

Designing a High-Performance Deep Learning Theoretical Model for Biomedical Image Segmentation by Using Key Elements of the Latest U-Net-Based Architectures

Andreea Roxana Luca^{1,2}, Tudor Florin Ursuleanu^{1,3,4*}, Liliana Gheorghe^{1,5*}, Roxana Grigorovici¹, Stefan Iancu¹, Maria Hlusneac¹, Cristina Preda^{1,6}, Alexandru Grigorovici^{1,3}

¹Faculty of General Medicine, “Grigore T. Popa” University of Medicine and Pharmacy, Iasi, Romania

²Department Obstetrics and Gynecology, Integrated Ambulatory of Hospital “Sf. Spiridon”, Iasi, Romania

³Department of Surgery VI, “Sf. Spiridon” Hospital, Iasi, Romania

⁴Iasi, Department of Surgery I, Regional Institute of Oncology, Iasi, Romania

⁵Department of Radiology, “Sf. Spiridon” Hospital, Iasi, Romania

⁶Department of Endocrinology, “Sf. Spiridon” Hospital, Iasi, Romania

Email: *tudorursuleanu@yahoo.com, *lgheorghe123@gmail.com

How to cite this paper: Luca, A.R., Ursuleanu, T.F., Gheorghe, L., Grigorovici, R., Iancu, S., Hlusneac, M., Preda, C. and Grigorovici, A. (2021) Designing a High-Performance Deep Learning Theoretical Model for Biomedical Image Segmentation by Using Key Elements of the Latest U-Net-Based Architectures. *Journal of Computer and Communications*, 9, 8-20.

<https://doi.org/10.4236/jcc.2021.97002>

Received: May 17, 2021

Accepted: July 20, 2021

Published: July 23, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Deep learning (DL) has experienced an exponential development in recent years, with major impact in many medical fields, especially in the field of medical image and, respectively, as a specific task, in the segmentation of the medical image. We aim to create a computer assisted diagnostic method, optimized by the use of deep learning (DL) and validated by a randomized controlled clinical trial, is a highly automated tool for diagnosing and staging precancerous and cervical cancer and thyroid cancers. We aim to design a high-performance deep learning model, combined from convolutional neural network (U-Net)-based architectures, for segmentation of the medical image that is independent of the type of organs/tissues, dimensions or type of image (2D/3D) and to validate the DL model in a randomized, controlled clinical trial. We used as a methodology primarily the analysis of U-Net-based architectures to identify the key elements that we considered important in the design and optimization of the combined DL model, from the U-Net-based architectures, imagined by us. Secondly, we will validate the performance of the DL model through a randomized controlled clinical trial. The DL model designed by us will be a highly automated tool for diagnosing and staging precancers and cervical cancer and thyroid cancers. The combined model we designed takes into account the key features of each of the architectures Overcomplete Convolutional Network Kite-Net (Kite-Net), Attention gate

mechanism is an improvement added on convolutional network architecture for fast and precise segmentation of images (Attention U-Net), Harmony Densely Connected Network-Medical image Segmentation (HarDNet-MSEG). In this regard, we will create a comprehensive computer assisted diagnostic methodology validated by a randomized controlled clinical trial. The model will be a highly automated tool for diagnosing and staging precancers and cervical cancer and thyroid cancers. This would help drastically minimize the time and effort that specialists put into analyzing medical images, help to achieve a better therapeutic plan, and can provide a “second opinion” of computer assisted diagnosis.

Keywords

Combined Model of U-Net-Based Architectures, Medical Image Segmentation, 2D/3D/CT/RMN Images

1. Introduction

Deep learning (DL) has experienced an exponential development in recent years, with major impact in many medical fields, especially in the field of medical image and, respectively, as a specific task, in the segmentation of the medical image.

We aim to create a computer assisted diagnostic method, optimized by the use of deep learning (DL) and validated by a randomized controlled clinical trial, over a period of 17 months, is a highly automated tool for diagnosing and staging precancerous and cervical cancer and thyroid cancers that would drastically minimize the time and effort that specialists put in analyzing medical images and that makes the right tool as support in the diagnostic process specialists and to achieve a better therapeutic plan.

We aim to:

- Design a high-performance deep learning model, combined from convolutional neural network (U-Net)-based architectures, for segmentation of the medical image that is independent of the type of organs/tissues, dimensions, or type of image (2D/3D);
- To validate the DL model in a randomized controlled clinical trial over a period of 17 months.

DL architectures designed for diagnosis—segmentation of medical images, three categories can be exemplified:

- FCN-based models (fully convolutional network) [1] [2];
- Convolutional Neural Network (U-Net)-based models (convolutional neural network-images segmentation) [3];
- GAN-based models (generative adversarial network) [4].

FCN achieves goals of segmenting the medical image with good results [5]. Types of FCN: Cascading FCN [6], parallel FCN [7] and recurrent FCN [8] also achieve medical image segmentation goals with good results.

U-Net [3] and its derivatives segment the medical image with good results. U-Net is based on the FCN structure, consisting of a series of convolutional and deconvolution layers and with short connections between equal resolution layers. U-Net and its variants such as UNet++ [9] and recurrent U-Net [10] perform well in many medical image segmentation tasks [11] [12].

GAN is a type of mixed architecture (supervised and unsupervised) called semi-supervised architecture, architecture composed of two neural networks, a generator and a discriminator or classifier, which competition with each other in an adversarial formation process [4]. In models, the generator is used to predict the target mask based on encoder-decoder structures (such as FCN or U-Net) [12]. The discriminator serves as a form regulator that helps the generator achieve satisfactory segmentation results [13] [14]. GAN has been used in the generation of synthetic instances of different classes.

The main core of the solution of this task, segmentation of medical images, is the approach based on convolutional neural networks because they are ideal for capturing the structure in data.

A certain NNC architecture has proven particularly effective at segmentation, namely U-Net, a type of encoder-decoding network that reduces image feature dimensions, maps and then tries to accurately reconstruct the image to learn key (key) key features. However, the basic U-Net has some drawbacks and for this reason, many architectures have been built over U-Net to make it stronger. We will analyze some of the most interesting or recent U-Net-based architectures and make a synthesis of their key advantages based on the main features and their performance in segmenting the medical image, to have a starting point for the model development imagined by us.

2. Methodology

Next, we will present and describe (2.1.) the U-Net-based architectures and then we will present (2.2.) the key elements that we considered important in the design, optimization and validation of the combined DL model, from the U-Net-based architectures, imagined by us.

2.1. U-Net-Based Architectures

2.1.1. Attention U-Net

One of the architectures investigated is Attention U-Net, developed in 2018. Usually, for a segmentation task, there is only a part or a few parts of the image that are relevant for the problem. However, the basic U-Net is not capable of focusing on a specific region of interest, and that results in excessive processing of irrelevant areas.

The Attention U-Net architecture is visually provided in **Figure 1**. Both the image and its description are taken from the original paper [15].

Attention gate mechanism is an improvement added on U-Net which suppresses irrelevant regions and highlights key features that are useful for segmentation. Another advantage of attention gates is that they do not add significant

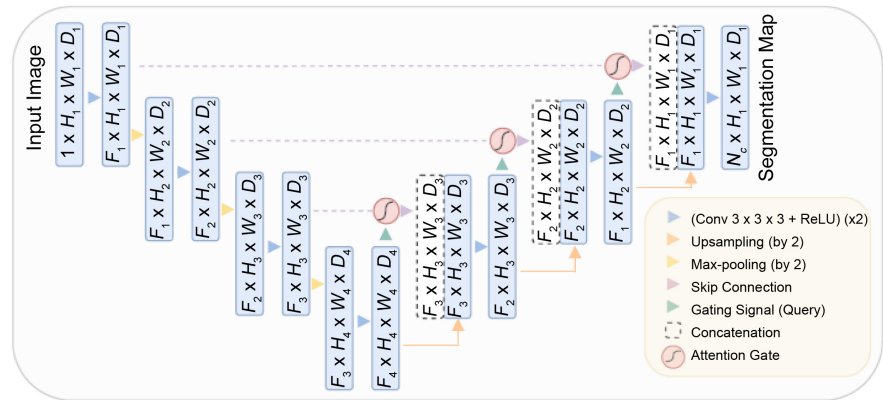


Figure 1. A block diagram of the proposed Attention U-Net segmentation model. Input image is progressively filtered and downsampled by factor of 2 at each scale in the encoding part of the network (e.g. $H_4 = H_1/8$). N_c denotes the number of classes. Attention gates (AGs) filter the features propagated through the skip connections. Feature selectivity in AGs is achieved by use of contextual information (gating) extracted in coarser scales.

computational overhead when integrated into U-Net. The authors of the architecture presented in [15] propose input features of each layer to be scaled by attention coefficients that are computed in the attention gate. Each pixel from the input has its own attention coefficient in order to enhance the important regions and suppress the irrelevant ones. The attention mechanism is shown in Figure 2. Both the schema and the description are taken entirely from the original paper [15].

The architecture has been trained on two 3D datasets, one for which the task was multi-class segmentation (pancreas, spleen, kidney) and another for one-class segmentation (only pancreas). Both datasets contain CT scans and can be found in the publicly available NIH-TCIA dataset. Table 1 below describes the performance of the network in terms of Dice score coefficient, in comparison with the classical U-Net.

2.1.2. KiU-Net

Classical U-Net performs poorly in detecting small structures and does not segment boundaries of regions precisely. This happens because the deeper we go in the layers of the network, the larger the receptive field is, and this results in reduced attention to details. A solution to this drawback came with the development of the KiU-Net in 2020 in [16]. The architecture consists of two networks, a Kite-Net and a U-Net that run in parallel having their results combined.

The Kite-Net can be thought of as the opposite of U-Net. While U-Net reduces the image dimensions in the encoder and reconstructs it in the decoder, the Kite-Net up samples the image in the encoder and reduces it back in the decoder. This way the receptive field will not increase in the deeper layers as in U-Net and hence, the desired fine details are obtained. Since Kite-Net alone is only focusing on extracting small structures and the dataset could have both large and small regions to be segmented, it has been put together with U-Net, which performs well at segmenting high-level features, *i.e.* large regions. Their outputs are

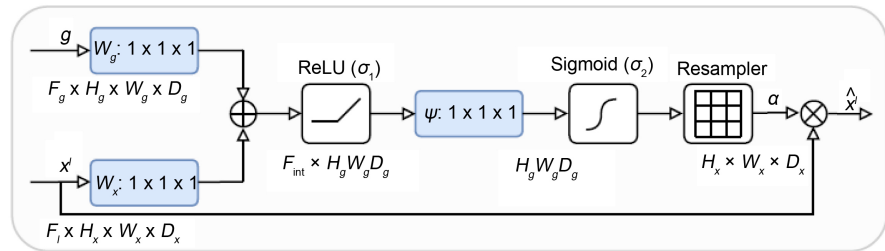


Figure 2. Schematic of the proposed additive attention gate (AG). Input features (x^l) are scaled with attention coefficients (α) computed in AG. Spatial regions are selected by analyzing both the activations and contextual information provided by the gating signal (g) which is collected from a coarser scale. Grid resampling of attention coefficients is done using trilinear interpolation.

Table 1. Comparison of the quantitative metric dice score coefficient for pancreas, spleen, kidney dataset and pancreas dataset between U-Net and Attention U-Net.

	U-Net	Attention U-Net	Number of CT scans in dataset
Pancreas CT (3D)	0.814	0.840	
Spleen CT (3D)	0.962	0.965	150
Kidney CT (3D)	0.963	0.964	
Pancreas 2 CT (3D)	0.820	0.831	82

concatenated matching their dimensions accordingly with the help of a Cross-Residual-Fusion-Block. **Figure 3** provides visual details regarding the architecture. The images and their descriptions have been taken entirely from the original paper [16].

The architecture has been trained on various datasets, both with 2D and 3D images, in order to prove the character of independency of the data type. The results are presented in **Table 2** below in terms of Dice score coefficient.

2.1.3. U-Net with Context Aggregation Blocks

Another improvement to the classical U-Net has been proposed in [17] and it consists of replacing some convolutional layers in the U-Net with Context Aggregation Blocks. These blocks contain dilated convolutional layers and normal convolutional layers. Dilation convolution helps detecting features in large receptive fields without increasing computational costs. However, this type of convolution has been reported to cause “gridding artefacts” which can affect model’s performance. In order to overcome this, the dilation convolutions are combined with normal convolutions and their output features are aggregated. The task performed in [17] is multi-class segmentation, since the images are multi-channelled—each channel corresponds to an organ. After applying the Context Aggregation Blocks, a Squeeze-Extract block is used to assign different weights to each channel for reweighting each organ mask importance. The Context Aggregation Block and the Squeeze-Extract block are depicted in **Figure 4** and **Figure 5**. Images and descriptions are taken entirely from the original paper [17].

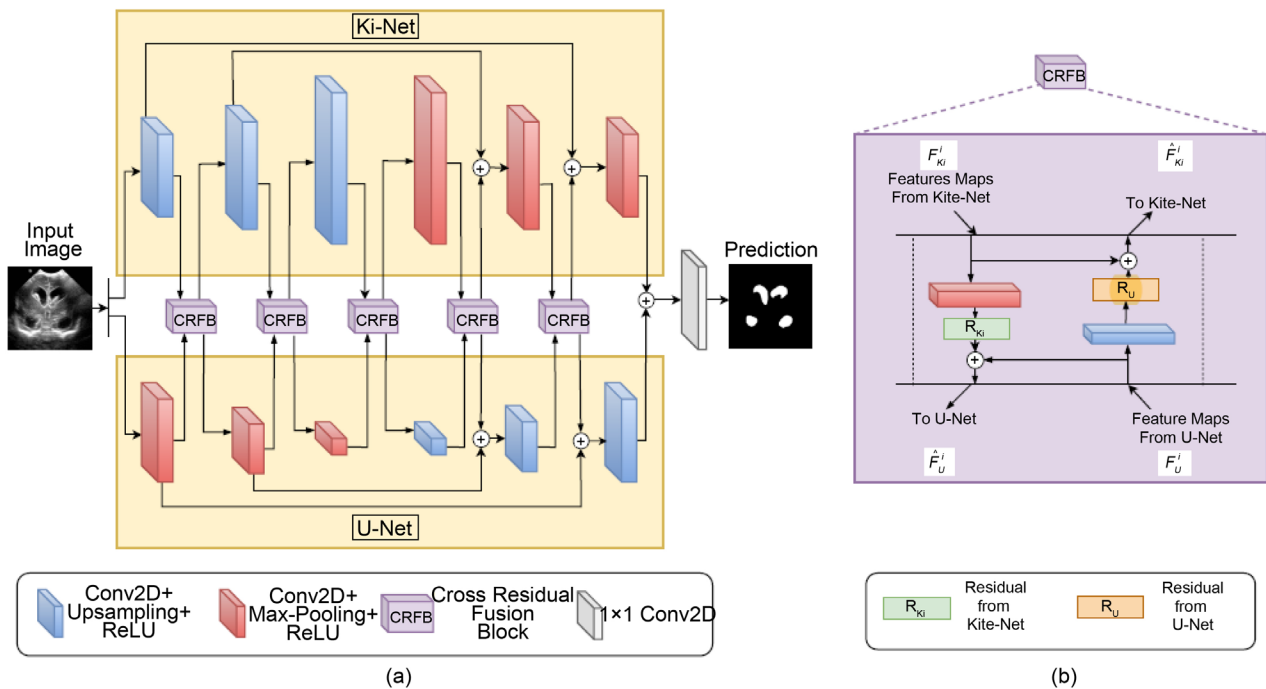


Figure 3. Architecture details of KiU-Net for 2D image segmentation. In KiU-Net, the input image is forwarded to the two branches of KiU-Net: Kite-Net and U-Net which have CRFB blocks connecting them at each level. The feature maps from the last layer of both the branches are added and passed through 1×1 2D conv to get the prediction. In CRFB, residual features of Kite-Net are learned and added to the features of U-Net to forward the complementary features and vice-versa. (b) Details of Cross Residual Fusion Block (CRFB).

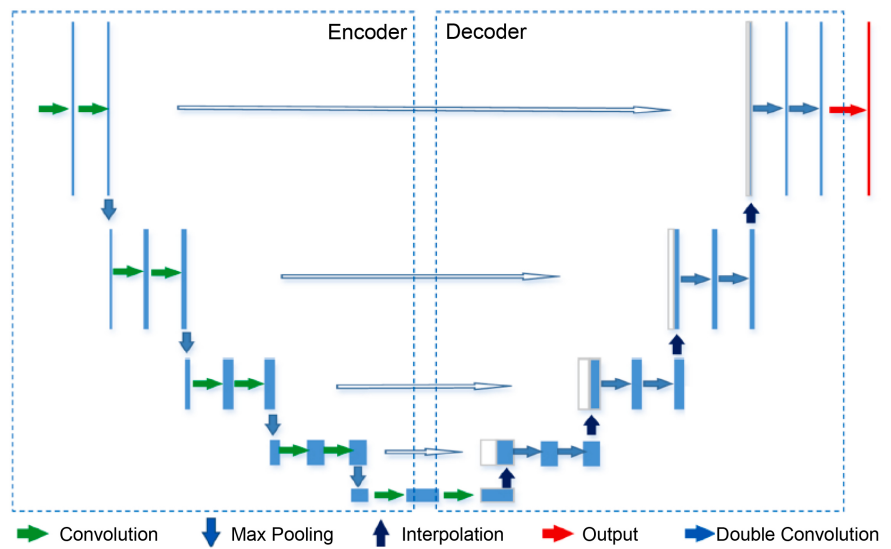


Figure 4. Squeeze-extract block used in the proposed model.

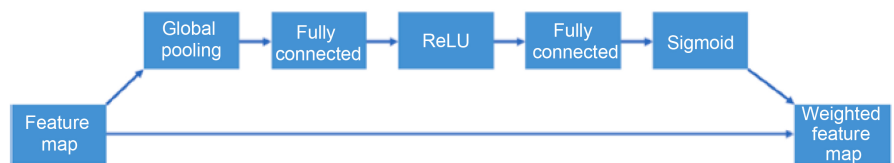


Figure 5. Overall architecture of the proposed model.

Table 2. Comparison of the quantitative metric dice score coefficient for Brain US, GLAS, RITE, BraTS and LiTS datasets between U-Net and KiU-Net.

	U-Net	KiU-Net	Number of 2D images/3D scans in dataset
Brain Anatomy Segmentation (US) (2D)	0.853	0.894	1629
Gland Segmentation (Microscopic) (2D)	0.797	0.832	165
Retinal Nerve Segmentation (Fundus) (2D)	0.552	0.751	40
Brain Tumor Segmentation MRI (3D)	0.808	0.876	335 (155 slices in each scan)
Liver segmentation CT (3D)	0.844	0.942	200

The architecture proposed has been trained on a dataset containing CT scans and the task consisted of multi-class segmentation (the parts segmented: bladder, bone marrow, femoral head left, femoral head right, rectum, small intestine, spinal cord). The results are shown in **Table 3** below in terms of Dice score coefficient.

2.1.4. HarDNet-MSEG

This architecture has been described in [18] and it has achieved the state-of-the-art on two datasets so far. It appeared as a solution to the failure of U-Net at segmenting small blurry areas, to the lack of coverage for broken image areas and the time-consuming training. It consists of a HarDNet encoder and a partial decoder, which reduces the training time. The encoder is based on a DenseNet but has significantly less connections for cutting computation costs and smaller channel width in order to recover the accuracy lost from connection pruning. The HarDNet block used in encoder is shown in **Figure 6**, as an evolution from DenseNet Block. The figure and its description are taken entirely from the original paper [18].

Talking about the decoder, the classical U-Net's decoder produces in the shallower layers high resolution low-level features, hence they require large computation costs. The good part is that the high-level features produced by the deeper layers also include a sort of low-level structure, so it followed that shallow layers from the encoder could be eliminated when connecting with the decoder's layers, resulting in a cascaded partial decoder. Its architecture is presented in **Figure 7**, which was entirely taken from [19].

The architecture has been trained on several datasets containing 2D colonoscopy images. **Table 4** below contains details about the results in terms of Dice coefficient score for classical U-Net and for HarDNet-MSEG. The scores on Kvasir-SEG and CVC-ClinicDB datasets are currently SOTA in biomedical image segmentation.

Table 3. Comparison of the quantitative metric dice score coefficient for abdominal dataset between U-Net and U-Net with context aggregation blocks.

	U-Net	U-Net with CAB	Number of CT scans in dataset
Bladder CT (3D)	0.902	0.924	
Bone Marrow CT (3D)	0.846	0.854	
Femoral Head Left CT (3D)	0.893	0.906	
Femoral Head Right CT (3D)	0.897	0.900	105 (from 71 to 119 slices in scan)
Rectum CT (3D)	0.778	0.791	
Small Intestine CT (3D)	0.813	0.833	
Spinal Cord CT (3D)	0.823	0.827	

Table 4. Comparison of the quantitative metric Dice score coefficient for Kvasir-SEG, CVC-ColonDB, ETIS-Larib Polyp DB and CVC-ClinicDB datasets between U-Net and HarDNet-MSEG.

	U-Net	HarDNet-MSEG	Data dimensions
Gastrointestinal polyps: Kvasir-SEG (2D)	0.818	0.912	900
Colon polyps: CVC-ColonDB (2D)	0.512	0.731	Unknown
Colon polyps: ETIS-Larib PolypDB (2D)	0.398	0.677	Unknown
Colon polyps: CVC-ClinicDB (2D)	0.823	0.932	550

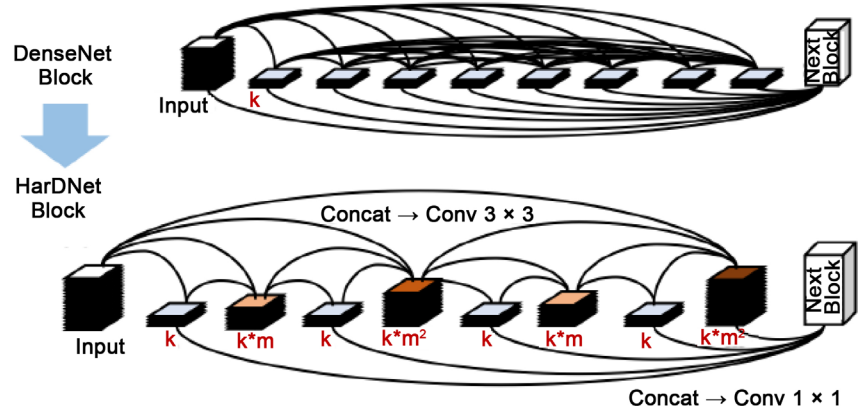


Figure 6. HarDNet block overview.

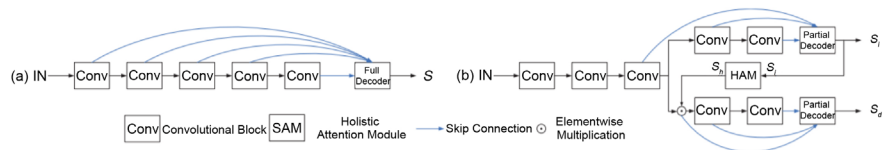


Figure 7. (a) Traditional encoder-decoder framework; (b) The proposed cascaded partial decoder framework. We use Visual Geometry Group Network (VGG16) as the backbone network. Traditional framework generates saliency map S by adopting full decoder which integrates all level features. The proposed framework adopts partial decoder, which only integrates features of deeper layers, and generates an initial saliency map S_1 and the final saliency map S_i .

2.2. Design, Optimization and Validation of the Combined DL Model

2.2.1. Design and Optimization of the Combined DL Model

We analyzed some of the most interesting and powerful architectures, suitable for segmenting the biomedical image with the objective of creating a high-performance, combined model of deep learning architectures (DL) aimed at segmenting the medical image that is independent of the type of organs/tissues, dimensions or type of image (2D/3D).

The performance of the model will also depend on the size, annotation and tagging of images in the data sets that we will use:

- Datasets containing 2D colonoscopy images-Data sets Kvasir-SEG, CVC-ColonDB, ETIS-Larib Polyp DB and CVC-ClinicDB;
- Data set containing ct scans abdomen, and the task consisted of multi-class segmentation (segmented parts: bladder, bone marrow, left femoral head, right femoral head, rectum, small intestine, spinal cord);
- Datasets, both with 2D images and 3D images-Brain US, GLAS, RITES, BraTS and LiTS datasets;
- Two sets of 3D data, one for which the task was multi-class segmentation (pancreas, spleen, kidney) and another for segmentation of a class (pancreas only). Both datasets contain CT scans and can be found in the publicly available NIH-TCIA dataset.

The DL model we imagined will combine the following DL architectures: Kite-Net, Attention U-Net, HarDNet-MSEG [15] [16] [17] [18].

The combined model we designed taking into account the key features that each of the architectures mentioned as follows (**Figure 8**):

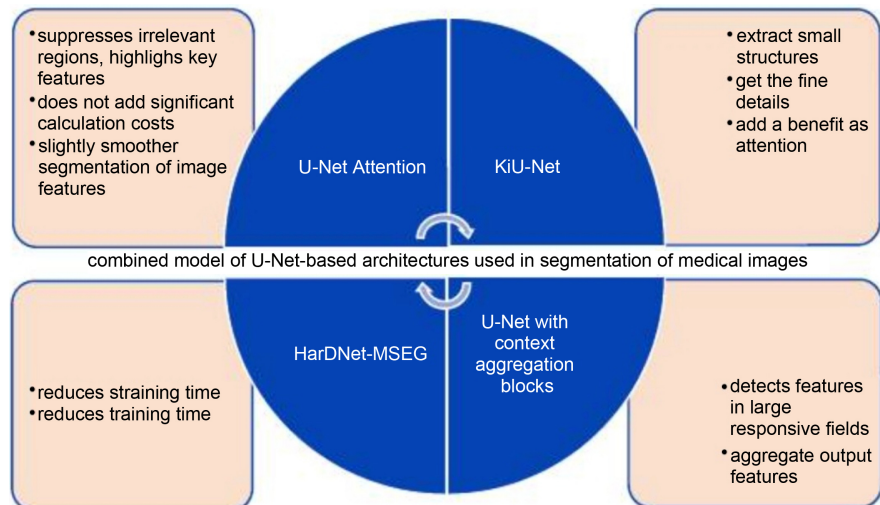


Figure 8. Combined model of U-Net-based architectures used in segmentation of medical images. Acronyms: Fully Convolutional Neural Network (U-Net), Overcomplete Convolutional Network Kite-Net (KiU-Net), Attention gate mechanism is an improvement added on convolutional network architecture for fast and precise segmentation of images (Attention U-Net), Harmony Densely Connected Network-Medical image Segmentation (HarD-Net-MSEG).

- U-Net will be enhanced by having a context aggregation block encoder and we will still retain the low-level image features resulting from The U-Net, but we will have slightly finer segmentation of them without adding costs due to context aggregation blocks;
- Kite-Net will have a unit with attention gates and a Kite-Net decoder, this way we add a benefit of attention to the details of Kite-Net;
- A partial decoder like the one in the HarDNet-MSEG architecture used as the new U-Net decoder to reduce training time.

2.2.2. Validation of the DL Combined Model

In addition to comparing the performance achieved by our model in terms of quantitative evaluation measurements, we also want to have an overview of the qualitative results. To do this, we aim to compare the results obtained by the model on images with the evaluation by a specialist of images (raw, non-segmented), regarding the diagnosis and colposcopic staging of cervical precancers and also the diagnosis and staging of cervical and thyroid cancers. We will validate the qualitative results through a randomized, controlled clinical trial over a period of 17 months [20] [21].

3. Discussions

Network Architecture Search Technique (NAS) can automatically identify a certain network architecture in computer vision tasks [22] and promises its use and performance in the medical field [12] [23].

Another problem is the lack of clinical trials demonstrating the benefits of using DL's medical applications in reducing morbidity and mortality and improving the quality of life of patients.

DL can be a support in solving complex problems, with uncertainties of options in investigations and therapy and could help medically and by filtering, providing data from literature. This aspect leads to a personalized medicine of the patient's die with diagnosis and therapeutic options based on scientific evidence. Another aspect is represented by the time encoded by the doctor in patient care, time gained by the constructive and effective support of DL in medical decision-making and synthesis activities.

4. Conclusion

We analyzed the best U-Net-based architectures suitable for biomedical image segmentation. We've specified the most important features we want to fit into a new, high-performance model. In this regard, we will create a comprehensive computer assisted diagnostic methodology validated by a randomized controlled clinical trial. The model will be a highly automated tool for diagnosing and staging precancers and cervical cancer and thyroid cancers. This would help drastically minimize the time and effort that specialists put into analyzing medical images, help to achieve a better therapeutic plan, and can provide a "second

opinion” of computer assisted diagnosis.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

Funding

Scientific research funded by the University of Medicine and Pharmacy “Gr. T. Popa” in Iasi, under contract No. 4714.

References

- [1] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, USA, 7-12 June 2015, 3431-3440. <https://doi.org/10.1109/CVPR.2015.7298965>
- [2] Ursuleanu, T.F., Luca, A.R., Gheorghe, L., Grigorovici, R., Iancu, S., Hlusneac, M., *et al.* (2021) Unified Analysis Specific to the Medical Field in the Interpretation of Medical Images through the Use of Deep Learning. *E-Health Telecommunication Systems and Networks*, **10**, 41-74. <https://doi.org/10.4236/etsn.2021.102003>
- [3] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi A., Eds., *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [4] Goodfellow, I., Pouget-Abadie, J., Mehdi, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2020) Generative Adversarial Networks. *Communications of the ACM*, **63**, 139-144. <https://doi.org/10.1145/3422622>
- [5] Gibson, E., *et al.* (2017) Towards Image-Guided Pancreas and Biliary Endoscopy: Automatic Multi-Organ Segmentation on Abdominal CT with Dense Dilated Networks. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D., Duchesne, S., Eds., *Medical Image Computing and Computer Assisted Intervention—MICCAI 2017*. Springer, Cham. https://doi.org/10.1007/978-3-319-66182-7_83
- [6] Christ, P.F., *et al.* (2016) Automatic Liver and Lesion Segmentation in CT Using Cascaded Fully Convolutional Neural Networks and 3D Conditional Random Fields. In: Ourselin, S., Joskowicz, L., Sabuncu, M., Unal, G., Wells, W., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016*. Springer, Cham. https://doi.org/10.1007/978-3-319-46723-8_48
- [7] Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., *et al.* (2017) Efficient Multi-Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation. *Medical Image Analysis*, **36**, 61-78. <https://doi.org/10.1016/j.media.2016.10.004>
- [8] Yang, X., Yu, L.Q., Wu, L.Y., Wang, Y., Ni, D., Qin, J. and Heng, P.-A. (2017) Fine-Grained Recurrent Neural Networks for Automatic Prostate Segmentation in Ultrasound Images. *Proceedings of the AAAI Conference on Artificial Intelligence*, **31**, 1633-1639. <https://ojs.aaai.org/index.php/AAAI/article/view/10761> .
- [9] Zhou, Z.W., Siddiquee, M.R., Tajbakhsh, N. and Liang, J.M. (2018) UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In: Stoyanov, D.,

- Stoyanov, D., Taylor, Z., Carneiro, G., Syeda-Mahmood, T., Martel, A., Maier-Hein, L., Tavares, J.M.R.S., Bradley, A., Papa, J.P., Belagiannis, V., Nascimento, J.C., Lu, Z., Conjeti, S., Moradi, M., Greenspan, H. and Madabhushi, A., Eds., *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLM*, Springer, Cham.
https://doi.org/10.1007/978-3-030-00889-5_1
- [10] Alom, M.Z., Hasan, M., Yakopcic, C., Taha, T. and Asari, V. (2019) Recurrent Residual Convolutional Neural Network Based on U-Net (R2U-Net) for Medical Image Segmentation. *Journal of Medical Imaging*, **6**, Article No. 014006.
<https://www.spiedigitallibrary.org/journals/journal-of-medical-imaging/volume-6/issue-01/014006/Recurrent-residual-U-Net-for-medical-image-segmentation/10.1117/1.JMI.6.1.014006.full?SSO=1>
- [11] Gordienko, Y., Peng, G., Jiang, H., Wei, Z., Kochura, Y., Alienin, O., Rokovy, O. and Stirenko, S. (2018) Deep Learning with Lung Segmentation and Bone Shadow Exclusion Techniques for Chest X-Ray Analysis of Lung Cancer. In: Hu, Z., Petoukhov, S., Dychka, I., He, M., Eds., *Advances in Computer Science for Engineering and Education, Advances in Intelligent Systems and Computing*, Springer, Cham, 638-647. https://doi.org/10.1007/978-3-319-91008-6_63
- [12] Xie, X.Z., Niu, J.W., Liu, X.F., Chen, Z.S., Tang, S.J. and Yu, S. (2021) A Survey on Incorporating Domain Knowledge into Deep Learning for Medical Image Analysis. *Medical Image Analysis*, **69**, Article ID: 101985.
<https://doi.org/10.1016/j.media.2021.101985>
- [13] Yang, D., Xu, D.G., Zhou, S.K., Georgescu, B., Chen, M.Q., Grbic, S., Metaxas, D. and Comaniciu, D. (2017) Automatic liver segmentation using an adversarial image-to-image network. *Medical Image Computing and Computer Assisted Intervention—MICCAI 2017*. Springer, Cham, 507-515.
https://doi.org/10.1007/978-3-319-66179-7_58
- [14] Zhao, Y, Dong, Q.L., Zhang, S., Zhang, W., Chen, H., Jiang, X., Guo, L., Hu, X., Han, J. and Liu, T. (2018) Automatic Recognition of fMRI-Derived Functional Networks Using 3-D Convolutional Neural Networks. *IEEE Transactions on Biomedical Engineering*, **65**, 1975-1984. <https://doi.org/10.1109/TBME.2017.2715281>
- [15] Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Rueckert, D., et al. (2018) Attention U-Net: Learning Where to Look for the Pancreas. arXiv:1804.03999. <https://arxiv.org/abs/1804.03999v3>
- [16] Valanarasu, J.M.J., Sindagi, V.A., Hacıhaliloglu, I. and Patel, V. M. (2020) Kiu-Net: Overcomplete Convolutional Architectures for Biomedical Image and Volumetric Segmentation. arXiv:2010.01663. <https://arxiv.org/abs/2010.01663v1>
- [17] Liu, Z., Liu, X., Xiao, B., Wang, S., Miao, Z., Sun, Y. and Zhang, F. (2020) Segmentation of Organs-At-Risk in Cervical Cancer CT Images with a Convolutional Neural Network. *European Journal of Medical Physics*, **69**, 184-191.
<https://doi.org/10.1016/j.ejmp.2019.12.008>
- [18] Huang, C.H., Wu, H.Y. and Lin, Y.L. (2021) HarDNet-MSEG: A Simple Encoder-Decoder Polyp Segmentation Neural Network that Achieves Over 0.9 Mean Dice and 86 FPS. arXiv:2101.07172. <https://arxiv.org/abs/2101.07172v2>
- [19] Wu, L., Xin, Y., Li, S., Wang, T., Heng, P.A. and Ni, D. (2017) Cascaded Fully Convolutional Networks for Automatic Prenatal Ultrasound Image Segmentation. 2017 *IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, Melbourne, Australia, 18-21 April 2017, 663-666.
<https://doi.org/10.1109/ISBI.2017.7950607>

- [20] Luca, A.R., Ursuleanu, T.F., Gheorghe, L., Grigorovici, R., Iancu, S., Hlusneac, M., *et al.* (2021) The Use of Artificial Intelligence on Colposcopy Images, In the Diagnosis and Staging of Cervical Precancers: A Study Protocol for a Randomized Controlled Trial. *Journal of Biomedical Science and Engineering*, **14**, 266-270. <https://doi.org/10.4236/jbise.2021.146022>
- [21] Ursuleanu, T.F., Luca, A.R., Gheorghe, L., Grigorovici, R., Iancu, S., Hlusneac, M., *et al.* (2021) The Use of Artificial Intelligence on Segmental Volumes, Constructed from MRI and CT Images, In the Diagnosis and Staging of Cervical Cancers and Thyroid Cancers: A Study Protocol for a Randomized Controlled Trial. *Journal of Biomedical Science and Engineering*, **14**, 300-304. <https://doi.org/10.4236/jbise.2021.146025>
- [22] Wistuba, M., Rawat, A. and Pedapati, T. (2019) A Survey on Neural Architecture Search. arXiv:1905.01392. <https://arxiv.org/abs/1905.01392v2>
- [23] Guo, D., Jin, D., Zhu, Z., Ho, T.Y., Harrison, A.P., Chao, C. H., *et al.* (2020) Organ at Risk Segmentation for Head and Neck Cancer Using Stratified Learning and Neural Architecture Search. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 13-19 June 2020, 4223-4232. <https://doi.org/10.1109/CVPR42600.2020.00428>