DOI: https://doi.org/10.17816/DD629866

Comparative analysis of modifications of U-Net neuronal network architectures in medical image segmentation

Anastasia M. Dostovalova[1,2], Andrey K. Gorshenin[1,2], Julia V. Starichkova[1], Kirill M. Arzamasov[1,3]

[1] MIREA — Russian Technological University, Moscow, Russia;
[2] Federal Research Center Computer Science and Control of the Russian Academy of Sciences, Moscow, Russia;
[3] Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow, Russia

ABSTRACT

Data processing methods based on neural networks are becoming increasingly popular in medical diagnostics. They are most commonly used to evaluate medical images of human organs using computed tomography, magnetic resonance imaging, ultrasound, and other non-invasive diagnostic methods. Disease diagnosis involves solving the problem of medical image segmentation, i.e. finding groups (regions) of pixels that characterize specific objects in the image. The U-Net neural network architecture developed in 2015 is one of the most successful tools to solve this issue. This review evaluated various modifications of the classic U-net architecture. The papers considered were divided into several key categories, such as modifications of the encoder and decoder; use of attention blocks; combination with elements of other architectures; methods for introducing additional attributes; transfer learning; and approaches for processing small sets of real-world data. Different training sets with the best parameters found in the literature were evaluated (Dice similarity score; Intersection over Union; overall accuracy, etc.). A summary table was developed showing types of images evaluated and abnormalities detected. Promising directions for further modifications to improve the quality of the segmentation are identified. The results can be used to detect diseases, especially cancer. Intelligent medical assistants can implement the presented algorithms.

Keywords: U-Net architecture; segmentation; computed tomography; magnetic resonance imaging; medical diagnostics; oncology diseases.

To cite this article:
Dostovalova AM, Gorshenin AK, Starichkova JuV, Arzamasov KM. Comparative analysis of modifications of U-Net neuronal network architectures in medical image segmentation. *Digital Diagnostics.* 2024;5(4):833–853. DOI: https://doi.org/10.17816/DD629866

Received: 03.04.2024 **Accepted:** 06.06.2024 **Published online:** 21.11.2024

DOI: https://doi.org/10.17816/DD629866

Сравнительный анализ модификаций нейросетевых архитектур U-Net в задаче сегментации медицинских изображений

А.М. Достовалова[1,2], А.К. Горшенин[1,2], Ю.В. Старичкова[1], К.М. Арзамасов[1,3]

[1] МИРЭА — Российский технологический университет, Москва, Россия;
[2] Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Россия;
[3] Научно-практический клинический центр диагностики и телемедицинских технологий, Москва, Россия

АННОТАЦИЯ

Методы обработки данных с использованием нейронных сетей завоёвывают всё большую популярность в области медицинской диагностики. Наиболее часто их применяют при исследовании медицинских изображений органов человека с использованием компьютерной и магнитно-резонансной томографии, ультразвуковых и иных средств неинвазивных исследований. Диагностирование патологии в таком случае сводится к решению задачи сегментации медицинского изображения, то есть поиска групп (областей) пикселов, характеризующих некоторые объекты на снимке. Один из наиболее успешных методов решения данной задачи — разработанная в 2015 году нейросетевая архитектура U-Net. В настоящем обзоре авторы проанализировали разнообразные модификации классической архитектуры U-Net. Рассмотренные работы разделены на несколько ключевых направлений: модификации кодировщика и декодировщика; использование блоков внимания; комбинирование с элементами других архитектур; методы внедрения дополнительных признаков; трансферное обучение и подходы для обработки малых наборов реальных данных. Изучены различные обучающие наборы, для которых приведены лучшие достигнутые в литературе значения метрик (показатель сходства Dice; пересечение над объединением Intersection over Union; общая точность и др.). Также создана сводная таблица с указанием типов анализируемых изображений и выявляемых патологий на них. Обозначены перспективные направления дальнейших модификаций для повышения качества решения задач сегментации. Результаты могут быть полезны в области выявления заболеваний, прежде всего, онкологических. Представленные алгоритмы могут стать частью профессиональных медицинских интеллектуальных ассистентов.

Ключевые слова: архитектура U-Net; сегментация; компьютерная томография; магнитно-резонансная томография; медицинская диагностика; онкологические заболевания.

Как цитировать:
Достовалова А.М., Горшенин А.К., Старичкова Ю.В., Арзамасов К.М. Сравнительный анализ модификаций нейросетевых архитектур U-Net в задаче сегментации медицинских изображений // Digital Diagnostics. 2024. Т. 5, № 4. С. 833–853. DOI: https://doi.org/10.17816/DD629866

Рукопись получена: 03.04.2024 **Рукопись одобрена:** 06.06.2024 **Опубликована online:** 21.11.2024

DOI: https://doi.org/10.17816/DD629866

U-Net神经网络架构在医学图像分割任务中的改型比较分析

Anastasia M. Dostovalova[1,2], Andrey K. Gorshenin[1,2], Julia V. Starichkova[1], Kirill M. Arzamasov[1,3]

[1] MIREA-Russian Technological University, Moscow, Russia;
[2] Federal Research Center Computer Science and Control of the Russian Academy of Sciences, Moscow, Russia;
[3] "Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department", Moscow, Russia

摘要

在医学诊断领域，使用神经网络的数据处理方法越来越受欢迎。它们最常用于使用计算机断层扫描和磁共振成像、超声波和其他非侵入性研究工具来研究人体器官的医学图像。在这种情况下，病理诊断归结为解决医学图像分割的问题，即搜索表征图像中某些对象的像素组（区域）。2015年开发的U-Net神经网络架构是解决这一问题的最成功方法之一。在本文中，作者分析了典型的 U-Net架构的各种改型。研究工作分为几个关键领域：编码器和解码器的改型；注意力模块的使用；与其他架构元素的结合；引入附加特征的方法；迁移学习和处理小型现实数据集的方法。研究了各种训练集，列出了文献中实现的最佳度量值（Dice相似度指标；交集大于联合Intersection over Union；总体准确性等等）。还创建了一份汇总表，说明所分析的图像类型和在这些图像上检测到的病理情况。概述了进一步改型以提高分割任务质量的有前景的方向。这些结果可能有助于疾病检测，主要是肿瘤检测。 所提出的算法可以成为专业医疗智能助手的一部分。

关键词： U-Net架构；分割；计算机断层扫描；磁共振成像；医学诊断；肿瘤病。

引用本文：
Dostovalova AM, Gorshenin AK, Starichkova JuV, Arzamasov KM. U-Net神经网络架构在医学图像分割任务中的改型比较分析. *Digital Diagnostics.* 2024;5(4):833–853. DOI: https://doi.org/10.17816/DD629866

收到： 03.04.2024 **接受：** 06.06.2024 **发布日期：** 21.11.2024

INTRODUCTION

Image processing using artificial-intelligence (AI)-based software plays a central role in modern medical diagnosis. Advancements in computational technology and machine learning algorithms have considerably expanded the capabilities of image analysis in recent decades. Comprehensive clinical decision support systems, including autonomous models, have replaced the previous generation of simple classification frameworks.

Medical image processing initially relied on basic imaging modalities such as radiography and mammography. These modalities have since evolved, and computed tomography (CT) and magnetic resonance imaging (MRI) data are now processed with high efficiency. In the context of diagnostic radiology, AI-based software is applied to a range of tasks, including data visualization, segmentation, recording, classification, and interpretation.

Among these, medical image segmentation remains among the most challenging tasks, as it involves identifying clusters of pixels that correspond to specific image objects, particularly in CT and MRI scans. Deep learning algorithms have demonstrated promising performance in segmenting abnormal regions (selecting target regions) and subsequently classifying them. These algorithms notably outperform conventional approaches in both processing accuracy and speed [1]. Various neural network architectures have been employed for segmentation tasks. These models differ in structural characteristics, including the number of layers, neurons per layer, activation functions, and optimization algorithms. Among these architectures, frameworks such as U-Net, V-Net, DenseNet, and Mask R-CNN have demonstrated strong performance in segmentation tasks [2–6].

Since its introduction in 2015, the U-Net segmentation network has become a standard tool in biomedical image processing. Nevertheless, the basic U-Net architecture continues to demonstrate strong performance in analyzing medical images for detecting organ abnormalities, such as those seen in kidney CT scans and lung changes associated with COVID-19 or obstructive pulmonary disease [7–9]. The U-Net3D architecture extends the conventional U-Net by replacing two-dimensional (2D) convolutions with 3D convolutions [10]. It is employed for the segmentation of 3D medical images. For instance, Pantovic et al. used U-Net3D to analyze CT scans of a brain containing neural implants to identify surgical sites for epileptogenic zone removal [11]. Han et al. used the same architecture to segment liver MRI scans and delineate both the contours and internal structures of the liver [12].

The standard architecture of the U-Net neural network consists of two primary components: the encoder and decoder. The encoder compresses the input data and extracts the most relevant features for subsequent recognition.

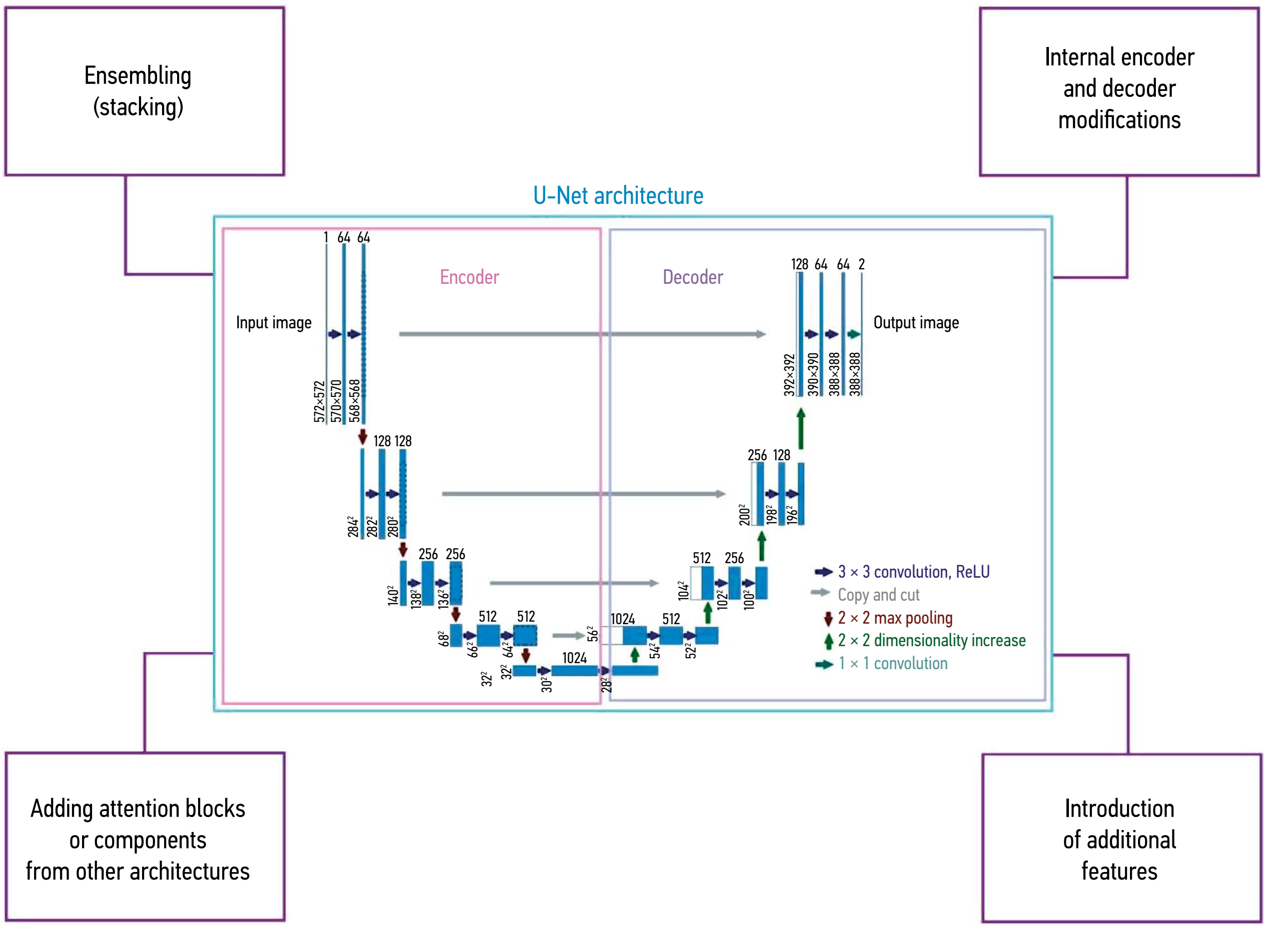

Fig. 1. Classic U-Net architecture proposed in 2015 and the main categories of its modification methods.

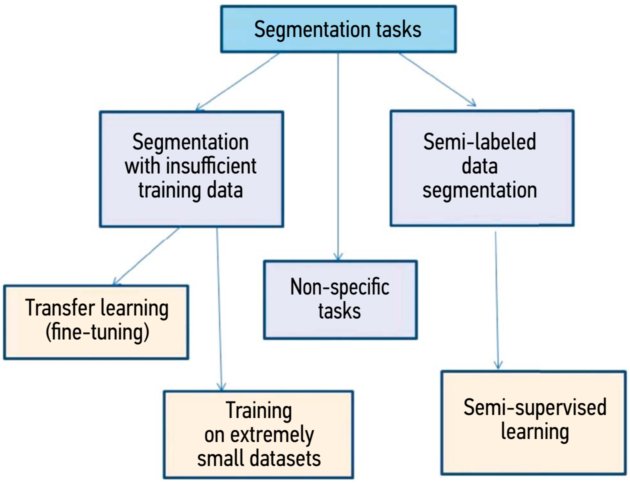

Fig. 2. Segmentation tasks categorized by the availability and type of training data.

Meanwhile, the decoder reconstructs a segmented image from the compressed data generated by the encoder. Since 2015, numerous modifications to the standard U-Net architecture (referred to as the U-architecture) (Fig. 1) have been developed to enhance its accuracy, speed, and robustness. These modifications can be grouped into four main categories: (1) modifying the encoder and decoder while preserving the overall network structure; (2) combining multiple U-architecture models through ensembling; (3) integrating additional architectural components, such as attention blocks; and (4) incorporating supplementary features into the model.

These modifications have also been applied to address image segmentation challenges arising during semi-supervised learning or when training data are limited (Fig. 2). The limited-training-data scenario can be further categorized into cases involving small and extremely small datasets. Specifically, when training on small datasets, transfer learning and fine-tuning are typically applied to networks pretrained on more diverse datasets.

Meanwhile, when training on extremely small datasets (few-shot learning), pretraining is inadequate; such cases generally require original architectures and data models.

This review explores the application of U-Net architecture modifications in medical image processing. Section 1 outlines the main modification strategies for the U-architecture, including (1) changing the encoder and decoder internally, (2) integrating additional architectural components such as attention blocks, and (3) altering the network's learning process. Section 2 explores how these modifications can be applied to address specific challenges in medical image segmentation. The conclusion in Section 3 summarizes the key findings of the review.

DATA SEARCH METHODOLOGY

The authors conducted a literature search using the Web of Science, Scopus, and PubMed databases, covering publications from 2018 to 2024. The search results were comparable across databases and reflected the primary trends in U-Net architecture modification methods. The search keywords included U-Net, medical images, and modification. The initial search returned approximately 5,000 sources. This was subsequently refined using additional terms, including attention, few-shot, unsupervised, semi-supervised, ensemble, stack, additional features, metadata, and DICOM data.

The selected publications were reviewed with a focus on the use of specific architectures for medical image processing. The inclusion criteria were as follows:

- Quality of result validation (e.g., comparison with other architectures, use of established evaluation metrics, and study completeness);
- Originality of the architectural modification in relation to its intended application;
- Specificity of the task (e.g., type of abnormality detected or organ segmented);
- Use of open datasets.

The U-Net architecture has substantially impacted medical image segmentation owing to its effectiveness. Originally proposed by Ronneberger et al. [2], U-Net has since evolved into several notable variants, including U-Net++ [13], Attention U-Net [5], 3D U-Net [10], EU-Net [14], NAS-U-Net [15], U-Net 3+ [16], and SwinAttU-Net [17]. Appendix 1 provides an overview of key studies on U-Net modification methods and segmentation accuracy evaluations, as well as datasets used for testing. It also includes studies wherein U-Net was applied to address specific segmentation challenges. The following abbreviations are used for performance metrics: DC, Dice coefficient; IoU, intersection over union; OA, overall accuracy [18, 19].

U-NET ARCHITECTURE MODIFICATIONS

Internal encoder and decoder modifications

This section discusses structural elements that are altered by internal modifications to the encoder and decoder of the U-architecture.

Encoder and decoder convolution blocks. To process spinal cord images (Verse2019 and Verse2020 datasets), Xu et al. replaced convolution layers with linear layers in the encoder and with octave convolutions in the decoder. Octave convolutions combine standard convolution blocks with pooling operations to extract frequency-based data [57]. Ayalew et al. reduced the number of convolution channels and incorporated batch normalization into the original U-architecture to detect liver tumors in CT scans [58]. This modification improved network accuracy on datasets with considerable class imbalance. Guan et al. proposed an architecture with modified convolution blocks wherein the outputs of each layer were concatenated and jointly processed to minimize distortion in photoacoustic images, such as brain scans [59].

Connections between encoder and decoder blocks. Özcan et al. used a U-Net variant to identify tumor regions in liver CT scans. In this variant, connections between encoder and decoder blocks passed through an inception block composed of convolutions with different kernel sizes, whose outputs were concatenated. In other studies, these connections passed through a pyramid of pooling layers (consisting of multiple pooling layers with different kernel sizes applied to the same data) [61]. This approach was used to accelerate the segmentation of liver ultrasound images.

Encoder or decoder regularization blocks. Omarov et al. applied a modified U-Net architecture to detect brain regions affected by ischemic stroke on CT scans. In this architecture, dropout and L2 regularization layers were incorporated into the decoder [62].

Ensembling U-Net architectures. A concatenated ensemble of U-Net networks trained on ImageNet images converted to sinograms was used to reconstruct CT images from projection data obtained by rotating an object [63]. In another example, an ensemble of two U-Net3D networks pretrained on the LiTS dataset was applied for detecting liver tumors in 3D CT scans [24, 64]. The first network processed low-resolution images (reduced source images), and its segmentation output was passed to the second network. A combined loss function incorporating the DC and cross-entropy was used. In another study, a two-stage U-Net ensemble was developed for liver tumor detection. Here, one network functioned as a post-processing and refinement stage [65].

Koirala et al. used an ensemble of U-Net3D, ONet3D, and SphereNet3D networks to locate brain tumors. Ensembling was achieved by weighing (summing and multiplying by a number reflecting the network's contribution to the overall result, i.e., its weight) the outputs of all models to determine the most probable class.

Li et al. used an unmodified U-Net architecture to select the optimal model for their application [67].

Overall, existing studies suggest that even minor architectural changes to U-Net can improve its effectiveness in medical imaging tasks.

Modifications using attention mechanisms

This section outlines how previous studies modified the standard U-Net architecture by integrating spatial and channel attention blocks [68]. In one study, a U-Net3D-based variant incorporating efficient channel attention in the encoder blocks was applied to detect COVID-19-related abnormalities in chest CT scans [69]. In another study, a pyramid fusion module was implemented at the lower layer of the U-architecture. In this module, features extracted using neural networks with varying window sizes were concatenated, and the resulting data were processed using a pooling layer with a global mean value. The Tversky loss function was used for optimization [70].

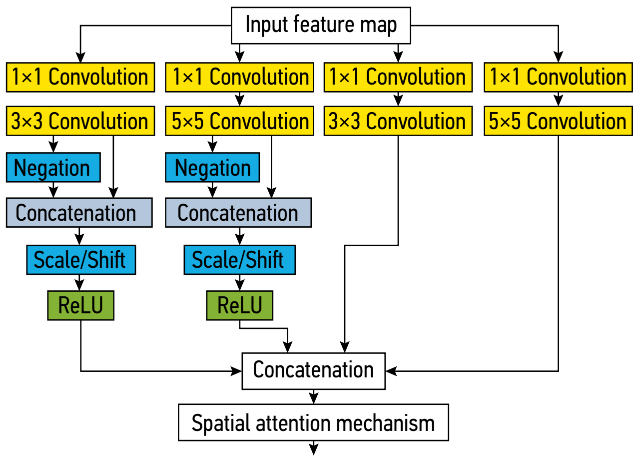

Fig. 3. Spatial attention block positioned between encoder elements [75].

Another study focused on the simultaneous segmentation of multiple organs using CT scans [71]. The proposed U-Net architecture included an attention block that took the outputs of both the encoder and decoder as input. These outputs were concatenated and processed using a 1D convolution operation with ReLU activation sigmoid functions [72].

One study [73] employed a U-Net architecture with spatial multi-scale attention blocks to segment liver tumors in CT scans. These attention blocks were placed at multiple points in the architecture, including within the encoder and decoder, as well as along the connections between them.

Zhang et al. applied pyramid pooling in the lower part of the U-architecture (corresponding to the point of maximum data reduction) and used efficient channel attention blocks on the connections between the encoder and decoder blocks [74]. Another study proposed a U-Net architecture with spatial attention between encoder blocks, incorporating convolutions with multiple receptive fields (Fig. 3). This architecture was trained using the Tversky loss function for breast cancer detection [75].

Subhan Akbar et al. introduced attention blocks into the connections between the encoder and decoder blocks of the U-architecture. For feature extraction, they also added a positional attention block and a self-attention block to each layer of the decoder [76, 77].

Thus, in U-Net, various attention blocks have been used to capture spatial relationships between image elements at different scales. Notably, these relationships cannot be detected using the basic architecture.

Modifications through the integration of elements from other architectures

A common approach to modifying the U-Net architecture involves incorporating elements from other networks, such as ResNet or transformers. Several variations of this approach have been proposed.

Full modification of the encoder and/or decoder. Xingfei et al. modified the U-Net architecture by replacing the encoder with ResNet50 for segmenting COVID-19-related abnormalities in the lungs [78, 79]. A channel attention block combined with a pyramid pooling module was applied following the encoder. Alternatively, a transformer encoder can be integrated, with its output upsampled via deconvolution for use in different parts of the U-architecture [80].

Modification of encoder and decoder blocks while maintaining the general U-Net architecture. Eskandari et al. focused on segmenting liver structures in CT scans [81]. To account for the considerable variability in liver shape, size, and position, they used a position-determining classifier network in combination with a modified U-Net architecture. This modification replaced standard convolution blocks with ConvLSTM blocks, which were also incorporated into the connections between encoder and decoder blocks [82].

In another study, a hybrid architecture combining efficient transformer blocks with the U-Net architecture was proposed for identifying skin abnormalities in medical images (Fig. 4) [83]. This architecture outperformed the classic U-Net, Attention U-Net, TransU-Net, FAT-Net, and Swin U-Net in terms of DC, sensitivity, specificity, and accuracy on the ISIC 2018 skin lesion dataset.

Ghofrani et al. applied a combination of an unmodified U-Net and transformer blocks to segment polyp images, achieving higher accuracy than U-Net, ResU-Net++, and DoubleU-Net [36, 37, 84–86].

For 3D liver image segmentation, U-Net was combined with Swin Transformer, BTSwin Transformer, and DenseNet components [87–89].

In summary, similar to attention-based modifications, integrating elements from other architectures enhances image processing quality by identifying subtle relationships between image regions. Transformer blocks that employ self-attention mechanisms to extract latent features are frequently used in this context.

Introducing additional features into u-net

Researchers often use metadata from DICOM files as supplementary features in medical image analysis. These data are typically tabulated and include both continuous and categorical variables. The metadata are often input into a separate network, which may be trained either jointly with or independently from the main segmentation model. This supplementary information is generally incorporated into the base network using attention mechanisms. For instance, in a previous study on spinal tumor segmentation, metadata were integrated into a U-Net-based segmentation model. Each block included a linear transformation block applied to the output of the preceding convolutional layer [90]. The U-Net-based generator computed transformation parameters (shift and scale) after receiving metadata related to the segmented image. In another study, Du et al. proposed a channel attention mechanism wherein metadata were used to train the 3D-RADNet network to detect image slices containing the target organ (liver) [91]. Slices selected using metadata were processed by a U-Net-based segmentation model. In kidney tumor segmentation, channel attention has been used to incorporate metadata into the network,

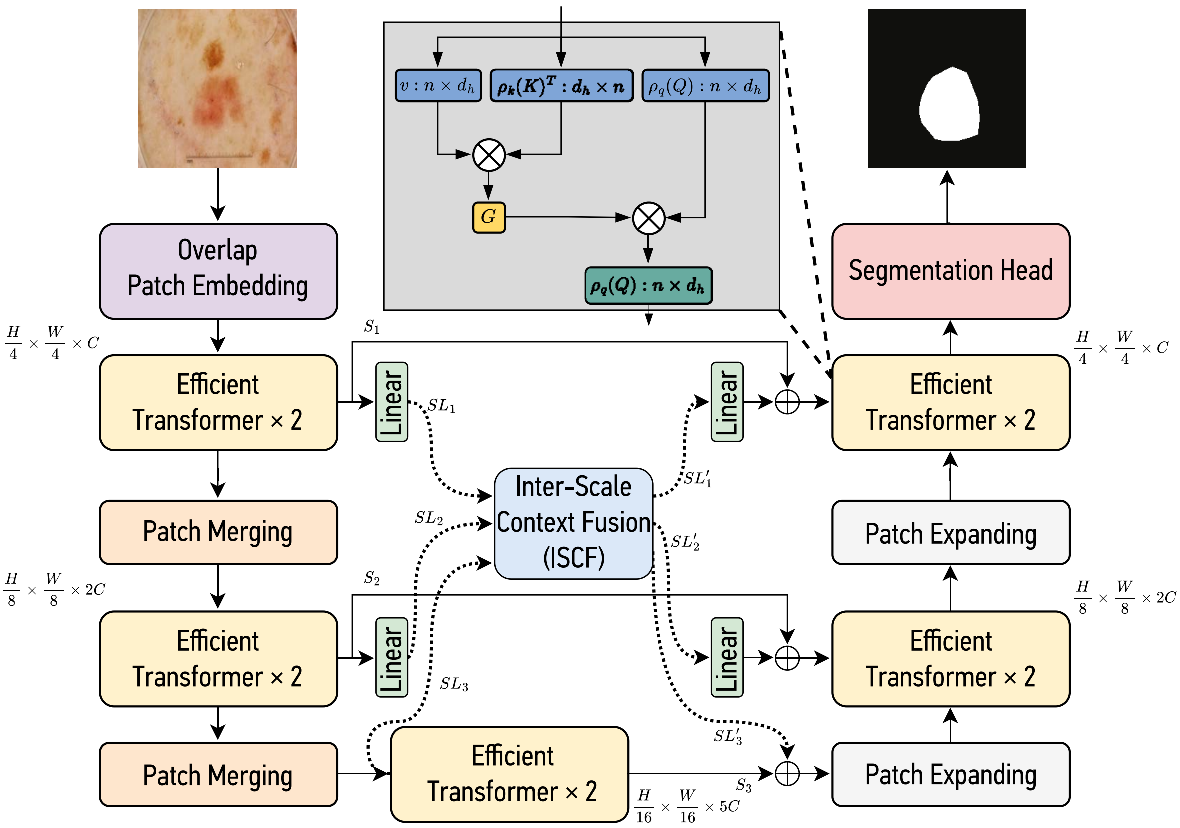

Fig. 4. Architecture integrating transformer blocks into the U-Net framework [81].

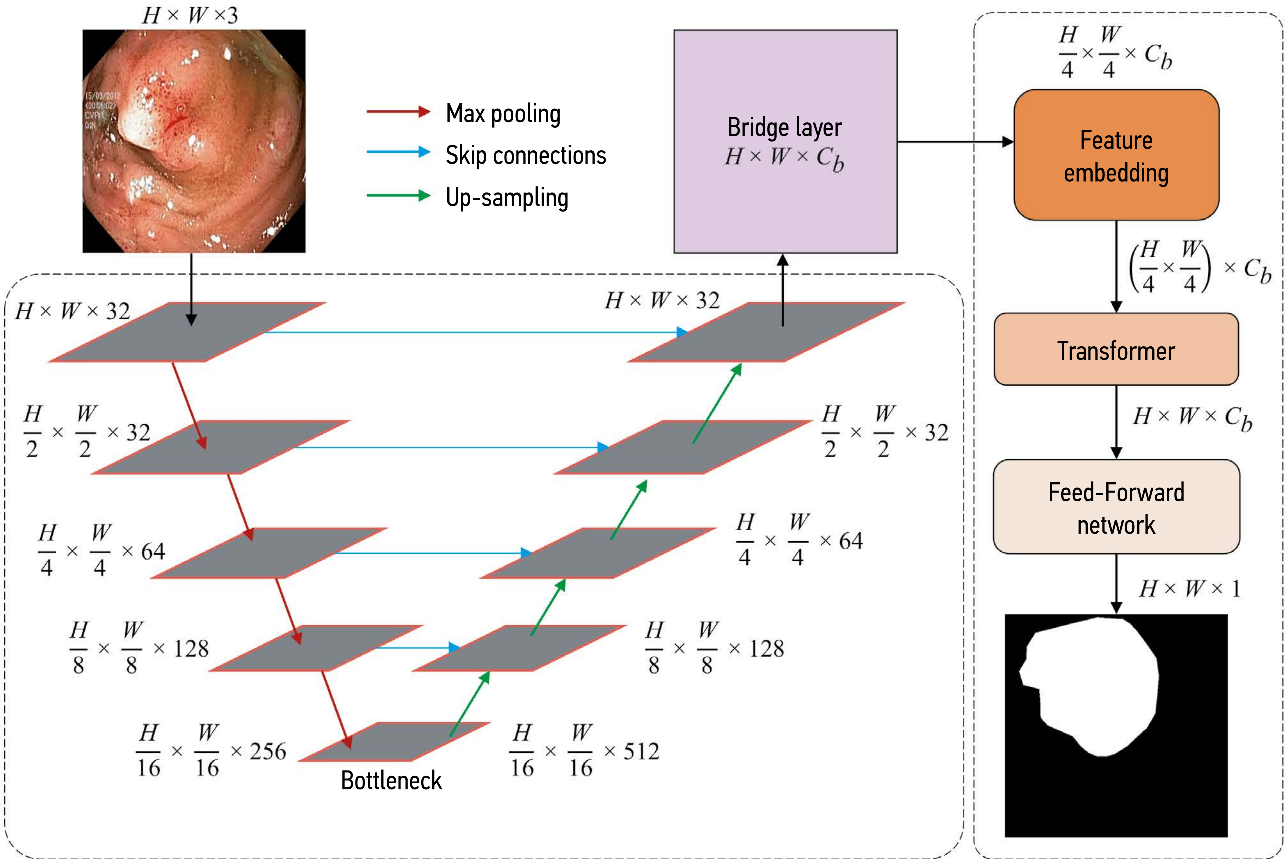

Fig. 5. Combined U-Net and transformer-based architecture [83].

allowing the data to serve as the outputs of U-Net blocks [92]. After the final convolutional layer of each block, both image data and metadata are passed to a layer where the metadata are input into a multi-layer perceptron (MLP) with a sigmoid activation function. The MLP outputs are then multiplied, on a per-channel basis, with the image data from the preceding convolutional layer.

In addition to metadata, other sources of auxiliary information have been used to enhance U-Net models:

- A two-branch architecture based on a convolutional network [93];
- CNNFormer for liver segmentation, which accounts for both intra-slice spatial relationships and inter-slice hierarchical structures [94];
- Additional features, such as spine, lung, and skin segmentation results obtained using the Python library Body Navigation [95].

These data have been concatenated with the input images to enhance the localization of the target organ. This approach has been applied to liver CT segmentation using both U-Net and U-Net3D architectures, depending on whether individual slices or entire scans were processed.

Many modifications to U-Net training involve the iterative reuse of features. For example, Ernst et al. focused on reconstructing CT images from sinograms [96]. They employed a combination of U-Net3D and Primal-Dual networks with iterative learning, where the output at each step was combined with the results of the previous iteration. Another study proposed a method to improve segmentation accuracy by reusing features extracted during learning [97]. RecycleNet, an architecture derived from U-Net, comprises three main blocks:

- I: input data block;
- R: latent feature reuse block;
- O: outcome block (Fig. 6).

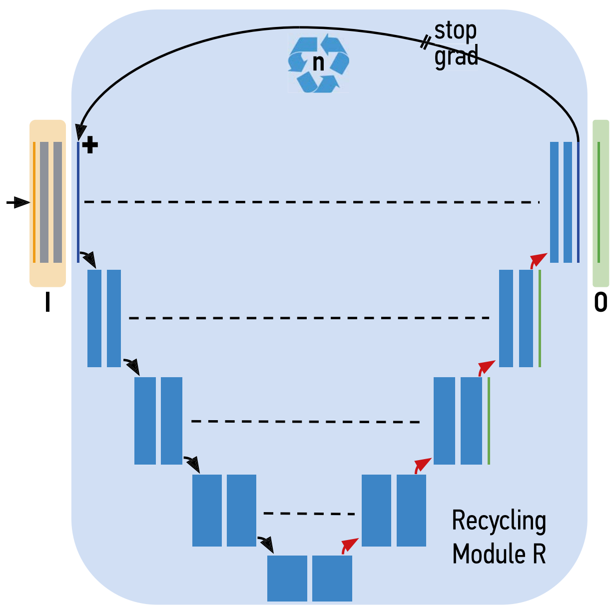

Fig. 6. Structural blocks of the U-Net architecture [97].

DOI: https://doi.org/10.17816/DD629866

The feature reuse algorithm is illustrated in Fig. 6. First, the number of iterations to be used for decision-making is randomly selected from a predefined range. The features extracted in the previous iteration are normalized and added to those from the current iteration, incorporating spatial embedding. After completing the selected number of iterations, the network generates the final output. In a previous study, RecycleNet was experimentally evaluated on the KiTS 2019 (kidney cancer), LiTS, BTCV, AMOS (multi-organ segmentation), and CHAOS (MRI) datasets [23, 24, 33, 40]. The proposed architecture was compared with a DC-optimized variant of nnU-Net and the DRU network [98]. RecycleNet outperformed the compared architectures on all evaluated datasets.

Thus, incorporating additional features can improve the accuracy of image processing using U-Net. Such supplementary data often reveal patterns that are not present or are only weakly expressed in the image itself.

ADDRESSING SPECIFIC SEGMENTATION CHALLENGES USING THE U-NET ARCHITECTURE

Transfer learning and fine-tuning of U-Net

In medical image processing, available training datasets are often small and structurally complex. This limitation arises from the difficulties encountered during data labeling and restrictions imposed by privacy agreements. A common approach in such cases is to employ pretrained models and fine-tune them on the available datasets.

Heker et al. investigated liver tumor segmentation using a small dataset of CT scans [99]. To this end, they first trained the U-Net architecture on the LiTS dataset and applied a hierarchical freezing strategy to its encoder weights. Initially, the encoder weights were frozen, meaning they were not updated during training. The rest of the network was trained for a set number of iterations. Afterward, the frozen encoder weights were gradually unfrozen and fine-tuned one by one.

Several researchers employed a U-Net architecture with a ResNet32-based encoder, initially pretrained on ImageNet and subsequently fine-tuned using optical coherence tomography images [100]. Meanwhile, others have explored fine-tuning techniques for U-Net and U-Net3D in the segmentation of various organs and diseases, including approaches involving a variable number of trainable layers [101, 102].

Moreover, transfer learning with U-Net and EfficientNet architectures—both originally developed for 2D image segmentation—has been applied to facilitate data transfer during 3D image processing [103, 104]. The authors of the aforementioned paper proposed two approaches: 1) increasing the sampling rate of 2D weights in the corresponding blocks of 3D architectures and 2) obtaining plane projections of 3D data and subsequently processing them using a network trained on 2D data (Fig. 7).

Another approach to training involves using U-Net for post-processing image segmentation results. Hong et al. applied this strategy for liver segmentation in CT scans. In their proposed modification, U-Net's segmentation output underwent post-processing through the optimization of an energy functional. This functional included two components: one for contour delineation in an image and another for optimizing voxel class labels within the evaluated region.

The effectiveness of fine-tuning and transfer learning strategies strongly depends on the datasets used during pretraining. The closer the training and target datasets are in terms of the types of objects assessed, the more effective fine-tuning and transfer learning become. However, achieving this similarity is not always feasible, particularly for specialized tasks. Large datasets are often unavailable—especially for 3D data. A promising alternative is to fine-tune using simpler, lower-dimensional data, which are generally easier to collect in sufficient quantities.

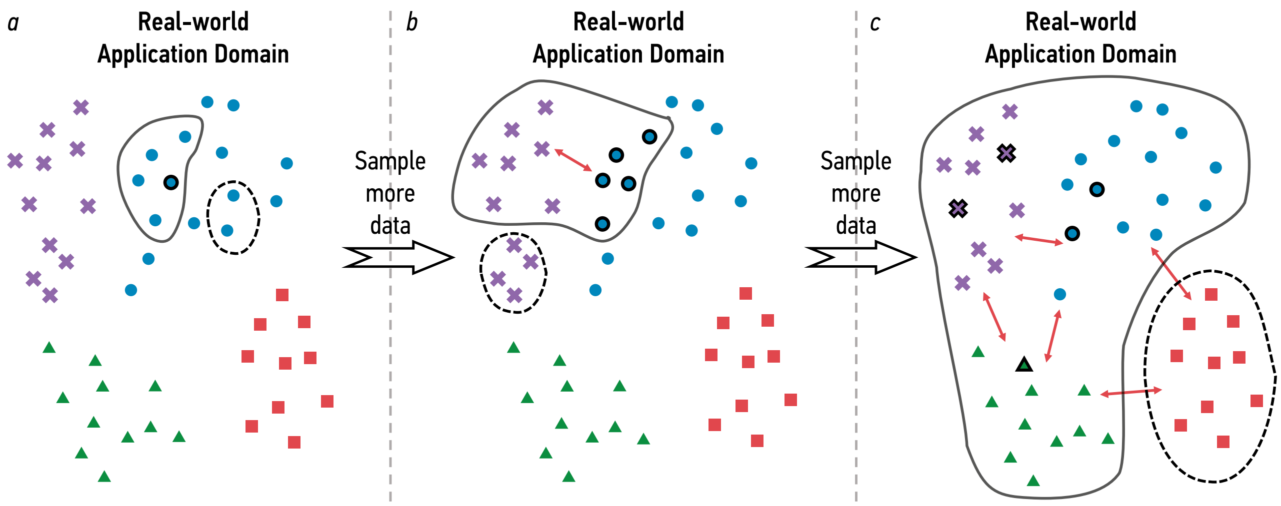

Fig. 7. Ratios of labeled and unlabeled data in network training and testing: (*a*) semi-supervised learning (SSL), (*b*) unsupervised domain adaptation (UDA), and (*c*) semi-supervised domain generalization (SemiDG) [106].

Semi-supervised learning methods

The shortage of sufficient training data for complex architectures is often due to the lack of expert annotation of raw data—a task that requires substantial domain-specific knowledge and expertise. To address this limitation, various training strategies based on the U-Net architecture have been developed to leverage unlabeled data and semi-supervised learning approaches.

Wang et al. explored the training of segmentation networks for 3D organ models using semi-supervised learning techniques [106]. They developed a framework capable of handling different proportions of labeled and unlabeled data during both training and testing phases (See Fig. 7):

- Fig. 7(*a*): labeled and unlabeled data, as well as testing data, are of the same type (testing data indicated with a dotted line);
- Fig. 7(*b*): labeled and unlabeled data are of different types;
- Fig. 7(*c*): the training set contains labeled and unlabeled data of different types, while the testing data are entirely distinct from both.

The resulting framework consists of two main components (Fig. 8): an aggregation block and a decoupling block. The aggregation block includes the encoder of the proposed Diffusion VNet, which performs image segmentation for type 1 relationships. The decoupling block contains three VNet decoders, each responsible for generating class labels of a specific type. The first decoder produces labels that are unbiased with respect to the type of labeled data, using a loss function that combines cross-entropy and DC. These labels are then used to generate re-weighted class labels, where the weights are applied in a loss function consisting of the sum of DCs across all labeled data classes. This weighting strategy enhances the training effectiveness for classes that perform poorly. The second decoder generates class pseudo-labels for unlabeled data, which are subsequently used to train the third decoder in an unsupervised manner.

In a previous study, the above framework was trained using the LASeg (brain MRI), Synapse (various organs), MMWHS, and M&Ms (heart) datasets [47–50]. Its performance was evaluated against that of UA-MT, LMISA-3D, vMFNet, SS-Net, and other architectures using metrics such as DC, Jaccard index, and HD95. In several cases, the framework demonstrated performance that was either superior to or comparable with that of specialized architectures.

Wang et al. investigated trained network adaptation for segmenting a small target dataset focused on polyp detection [107]. The study evaluated a scenario wherein the target dataset consisted of images similar to those used for network training but lacked labels. Two techniques were applied for training: contrastive learning and pseudo-labeling with calibration.

In the contrastive learning phase, unlabeled images were labeled as either positive (consistent with a given image) or negative. Images obtained through augmentation were treated as positive, while others were treated as negative. A network trained on a different dataset generated pseudo-masks for the target dataset. These predicted masks were then used to calculate entropy and determine class centers within the target scans.

To improve the reliability of the generated pseudo-masks, a per-pixel calibration block was introduced. This block incorporated previous predictions to refine the mask quality. To evaluate the effectiveness of the proposed method in polyp segmentation, experiments were conducted using the ClinicDB, ETIS-LARIB, and Kvasir-SEG datasets. The proposed

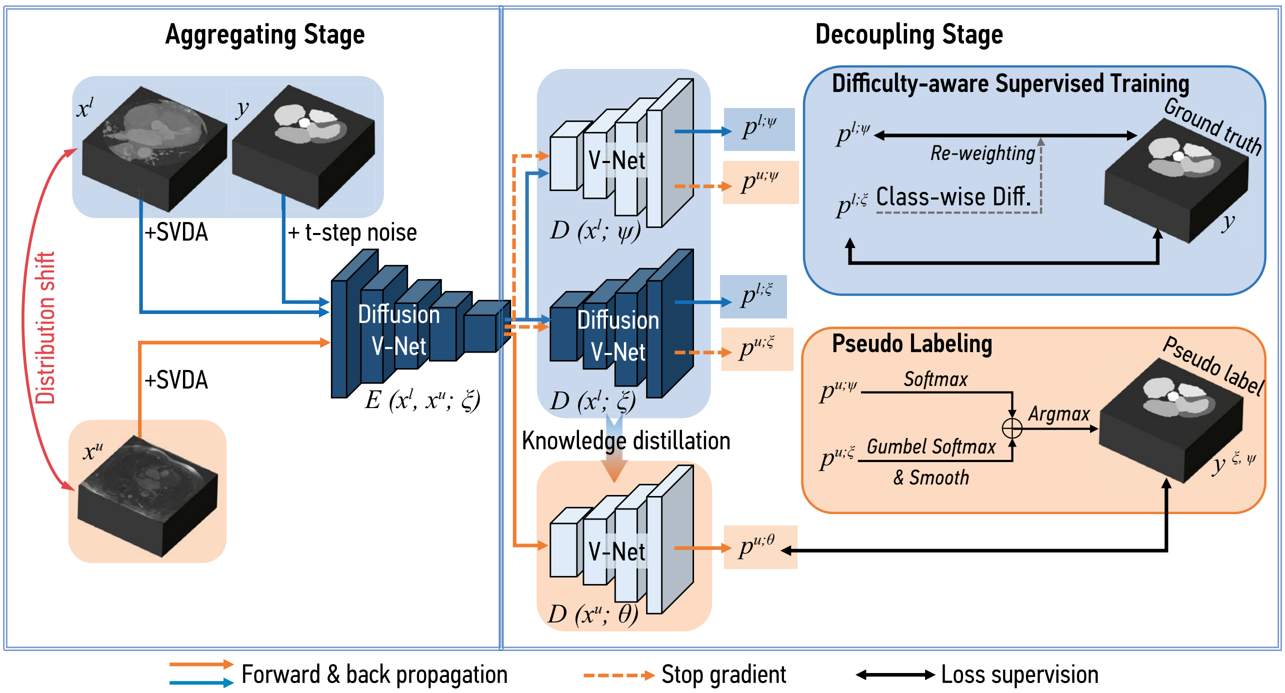

Fig. 8. A&D framework [106].

architecture was compared with other networks employing techniques such as bidirectional learning (BDL), Fourier domain adaptation, historical contrastive learning, and denoised pseudo-labeling. The proposed architecture outperformed these alternatives in terms of DC and IoU variations.

Wang et al. also proposed a method for segmenting human organ images, including those captured during surgery, using semi-labeled datasets.

For unlabeled data processing, a dual-network configuration was used (Fig. 9), in which two networks with identical architectures received the same image input. Although the networks were initialized differently, aggregating their outputs enabled more accurate predictions than either network could achieve independently. To avoid distortion when assigning pseudo-labels to unlabeled data in cases where the training dataset exhibited heterogeneous class representations, individual class distributions were reconstructed rather than relying on the overall data distribution.

To align individual class densities, an exponential moving average transformation was applied to class alignment matrices of both labeled and unlabeled data. The effectiveness of the proposed method was evaluated using the CaDIS (surgical images), LGE-MRI, and ACDC (heart disease) datasets. Its performance was compared with that of the URPC, UAMT, CLD, and CPS architectures using the DC, Jaccard index, and additional metrics. The proposed method outperformed all of these architectures across the evaluated parameters.

Thus, a properly selected architecture enables the effective use of unlabeled data in training U-Net-based models, even in the presence of class imbalance.

U-Net training using extremely small sets of real-world data

Developing AI-based software for specific medical tasks is hindered by the challenge of assembling a sufficiently large training dataset [109]. In many cases, dedicated tools are required to process and structure text-based protocols [110–112]. Combined with the high cost of data annotation, these challenges frequently force developers to work with limited amounts of labeled data for machine learning. Consequently, few-shot learning has become a widely adopted approach in medical image processing.

A study investigated the use of CT and positron emission tomography scans for lung cancer detection [113]. A standard U-Net architecture without modifications was trained using data augmentation, with additional data introduced during both training and testing phases based on feedback from an expert evaluating the model's performance. A similar approach was later applied to COVID-19 data [114]. In another study, the encoder of the U-Net architecture was modified using a Siamese-Net-type structure to enhance segmentation quality. A second encoder branch was introduced; it received the image multiplied by its corresponding mask (segment). The weights from this branch were then combined with those of the primary encoder branch, which processed the original, unmodified image [115].

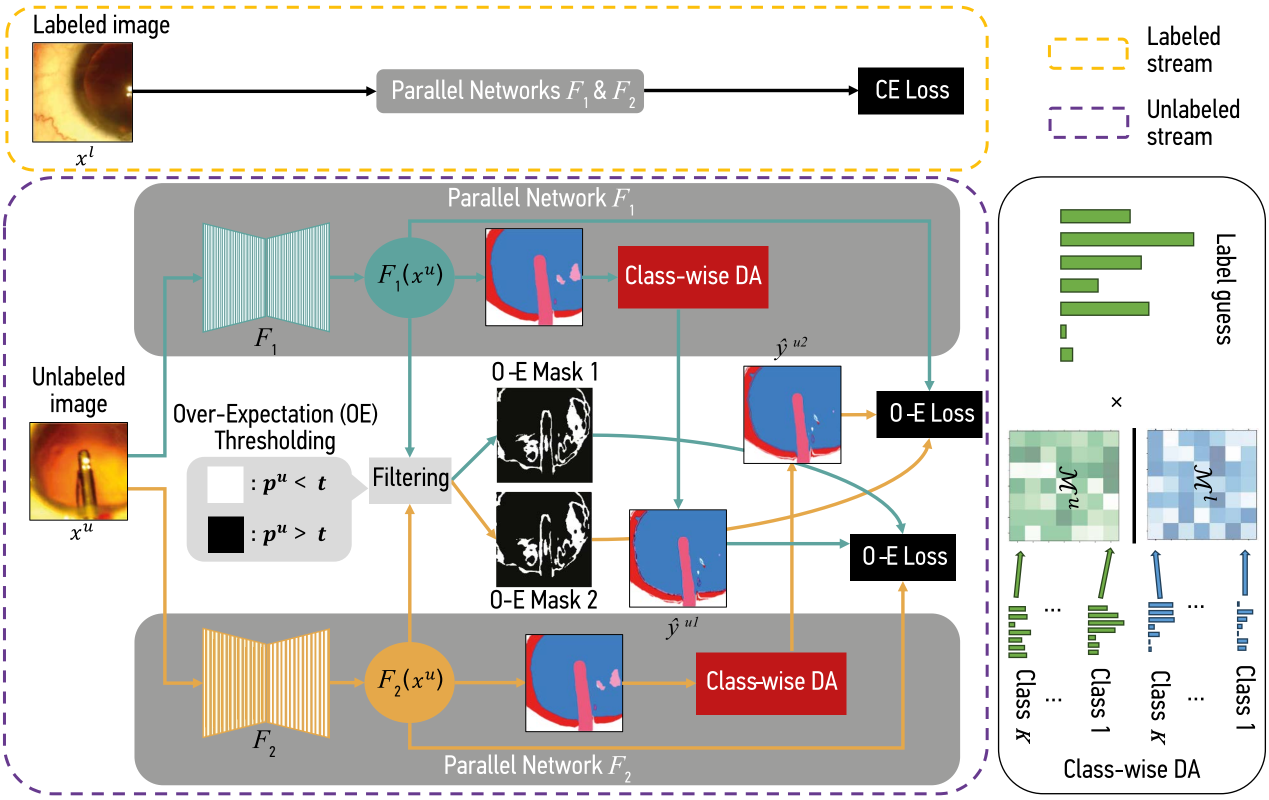

Fig. 9. Dual-network architecture trained on datasets with heterogeneous class representation [108].

In the context of medical imaging, this approach is more frequently applied to architectures other than U-Net, which may be due to the network's size and the number of neurons it contains.

CONCLUSION

The classic U-Net architecture has proven highly effective for medical image segmentation, which explains its widespread use and the ongoing development of various performance-enhancing modifications. These modifications are designed to improve the interpretation of available data and to pool features obtained during pretraining on diverse datasets, including those that are unlabeled. U-Net modifications can also be categorized according to their intended tasks—such as segmentation or the detection of affected tissues—as well as by the types of datasets used, particularly those representing specific diseases. Additionally, the diagnostic accuracy of U-Net-based solutions can be further enhanced by incorporating supplementary training features derived from text, tabular data, or mathematical models.

U-Net architectures are applied across a wide range of medical image segmentation tasks, which vary in both problem formulation and data type (various types of images and diseases). Each task presents its own unique challenges, making it difficult to define a single, universally effective architecture or even a universally applicable class of models. However, among the approaches assessed, U-Net modifications incorporating elements from other architectures demonstrate the strongest performance. These hybrid models are effective for standard image segmentation tasks—particularly when integrating transformer blocks—as well as for situations where training data are limited, such as through pretraining with networks of lower dimensionality than the target data. The integration of additional features into neural network architectures also shows promise. Similarly, the application of physics-informed neural networks, which incorporate information on object models or image structure, is another promising direction [116–119].

ADDITIONAL INFORMATION

Appendix 1. Ways to modify the U-Net architecture. doi: 10.17816/DD629866-4224037

Funding source. This article was prepared by a group of authors as a part of the research and development effort titled «Development of a platform for improving the quality of AI services for clinical diagnostics» (USIS No.: 123031400006-0) in accordance with the Order No. 1196 dated December 21, 2022 «On approval of state assignments funded by means of allocations from the budget of the city of Moscow to the state budgetary (autonomous) institutions subordinate to the Moscow Health Care Department, for 2023 and the planned period of 2024 and 2025» issued by the Moscow Health Care Department. The research was carried out using the infrastructure of the federal state budgetary educational institution of higher education «MIREA – Russian Technological University» within the framework of additional agreement No. 1 dated November 24, 2023 to the cooperation agreement No. 1 dated 07.07.2022, (Moscow).

Competing interests. The authors declare that they have no competing interests. Figures 1 and 2 are original and made by the authors. Figures 3-9 are distributed under the CC BY 4.0 license and are presented in this work unchanged with reference to the original works where they were first presented.

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work. A.M. Dostovalova — collection and processing of materials, writing the text of the article; A.K. Gorshenin — problem statement, analysis and systematization of approaches, conceptualization, writing the text of the article; Ju.V. Starichkova, K.M. Arzamasov — concept of the work, writing the text of the article.

REFERENCES

1. Shen D, Wu G, Suk HI. Deep Learning in Medical Image Analysis. *Annual Review of Biomedical Engineering.* 2017;19:221–248. doi: 10.1146/annurev-bioeng-071516-044442

2. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015.* 2015:9351. doi: 10.1007/978-3-319-24574-4_28

3. Milletari F, Navab N, Ahmadi SA. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *Fourth International Conference on 3D Vision (3DV).* 2016:565–571. doi: 10.48550/arXiv.1606.04797

4. Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille A. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2017;40(4):834–848. doi: 10.1109/TPAMI.2017.2699184

5. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 2017:2261–2269. doi: 10.1109/CVPR.2017.243

6. He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. *IEEE International Conference on Computer Vision (ICCV).* 2017:2980–2988. doi: 10.1109/ICCV.2017.322

7. Khalal DM., Azizi H, Maalej N. Automatic segmentation of kidneys in computed tomography images using U-Net. *Cancer/Radiothérapie.* 2023;27(2):109–114. doi: 10.1016/j.canrad.2022.08.004

8. Bernardo Gois FN, Lobo Marques JA. Segmentation of CT-Scan Images Using UNet Network for Patients Diagnosed with COVID-19. *Computerized Systems for Diagnosis and Treatment of COVID-192023.* 2023:29–44. doi: 10.1007/978-3-031-30788-1_3

9. Sarsembayeva T, Shomanov A, Sarsembayev M, et al. UNet Model for Segmentation of COPD Lung Lesions on Computed Tomography Images. *Proceedings of the 7th International Conference* *on Digital Technologies in Education, Science and Industry (DTESI 2022).* 2022. Available at: https://ceur-ws.org/Vol-3382/Short5.pdf. Accessed: November 9, 2024.

10. Çiçek Ö, Abdulkadir A, Lienkamp S, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2016.* 2016:424–432. doi: 10.1007/978-3-319-46723-8_4

11. Pantovic A, Ollivier I, Essert C. 2D and 3D-UNet for segmentation of SEEG electrode contacts on post-operative CT scans. *Medical Imaging 2022: Image-Guided Procedures, Robotic Interventions, and Modeling.* 2022. doi: 10.1117/12.2606538

12. Han X, Wu X, Wang S, et al. Automated segmentation of liver segment on portal venous phase MR images using a 3D convolutional neural network. *Insights Imaging.* 2022;13(26). doi: 10.1186/s13244-022-01163-1

13. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support.* 2018:3–11. doi: 10.1007/978-3-030-00889-5_1

14. Yu C, Wang Y, Tang C, Feng W, Lv J. EU-Net: Automatic U-Net neural architecture search with differential evolutionary algorithm for medical image segmentation. *Computers in Biology and Medicine.* 2023;167:107579. doi: 10.1016/j.compbiomed.2023.107579

15. Weng Y, Zhou T, Li Y, Qiu X. NAS-Unet: Neural Architecture Search for Medical Image Segmentation. *IEEE Access.* 2019;7:44247–44257. doi: 10.1109/ACCESS.2019.2908991

16. Huang H, Lin L, Tong R, et al. UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation. *ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* 2020:1055–1059. doi: 10.1109/ICASSP40776.2020.9053405

17. Li C, Bagher-Ebadian H, Sultan RI, et al. A new architecture combining convolutional and transformer-based networks for automatic 3D multi-organ segmentation on CT images. *Med Phys.* 2023;50(11):6990–7002. doi: 10.1002/mp.16750

18. Müller D, Soto-Rey I, Kramer F. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes.* 2022;15(210). doi: 10.1186/s13104-022-06096-y

19. Alberg AJ, Park JW, Hager BW, Brock MV, Diener-West M. The use of «overall accuracy» to evaluate the validity of screening or diagnostic tests. *Journal of General Internal Medicine.* 2004;19:460–465. doi: 10.1111/j.1525-1497.2004.30091.x

20. Soler L, Hostettler A, Agnus V, et al. 3D image reconstruction for comparison of algorithm database: A patient specific anatomical and medical image database. *IRCAD.* 2010. Available at: https://www-sop.inria.fr/geometrica/events/wam/abstract-ircad.pdf. Accessed: November 9, 2024.

21. Löffler M, Sekuboyina A, Jakob A, et al. A Vertebral Segmentation Dataset with Fracture Grading. *Radiology: Artificial Intelligence.* 2020;2(4). doi: 10.1148/ryai.2020190138

22. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing.* 2004;13(4):600–612. doi: 10.1109/TIP.2003.819861

23. Kavur AE, Gezer NS, Barıs M, et al. CHAOS Challenge – combined (CT-MR) healthy abdominal organ segmentation. *Medical Image Analysis.* 2021;69:101950. doi: 10.1016/j.media.2020.101950

24. Bilic P, Christ P, Li HB, et al. The Liver Tumor Segmentation Benchmark (LiTS). *Medical Image Analysis.* 2023;84:102680. doi: 10.1016/j.media.2022.102680

25. Petrusca L, Cattin P, De Luca V, et al. Hybrid ultrasound/magnetic resonance simultaneous acquisition and image fusion for motion monitoring in the upper abdomen. *Investigative Radiology.* 2013;48(5):333–340. doi: 10.1097/RLI.0b013e31828236c3

26. Jun M, Cheng G, Yixin W, et al. Covid-19 CT lung and infection segmentation dataset. *Zenodo.* 2020. Available at: https://zenodo.org/records/3757476#.YLov8vkzaUk. Accessed: November 9, 2024.

27. Morozov SP, Andreychenko AE, Blokhin IA, et al. MosMedData: data set of 1110 chest CT scans performed during the COVID-19 epidemic. *Digital Diagnostics.* 2020;1(1):49–59. doi: 10.17816/DD46826

28. Roth HR, Oda H, Hayashi Y, et al. Hierarchical 3D fully convolutional networks for multi-organ segmentation. *ArXiv.* 2017. Available at: https://arxiv.org/abs/1704.06382v1. Accessed: November 9, 2024.

29. Roth H, Farag A, Turkbey EB, et al. Data from Pancreas-CT. Data From Pancreas-CT (Version 2) [Data set]. *The Cancer Imaging Archive.* 2016. doi: 10.7937/K9/TCIA.2016.tNB1kqBU

30. Heimann T, Styner M, van Ginneken B. 3D Segmentation in the Clinic: A Grand Challenge. *MICCAI 2007, the 10th Intel Conf. on Medical Image Computing and Computer Assisted Intervention.* 2007:7–15. Available at: https://www.diagnijmegen.nl/publications/ginn07/. Accessed: November 9, 2024.

31. Suckling J. The Mammographic Image Analysis Society Digital Mammogram Database. *International Congress Series.* 1994:375–378. Available at: http://peipa.essex.ac.uk/info/mias.html. Accessed: November 9, 2024.

32. WHO Director-General's opening remarks at the media briefing on COVID-19 — 11 March 2020 [Internet]. 2020. Available at: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020. Accessed: November 9, 2024.

33. Landman B, Xu Z, Igelsias J, et al. Miccai multi-atlas labeling beyond the cranial vault–workshop and challenge. *Proceedings of the MICCAI Multi-Atlas Labeling Beyond Cranial Vault — Workshop Challenge.* 2015;5:12.

34. Simpson AL, Antonelli M, Bakas S, et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *ArXiv.* 2019. doi: 10.48550/arXiv.1902.09063

35. Gutman D, Codella NCF, Celebi E, et al. Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC). *ArXiv.* 2016. doi: 10.48550/arXiv.1605.01397

36. Jha D, Smedsrud PH, Riegler MA, et al. Kvasir-SEG: A Segmented Polyp Dataset. *MultiMedia Modeling.* 2020;11962:451–462. doi: 10.1007/978-3-030-37734-2_37

37. Bernal J, Sánchez FJ, Fernández-Esparrach G, et al. WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Computerized Medical Imaging and Graphics.* 2015;43:99–111. doi: 10.1016/j.compmedimag.2015.02.007

38. Grove O, Berglund AE, Schabath MB, et al. Quantitative Computed Tomographic Descriptors Associate Tumor Shape Complexity and Intratumor Heterogeneity with Prognosis in Lung Adenocarcinoma. *PLOS ONE.* 2015;10(3):e0118261. doi: 10.1371/journal.pone.0118261

39. Heller N, Sathianathen N, Kalapara A, et al. The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes. *ArXiv.* 2019:13. doi: 10.48550/arXiv.1904.00445

40. Ji Y, Bai H, Yang J, et al. AMOS: A Large-Scale Abdominal Multi-Organ Benchmark for Versatile Medical Image Segmentation. *ArXiv.* 2022. doi: 10.48550/arXiv.2206.08023

41. Lemay A, Gros C, Zhuo Z, et al. Multiclass Spinal Cord Tumor Segmentation on MRI with Deep Learning. *ArXiv.* 2021. doi: 10.48550/arXiv.2012.12820

42. Ali MAS, Misko O, Salumaa SO, et al. Evaluating Very Deep Convolutional Neural Networks for Nucleus Segmentation from Brightfield Cell Microscopy Images. *SLAS Discovery.* 2021;26(9):1125–1137. doi: 10.1177/24725552211023214

43. Gibson E, Giganti F, Hu Y, et al. Automatic Multi-Organ Segmentation on Abdominal CT With Dense V-Networks. *IEEE Transactions on Medical Imaging.* 2018;37(8):1822–1834. doi: 10.1109/TMI.2018.2806309

44. Jimenez-del Toro O, Müller H, Krenn M, et al. Cloud-Based Evaluation of Anatomical Structure Segmentation and Landmark Detection Algorithms: VISCERAL Anatomy Benchmarks. *IEEE Transactions on Medical Imaging.* 2016;35(11):2459–2475. doi: 10.1109/TMI.2016.2578680

45. Regan EA, Hokanson JE., Murphy JR, et al. Genetic Epidemiology of COPD (COPDGene) Study Design. *COPD: Journal of Chronic Obstructive Pulmonary Disease.* 2010;7(1):32–43. doi: 10.3109/15412550903499522

46. Litjens G, Toth R, van de Ven W, et al. Evaluation of prostate segmentation algorithms for MRI: The PROMISE12 challenge. *Medical Image Analysis.* 2014;18(2):359–373. doi: 10.1016/j.media.2013.12.002

47. Xiong Z, Xia Q, Hu Z, et al. A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging. *Medical Image Analysis.* 2021;67:101832. doi: 10.1016/j.media.2020.101832

48. Landman B, Xu Z, Igelsias J, et al. 2015 MICCAI multi-atlas labeling beyond the cranial vault–workshop and challenge. *MICCAI Multi-Atlas Labeling Beyond Cranial Vault — Workshop Challenge.* 2015;5:12.

49. Zhuang X, Shen J. Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI. *Medical Image Analysis.* 2016;31:77–87. doi: 10.1016/j.media.2016.02.006

50. Campello VM, Gkontra P, Izquierdo C, et al. Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge. *IEEE Transactions on Medical Imaging.* 2021;40(12):3543–3554. doi: 10.1109/TMI.2021.3090082

51. Silva J, Histace A, Romain O, Dray X, Granado B. Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer. *International Journal of Computer Assisted Radiology and Surgery.* 2014;9:283–293. doi: 10.1007/s11548-013-0926-3

52. Trikha S, Turnbull A, Morris R, Anderson D, Hossain P. The journey to femtosecond laser-assisted cataract surgery: New beginnings or a false dawn? *Eye.* 2013;27(4):461–473. doi: 10.1038/eye.2012.293

53. Xiong Z, Xia Q, Hu Z, et al. A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging. *Medical Image Analisys.* 2021;67:101832. doi: 10.1016/j.media.2020.101832

54. Bernard O, Lalande A, Zotti C, et al. Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved? *IEEE Transactions on Medical Imaging.* 2018;37(11):2514–2525. doi: 10.1109/TMI.2018.2837502

55. Li P, Wang S, Li T, et al. A Large-Scale CT and PET/CT Dataset for Lung Cancer Diagnosis (Lung-PET-CT-Dx) [Data set]. *The Cancer Imaging Archive.* 2020. doi: 10.7937/TCIA.2020.NNC2-0461

56. Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *Journal of Digital Imaging.* 2013;26:1045–1057. doi: 10.1007/s10278-013-9622-7

57. Xu Z, Jia Z, Sun J, Dong W, Li Z. DO-U-Net: Improved U-Net Model for CT Image Segmentation using DBB and Octave Convolution. *Proceedings of the 2023 International Conference on Computer, Vision and Intelligent Technology (ICCVIT '23).* 2023:1–8. doi: 10.1145/3627341.3630403

58. Ayalew Y, Fante K, Aliy M. Modified U-Net for liver cancer segmentation from computed tomography images with a new class balancing method. *BMC Biomedical Engineering.* 2021;3(4). doi: 10.1186/s42490-021-00050-y

59. Guan S, Khan AA, Sikdar S, Chitnis PV. Fully Dense UNet for 2-D Sparse Photoacoustic Tomography Artifact Removal. *IEEE Journal of Biomedical and Health Informatics.* 2020;24(2):568–576. doi: 10.1109/JBHI.2019.2912935

60. Özcan F, Uçan ON, Karaçam S, Tunçman D. Fully Automatic Liver and Tumor Segmentation from CT Image Using an AIM-UNet. *Bioengineering.* 2023;10(2). doi: 10.3390/bioengineering10020215

61. Ansari MY, Yang Y, Meher PK, Dakua SP. Dense-PSP-UNet: A neural network for fast inference liver ultrasound segmentation. *Computers in Biology and Medicine.* 2023;153:106478. doi: 10.1016/j.compbiomed.2022.106478

62. Omarov B, Tursynova A, Postolache O, et al. Modified UNet Model for Brain Stroke Lesion Segmentation on Computed Tomography Images. *Computers, Materials and Continua.* 2022;71(3):4701–4717. doi: 10.32604/cmc.2022.020998

63. Mizusawa S, Sei Y, Orihara R, Ohsuga A. Computed tomography image reconstruction using stacked U-Net. *Computerized Medical Imaging and Graphics.* 2021;90:101920. doi: 10.1016/j.compmedimag.2021.101920

64. Golts A, Khapun D, Shats D, Shoshan Y, Gilboa-Solomon F. An Ensemble of 3D U-Net Based Models for Segmentation of Kidney and Masses in CT Scans. *Kidney and Kidney Tumor Segmentation (KiTS 2021).* 2022;13168:103–115. doi: 10.1007/978-3-030-98385-7_14

65. Araújo JDL, da Cruz LB, Diniz JOB, et al. Liver segmentation from computed tomography images using cascade deep learning. *Computers in Biology and Medicine.* 2022;140:105095. doi: 10.1016/j.compbiomed.2021.105095

66. Koirala CP, Mohapatra S, Gosai A, Schlaug G. Automated Ensemble-Based Segmentation of Adult Brain Tumors: A Novel Approach Using the BraTS AFRICA Challenge Data. *ArXiv.* 2023. doi: 10.48550/arXiv.2308.07214

67. Li Z, Zhu Q, Zhang L, et al. A deep learning-based self-adapting ensemble method for segmentation in gynecological brachytherapy. *Radiation Oncology.* 2022;17(152). doi: 10.1186/s13014-022-02121-3

68. Woo S, Park J, Lee J-Y, Kweon IS. CBAM: Convolutional Block Attention Module. *Proceedings of the European conference on computer vision (ECCV).* 2018:3–19. doi: 10.48550/arXiv.1807.06521

69. Nazir S, Zheng R, Zheng Y, Dong-Ye C. Improved 3D U-Net for COVID-19 Chest CT Image Segmentation. *Scientific Programming.* 2021;2021(9999368):9. doi: 10.1155/2021/9999368

70. Salehi SSM, Erdogmus D, Gholipour A. Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks. *Machine Learning in Medical Imaging.* 2017;10541:379–387. doi: 10.1007/978-3-319-67389-9_44

71. Oktay O, Schlemper J, Folgoc LL, et al. Attention U-Net: Learning Where to Look for the Pancreas. *ArXiv.* 2018. doi: 10.48550/arXiv.1804.03999

72. Agarap AF. Deep Learning using Rectified Linear Units (ReLU). *ArXiv.* 2018:7. doi: 10.48550/arXiv.1803.08375

73. Wu J, Zhou S, Zuo S, et al. U-Net combined with multi-scale attention mechanism for liver segmentation in CT images. *BMC Medical Informatics and Decision Making.* 2021;21(283). doi: 10.1186/s12911-021-01649-w

74. Zhang L, Liu Y, Li Z, Li D. Epa-unet:automatic Segmentation of Liver and Tumor in Ct Images Based on Residual U-net and Efficient Multiscale Attention Methods. *Research Square.* 2023. doi: 10.21203/rs.3.rs-3273964/v1

75. Zarbakhsh P. Spatial Attention Mechanism and Cascade Feature Extraction in a U-Net Model for Enhancing Breast Tumor Segmentation. *Applied Sciences.* 2023;13(15):8758. doi: 10.3390/app13158758

76. Subhan Akbar A, Fatichah C, Suciati N. UNet3D with Multiple Atrous Convolutions Attention Block for Brain Tumor Segmentation. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries.* 2022:182–193. doi: 10.1007/978-3-031-08999-2_14

77. Yu Z, Han S, Song Z. 3D Medical Image Segmentation based on multi-scale MPU-Net. *ArXiv.* 2023. doi: 10.48550/arXiv.2307.05799

78. Xingfei F, Chaobing H. CAE-UNet: An Effective Automatic Segmentation Model for CT Images of COVID-19. *2022 6th International Conference on Communication and Information Systems (ICCIS).* 2022:113–117. doi: 10.1109/ICCIS56375.2022.9998131

79. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 2016:770–778. doi: 10.1109/cvpr.2016.90

80. Hatamizadeh A, Tang Y, Nathet V, et al. U-NETR: Transformers for 3D Medical Image Segmentation. *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV).* 2022:1748–1758. doi: 10.1109/WACV51458.2022.00181

81. Eskandari S, Lumpp J. Inter-Scale Dependency Modeling for Skin Lesion Segmentation with Transformer-based Networks. *ArXiv.* 2023. doi: 10.48550/arXiv.2310.13727

82. Shi X, Chen Z, Wang H, et al. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *Neural Information Processing Systems.* 2015. doi: 10.48550/arXiv.1506.04214

83. Pham TH, Li X, Nguyen KD. SeU-Net-Trans: A Simple yet Effective UNet-Transformer Model for Medical Image Segmentation. *ArXiv.* 2023. doi: 10.48550/arXiv.2310.09998

84. Ghofrani F, Behnam H, Motlagh HDK. Liver Segmentation in CT Images Using Deep Neural Networks. *2020 28th Iranian Conference on Electrical Engineering (ICEE).* 2020:1–6. doi: 10.1109/ICEE50131.2020.9260809

85. Diakogiannis FI, Waldner F, Caccetta P, Wuet C, et al. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing.* 2020;16(2):94–114. doi: 10.1016/j.isprsjprs.2020.01.013

86. Jha D, Riegler MA, Johansen D, Halvorsen P, Johansen HD. Doubleu-net: DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation. *IEEE 33rd International symposium on computer-based medical systems (CBMS).* 2020:558–564. doi: 10.1109/CBMS49503.2020.00111

87. Lee HH, Bao S, Huo Y, Landman BA. 3D UX-Net: A Large Kernel Volumetric ConvNet Modernizing Hierarchical Transformer for Medical Image Segmentation. *International Conference on Learning Representations.* 2023. doi: 10.48550/arXiv.2209.15076

88. Liang J, Yang C, Zhong J, Ye X. BTSwin-U-Net: 3D U-shaped Symmetrical Swin Transformer-based Network for Brain Tumor Segmentation with Self-supervised Pre-training. *Neural Processing Letters.* 2022;55:3695–3713. doi: 10.1007/s11063-022-10919-1

89. Alalwan N, Abozeid A, ElHabshy AA, Alzahrani A. Efficient 3D Deep Learning Model for Medical Image Semantic Segmentation. *Alexandria Engineering Journal.* 2021;60(1):1231–1239. doi: 10.1016/j.aej.2020.10.046

90. Lemay A, Gros C, Vincent O, et al. Benefits of Linear Conditioning with Metadata for Image Segmentation. *ArXiv.* 2021. doi: 10.48550/arXiv.2102.09582

91. Du R, Vardhanabhuti V. 3D-RADNet: Extracting labels from DICOM metadata for training general medical domain deep 3D convolution neural networks. *International Conference on Medical Imaging with Deep Learning.* 2020;121:174–192. Available at: https://proceedings.mlr.press/v121/du20a/du20a.pdf. Accessed: November 9, 2024.

92. Plutenko I, Papkov M, Palo K, Parts L, Fishman D. Metadata Improves Segmentation Through Multitasking Elicitation. *Domain Adaptation and Representation Transfer.* 2023:147–155. doi: 10.1007/978-3-031-45857-6_15

93. Jiang J, Peng Y, Hou Q, Wang J. MDCF_Net: A Multi-dimensional hybrid network for liver and tumor segmentation from CT. *Biocybernetics and Biomedical Engineering.* 2023;43(2):494–506. doi: 10.1016/j.bbe.2023.04.004

94. Fu T, Yu Q, Lao H, Liu P, Wan S. Traffic Safety Oriented Multi-Intersection Flow Prediction Based on Transformer and CNN. *Security and Communication Networks.* 2023:1–13. doi: 10.1155/2023/1363639

95. Chen X, Wei X, Tang M, et al. Liver segmentation in CT imaging with enhanced mask region-based convolutional neural networks. *Annals of Translational Medicine.* 2021;9(24):1768. doi: 10.21037/atm-21-5822

96. Ernst P, Chatterjee S, Rose G, Nürnberger A. Primal-Dual U-Net for Sparse View Cone Beam Computed Tomography Volume Reconstruction. *ArXiv.* 2022. doi: 10.48550/arXiv.2205.07866

97. Koehler G, Wald T, Ulrichet C, et al. RecycleNet: Latent Feature Recycling Leads to Iterative Decision Refinement. *ArXiv.* 2023. doi: 10.48550/arXiv.2309.07513

98. Jafari M, Auer D, Francis S, Garibaldi J, Chen X. DRU-net: An Efficient Deep Convolutional Neural Network for Medical Image Segmentation. *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI).* 2020:1144–1148. doi: 10.48550/arXiv.2004.13453

99. Heker M, Ben-Cohen A, Greenspan H. Hierarchical Fine-Tuning for joint Liver Lesion Segmentation and Lesion Classification in CT. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* 2019:895–898. doi: 10.1109/EMBC.2019.8857127

100. Matovinovic IZ, Loncaric S, Lo J, Heisler M, Sarunic M. Transfer Learning with U-Net type model for Automatic Segmentation of Three Retinal Layers In Optical Coherence Tomography Images. *2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA).* 2019:49–53. doi: 10.1109/ISPA.2019.8868639

101. Kora P, Ooi CP, Faust O, et al. Transfer learning techniques for medical image analysis: A review. *Biocybernetics and Biomedical Engineering.* 2022;42(1):79–107. doi: 10.1016/j.bbe.2021.11.004

102. Humpire-Mamani GE, Jacobs C, Prokop M, van Ginneken B, Lessmann N. Transfer learning from a sparsely annotated dataset of 3D medical images. *ArXiv.* 2023. doi: 10.48550/arXiv.2311.05032

103. Messaoudi H, Belaid A, Salem DB, Conze P-H. Cross-dimensional transfer learning in medical image segmentation with deep learning. *Medical Image Analysis.* 2023;88:102868. doi: 10.1016/j.media.2023.102868

104. Tan M, Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *International conference on machine learning (PMLR).* 2019:6105–6114. doi: 10.48550/arXiv.1905.11946
105. Hong Y, Mao X, Hui Q. et al. Automatic liver and tumor segmentation based on deep learning and globally optimized refinement. *Applied Mathematics-A Journal of Chinese Universities.* 2021;36:304–316. doi: 10.1007/s11766-021-4376-3
106. Wang H, Li X. Towards Generic Semi-Supervised Framework for Volumetric Medical Image Segmentation. *ArXiv.* 2023. doi: 10.48550/arXiv.2310.11320
107. Wang J, Chen C. Unsupervised Adaptation of Polyp Segmentation Models via Coarse-to-Fine Self-Supervision. *Information Processing in Medical Imaging.* 2023:250–262. doi: 10.1007/978-3-031-34048-2_20
108. Wang T, Huang Z, Wu J, Cai Y, Li Z. Semi-Supervised Medical Image Segmentation with Co-Distribution Alignment. *Bioengineering.* 2023;10(7):869. doi: 10.3390/bioengineering10070869
109. Vasilev YA, Bobrovskaya TM, Arzamasov KM, et al. Medical datasets for machine learning: fundamental principles of standartization and systematization. *Manager Zdravoochranenia.* 2023(4):28–41. doi: 10.21045/1811-0185-2023-4-28-41
110. Kokina DYu, Gombolevskiy VA. Arzamasov KM, Andreychenko AE, Morozov SP Possibilities and limitations of using machine text-processing tools in Russian radiology reports. *Digital Diagnostics.* 2022;3(4):374–383. doi: 10.17816/DD101099
111. Ronzhin LV, Astanin PA, Kokina DYu, et al Semantic analysis methods in the system for authomated marking of the unstructured radiological chest examination protocols. *Social'nye aspekty zdorov'a naselenia.* 2023;69(1):12. doi: 10.21045/2071-5021-2023-69-1-12
112. Tomashevskaya VS, Yakovlev DA. Research of unstructured data interpretation problems. *Russian Technological Journal.* 2021;9(1):7–17. doi: 10.32362/2500-316X-2021-9-1-7-17
113. Protonotarios N, Katsamenis I, Sykiotis S, et al. A few-shot U-Net deep learning model for lung cancer lesion segmentation via PET/CT imaging. *Biomedical Physics and Engineering Express.* 2022;8:025019. doi: 10.1088/2057-1976/ac53bd
114. Voulodimos A, Protopapadakis E, Katsamenis I, Doulamis A, Doulamis N. A Few-Shot U-Net Deep Learning Model for COVID-19 Infected Area Segmentation in CT Images. *Sensors.* 2021;21(6):2215. doi: 10.3390/s21062215
115. Zhao G, Zhao H. One-Shot Image Segmentation with U-Net. *Journal of Physics: Conference Series.* 2021;1848(1):012113. doi: 10.1088/1742-6596/1848/1/012113

СПИСОК ЛИТЕРАТУРЫ

1. Shen D., Wu G., Suk H.I. Deep Learning in Medical Image Analysis // Annual Review of Biomedical Engineering. 2017. Vol. 19. P. 221–248. doi: 10.1146/annurev-bioeng-071516-044442
2. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015. 2015. Vol. 9351. doi: 10.1007/978-3-319-24574-4_28
3. Milletari F., Navab N., Ahmadi S.A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation // Fourth International Conference on 3D Vision (3DV). 2016. P. 565–571. doi: 10.48550/arXiv.1606.04797
4. Chen L.C., Papandreou G., Kokkinos I., Murphy K., Yuille A. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. Vol. 40, N 4. P. 834–848. doi: 10.1109/TPAMI.2017.2699184
5. Huang G., Liu Z., Van Der Maaten L., Weinberger K.Q. Densely Connected Convolutional Networks // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 2261–2269. doi: 10.1109/CVPR.2017.243
6. He K., Gkioxari G., Dollár P., Girshick R. Mask R-CNN // IEEE International Conference on Computer Vision (ICCV). 2017. P. 2980–2988. doi: 10.1109/ICCV.2017.322
7. Khalal D.M., Azizi H., Maalej N. Automatic segmentation of kidneys in computed tomography images using U-Net // Cancer/Radiothérapie. 2023. Vol. 27, N 2. P. 109–114. doi: 10.1016/j.canrad.2022.08.004
8. Bernardo Gois F.N., Lobo Marques J.A. Segmentation of CT-Scan Images Using UNet Network for Patients Diagnosed with COVID-19 // Computerized Systems for Diagnosis and Treatment of COVID-192023. 2023. P. 29–44. doi: 10.1007/978-3-031-30788-1_3
9. Sarsembayeva T., Shomanov A., Sarsembayev M., et al. UNet Model for Segmentation of COPD Lung Lesions on Computed Tomography Images // Proceedings of the 7th International Conference on Digital Technologies in Education, Science and Industry (DTESI 2022). 2022. Available at: https://ceur-ws.org/Vol-3382/Short5.pdf. Accessed: November 9, 2024.
10. Çiçek Ö., Abdulkadir A., Lienkamp S., Brox T., Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. 2016. P. 424–432. doi: 10.1007/978-3-319-46723-8_4
11. Pantovic A., Ollivier I., Essert C. 2D and 3D-UNet for segmentation of SEEG electrode contacts on post-operative CT scans // Medical Imaging 2022: Image-Guided Procedures, Robotic Interventions, and Modeling. 2022. doi: 10.1117/12.2606538
12. Han X., Wu X., Wang S., et al. Automated segmentation of liver segment on portal venous phase MR images using a 3D convolutional neural network // Insights Imaging. 2022. Vol. 13, N 26. doi: 10.1186/s13244-022-01163-1
13. Zhou Z., Rahman Siddiquee M.M., Tajbakhsh N., Liang J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation // Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. 2018. P. 3–11. doi: 10.1007/978-3-030-00889-5_1
14. Yu C., Wang Y., Tang C., Feng W., Lv J. EU-Net: Automatic U-Net neural architecture search with differential evolutionary algorithm for medical image segmentation // Computers in Biology and Medicine. 2023. Vol. 167. P. 107579. doi: 10.1016/j.compbiomed.2023.107579
15. Weng Y., Zhou T., Li Y., Qiu X. NAS-Unet: Neural Architecture Search for Medical Image Segmentation // IEEE Access. 2019. Vol. 7. P. 44247–44257. doi: 10.1109/ACCESS.2019.2908991
16. Huang H., Lin L., Tong R., et al. UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation // ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2020. P. 1055–1059. doi: 10.1109/ICASSP40776.2020.9053405
17. Li C., Bagher-Ebadian H., Sultan R.I., et al. A new architecture combining convolutional and transformer-based networks for

automatic 3D multi-organ segmentation on CT images // Med Phys. 2023. Vol. 50, N 11. P. 6990–7002. doi: 10.1002/mp.16750
18. Müller D., Soto-Rey I., Kramer F. Towards a guideline for evaluation metrics in medical image segmentation // BMC Research Notes. 2022. Vol. 15, N 210. doi: 10.1186/s13104-022-06096-y
19. Alberg A.J., Park J.W., Hager B.W., Brock M.V., Diener-West M. The use of «overall accuracy» to evaluate the validity of screening or diagnostic tests // Journal of General Internal Medicine. 2004. Vol. 19. P. 460–465. doi: 10.1111/j.1525-1497.2004.30091.x
20. Soler L., Hostettler A., Agnus V., et al. 3D image reconstruction for comparison of algorithm database: A patient specific anatomical and medical image database // IRCAD. 2010. Available at: https://www-sop.inria.fr/geometrica/events/wam/abstract-ircad.pdf. Accessed: November 9, 2024.
21. Löffler M., Sekuboyina A., Jakob A., et al. A Vertebral Segmentation Dataset with Fracture Grading // Radiology: Artificial Intelligence. 2020. Vol. 2, N 4. doi: 10.1148/ryai.2020190138
22. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: from error visibility to structural similarity // IEEE Transactions on Image Processing. 2004. Vol. 13, N 4. P. 600–612. doi: 10.1109/TIP.2003.819861
23. Kavur A.E., Gezer N.S., Barıs M., et al. CHAOS Challenge – combined (CT-MR) healthy abdominal organ segmentation // Medical Image Analysis. 2021. Vol. 69. P. 101950. doi: 10.1016/j.media.2020.101950
24. Bilic P., Christ P., Li H.B., et al. The Liver Tumor Segmentation Benchmark (LiTS) // Medical Image Analysis. 2023. Vol. 84. P. 102680. doi: 10.1016/j.media.2022.102680
25. Petrusca L., Cattin P., De Luca V., et al. Hybrid ultrasound/magnetic resonance simultaneous acquisition and image fusion for motion monitoring in the upper abdomen // Investigative Radiology. 2013. Vol. 48, N 5. P. 333–340. doi: 10.1097/RLI.0b013e31828236c3
26. Jun M., Cheng G., Yixin W., et al. Covid-19 CT lung and infection segmentation dataset // Zenodo. 2020. Available at: https://zenodo.org/records/3757476#.YLov8vkzaUk. Accessed: November 9, 2024.
27. Morozov S.P., Andreychenko A.E., Blokhin I.A., et al. MosMedData: data set of 1110 chest CT scans performed during the COVID-19 epidemic // Digital Diagnostics. 2020. Vol. 1, N 1. P. 49–59. doi: 10.17816/DD46826
28. Roth H.R., Oda H., Hayashi Y., et al. Hierarchical 3D fully convolutional networks for multi-organ segmentation // ArXiv. 2017. Available at: https://arxiv.org/abs/1704.06382v1. Accessed: November 9, 2024.
29. Roth H., Farag A., Turkbey E.B., et al. Data from Pancreas-CT // The Cancer Imaging Archive. 2016. doi: 10.7937/K9/TCIA.2016.tNB1kqBU
30. Heimann T., Styner M., van Ginneken B. 3D Segmentation in the Clinic: A Grand Challenge // MICCAI 2007, the 10th Intel Conf. on Medical Image Computing and Computer Assisted Intervention. 2007. P. 7–15. Available at: https://www.diagnijmegen.nl/publications/ginn07/. Accessed: November 9, 2024.
31. Suckling J. The Mammographic Image Analysis Society Digital Mammogram Database // International Congress Series. 1994. P. 375–378. Available at: http://peipa.essex.ac.uk/info/mias.html. Accessed: November 9, 2024.
32. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020 [Internet]. 2020. Доступ по ссылке: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 [Дата обращения: 9.11.2024].
33. Landman B., Xu Z., Igelsias J., et al. Miccai multi-atlas labeling beyond the cranial vault–workshop and challenge // Proceedings of the MICCAI Multi-Atlas Labeling Beyond Cranial Vault — Workshop Challenge. 2015. Vol. 5. P. 12.
34. Simpson A.L., Antonelli M., Bakas S., et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms // ArXiv. 2019. doi: 10.48550/arXiv.1902.09063
35. Gutman D., Codella N.C.F., Celebi E., et al. Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC) // ArXiv. 2016. doi: 10.48550/arXiv.1605.01397
36. Jha D., Smedsrud P.H., Riegler M.A., et al. Kvasir-SEG: A Segmented Polyp Dataset // MultiMedia Modeling. 2020. Vol. 11962. P. 451–462. doi: 10.1007/978-3-030-37734-2_37
37. Bernal J., Sánchez F.J., Fernández-Esparrach G., et al. WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians // Computerized Medical Imaging and Graphics. 2015. Vol. 43. P. 99–111. doi: 10.1016/j.compmedimag.2015.02.007
38. Grove O., Berglund A.E., Schabath M.B., et al. Quantitative Computed Tomographic Descriptors Associate Tumor Shape Complexity and Intratumor Heterogeneity with Prognosis in Lung Adenocarcinoma // PLOS ONE. 2015. Vol. 10, N 3. P. e0118261. doi: 10.1371/journal.pone.0118261
39. Heller N., Sathianathen N., Kalapara A., et al. The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes // ArXiv. 2019. P. 13. doi: 10.48550/arXiv.1904.00445
40. Ji Y., Bai H., Yang J., et al. AMOS: A Large-Scale Abdominal Multi-Organ Benchmark for Versatile Medical Image Segmentation // ArXiv. 2022. doi: 10.48550/arXiv.2206.08023
41. Lemay A., Gros C., Zhuo Z., et al. Multiclass Spinal Cord Tumor Segmentation on MRI with Deep Learning // ArXiv. 2021. doi: 10.48550/arXiv.2012.12820
42. Ali M.A.S., Misko O., Salumaa S.O., et al. Evaluating Very Deep Convolutional Neural Networks for Nucleus Segmentation from Brightfield Cell Microscopy Images // SLAS Discovery. 2021. Vol. 26, N 9. P. 1125–1137. doi: 10.1177/24725552211023214
43. Gibson E., Giganti F., Hu Y., et al. Automatic Multi-Organ Segmentation on Abdominal CT With Dense V-Networks // IEEE Transactions on Medical Imaging. 2018. Vol. 37, N 8. P. 1822–1834. doi: 10.1109/TMI.2018.2806309
44. Jimenez-del Toro O., Müller H., Krenn M., et al. Cloud-Based Evaluation of Anatomical Structure Segmentation and Landmark Detection Algorithms: VISCERAL Anatomy Benchmarks // IEEE Transactions on Medical Imaging. 2016. Vol. 35, N 11. P. 2459–2475. doi: 10.1109/TMI.2016.2578680
45. Regan E.A., Hokanson J.E., Murphy J.R., et al. Genetic Epidemiology of COPD (COPDGene) Study Design // COPD: Journal of Chronic Obstructive Pulmonary Disease. 2010. Vol. 7, N 1. P. 32–43. doi: 10.3109/15412550903499522
46. Litjens G., Toth R., van de Ven W., et al. Evaluation of prostate segmentation algorithms for MRI: The PROMISE12 challenge // Medical Image Analysis. 2014. Vol. 18, N 2. P. 359–373. doi: 10.1016/j.media.2013.12.002
47. Xiong Z., Xia Q., Hu Z., et al. A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging // Medical Image Analysis. 2021. Vol. 67. P. 101832. doi: 10.1016/j.media.2020.101832

48. Landman B., Xu Z., Igelsias J., et al. 2015 MICCAI multi-atlas labeling beyond the cranial vault–workshop and challenge // MICCAI Multi-Atlas Labeling Beyond Cranial Vault — Workshop Challenge. 2015. Vol. 5. P. 12.

49. Zhuang X., Shen J. Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI // Medical Image Analysis. 2016. Vol. 31. P. 77–87. doi: 10.1016/j.media.2016.02.006

50. Campello V.M., Gkontra P., Izquierdo C., et al. Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge // IEEE Transactions on Medical Imaging. 2021. Vol. 40, N 12. P. 3543–3554. doi: 10.1109/TMI.2021.3090082

51. Silva J., Histace A., Romain O., Dray X., Granado B. Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer // International Journal of Computer Assisted Radiology and Surgery. 2014. Vol. 9. P. 283–293. doi: 10.1007/s11548-013-0926-3

52. Trikha S., Turnbull A., Morris R., Anderson D., Hossain P. The journey to femtosecond laser-assisted cataract surgery: New beginnings or a false dawn? // Eye. 2013. Vol. 27, N 4. P. 461–473. doi: 10.1038/eye.2012.293

53. Xiong Z., Xia Q., Hu Z., et al. A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging // Medical Image Analisys. 2021. Vol. 67. P. 101832. doi: 10.1016/j.media.2020.101832

54. Bernard O., Lalande A., Zotti C., et al. Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved? // IEEE Transactions on Medical Imaging. 2018. Vol. 37, N 11. P. 2514–2525. doi: 10.1109/TMI.2018.2837502

55. Li P., Wang S., Li T., et al. A Large-Scale CT and PET/CT Dataset for Lung Cancer Diagnosis (Lung-PET-CT-Dx) [Data set] // The Cancer Imaging Archive. 2020. doi: 10.7937/TCIA.2020.NNC2-0461

56. Clark K., Vendt B., Smith K., et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository // Journal of Digital Imaging. 2013. Vol. 26. P. 1045–1057. doi: 10.1007/s10278-013-9622-7

57. Xu Z., Jia Z., Sun J., Dong W., Li Z. DO-U-Net: Improved U-Net Model for CT Image Segmentation using DBB and Octave Convolution // Proceedings of the 2023 International Conference on Computer, Vision and Intelligent Technology (ICCVIT '23). 2023. P. 1–8. doi: 10.1145/3627341.3630403

58. Ayalew Y., Fante K., Aliy M. Modified U-Net for liver cancer segmentation from computed tomography images with a new class balancing method // BMC Biomedical Engineering. 2021. Vol. 3, N 4. doi: 10.1186/s42490-021-00050-y

59. Guan S., Khan A.A., Sikdar S., Chitnis P.V. Fully Dense UNet for 2-D Sparse Photoacoustic Tomography Artifact Removal // IEEE Journal of Biomedical and Health Informatics. 2020. Vol. 24, N 2. P. 568–576. doi: 10.1109/JBHI.2019.2912935

60. Özcan F., Uçan O.N., Karaçam S., Tunçman D. Fully Automatic Liver and Tumor Segmentation from CT Image Using an AIM-UNet // Bioengineering. 2023. Vol. 10, N 2. doi: 10.3390/bioengineering10020215

61. Ansari M.Y., Yang Y., Meher P.K., Dakua S.P. Dense-PSP-UNet: A neural network for fast inference liver ultrasound segmentation // Computers in Biology and Medicine. 2023. Vol. 153. P. 106478. doi: 10.1016/j.compbiomed.2022.106478

62. Omarov B., Tursynova A., Postolache O., et al. Modified UNet Model for Brain Stroke Lesion Segmentation on Computed Tomography Images // Computers, Materials and Continua. 2022. Vol. 71, N 3. P. 4701–4717. doi: 10.32604/cmc.2022.020998

63. Mizusawa S., Sei Y., Orihara R., Ohsuga A. Computed tomography image reconstruction using stacked U-Net // Computerized Medical Imaging and Graphics. 2021. Vol. 90. P. 101920. doi: 10.1016/j.compmedimag.2021.101920

64. Golts A., Khapun D., Shats D., Shoshan Y., Gilboa-Solomon F. An Ensemble of 3D U-Net Based Models for Segmentation of Kidney and Masses in CT Scans // Kidney and Kidney Tumor Segmentation (KiTS 2021). 2022. Vol. 13168. P. 103–115. doi: 10.1007/978-3-030-98385-7_14

65. Araújo J.D.L., da Cruz L.B., Diniz J.O.B., et al. Liver segmentation from computed tomography images using cascade deep learning // Computers in Biology and Medicine. 2022. Vol. 140. P. 105095. doi: 10.1016/j.compbiomed.2021.105095

66. Koirala C.P., Mohapatra S., Gosai A., Schlaug G. Automated Ensemble-Based Segmentation of Adult Brain Tumors: A Novel Approach Using the BraTS AFRICA Challenge Data // ArXiv. 2023. doi: 10.48550/arXiv.2308.07214

67. Li Z., Zhu Q., Zhang L., et al. A deep learning-based self-adapting ensemble method for segmentation in gynecological brachytherapy // Radiation Oncology. 2022. Vol. 17, N 152. doi: 10.1186/s13014-022-02121-3

68. Woo S., Park J., Lee J.-Y., Kweon I.S. CBAM: Convolutional Block Attention Module // Proceedings of the European conference on computer vision (ECCV). 2018. P. 3–19. doi: 10.48550/arXiv.1807.06521

69. Nazir S., Zheng R., Zheng Y., Dong-Ye C. Improved 3D U-Net for COVID-19 Chest CT Image Segmentation // Scientific Programming. 2021. Vol. 2021, N 9999368. P. 9. doi: 10.1155/2021/9999368

70. Salehi S.S.M., Erdogmus D., Gholipour A. Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks // Machine Learning in Medical Imaging. 2017. Vol. 10541. P. 379–387. doi: 10.1007/978-3-319-67389-9_44

71. Oktay O., Schlemper J., Folgoc L.L., et al. Attention U-Net: Learning Where to Look for the Pancreas // ArXiv. 2018. doi: 10.48550/arXiv.1804.03999

72. Agarap A.F. Deep Learning using Rectified Linear Units (ReLU) // ArXiv. 2018. P. 7. doi: 10.48550/arXiv.1803.08375

73. Wu J., Zhou S., Zuo S., et al. U-Net combined with multi-scale attention mechanism for liver segmentation in CT images // BMC Medical Informatics and Decision Making. 2021. Vol. 21, N 283. doi: 10.1186/s12911-021-01649-w

74. Zhang L., Liu Y., Li Z., Li D. Epa-unet:automatic Segmentation of Liver and Tumor in Ct Images Based on Residual U-net and Efficient Multiscale Attention Methods // Research Square. 2023. doi: 10.21203/rs.3.rs-3273964/v1

75. Zarbakhsh P. Spatial Attention Mechanism and Cascade Feature Extraction in a U-Net Model for Enhancing Breast Tumor Segmentation // Applied Sciences. 2023. Vol. 13, N 15. P. 8758. doi: 10.3390/app13158758

76. Subhan Akbar A., Fatichah C., Suciati N. UNet3D with Multiple Atrous Convolutions Attention Block for Brain Tumor Segmentation // Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. 2022. P. 182–193. doi: 10.1007/978-3-031-08999-2_14

77. Yu Z., Han S., Song Z. 3D Medical Image Segmentation based on multi-scale MPU-Net // ArXiv. 2023. doi: 10.48550/arXiv.2307.05799

78. Xingfei F., Chaobing H. CAE-UNet: An Effective Automatic Segmentation Model for CT Images of COVID-19 // 2022 6th

International Conference on Communication and Information Systems (ICCIS). 2022. P. 113–117. doi: 10.1109/ICCIS56375.2022.9998131
79. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 770–778. doi: 10.1109/cvpr.2016.90
80. Hatamizadeh A., Tang Y., Nathet V., et al. U-NETR: Transformers for 3D Medical Image Segmentation // 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). 2022. P. 1748–1758. doi: 10.1109/WACV51458.2022.00181
81. Eskandari S., Lumpp J. Inter-Scale Dependency Modeling for Skin Lesion Segmentation with Transformer-based Networks // ArXiv. 2023. doi: 10.48550/arXiv.2310.13727
82. Shi X., Chen Z., Wang H., et al. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting // Neural Information Processing Systems. 2015. doi: 10.48550/arXiv.1506.04214
83. Pham T.H., Li X., Nguyen K.D. SeU-Net-Trans: A Simple yet Effective UNet-Transformer Model for Medical Image Segmentation // ArXiv. 2023. doi: 10.48550/arXiv.2310.09998
84. Ghofrani F., Behnam H., Motlagh H.D.K. Liver Segmentation in CT Images Using Deep Neural Networks // 2020 28th Iranian Conference on Electrical Engineering (ICEE). 2020. P. 1–6. doi: 10.1109/ICEE50131.2020.9260809
85. Diakogiannis F.I., Waldner F., Caccetta P., Wuet C., et al. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data // ISPRS Journal of Photogrammetry and Remote Sensing. 2020. Vol. 16, N 2. P. 94–114. doi: 10.1016/j.isprsjprs.2020.01.013
86. Jha D., Riegler M.A., Johansen D., Halvorsen P., Johansen H.D. Doubleu-net: DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation // IEEE 33rd International symposium on computer-based medical systems (CBMS). 2020. P. 558–564. doi: 10.1109/CBMS49503.2020.00111
87. Lee H.H., Bao S., Huo Y., Landman B.A. 3D UX-Net: A Large Kernel Volumetric ConvNet Modernizing Hierarchical Transformer for Medical Image Segmentation // International Conference on Learning Representations. 2023. doi: 10.48550/arXiv.2209.15076
88. Liang J., Yang C., Zhong J., Ye X. BTSwin-U-Net: 3D U-shaped Symmetrical Swin Transformer-based Network for Brain Tumor Segmentation with Self-supervised Pre-training // Neural Processing Letters. 2022. Vol. 55. P. 3695–3713. doi: 10.1007/s11063-022-10919-1
89. Alalwan N., Abozeid A., ElHabshy A.A., Alzahrani A. Efficient 3D Deep Learning Model for Medical Image Semantic Segmentation // Alexandria Engineering Journal. 2021. Vol. 60, N 1. P. 1231–1239. doi: 10.1016/j.aej.2020.10.046
90. Lemay A., Gros C., Vincent O., et al. Benefits of Linear Conditioning with Metadata for Image Segmentation // ArXiv. 2021. doi: 10.48550/arXiv.2102.09582
91. Du R., Vardhanabhuti V. 3D-RADNet: Extracting labels from DICOM metadata for training general medical domain deep 3D convolution neural networks // International Conference on Medical Imaging with Deep Learning. 2020. Vol. 121. P. 174–192. Available at: https://proceedings.mlr.press/v121/du20a/du20a.pdf. Accessed: November 9, 2024.
92. Plutenko I., Papkov M., Palo K., Parts L., Fishman D. Metadata Improves Segmentation Through Multitasking Elicitation // Domain Adaptation and Representation Transfer. 2023. P. 147–155. doi: 10.1007/978-3-031-45857-6_15
93. Jiang J., Peng Y., Hou Q., Wang J. MDCF_Net: A Multi-dimensional hybrid network for liver and tumor segmentation from CT // Biocybernetics and Biomedical Engineering. 2023. Vol. 43, N 2. P. 494–506. doi: 10.1016/j.bbe.2023.04.004
94. Fu T., Yu Q., Lao H., Liu P., Wan S. Traffic Safety Oriented Multi-Intersection Flow Prediction Based on Transformer and CNN // Security and Communication Networks. 2023. P. 1–13. doi: 10.1155/2023/1363639
95. Chen X., Wei X., Tang M., et al. Liver segmentation in CT imaging with enhanced mask region-based convolutional neural networks // Annals of Translational Medicine. 2021. Vol. 9, N 24. P. 1768. doi: 10.21037/atm-21-5822
96. Ernst P., Chatterjee S., Rose G., Nürnberger A. Primal-Dual U-Net for Sparse View Cone Beam Computed Tomography Volume Reconstruction // ArXiv. 2022. doi: 10.48550/arXiv.2205.07866
97. Koehler G., Wald T., Ulrichet C., et al. RecycleNet: Latent Feature Recycling Leads to Iterative Decision Refinement // ArXiv. 2023. doi: 10.48550/arXiv.2309.07513
98. Jafari M., Auer D., Francis S., Garibaldi J., Chen X. DRU-net: An Efficient Deep Convolutional Neural Network for Medical Image Segmentation // 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). 2020. P. 1144–1148. doi: 10.48550/arXiv.2004.13453
99. Heker M., Ben-Cohen A., Greenspan H. Hierarchical Fine-Tuning for joint Liver Lesion Segmentation and Lesion Classification in CT // 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2019. P. 895–898. doi: 10.1109/EMBC.2019.8857127
100. Matovinovic I.Z., Loncaric S., Lo J., Heisler M., Sarunic M. Transfer Learning with U-Net type model for Automatic Segmentation of Three Retinal Layers In Optical Coherence Tomography Images // 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA). 2019. P. 49–53. doi: 10.1109/ISPA.2019.8868639
101. Kora P., Ooi C.P., Faust O., et al. Transfer learning techniques for medical image analysis: A review // Biocybernetics and Biomedical Engineering. 2022. Vol. 42, N 1. P. 79–107. doi: 10.1016/j.bbe.2021.11.004
102. Humpire-Mamani G.E., Jacobs C., Prokop M., van Ginneken B., Lessmann N. Transfer learning from a sparsely annotated dataset of 3D medical images // ArXiv. 2023. doi: 10.48550/arXiv.2311.05032
103. Messaoudi H., Belaid A., Salem D.B., Conze P.-H. Cross-dimensional transfer learning in medical image segmentation with deep learning // Medical Image Analysis. 2023. Vol. 88. P. 102868. doi: 10.1016/j.media.2023.102868
104. Tan M., Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks // International conference on machine learning (PMLR). 2019. P. 6105–6114. doi: 10.48550/arXiv.1905.11946
105. Hong Y., Mao X., Hui Q. et al. Automatic liver and tumor segmentation based on deep learning and globally optimized refinement // Applied Mathematics-A Journal of Chinese Universities. 2021. Vol. 36. P. 304–316. doi: 10.1007/s11766-021-4376-3
106. Wang H., Li X. Towards Generic Semi-Supervised Framework for Volumetric Medical Image Segmentation // ArXiv. 2023. doi: 10.48550/arXiv.2310.11320
107. Wang J., Chen C. Unsupervised Adaptation of Polyp Segmentation Models via Coarse-to-Fine Self-Supervision // Information Processing in Medical Imaging. 2023. P. 250–262. doi: 10.1007/978-3-031-34048-2_20
108. Wang T., Huang Z., Wu J., Cai Y., Li Z. Semi-Supervised Medical Image Segmentation with Co-Distribution Alignment // Bioengineering. 2023. Vol. 10, N 7. P. 869. doi: 10.3390/bioengineering10070869

109. Васильев Ю.А., Бобровская Т.М., Арзамасов К.М., и др. Основополагающие принципы стандартизации и систематизации информации о наборах данных для машинного обучения в медицинской диагностике // Менеджер здравоохранения. 2023. № 4. С. 28–41. doi: 10.21045/1811-0185-2023-4-28-41

110. Кокина Д.Ю., Гомболевский В.А., Арзамасов К.М., Андрейченко А.Е., Морозов С.П. Возможности и ограничения использования инструментов машинной обработки текстов в лучевой диагностике // Digital Diagnostics. 2022. Т. 3, № 4. С. 374–383. doi: 10.17816/DD101099

111. Ронжин Л.В., Астанин П.А., Кокина Д.Ю., и др. Система автоматической разметки неструктурированных протоколов рентгенологических исследований грудной клетки с использованием методов семантического анализа // Социальные аспекты здоровья населения. 2023. Т. 69, № 1. С. 12. doi: 10.21045/2071-5021-2023-69-1-12

112. Томашевская В.С., Яковлев Д.А. Способы обработки неструктурированных данных // Российский технологический журнал. 2021. Т. 9, № 1. С. 7–17. doi: 10.32362/2500-316X-2021-9-1-7-17

113. Protonotarios N., Katsamenis I., Sykiotis S., et al. A few-shot U-Net deep learning model for lung cancer lesion segmentation via PET/CT imaging // Biomedical Physics and Engineering Express. 2022. Vol. 8. P. 025019. doi: 10.1088/2057-1976/ac53bd

114. Voulodimos A., Protopapadakis E., Katsamenis I., Doulamis A., Doulamis N. A Few-Shot U-Net Deep Learning Model for COVID-19 Infected Area Segmentation in CT Images // Sensors. 2021. Vol. 21, N 6. P. 2215. doi: 10.3390/s21062215

115. Zhao G., Zhao H. One-Shot Image Segmentation with U-Net // Journal of Physics: Conference Series. 2021. Vol. 1848, N 1. P. 012113. doi: 10.1088/1742-6596/1848/1/012113

AUTHORS' INFO

*** Kirill M. Arzamasov,** MD, Cand. Sci. (Medicine);
address: 24 bldg. 1 Petrovka str., 127051, Moscow, Russia;
ORCID: 0000-0001-7786-0349;
eLibrary SPIN: 3160-8062;
e-mail: ArzamasovKM@zdrav.mos.ru

Anastasia M. Dostovalova;
ORCID: 0009-0004-9420-4182;
eLibrary SPIN: 3784-0791;
e-mail: ADostovalova@frccscs.ru

Andrey K. Gorshenin, Dr. Sci. (Physics and Mathematics), Assistant Professor;
ORCID: 0000-0001-8129-8985;
eLibrary SPIN: 1512-3425;
e-mail: AGorshenin@frccsc.ru

Jullia V. Starichkova, Cand. Sci. (Engineering), Assistant Professor;
ORCID: 0000-0003-1804-9761;
eLibrary SPIN: 3001-6791;
e-mail: starichkova@mirea.ru

ОБ АВТОРАХ

*** Арзамасов Кирилл Михайлович,** канд. мед. наук;
адрес: Россия, 127051, Москва, ул. Петровка, д. 24, стр. 1;
ORCID: 0000-0001-7786-0349;
eLibrary SPIN: 3160-8062;
e-mail: ArzamasovKM@zdrav.mos.ru

Достовалова Анастасия Михайловна;
ORCID: 0009-0004-9420-4182;
eLibrary SPIN: 3784-0791;
e-mail: ADostovalova@frccscs.ru

Горшенин Андрей Константинович, д-р. физ.-мат. наук, доцент;
ORCID: 0000-0001-8129-8985;
eLibrary SPIN: 1512-3425;
e-mail: AGorshenin@frccsc.ru

Старичкова Юлия Викторовна, канд. техн. наук, доцент;
ORCID: 0000-0003-1804-9761;
eLibrary SPIN: 3001-6791;
e-mail: starichkova@mirea.ru

* Corresponding author / Автор, ответственный за переписку