Journal of COMPUTER AND FORENSIC SCIENCES

CFS 2024, Vol. 3, Issue 1, pp. 5–20
https://doi.org/ 10.5937/jcfs3-48848
Review article

An Overview of Image Processing in Biomedicine Using U-Net Convolutional Neural Network Architecture

Aleksa Komosar[1*], Darko Stefanović[2], and Srđan Sladojević[3]

[1] Faculty of Technical Sciences, University of Novi Sad, Serbia; aleksakomosar@uns.ac.rs

[2] Faculty of Technical Sciences, University of Novi Sad, Serbia; darko.stefanovic@uns.ac.rs

[3] Faculty of Technical Sciences, University of Novi Sad, Serbia; sladojevic@uns.ac.rs

* Corresponding author: aleksakomosar@uns.ac.rs

Received: January 22, 2024 • Accepted: January 30, 2024 • Published: February 22, 2024.

Abstract: Image processing in biomedicine is a very broad field, which includes both medical and technical significance. The aim of this paper is to investigate the current trends in the domain of application of U-Net architecture in the period from 2018 to 2023. The PRISMA framework was used for the systematic literature review, and 4 research questions were asked. For the most part, U-Net architectures are used that can process complex high-resolution images in the fastest way in the context of semantic segmentation. Previous work in image processing has focused on overcoming problems such as the complexity of different architectures, image loss, image resolution, and quality, as well as the size of datasets and noise reduction. The most frequently used groups of datasets are BraTS, Data Science Bowl, and ISIC Challenge. The best general Dice score was obtained for the LUNA16, VESSEL12, and Kaggle Lung datasets with 0.98. It is concluded that the application of the U-Net network is growing, with a focus on solving specific challenges in the context of a certain modality and segment of biomedicine.

Keywords: U-Net; convolution neural network (CNN); biomedicine; medical image.

1. INTRODUCTION

In the last century, the development of biomedicine has rapidly advanced in the fields of diagnostics and health care. The way of observing different types and forms of diseases, conditions, and anomalies when it comes to human health relied on achievements in information technology and equipment. Atam P. Dhawan [1] highlighted radiography (X-rays), computerized tomography (CT), nuclear magnetic resonance (NMR), and more specifically, magnetic resonance imaging (MRI) as conventional bases of biomedicine image techniques in the early stage. In that period, the equipment used for obtaining different recordings was also developed, with a focus on enhancing image quality, noise reduction, and basic feature extraction. Over time, large amounts of data were collected and enabled more complex use in the field of biomedicine, where a great synergy between computer

Citation: A. Komosar, D. Stefanović, and S. Sladojević, "An Overview of Image Processing in Biomedicine Using U-Net Convolutional Neural Network Architecture," *Journal of Computer and Forensic Sciences*, vol. 3, no. 1, pp 5–20, 2024, https://doi.org/ 10.5937/jcfs3-48848. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

vision and medical imaging was observed [1], [2]. Therefore, over time, different groups of obstacles appeared in the development of accurate and precise techniques in this field [3]. There are many applications of different architectures in image processing.

For instance, P. Chinmayi et al. [4] highlighted image segmentation as the most significant stage of image processing, due to extracting the Region of Interest (ROI) for further analysis of an image. To explain the complexity of the image segmentation technical obstacles, the problem of identifying different types of tissue is doubled. Additionally, by way of illustration, different image sizes can cause incorrect outputs and demand manual intervention [4]. Furthermore, some methods regarding segmentation cannot be done for MRI 3D data. On the other hand, as a security obstacle, some methods cannot be used on limited datasets [3], [4]. Regarding privacy and ethics, taking images as medical data from patients needs to be prepared in a specific way. Katharina Grünberg et al. [5] present a protocol for data preparation. The first step is gaining ethical and privacy approval from the patient and reviewing it by the Medical Ethics Committee, which results in the handling of informed consent procedures. In addition to this challenge, neural networks for image processing are prone to overfitting due to the size of the dataset. On the other hand, some technical difficulties can be described as the unbalance of pixels that belong to different classes.

The last decade has been characterized by the acceptance of deep learning techniques, especially convolutional neural networks (CNN), to solve complex challenges [6]. Indeed, CNN can solve many problems regarding factors that influence accuracy and precision in the fields of image segmentation, classification, and biomedicine image processing. As a case in point, there is a U-Net architecture that resolves many challenges nowadays. Many variations on its architecture, as well as its wide use, is an emerging field.

This article analyzes the diverse forms of the U-Net convolutional neural network architecture in the domain of biomedical image processing. In this article, four research questions due to related work are presented, focusing on use and current trends. For conducting and reporting a systematic literature review, Preferred Reporting Items of Systematic Reviews and Meta-Analyses (PRISMA) are used, and a flow diagram is adopted [7].

This article is organized as follows: Section 2 centers on the base U-Net architecture. Section 3 provides a state-of-the-art systematic literature review based on U-Net architecture in the domain of biomedicine. Section 4 presents the methodology, as well as the criteria and query string. Data results are discussed in Section 4, including an overview of domains, datasets, and architectures. Lastly, Section 5 concludes the research, results, limitations, and future research directions.

2. U-NET ARCHITECTURE

The U-Net architecture was first launched in 2015 with the main purpose of segmentation in the field of biomedicine image processing [8]. Figure 1 represents the first and original representations of U-Net CNN with an encoder-decoder type architecture. The left part consists of contracting paths, which are responsible for extracting relevant features from the input image. On the other hand, the right part of the architecture, the decoder, is responsible for upsampling and also finalizing the output.

The encoder consists of repeated 3x3 convolutions followed by a rectified linear unit (ReLU) layer. Between stages in the encoder, there is a max pooling 2x2 operation to downsample the features of the input image. After max-pooling, the dimensions of the features are reduced. After each max-pooling operation, channels are doubled. The decoder also consists of repeated 3x3 convolutions followed by a (ReLU) layer. The point reversing part, for upsampling a 2x2 convolution layer, is used to restore the spatial resolution of the features that were downsampled during the encoding part. Connection paths are responsible for taking copies of the feature from the symmetrical part of the encoder and concatenating them into their stage in the decoder. The encoder's feature is based on more spatial information, regarding pixels and position, while on the other hand, the decoder's feature can include more semantic information, such as the area and the object created. At the very bottom of the U-Net architecture, channels are doubled after the max-pooling operation, which is followed by upsampling with 2x2 convolutions.

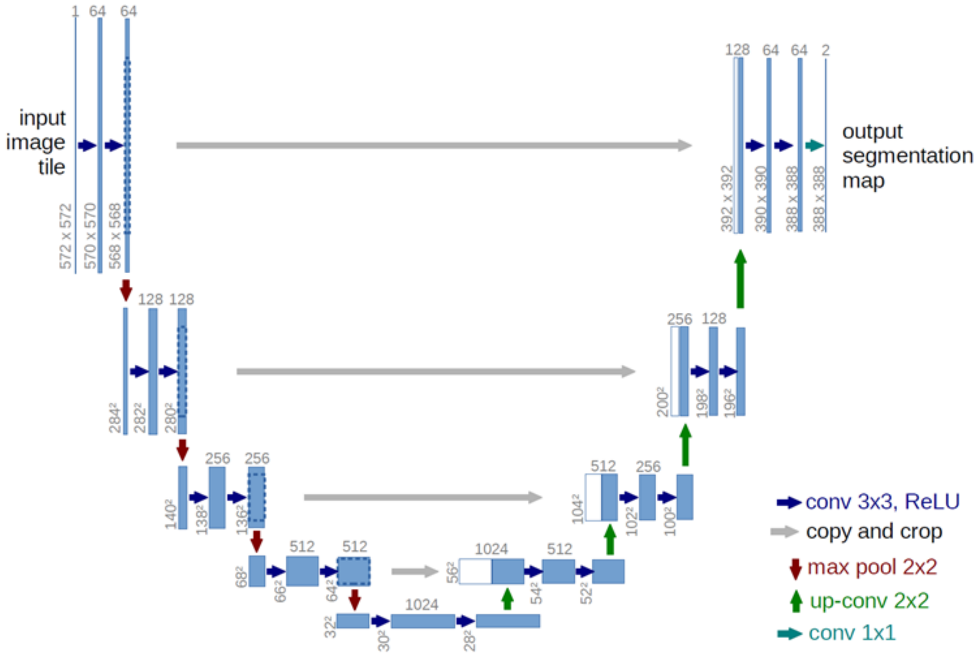

Figure 1. *Base U-Net architecture. [8]*

3. RELATED WORKS

Over time, the U-Net-based architecture has undergone modifications, leading to the emergence of various variants. These adaptations cater to specific research goals, datasets, biomedical domains, and broader purposes such as segmentation. The authors have extended and customized convolutional neural networks based on the U-Net architecture to address the diverse needs within the field of biomedicine. To explain further, Nahian Siddique et al. [9] proposed a review of variants of U-Net for medical image segmentation. According to their research findings, the number of papers on this topic has exhibited a consistent increase over time, comprising 135 papers focused on the MR image modality and 86 papers on CT. Furthermore, the predominant application areas were the brain organ, with 82 papers, followed by pathology, encompassing 55 papers, and cardiovascular applications, with 44 papers. The survey was specifically confined to publications from three prominent publishers: IEEE, Springer, and Elsevier, within the timeframe spanning 2017 to 2020.

They [9] introduced 3D U-Net as the network that uses 3D convolutions, 3D max-pooling, and 3D up-convolutions that result in 3D images and enable a faster training process mainly used in volumetric CT and MRI segmentation. Furthermore, they summed up Attention U-Net as the architecture that trims the features that are not important and relevant to the task with the pros of providing localized classification information. Additionally, the authors pointed to Inception U-Net as the solution for variations in the image's shape and size. Due to performance degradation, Residual U-Net uses a skip connection to take one layer and add it to a different, deeper layer, but it does not resolve the problem of vanishing the gradients. On the other hand, Recurrent U-Net is known for recurrent connection that relies on changing the output's node based on the previous output for the same node. Dense U-Net is characterized by every layer receiving a feature or map from all preceding layers. In U-Net++, each skip connection unit gathers feature maps from all preceding units at the same level, along with an upsampled feature map from the unit directly below it.

Soroush Baseri Saadi et al. [10] highlighted in their literature review that U-Net is one of the convolutional neural networks most efficient for detecting and analyzing osteolytic lesions, with a recognition rate above 81%. Furthermore, the authors involved Seg-Unet as the architecture for detecting knee bone tumors from X-ray images, with a mean classification accuracy of 99.05% in their original research [11].

Furthermore, Zeeshan Shaukat et al. [12] directed their attention to glioma and brain tumor segmentation, employing a 3D U-Net for semantic segmentation. The research landscape in this domain has expanded significantly, with Google Scholar indexing over 3020 papers related to brain tumor segmentation, while PubMed has documented 1495 papers spanning the years 2017 to 2021.

On the other hand, the literature review in the field of External Beam Radiation Therapy (EBRT) and prostate cancer was conducted by Bruno Mendes et al. [13]. They identified 16 papers from 2015 to February 2022. Most techniques are applied to MRI, while six of them are CT-based. They pointed to the smallest dataset with 11 patients and the largest with 2226 patients. The main architecture was based on U-Net, followed by one paper based on V-Net, R-CNN U-Net, ProstAttention-Net, U^2-Net, and DAUNet. They obtained the results that show the highest Dice Similarity Coefficient (DSC) in research [14] with a 0.94 score in the 3D approach. The authors claim 3D U-Net, Adversial U-Net, and Attention U-Bet as the main variants.

In the research [15], the authors point out 2.5D U-Net, 3D U-Net, Context Nested U-Net, all connection U-Net, RU-Net, and VGG16 U-Net as the main variants. They review performances and features in the context of segmentation disadvantages. For instance, the 2.5D U-Net does not obtain as good accuracy as the 3D U-Net, while the 3D U-Net obtains dependencies between picture slices. Despite U-Net being renowned for addressing overfitting issues, the proposed VGG16 U-Net method exhibits challenges related to overfitting.

4. MATERIALS AND METHODS

This systematic literature review is conducted based on the PRISMA framework [7]. According to PRISMA, the systematic literature review is based on three phases: 1. identification, 2. screening, and 3. including papers. This systematic literature review was performed using the Scopus database.

4.1. Search Strategy

Before the first phase, we form variations of keywords to search the scientific database. Based on related work mentioned in Section 3, we used "U-Net" as the main phrase and its variants. Synonyms used for the term "U-Net" were "U-Net++" OR "3D U-Net" OR "Adversarial U-Net" OR "Inception U-Net" OR "Residual U-Net" OR "Dense U-Net" OR "U-Net++". There are limitations in this part of the query, which refers to all variants and names, but a current overview of the state in the area revealed that these architectures are the most prominent. On the other hand, the names of the CNN architects have just been revealed. For the second part of the query, terms "biomedicine" OR "biomedical" are used. The final query for searching the Scopus database is TITLE-ABS-KEY (("u-net" OR "u-net++" OR "3d u-net" OR "adversarial u-net" OR '3d u-net" OR "inception u-net" OR "residual u-net" OR "dense u-net" OR "u-net++") AND ("biomedicine" OR "biomedical")).

4.2. Research Questions

To conduct a concise systematic literature review, we proposed three research questions as follows:

RQ1: Which U-Net architectures are used in biomedicine image processing?

RQ2: Which human organs are obtained in the studies?

RQ3: Which datasets are used to perform and evaluate the results of the studies?

RQ3.1. What are the best Dice scores for specific architectures and their datasets?

4.3. Study Selection Criteria

Due to the related work, the period for the papers that were taken into account is between 2018 and 2023. In the following text, we will present inclusion and exclusion criteria for this systematic literature review based on research questions.

IC1 – The main objective of the paper must be related to U-Net architecture in biomedicine related to humans

IC2 – The paper must be an article or conference paper

IC3 – Papers published between 2018 and 2023

IC4 – The paper must be written in English

IC5 – Only papers that authors managed to access should be accepted

IC6 – Datasets must be publicly available
IC7 – There must be a defined domain of the study in the context of organ/tissue
IC8 – The focus of the study must be U-Net architecture
IC9 – The paper must be in the domain of image processing Exclusion criteria:
EC1 – Papers that are not related to the research questions
EC2 – Duplicate papers
EC3 – Demonstrations, preliminary studies, technical reports, posters, and proof-of-concept papers were excluded
EC4 – Theoretical papers

4.4. PRISMA Workflow

The PRISMA workflow conducted in this review is shown in Figure 2. The query result was 667 papers. Before the process of screening, 3 papers were excluded based on the year, 40 based on paper type, 11 on language, and 383 based on access. In the second phase, data was extracted from each paper: authors, title, the year when the paper was published, link to Scopus, DOI, abstract, author keywords, publisher, document type, and country.

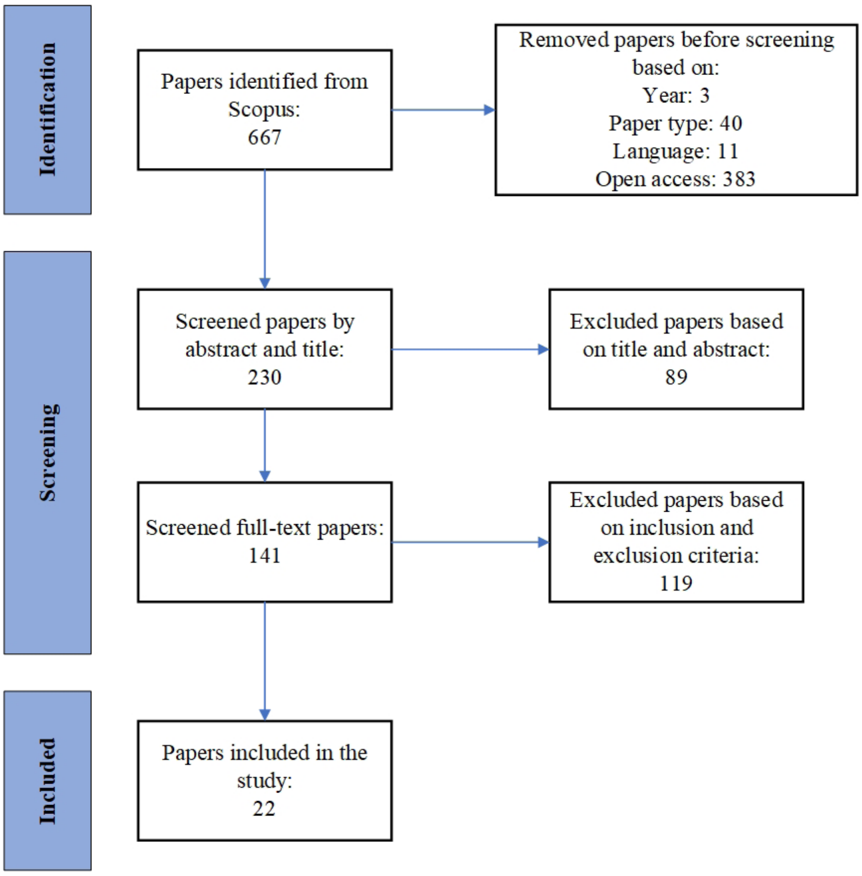

Figure 2. *PRISMA workflow.*

In the phase of screening, 230 papers were included, where 89 of them were excluded as nonrelevant based on title and abstract. For the second part of the screening, 89 papers were excluded based on inclusion and exclusion criteria. The main reason for excluding the papers was IC6.

5. RESULTS AND DISCUSSION

In the concluding phase of the study, a total of 22 papers have been incorporated. Only three of them are conference papers, while the rest are articles from journals. In 2023, 13 papers are included in the study, followed by 2 in 2022, 5 in 2021, and 3 in 2020. To answer **RQ1** and **RQ2,** we conducted Table 1. Based on common properties, we grouped the results into 7 categories as follows: brain, lung, liver, nuclei, breast, skin, heart and veins, polyp, and gland. The predominant focus among the papers lies in U-Net architectures, with a significant emphasis on brain tumor segmentation, encompassing the highest number of publications – specifically, 8 papers. Lung, liver, and nuclei are the focus of five papers. From this perspective, categories such as glands, hearts, and veins are relatively less explored in the domain of biomedicine image processing.

Table 1. *Overview of U-Net architectures based on category.*

Category	U-Net Architectures	Papers
Brain	SAB-Net, EMED-Unet, GAU-Net, U-Net++, MILD-Net, SD-Unet, KiU-Net, 3D U-Net	[16], [17], [18], [19], [20], [21], [22], [23]
Lung	U-Net11, EMED-Unet, SMR-Unet, GA-Unet, Sharp U-Net, Recurrent Residual 3D U-Net	[18], [24], [25], [26], [27]
Liver	U-Net, SAB-Net, GA-Unet, 3 D RP-Unet, ELU-Net	[28], [17], [25], [29], [30],
Nuclei	U-Net11, EG-TransUNet, FRUNet, GA-Unet, Sharp U-Net	[31], [32], [33], [25], [26]
Breast	U-Net11, EG-TransUNet, ResUNet, Dense UNet, DUNet, Attention U-Net, UNet++, MultiResUNet, RAUNet, Inception U-Net and U-Net GAN, ELU-Net	[31], [34], [32], [30]
Skin	MAAU, Residual Attention U-Net, Sharp U-Net, MOLD-Net	[35], [36], [26], [21]
Heart and veins	U-Net, SAB-Net	[28], [17]
Polyp	EG-TransUNet, SAB-Net, Sharp U-Net	[32], [17], [26]
Gland	FRUNet, Spatial-Channel Attention U-Net	[33], [37]

Priscilla Benedetti et al. [28] use the base structure of U-Net in the category of liver and veins segmentation, while Hasib Zunair and A. Ben Hamza [26] modify the base U-Net with a sharpening filter to sharpen the encoder features before their fusion with the decoder features by incorporating a depthwise convolution. Furthermore, for minor modification, Umut Tatli and Cafer Budak [31] use U-Net with 11 convolutional layers to improve the general performance of segmentation. Therefore, Neil Micalleef et al. [20] use U-Net++ variation based on additional upsampling layers along the skip connections between the encoder and decoder for brain tumor segmentation. On the other hand, Weihao Weng et al. [17] propose SAB-Net that represents a smooth attention branch regarding modified attention operation.

Saleh Naif Almuayqil et al. [36] developed a framework with specific noise reduction in the Residual Attention U-Net model. The authors use a variational autoencoder for removing the hair noises, DGAN-Net, D-U-NET, and Br-U-NET for removing the speckle noise, and the Laplacian Vector Median Filter MLVMF. Fu Zou et al. [33] introduce FRU-Net, which integrates the strengths of Fourier channel attention (FCA Block) and Residual units to enhance accuracy. The FCA Block dynamically allocates weights to learned frequency information, emphasizing precise high-frequency details in various biomedical images for segmentation based on different types of datasets.

For volumetric data, the authors transform the base U-Net architecture into 3D variants. Dhaval D. Kadia et al. [27] introduced Recurrent Residual 3D U-Net with Squeeze-and-Excitation Residual modules, Soft-DSC and Exponential Logarithmic Loss functions, and the Adam optimizer for processing volumetric data. On the other hand, Vanda Czipczer and Andrea Manno-Kovacs [29] modified U-Net with an upsampling operation to expand the path to reduce the number of total parameters. Additionally, the authors introduce ResPath instead of skip connection to enable propagating spatial information loss during the pooling operation. Rongsheng Liu et al. [16] focus on integrating serial and parallel multi-attention modules in a 3D U-Net, with a specific emphasis on scale attention and dual attention modules. Those three architectures are used for lung, heart, and brain tumor segmentation. Based on a high number of parameters, Floating-Point Operations Per Second (FLOPS), and not static salient regions, Kashish D. Shah [18] et al. propose an Efficient Multi-Encoder-Decoder (EMED-UNet).

Sudarshan Saikia et al. [34] introduced an empirical study for the segmentation of breast cancer using variations. The authors conclude that ResUNet has the best score, followed by base U-Net, MulitiResUNet, and Attention U-Net. The loss function used in their study is binary cross-entropy, with the same optimizer – Adam, and a learning rate of 0.01. U-Net GAN used the highest number of epochs – 1150, but on the other hand, the lowest number for epoch was 85 for Inception U-Net. Jiachen Hou et al. [24] propose SMR-UNet for lung nodule segmentation that integrates attention, multi-scale features, and residual structures. It replaces U-Net's convolutional units with residual units for faster convergence, incorporates a Transformer for improved global modeling, utilizes PixelShuffle to restore detailed information, and includes a multi-scale feature fusion module to enlarge the receptive field before upsampling. Additionally, GAU-Net, a Global Attention Mechanism that combines channel attention module and spatial attention module and integrates different convolutions used on a brain tumor segmentation dataset, is research in [19].

Phuong Thi Le et al. [35] introduce the mobile anti-aliasing attention U-Net model (MAAU). The model's encoder includes an anti-aliasing layer and convolutional blocks to decrease the spatial resolution of input images while preventing shift equivariance. On the other hand, the decoder utilizes an attention block and decoder module to capture significant features in each channel. This model is used on the skin lesion segmentation dataset. Furthermore, Shaoming Pan et al. [32] use EG-TransUNet on various datasets. The authors emphasized challenges related to extracting features and integrating spatial and semantic information within the encoder process. Consequently, they enhanced EG-TransUNet by introducing three transformer-based modules: the progressive enhancement module, channel spatial attention, and semantic guidance attention.

On the other hand, Mohammed Khouy et al. [25] introduce genetic algorithms for base U-Net architecture to minimize the complexity and evaluate it on three different datasets. Furthermore, in [30], the base U-Net is modified with a deep skip connection that includes the same and large scale from the encoder to extract features of the encoder in ELU-Net. Stripped-Down UNet (SD-UNet) presented in [22] has five blocks of encoding as well as decoding. Both blocks incorporate depthwise separable convolution layers, SD-UNet blocks, dropout layers for regularization, and max-pooling layers.

Naga Raju Gudhe et al. [21] present MILD-Net, a multi-level dilated residual network. This architecture integrates non-linear multi-level residual blocks into skip connections. The primary aim is to address the semantic gap by introducing these blocks, which help restore information lost during the concatenation of features from the encoder to the decoder units. On the other hand, the Spatial-Channel Attention U-Net (SCAU-Net) adopts a symmetrical structure with an encoder-decoder style [37]. It incorporates spatial and channel attention as modular components, allowing them to be seamlessly integrated into the architecture. The primary objective is to improve the significance of locally related features while suppressing irrelevant features at both spatial and channel levels.

To answer **RQ3** and **RQ3.1,** we conducted Table 2. The table consists of publicly available datasets that are used to evaluate different U-Net-based architectures and Dice similarity coefficient scores. The most frequently used datasets for evaluating the U-Net-based architectures are from the BraTS and Data Science Bowl group, followed by the ISIC Challenge and GlaS. The highest Dice score – 0.82 for the BraTS datasets group has SD-UNet architecture. For the DataScience Bowl, MILDNet and U-Net11 have the highest Dice score of 0.95. For the 3D-IRCADb-01 dataset, base U-Net has a Dice score of 0.97. Four U-Net-based architectures were evaluated on the ISIC Challenge dataset and MILDNet achieved the highest score of 0.94. Additionally, SAB-Net was evaluated on four different datasets, with the highest Dice score of 0.78 on the Liver CT dataset.

Table 2. *Overview of datasets and Dice scores.*

Dataset	Paper	Used U-Net Architecture	Dice
BraTS 2018, 2019 [38]	[16]	3D U-Net	0.79
	[18]	EMED-UNet	0.74
	[30]	ELU-Net	0.81
	[19]	GAU-Net	0.76
	[20]	U-Net++	0.69
	[22]	**SD-UNet**	**0.82**
3D-IRCADb-01 [39]	[28]	**U-Net based**	**0.97**
	[29]	3D RP-UNet	0.94
Data Science Bowl 2018 [40]	[31]	**U-Net11**	**0.95 lung** 0.79 nuclei 0.69 breast cancer
	[26]	Sharp U-Net	0.91
	[32]	EG-TransUNet	0.93
	[33]	FRUNet	0.92
	[25]	GA-UNet	0.90
	[21]	**MILDNet**	**0.95**
TCGA-BRCA [41]	[34]	Best is RAUNet	0.85 breast cancer
ISIC Challenge 2017, 2018, 2019 [42]	[35]	MAAU-Net	0.88 skin melanoma
	[26]	Sharp U-Net	0.84
	[21]	**MILDNet**	**0.94**
	[32]	EG-TransUNet	0.90
PH^2 [43]	[36]	Residual Attention U-Net	0.94
GlaS [44]	[33]	FRUNet	0.84
	[32]	**EG-TransUNet**	**0.90**
	[23]	KiU-Net	0.83
Heart MRI [45]	[17]	SAB-Net	0.57
Liver CT [46]			0.78
Spleen CT [45]			0.74
Colonoscopy [47]			0.61
LIDC [48]	[24]	SMR-UNet	0.91
The Kaggle Lung [49]	[25]	GA-UNet	0.98
The Kaggle Liver [50]			0.97
CVC-ClinicDB [51]	[32]	EG-TransUNet	0.95
	[26]	Sharp U-Net	0.83
MICCAI Sliver07 [52]		3D RP-UNet	0.94
LUNA16 [53]	[27]	Recurrent Residual 3D U-Net	0.98
VESSEL12 [54]			0.98
CRAG [55]	[37]	SCAU-Net	0.91

The general highest Dice score – 0.98, was obtained for the VESSEL12, Kaggle Lung, and LUNA16 datasets. The same type of U-Net architecture, Recurrent Residual 3D U-Net was evaluated on the LUNA16 and VESSEL12 datasets. Furthermore, GA-UNet was evaluated on the Kaggle Lung dataset.

In their review from 2022, Reza Azad et al. [56] proposed various forms of U-Net and some parts of popular datasets to provide a model library for future research. Since the convolutional neural network U-Net architecture is constantly evolving, the focus of this literature review is the time frame from 2018 to 2023. In this paper, based on the research questions of the U-Net architecture and its variants, specific groups of organs and data sets are correlated. In order to gain complete insight, the systematic review follows strict inclusion and exclusion criteria for the inclusion of papers with the aim of being able to compare architectures, datasets, and results.

6. CONCLUSIONS

The use of U-Net convolutional neural networks in the field of biomedicine is extremely important for the identification and diagnosis of various health conditions. This literature review examines the development of a base U-Net architecture specified for certain emerging challenges. Different variants solve the problems of noise reduction, complexity, processing time, small datasets, low feature extraction, and image quality preservation. Furthermore, the number of variants related to complex images, with a focus on 3D modalities, is growing. The most frequently used datasets are the areas of semantic segmentation of brain, lung, liver, and breast tumors. On the other hand, the door related to segmentation at the level of the cell nucleus is open. The limitations of this review refer to the scientific database, Scopus.

In further work, it should be proposed to expand the scientific bases during the literature review, as well as evaluate the U-Net network on different datasets, and find patterns for the best Dice score. The development of this architecture should refer to specialized parts of modalities and medicine, due to the specificity of the challenges that appear. Furthermore, future steps may be to search for challenges for U-Net architecture.

FUNDING

This research received no external funding.

INSTITUTIONAL REVIEW BOARD STATEMENT:

Not applicable.

INFORMED CONSENT STATEMENT:

Not applicable.

CONFLICTS OF INTEREST:

The author declares no conflict of interest.

REFERENCES

[1] A. P. Dhawan, "A review on biomedical image processing and future trends," *Computer Methods and Programs in Biomedicine*, vol. 31, no. 3–4, pp. 141–183, Mar. 1990, doi: 10.1016/0169-2607(90)90001-P.

[2] P. Mamoshina, A. Vieira, E. Putin, and A. Zhavoronkov, "Applications of Deep Learning in Biomedicine," *Mol. Pharmaceutics*, vol. 13, no. 5, pp. 1445–1454, May 2016, doi: 10.1021/acs.molpharmaceut.5b00982.

[3] A. O. A. Deheyab *et al.*, "AN OVERVIEW OF CHALLENGES IN MEDICAL IMAGE PROCESSING," in *Proceedings of the 6th International Conference on Future Networks & Distributed Systems*, Tashkent TAS Uzbekistan: ACM, Dec. 2022, pp. 511–516. doi: 10.1145/3584202.3584278.

[4] P. Chinmayi, L. Agilandeeswari, and M. Prabukumar, "Survey of Image Processing Techniques in Medical Image Analysis: Challenges and Methodologies," in *Proceedings of the Eighth International Conference on Soft Computing and Pattern Recognition (SoCPaR 2016)*, vol. 614, A. Abraham, A. K. Cherukuri, A. M. Madureira, and A. K. Muda, Eds., in Advances in Intelligent Systems and Computing, vol. 614, Cham: Springer International Publishing, 2018, pp. 460–471. doi: 10.1007/978-3-319-60618-7_45.

[5] K. Grünberg *et al.*, "Ethical and Privacy Aspects of Using Medical Image Data," in *Cloud-Based Benchmarking of Medical Image Analysis*, A. Hanbury, H. Müller, and G. Langs, Eds., Cham: Springer International Publishing, 2017, pp. 33–43. doi: 10.1007/978-3-319-49644-3_3.

[6] A. Z. Mustaqim, N. A. Fadil, and D. A. Tyas, "Artificial Neural Network for Classification Task in Tabular Datasets and Image Processing: A Systematic Literature Review," *join*, vol. 8, no. 2, pp. 158–168, Dec. 2023, doi: 10.15575/join.v8i2.1002.

[7] M. J. Page *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, p. n71, Mar. 2021, doi: 10.1136/bmj.n71.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, vol. 9351, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., in Lecture Notes in Computer Science, vol. 9351, Cham: Springer International Publishing, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[9] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications," *IEEE Access*, vol. 9, pp. 82031–82057, 2021, doi: 10.1109/ACCESS.2021.3086020.

[10] S. B. Saadi *et al.*, "Osteolysis: A Literature Review of Basic Science and Potential Computer-Based Image Processing Detection Methods," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–21, Oct. 2021, doi: 10.1155/2021/4196241.

[11] N.-T. Do, S.-T. Jung, H.-J. Yang, and S.-H. Kim, "Multi-Level Seg-Unet Model with Global and Patch-Based X-ray Images for Knee Bone Tumor Detection," *Diagnostics*, vol. 11, no. 4, p. 691, Apr. 2021, doi: 10.3390/diagnostics11040691.

[12] Z. Shaukat, Q. U. A. Farooq, S. Tu, C. Xiao, and S. Ali, "A state-of-the-art technique to perform cloud-based semantic segmentation using deep learning 3D U-Net architecture," *BMC Bioinformatics*, vol. 23, no. 1, p. 251, Dec. 2022, doi: 10.1186/s12859-022-04794-9.

[13] B. Mendes, I. Domingues, and J. Santos, "U-Net Architectures for Prostate Cancer Radiation Therapy: A Literature Review," *UPjeng*, vol. 9, no. 1, pp. 177–190, Jan. 2023, doi: 10.24840/2183-6493_009-001_001551.

[14] A. Meyer *et al.*, "Anisotropic 3D Multi-Stream CNN for Accurate Prostate Segmentation from Multi-Planar MRI," *Computer Methods and Programs in Biomedicine*, vol. 200, p. 105821, Mar. 2021, doi: 10.1016/j.cmpb.2020.105821.

[15] J. Shao, S. Chen, J. Zhou, H. Zhu, Z. Wang, and M. Brown, "Application of U-Net and Optimized Clustering in Medical Image Segmentation: A Review," *Computer Modeling in Engineering & Sciences*, vol. 136, no. 3, pp. 2173–2219, 2023, doi: 10.32604/cmes.2023.025499.

[16] R. Liu, X. Liu, C. Peng, A. Li, and Y. Liao, "Automatic Brain Tumour Subregion Segmentation from Multimodal MRIs Fusing Muti-channel and Spatial Features," *J. Phys.: Conf. Ser.*, vol. 2449, no. 1, p. 012034, Mar. 2023, doi: 10.1088/1742-6596/2449/1/012034.

[17] W. Weng, X. Zhu, L. Jing, and M. Dong, "Attention Mechanism Trained with Small Datasets for Biomedical Image Segmentation," *Electronics*, vol. 12, no. 3, p. 682, Jan. 2023, doi: 10.3390/electronics12030682.

[18] K. D. Shah, D. K. Patel, M. P. Thaker, H. A. Patel, M. J. Saikia, and B. J. Ranger, "EMED-UNet: An Efficient Multi-Encoder-Decoder Based UNet for Medical Image Segmentation," *IEEE Access*, vol. 11, pp. 95253–95266, 2023, doi: 10.1109/ACCESS.2023.3309158.

[19] X. Gan, L. Wang, Q. Chen, Y. Ge, and S. Duan, "GAU-Net: U-Net Based on Global Attention Mechanism for brain tumor segmentation," *J. Phys.: Conf. Ser.*, vol. 1861, no. 1, p. 012041, Mar. 2021, doi: 10.1088/1742-6596/1861/1/012041.

[20] N. Micallef, D. Seychell, and C. J. Bajada, "Exploring the U-Net++ Model for Automatic Brain Tumor Segmentation," *IEEE Access*, vol. 9, pp. 125523–125539, 2021, doi: 10.1109/ACCESS.2021.3111131.

[21] N. R. Gudhe *et al.*, "Multi-level dilated residual network for biomedical image segmentation," *Sci Rep*, vol. 11, no. 1, p. 14105, Jul. 2021, doi: 10.1038/s41598-021-93169-w.

[22] P. K. Gadosey *et al.*, "SD-UNet: Stripping down U-Net for Segmentation of Biomedical Images on Platforms with Low Computational Budgets," *Diagnostics*, vol. 10, no. 2, p. 110, Feb. 2020, doi: 10.3390/diagnostics10020110.

[23] J. M. Jose, V. Sindagi, I. Hacihaliloglu, and V. M. Patel, "KiU-Net: Towards Accurate Segmentation of Biomedical Images using Over-complete Representations." arXiv, Jul. 08, 2020. Accessed: Jan. 12, 2024. [Online]. Available: http://arxiv.org/abs/2006.04878

[24] J. Hou, C. Yan, R. Li, Q. Huang, X. Fan, and F. Lin, "Lung Nodule Segmentation Algorithm With SMR-UNet," *IEEE Access*, vol. 11, pp. 34319–34331, 2023, doi: 10.1109/ACCESS.2023.3264789.

[25] M. Khouy, Y. Jabrane, M. Ameur, and A. Hajjam El Hassani, "Medical Image Segmentation Using Automatic Optimized U-Net Architecture Based on Genetic Algorithm," *JPM*, vol. 13, no. 9, p. 1298, Aug. 2023, doi: 10.3390/jpm13091298.

[26] H. Zunair and A. Ben Hamza, "Sharp U-Net: Depthwise convolutional network for biomedical image segmentation," *Computers in Biology and Medicine*, vol. 136, p. 104699, Sep. 2021, doi: 10.1016/j.compbiomed.2021.104699.

[27] D. D. Kadia, M. Z. Alom, R. Burada, T. V. Nguyen, and V. K. Asari, "R ^{2} U3D: Recurrent Residual 3D U-Net for Lung Segmentation," *IEEE Access*, vol. 9, pp. 88835–88843, 2021, doi: 10.1109/ACCESS.2021.3089704.

[28] P. Benedetti, M. Femminella, and G. Reali, "Mixed-Sized Biomedical Image Segmentation Based on U-Net Architectures," *Applied Sciences*, vol. 13, no. 1, p. 329, Dec. 2022, doi: 10.3390/app13010329.

[29] V. Czipczer and A. Manno-Kovacs, "Adaptable volumetric liver segmentation model for CT images using region-based features and convolutional neural network," *Neurocomputing*, vol. 505, pp. 388–401, Sep. 2022, doi: 10.1016/j.neucom.2022.07.024.

[30] Y. Deng, Y. Hou, J. Yan, and D. Zeng, "ELU-Net: An Efficient and Lightweight U-Net for Medical Image Segmentation," *IEEE Access*, vol. 10, pp. 35932–35941, 2022, doi: 10.1109/ACCESS.2022.3163711.

[31] U. Tatli and C. Budak, "Biomedical Image Segmentation with Modified U-Net," *TS*, vol. 40, no. 2, pp. 523–531, Apr. 2023, doi: 10.18280/ts.400211.

[32] S. Pan, X. Liu, N. Xie, and Y. Chong, "EG-TransUNet: a transformer-based U-Net with enhanced and guided models for biomedical image segmentation," *BMC Bioinformatics*, vol. 24, no. 1, p. 85, Mar. 2023, doi: 10.1186/s12859-023-05196-1.

[33] F. Zou, Y. Liu, Z. Chen, K. Zhanghao, and D. Jin, "Fourier Channel Attention Powered Lightweight Network for Image Segmentation," *IEEE J. Transl. Eng. Health Med.*, vol. 11, pp. 252–260, 2023, doi: 10.1109/JTEHM.2023.3262841.

[34] S. Saikia *et al.*, "Lesion detection in women breast's dynamic contrast-enhanced magnetic resonance imaging using deep learning," *Sci Rep*, vol. 13, no. 1, p. 22555, Dec. 2023, doi: 10.1038/s41598-023-48553-z.

[35] P. T. Le *et al.*, "Anti-Aliasing Attention U-net Model for Skin Lesion Segmentation," *Diagnostics*, vol. 13, no. 8, p. 1460, Apr. 2023, doi: 10.3390/diagnostics13081460.

[36] S. Naif Almuayqil, R. Arnous, N. Sakr, and M. M. Fadel, "A New Hybrid Model for Segmentation of the Skin Lesion Based on Residual Attention U-Net," *Computers, Materials & Continua*, vol. 75, no. 3, pp. 5177–5192, 2023, doi: 10.32604/cmc.2023.038625.

[37] P. Zhao, J. Zhang, W. Fang, and S. Deng, "SCAU-Net: Spatial-Channel Attention U-Net for Gland Segmentation," *Front. Bioeng. Biotechnol.*, vol. 8, p. 670, Jul. 2020, doi: 10.3389/fbioe.2020.00670.

[38] B. H. Menze *et al.*, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Trans. Med. Imaging*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: 10.1109/TMI.2014.2377694.

[39] “Liver segmentation – 3D-ircadb-01,” IRCAD. Accessed: Jan. 12, 2024. [Online]. Available: https://www.ircad.fr/research/data-sets/liver-segmentation-3d-ircadb-01/

[40] “2018 Data Science Bowl.” Accessed: Jan. 12, 2024. [Online]. Available: https://kaggle.com/competitions/data-science-bowl-2018

[41] “Projects.” Accessed: Jan. 12, 2024. [Online]. Available: https://portal.gdc.cancer.gov/projects/TCGA-BRCA

[42] N. Codella *et al.*, “Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC).” arXiv, Mar. 29, 2019. Accessed: Jan. 13, 2024. [Online]. Available: http://arxiv.org/abs/1902.03368

[43] “ADDI – Automatic computer-based Diagnosis system for Dermoscopy Images.” Accessed: Jan. 13, 2024. [Online]. Available: https://www.fc.up.pt/addi/ph2%20database.html

[44] K. Sirinukunwattana *et al.*, “Gland segmentation in colon histology images: The glas challenge contest,” *Medical Image Analysis*, vol. 35, pp. 489–502, Jan. 2017, doi: 10.1016/j.media.2016.08.008.

[45] A. L. Simpson *et al.*, “A large annotated medical image dataset for the development and evaluation of segmentation algorithms,” 2019, doi: 10.48550/ARXIV.1902.09063.

[46] P. Bilic *et al.*, “The Liver Tumor Segmentation Benchmark (LiTS),” 2019, doi: 10.48550/ARXIV.1901.04056.

[47] J. Bernal *et al.*, “Comparative Validation of Polyp Detection Methods in Video Colonoscopy: Results from the MICCAI 2015 Endoscopic Vision Challenge,” *IEEE Trans. Med. Imaging*, vol. 36, no. 6, pp. 1231–1249, Jun. 2017, doi: 10.1109/TMI.2017.2664042.

[48] S. G. Armato *et al.*, “The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans,” *Medical Physics*, vol. 38, no. 2, pp. 915–931, Feb. 2011, doi: 10.1118/1.3528204.

[49] “Finding and Measuring Lungs in CT Data.” Accessed: Jan. 13, 2024. [Online]. Available: https://www.kaggle.com/datasets/kmader/finding-lungs-in-ct-data

[50] “Liver Dataset.” Accessed: Jan. 13, 2024. [Online]. Available: https://www.kaggle.com/datasets/stevenazy/liver-dataset

[51] J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilariño, “WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians,” *Computerized Medical Imaging and Graphics*, vol. 43, pp. 99–111, Jul. 2015, doi: 10.1016/j.compmedimag.2015.02.007.

[52] T. Heimann *et al.*, “Comparison and Evaluation of Methods for Liver Segmentation from CT Datasets,” *IEEE Trans. Med. Imaging*, vol. 28, no. 8, pp. 1251–1265, Aug. 2009, doi: 10.1109/TMI.2009.2013851.

[53] “LUNA16 – Grand Challenge,” grand-challenge.org. Accessed: Jan. 13, 2024. [Online]. Available: https://luna16.grand-challenge.org/

[54] “VESSEL12 – Grand Challenge,” grand-challenge.org. Accessed: Jan. 13, 2024. [Online]. Available: https://vessel12.grand-challenge.org/

[55] S. Graham *et al.*, “MILD-Net: Minimal information loss dilated network for gland instance segmentation in colon histology images,” *Medical Image Analysis*, vol. 52, pp. 199–211, Feb. 2019, doi: 10.1016/j.media.2018.12.001.

[56] R. Azad *et al.*, “Medical Image Segmentation Review: The success of U-Net.” arXiv, Nov. 27, 2022. Accessed: Jan. 30, 2024. [Online]. Available: http://arxiv.org/abs/2211.14830