 50.01% NiN Accuracy 55.7%
Mondal et al. [20]	Brain MRI dataset	GAN	Using generative adversarial learning, the authors used a few shot medical image segregation in 3D Multi-model. They used Brain MRI datasets for the segmentation task.	DICE Score 0.75
Milletari et al. [17]	Prostate MRI images	V-Net	Proposed volumetric CNN for the segmentation of MRI prostate volumes by optimizing dice overlap coefficient between predicted segmentation and ground truth annotations.	DICE Score 82.39
Xu, Yuan, et al. [16]	3D US	CNN	Proposed segmentation of breast Ultrasound (US) images into four tissues which include skin, mass, fibro epithelial tissues, and fatty tissues, by using CNN, which indicates a centered pixel in an image block	Accuracy, Precision, recall, and F1 score all metrics are above 80%
Zhang N, et al. [18]	Dermquest and DermIS databases	Whale optimization algorithm	For the early detection of skin cancer, the authors used a whale optimization algorithm to optimize the CNN. They used half value precision function for validation of optimized skin cancer system	Specificity, Sensitivity, Accuracy, NPV, PPV All metrics are between 65% - 90%
Q Abbas et al. [24]	DIARETDB1, FAZ, MESSIDOR, Prv-DR	Multilayer deep learning neural network (DLNN)	An automatic recognition system for five severity levels of diabetic retinopathy is developed through deep visual features. Using a semi-supervised multilayer deep learning algorithm.	AUC 0.924 Sensitivity 92.18
Kemnitz, Jana, et al. [25]	Thigh MRI Images	U-Net	The authors applied a fully automated approach to segment muscles and adipose tissues in cross-sectional area compared with manual segmentation to understand musculoskeletal diseases.	Dice Similarity (0.96 ± 0.01)
Orlando, Nathan, et al. [26]	3D TRUS prostate images	U-Net	They proposed a supervised deep learning method for prostate segmentation in 3D TRUS images from different facilities. This method is helpful for needle-based prostate cancer procedures.	Absolute DSC, Recall, Precision, VPD, MSD, and HD Above 90%
Ngo et al. [23]	MRI	RBM	A segmentation system was proposed to get borders of the left ventricle (LV). All slices of end-diastole (ED) and end-systole (ES) are used in the segmentation process. The user manually selected the ED and ES volumes.	DICE Score: 90%
Skourt et al. [27]	Lung CT images	U-Net	They suggested a lung segregation method using deep learning U-Net architecture. To erase unnecessary information from lung CT images, their approach achieves excellent results.	DICE coefficient index: 0.9502

93%, and dice score of 92.9%. Many authors applied semantic segmentation methods to brain images using different modalities, including MRI, CT, and PET [71–75]. Table 5 summarizes some state-of-the-art methods using deep learning techniques for biomedical image segmentation. Similarly, Table 6 shows the list of some state-of-the-art papers using segmentation methods for Organs and substructure segmentations with different image modalities.

4. Deep learning-based architectures

The authors have covered deep-learning architectures (DLAs) in this part, focusing on current machine learning methods. Several research studies indicate that DLAs-based algorithms achieved high accuracy in detecting retinal features to determine appropriateness for other authors. We discussed the principles, structures, and strategies typically used to detect lesions using visual cues in this part. DLAs are relatively new ways of performing imaging analysis in real-time situations. These DLA approaches were used primarily on images for pattern recognition and feature learning. DLAs have been employed in cutting-edge systems for the segmentation of lesions in a variety of different ways. A comparison of DLA systems may be found in Table 5, which explains each and compares them. To assist potential readers, the authors have split DLAs into three main categories to perform medical image segmentation, i.e., Model compression, Attention, Encoder-Decoder-based, and sequenced based on their goal.

4.1. Model compression-based image segmentation

For the segmentation of larger medical images like high-definition resolution and volumetric, 2 Dimensional images, e.g., CT, MRI, and ultrasound (US) images, several methods have been proposed to be compressed. The author in [28] applied a neural network-based search

method for improved organ segmentation on MRI, CT, and US images by using reduced U-Net architecture. Authors in [29] redesigned U-Net architecture to make it memory efficient for 3D imaging segmentation by applying group normalization and leaky ReLU activation function. Similarly, authors in [30] proposed dilated CNN with fewer parameters than previous methods. Authors in [31,32] focused on the weight's quantization of deep neural networks to compress segmentation models.

4.2. Attention-based image segmentation

Attention-based models were proposed for the segmentation of biomedical images in past years. Authors in [33] presented an optimized attention-based model for the segmentation of prostate from MRI pictures with a better score in accuracy as set parallel to some baseline's methods, e.g., FCN [35] and V-Net [17]. Similarly, authors in [36] designed an attention-based method that is multi-level to segregate abdominal organs from MRI pictures. A dilated convolution base block for 3D medical image segmentation is used to preserve more dilated attention [37]. Similarly, some other papers have also used attention-based architecture for semantic segmentation of biomedical images [38–40].

4.3. Encoder-Decoder-based image segmentation

In these methods, images are normalized using CNN-based pre-processing arts before going through image segmentation tasks. Encoder-decoder architectures in medical image segmentation have shown better MRI, electron microscopy segregation, CT, liver, and prostate segregation scores from MRI scans and CT, respectively [41]. Authors in [42] designed a convolution dilated block to preserve the contextual information for the segmentation of medical images. Authors in [43] have proposed a method to compare a tumor image with a

Table 6

List of some state-of-the-art papers using segmentation methods for organs and substructure segmentations with different image modalities.

Organ	Imaging Modalities	Papers using Segmentation Methods
Brain	MRI, CT, PET	de Brebisson et al. [71], Chen, Hao, et al. [72], Roy, Abhijit Guha, et al. [73], Wang, Li, et al. [74], Devuonoor, Sindhu, et al. [75]
Chest	X-ray, CT, MRI	Islam, Jyoti, et al. [76], Gordienko, Yu, et al. [77], Skourt, Brahim Ait et al. [27], Chen et al. [78], Mittal, Ajay. [79]
Breast	Mammography, MRI, US, CT, Optical imaging	Kallenberg, Michiel, et al. [80], Almajalid, Rania, et al. [81], Caballo, Marco, et al. [82], Cheng, Jie-Zhi, et al. [83], Zelezniak, Roman, et al. [84] Cheng, Ruidaa, et al. [85], Guo, Yanrong, et al. [86], Anas, Emran, et al. [87], Tian, Zhiqiang, et al. [88]
Prostate	MRI, PET, US	Jha, Debesh, et al. [89], Brandao, Patrick, et al. [90], Fan, Deng-Ping, et al. [91], Guo, Yunbo, et al. [92]
polyp	Colonoscopy, CT, MRI	Xu, Lang, et al. [93], Tang, Jinshan. [94], Oliveira, Roberta B., et al. [95], Jin, Qiangguo, et al. [96]
Skin	Surface Microscopy, US, MRI, Confocal Microscopy	Thomas, Simon M., et al. [97]

healthy image, like the presence of a domain compared with the absence of a domain. In the next step, their proposed method learned to add removed tumors to healthy images. They transformed an image into an object of interest. Authors in [44] redesigned a method by skipping long connections in U-Net architecture. They performed nodule segmentation in CT scans of chest images, liver segregation from CT images of abdomen, polyp segmentation from colonoscopy images and videos, and nuclei segregation in microscopic images.

4. 4. Sequenced models

Recurrent Neural Networks (RNN) architecture was proposed in the past to handle the sequences. For example, the Long Short-Term Memory (LSTM) model based on RNN is used for self-loop to enable the gradient flow for more time [46]. RNN has been widely used in medical image analysis to shape the temporal dependency in image sequences. Authors in [47] proposed a method to combine the spatial and temporal information for the image segmentation task by using a fully convolutional network (FCN) combined with RNN. Similarly, for brain MRI, authors in [48] used LSTM with the combination of CNN to enhance the segmentation and temporal relationship in 4D volumes. For pancreas segmentation from CT images, authors in [49] used U-Net for initial segmentation probability and later improved them using LSTM. Some other authors have also applied RNN, LSTM for medical image segmentation tasks [50–52].

5. Optimization functions used for image segmentation

5.1. Cross-entropy

Pixel-wise cross-entropy loss is commonly used for image segmentation. The benefit of using this loss function is that it focuses on each pixel individually. It compares the class prediction values with one-hot-encoder target vector or ground reality values. In the situation of binary

segregation, if $P(Y = 0) = p$ and $P(Y = 1) = 1 - p$. Logistic/Softmax function can be used for prediction and defined as:

$$P(\hat{Y} = 0) = \frac{1}{1 + e^{-x}} = \hat{p} \text{ and } P(\hat{Y} = 1) = 1 - \frac{1}{1 + e^{-x}} = 1 - \hat{p} \quad (1)$$

Where the output is x . Now cross-entropy is to be written as:

$$\text{Cross-Entropy}(p, \hat{p}) = -(p \log(\hat{p}) + (1 - p) \log(1 - \hat{p})) \quad (2)$$

For multiclass (multi-region) segmentation, this equation is to be defined as:

$$\text{Cross Entropy} = -\sum_{\text{classes}} p \log \hat{p} \quad (3)$$

5.2. Focal loss

If the emphasis of the CNN is more on complex examples and to reduce the impact of easy examples, Author [53] added $(1 - \hat{p})^\gamma$ term to the cross-entropy loss function is:

$$\text{FL}(p, \hat{p}) = -(\alpha(1 - \hat{p})^\gamma p \log(\hat{p}) + (1 - \alpha)\hat{p}^\gamma(1 - p) \log(1 - \hat{p})) \quad (4)$$

Yields the balanced cross-entropy (BCE) if we set $\gamma = 0$.

5.3. Dice loss

In image segmentation, Dice loss based on dice coefficient is commonly used. It measures the intersection between two different samples, and F1 score is equivalent to it, which is very commonly used in image classification and segmentation tasks. The range of dice coefficient is between 0 and 1, where Eq. 1 shows that samples overlap entirely. The Dice coefficient (DC) can be calculated as:

$$\text{DC} = \frac{2TP}{2TP + FP + FN} = \frac{2|X \cap Y|}{|X| + |Y|} \quad (5)$$

Similarly, intersection over the union (IoU) known as Jaccard metric [4] can be calculated as:

$$\text{IoU} = \frac{TP}{FN + TP + FN} = \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|} \quad (6)$$

Where TP is the true positive, FN is the false negative, and FP is the false positive. X and Y are anticipated and ground reality segregation values, respectively. We can observe that $\text{IoU} \leq \text{DC}$. When we use DC as a loss function, this can be written as the Dice loss (DL) function [17].

$$\text{DL}(p, \hat{p}) = \frac{2\langle p, \hat{p} \rangle}{\|p\|_1 + \|\hat{p}\|_1} \quad (7)$$

Where p is the ground reality and \hat{p} is the forecasted segregation. And $p \in \{0, 1\}^n$ and $0 \leq \hat{p} \leq 1$, where $\langle \cdot, \cdot \rangle$ is the dot product between p , \hat{p} .

5.4. Tversky loss

It is a generalization of Dice loss (DL) to address the issue of data imbalance, FN and FP level to be controlled, Tversky loss (TL) [55] weighs them as the equation given below:

$$\text{TL}(p, \hat{p}) = \frac{\langle p, \hat{p} \rangle}{\langle p, \hat{p} \rangle + \beta(1 - p, \hat{p}) + (1 - \beta)(p, 1 - \hat{p})} \quad (8)$$

If $\beta = 0.5$ untangle the equation to $\text{DL}(p, \hat{p})$ as mentioned above.

5.5. Exponential logarithmic loss

Authors [56] proposed exponential, logarithmic dice loss to improve the segmentation accuracy for the samples with a large variability in the sample size and small structures. They computed the weighted sum of

exponential, logarithmic dice loss (\mathcal{L}_{eld}) and weighted cross-entropy loss ($\mathcal{L}_{\text{wece}}$).

$$\mathcal{L} = w_{\text{eld}} \mathcal{L}_{\text{eld}} + w_{\text{wece}} \mathcal{L}_{\text{wece}} \quad (9)$$

Where

$$\mathcal{L}_{\text{eld}} = \mathbf{E} [(-\ln(D_i))^{\gamma_D}] \text{, and} \quad (10)$$

$$\mathcal{L}_{\text{wece}} = \mathbf{E} [(-\ln(p_l(\mathbf{x})))^{\gamma_{CE}}]. \quad (11)$$

x Indicates the pixel's position, i is the projected label, and l is the ground reality label. D_i shows the flattened dice loss, γ_{CE} and γ_D used to control the nonlinear behavior of corresponding loss functions.

In medical image segmentation, the issue of class imbalance has decreased the performance of image segmentation tasks. But measures as mentioned above of overlap methods (Dice loss, Tversky loss, and Exponential Logarithmic loss) have shown good performance to overcome class imbalance issues.

6. Applications of semantic-based segmentation

This section illustrates a comprehensive overview of the recently published semantic segmentation arts that use multimodal medical images to diagnose significant diseases, as shown in Fig. 6. The discussed literature will aid a rigorous value to open new research ways to improve researchers' interest in presenting highly efficient products for the severe diseases lays down in the medical image analysis domain.

6.1. Semantic application to skin cancer

Skin lesion segmentation considers a vital indicator for diagnosing skin cancer in dermoscopy samples. However, the variance in skin lesion shapes, color, texture, and the existence of poor contrast between regular features and lesions make crucial the process of skin lesion segmentation. In literature, the involvement of a deep learning technique called convolution neural network (CNN) has shown phenomenal performance for skin lesion segmentation. Hasan et al. [98] projected learned discriminated features on pixel-level using depth-wise separable convolution instead of standard convolution to segment skin lesions, as shown in Fig. 7. Their network outperformed U-Net and FCN8s having 3.6% and 6.8% mean intersection over union (mIOU) on the ISIC-2017 dataset of dermoscopic images.

While in a previous study [99], a topology named fully convolution and dual path (FC-DPN) was implemented to perform skin lesion segmentation, as shown in Fig. 8. The authors fine-tuned FC-DPN by employing sub-DPN projection and sub-DPN processing blocks and captured the most representative and distinguishable features to achieve robust segmentation results. The experimental study incorporated two datasets named ISBI 2017 and PH2 and obtained 88.13%, 80.02% of Dice coefficient and Jaccard index values on ISBI 2017 and 90.26%, 83.51% on PH2, respectively.

Another alarming factor called hair existence in dermoscopy samples affects the diagnostic accuracy of skin lesions. Li et al. in [100] proposed a specialized deep learning framework to remove digital hair with high accuracy. They trained U-Net on the train set of the ISIC 2018 benchmark to achieve precise hair masks and then presented a novel free-form image inpainting network (Gated Convolution and SN-PatchGAN) for the inpainting of any hair gap as exhibited in Fig. 9. The hair gap inpainting task is evaluated using their proposed structured similarity (SSIM)-based single-image digital hair removal art called Intra-SSIM and outperformed existing approaches.

With the application of images of dermoscopic skin laceration, Jin et al. [101] presented a new idea of waterfall dissemination of knowledge in various sub-networks. When combined with cutting-edge skin laceration diagnosis and segregation algorithms, the proposed CKDNet architecture produced competitive results without any additional data. Compared to the ISIC 2017 and 2018 datasets, the provided model had a

superior JA of 0.800 and ACC of 0.946, contrasting the techniques. Fig. 10 represents the segmentation results.

Gu et al. [102] implemented a technique that combines an advanced network structure DE-Net with a unique loss function to look deeply into the skin laceration boundaries to obtain more refined segregation and to enhance further the skin laceration boundaries on necessary segmentation results as illustrated in Fig. 11, the novel entirety-center-edge loss function is suggested. The EIGM, which includes all of these procedures, has a Dice score of 0.8662 and a Jac of 0.7887, which is higher than the prior methods' statistics.

By applying FCNs, Kaymak et al. [103] suggested an extensive investigation for a skin laceration segregation application that will lead to melanoma identification, the fatal type of skin cancer. The application used the FCNAlexNet, FCN-8s, FCN-16s, and FCN-32s architectures, all well-known FCN architectures. The FCN-AlexNet, FCN8s, FCN-16, and FCN-32s models perform much better, with accuracy rates of 94.461 percent, 94.811 percent, 94.5818 percent, and 94.5202 percent, respectively. Fig. 12 demonstrates the visual example of segmentation results the .

Sarker et al. [104] proposed SLSNet, a GAN-based model for skin laceration segregation that is effective and easy. A 1-D kernel factorized network, variable scale aggregation, channel attention, and position mechanisms are included in the proposed model. SLSNet summarises the findings and contributions on the ISBI 2017 test and ISIC 2018 validation datasets. It provides precise segregation results with sensitivity, accuracy, specificity, Jaccard index, and Dice coefficient of 87.81 percent, 97.61 percent, 99.92 percent, 81.98 percent, and 90.63 percent. Moreover, the Jaccard index of 87.81 percent, 97.61 percent, 99.92 percent, 81.98 percent, and 90.63 On the ISIC 2018 validation dataset, the result shows a threshold JSC score of 78.4 percent was achieved. A visual representation of their lesion segmentation outputs is given in Fig. 13.

Lei et al. [105] presented an unconventional yet effective GAN model for skin laceration segregation using dermoscopy pictures. This model comprises two modules. The first bypass the connection, and the other uses the dense dilated convolution block in the UNetSCDC module to improve distinctive feature representations and save more detailed information on authentic segregation. The significant conclusions from the initial review are that the DAGAN model, with 0.935 accuracies, outperforms the existing UNet-SCDCcon and UNet-SCDCgen models, which have 0.900 and 0.889 accuracies. Fig. 14 illustrates the proposed segmented method results.

Tang et al. in [106] demonstrated a unique DL-based model for skin laceration segmentation, as seen in Fig. 15 on a dermoscopy picture collection. In this model, the AFLN is adjusted to increase variable scale feature learning capacity, while DGCL is employed to overcome the problem of overfitting. Additionally, the STBCRs address the issue of over-segregation. Compared to existing models, the combination of AFLN DGCL and AFLN DGCL framework produces a significantly higher accuracy score of 0.963 and 0.966, respectively.

Arora et al. [107] introduced an AG-based customized U-Net model for automated skin laceration segregation, as shown in Fig. 16. The Attention Gates (AG), which eventually adds Tversky Loss (TL) as the output loss function, pays close attention to the skip connection's tiniest and most subtle elements. The GN efficiently extracts the feature maps, and the AGs are in charge of digging out high-dimensional information from low-dimensional data. Extensive testing on the ISIC 2018 dataset revealed that this strategy achieved superior accuracy and a Jaccard score of 0.95 and 0.83, outperforming cutting-edge segregation strategies.

6.2. Semantic application to breast cancer

The most common type of cancer in women worldwide is breast cancer. It is also the leading cause of death. If identified early, it could save the patient's life. Breast ultrasound (BUS) has a poor detection rate

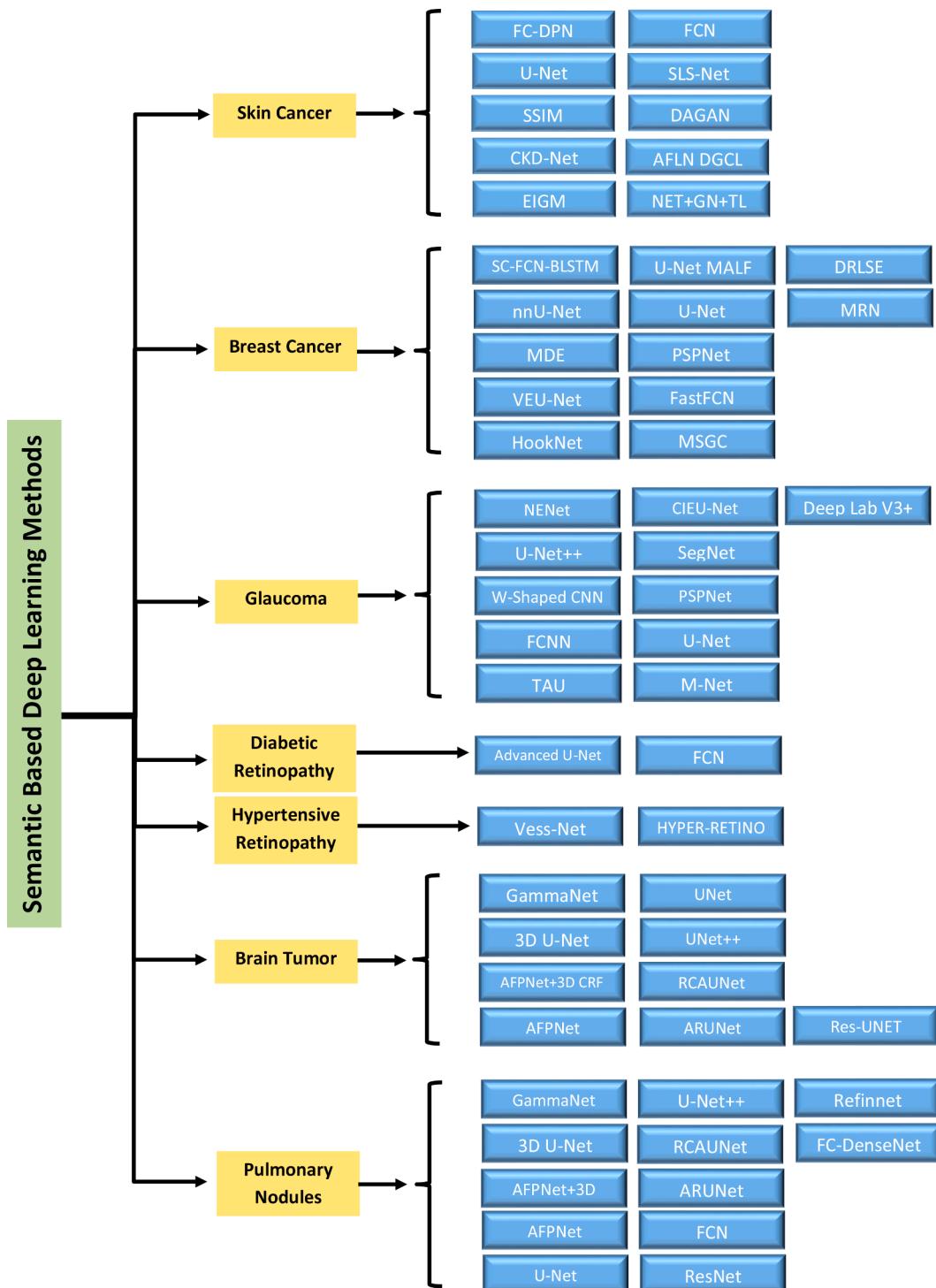


Fig 6. Representation of different deep-based semantic segmentation models described in this review paper.

and relies highly on the operator's abilities. Automated breast ultrasound (ABUS) has grown in popularity as a result of the potential limitations of ultrasound. Recent theoretical advances have been achieved in classifying lesions based on area and contour features. The SC-FCN-BLSTM network was presented by Pan et al. [108] for ABUS tumor segregation. The SC-attention module is customized to account for usable, finer-grained geographical information and semantic information. SCBLSTM module is available on top, which helps decrease false positives and improve segmentation efficiency. This method was notable since it outperformed most cutting-edge approaches for tumor

segregation, with a DSC of 0.8178 and an accuracy rate of 0.8292. Fig. 17 depicts the visual example of the original image and segmented result using the proposed method.

A framework for FGT and breast segmentation was designed by Huo et al. [109] on the principle of deep-learning through nnU-Net. This model was used to measure the volume of breast gland tissue. The method provided high accuracy and segregation across various Mr imaging and breast categories. It did so without relying on additional pre and post-processing procedures. Conclusively, a model based on machine learning deduced significant outcomes for FGT and breast

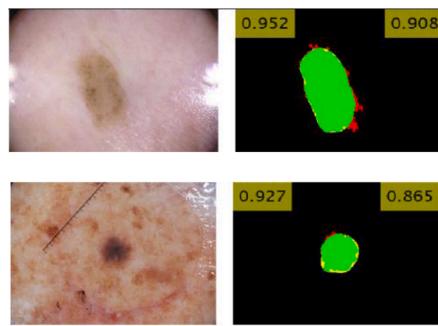


Fig 7. Skin lesion segmentation results by [98] where the left column represents the original samples and the right column show the segmented masks having Dice coefficient (top-left) and IoU (top-right).

segregation. [Fig. 18](#) shows the segmentation results.

A digital mammogram is used to detect breast cancer and its segregation, as shown in [Fig. 19](#). Modified Differential Evolution (MDE) is an upgraded form of the Slime Mold Algorithm proposed by Liu et al. [110] where MDE-based multilevel models produced unprecedented image segregation results. It was extensively validated on publicly available databases. In comparison to other imaging models, the advanced model produced notable outcomes for duct breast invasive ductal carcinoma image datasets.

Pawar et al. [111] used this technique to analyze the pectoral muscle. The heuristic approach computed significant results of 86.16% in contrast to the non-heuristic approach with an 82.23% accuracy rate. This algorithm works with mammograms of various shapes, sizes, and textures and generates noteworthy outcomes. [Fig. 20](#) shows the different segmentation results.

Another segregation model that can detect and reduce noise was presented by Zou et al. [112]. During model training, the network was monitored with the probability distribution of samples to differentiate noises. The model's noise index and output for the two parameters generate output like the previous version. The function of the technique is to get more feature information on segmentation. In contrast to models in a publicly available dataset BUSI, the method's output presented in this paper is more robust and superior, with 87.2 % precision and 88.6 % recall, respectively. [Fig. 21](#) demonstrates the segmentation results.

Yan et al. [113] presented a two-stage model based on coarse-scale mass detection with a multi-stage detection strategy. It uses dense and nested skip connections to incorporate fine-scale mass segregation. The outputs on the INbreast dataset confirmed that the presented model surpasses other techniques in terms of robustness and generalizability.

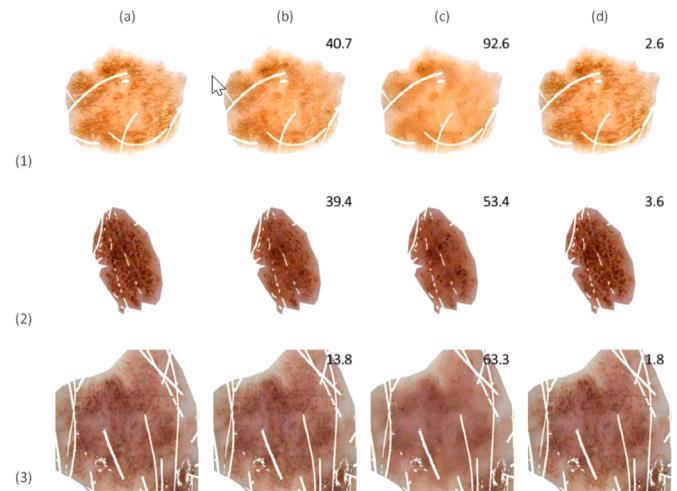


Fig 9. A visual example to show the texture change inside the tumor region, where figure (a) shows the input dermoscopic image, (b) represents the image inpainting result for hairs removal after using (b) DullRazor [8], (c) Huang et al. method [9], and (d) their proposed system. The value of MSE [100] is also calculated.

The model generated an outcome of 80.44% Dice, equal to avant-garde performance in mass segregation on the publicly available database. [Fig. 22](#) exhibits the segmentation results.

A model based on VEU-Net for semantic segregation is proposed by Ilesanmi et al. [114]. It works by using CLAHE to remove noise, followed by semantic segregation with the Ve block and convolution method to segregate tumor cells. Two datasets of authentic breast ultrasound images were used to measure HD, JM, and DM values. In comparison to state-of-art methods, this model resulted in highly significant results. The segmentation results of the proposed method are shown in [Fig. 23](#).

Huang et al. [115] presented an advanced method with three types of features to extract the semantics of superpixels. It uses the BoW model to label the superpixels as tumors. The proposed model's efficacy outran existing models with a significant outcome of 90.20 % and 89.53 F1-Score for Benign and Malignant tumor datasets, respectively. The segmentation outputs can be shown in [Fig. 24](#).

Rijthoven et al. in [116] presented HookNet, a model that measures high dimensional resolution tissue segregation. The framework was applied to two different publicly available datasets of high-resolution tissues. Contrary to single-resolution models, the suggested framework raises overall efficiency and can deal with high-resolution and contextual information with narrow differences. The HookNet model outran

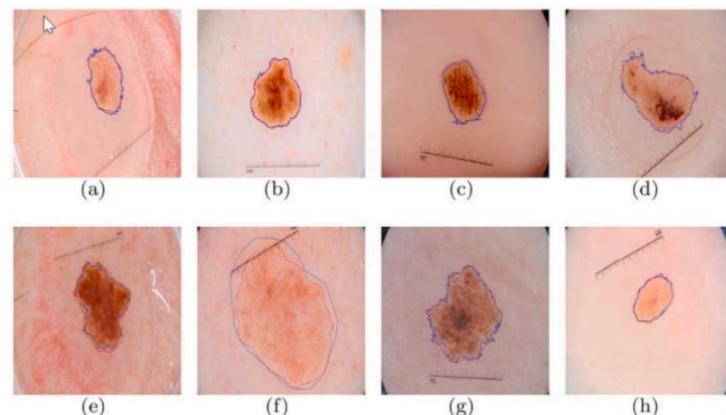


Fig 8. Given are eight samples of skin lesion segmentation results on test images, where (a)–(d) are nonmelanoma lesions, and (e)–(h) are melanoma lesions. The red contour shows the segmentation result, and the blue depicts the ground truths [99].

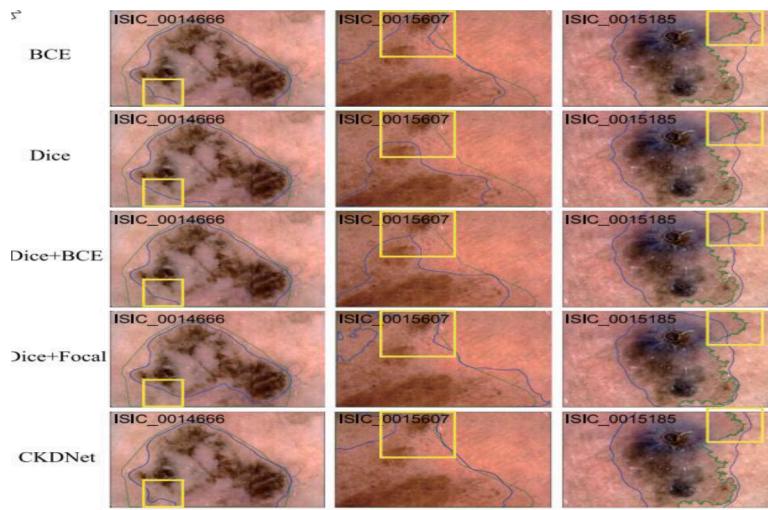


Fig 10. An example of tumor segmentation results proposed by different loss functions and the CKDNet [101] method.

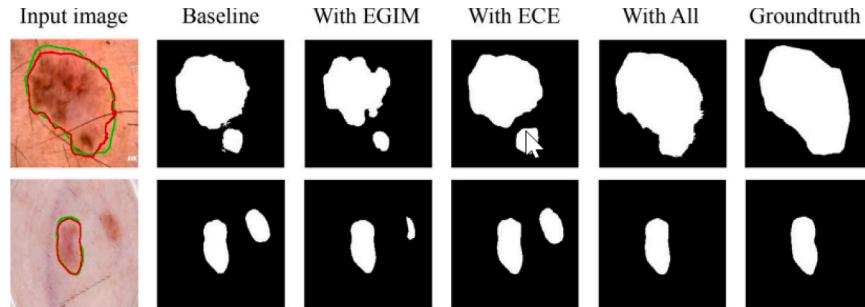


Fig 11. Another visual example of segmentation results obtained by the DE-Net technique [102]. In this figure, the green contour is ground truth (GT), and the red contour is the segmentation results of the ISIC-2017 dataset.

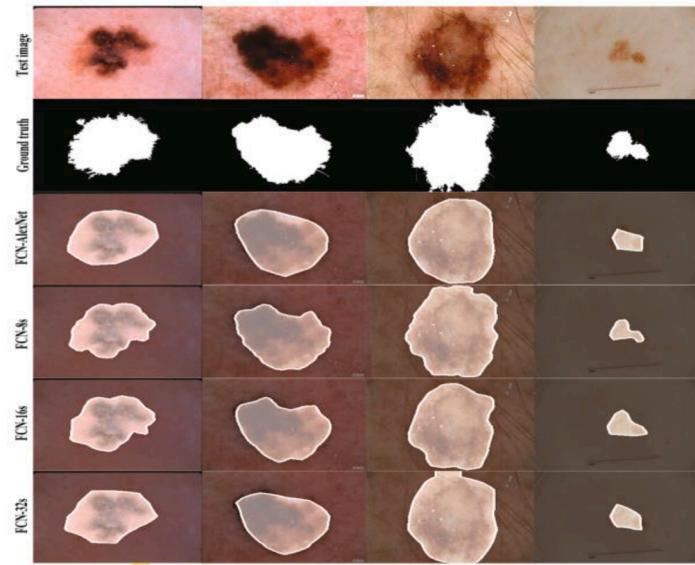


Fig 12. A visual example of FCN network architecture [103] for dermoscopic lesion delineation contour.

other tools with a 0.91 F1-Score overall. Fig. 25 shows the segregation results using the HookNet model.

Tong et al. [117] proposed a solution to improve brightness, contrast, and low-quality ultrasound images of breast tumor segregation. It involves replacing the convolution module in the left encoding

path with the residual convolution module and extended residual convolution module. The features were extracted efficiently from ultrasounds of breast tumors. Four attention loss functions were integrated into the cross-entropy loss feature to attain more efficient network loss values and precisely locate tumor targets. The model performed its

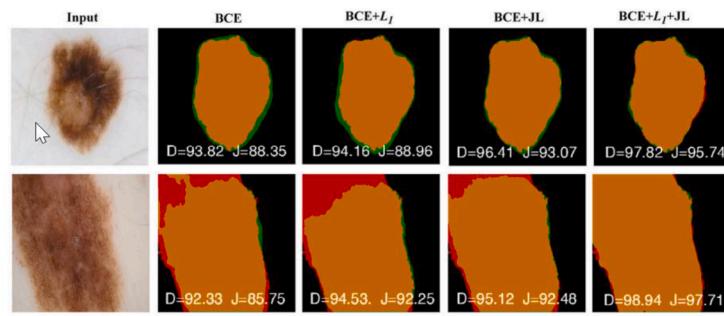


Fig 13. An example of various loss functions when calculated on ISBI 2017 dataset for lesion segmentation by [104].

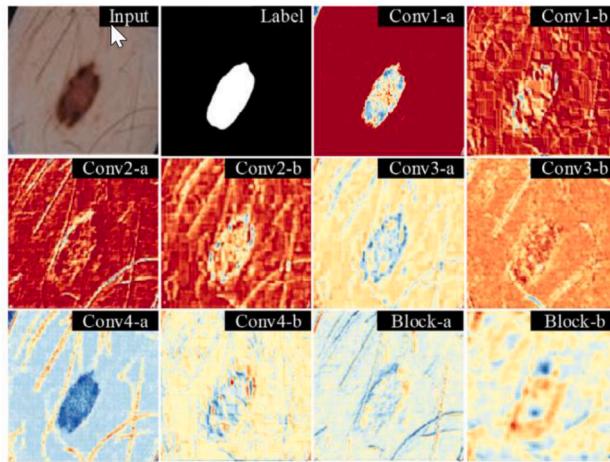


Fig 14. Intermediate feature visualization of proposed DAGAN. The first and second images are the original input and ground truth, respectively. Two feature maps are randomly extracted from each layer. These feature maps illustrate that the network can extract adequate information from target regions [105].

functions without increasing the overall cost. Three hundred sixteen images were tested to better evaluate the model's applicability. The segregation performance was significantly better and generated a 0.873 F1-Score and 0.959 accuracy value. Fig. 26 shows the segregation results of breast ultrasounds images.

The most significant field of mammography research is the early identification of breast cancer. Therefore, the study in [155] investigated the cutting-edge deep segmentation model for the detection of breast tumors. They evaluated several deep-based segmentation models, namely, Dilatation 10, Deep lab v3, FCN, and U-Net on newly prepared datasets. The result shows that dilation 10 outperforms the other three

segmentation models by obtaining pixel accuracy of 92.98% in the comparative test.

6.3. Semantic application to retinograph images

In this portion, different retinal disorders such as glaucoma, diabetic retinopathy, and hypertensive retinopathy are demonstrated by employing semantic-based scientific algorithms.

6.3.1. Semantic application to glaucoma

Glaucoma is an eye disease and the second most common eye-related anomaly [164]. It is also called optic neuropathy causes retinal ganglion cell apoptosis and damages the optic nerve. Loss of vision is one of the symptoms of glaucoma that can gradually lead to blindness. The chances of glaucoma can be reduced through treatment if the disease is diagnosed at an early stage. However, the pathological condition becomes irreversible in case of severe nerve damage. Symptoms of glaucoma only appear during the advanced stage of the disease. Hence, several researchers have made significant contributions using deep learning techniques to restrict the damage to the eyes in the initial stages.

A deep learning-based algorithm to semantically segment OD and OC was presented by Imtiaz et al. [118]. The proposed model illustrated improved and significant outcomes on two public datasets: Drishti and Rim-one. VGG-16 was used to classify three categories, i.e., OD, OC, and background. Next, feature vectors are applied to a softmax classifier that categorizes them on the base of pixels. The algorithm calculated comparable results on the Drishti dataset for segmented OD and OC. Fig. 27 shows the segmented results of glaucoma features.

The NENet patch-based adversarial model was proposed by Pachade et al. [119] to study OD and OC segregation, as shown in Fig. 28. It was used to perform extensive experimentation on three datasets, i.e., REFUGE, Drishti, and RIM-ONE-r3. The algorithms showed significant findings related to OD and OC segregation. In contrast to NENet-seg, the newly designed model accurately determined OD and OC boundaries.

An advanced model is suggested by Bian et al. [120]. It works by

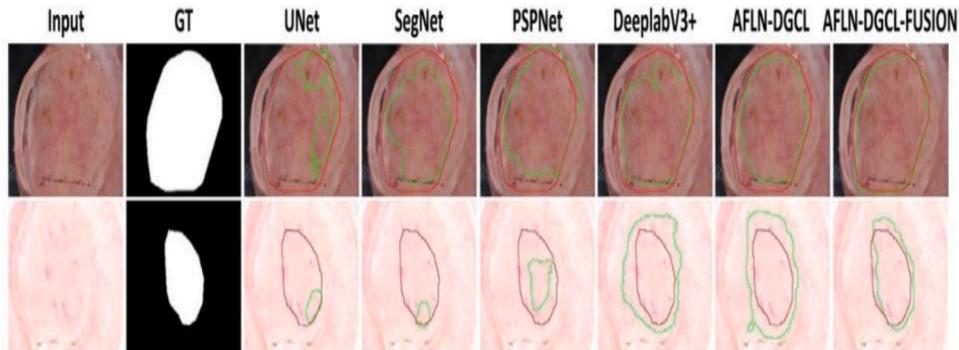


Fig 15. The segmentation results obtained by different segmentation models. The red boundary is obtained by ground truth, and the green boundary is obtained by each prediction [106].

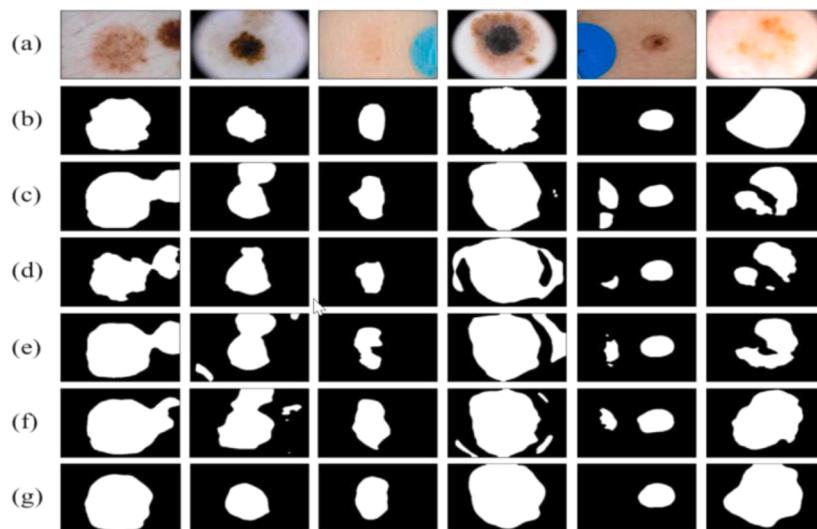


Fig 16. An example of segmentation results on dermoscopic images by the proposed model [107] and compared with the conventional methods, where figure (a) input dermoscopy image, (b) Given ground truth image, (c) segmentation results of U-Net, (d) the segmentation results of SE block on basic U-Net, (e) results of BCDU network (with 1 dense unit), (f) results of U-Net (with all 64 filters) network and (g) results of the proposed Att_U-Net+GN+TL model.

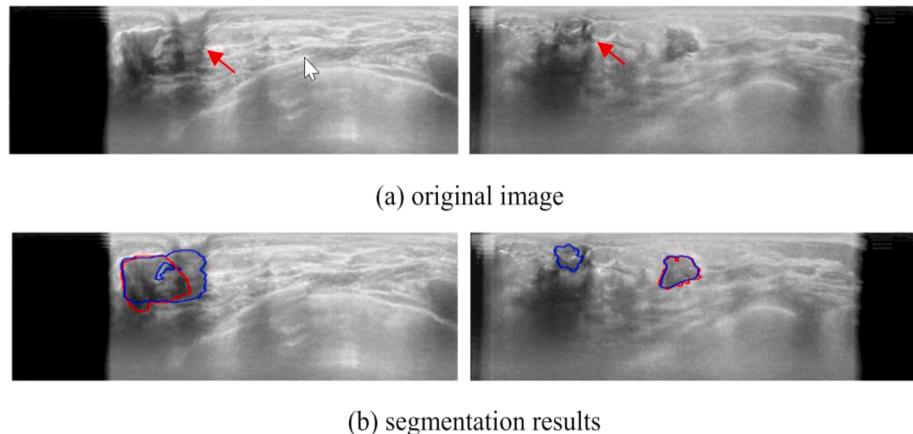


Fig 17. A visual example of wrong detection results, where figure (a) shows the original input image. (b) is the segmentation results using the SC-FCN-BLSTM method [108].

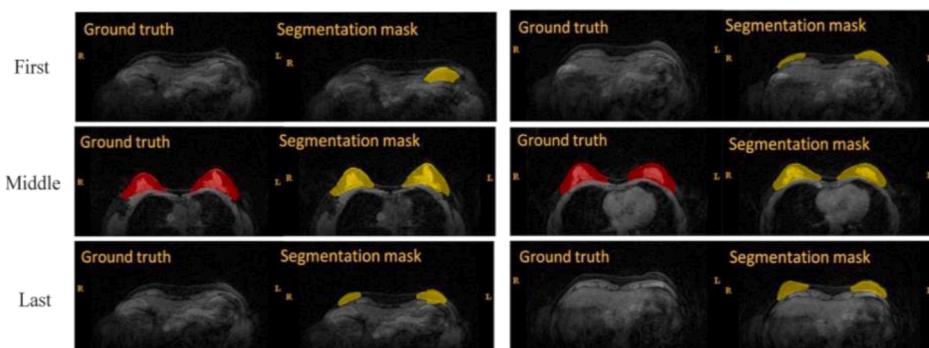


Fig 18. Segmentation result produced in [109] by automatic segmentation technique compared with ground truth of MR scan from the three slices such as first, middle, and last slices.

segmenting optic disc (OD) and optic cup (OC) without relying on post-processing or ROI spotting, as shown in Fig. 29. Segmentation is further improved using the cGAN structure, and the overfitting issue is mitigated to lower learning rates. Moreover, it is an advanced way to automate OC and OD segregation. The model outran existing tools with

0.8887 and 0.9367 accuracy rates and segregation values.

Tulsani et al. [121] displayed a new segregation-based approach on the optic cups and optic disc to identify glaucoma. It is a custom model based on the UNET++ custom loss function and built to conduct segregation experiments. The model was extensively tested on

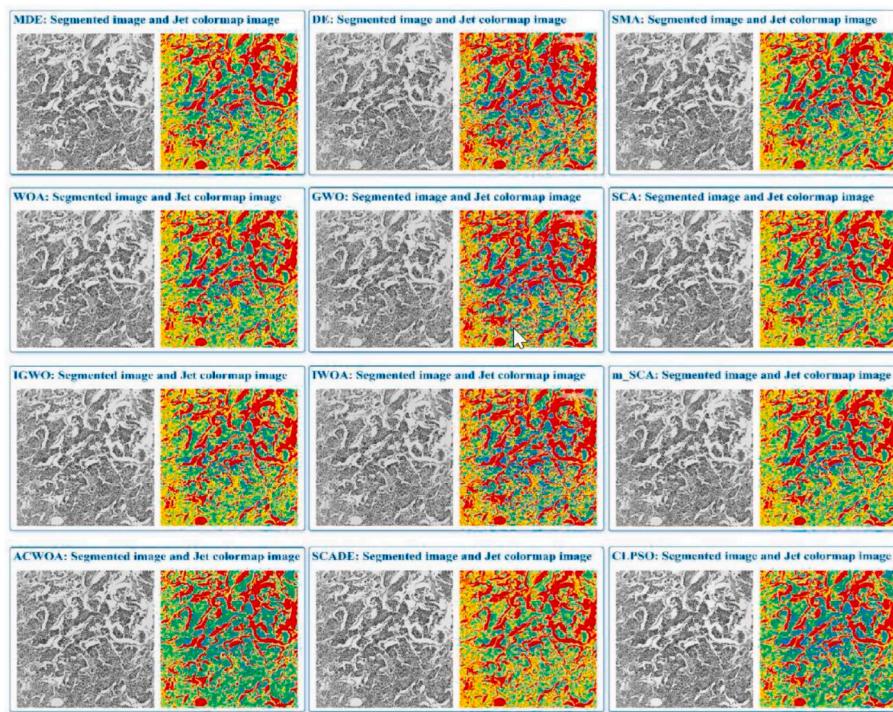


Fig 19. The segmented results at threshold value 20 [110].

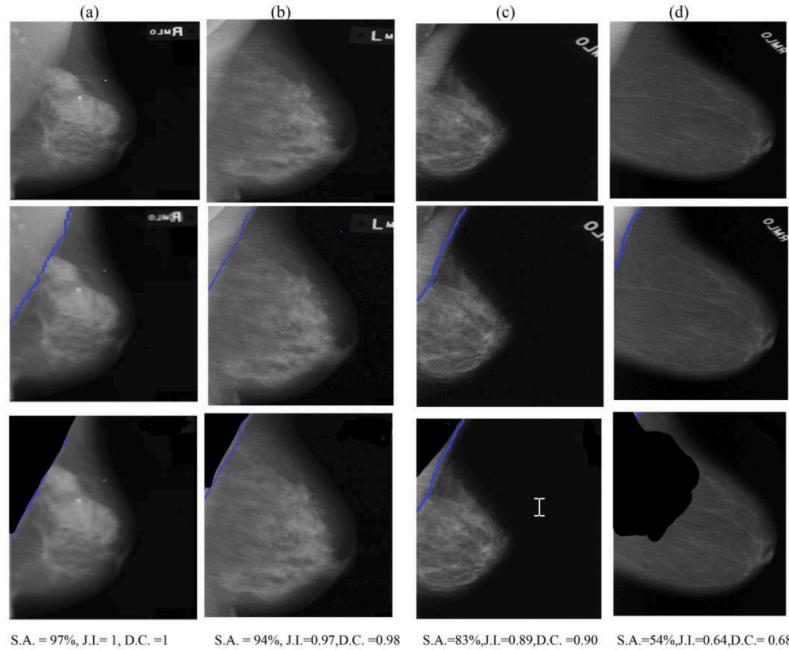


Fig 20. Subjective and objective assessment of the proposed algorithm (a) Correct segmentation (b) Acceptable segmentation (c) Under segmentation (d) Over segmentation [111].

open-source datasets such as RIM-ONE, DRIONS-DB, and ORIGA. It deduced comparatively significant results parallel to the avant-garde algorithms. Fig. 30 shows the segmentation results.

Zhou et al. [122] proposed a model to improve the retinal vessel segregation capability, especially in low-contrast background and laceration regions. It uses SEGAN and MSFRB to refine the process of extracting information. AM improves the overall performance by distributing more attention to the discriminative feature. Extensive studies with the model produced highly satisfactory outcomes. It

surpassed the efficacy of existing methods with 0.9083 and 0.8211 accuracy rates on the CHASEDB1 and HRF datasets, respectively. Fig. 31 shows the segmentation results.

Yin et al. [123] tailored a level set loss to cater to the optic disc and cup segregation issue, as shown in Fig. 32. Through CNN output as a level set, you can add smooth boundary and region consistency and other constraints for the segregation. The main advantage of this model is the ease of adding constraints. The level set loss also gives more supervision. The Dice Score on the DRISHTI-GS dataset gave comparable

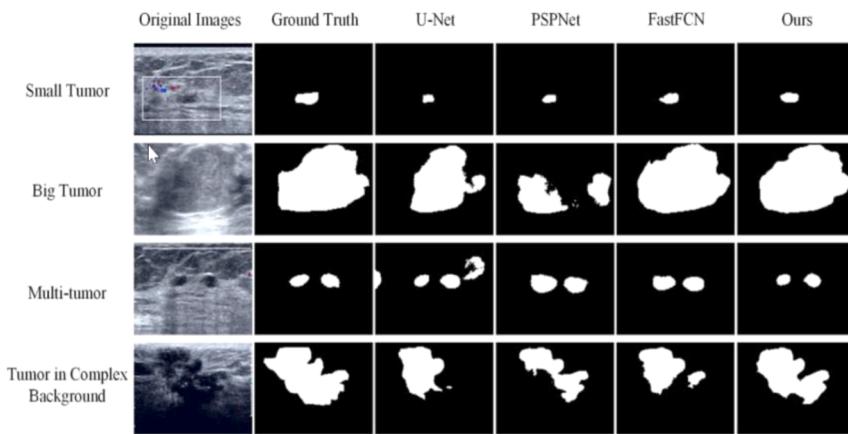


Fig 21. The segmentation results were compared under four conditions on BUID [112].

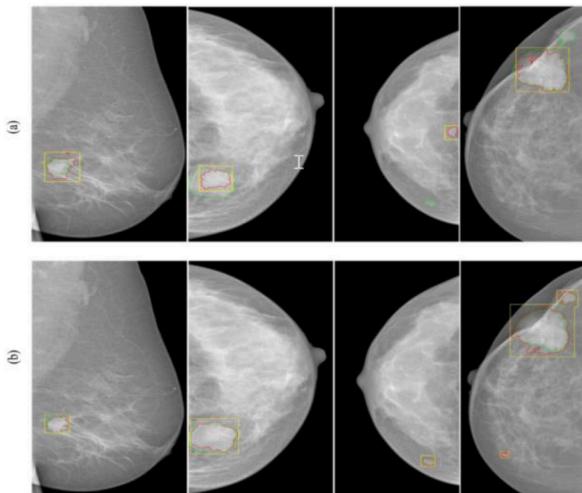


Fig 22. Mass segmentation using two-stage method without (a) and with (b) multi-scale fusion (MSF). Yellow, red, and green stand for final detection, segmentation, and ground truth [113].

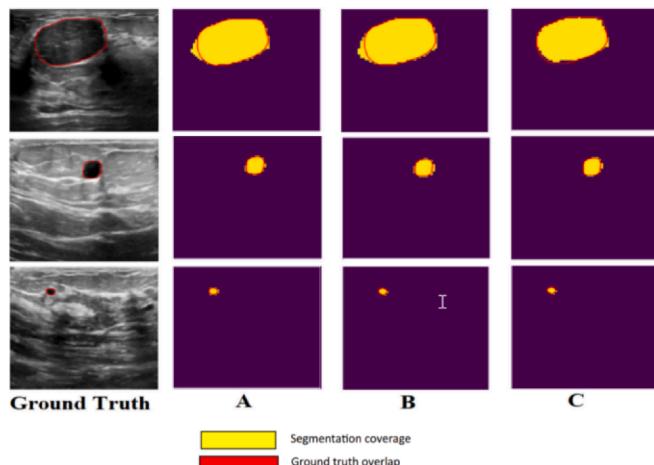


Fig 23. Results of augmentation and ablation (LOGIQ E9 and LOGIQ E9 Agile): (A) Results of Criteria 1; (B) Results of Criteria 2; (C) Results of the proposed method [114].

results to the other existing designs.

Yuan et al. [124] proffered a deep learning CNN structure on a multi-scale for segregating OC and OD collectively as a multi-label single-stage process visualized in Fig. 33. The given model with W-shaped CNN extracts more features, reduces the semantic distances between shallow and deep features, and uplifts the merging of features. After testing four publicly available datasets, the presented model gave sufficiently excellent, and comparable results. While during OC segregation, this model has an upper edge over others.

Morano et al. [125] published a modern approach for the simultaneous segregation and classification of the retinal arteries and veins (SSCAV) using FCNNs. This approach segregates arteries, veins, and the whole vascular tree separately as three different tasks. Extensive results have been compared with the state-of-the-art models by giving comparable figures in the segmentation process. The segmentation results as shown in Fig. 34.

Zhang et al. [126] presented the TAU model to address domain adaptation issues in the OD and OC segregation tasks in Fig. 35. The proposed model was tested on DRISHTIGS, RIM-ONE v3, and REFUGE datasets to check its effectiveness. The experimental results show that TAU outperformed the existing methods in the segregation performance and CDR MAE.

Sun et al. [127] suggested a CIEU-Net for segregation of retinal vessels, being a deep-learning segmentation framework. It gained satisfactory results for retinal vessel segregation that illustrate the benefit of combining semantic segregation modules and basic medical image segregation methods. The model was tested on the CIEU-Net with BN and CHASE_DB1 Dataset. It attained the avant-garde performance in F1-score, Recall, MCC, and AUC. While on dataset named STARE, it gained the avant-garde performance in F1-score, Recall, MCC and AUC. The segmentation results are illustrated in Fig. 36.

6.3.2. Semantic application to diabetic retinopathy

One of the most known forms of diabetic eye lesion is Diabetic retinopathy (DR). It can harm the patients and become extremely dangerous, leading to blindness if neglected. This can be cured only if diagnosed at an early stage. The automatic diagnosis systems aided with computers, could help in the early and fast diagnosis of its possible symptoms with a reasonable amount of expenditure and extended coverage. Researchers have given their time and focus on this disease to help diagnose DR patients at an early stage.

Sambhyal et al. [128] presented a modern and advanced UNet architecture that consists of a residual network and employs shuffling periodically with sub-pixel convolution. The presented framework has been trained and tested for microaneurysm and hard exudate segregation on IDRID and e-ophtha. The given model achieves avant-garde results for retinal laceration segregation. The segmentation results are shown in Fig. 37.

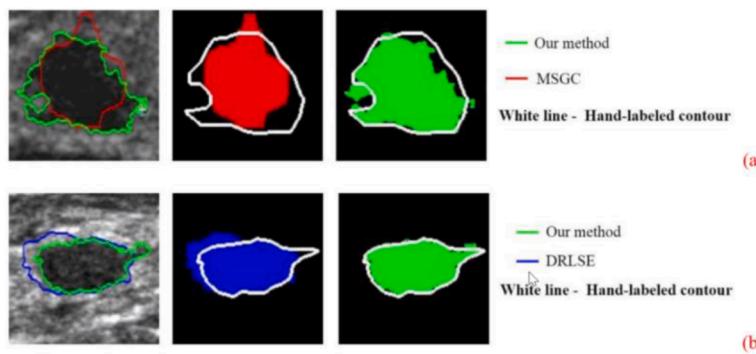


Fig 24. Suggested method being compared with MSGC and DRLSE methods. (a) Proposed method: ARE, F1 Score, TP and FP score is 13.53%, 92.26%, 90.17% and 5.29%, respectively. MSGC: ARE, F1 Score, TP and FP score is 15.75%, 79.90%, 69.50% and 4.47%, respectively. The FP score of MSGC achieved better, but TP was worse, and (b) proposed method: ARE, F1 Score, TP, and FP score is 6.26%, 92.54%, 95.95%, and 11.41%. DRLSE: ARE, F1 Score, TP and FP score is 12.06%, 83.75, 97% and 34.63%. The TP score of DRLSE achieved better, but FP was worse.[115].

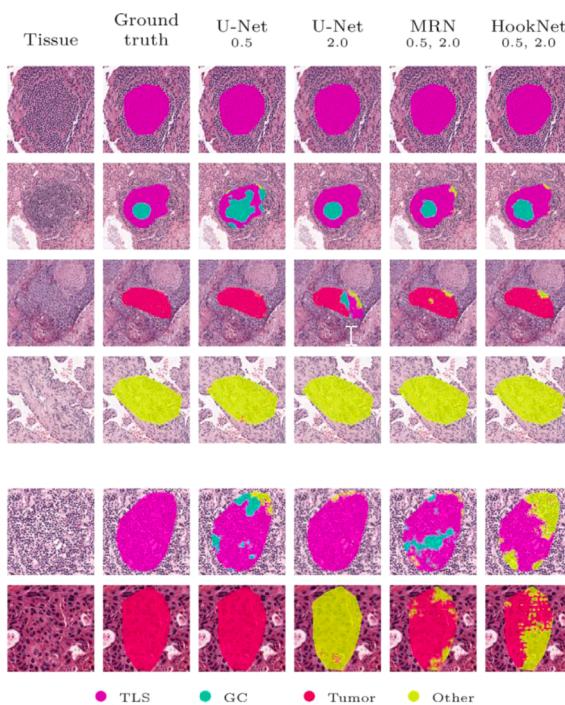


Fig 25. Segregation results of GC, Tumor, TLS, and Other on lung tissue. HookNet results are shown for Lambda is set $\lambda = 1.0$ for HookNet giving results along with failure examples of HootNet in the last two rows [116].

Saha et al. [129] published a method that solves lacerations' segregation issue. The severity level can also be found by segmented output, as shown in Fig. 38. This method solves six subtasks in a single solution. Whereas previous work in automatic detection of diabetic retinopathy deals with identifying the disease stage. After extensive preprocessing and several steps, it is achieved to reach the result finally. The evaluation of the localization of the optic disk for this proposed method gave substantial results compared to the results of the existing models.

Saha et al. [130] presented a multi-adversary-based fully convolutional neural network for retinal anatomy and pathology segregation by taking support of weakly labelled fundus images. An FCN with skip connections passes the context information to higher resolution layers for more precise segregation. The model utilizes two discriminators. The first makes the result as realistic as possible, while the second discriminator enforces segregation of classes not annotated for a particular image. The proposed method surpassed the existing avant-garde techniques by SE, SP, ACC, and AUC scores with 0.7906, 0.9839, 0.9641, and 0.9812 values. The result of segmentation is demonstrated in Fig. 39.

A novel active deep learning strategy is utilized in [163] with a traditional CNN network, comprising only 7 layers to accurately

segment diabetic retinopathy lesions in colored retinal samples, as shown in Fig. 40. Their proposed art obtained promising results on 80 thousand images from various retinal benchmarks.

Ayoub et al. [165] proposed a modified CNN U-Net network for identifying retinal hemorrhages in fundus images. The proposed UNet was trained using the GPU and the IDRiD dataset to segment and identify possible regions that might contain retinal hemorrhages. The proposed methodology's sensitivity, specificity, and accuracy were 80.49%, 99.68%, and 98.68%, respectively. The experimental outputs also produced an IoU of 76.61% and a Dice value of 86.51%, demonstrating the effectiveness of the network's predictions and their potential to lessen ophthalmologists' workloads significantly. Fig. 41 shows the segmentation result using the proposed method.

6.3.3. Semantic application to hypertensive retinopathy

Retinal damage is most caused by Hypertensive Retinopathy (HR). It is pertinent to know that the number of people projected by hypertension is 1.56 billion worldwide. Out of which, around 66% of them hail from impoverished countries, where there are no proper healthcare facilities to identify and tackle hypertension. Hypertension also causes headaches, and nosebleeds and may also cause vision loss. Moreover, if hypertension persists for a long time, it can damage the lungs, heart, kidneys, and eyes and even lead to cardiovascular pathologies resulting in death. The risk may be reduced if the HR could be identified and treated early. Only two HR stages can be recognized using a few systems developed using deep learning techniques. It is difficult to design a system that could identify the five stages of Hypertensive Retinopathy. In addition, researchers have used deep features in the past, but the accuracy score to classify is not as expected.

Arsalan et al. [131] published a dual-residual-stream-based vessel segregation network (Vess-Net). The suggested model does semantic segregation to aid the diagnosis of hypertensive retinopathy using artificial intelligence. The suggested Vess-Net model has been extensively validated on publicly available datasets for vessel segregation. Pragmatic results illustrate that Vess-Net attained better results for all given datasets giving the most robust results. Fig. 42 shows the vessel segmentation results.

Abbas et al. [132] presented advanced hypertensive retinopathy (HYPER-RETINO) framework, which segments the HR based on five grades. The HYPER-RETINO system is implemented based on pre-trained HR-related lesions. Several steps are implemented to develop the framework: preprocessing, HR-related lacerations detection by instance-based and semantic segregation, and a DenseNet framework to identify the HR classes. The HR system obtained satisfactory results to segment HR based on five grades. Fig. 43 gives the visual example of semantic and instance segmentation of HR lesions.

6.4. Semantic application to brain tumor

Gliomas are the most common malignant brain tumors originating from abnormal glial cell growth. Gliomas are responsible for 80% of

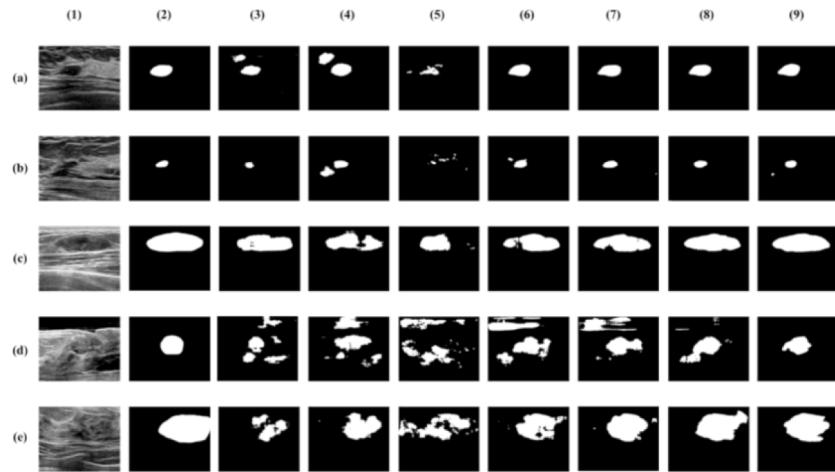


Fig 26. Breast ultrasound picture segregation results. The Fig. (a1)–(e1) are the original ultrasound breast images and Fig. (a2)–(e2) are the corresponding ground reality annotation. The Fig. 5(a3)–(e3), (a4)–(e4), (a5)–(e5), (a6)–(e6), (a7)–(e7), (a8)–(e8), (a9)–(e9) are FCN16s, FCN8s, SegNet, U-net, Attention U-net, Attention U-net with mixed attention loss, Improved U-net MALF network segregation results, respectively [117].

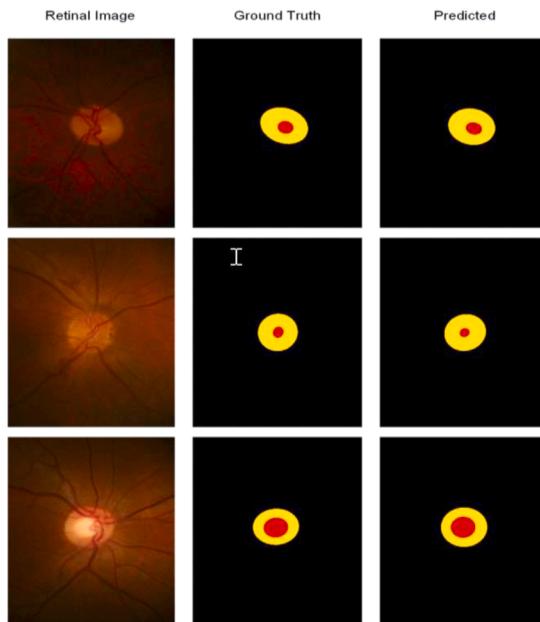


Fig 27. The segmented OD and OC results for Rim-one [118].

malignant brain tumors with a high mortality rate. Many researchers have spent their day and night helping to develop a computer-aided brain tumor diagnosis system to improve the survival rate of glioma patients. Since gliomas can originate at various locations with various sizes, shapes, and intensity distribution, it is challenging to segment them efficiently and correctly.

Huang et al. [133] published a computerized brain tumor segregation system that seeks help from the adaptive gamma correction neural network (GammaNet). It focuses on significant and malignant regions and recognizes intensity invariance. In addition, the feature maps are distributed into many proposed regions and emphasize the local image characteristics. Furthermore, combining pooling (DenseASPP) modules and AGC blocks reduce the information loss by enlarging the receptive field and improving the segmentation performance. The extensive experimental results conclude that the GammaNet can achieve avant-garde performance. Fig. 44 shows the different model segmentation results.

Ye et al. [134] presented a model for medical images which portrays high-level supervision into lower layers of neural networks and is a 3D Center-crop Dense. The two pathways are the attention pathway and the context pathway. The first one is of acceptable resolution focusing on detailed information of every voxel, while the context pathway works with a low-resolution focused on the extra surrounding details. Furthermore, the missing information is due to downsampling and dealt with by introducing cross-pathway connections from the attention pathway to the context pathway involving weights. The proposed model produces the segregation performance, which corroborates the avant-garde results.

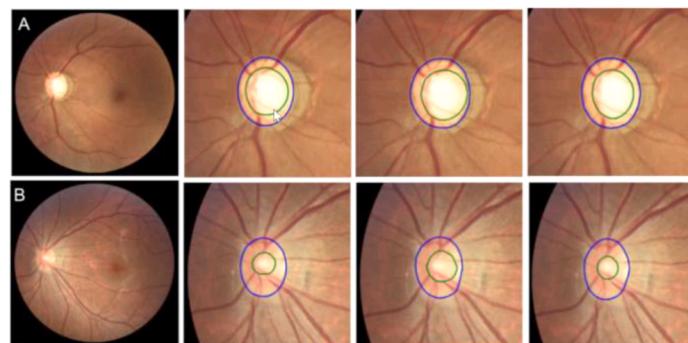


Fig 28. Qualitative results on REFUGE dataset. First column: original fundus images; Second column: ground truth; Third column: results of NENet-seg framework; Forth column: results of NEENet framework. The blue and green colors indicate the boundaries of OD and OC, respectively [119].

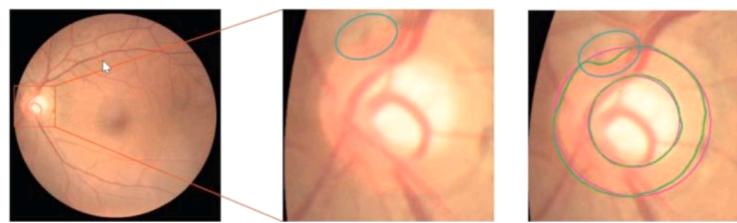


Fig 29. Typical sample in REFUGE. The left image is the RGB retinal fundus image. The middle one is the partially enlarged drawing of the left image framed by a red square. The right one is the comparison between the reference mask and our result. The ground truth provides the red circle, and the green one is predicted by the proposed model [120].

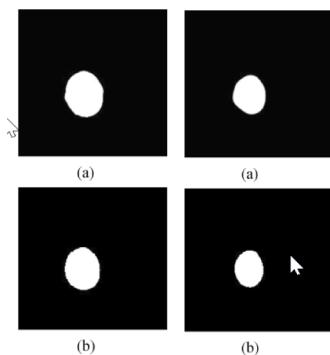


Fig 30. Displaying segmentation results, (a) Predicted binary mask and (b) Original binary mask [121].

Zhou et al. [135] presented a model with a three-stage network. The first stage produces extra context constraints for every tumor region using a 3D U-Net segregation network. The attention mechanism is then used to fuse with the Multi-sequence MRI to achieve three independent tumor regions. A new loss function is involved per the location to make the segmentation issue more efficient. Finally, all three results are combined by another 3D U-Net. The resulting outcome was promising and comparable by considering dice score, Hausdorff distance, and the

amount of memory required for training could be visually represented in Fig. 45.

Zhou et al. [136] proposed an advanced model that integrates with the backbone and uses a 3D atrous-convolution feature pyramid to learn the tumor's essential features and substructures. Using a 3D fully linked conditional random field, accurate predictions along borders are obtained by performing post-processing steps. The findings of the experimental study show that the proposed method achieved comparable results compared to cutting-edge methodologies, notably in terms of tumor core and improving core segmentation. Fig. 46 shows the visual segmentation results.

Ding et al. [137] suggested a new multi-path combined network with adaptive nature that uses ResNets "skip connection" to reserve and send additional low-level information accurately. A multi-path adaptive fusion dense block fuses low-level information with high-level semantic information. Compared to other current methodologies, the extensive experimental data leads to the conclusion that the proposed framework provides avant-garde results. Brain tumor segmentation results can be seen in Fig. 47.

U-Nets, which are multi-task driven, were presented by Zhang et al. [138]. To efficiently enhance difficulty-classified pixels for training, BU-Net is employed to reduce noise disturbance, which is optimized using an S-CE loss. The edge branch mechanism is employed in glioma borders to increase edge information. When comparing experimental results with cutting-edge 2D approaches, the presented model produces

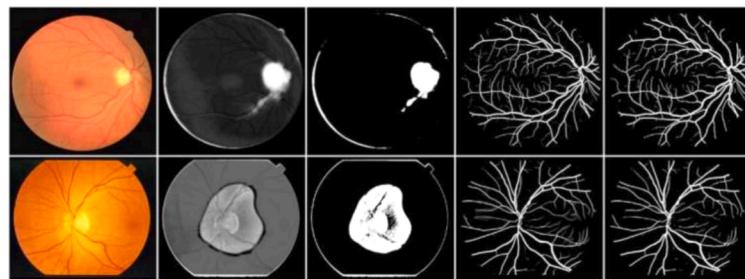


Fig 31. Segmentation map and segregated vessel of only GAN loss (column 2 to 3) and GAN loss along with BCE (column 4 to 5) [122].

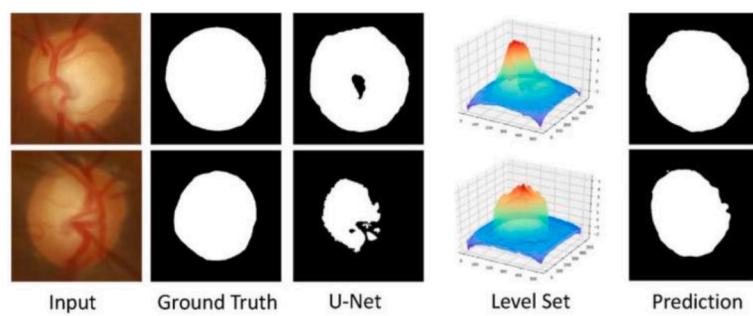


Fig 32. The example results for the DRISHTI-GS dataset. The first row is for the OD segmentation, and the second row is for the OC segmentation [123].

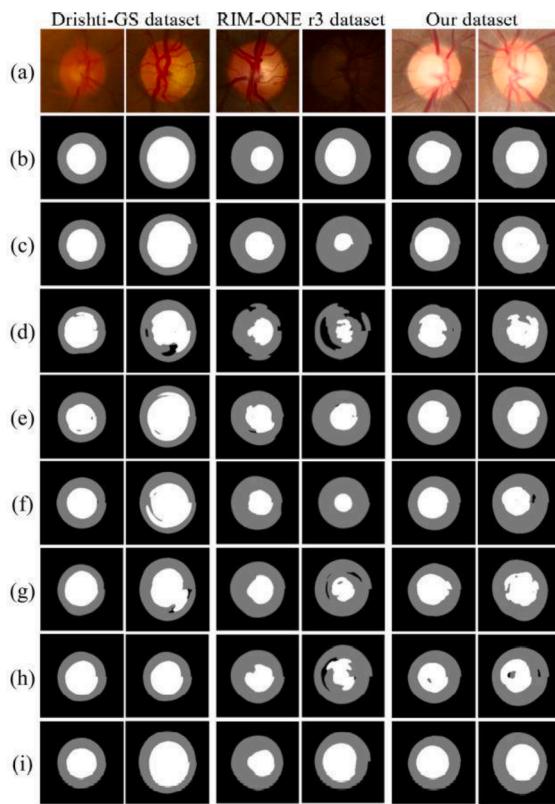


Fig 33. Drishti-GS dataset, RIM-ONE r3 dataset, and their datasets are tested on Network segregation: (a) fundus pictures color, (b) Ground reality, (c) FCNs, (d) SegNet, (e) PSPNet, (f) DeepLab V3+, (g) U-Net, (h) M-Net and (i) the suggested method [124].

a Dice score of 0.82 for entire tumor regions and 0.9950 for the BRTAS2015 and BrainWeb datasets, respectively. Fig. 48 shows the visual comparison of brain tissue segmentation results.

Sun et al. [139] proposed a 3D contemporary convolutional network. Multi-pathway architecture is used to extract characteristics from MRI pictures successfully. A different receptive field of the feature can be extracted using a 3D model, as shown in Fig. 49. We found that our suggested and the one-pathway models produced equivalent findings after a thorough evaluation with a collection of publically available datasets.

Zhou et al. [140] published an unorthodox brain tumor segregation network. The three sub-network model given here uses the existing modalities to create a 3D feature-refined picture that depicts the missing part. The published technique is thoroughly tested on the BraTS 2018 dataset, concluding that this model achieves significant results and outperforms all other methods. Fig. 50 shows the segregation result of the proposed method.

Zhou et al. [141] introduced (ERV-Net), which is a neural network and 3D in nature, that effectively segments brain tumors while

consuming less GPU memory and requiring less computational complexity. In this model, 3D ShuffleNetV2 is first used as an efficient encoder. Then a decoder (Res-decoder) with residual blocks to abstain deterioration, followed by a loss function fusion consisting of Dice loss and Cross-entropy loss. Finally, post-processing refines the ERV-Net segmentation result, as shown in Fig. 51. On (BRATS 2018), a multi-modal brain lesion segregation dataset, the experimental results yielded the best results, with Dice of 81.8 percent, 91.21 percent, and 86.62 percent, respectively, and ERV-Net also produced comparatively substantial results compared to the avant-garde methods.

Shehab et al. [142] published a model that classifies brain tumors automatically. When compared to the DNN equivalent, ResNets performs better. ResNets has a shortcut skip connection that works in conjunction with the layers of convolutional neural networks. The findings show that the suggested model has superior accuracy in the complete, core, and enhancing regions, with 83 percent, 90 percent, and 85 percent, respectively. Furthermore, it is three times faster than the DNN model in computing. Yogananda et al. in [167] developed six disparity autoencoders to segment different types of brain tumors using MRI samples. The authors apply each pair of two DAE for segmenting whole-tumor, tumor-core, and enhancing tumors with stable dice scores of 0.89, 0.82, and 0.81 for WT, TC, and ET on the validation dataset from the RSNA-ANSR-MICCAI benchmark.

6.5. Semantic application to pulmonary nodules

Lung cancer is the most hazardous and common cancer-related fatality. Lung cancer has a five-year survival rate of 10 to 16 percent. The five-year survival rate rises to 70% when cancer is detected early. CT

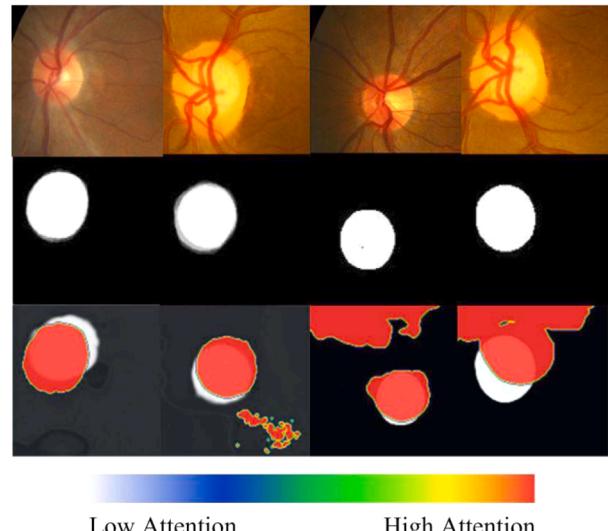


Fig 35. Top row: fundus pictures of the eye. Center row: ground-reality segregation results. Last row: visualized heatmaps [126].

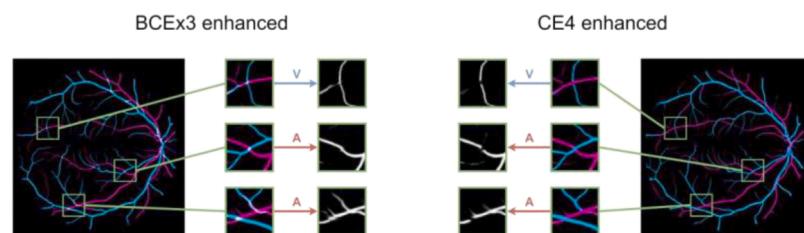


Fig 34. Examples of arteries, veins, and vessels probability maps (RGB compositions) generated by the models trained using the MS and the traditional approaches with preprocessed images [125].

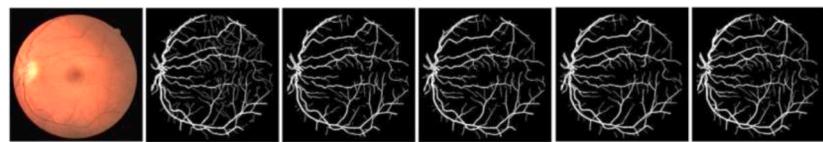


Fig 36. The illustrations of the improvements of different modules. From left to right: (1). Original image, (2). Label, (3). Result of U-net, (4). Result of Unet + CDM, (5). Result of Unet + CDM + Contextual module, (6). Result of CIEU-net. The result gets better in detail with the increase of modules [127].

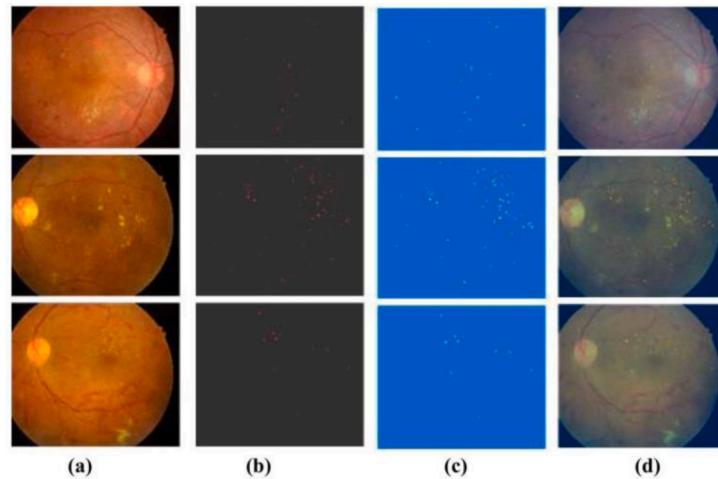


Fig 37. Results of microaneurysm detection on e-ophtha as training set and IDRiD as validation and test set: (a) Original image from the IDRiD test set; (b) Microaneurysm label; (c) Segmented image; (d) Segmented image superimposed on the original image [128].

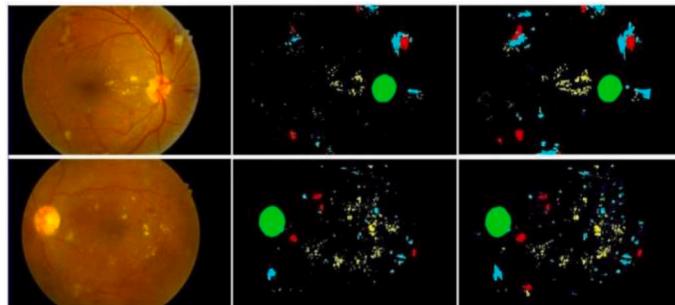


Fig 38. The first column shows the retina fundus images, the second the predicted segmented masks, and the third column shows the ground truth segmented masks. Green: Optic Disk, Red: Soft Exudates, Blue: Microaneurysms, Cyan: Hemorrhages, Yellow: Hard Exudates [129].

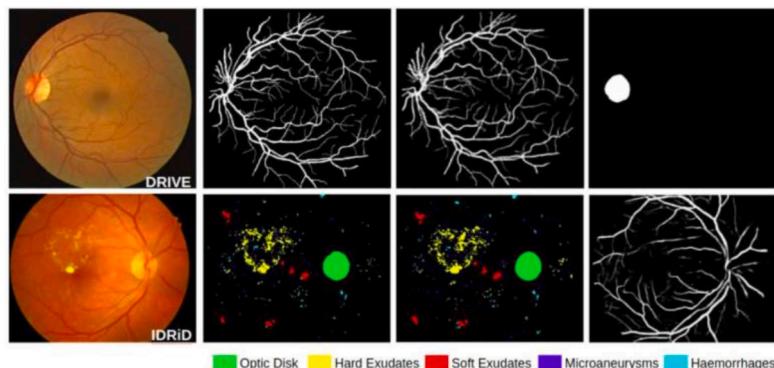


Fig 39. Qualitative results: First column shows the retinal image, the second column shows the manual annotation used as ground truth, the third column shows the results obtained by the proposed method for the annotated classes in the original dataset, and the fourth column shows a prediction of class not annotated but present in the image [130].

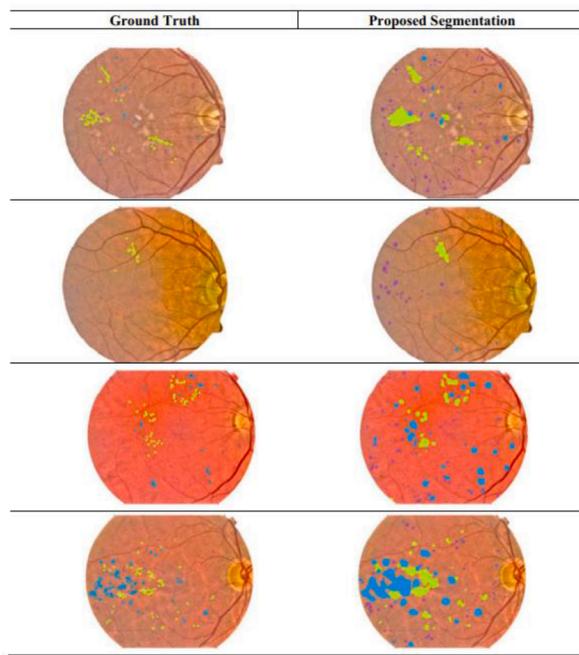


Fig 40. An example of DR lesions segmentation outputs. The ground truth annotated samples are on the left. Our ADL-CNN results are on the right. Light green marks represent EX's, light blue marks denote HEM's and purple marks represent MA's [163]

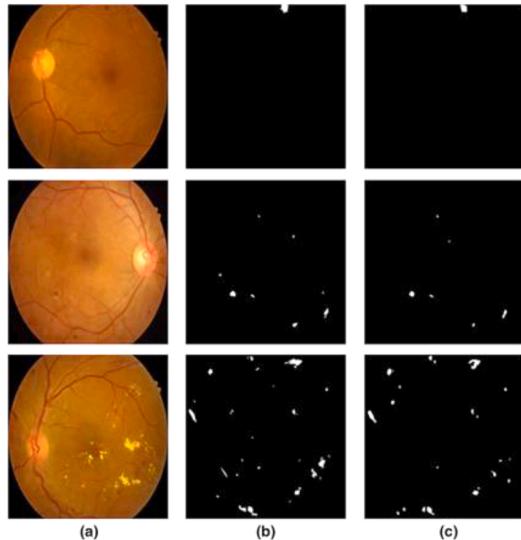


Fig 41. Results of hemorrhages segmentation: an original image, b hemorrhages label, and c segmented image [165]

scans are now the most effective and sensitive screening method available. Computer-aided diagnosis (CAD) systems provide detailed laceration information to help radiologists discover and diagnose lacerations. Many researchers have made contributions to this field.

Huang et al. [143] suggested a method for segmenting accurate lung nodule locations from raw CT scans that is both automatic and efficient. Detection of Nodule with quicker regional-CNN (R-CNN), lowering and merging falsely positive (FP) findings with Convolutional Neural Networks. Furthermore, segregation of nodule with improvised Fully CNN is the proposed model's four key sub-portions (FCN). Human intervention is not required in this model. The nodule detection gave the accuracy score of 91.4 percent and 94.6 percent, with 1 to 4 false positives

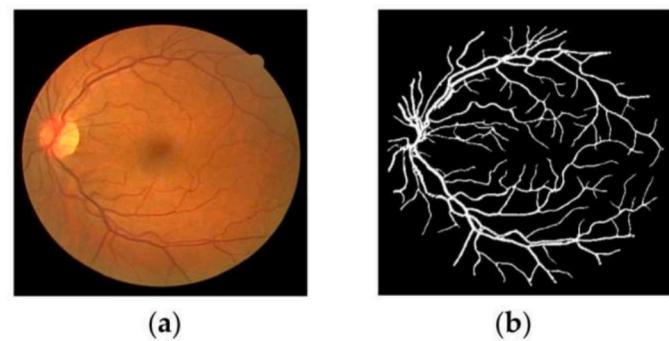


Fig 42. Vessel segregation results on both (a) original picture and (b) anticipated mask by Vess-Net on sample images [131].

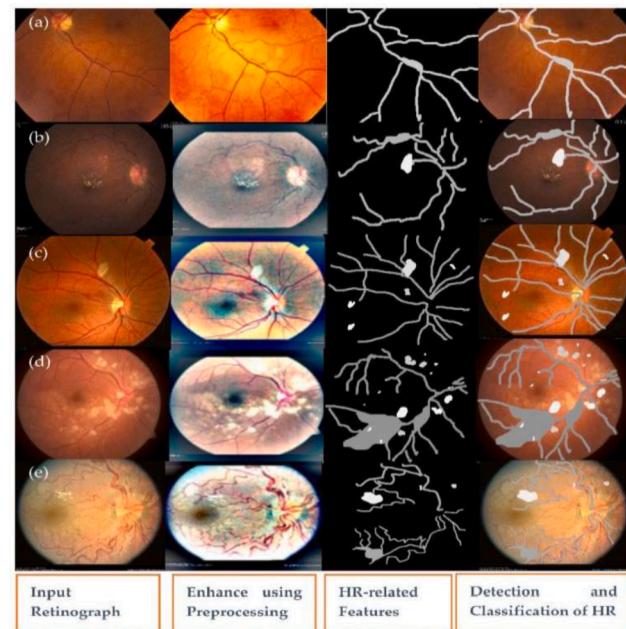


Fig 43. A visual example of the proposed semantic and instance segmentation of HR lesions where there is (a) no sign of abnormality, (b) mild HR, (c) moderate HR, (d) severe HR, and (e) malignant HR [132].

per scan on average, which is an excellent result. The visual results of the proposed segmenting method are given in Fig. 52.

Zhu et al. [144] proposed a completely automatic lung CT diagnosis method based on deep learning, divided into nodule identification and categorization. The proposed model uses two deep 3D convolutional networks, which run on a 3D dual-path framework. A 3D Faster R-CNN and a U-netlike encoder-decoder structure are used to detect nodules. To categorise nodules into malignant or benign, a deep 3D dual-path architecture is employed to extract classification features, and a gradient boosting machine is trained. Extensive results show that the model outperforms the DeepLung method, attaining 81.41 percent diagnosis accuracy and 99 percent average performance compared to four highly experienced radiologists. Fig. 53 shows the detection results.

Bana et al. [145] published a model split into two parts: segmentation and classification. The recently published DeepLab model was compared to this model for semantic segregation. The findings show that the accuracy of nodule identification outperforms the primitive U-Net model and its most recent variations. Fig. 54 presents a visual example of the segmentation results of the proposed model.

Wu et al. [146] introduced a CNN that functions as a multi-task learning system for detecting maliciousness and pulmonary tumor

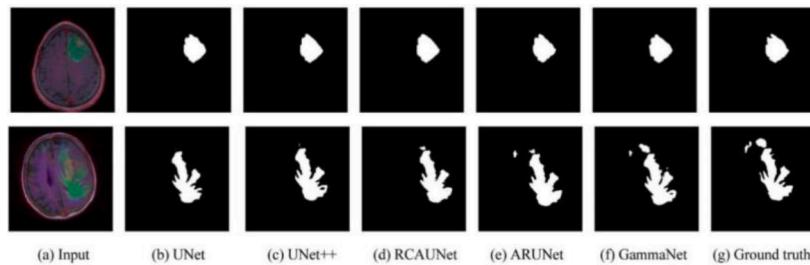


Fig 44. The prediction results of UNet, UNet++, RCAUNet, ARUNet, and GammaNet [133].

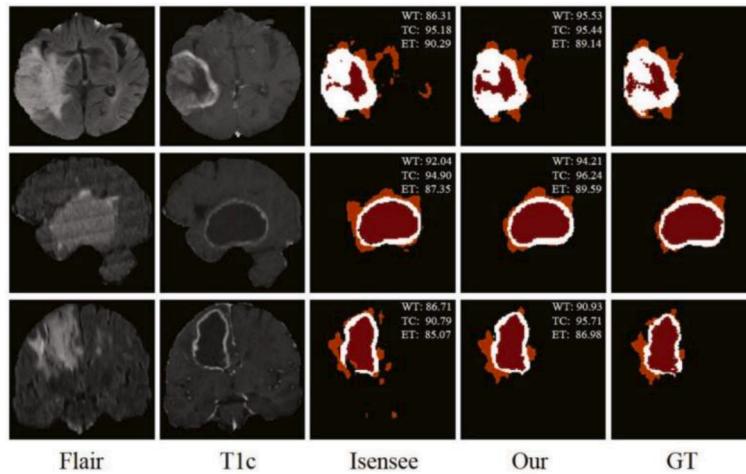


Fig 45. Qualitative experiment results between proposed method (8 initial filters) and Isensee et al. (2017) (16 initial filters) on several examples. We denote the Dice Score on each result. Net&Ncr is shown in red, edema in orange and enhancing tumor in white [135].

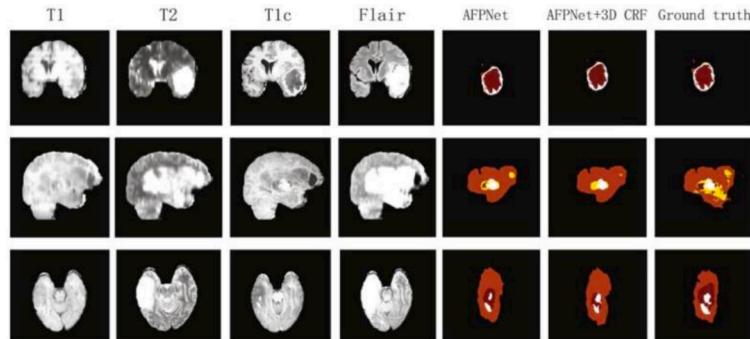


Fig 46. Visual segmentation results using the trained models on the BRATS 2013 (first row), 2015 datasets (second row), and 2018 (third row). The color codes are as follows: edema(orange), necrosis(maroon), enhancing core(white), non-enhancing core(yellow) [136].

segregation attributes (PN-SAMP), as shown in Fig. 55. It can accurately diagnose the malignancy of a lung tumor. Furthermore, the combination of tumor segregation, characteristics, and maliciousness identification aids in improving task proficiency. The published model was tested extensively using the publicly available dataset, one of the largest for malignancy, named LIDC-IDRI. Compared to cutting-edge approaches, the results accurately interpret that the proposed model achieved considerable improvements in lung tumor categorization and spectacular lung tumor segregation performance and learning of attributes.

Meraj et al. [147] proposed a system for predicting lung cancer and distinguishing between benign and malignant tumors. After pre-processing to eliminate noise, the proposed model was extensively tested on a publicly available dataset. In addition to the procedure, four key characteristics are chosen based on the classification. According to a comprehensive investigation, the provided method surpassed other

methods with an accuracy score of 99.23% using the logit boost classifier, as represented in table 7.

Qin et al. [148] published a model mostly made up of two sections. The first phase balances the dataset by increasing the variety of samples, while the conditional generative adversarial network (cGAN) is used to create synthetic CT images. To refine synthesized samples, reconstruction error loss is added to cGAN. The tumor segregation network is trained on the dataset in the second half. Finally, a three-dimensional (3D) CNN model is created to inform the model about tumor texture patterns and border features, assisting in high-level feature learning for segregation. The outcomes of this model have been empirically validated, yielding results comparable to existing methodologies. Table 7 illustrates a summary of the developed deep semantic-based arts to detect diseases. Fig. 56 shows the qualitative segmented results using the validation samples.

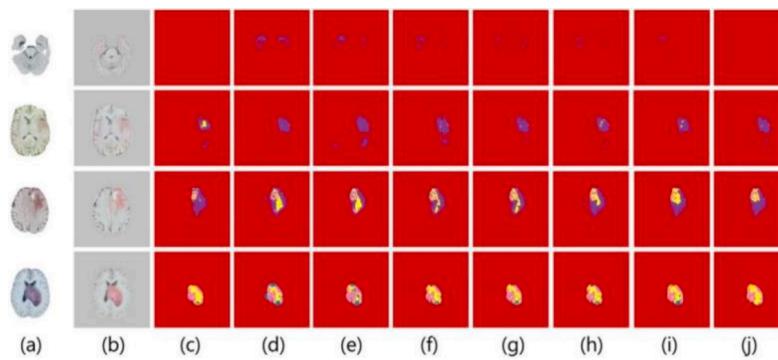


Fig 47. Samples of Multimodal Brain Tumor Segmentation on BRATS 2015. From left to right are: (a) Input, (b) pre-processed image, (c) ground truth, (d) Fcn16s, (e) Fcn8s, (f) U-Net, (g) Resnet, (h) Refinnet, (i) FC-Densenet, (j) proposed method [137].

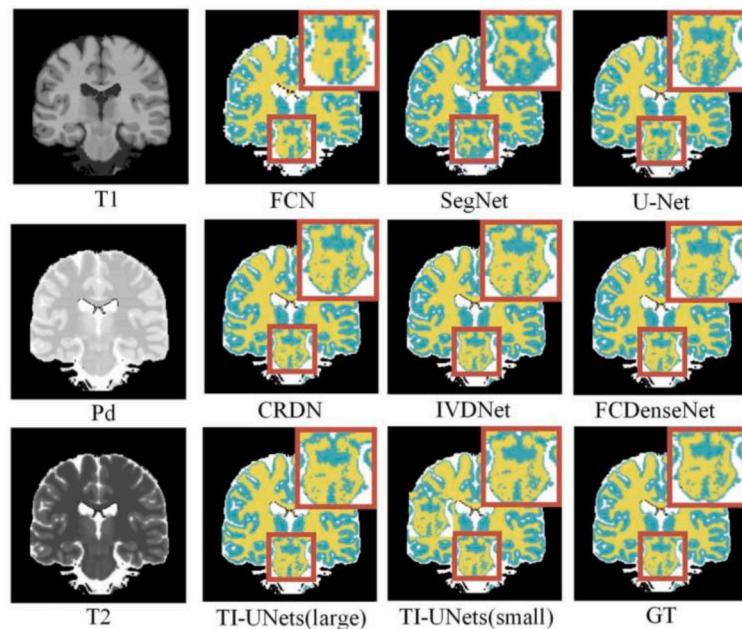


Fig 48. The visual comparisons of different brain tissue segmentation results from the BrainWeb dataset [138].

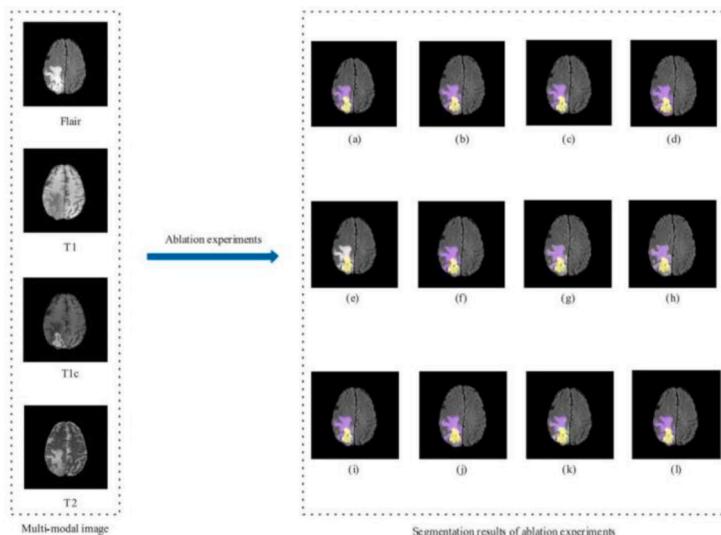


Fig 49. The result of the suggested framework is displayed in the first row. (a) Segregation results without using convolution being dilated. (b) Segregation result with convolution along with dilation and kernel size = 3. (c) Segregation result with convolution along with dilation = 4 and kernel size = 3. (d) Segregation result with convolution along with dilation = 2 and kernel size = 5. (e) Segregation results without pooling. (f) Segregation results in average pooling. (g) Segregation results with convolution along with kernel size = 5. (h) Segregation results with convolution along with kernel size = 7. (i) Segregation results with ReLU activation function. (j) Segregation result with Leaky ReLU. (k) Segregation results without pre-processing. (l) Segregation result in one pathway [139].

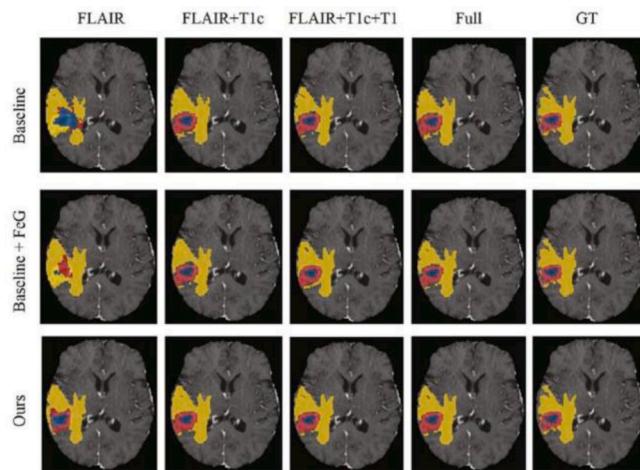


Fig 50. Segregation results of proposed framework where the First row depicts the four MR modalities, FLAIR, T1c, T1, and T2. The segregation results of different models been shown in the bottom three rows. The last column displays the ground truth segregation, while the first four columns show a situation of different missing modality. Blue shows the Net&ncr, yellow depicts the edema, and red represents the enhancing tumor [140].

The automatic segmentation of brain tumors greatly aids accurate diagnosis and treatment planning in magnetic resonance imaging (MRI). Therefore, the study in [162] presented the Attention Res-UNet with Guided Decoder (ARU-GD) as a new deep learning generator architecture for segmenting brain tumors. The architecture incorporates attention gating and a novel guided decoder. These modifications to the

Res-UNet base network enhanced learning by producing superior feature maps at the decoder and allowing only activations from relevant regions at the encoder side. Together, these factors increased segmentation performance. On unseen test data, the proposed ARU-GD successfully achieved Dice Scores of 0.911, 0.876, and 0.801 and mean IoU of 0.838, 0.781, and 0.668 on the entire tumor, tumor core, and enhancing tumor segmentations, respectively. Fig 57 represents the segmentation results. The current investigation has demonstrated that their model performs better than its baseline and existing models. However, the presented model's training is computationally intensive.

7. Discussion

This section discusses current architectures' strengths, limitations, and future directions for medical image-based semantic segmentation (Semseg), datasets, and optimization functions. Different deep-based semantic segmentation arts have been investigated and discussed to diagnose various potential diseases. After careful analysis, this survey has a unique contribution to the biomedical imaging segmentation domain because no such study can cover the broader sense of semantic-based segmentation models developed for the accurate detection of medical ailments. Fig. 59 presents a visual example of the actual distribution of deep semantic segmentation models utilized for detecting and diagnosing different ophthalmic diseases. According to some state-of-the-art methods, Encoder-decoder architecture is a promising method for semantic segmentation of medical images as it contains long and short skip connections. These deep neural networks with skip connections have comprehensively enhanced the performance for categorization and segmentation tasks because they reduce the training time and mitigate the risk of vanishing the gradient. Despite performing better, there are still issues like computation cost, higher memory usage cost,

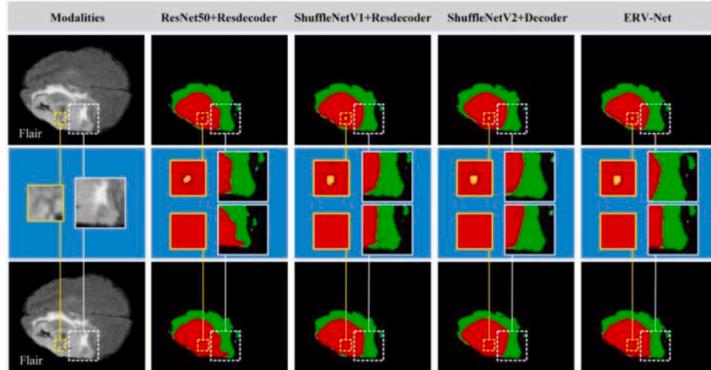


Fig 51. Flair's original picture is shown in the first column. Segregation results on case Brats18_TCIA13_646_1 of the BRATS 2018 validation dataset with and then without post-processing are visualized in the first and third row, respectively. While the second row displays the magnified images of the details with and without post-processing. Different network frameworks and architecture and their segregation results are displayed from left to right. Different color indicates different lesion, i.e., yellow, green, and red represent edema, enhancing tumor, edema, and necrotic and non-enhancing tumor core, respectively [141].

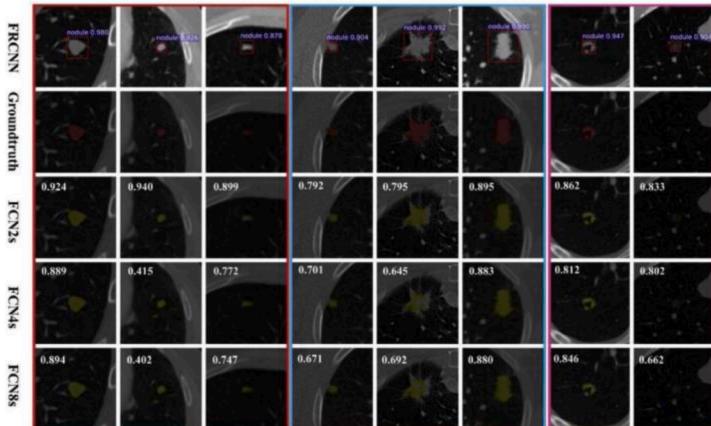


Fig 52. Visualization results of proposed nodule segregation system gave results with several anatomical particulars. The first three columns show three isolated nodules, the following 3 columns marked in blue represents one juxta-pleural (4th column) and remaining columns in blue denotes juxta-vascular (5th and 6th column) nodules, and the last two columns highlighted in the light purple show one subsolid nodule with center excavation and one ground glass opacity (GGO) nodule. Original nodule patches obtained after Faster R-CNN identification and reduction by FP are shown in the first row. While the annotations by the specialist doctors in radiology corresponding to the first row are shown in the second row. The manually segregated parts are emphasized by masking them in red. The last three rows display the nodule segregation results by FCN2s, FCN4s, and FCN8s, respectively. The dice coefficients for each segregation parallel to ground reality markings are shown in white decimals [143].

Ground truth center slice								
Detection center slice								
Detection prob. (%)	99.46	99.93	100.00	99.86	99.95	99.90	99.61	99.52

Fig 53. Visualization of central slices for nodule ground truths and detection results. We randomly choose nodules (red circle boxes in the first row) from test fold 1. Detection results are shown in the blue circles of the second row. The center slice numbers are shown below the images. The last row shows detection probability. The DeepLung performs well for nodule detection [144].

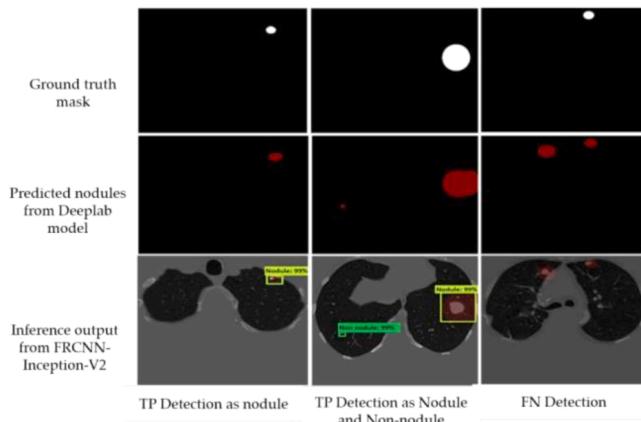


Fig 54. Examples of the inference produced by the Faster-RCNN Inception-V2. The first column shows TP detection as a nodule with IOU = 99%. The second column shows TP detection as a nodule with IOU = 99% and an FP with IOU = 99%. Finally, the third column shows an FN where the model failed to detect the nodule lesion [145].

and transferring non-discriminative feature maps for better feature representation.

Due to the increase in medical image modalities and datasets for segmentation tasks, it is a demanding and tiresome task to summarize the result of all medical image segregation methods. But in this review article, we have tried to summarize the necessary imaging modalities used in cited papers, as shown in Fig. 4. We have also summarized some state-of-the-art methods used for medical image segmentation, as shown in Table 5. In Fig. 58, we have shown the number of images available in each dataset for a specific image modality. MRI, CT, and electron microscopy (EM) are difficult to label and exist in less amount in their benchmarks than the frequently available image modalities like ultrasound (US) and X-ray images, which are available in a large number of samples in their datasets.

Sequenced models can segment 3D medical image datasets such as videos and volumetric data. Volumetric data are processed more efficiently using 3D convolutions than by processing the volume using 2D models at each level. In addition, sequence models capture an object's 3D volume geometrical information easily compared to slice-by-slice 2D sequenced models where some information may not be adequately captured. There could be a future direction in which comparisons and analyses of volumetric convolutional methods with the sequenced

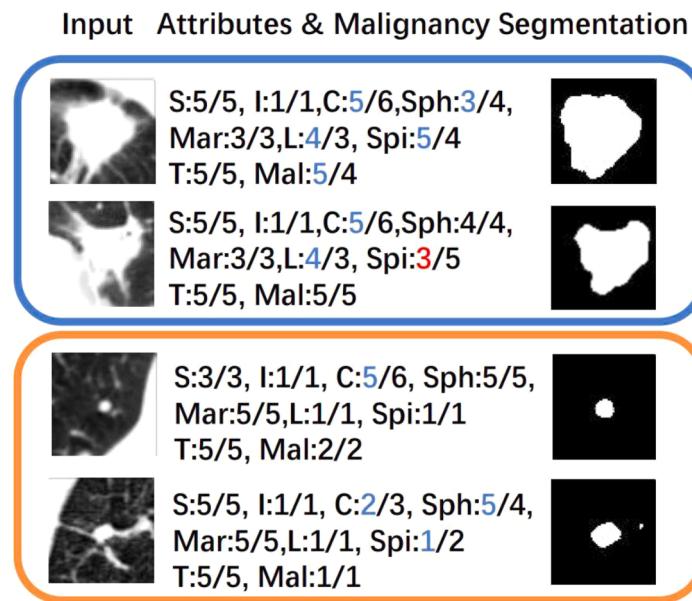


Fig 55. Pictures on the right show segregation results with malignancy rating identification, attributes shown by qualitative results. The top part depicts the malignant, while the bottom part in the figure shows the benign nodules. Identification of the suggested framework is denoted by A, and the ground reality is denoted by B. The in-range prediction of ground truth in range 1 is shown in blue color integer and red shows out of range prediction [146].

Table 7
Summary of developed deep semantic method [147].

Work	Classifier	Accuracy	Sensitivity	Specificity
Akram- 2015	SVM	96.6	96.7	96.3
El-Regailly-2017	Simple rule-based classifier	70.5	77.7	69.5
Hancock-2017	Nonlinear	88.0	84.6	NA
Jaffar-2018	Random Forest	98.9	98.4	98.7
Tran GS-2019	LdcNet-FL	97.2	96.0	97.3
Shaukat-2019	RGBPCANet	93.25	93.12	91.37
Framework- 2019[147]	Logi boost	99.23	96.88	100

model can be made.

Every method for semantic segregation is limited based on deep learning in medical image processing. As region-based segmentation performs better in the region growing-based techniques and performs better in the noisy image and is easy to compute, there are still some issues like seed points must be specified, costly, and the image may be under-segmented or over-segmented. Similarly, in region merging and splitting methods, the image can be split according to the demanded resolution, but it may produce blocky segments. Due to the scarcity of annotated medical-related datasets of images, the use of supervised learning methods in medical-related segmentation of images is limited. That is why weakly supervised, or unsupervised methods have been applied to semantic segmentation tasks. As authors in [58] proposed a sickly supervised method that decreases the cost of computation for training and achieved similar results to fully supervised segmentation of cardiac images. Similarly, many other authors used weakly supervised approaches for different image modalities [59–61]. These methods used automated algorithms for segmentation tasks with a little interaction of the domain experts with the systems to accurately identify the results produced by these methods [7]. Domain experts may be required when selecting ROI, which will be generalized to the whole image using weakly supervised algorithms.

Fully convolutional networks (FCN) are fully automatic segmentation techniques because they do not need domain experts to select ROI. But most of these methods involve supervised learning and require training data, e.g., deep neural networks, shape models, and random forests. These methods are applied to various problems in medical image semantic segmentation. But there are some issues with these methods in their use as a lack of profound theoretical basis for ANNs, problems in choosing the best architecture, and these are black-box problems.

In medical image segmentation, many researchers used cross entropy as a loss function. And then applied some distance or overlap function

suitable for medical image segmentation, because the purpose of using overlap function is to separate specific regions (such as polyp, tumor, and cancer) as a part of the semantic segmentation process. Authors in [62] claimed that if we use only an overlap or distance function in a deep neural network with a sigmoid function used in the final layer then the chances of the gradient vanishing problem may increase. As discussed in the optimization functions section, overlap loss functions are helpful in case of imbalanced data but with small foreground medical image segmentation. In the future, a function can be designed with a single loss that computes the distance based on cross entropy and overlaps simultaneously. This can be possible by reviewing the previously functional overlap-based and distance functions. Table 8 presents a brief overview of the literature on deep semantic-based arts for multiple medical imaging problems.

7.1. Current limitations and computational challenges

In this part of the paper, the current limitations and challenges faced during the training and application of deep learning techniques and models are thoroughly discussed and used to detect various ailments (as stated in outlined Section 6).

Researchers have applied different techniques in diagnosing skin lesion cancer, facing different sorts of limitations in their models and frameworks. An apparent limitation of the FCN is that it is heavier compared to others because of more parameters which lead to more time being taken for inference and training of networks. Another limitation SLSNet faced was the inaccuracy of lesion boundary segmentation because of the improper boundary between lacerated and normal skin. A Few limitations in applying the GAN model were low and poor contrast, inappropriate boundaries, and unwanted artificial noise, which makes it even more challenging to differentiate between the lacerated region and the normal portion. The poor and low contrast also affected the segmentation performance of the AFLN-DGCL-FUSION model, ultimately resulting in compromised accuracy scores. The SC-FCN-BLSTM approach utilized in research suffers from the limitation of GPU memory, for which images had to be cropped to compensate for the GPU memory issue. One limitation of the DFS model was the high time complexity in its computation due to the involvement of different models for its segmentation. Whereas the poor quality of training data has been a significant reason to restrain the performance of other models. An apparent setback in the semantic classification for breast tumor segmentation was the lack of automation because the images must be cropped, which restrains the model from fully automatic working. One concern while using the U-Net MALF model was fuzzy tumor boundaries which led to the cause of poor detection and

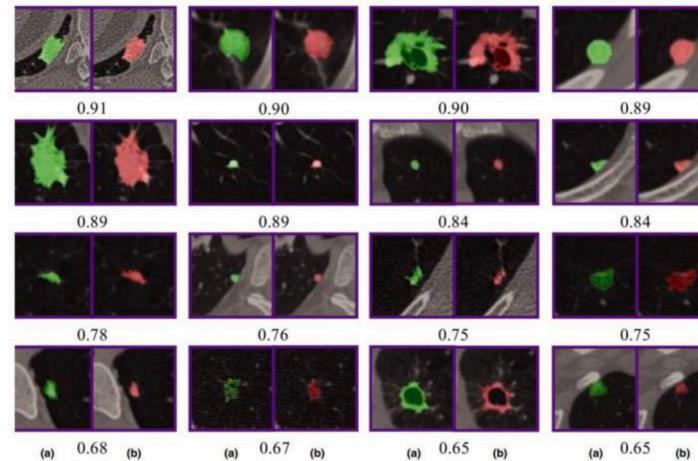


Fig 56. Qualitative segmentation results of validation samples. (a) Groundtruth labels are green; (b) predicted nodules are red. The score beneath each pair is the Dice coefficient of the result. The central slice of each volume of the interest cube is displayed for simplicity [148].

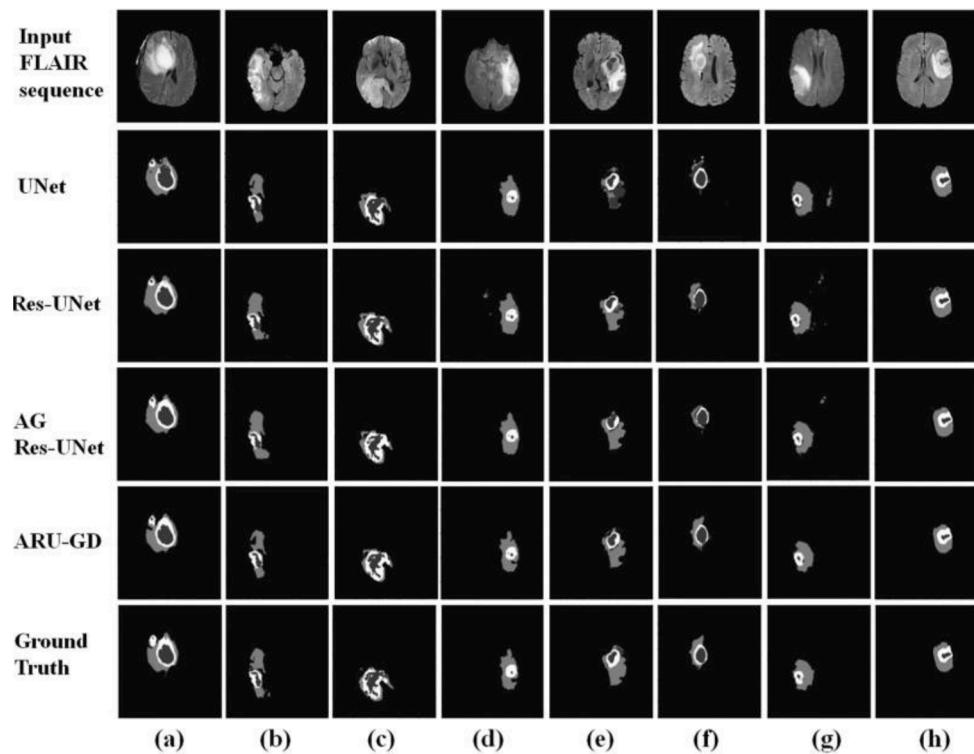


Figure 57. Predictions of the proposed model and the baseline models on randomly chosen eight unseen images from the test set; (row-wise) 1: flair images, 2: U-Net, 3: Res-UNet, 4: Res-UNet with AG, 5:Attention Res-UNet with Guided Decoder, 6: ground truth. White color in the segmented image corresponds to enhancing tumor, dark grey corresponds to necrosis, and light grey corresponds to edema.

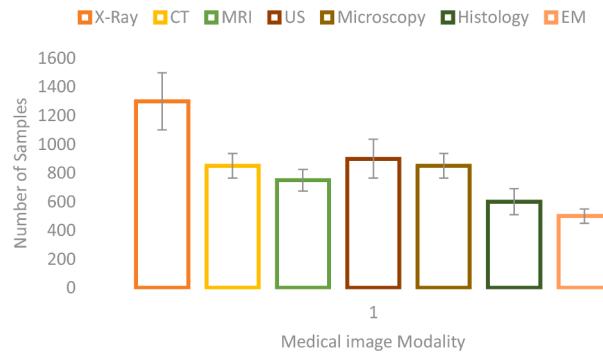


Fig 58. Average Number of Samples for different image modalities

segmentation of tumors. The drawback of semantic pixel-based segmentation naturally includes the problem of OD and OC being distinguished because each pixel is labeled separately and independently based on its category; all the pixels are labeled independently based on the category of that pixel. As in previous models, poor quality and contrast of fundus images had also been the point of concern in the W-shaped backbone network to give an effective and accurate location of the OD center. Regarding the setbacks in the TAU model, it could be argued that the imbalance markers of small objects put the model performance in a significant challenge. Secondly, it is unable to adapt to third-party datasets.

The deep learning-based segmentation model is widely accepted, but it suffers from the limitation of being computationally heavy, which leads to the fact that it takes more time in parallel to the U-Net baseline framework, and it also has the setback of not being robust in terms of unhealthy images in the dataset. The modified UNet approach utilized in the diabetic retinopathy segmentation suffers from an apparent drawback: it is only validated and tested on publicly available datasets, which

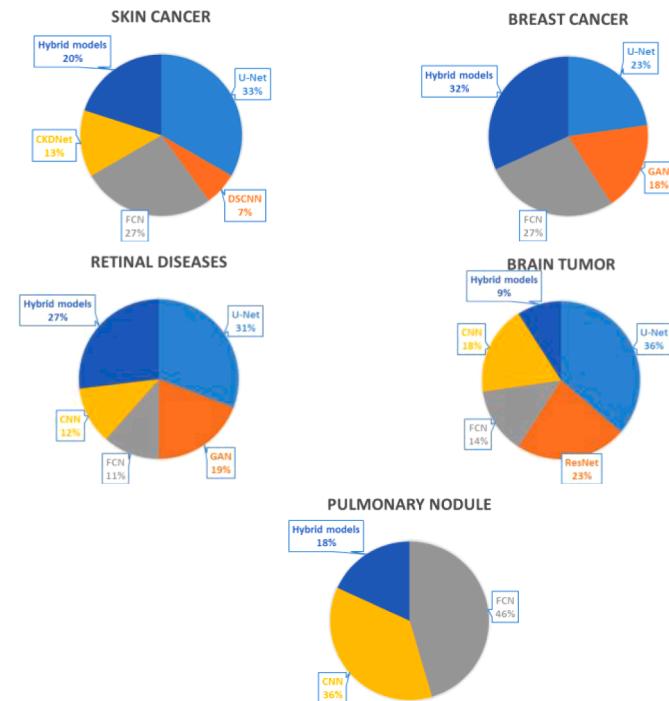


Fig 59. Conveyance of different deep-based semantic segmentation arts for detecting various diseases.

limits the performance of the model as rigorous as it should be extended to the clinically acquired database. Not much research is being done in this domain, but the HYPER-RETINO system has an apparent limitation the difficulty in extracting relevant HR-related lacerated features from

the retinopathy dataset. Secondly, the lack of availability of labeled datasets by clinical experts to train and test the model is the central point of concern which could be one of the main reasons this domain has not been under the focus of many researchers. The multi-encoder-based U-Net model limitations are apparent in many ways, and the main setback is that it only validates multi-MR modalities, along with the room for improvement still needed in the correlation constraint block. Another limitation is that the model has only been evaluated on a publicly available brain tumor dataset. The ERV-Net framework presents some limitations of being computationally complex and the number of parameters, which could be improved further to improve the performance results. 3D R-CNN utilized suffers from the same limitation as previously discussed: the GPU memory constraint, which limits the researcher to split the CT image into small patches to recover the resource and time limitation for training, also leading to the use of 5-fold validation. CNN-based network, although widely accepted, it suffers from some limitations, including the inability of the model to consider other structures such as bronchi and bones. Another drawback of this model is the uneven distribution of the training dataset, and lastly, the apparent setback of the method is to find out the optimal form of introducing random noises into cGAN.

7.2. Future directions

This subsection presents a detailed view, highlighting the future perspectives and suggestions to further improve the scientific deep learning-based arts for various medical diseases.

7.2.1. Future suggestions to optimize skin cancer arts

One significant future research suggested in [98] should consider the potential effect of ROIs to extract valuable features for different types of laceration classification coupled with the DSNet and loss function to verify the versatility and generality of medical contexts. It will be essential to investigate to improve the ability to extract significant features in the segmentation model to make better use of medical images [99]. The assumption of diagnosing skin cancer using the proposed framework [100] may be addressed in the future as it precisely removes specific objects, e.g., watermarks and images. Future studies could fruitfully explore the time complexity issue by reducing it with an end-to-end platform for clinical research. The CKDNet cascade knowledge diffusion framework could be generalized to be implemented on image segmentation and classification datasets [101]. Future research could be completed on enhancing the effectiveness of the proffered framework [102] by refining the optimization of laceration boundary details and the generalizability of different networks on dermoscopic image datasets. What is desirable for the future is to increase the number of lesion images which may help in data augmentation techniques to increase efficiency [54]. At the same time, the computation time for the training of data can also be curtailed by increasing the GPU memory.

Lastly, the inference performance can also be improved by increasing the epoch numbers for the training phase [103]. Future research will be essential to include the SLSNet model implementation in the form of a mobile application to segment and classify skin lacerations in images captured by handheld phone cameras with low resolution. Contrary to that, parameters like efficiency, performance, deployment, and upgradeability may be assessed using different embedded systems like FPGAs, NVIDIA Jetson GPUs, and Mobile SoCs [104]. In the future, working on blurred boundaries and low contrast images using transfer learning to make skin laceration recognition efficient and make it more sensitive to the dataset of blurred boundaries and low contrast images [105]. Future studies could fruitfully explore the issue of low contrast and blurred images by working on the GAN strategy and qualification tests to improve the framework's robustness [106]. Future research is needed to delimitate the use of supervised learning and unsupervised models for skin laceration segmentation to efficiently increase the dice coefficient and accuracy score [107].

7.2.2. Future suggestions to optimize breast cancer arts

In the future, investigating and working on an orthogonality view to locate the nipple region automatically could be of great use to increase efficiency. As in the proposed model [108], the nipple region cannot be differentiated from the tumor. Future research on applying the presented framework to multivendor and multicenter datasets could make the proposed framework more robust. Secondly, working on a three-class one-stage nnU-Net for FGT and breast segmentation could get us desired improved results. Thirdly, rigorous work can be done to reduce the inference time by focusing on model compression [109]. It will be important that future work should improve cancer detection in a more targeted manner. In addition, the MDE technique is proposed in this paper as a complex framework that could be applied using high-performance computing and distributed computing in particular segmentation processes.

Moreover, this model is a high performer and can also be applied in other domains, such as engineering optimization, feature recognition, and medical diagnostics [110]. The future result could examine the simplicity of the proposed algorithm and improve the segmentation process concerning the failed cases. And be tested and validated on different publicly available and clinical datasets to make its results more generalized [111].

Desirable future work can be the automation of the annotation of ultrasonic segmentation data [112]. Moreover, more novel and advanced models could be investigated to improve the segmentation. Future research could fruitfully explore feature information's multi-dimensionality to increase breast laceration identification's robustness. Secondly, because of its generic framework, many other medical imaging modalities could be tested and trained on the presented model [113]. Future research should apply advanced deep learning algorithms to segregate breast ultrasound images correctly. However, it could be greatly used in clinical medicine [114]. Future research should certainly further improve tumor supervised learning on the classification and reclassification of superpixels. Further research might help make a closed loop to provide the CAD output as input to the segmentation and the result as input to CAD for tumor classification [115]. The proposed approach limits the performance evaluation of different tissue types. Further extending the scope for the assessment to the set of densely annotated data could be the interest of future research along with the generation of such manual annotation [116]. Further research is needed to study custom loss function being incorporated with a deep learning model for encoding prior results and contributing with medical peers to get annotated images with breast tissue layer and integrate it into a single segmentation model [117].

7.2.3. Future suggestions to optimize retinopathy diagnosis arts

Future research may further be needed to investigate the performance of an algorithm on a diverse dataset. Furthermore, this model can also be applied to other retinal diseases, which could be a good addition to the research domain [118]. The authors in [119] recommended that future research be devoted to detecting other diseases related to optic discs, such as disc anomalies, tilted optic disc, and anterior ischemic optic disc neuropathy. Additionally, the proposed framework could be extended to other biomedical image analysis problems. It will be important that future research investigates to find the perfect balance between efficiency and accuracy of the presented framework, ensuring the effective segmentation of optic cup and optic disc. Moreover, future studies could also focus on automatic glaucoma diagnosis [120].

Future work should examine restraining the FP pixels production supported by using prior knowledge of vessel structure and extracted information from fundus images. Exploration of further constraints may constitute the objective of future studies, such as convexity. Furthermore, the distances between two level sets can be considered in the future, such as the manifold distance [121]. Future research should fully utilize the proposed framework to grasp the information of its semantics of fundus pictures and combine the segmentation framework with the

Table 8

A comprehensive summary of deep-based semantic segmentation approaches for diagnosis of different diseases.

Cited Paper	Modality	Description	Database	Performance metrics & Results
Hasan et al. [98]	Dermoscopic Images	Projected learned discriminated features on pixel-level using depth-wise separable convolution instead of standard convolution to segment skin lesions.	ISIC 2-17	Dice Score 0.927 and IoU Score 0.865
Shan et al. [99]	Dermoscopic Images	A topology named fully convolution and dual path (FC-DPN) was implemented to perform skin lesion segmentation.	ISBI 2017 and PH2	DICE Score and Jaccard Index of 88.13% and 90.26% on ISBI 2017 and 80.02% and 83.51%, respectively
Li et al. [100]	Dermoscopic Images	Proposed a specialized deep learning framework to remove digital hair with high accuracy.	ISIC 2018	Accuracy 92.59% and Dice Score 86.51%
Jin et al. [101]	Dermoscopic Images	The author proposed a CKDNet architecture combining cutting-edge skin laceration diagnosis and segregation.	ISIC 2017 and 2018	Accuracy 94.6% and JA Score 80.0%
Gu et al. [102]	Dermoscopic Images	The author implemented a technique that combines an advanced network structure DE-Net with a unique loss function to look deeply into the skin laceration boundaries called EIGM.	ISIC 2017	Dice Score 0.8662 and Jaccard Coefficient 0.7887
Kaymak et al. [103]	Dermoscopic Images	The researcher suggested an extensive investigation for skin laceration segregation using all well-known FCN architectures. The FCN-AlexNet performed best.	ISIC 2017	Accuracy 94.811%
Sarker et al. [104]	Dermoscopic Images	The author proposed SLSNet, a GAN-based model for skin laceration segregation that is effective and easy.	ISBI 2017 and ISIC 2018	Sensitivity, Accuracy, Specificity, Jaccard index, and Dice coefficient of 87.81%, 97.61%, 99.92%, 81.98%, and 90.63% on ISIB 2017 and 87.81%, 97.61%, 99.92%, 81.98%, and 90.63% on the ISIC 2018
Lei et al. [105]	Dermoscopic Images	The author proposed an unconventional GAN model for skin laceration segregation made up of two modules called the DAGAN model.	ISIC 2017	Accuracy 93.5%
Tang et al. [106]	Dermoscopic Images	Demonstrated a unique DL-based model for skin laceration segmentation which is a combination of AFLN DGCL and AFLN DGCL framework.	ISIC 2017 and ISIC 2018	Accuracy Score 0.963 and 0.966
Arora et al. [107]	Dermoscopic Images	Introduced an AG-based customized U-Net model for automated skin laceration segregation supported by Tversky Loss (TL) as the output loss function.	ISIC 2018	Accuracy and Jaccard Score of 0.95 and 0.83
Pan et al. [108]	Ultrasound Images	The author proposed a customized SC-attention module to get finer-grained geographical information along with the SCBLSTM module on top.	ABUS image dataset was acquired from Peking University People's Hospital	DSC 0.8178 and Accuracy 0.8292
Huo et al. [109]	Ultrasound Images	The author proposed a framework for FGT and breast segmentation on the principle of deep-learning through nnU-Net.	Own dataset	DSC 0.734 and 0.823
Liu et al. [110]	Ultrasound Images	The author introduced a Modified Differential Evolution (MDE), which is an upgraded form of the Slime Mold.	Own dataset	
Pawar et al. [111]	Ultrasound Images	The researcher proposed the technique using heuristic and non-heuristic approaches for the analysis of pectoral muscle.	DDSM	Accuracy of 86.16% for heuristic and 82.23% for non-heuristic approach
Zou et al. [112]	Ultrasound Images	The author proposed a segregation model that can detect and reduce noise. The network is monitored with the probability distribution of samples to differentiate noises.	BUSI	Precision 87.25 and Recall 88.6%
Yan et al. [113]	Ultrasound Images	The researcher presented a two-stage model that is based on coarse-scale mass detection with a multi-stage detection strategy.	DDSM-CBIS	Dice Score 80.44%
Ilesanmi et al. [114]	Ultrasound Images	The author proposed a model using CLAHE to remove noise, followed by semantic segregation with the Ve block and convolution method to segregate tumor cells.	Dataset was obtained from the Thammasat University Hospital and Baheya Hospital for Early Detection & Treatment of Women's Cancer, Cairo, Egypt	HD, JM and DM of 8.40, 78.87 and 89.73, respectively
Huang et al. [115]	Ultrasound Images	The author presented an advanced method with three types of features to extract the semantics of superpixels, which uses the BoW model to label the superpixels as tumors.	Own dataset	F1-Score 90.20% on Benign and 89.53% on Malignant tumor dataset
Rijthoven et al. [116]	Ultrasound Images	The author presented HookNet, a model that measures high-dimensional resolution tissue segregation.	Own dataset	F1-Score 0.91
Tong et al. [117]	Ultrasound Images	The author proposed a solution to improve brightness, contrast, and low-quality in ultrasound images of breast tumor segregation by integrating four attention loss functions into the cross-entropy loss feature.	data from Jiangsu Traditional Chinese Medicine Hospital	F1-Score and Accuracy of 0.873 and 0.959

(continued on next page)

Table 8 (continued)

Cited Paper	Modality	Description	Database	Performance metrics & Results
Imtiaz et al. [118]	Fundus Images	A deep learning-based algorithm to semantically segment OD and OC. Next, feature vectors are applied to a softmax classifier that categorizes based on pixels.	Drishti and Rim-one	Accuracy and Dice Score of 99.03% and 85.94%
Pachade et al. [119]	Fundus Images	The NENet patch-based adversarial model was proposed to study OD and OC segregation.	REFUGE, Drishti and Rim-One-r3	
Bian et al. [120]	Fundus Images	The author proposed a model which segments optic disc (OD) and optic cup (OC) using the cGAN structure, and the overfitting issue is mitigated to lower learning rates.	REFUGE	Accuracy and Segregation values 0.8887 and 0.9367
Tulsani et al. [121]	Fundus Images	The author proposed a custom model based on the UNET++ custom loss function and built to conduct segregation experiments.	RIM-ONE, DRIONS-DB, and ORIGA	IOU of OD and OC 0.9477 and 0.9321 and DC of OD and OC 0.97 and 0.95, respectively
Zhou et al. [122]	Fundus Images	The author proposed a model which uses SEGAN and MSFRB to refine the process of extracting information.	CHASEDB1 and HRF	Accuracy rates 0.9083 and 0.8211
Yin et al. [123]	Fundus Images	Author tailored a level set loss through CNN output as a level set, and you can add smooth boundary and region consistency and other constraints for the segregation.	DRISHTI-GS dataset	Dice Score 96.23
Yuan et al. [124]	Fundus Images	The author proffered deep learning with a W-shaped CNN structure on a multi-scale for segregating OC and OD collectively as a multi-label single-stage process.	DRISHTI-GS and RIM-ONE r3	F1 Score of OD and OC is 0.9644 and 0.9003
Morano et al. [125]	Fundus Images	The author published a modern approach for the simultaneous segregation and classification of the retinal arteries and veins (SSCAV) using FCNNs.	RITE dataset	Accuracy and Specificity of Artery and Vein detection is 89.24%, 90.89% and 96.16%, 98.65% respectively
Zhang et al. [126]	Fundus Images	The researcher presented the TAU model to address domain adaptation issues in OD and OC segregation tasks.	REFUGE, Drishti and Rim-One-r3	Dice Score and IOU on OD and OC segmentation are 0.87, 0.78 and 0.63, 0.48, respectively
Sun et al. [127]	Fundus Images	The author suggested a CIEU-Net for segregation of retinal vessels, which combines semantic segregation modules and basic medical image segregation methods.	Drive, CHASE_DB1 and STARE Dataset	F1-score of 0.8227, 0.8037 and 0.8230 respectively
Sambyal et al. [128]	Fundus Images	The author presented a modern and advanced UNet architecture that consists of a residual network and employs shuffling periodically with sub-pixel convolution.	IDRiD	Accuracy and Dice Score of 99.98% and 0.9998
Saha et al. [129]	Fundus Images	The researcher published a method that finds the severity level by using segmented output. It solves six subtasks in a single solution.	Drishti-GS	Jaccard Index 0.8572
Saha et al. [130]	Fundus Images	The author proposed a multi-adversary-based fully convolutional neural network utilizing two discriminators for retinal anatomy and pathology segregation.	DRIVE and IDRiD dataset	Accuracy of 0.9641 on the Drishti-GS dataset
Arsalan et al. [131]	Fundus Images	Author published a dual-residual-stream-based vessel segregation network (Vess-Net) which uses artificial intelligence to aid the diagnosis of retinopathy.	DRIVE, STARE and CHASE_DB1 dataset	sensitivity (Se), specificity (Sp), area under the curve (AUC), and accuracy (Acc) of 80.22%, 98.1%, 98.2%, and 96.55% for DRIVIE; 82.06%, 98.41%, 98.0%, and 97.26% for CHASE-DB1; and 85.26%, 97.91%, 98.83%, and 96.97% for STARE dataset
Abbas et al. [132]	Fundus Images	The researcher presented advanced hypertensive retinopathy (HYPER-RETINO) framework divided into several steps, segments the HR based on five grades.	STARE dataset	F1-Score and Accuracy of 92.0% and 92.6%, respectively
Huang et al. [133]	MRI, CT, PET	The author published a computerized brain tumor segregation system using (GammaNet) and combining pooling (DenseASPP) modules and AGC blocks.	TCGA-LGG	DSC and Sensitivity of 85.8% and 87.8%
Ye et al. [134]	MRI, CT, PET	The author presented a model for medical images which portrays high-level supervision into lower layers of neural networks and is a 3D Center-crop Dense.	BraTS 2017	DSC and Sensitivity of the whole are 88.7% and 84.3%
Zhou et al. [135]	MRI, CT, PET	The researcher presented a model with a three-stage network combined by 3D U-Net, producing extra context, fusing with the Multi-sequence MRI, and making the segmentation issue more efficient.	BraTS 2017	Dice Score of WT, TC and ET are 89.4%, 81.6% and 73.0%, respectively
Zhou et al. [136]	MRI, CT, PET	The author proposed an advanced model that integrates with the backbone and uses a 3D atrous-convolution feature pyramid and utilizing a 3D fully linked conditional random field, followed by post-processing steps.	BraTS 2013, 2015 and 2018	Validated on BraTS 2018 with Dice score of 0.8658, 0.7688, and 0.74434 against WT, TC, and ET, respectively

(continued on next page)

Table 8 (continued)

Cited Paper	Modality	Description	Database	Performance metrics & Results
Ding et al. [137]	MRI, CT, PET	The author suggested a new multi-path combined network with adaptive nature that uses ResNets "skip connection."	BraTS 2015	Dice Score for the whole tumor with 10 Epoch is 0.827
Zhang et al. [138]	MRI, CT, PET	The researcher proposed a multi-task driven U-Nets, BU-Net is also employed to reduce noise disturbance, followed by an S-CE loss.	BraTS 2015 and BrainWeb	Dice Score of 0.82 and 0.9950, respectively
Sun et al. [139]	MRI, CT, PET	The author proposed a 3D contemporary convolutional network using multi-pathway architecture is used to extract characteristics from MRI pictures successfully.	BraTS 2019	Dice Score for WT, TC and ET is 0.89, 0.78 and 0.76, respectively
Zhou et al. [140]	MRI, CT, PET	The author presented a three sub-network model making use of the existing modalities to create a 3D feature-refined picture that depicts the missing part.	BraTS 2018	Avg Dice Score of 72.3
Zhou et al. [141]	MRI, CT, PET	The researcher proposed a model which encodes using a 3D ShuffleNetV2 and then a decoder, followed by a loss function fusion and refined by post-processing.	BraTS 2018	Dice Score of ET, WT and TC is 81.81%, 91.21% and 86.62%, respectively
Shehab et al. [142]	MRI, CT, PET	The author presented a ResNets model that has a shortcut skip connection that works in conjunction with the layers of convolutional neural networks.	BraTS 2015	Accuracy of complete, core, and enhancing regions is 83%, 90%, and 85%, respectively
Huang et al. [143]	CT	The researcher presented a model which detects Nodule with quicker regional-CNN (R-NN) with improvised Fully CNN.	LIDC dataset	IOU score 70.24%
Zhu et al. [144]	CT	The author proposed a completely automatic lung CT diagnosis method using two deep 3D convolutional networks. A 3D Faster R-CNN and a U-netlike encoder-decoder structure.	LIDC-IDRI	Diagnosis accuracy of 81.41% and avg accuracy of 99%
Bana et al. [145]	CT	The author published a model split into two parts: segmentation and classification.	LUNA16	IOU with TP and FP detection are 99% and 99%, respectively
Wu et al. [146]	CT	The researcher introduced a CNN that functions as a multi-task learning system to diagnose the malignancy of a lung tumor accurately.	LIDCIDRI	Accuracy and Dice score of 97.58% and 73.89%, respectively
Meraj et al. [147]	CT	The author proposed a system for predicting lung cancer and distinguishing between benign and malignant tumors.	LIDC	Accuracy of 99.23%
Qin et al. [148]	CT	Researcher published a model with two phases. The first phase increases the variety of samples, while the other creates synthetic CT images to make a three-dimensional (3D) CNN model.	LIDC-IDRI	Accuracy and DSC score of 0.9904 and 0.8483

localization of the OD center. Secondly, the transfer learning method can bring a lot of improvement to overcome domain transfer. An approach driven by data for constructing a fundus picture database and a process driven by the model could be incorporated with proficient data to conquer the transfer of domain [122]. Looking forward, future research could bring further proficiency and a better approach by using pre-trained networks or deep supervision to improve the backpropagation of gradients. The proposed model provides a good starting point for generalizing the framework with the available third-party dataset. This framework could also lead to automatic glaucoma detection by using an additional classifier [123]. Future attempts could be made to explore a more efficient and lighter version of the proposed framework to be easily embedded on commercial medical devices. Future work should investigate the proposed model's extension to detect other retinal abnormalities like soft exudates and haemorrhages on clinically arranged and other fundus image datasets [124]. Future research should further develop a similar framework to be applied to different mapping options, which could obtain good segmentation results with lesser trainable parameters performance. In addition, this framework can also be utilized for semantic segmentation in other domains [125]. A recommended future research is to develop a framework that could differentiate between the two, hypertensive retinopathy (HR) and diabetic retinopathy (DR) [131].

7.2.4. Future suggestions to optimize brain tumor arts

Future work is undoubtedly required to collect the images and make a new dataset to curtail the limitation of segmenting gliomas and evaluate the performance [133]. As the proposed model only validates the

brain tumor segmentation dataset; this model can be tested on various other types of modalities in the future. Secondly, the training stages could be integrated and merged into one, rather than two, in the proposed model [134]. Future research could be carried out on identifying the small and tiny segments of a tumor by taking several potential directions for tackling this challenge [135]. An interesting future topic can be segmentation based on multi-modality MR images simultaneously and fusion on the pillars of the deep neural network [136].

Further research should dig into different functions' performance which is normalized, for example, the batch normalizing with different sized of batch, instance normalization, normalization of groups, local response normalization, and normalization of layer [137]. Future work is certainly required to expand the proposed model on other multi-modal segregation issues, for example, MRI and CT, and compare it with other existing models to improve the performance. In addition, the model could also be validated on different publicly or privately available datasets to make a robust framework [138]. Future research might apply lightweight processes in the decoder to decrease the network parameters. Contemporary to that, ERV-Net layers could be increased to get a larger receptive field, which could ultimately improve the performance of ERV-Net. This model could also be replicated on an open-source deep learning platform [139]. Future research should focus on the modification and up gradations in the model in terms of performance for feature extraction of LGG brain tumors. With this, we can improve the robustness, validity, and accuracy of MRI-based brain lesion segregation [140].

7.2.5. Future suggestions to optimize pulmonary nodules arts

Future research might evaluate the proposed model on independent training and testing datasets to perform a comprehensive and detailed comparison. The model performance could also be assessed by applying it to the entire volume as it is a critical concern in terms of performance [143]. Future research should strategically examine the combination of speed and effectiveness and then further be applied across different imaging modalities and commercial medical applications [144]. In the future, nodule malignancy identification of lungs with unlabelled data obtained from pathology needs to be focused upon. Moreover, neural networks can further increase the model interpretability [145]. Future work might be undertaken to investigate the application of the proposed model in many other applications related to medical images for segregation problems, such as MRI for the detection of brain lesions [146]. Future research should certainly include designing new ideas and techniques to solve the issue of an imbalanced dataset. The datasets are obtained from different radiologists, with the subjective judgment of each doctor leading to additional analyses and conclusions. Hence, affects the performance of the model drastically [147].

7.2.6. Future directions for improving deep learning arts

- 1 Create deeper DL architecture by stacking the smaller kernels to each layer which reduces higher memory usage cost and improve the computation cost for semantic segmentation of medical images.
- 2 A comprehensive comparison and analysis of volumetric convolutional methods with sequenced models are possible future directions.
- 3 In region-based approaches, how to avoid blocky segments when using region merging and splitting methods?
- 4 For supervised learning of medical images what could be the possible image annotation methods to improve the semantic segmentation of medical images?
- 5 What could be the possible measures in choosing ANN architectures for semantic segmentation of medical images as these are kind of black box problems?
- 6 Designing a function with a single loss that computes distance function based on cross entropy and overlap simultaneously. This can be possible by reviewing the previously available overlap-based and distance functions.
- 7 How can multimodal features be helpful for semantic segmentation of medical images?

8. Conclusion

Semantic segmentation (Semseg) can be used by many medical experts in the domain of radiology, ophthalmologists, dermatologist, and image-guided radiotherapy. In this review article, we have presented perspectives on the development of an architectural and operational mechanism for each machine learning-based semantic segmentation approach with merits and demerits. Moreover, many authors have developed several Semseg methods and examined their performance in a variety of applications, especially in medical image analysis (e.g., medical image classification and segmentation). This article is further presented a comprehensive investigation of how different architectures are helpful for medical image segmentation. We discussed three major categories, which include fully convolutional network (FCN), region-based and sickly supervised segregation frameworks, and the optimization functions used for semantic segmentation to improve the performance measures. According to some state-of-the-art methods, the Encoder-decoder framework is among the most promising method for semantic segmentation of medical images as it contains long and short skip connections. We observed that due to the scarcity of annotated medical image datasets the use of supervised learning techniques in medical image segmentation is limited. That is why weakly supervised or unsupervised methods have been applied to semantic segmentation tasks. We also discussed the advantages of deep learning techniques and

their limitations for segregation related to biomedical pictures and further discussed some possible future research directions. Finally, advantages, open challenges, and possible future directions are elaborated in the discussion part, beneficial to the research community to understand the significance of the available medical imaging segmentation technology based on Semseg and thus deliver robust segmentation solutions.

CRediT authorship contribution statement

Imran Qureshi: Conceptualization, Methodology, Software, Data curation, Writing – original draft, Visualization, Investigation. **Junhua Yan:** Supervision. **Qaisar Abbas:** Conceptualization, Methodology, Writing – review & editing, Investigation, Validation. **Kashif Shaheed:** Data curation, Writing – review & editing. **Awais Bin Riaz:** Writing – review & editing. **Abdul Wahid:** Visualization. **Muhammad Waseem Jan Khan:** Visualization. **Piotr Szczuko:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

Acknowledgements

The authors extend their appreciation to the Fundamental Research Funds for the Central Universities (Grant No. NJ2020021).

References

- [1] F. Jiang, A. Grigorev, S. Rho, Z. Tian, Y.S. Fu, W. Jifara, K. Adil, Medical image semantic segmentation based on deep learning, *Neural Comput. Appl.* 29 (5) (2018) 1257–1265.
- [2] S.A. Taghanaki, K. Abhishek, J.P. Cohen, J.C. Adad, G. Hamarneh, Deep semantic segmentation of natural and medical images: a review, *Artif. Intell. Rev.* 54 (1) (2021) 137–178.
- [3] D.L. Pham, C. Xu, J.L. Prince, Current methods in medical image segmentation, *Annu. Rev. Biomed. Eng.* 2 (1) (2000) 315–337.
- [4] J. Bertels, T. Eelbode, M. Berman, D. Vandermeulen, F. Maes, R. Bisschops, M. B. Blaschko, Optimizing the dice score and jaccard index for medical image segmentation: Theory and practice, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, 2019.
- [5] A. Isin, C. Direkoglu, M. Sah, Review of MRI-based brain tumor image segmentation using deep learning methods, *Procedia Comput. Sci.* (2016) 317–324.
- [6] R. Millioni, S. Sbrignadello, A. Tura, E. Iori, E. Murphy, P. Tessari, The inter- and intra-operator variability in manual spot segmentation and its effect on spot quantitation in two-dimensional electrophoresis analysis, *Electrophoresis* 31 (10) (2010) 1739–1742.
- [7] J.E. Iglesias, Globally optimal coupled surfaces for semi-automatic segmentation of medical images, in: International Conference on Information Processing in Medical Imaging, Springer, Cham, 2017, pp. 610–621.
- [8] M. Fan, T.C.M. Lee, Variants of seeded region growing, *IET Image Process* 9 (6) (2014) 478–485.
- [9] J. Fan, R. Wang, S. Li, C. Zhang, Automated cervical cell image segmentation using level set based active contour model, in: 2012 12th int. Conf. Control. Autom. Robot. Vision, ICARCV, 2012, 2012, pp. 877–882.
- [10] Y.J. Kim, S.H. Lee, C.M. Park, K.G. Kim, Evaluation of semi-automatic segmentation methods for persistent ground glass nodules on thin-section CT scans, *Healthc. Inform. Res.* 22 (4) (2016) 305–315.
- [11] H.R. Roth, C. Shen, H. Oda, M. Oda, Y. Hayashi, K. Misawa, K. Mori, Deep learning and its application to medical image segmentation, *Med. Imaging Technol.* 36 (2) (2018) 63–71.
- [12] X. Zhou, K. Yamada, T. Kojima, R. Takayama, S. Wang, X. Zhou, T. Hara, H. Fujita, Performance evaluation of 2D and 3D deep learning approaches for automatic segmentation of multiple organs on CT images. *Medical imaging 2018, Computer-Aided Diagnosis*, 2018, p. 83.
- [13] D. Shen, G. Wu, H.I. Suk, Deep learning in medical image analysis, *Annu. Rev. Biomed. Eng.* 19 (1) (2017) 221–248.

- [14] O. Jamshidi, A.H. Pilevar, Automatic segmentation of medical images using fuzzy c-means and the genetic algorithm, *J. Comput. Med.* (2013).
- [15] Y. Yang, C. Feng, R. Wang, Automatic segmentation model combining U-Net and level set method for medical images, *Expert Syst. Appl.* 153 (2020) 113419.
- [16] Y. Xu, Y. Wang, J. Yuan, Q. Cheng, X. Wang, P.L. Carson, Medical breast ultrasound image segmentation by machine learning, *Ultrasonics* 91 (2019) 1–9.
- [17] F. Milletari, N. Navab, S.A. Ahmadi, V-net: Fully convolutional neural networks for volumetric medical image segmentation, in: 2016 Fourth International Conference on 3D Vision (3DV), IEEE, 2016, pp. 565–571.
- [18] N. Zhang, Y.X. Cai, Y.Y. Wang, Y.T. Tian, X.L. Wang, B. Badami, Skin cancer diagnosis based on optimized convolutional neural network, *Artif. Intell. Rev.* (2020), 101756.
- [19] Badea M.S., Felea I.I., Florea L.M., Vertan C., The use of deep learning in image segmentation, classification and detection, arXiv preprint arXiv:1605.09612.
- [20] Mondal A.K., Dolz J., Desrosiers C., Few-shot 3d multi-modal medical image segmentation using generative adversarial learning, arXiv preprint arXiv: 1810.12241.
- [21] T.A. Ngo, G. Carneiro, Fully automated non-rigid segmentation with distance regularized level set evolution initialized and constrained by deep-structured inference, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 3118–3125.
- [22] Ngo T.A., Carneiro G., Lung segmentation in chest radiographs using distance regularized level set and deep-structured learning and inference, in: 2015 IEEE International Conference on Image Processing (ICIP), pp. 2140–2143.
- [23] T.A. Ngo, Z. Liu, G. Carneiro, Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance, *Med. Image Anal.* 35 (2017) 159–171.
- [24] Q. Abbas, I. Fondon, A. Sarmiento, S. Jimenez, P. Alemany, Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features, *Med. Biol. Eng. Comput.* 55 (11) (2017) 1959–1974.
- [25] J. Kemnitz, F.E.C.F.Baumgartner, A. Chaudhari, R. A. W. Wirth, S.K. Eder, E. Konukoglu, Clinical evaluation of fully automated thigh muscle and adipose tissue segmentation using a U-Net deep learning architecture in context of osteoarthritic knee pain, *Magnetic Resonance Materials in Physics, BIOL. MED.* 33 (4) (2020) 483–493.
- [26] N. Orlando, D.J. Gillies, I. Gyakas, C. Romagnoli, D. D’Souza, A. Fenster, Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images, *Med. Phys.* 47 (6) (2020) 2413–2426.
- [27] B.A. Skourt, A. El Hassan, A. Majda, Lung CT image segmentation using deep neural networks, *Procedia Comput. Sci.* 127 (2018) 109–113.
- [28] Yu Weng, T. Zhou, Y. Li, X. Qiu, Nas-unet: Neural architecture search for medical image segmentation, *IEEE Access* 7 (2019) 44247–44257.
- [29] R. Brügger, C.F. Baumgartner, E. Konukoglu, A partially reversible U-Net for memory-efficient volumetric image segmentation, in: International conference on medical image computing and computer-assisted intervention, Springer, Cham, 2019, pp. 429–437.
- [30] CS. Perone, E. Calabrese, J. Cohen-Adad, Spinal cord gray matter segmentation using deep dilated convolutions, *Sci. Rep.* 8 (1) (2018) 1–13.
- [31] X. Xu, Q. Lu, L. Yang, S. Hu, D. Chen, Yu Hu, Y. Shi, Quantization of fully convolutional networks for accurate biomedical image segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 8300–8308.
- [32] M. Paschali, S. Gasperini, A.G. Roy, MY-S. Fang, N. Navab, 3DQ: Compact quantized neural networks for volumetric whole brain segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, 2019, pp. 438–446.
- [33] D. Nie, Y. Gao, L. Wang, D. Shen, ASDNet: attention based semi-supervised deep networks for medical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, 2018, pp. 370–378.
- [34] S.M. Anwar, M. Majid, A. Qayum, M. Awais, M. Alnowami, M.K. Khan, Medical image analysis using convolutional neural networks: a review, *J. Med. Syst.* 42 (11) (2018) 1–13.
- [35] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
- [36] Sinha A., Dolz J., Multi-scale guided attention for medical image segmentation, arXiv preprint arXiv:1906.02849.
- [37] Y. Qin, K. Kamnitsas, S. Ancha, J. Nanavati, G. Cottrell, A. Criminisi, A. Nori, Autofocus layer for semantic segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 603–611.
- [38] S. Lian, Z. Luo, Z. Zhong, X. Lin, S. Su, S. Li, Attention guided U-Net for accurate iris segmentation, *J. Visual Commun. Image Represent.* 56 (2018) 296–304.
- [39] F. Isensee, J. Petersen, A. Klein, D. Zimmerer, P.F. Jaeger, S. Kohl, J. Wasserthal, G. Koehler, T. Norajitra, S. Wirkert, nnU-Net: Self-adapting framework for U-Net-based medical image segmentation. *Bildverarbeitung fur die Medizin*, Springer, 2019, 22–22.
- [40] J. Schlemper, O. Oktay, M. Schaap, M. Heinrich, B. Kainz, B. Glocker, D. Rueckert, Attention gated networks: learning to leverage salient regions in medical images, *Med. Image Anal.* 53 (2019) 197–207.
- [41] M. Drozdzal, G. Chartrand, E. Vorontsov, M. Shakeri, L. Di Jorio, A. Tang, A. Romero, Y. Bengio, C. Pal, S. Kadoury, Learning normalized inputs for iterative estimation in medical image segmentation, *Med. Image Anal.* 44 (2018) 1–13.
- [42] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, J. Liu, Ce-net: Context encoder network for 2d medical image segmentation, *IEEE Trans. Med. Imaging* 38 (10) (2019) 2281–2292.
- [43] Vorontsov E., Molchanov P., Beckham C., Byeon W., Mello S.D., Jampani V., Liu M.Y., Kadoury S., Kautz J., Towards semi-supervised segmentation via image-to-image translation, arXiv preprint arXiv:1904.01636.
- [44] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, UNet++: A nested U-Net architecture for medical image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 2018, pp. 3–11.
- [45] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, 2015.
- [46] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (8) (1997) 1735–1780, 1997.
- [47] W. Bai, H. Suzuki, C. Qin, G. Tarroni, O. Oktay, P.M. Matthews, D. Rueckert, Recurrent neural networks for aortic image sequence segmentation with sparse annotations, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 586–594.
- [48] Y. Gao, J.M. Phillips, Y. Zheng, R. Min, P.T. Fletcher, G. Gerig, Fully convolutional structured LSTM networks for joint 4D medical image segmentation, in: 2018 IEEE 15th International Symposium on Biomedical Imaging, IEEE, 2018, pp. 1104–1108.
- [49] Li H., Li J., Lin X., Qian X., Pancreas segmentation via spatial context based u-net and bidirectional lstm, arXiv preprint arXiv:1903.00832.
- [50] M.Z. Alom, C. Yakopcic, M. Hasan, T.M. Taha, V.K. Asari, Recurrent residual U-Net for medical image segmentation, *J Med. Imag.* 6 (1) (2019) 14006.
- [51] M. Zhao, G. Hamarneh, Retinal image classification via vasculature-guided sequential attention, in: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019.
- [52] M. Zhao, G. Hamarneh, Tree-LSTM: using LSTM to encode memory in anatomical tree prediction from 3D Images. *International Workshop on Machine Learning in Medical Imaging*, Springer, Cham, 2019, pp. 637–645.
- [53] T-Y Lin, P. Goyal, R. Girshick, Kaiming He, and Piotr Dollár, Focal loss for dense object detection, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2980–2988.
- [54] A. Buslaev, V.I. iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, A. A. Kalinin, Albumentations: fast and flexible image augmentations, *Information* 11 (2) (2020) 125.
- [55] S.S.M Salehi, D. Erdogmus, A. Gholipour, Tversky loss function for image segmentation using 3D fully convolutional deep networks. *International workshop on Machine Learning in Medical Imaging*, Springer, 2017, pp. 379–387.
- [56] K.C.L. Wong, M. Moradi, H. Tang, T.S. Mahmood, 3D segmentation with exponential logarithmic loss for highly unbalanced object sizes, in: International Conference on Medical Image Computing and Computer Assisted Intervention, Springer, 2018, pp. 612–619.
- [57] M. Buda, A. Saha, M.A. Mazurowski, Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm, *Comput. Biol. Med.* 109 (2019) 218–225.
- [58] H. Kervadec, J. Dolz, M. Tang, E. Granger, Y. Boykov, I.B. Ayed, Constrained-CNN losses for weakly supervised segmentation, *Med. Image Anal.* 54 (2019) 88–99.
- [59] C.S. Perone, J. Cohen-Adad, Deep semi-supervised segmentation with weight-averaged consistency targets. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, Cham, 2018, pp. 12–19.
- [60] C.S. Perone, J. Cohen-Adad, Promises and limitations of deep learning for medical image segmentation, *J. Med. Artif. Intell.* 2 (1) (2019) 1–2.
- [61] K.B. Girum, G. Crehange, R. Hussain, A. Lalande, Fast interactive medical image segmentation with weakly supervised deep learning method, *Int. J. Comput. Assist. Radiol. Surg.* 15 (9) (2020) 1437–1444.
- [62] S.A Taghanaki, Y. Zheng, S.K. Zhou, B. Georgescu, P. Sharma, D. Xu, D. Comaniciu, G. Hamarneh, Combo loss: handling input and output imbalance in multi-organ segmentation, *Comput. Med. Imaging Graph.* 75 (2019) 24–33.
- [63] A. Chaurasia, E. Culurciello, Linknet: Exploiting encoder representations for efficient semantic segmentation, in: 2017 IEEE Visual Communications and Image Processing (VCIP), IEEE, 2017.
- [64] X. Zhu, Z. Cheng, S. Wang, X. Chen, G. Lu, Coronary angiography image segmentation based on PSPNet, *Comput. Methods Programs Biomed.* 200 (2021), 105897.
- [65] Kirillov A., K.He, R.G. Dollar P., A unified architecture for instance and semantic segmentation, (2017).
- [66] X. Liu, Z. Deng, Y. Yang, Recent progress in semantic image segmentation, *Artif. Intell. Rev.* 52 (2) (2019) 1089–1106.
- [67] Z. Huang, X. Wang, L. Huang, C. Huang, Y. Wei, W. Liu, Ccnet: Criss-cross attention for semantic segmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019.
- [68] T.D. Tô, D.T. Lan, T.T.H. Nguyen, H.P.N T.T.N.Nguyen, L. Phuong, T.Z. Nguyen, Ensembled Skin Cancer Classification, ISIC 2019 Challenge Submission, Diss. ISIC, 2019.
- [69] T. Zhou, S. Ruan, S. Canu, A review: Deep learning for medical image segmentation using multi-modality fusion, *Array* 3 (2019), 100004.
- [70] N. Tajbakhsh, L. Jeyaseelan, Q. Li, J.N. Chiang, Z. Wu, X. Ding, Embracing imperfect datasets: a review of deep learning solutions for medical image segmentation, *Med. Image Anal.* 63 (2020), 101693.

- [71] A. de Brebinson, G. Montana, Deep neural networks for anatomical brain segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015.
- [72] H. Chen, Q. Dou, L. Yu, J. Qin, P.A. Heng, VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images, *Neuroimage* 170 (2018) 446–455.
- [73] A.G. Roy, S. Conjeti, N. Navab, C. Wachinger, Bayesian quicknat: model uncertainty in deep whole-brain segmentation for structure-wise quality control, *Neuroimage* 195 (2019) 11–22.
- [74] L. Wang, D. Nie, G. Li, E. Puylbareau, J. Dolz, Q. Zhang, F. Wang, J. Xia, Z. Wu, J. W. Chen, K.H. Thung, T.D. Bui, J. Shin, Z. G, G. Zeng, V.S. Fonov, A. Doyle, Y. Xu, P. Moeskops, J.P.W. Pluim, C. Desrosiers, I.B. Ayed, G. Sanroma, O.M. Benkarim, A. Casamitjana, V. Vilaplana, W. Lin, G. Li, D. Shen, Benchmark on automatic six-month-old infant brain segmentation algorithms: the iSeg-2017 challenge, *IEEE Trans. Med. Imaging* 38 (9) (2019) 2219–2230.
- [75] S. Devunooru, A. Alsadoon, P.W.C. Chandana, A. Beg, Deep learning neural networks for medical image segmentation of brain tumours for diagnosis: a recent review and taxonomy, *J. Ambient Intell Humaniz. Comput* 12 (1) (2021) 455–483.
- [76] Islam J., Zhang Y., Towards robust lung segmentation in chest radiographs with deep learning, arXiv preprint arXiv:1811.12638.
- [77] Y. Gordienko, P. Gang, J. Hui, W. Zeng, Y. Kochura, O. Alienin, O. Rokovy, S. Stirenko, Deep learning with lung segmentation and bone shadow exclusion techniques for chest X-ray analysis of lung cancer, in: International Conference on Computer Science, Engineering and Education Applications, Springer, Cham, 2018, pp. 638–647.
- [78] Chen X., Yao L., Zhang Y., Residual attention u-net for automated multi-class segmentation of covid-19 chest ct images, arXiv preprint arXiv:2004.05645.
- [79] A. Mittal, R. Hooda, S. Sofat, Lung field segmentation in chest radiographs: a historical review, current status, and expectations from deep learning, *IET Image Proc.* 11 (11) (2017) 937–952.
- [80] M. Kallenberg, K. Petersen, M. Nielsen, A.Y. Ng, P. Diao, C. Igel, C.M. Vachon, K. Holland, R.R. Winkel, N. Karssemeijer, M. Lillholm, Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring, *IEEE Trans. Med. Imaging* 35 (5) (2016) 1322–1331.
- [81] R. Almajid, J. Shan, Y. Du, M. Zhang, Development of a deep-learning-based method for breast ultrasound image segmentation, in: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, 2018, pp. 1103–1108.
- [82] M. Caballo, D.R. Pangallo, R.M. Mann, I. Sechopoulos, Deep learning-based segmentation of breast masses in dedicated breast CT imaging: radiomic feature stability between radiologists and artificial intelligence, *Comput. Biol. Med.* 118 (2020), 103629.
- [83] J.Z. Cheng, D. Ni, Y.H. Chou, J. Qin, C.M. Tiu, Y.C. Chang, C.S. Huang, D. Shen, C. M. Chen, Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans, *Sci. Rep.* 6 (1) (2016) 1–13.
- [84] R. Zelezniak, J. Weiss, J. Taron, C. Guthier, D.S. Bitterman, C. Hancox, B.H. Kann, D.W. Kim, R.S. Punglia, J. Bredfeldt, B. Foldyna, P. Eslami, M.T. Lu, U. Hoffmann, R. Mak, H.J.W.L. Aerts, Deep-learning system to improve the quality and efficiency of volumetric heart segmentation for breast cancer, *NPJ Digit. Med.* 4 (1) (2021) 1–7.
- [85] R. Cheng, H.R. Roth, N.S. Lay, L. Lu, B. Turkbey, W. Gandler, E.S. McCready, T. J. Pohida, P.A. Pinto, P.L. Choyke, M.J. McAuliffe, R.M. Summers, Automatic magnetic resonance prostate segmentation by deep learning with holistically nested networks, *J. Med. Imaging* 4 (4) (2017), 041302.
- [86] Y. Guo, Y. Gao, D. Shen, Deformable MR prostate segmentation via deep feature learning and sparse patch matching, *IEEE Trans. Med. Imaging* 35 (4) (2015) 1077–1089.
- [87] E.M.A. Anas, P. Mousavi, P. Abolmaesumi, A deep learning approach for real time prostate segmentation in freehand ultrasound guided biopsy, *Med. Image Anal.* 48 (2018) 107–116.
- [88] Z. Tian, L. Liu, Z. Zhang, B. Fei, PSNet: prostate segmentation on MRI based on a convolutional neural network, *J. Med. Imaging* 5 (2) (2018), 021208.
- [89] D. Jha, P.H. Smedsrub, M.A. Riegler, P. Halvorsen, T. de Lange, D. Johansen, H. D. Johansen, in: Kvasir-seg: A segmented polyp dataset, International Conference on Multimedia Modeling, Springer, Cham, 2020, pp. 451–462.
- [90] P. Brando, E. Mazomenos, G. Ciuti, R. Calio, F. Bianchi, A. Menciassi, A.K.P. Dario, A. Arezzo, D. Stoyanov, Fully convolutional neural networks for polyp segmentation in colonoscopy, in: Medical Imaging, Computer-Aided Diagnosis, 10134, 2017, pp. 101–107.
- [91] D.P. Fan, G.P. Ji, T. Zhou, G. Chen, H. Fu, J. Shen, L. Shao, Planet: Parallel reverse attention network for polyp segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, 2020, pp. 263–273.
- [92] Y. Guo, J. Bernal, B.J. Matuszewski, Polyp segmentation with fully convolutional deep neural networks—extended evaluation study, *J. Imaging* 6 (7) (2020) 69.
- [93] L. Xu, M. Jackowski, A. Goshtasby, D. Roseman, S. Bines, C. Yu, A. Dhawan, A. Huntley, Segmentation of skin cancer images, *Image Vision Comput.* 17 (1) (1999) 65–74.
- [94] J. Tang, A multi-direction GVF snake for the segmentation of skin cancer images, *Pattern Recognit.* 42 (6) (2009) 1172–1179.
- [95] R.B. Oliveira, E.M. Filho, Z. Ma, J.P. Papa, A.S. Pereira, J.M.R.S. Tavares, Computational methods for the image segmentation of pigmented skin lesions: a review, *Comput. Methods Progr. Biomed.* 131 (2016) 127–141.
- [96] Q. Jin, H. Cui, C. Sun, Z. Meng, R. Su, Cascade knowledge diffusion network for skin lesion diagnosis and segmentation, *Appl. Soft Comput.* 99 (2021), 106881.
- [97] S.M. Thomas, J.G. Lefevre, G. Baxter, N.A. Hamilton, Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer, *Med. Image Anal.* 68 (2021), 101915.
- [98] M.K. Hasan, L. Dahal, P.N. Samarakoon, F.I. Tushar, R. Marti, DSNet: automatic dermoscopic skin lesion segmentation, *Comput. Biol. Med.* 120 (2020), 103738.
- [99] P. Shan, Y. Wang, C. Fu, W. Song, J. Chen, Automatic skin lesion segmentation based on FC-DPN, *Comput. Biol. Med.* 123 (2020), 103762.
- [100] W. Li, A.N.J. Raj, T. Tjahjadi, Z. Zhuang, Digital hair removal by deep learning for skin lesion segmentation, *Pattern Recognit.* 117 (2021), 107994.
- [101] Q. Jin, H. Cui, C. Sun, Z. Meng, R. Su, Cascade knowledge diffusion network for skin lesion diagnosis and segmentation, *Appl. Soft Comput.* 99 (2021), 106881.
- [102] R. Gu, L. Wang, L. Zhang, DE-Net: a deep edge network with boundary information for automatic skin lesion segmentation, *Neurocomputing* 468 (2022) 71–84.
- [103] R. Kaymak, C. Kaymak, A. Ucar, Skin lesion segmentation using fully convolutional networks: a comparative experimental study, *Exp. Syst. Appl.* 161 (2020), 113742.
- [104] M. Sarker, M. Kamal, H.A. Rashwan, F. Akram, V.K. Singh, S.F. Banu, F.U. H. Chowdhury, K.A. Chowdhury, SLSNet: skin lesion segmentation using a lightweight generative adversarial network, *Exp. Syst. Appl.* 183 (2021), 115433.
- [105] B. Lei, Z. Xia, F. Jiang, X. Jiang, Z. Ge, Y. Xu, J. Qin, S. Chen, T. Wang, S. Wang, Skin lesion segmentation via generative adversarial networks with dual discriminators, *Med. Image Anal.* 64 (2020), 101716.
- [106] P. Tang, X. Yan, Q. Liang, D. Zhang, AFLN-DGCL: adaptive feature learning network with difficulty-guided curriculum learning for skin lesion segmentation, *Appl. Soft Comput.* 110 (2021), 107656.
- [107] R. Arora, B. Raman, K. Nayyar, R. Awasthi, Automated skin lesion segmentation using attention-based deep convolutional neural network, *Biomed. Signal Process. Control* 65 (2020), 102358.
- [108] P. Pan, H. Chen, Y. Li, N. Cai, L. Cheng, S. Wang, Tumor segmentation in automated whole breast ultrasound using bidirectional LSTM neural network and attention mechanism, *Ultrasonics* 110 (2020), 106271.
- [109] L. Huo, X. Hu, Q. Xiao, Y. Gu, X. Chu, L. Jiang, Segmentation of whole breast and fibroglandular tissue using NuU-Net in dynamic contrast enhanced MR images, *Magn. Reson. Imaging* 82 (2021) 31–41.
- [110] L. Liu, D. Zhao, F. Yu, A.A. Heidari, J. Ru, H. Chen, M. Mafarja, H. Turabieh, Z. Pan, Performance optimization of differential evolution with slime mould algorithm for multilevel breast cancer image segmentation, *Comput. Biol. Med.* 138 (2021), 104910.
- [111] S.D. Pawar, K.K. Sharma, S.G. Sapate, G.Y. Yadav, Segmentation of pectoral muscle from digital mammograms with depth-first search algorithm towards breast density classification, *Probl. Biocybern. Biomed. Eng.* 41 (3) (2021) 1224–1241.
- [112] H. Zou, X. Gong, J. Luo, T. Li, A robust breast ultrasound segmentation method under noisy annotations, *Comput. Methods Programs Biomed.* 209 (2021), 106327.
- [113] Y. Yan, P.H. Conze, G. Quellec, M. Lamard, B. Cochener, G. Coatrieux, Two-stage multi-scale breast mass segmentation for full mammogram analysis without user intervention, *Probl. Biocybern. Biomed. Eng.* 41 (2) (2021) 746–757.
- [114] A.E. Ilesanmi, U. Chaumrattanakul, S.S. Makhanov, A method for segmentation of tumors in breast ultrasound images using the variant enhanced deep learning, *Probl. Biocybern. Biomed. Eng.* 41 (2) (2021) 802–818.
- [115] Q. Huang, Y. Huang, Y. Luo, F. Yuan, X. Li, Segmentation of breast ultrasound image with semantic classification of superpixels, *Med. Image Anal.* 61 (2020), 101657.
- [116] M.V. Rijthoven, M. Balkenhol, K. Siliqa, J.V.D. Laak, F. Ciompi, HookNet: multi-resolution convolutional neural networks for semantic segmentation in histopathology whole-slide images, *Med. Image Anal.* 68 (2021), 101890.
- [117] Y. Tong, Y. Liu, M. Zhao, L. Meng, J. Zhang, Improved U-Net MALF model for lesion segmentation in breast ultrasound images, *Biomed. Signal Process. Control* 68 (2021), 102721.
- [118] R. Intiaz, T.M. Khan, S.S. Naqvi, M. Arsalan, S.J. Nawaz, Screening of glaucoma disease from retinal vessel images using semantic segmentation, *Comput. Electr. Eng.* 91 (2020), 107036.
- [119] S. Pachade, P. Porwal, M. Kokare, L. Giancardo, F. Mériadeau, NENet: nested efficientnet and adversarial learning for joint optic disc and cup segmentation, *Med. Image Anal.* 74 (2021), 102253.
- [120] X. Bian, X. Luo, C. Wang, W. Liu, X. Lin, Optic disc and optic cup segmentation based on anatomy guided cascade network, *Comput. Methods Programs Biomed.* 197 (2020), 105717.
- [121] A. Tulsani, P. Kumar, S. Pathan, Automated segmentation of optic disc and optic cup for glaucoma assessment using improved UNET++ architecture, *Probl. Biocybern. Biomed. Eng.* 41 (2) (2021) 819–832.
- [122] Y. Zhou, Z. Chen, H. Shen, X. Zheng, R. Zhao, X. Duan, A refined equilibrium generative adversarial network for retinal vessel segmentation, *Neurocomputing* 437 (2021) 118–130.
- [123] P. Yin, Y. Xu, J. Zhu, J. Liu, C. Yi, H. Huang, Q. Wu, Deep level set learning for optic disc and cup segmentation, *Neurocomputing* 464 (2021) 330–341.
- [124] X. Yuan, L. Zhou, S. Yu, M. Li, X. Wang, X. Zheng, A multi-scale convolutional neural network with context for joint segmentation of optic disc and cup, *Artif. Intell. Rev.* 113 (2021), 102035.
- [125] J. Morano, Á.S. Hervella, J. Novo, J. Rouco, Simultaneous segmentation and classification of the retinal arteries and veins from color fundus images, *Artif. Intell. Rev.* 118 (2021), 102116.

- [126] Y. Zhang, X. Cai, Y. Zhang, H. Kang, X. Ji, X. Yuan, TAU: transferable attention U-Net for optic disc and cup segmentation, *Knowl.-Based Syst.* 213 (2021), 106668.
- [127] M. Sun, K. Li, X. Qi, H. Tang, G. Zhang, Contextual information enhanced convolutional neural networks for retinal vessel segmentation in color fundus images, *J. Visual Commun. Image Represent.* 77 (2020), 103134.
- [128] N. Sambaly, P. Saini, R. Syal, V. Gupta, Modified u-net architecture for semantic segmentation of diabetic retinopathy images, *Probl. Biocybern. Biomed. Eng.* 40 (3) (2020) 1094–1109.
- [129] Saha, O., Sathish, R. and Sheet, D., Fully convolutional neural network for semantic segmentation of anatomical structure and pathologies in colour fundus images associated with diabetic retinopathy, arXiv preprint arXiv:1902.03122.
- [130] O. Saha, R. Sathish, D. Sheet, Learning with multitask adversaries using weakly labelled data for semantic segmentation in retinal images, in: *International Conference on Medical Imaging with Deep Learning*, 2019, pp. 414–426.
- [131] M. Arsalan, M. Owais, T. Mahmood, S.W. Cho, K.R. Park, Aiding the diagnosis of diabetic and hypertensive retinopathy using artificial intelligence-based semantic segmentation, *J. Clinic. Med.* 8 (9) (2019) 1446.
- [132] Q. Abbas, I. Qureshi, M.E.A. Ibrahim, An automatic detection and classification system of five stages for hypertensive retinopathy using semantic and instance segmentation in densenet architecture, *Sensors* 21 (2021) 6936.
- [133] Z. Huang, Y. Liu, G. Song, Y. Zhao, Gammanet: an intensity-invariance deep neural network for computer-aided brain tumor segmentation, *Optik* 243 (2021), 167441.
- [134] F. Ye, Y. Zheng, H. Ye, X. Han, Y. Li, J. Wang, J. Pu, Parallel pathway dense neural network with weighted fusion structure for brain tumor segmentation, *Neurocomputing* 425 (2021) 1–11.
- [135] T. Zhou, S. Canu, S. Ruan, Fusion based on attention mechanism and context constraint for multi-modal brain tumor segmentation, *Comput. Med. Imaging Graph.* 86 (2020), 101811.
- [136] Z. Zhou, Z. He, Y. Jia, AFPNet: A 3D fully convolutional neural network with atrous-convolution feature pyramid for brain tumor segmentation via MRI images, *Neurocomputing* 402 (2020) 235–244.
- [137] Y. Ding, L. Gong, M. Zhang, C. Li, Z. Qin, A multi-path adaptive fusion network for multimodal brain tumor segmentation, *Neurocomputing* 412 (2020) 19–30.
- [138] J. Zhang, J. Zeng, P. Qin, L. Zhao, Brain tumor segmentation of multi-modality MR images via triple intersecting U-Nets, *Neurocomputing* 421 (2021) 195–209.
- [139] J. Sun, Y. Peng, Y. Guo, D. Li, Segmentation of the multimodal brain tumor image used the multi-pathway architecture method based on 3D FCN, *Neurocomputing* 423 (2021) 34–45.
- [140] T. Zhou, S. Canu, P. Vera, S. Ruan, Feature-enhanced generation and multi-modality fusion based deep neural network for brain tumor segmentation with missing MR modalities, *Neurocomputing* 466 (2021) 102–112.
- [141] X. Zhou, X. Li, K. Hu, Y. Zhang, Z. Chen, X. Gao, ERV-Net: an efficient 3D residual neural network for brain tumor segmentation, *Expert Syst. Appl.* 170 (2019), 114566.
- [142] L.H. Shehab, O.M. Fahmy, S.M. Gasser, M.S. El-Mahallawy, An efficient brain tumor image segmentation based on deep residual networks (ResNets), *J King Saud Univ Sci* 33 (6) (2021) 404–412.
- [143] X. Huang, W. Sun, T.L. Tseng, C. Li, W. Qian, Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic CT scans using deep convolutional neural networks, *Comput. Med. Imaging Graph.* 74 (2019) 25–36.
- [144] W. Zhu, C. Liu, W. Fan, X. Xie, DeepLung: Deep 3D dual path nets for automated pulmonary nodule detection and classification, in: *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV, 2018*, pp. 673–681.
- [145] S. EL-Bana, A. Al-Kabbany, M. Sharkas, A two-stage framework for automated malignant pulmonary nodule detection in CT scans, *Diagnostics* 10 (3) (2020) 1–19.
- [146] B. Wu, Z. Zhou, J. Wang, Y. Wang, Joint learning for pulmonary nodule segmentation, attributes and malignancy prediction, in: *Proceedings - International Symposium on Biomedical Imaging, 2018*, pp. 1109–1113.
- [147] T. Meraj, H.T. Rauf, S. Zahoor, A. Hassan, M.I.U. Lali, L. Ali, S.A.C. Bukhari, U. Shoaib, Lung nodules detection using semantic segmentation and classification with optimal features, *Neural Comput Appl* 33 (17) (2021) 10737–10750.
- [148] Y. Qin, H. Zheng, X. Huang, J. Yang, Y.M. Zhu, Pulmonary nodule segmentation with CT sample synthesis using adversarial networks, *Med. Phys.* 46 (3) (2019) 1218–1229.
- [149] F. Hooralı, H. Khosravi, B. Moradi, Automatic Bacillus anthracis bacteria detection and segmentation in microscopic images using UNet++, *J. Microbiol. Methods* 177 (2020), 106056.
- [150] F. Hooralı, H. Khosravi, B. Moradi, IRUNet for medical image segmentation, *Exp. Syst. Appl.* 191 (2022), 116399.
- [151] X. Liu, L. Yang, J. Chen, S. Yu, K. Li, Region-to-boundary deep learning model with multi-scale feature fusion for medical image segmentation, *Biomed. Signal Process. Control* 71 (2022), 103165.
- [152] J. Chen, N. Yang, M. Zhou, Z. Zhang, X. Yang, A configurable deep learning framework for medical image analysis, *Neural Comput. Appl.* 34 (10) (2022) 7375–7392.
- [153] J. Chen, K. Li, Z. Zhang, K. Li, PS. Yu, A survey on applications of artificial intelligence in fighting against COVID-19, *ACM Comput. Surv.* 54 (8) (2021) 1–32.
- [154] B. Pu, K. Li, S. Li, N. Zhu, Automatic fetal ultrasound standard plane recognition based on deep learning and IIoT, *IEEE Trans. Ind. Inf.* 17 (11) (2021) 7771–7780.
- [155] S.U. Khan, F. Wang, JJ. Liou, Y. Liu, Segmentation of breast tumors using cutting-edge semantic segmentation models, *Comput. Methods Biomed. Eng. Imaging Vis.* (2022) 1–11.
- [156] R. Wang, T. Lei, R. Cui, B. Zhang, H. Meng, AK. Nandi, Medical image segmentation using deep learning: A survey, *IET Image Proc.* 16 (5) (2022) 1243–1267.
- [157] P. Malhotra, S. Gupta, D. Koundal, A. Zagaria, W. Enbeyle, Deep neural networks for medical image segmentation, *J. Healthc. Eng.* (2022).
- [158] K. Shaheed, A. Mao, I. Qureshi, M. Kumar, Q. Abbas, I. Ullah, X. Zhang, A systematic review on physiological-based biometric recognition systems: Current and future trends, *Arch. Comput. Meth. Eng.* 28 (7) (2021) 4917–4960.
- [159] K. Shaheed, A. Mao, I. Qureshi, M. Kumar, S. Hussain, X. Zhang, Recent advancements in finger vein recognition technology: methodology, challenges and opportunities, *Inf. Fusion* 79 (2022) 84–109.
- [160] K. Shaheed, A. Mao, I. Qureshi, Q. Abbas, M. Kumar, X. Zhang, Finger-vein presentation attack detection using depthwise separable convolution neural network, *Expert Syst. Appl.* 198 (2022), 116786.
- [161] A. Lin, B. Chen, J. Xu, Z. Zhang, G. Lu, D. Zhang, Ds-transunet: Dual swim transformer u-net for medical image segmentation, *IEEE Trans. Instrum. Meas.* (2022).
- [162] D. Maji, P. Sigedar, M. Singh, Attention Res-UNet with Guided Decoder for semantic segmentation of brain tumors, *Biomed. Signal Process. Control* 71 (2022), 103077.
- [163] I. Qureshi, J. Ma, Q. Abbas, Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning, *Multimed. Tools Appl.* 80 (8) (2021) 11691–11721.
- [164] Q. Abbas, I. Qureshi, J. Yan, K. Shaheed, Machine learning methods for diagnosis of eye-related diseases: a systematic review study based on ophthalmic imaging modalities, *Arch. Comput. Meth. Eng.* (2022) 1–58.
- [165] A. Skouta, A. Elmoufidi, S. Jai-Andaloussi, O. Ouchetto, Hemorrhage semantic segmentation in fundus images for the diagnosis of diabetic retinopathy by using a convolutional neural network, *J. Big Data* 9 (1) (2022) 1–24.
- [166] T. Dhamija, A. Gupta, S. Gupta, R. Katarya, G. Singh, Semantic segmentation in medical images through transfused convolution and transformer networks, *Appl. Intell.* (2022) 1–17.
- [167] B. Yogananda, C. Ganesh, Y. Das, BC. Wagner, SS. Nalawade, D. Reddy, J. Holcomb, MC. Pinho, B. Fei, AJ. Madhuranthakam, JA. Maldjian, Disparity autoencoders for multi-class brain tumor segmentation. *International MICCAI Brainlesion Workshop*, Springer, Cham, 2022, pp. 116–124.