Cognitive Computation (2022) 14:1287–1302
https://doi.org/10.1007/s12559-022-10038-y

FF-UNet: a U-Shaped Deep Convolutional Neural Network for Multimodal Biomedical Image Segmentation

Ahmed Iqbal[1] · Muhammad Sharif[1] · Muhammad Attique Khan[2] · Wasif Nisar[1] · Majed Alhaisoni[3]

Received: 7 July 2021 / Accepted: 20 June 2022 / Published online: 27 June 2022

Abstract
Automatic multimodal image segmentation is considered a challenging research area in the biomedical field. U-shaped models have led to an enormous breakthrough in a large domain of medical image segmentation in recentyears. The receptive field plays an essential role in convolutionalneural networks because too small a receptive field limits context information, and too large loses localization accuracy. Despite outstanding overall performance in biomedical segmenting, classical UNet architecture uses a fixed receptive field in convolutions operations. This study proposes a few modifications in classical UNet architecture by adjusting the receptive field via feature-fused module and attention gate mechanism. Compared with baseline UNet, the numerical parameters of FF-UNet (3.94 million) is 51% of classical UNet architecture (7.75 million). Furthermore, we extended our model performance by introducing post-processing schemes. The tri-threshold fuzzy intensification-based contrast enhancement technique is utilized to improve the contrast of biomedical datasets. In the second tier, the black top-hat filtering-based method is employed to remove hair-like artifacts from the ISIC 2018 skin lesion dataset, which may create a barrier to correctly segmenting the images. The proposed models have been trained using fivefold cross-validation on five publicly available biomedical datasets and achieved the dice coefficients of 0.860, 0.932, 0.932, 0.925, and 0.894 on ETIS-LaribPolypDB, CVC-ColonDB, CVC-ClinicDB, DSB 2018, and ISIC 2018 datasets, respectively. To further verify our claims, comparative analysis based on dice results is conducted, proving the proposed model effectiveness. The FF-UNet implementation models and pre-trained weights are freely publicly available: https://github.com/ahmedeqbal/FF-UNet.

Keywords Biomedical image segmentation · Convolutional neural networks · U-shaped networks · Receptive fields · Attention mechanism

Introduction

Organ segmentation is a classical challenging research area in the medical imaging field. Image segmentation for colorectal cancer has become an emerging research topic in the past few years. Colorectal cancer is a primary cause of mortality in males and females worldwide [1]. American Cancer Society study has revealed that colorectal cancer is the 3rd major cause of mortality, and almost 104,270 new cases and 52,980 deaths are reported as standalone of gastrointestinal-type cancers 2021 in the USA [2]. This type of cancer patient survival relies on the stage of cancer which it is detected by medical experts. Nowadays, colonoscopy is widely used as a colon screening method. As an alternative, wireless capsule endoscopy technology is adopted a new method, where a patient can swallow a small WCE capsule to gather their intestine video images. After finishing this procedure, WCE capsule video images are downloaded and analyzed by a physician. One of the main challenges with WCE technology is that it requires highly skilled endoscopists and takes a lot of time to analyze the complete video [3]. Because polyps show itself only for a short duration in the complete video, few recently published studies have reflected that 22 to 28% of polyps are missed during the screening process [4]. This missed rate of the polyp by endoscopists causes

✉ Muhammad Sharif
sharif@ciitwah.edu.pk

✉ Muhammad Attique Khan
attique@ciitwah.edu.pk

[1] Department of Computer Science, COMSATS University Islamabad, Wah Campus, Pakistan

[2] Department of Computer Science, HITEC University Taxila, Taxila, Pakistan

[3] Computer Sciences Department, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia

the late identification of colorectal cancer and affects the survival rate of patients, which reduces as low as 10% [5]. Similarly, skin cancer is another deadliest type of cancer and the most rapidly spreading disease among other types of cancers with the highest mortality rate worldwide [6]. The skin cancers can be sub-categorized into melanoma and non-melanoma cancer. However, early diagnosis and treatment are more important because diagnosis at stage1, the survival rate is around 96%, compared to 5% at stage5 [7–9]. A dermatoscopy technique magnifies the skin area a dozen times, used by skin experts to observe skin pigmentation. Nuclei segmentation is another hot research topic for the researcher because it is the first step to quantitative analysis of biomedical imaging data [10]. Identifying nuclei cells helps the researcher know how the cells react to various treatments and diagnoses. To support the surgical planning and treatment, biomedical image segmentation extracts biological structures from tissues, organs, and pathologies. The manual medical image segmentation approach is timeconsuming to perform by some expert radiologists or pathologists. Moreover, when the number of medical images increases, manual segmentation becomes tedious and quite expensive, and more challenging to perform [11]. In recent years, automated segmentation methods have shown substantial advancement in biomedical data segmentation and accurate clinical decision-making. Furthermore, automated segmentation methods reduce the miss rate of polyp or lesion areas and help clinical experts localize the infected region more accurately. Different convolutional neural network models have recently emerged as a de factostandard for the biomedical field and dominate the outstanding performance in detection, classification, and semantic segmentation problems [12]. The encoder-decoder-based networksare one of the most popular types of networks that has been frequently utilized in the medical image segmentation problem. The UNet architecture is the most well-known encoder-decoder structure for medical image segmentation. But there are few limitations with classical UNet architecture; the optimal depth differs from application to application in the encoder and decoder parts because of the difficulty level and label data provided for the training task. One another challenge is classical UNet architecture used convolutions of the fixed receptive field [13]. And the receptive field is essential to optimize the accuracy of convolutional neural networks. The network receptive field plays a vital role for CNN networks because too small a receptive field will limit context information, and too large will lose localization accuracy [14]. The dynamic and adjustable receptive field provides a great advantage in U-shaped neural network. Our study improved the classical UNet architecture for biomedical segmentation, and the resulting model performs better than previously described models.

This paper contribution is summarized as follows.

- In post-processing, tri-threshold fuzzy intensification-based contrast enhancement technique is used to improve the contrast of different biomedical datasets. The black top-hat filtering-based method is utilized to remove hair-like artifacts from ISIC 2018 skin lesion dataset.
- To tackle the fixed receptive field limitation of the classical UNet architecture, we proposed an FF-UNet model, an improved version of the classical UNet model. The image segmentation network introduced in this research entirely relies on a feature-fused module that automatically adjusts the receptive fields and fused feature map extracted with dilated and standard convolution operations.
- Traditional UNet models were also developed, and the attention gate mechanism is stacked to compare our proposed model results with the baseline.
- To evaluate the proposed model performance, various publicly available datasets of different modalities are acquired, and FF-UNet delivers the best performance compared with previously proposed models.
- Our implementation code is freely available for researcher in the community of biomedical segmentation.

The rest of the work is organized as follows. The studies closely related to our proposed work are described in the next section. The materials and methods are presented in the third section, and detailed results are presented in the fourth section. Finally, the last two sections discuss and conclude our findings.

Related Work

In literature, several U-shaped convolutional networks are presented for automated classification, detection, and segmentation of multimodal biomedical images. UNet architecture achieves outstanding success in various modalities of medical image segmentation. The classical-UNet also won the championship with an exceptional victory in ISBI challenge 2012 and top-ranked in ISBI 2014 competition [15]. Despite initiating breakthroughs in medical image segmentation, the classical model also has some severe defects. Therefore, various researchers have presented different enhanced models to improve U-shaped models, such as dense mechanism, residual mechanism, attention mechanism, and ensemble mechanism, respectively. These proposed models call themselves successors of UNet architecture.

Various Encoder-Decoder-Based Models

Similar to UNet, [16]the proposed V-Net is 3D end-to-end network for biomedical organ segmentation. Dice coefficient is incorporated as an objective function, and residual

learning is used to accelerate convergence. A MultiResUNet [17] is an updated variant of classical UNet architecture [18]. The MultiRes blocks were introduced to reconcile incompatible sets of features and make two feature maps more homogenous. Five publicly available datasets of biomedical imaging are employed to validate the proposed MultiResU-Net performance, and results are compared with the classical U-Net. Similarly, Double-Unet [19] proposed a hybrid of two U-Net architectures and featured extraction improved with a pretrained VGG-19 network. Furthermore, the ASSP mechanism is adopted to obtain contextual information with a network more efficiently. Results are compared with classical U-Net architecture [18] on four publicly available biomedical datasets. A selective kernel-based U-Net model was developed by [20]. The network receptive fields are adjusted using an attention-based mechanism, and conventional and dilated convolutional feature maps are fused to obtain a common feature map. Iqbal and Sharif [21] proposed dual attention mechanism for segmentation of breast tumor images. The proposed dual attention mechanism is hybrid of lesion-based attention and channel-based attention. Reference [22] introduced channel and spatial attention in the FCN network to perform biomedical image segmentation more accurately and efficiently. This network obtains short- and long-range distance feature dependency and utilizes an attention mechanism to optimize feature maps. The experiments are demonstrated on various datasets, such as Chest X-rays and Nicole biomedical datasets. The researcher in [23] study presents a lightweight DSNet for the skin lesion segmentation task. To reduce parameter size, the depth-wise separable convolution has been utilized in the network decoder module. And the binary cross-entropy and intersection over union-based hybrid loss function have been proposed. The obtained results of the presented network were compared against the classical U-Net and FCN8s networks. PolypSegNet [24] is a modified variant of deep neural network architecture specially designed to segmentation colonoscopy images. Some architectural limitation of the classical U-Net model is tackled by adding a few building blocks such as DDI modules, DFSM, and DRM blocks. For the validation of the performance proposed network, multiple experimental steps are performed on four publicly available colonoscopy datasets. The obtained results are compared with bassline U-Net and some other previously proposed models. The novel cascade knowledge diffusion in multiple sub-network is presented for skin lesion analysis [25]. The proposed CKDNet uses two task-specific entanglement modules to improve network performance. Skin challenge data sets, ISIC2017 and ISIC2018, are used as benchmark datasets, and results are compared with other proposed methods. Nguyen et al. [26] have presented a multi-model encoder and decoder model known as MED-Net. The network captures multi-level contextual information and substantially up-sampled accurately to produce good results. A boundary-aware data augmentation approach is presented to enhance segmentation performance. Benchmark results are evaluated on four different colonoscopy medial images. In [27], researchers introduce a novel GAN-based model consisting of a deep encoder and decoder block (UNet-SDC) for the segmentation and dual discriminator block. The present model performance is validated on the ISIC skin dataset 2017 and 2018. Results are compared with models such as Mask RCNN, DCNN, and ensemble methods. Similarly, researcher [28] proposed encoder-decoder-based architecture with combination of ResNet, DenseNet, and ASSP block to segment skin lesion images. To extract the multi-scale contextual information of different dilation rates are used and fine-grained information is capture using dense skip connection in networks.

Materials and Methods

In the following section, the materials and methods for medical image segmentation using FF-UNet are proposed. Our presented method steps are based on contrast enhancement, skin lesion hair removal, and segmentation using FF-UNet models.

Dataset Description

To train and validate the proposed FF-UNet models in this paper, experimental steps are performed on five publicly available datasets. Details of these datasets are described as below.

Endoscopy Dataset

- ETIS-Larib [29] data generated by 34 endoscopy video sequences and converted in 196 RGB images varied from 1920×1072 resolution to 332×482. The dataset consists of 36 types of polyps with different sizes, textures, and shapes. The dataset is considered more challenging because of limited images and the diversity of polyp shapes and textures.
- Similarly, the CVC-ColonDB [30] dataset was generated by 15 endoscopy video sequences and contained 300 RGB images. These dataset images and their ground truth images initially have a resolution of 574×500. Hence, frames are selected to create visual dissimilarity between each image.
- The CVC-ClinicDB [31] dataset includes 612 RGB images, and the experts manually annotate their ground truth images. The images had a resolution of 384×288. These dataset images were originally extracted from 25 video sequences of colonoscopy videos, and all images show at least one polyp shape.

Nuclei Microscopy Dataset

The nuclei images have been acquired from the 2018Data-ScienceBowl. Dataset was contributed by numerous biological laboratories and presented by Booz Allen Hamilton and Kaggle. This dataset contains a total of 660 images with their original 256 × 256 resolution. The ground truth images of nuclei are segmented manually by expert biologists at the Broad Institute [10]. For our easiness, we also merged multiple nuclei images into single-cell segmented images.

Dermatoscopy Dataset

Skin lesion RBG images were acquired from ISIC 2018 challenge dataset. This dataset consists a total of 2594 RGB images and their corresponding ground truth with 4288 × 2848 resolution. Dermatoscopy experts annotate all ground truth images for the provided dataset. As compared with the previous ISIC-2017 dataset, the organizer did not provide ground truth images for validation or test images. So, the training dataset is utilized for training, validation, and test purpose.

Data Pre-processing

The pre-processing step directly affects the performance of represented image since it yields important features into account. Few necessary pre-processing steps are performed on the acquired dataset. All datasets are collected from various sources; hence, dataset images have different resolutions. Imaging resizing is an essential step in making image size smaller for the network to convolve faster and save compactional resources. First, all the images and their masks are resized to uniform input (256 × 256 × 3) in all experiments. The image pixels are normalized between 0 and 255.0.

Data Augmentation

Researchers rely on large database images to overcome overfitting and class imbalance problems in recent deep learning trends. Therefore, data augmentation methods are considered a useful solution to tackle such issues [32]. To prevent modification in provided image characteristic texture, augmentation strategies must be applied very carefully to understand the problem domain. We applied transpose and rotated the image into three fixed angles [90°, 180°, 270°]. Four images are acquired using the transpose and rotation process, and the newly created dataset was more heterogeneous. However, data augmentation is not implemented on the testing dataset.

Data Post-processing

This section describes the details of morphological operations which are applied on five datasets to enhance the contrast of images. Also, the method is proposed to remove hair-like artifacts from the ISIC2018 skin dataset. Figure 1 shows the proposed pipeline for post-processing steps. However, the hair removal method is an additional step and is only performed on the dermatoscopy dataset.

In the medical imaging domain, the contrast enhancement technique makes the image features more apparent and more visible using optimal use of color pixels. The level of contrast in a provided dataset is not sufficient for an accurate feature extraction process. And most of the provided images have weak edges and blurred visibility issues as well. A contrast enhancement technique is employed to segment the dataset images accurately to tackle such issues. In this research work, tri-threshold fuzzy intensification operator [33] is used with few optimizations for the contrast enhancement process.The secondary purpose of our post-processing phase is to remove noise from acquired images. Similar skin lesion images may contain hair-like artifacts, which may create a barrier to segmenting the images correctly. In hair removal steps, black top-hat (BTH) filtering [34] is employed to remove hair-like artifacts. The resulting images are illustrated in Fig. 2.

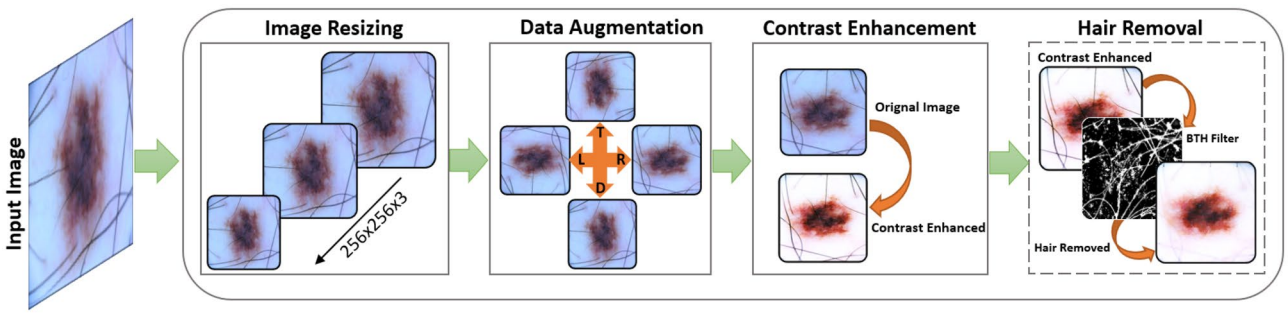

Fig. 1 Theproposed pipeline for data post-processing, followed by (**i**) contrast enhancement technique (**ii**) and hair removal method

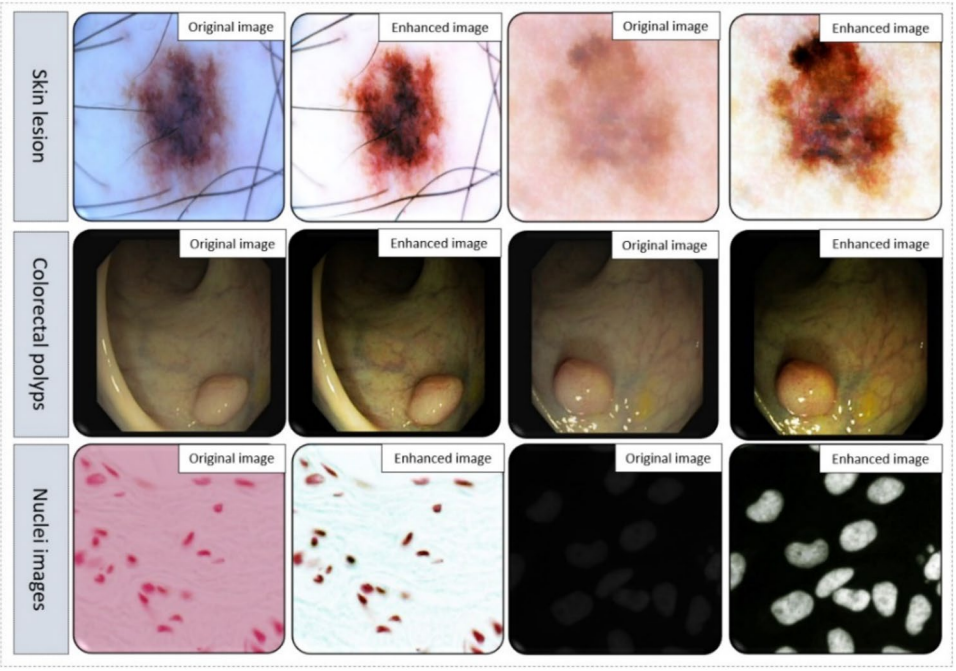

Fig. 2 Results of the contrast enhancement technique on skin lesion, colonoscopy, and nuclei datasets

Convolutional Neural Network

CNN has had a tremendous impact on many tasks, including localization, classification, and image segmentation. The classical UNet allows multiple-resolution composition and decomposition method by concatenating encoder and decoder with skip connections [18]. A novel variation of UNet classical architecture is proposed for multiple datasets of biomedical image segmentation. The following research proposed a convolutional neural network, called FF-UNet, which incorporates feature-fused module into classical UNet architecture, i.e., one of the well-recognized networks in the biomedical segmentation field. An encoder (contraction path) of the proposed model, convolutional and pooling, is implemented on the given input image to construct a compressed image representation. This compressed representation of the given input image is unsampled with deconvolutional in the decoder (or expansion path) to construct the mask showing the area of interest. Furthermore, a skip connection mechanism is incorporated to generate feature maps from the proposed model contraction path and expansion path. Additionally, classical UNet architecture is based on convolution operations of the fixed receptive field. Our proposed segmentation models introduced in this research completely rely on a feature-fused module that automatically adjusts model-receptive field using attention gate and fused feature map detail extracts with dilated convolutions and normal convolutions [35]. In proposed models, 16 filters are used at the first block, then 32, 64, 128, and 256 filters are used at each block in the network contraction path. The proposed architecture of FF-UNet is shown in Fig. 3.

Feature-Fused Module

The proposed feature-fused module is depicted in Fig. 3. The presented architecture proposed an improvement in classical UNet by replacing traditional convolution, batchsnormalization, and activation function with feature-fused module. The feature-fused module is embedded to adjust adaptivity network receptive fields and fused feature maps using two different convolution operations. The feature-fused module two branches are based on convolution operation with 3×3 kernel filter and convolutionoperation with dilation size 2 and 3×3 kernel filters. Feature maps are transformed to a single-feature vector, the resultingsfeature maps were summed, and global averaging pooling(GAP) was used. The vector was then compressed with a compression ratio of 0.5 (the number of features was half) using a fully connected layer. The compressed feature vector is de-compressed with FC layer by applying sigmoid function to decide attention coefficientforeach featuremap. The attention coefficient utilized to weight the featuremaps and compute output of feature-fused module used as

$$f^a = a^i f_d^c \oplus (1 - a^i) f^c \tag{1}$$

where Fi is ith feature map and $f_d^c f_i^c$ represent feature maps computed using 2-dilated convolution and standard convolutions. Here, a_i is a scalar $a_i \in [0, 1][0, 1]$ and a_i selection criteria are based on Eq. (2),

$$a_i = \begin{Bmatrix} f_i^d & if & a^i = 1 \\ f_i^c & if & a^i = 0 \end{Bmatrix} \tag{2}$$

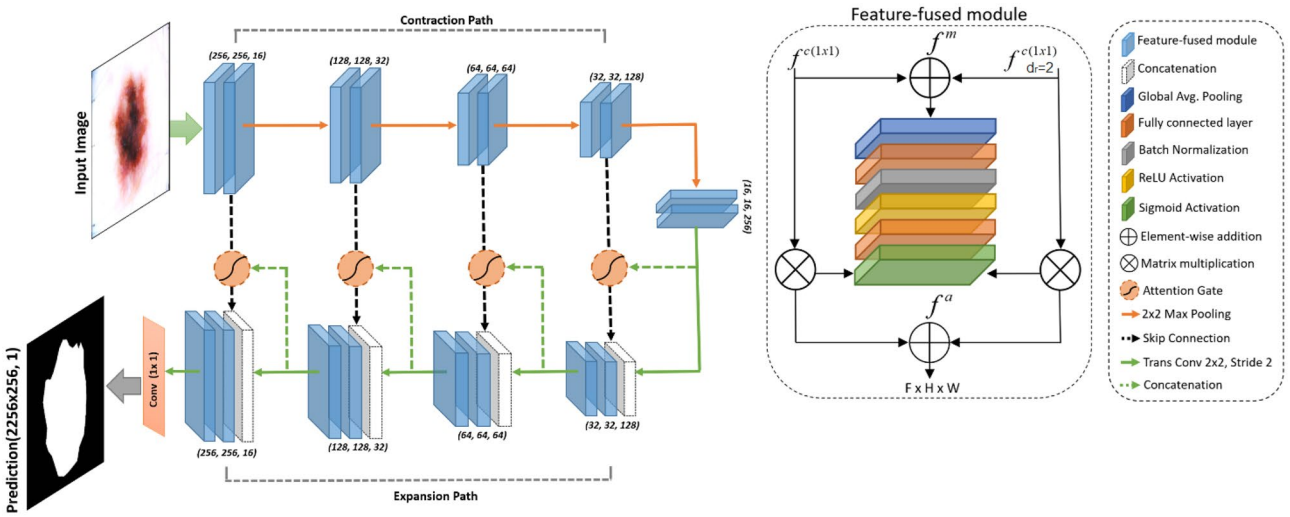

Fig. 3 The proposed model of FF-UNet with a feature-fused module for multimodal biomedical image segmentation

The average sample attention for each region of interest and feature-fused module is computed. The average sample attention was computed as follows:

$$a_k(x) = 1/n \sum_{i=1}^{n} a_i^k(x) \tag{3}$$

where x denotes number of sample image, k stands for the kth feature-fused module,and n is the number of featuremaps in the kth feature-fused module.

Attention Gate Mechanism

In various deep neural networks, deeper level of encoding consists of the highest feature representation possible. Even so, with cascaded convolution and non-linearities, spatial detail frequently vanishes in high-level output feature maps. This problem creates difficulty minimizing the smaller-object false detection rate that shows more substantial shape variability [36]. Such issues are tackled by introducing an attention gate (AG) in FF-UNet to improve the proposed model accuracy and sensitivity to foreground pixel [37]. The introduced attention gate (AG) rapidly reduced feature response and suppressed irrelevant background areas. In proposed models, the AG mechanism is applied skip connection before concatenation block only to combine relevant activations. The structure of the AG mechanism adopted in our proposed model is shown in Fig. 4.

Experimental Results

The overall experiments are performed on a system with the Windows's, 10thGeneration Intel5 Processors, 16 RAM memory, and a GeForce RTX 2060Super 8 GB graphic card. The training and testing for FF-UNet were developed on the Python Keras v2.2 library with the TensorFlow v2.1 backend. We use Adam Optimizer for stochastic gradient descent for the above five datasets optimizing the proposed model objective. The initial learning value was to "1e − 4" when train the ETIS-LaribPolypDB dataset and "1e − 3"was selected for the rest of the datasets. The training process started with a batch size of 16, and the max epoch size is selected to 300. However, early stopping is set after 50 epochs; it is utilized to prevent the overfitting problems when no improvement is observed during the training phase.

Loss Function

A "binary crossentropy is not an optimal solution for our problem domain because it may leadto bias, as the size of polyps, lesions, or nuclei is smaller than the background size." The dice coefficient-based loss is adopted rather than the cross-entropy loss. As l_{dice} is expressed in Eq. (24),here

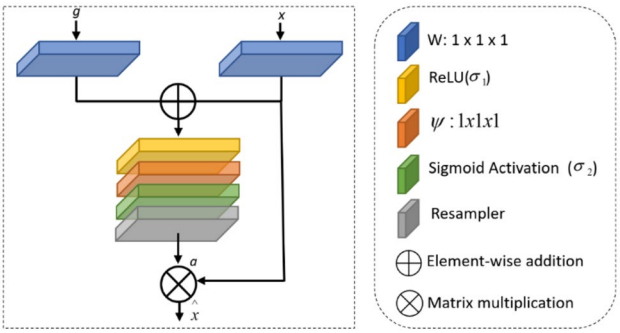

Fig. 4 The schematic illustration of the attention gate mechanism

N denotes the pixel numbers and k^n is the class "number. $p(k^n, i) \in (0, 1)$ and $g(k^n, i) \in (0, 1)$ are predicted probabilities and mask labels for the k class, respectively. The weight forclass k" is denoted as w_k,

$$l_{\text{dice}} = 1 - \sum_{k=0}^{k} \frac{2w_k \sum_i^N p(k^n, i)g(k^n, i)}{\sum_i^N [p^2(k^n, i) + g^2(k^n, i)]} \tag{4}$$

Evaluation Metrics

In this paper, several standard evaluation metrics are utilized to validate the usefulness of FF-UNet. The purity of our positive detection relative to the mask is computed using Eq. (5), termed as precision. The numbers tocorrectly allocate non-infected pixels in an image are computed using Eq. (6), termed as recall or sensitivity. Intersection over union (IoU), as calculated by Eq. (7), is adopted to count shift in the infected center of change in imageaxis. The rate of overlap betweenprediction and the given binary labels is computed using Eq. (8), termed a dice coefficient. The accurate classification of each pixel is computed by Eq. (9), denoted as accuracy.

$$\text{Precision} = \frac{TP_{\text{rate}}}{TP_{\text{rate}} + FP_{\text{rate}}} \tag{5}$$

$$\text{Recall} = \frac{TP_{\text{rate}}}{TP_{\text{rate}} + FN_{\text{rate}}} \tag{6}$$

$$\text{IoU} = \frac{TP_{\text{rate}}}{TP_{\text{rate}} + FP_{\text{rate}} + FN_{\text{rate}}} \tag{7}$$

$$\text{Dice} = \frac{2TP_{\text{rate}}}{2TP_{\text{rate}} + FP_{\text{rate}} + FN_{\text{rate}}} \tag{8}$$

$$\text{Accuracy} = \frac{TP_{\text{rate}} + TN_{\text{rate}}}{TP_{\text{rate}} + FP_{\text{rate}} + FN_{\text{rate}} + TN_{\text{rate}}} \tag{9}$$

where TP (true positive) is the correct-label infected region pixels, TN(true negative) reflects the correctly label non-infected region pixels, and FP(false positive) labels depict the non-infected pixels as infected region pixels. The infected-region pixels which are incorrectly identified as background are FN(false negatives).

k-Fold Cross-Validation

k-fold cross-validation test evaluates the overall usefulness of the proposed model separately on these datasets. In the k-fold scheme, dataset D is spilled randomly into subsets such as $[D_1, D_2, D_3, ..., D_k]$. Here, for training purposes, fourfoldisrandomly selected, and kth fold is utilized to validate the results of proposed models. The procedure is repeated five times, and values are reported, obtained from various evaluation metrics. The k-folds ensure a balance between bias and variance and reduce data dependency on performance evaluation.

Experiment Results on Endoscopy Datasets

The results were presented for the proposed FF-UNet, as well as with the attention gate mechanism. The classical UNet models were also developed, and the attention gate mechanism is stacked to compare our proposed model results with the baseline UNet. The performance measures recall, precision, IoU, dice, and accuracy are derived initially from the confusion matrix. The results consist of mean and standard division values of the fivefold cross-validation test scheme. The entries in the bold text represent the highest results obtained by proposed and classical UNet models.

Table 1 shows experimental results for the ETIS-LaribPolypDB dataset. The obtained results exhibit the advantages of the proposed FF-UNet outperforming the baseline UNet models, with the highest performance of dice 0.860 (±0.041) and IoU of 0.757 (±0.062). The consistent best results are also observed in the precision of 0.949 (±0.021), recall of 0.787 (±0.063), and highest segmentation accuracy of 0.991 (±0.160). With the attention gate mechanism, our proposed FF-UNet + AG achieves the dice-coefficient improvement of 0.083 and IoU improvement of 0.117. The consistent improvement is also recorded in the precision improvement of 0.057, recall improvement of 0.099, and in the term of accuracy 0.006, as compared with the classical UNet + AG.

Table 2 depicts the results of the CVC-ColonDB database. The results have shown that the proposed FF-UNet outperforms the baseline UNet models, with the highest performance of dice coefficient 0.932 (±0.007) and IoU of 0.873 (±0.012). Furthermore, our model also gains the best outperform with the precision of 0.946 (±0.018) and

Table 1 The performance evaluation on ETIS-Laribpolypdb dataset (mean ± standard deviation)

Method	Precision	Recall	IoU	Dice	Accuracy
UNet	0.881 (±0.121)	0.746 (±0.049)	0.684 (±0.106)	0.807 (±0.080)	0.987 (±0.680)
UNet + AG	0.857 (±0.072)	0.686 (±0.053)	0.616 (±0.058)	0.761 (±0.045)	0.984 (±0.550)
FF-UNet	**0.949 (±0.021)**	**0.787 (±0.063)**	**0.757 (±0.062)**	**0.860 (±0.041)**	**0.991 (±0.160)**
FF-UNet + AG	0.914 (±0.049)	0.785 (±0.055)	0.733 (±0.068)	0.844 (±0.045)	0.990 (±0.350)

Table 2 The performance evaluation on CVC-ColonDB dataset (mean ± standard deviation)

Method	Precision	Recall	IoU	Dice	Accuracy
UNet	0.935 (±0.017)	0.906 (±0.043)	0.852 (±0.030	0.919 (±0.017)	0.990 (±0.160)
UNet + AG	0.924 (±0.029)	0.906 (±0.024)	0.843 (±0.026	0.915 (±0.015)	0.990 (±0.150)
FF-UNet	**0.946 (±0.018)**	0.918 (±0.009)	**0.873 (±0.012)**	**0.932 (±0.007)**	**0.991 (±0.080)**
FF-UNet + AG	0.928 (±0.024)	**0.920 (±0.019)**	0.858 (±0.030)	0.923 (±0.017)	0.990 (±0.210)

the highest segmentation accuracy of 0.991 (±0.080). The UNet + AG achieved the highest recall of 0.920 (±0.019) and improvement of recall recorded 0.014, improvement in the IoU 0.015, and improvement in the dice coefficient 0.008, as compared with the classical UNet + AG.

Table 3 reports the experiment results of the CVC-ClinicDB dataset. The results reflect that the proposed FF-UNet model consistently outperforms as compared to bassline UNet with the highest precision of 0.966 (±0.012), recall of 0.886 (±0.026), IoU of 0.873 (±0.013), dice of 0.932 (±0.007), and the accuracy of 0.980 (±0.200). By using the attention gate, our model achieves the increment in precision 0.019, increment in recall 0.045, increment in IoU 0.054, and increment in the dice 0.033, and the increment in accuracy 0.007. It should be noticed that the proposed FF-UNet has shown consistent improvement in all evaluation matrices as compared with UNet + AG model.

The visual comparison of the FF-UNet with baseline UNet on three datasets is shown in Fig. 5. Here, UNet has shown satisfactory experimental results. In Fig. 5c column, polyp images affected from a lack of clear boundaries, and UNet seems under-segmented to have predicted small regions in the second and third rows. In Fig. 5d, results also indicate that our UNet + AG seems under-segmented in the second and third rows. Moreover, in Fig. 5e column, FF-UNet results better than the baseline model but still seems under-segmented on the second and third rows. Finally, FF-UNet + AG contains more semantic information, and archive segmentation accuracy is more accurate than the simple baseline UNet and FF-UNet models.

Validation Performance on Three Endoscopy Datasets

In Fig. 6a, the ETIS-LaribPolypDB dataset reflects that the proposed FF-UNet + AG achieved considerably lower validation IoU. In Fig. 6b, FF-UNet + AG achieved higher validation IoU throughout the training epochs. Similarly, Fig. 6c, d shows that FF-UNet and FF-UNet + AG results have shown that both result validation loss and validation IoU score arecomparatively better than baseline U-Net models on the CVC-ColonDB dataset. In Fig. 6e, FF-UNet seems more dominant on the CVC-ClinicDB dataset by achieving the lowest validation loss. In Fig. 6f, FF-UNet achieved the highest validation IoU score throughout the training epochs. Hence, results summarized that our proposed FF-UNet and FF-UNet + AG converge quicker and reach closer to best performance in fewer epochs.

ExperimentResults on Electron Microscopy Datasets

The results are shown for the proposed FF-UNet and classical UNet. Furthermore, the attention gate mechanism is also stacked in the proposed FF-UNet and classical UNet. The entries in bold represent the best results obtained by proposed and traditional UNet.

Table 4 depicts the experimental results of the 2018 DSB dataset. The mean and standard division results are calculated on fivefold cross-validation scheme. The results shown that FF-UNet + AG outperform the UNet + AG model, with the best performance of dice coefficient 0.925 (±0.012), IoU of 0.861 (±0.021), the accuracy of 0.951 (±0.960), and improvement of recall 0.005. The UNet achieves an improvement of 0.016 in the recall, an improvement of 0.005 in IoU, and an improvement of 0.002 in the dice coefficient, as compared with the classical UNet architecture. The UNet + AG achieves the highest precision of 0.948 (±0.014), and baseline UNet precision results are better than our proposed models. The accuracy measure consistency is observed in three models except for FF-UNet + AG.

Figure 7 demonstrates the result of each model on the DSB 2018 dataset.Fig. 7 first row results show that baseline UNet and FF-UNet performance is better than the other two sample images. Furthermore, the segmentation results on these test images show that nuclei images suffer from rough boundaries of segmented images obtained by FF-UNet and standard UNet. It seems our model decoder module is

Table 3 The performance evaluation on CVC-ClinicDB dataset (mean ± standard deviation)

Method	Precision	Recall	IoU	Dice	Accuracy
UNet	0.947 (±0.019)	0.818 (±0.064)	0.795 (±0.062)	0.884 (±0.039)	0.971 (±0.910)
UNet + AG	0.942 (±0.007)	0.819 (±0.039)	0.793 (±0.036)	0.884 (±0.022)	0.970 (±0.530)
FF-UNet	**0.966 (±0.012)**	**0.886 (±0.026)**	**0.873 (±0.013)**	**0.932 (±0.007)**	**0.980 (±0.200)**
FF-UNet + AG	0.961 (±0.014)	0.864 (±0.036)	0.847 (±0.023)	0.917 (±0.013)	0.977 (±0.240)

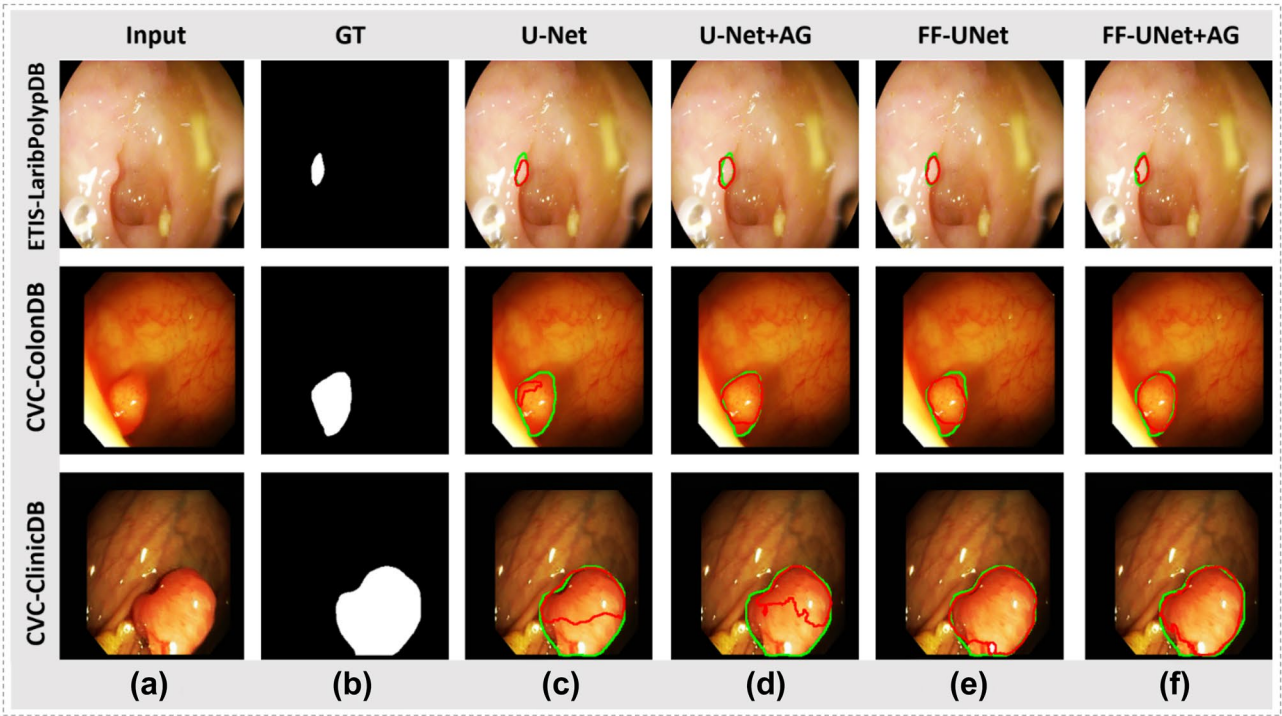

Fig. 5 Three examples present the predicted outputs by proposed models. **a**, **b** Input and ground truth of polyps; **c**–**f** the result of various models. The green curve reflects the mask image (GT), and the red curve represents the contour of the predicted results

lightweight and simpler to retrain less contextual information of input images.

Validation Performance on Microscopy Datasets

In Fig. 8a, DSB 2018 dataset illustrated that the proposed FF-UNet + AG achieved considerably lower validation IoU. In Fig. 8b, gain higher validation IoU throughout the training epochs. Similarly, FF-UNet results have shown that the performance results on validation loss and validation IoU score are comparatively better than baseline UNet and U-Net + AG models. Our results show the proposed FF-UNet and FF-UNet + AG converge quicker and reach closer to best performance in fewer epochs.

Experiment Results on Dermatoscopy Datasets

Similarly, the obtained results consist of the mean and standard division of the fivefold cross-validation test scheme. The bold entries represent the best results obtained by FF-UNet and traditional UNet. The proposed-model performance is monitored in the recall, precision, IoU, dice, and accuracy measures. Results are presented for FF-UNet and classical UNet architectures. And the attention gate mechanism is also stacked in the proposed FF-UNet and UNet architectures.

Table 5 reports the experimental ablation results of the ISIC 2018 challenge dataset. The results validate the efficacy of our FF-UNet model, which achieves the highest performance in IoU 0.809 (±0.029), performance in dice coefficient 0.894 (±0.018), as compared with baseline UNet. Furthermore, FF-UNet + AG achieves the highest performance in precision 0.894 (±0.077) and performance in accuracy 0.965 (± 1.760),additionally, improvement in 0.043, improvement in IoU 0.023, improvement in dice 0.012, and in terms of accuracy in 0.007, receptively. The regular UNet achieves the highest results in precision, IoU, dice, and accuracy measures, as compared with the UNet + AG model.

Figure 9 demonstrates the result of each model on the ISIC 2018 dataset. Figure 9c results reflect that baseline UNet poorly performed on first and second row images. Compared to baseline UNet, UNet + AG model performance is slightly better on Fig. 9d images, except the first-row image. Besides, Fig. 9e demonstrates that proposed FF-UNet results are more accurate than baseline models. Moreover, FF-UNet + AG model performance results are better than all three models, and images are segmented accurately.

Validation Performance on Dermoscopy Datasets

In Fig. 10a, ISIC 2018 dataset has shown that the proposed FF-UNet achieved considerably lower validation loss. In Fig. 10b, gain higher validation IoU throughout the

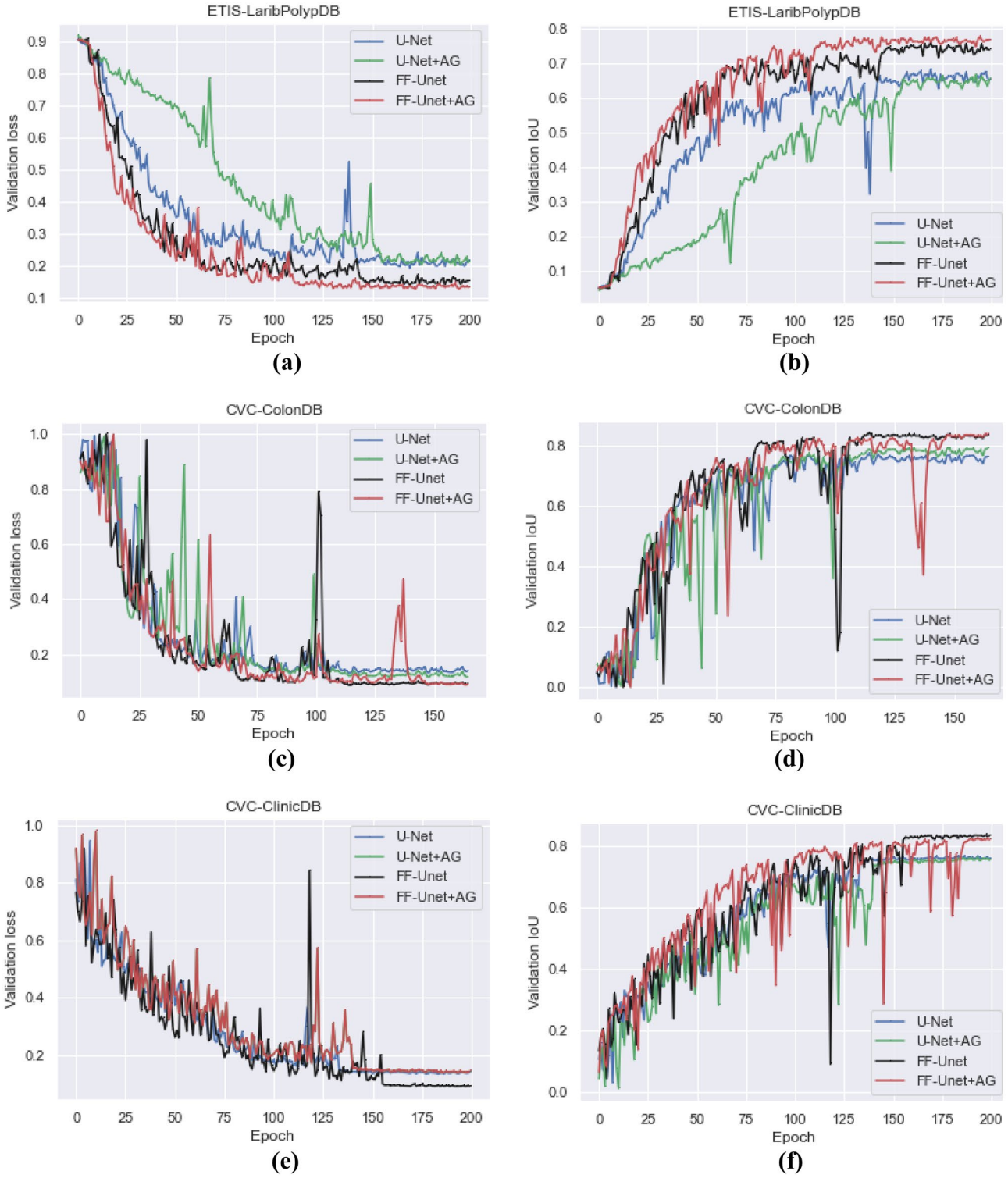

Fig. 6 The performance of validation loss and validation IoU of various models throughout training epochs are shown for three colonoscopy datasets

training epochs. Similarly, FF-UNet + AG results reflect that the performance results on validation loss and IoU score are comparatively better than baseline UNet and UNet + AG models. Thus, our results validate that our proposed FF-UNet and FF-UNet + AG converge quicker and reach closer to the best performance in fewer epochs.

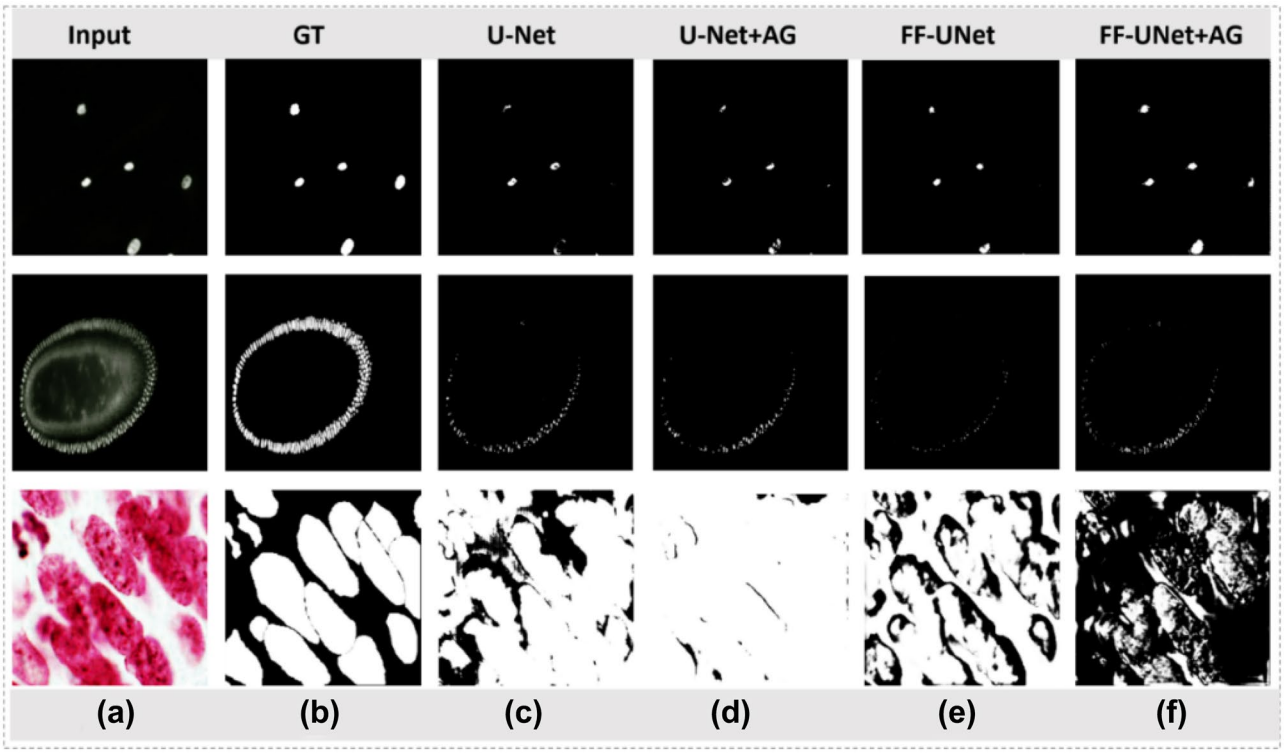

Fig. 7 Three examples presenting the predicted outputs by proposed models. **a**, **b**Input and ground truth of nuclei images; **c–f** the result of various models

Results Comparison with Other Proposed Models

Except comparing our model results with classical UNet, a comparative analysis based on dice results is conducted with other previously proposed models to evaluate our work effectiveness.

Although the classical U-shaped network has shown breakthroughs in segmenting biomedical images, the model has some serious weaknesses in classical architecture. To reconcile such issues, different studies have proposed numerous improved versions to enhance the classical UNet model, such as embedded dense mechanism,

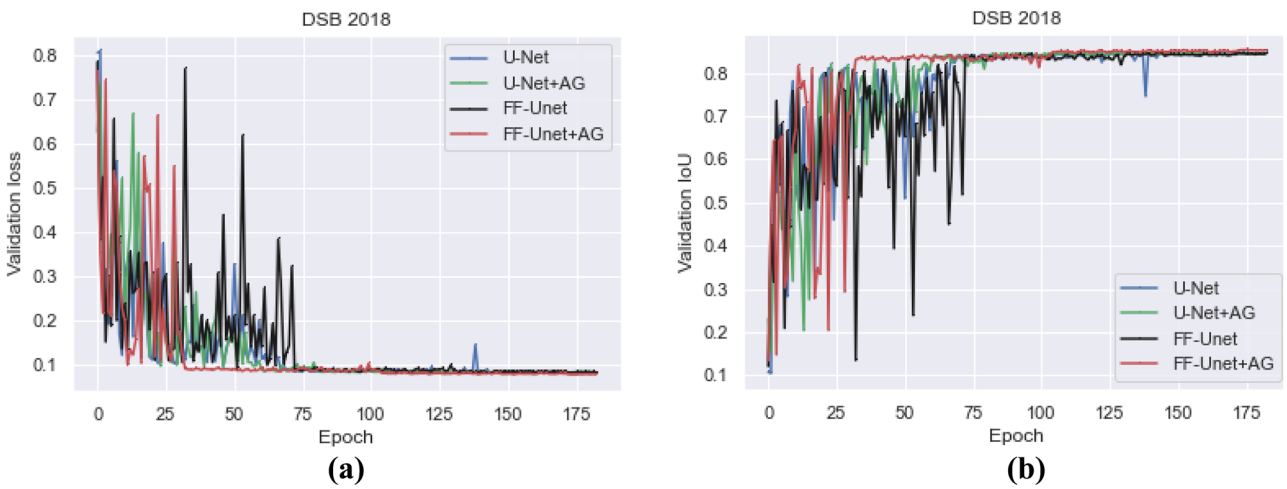

Fig. 8 Performance of validation loss (**a**) and IoU (**b**) of various models throughout training epochs are shown for three DSB 2018 datasets

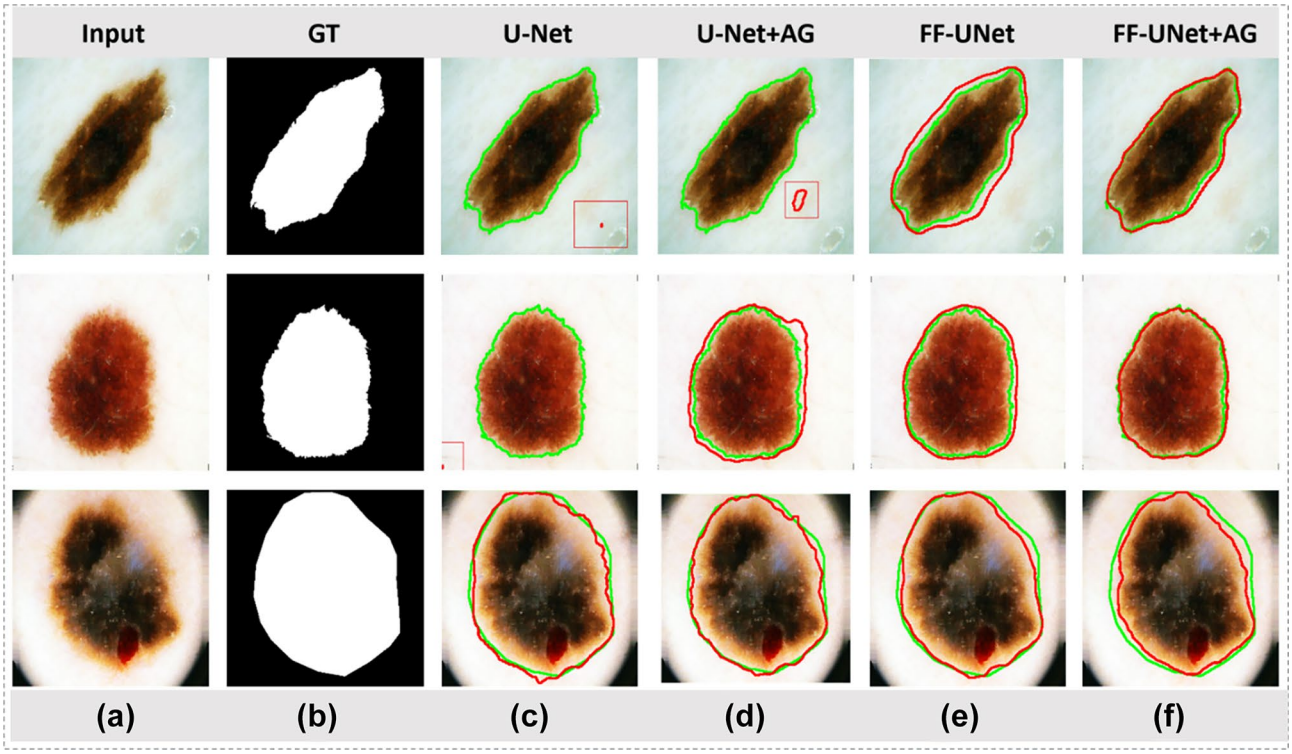

Fig. 9 Three examples presenting the predicted outputs by proposed models. **a**, **b** Input and ground truth of lesion images; **c**–**f** the result of various models. The green curve reflects the mask image (GT), and the red curve represents the contour of the predicted results

residual block, attention mechanism, and ensemble block. These proposed models are claimedto be successors of classical UNet architecture. The corresponding comparative results are illustrated in Table 6. The proposed model dice results are compared with previous models such as regular UNet [18], Unet + AG [18], UNet + + [44], ResUNet + + [45], Double-Unet [19], MultiResUnet [17], and PolypSegNet [24]. The results have indicated that our proposed FF-UNet shows the state-of-the-art results than other models, with the highest dice coefficient 0.860, 0.932, and 0.932 on ETIS-LaribPolypDB, CVC-ColonDB, and CVC-ClinicDB datasets, respectively. Similarly, FF-UNet + AG also achieved the highest dice coefficient 0.925 on DSB 2018 dataset. Result comparison is conducted with other models such as UNet [18], UNet + AG [18], MRCNN [46], LinkNet [47], FCANet [22], Double-Unet

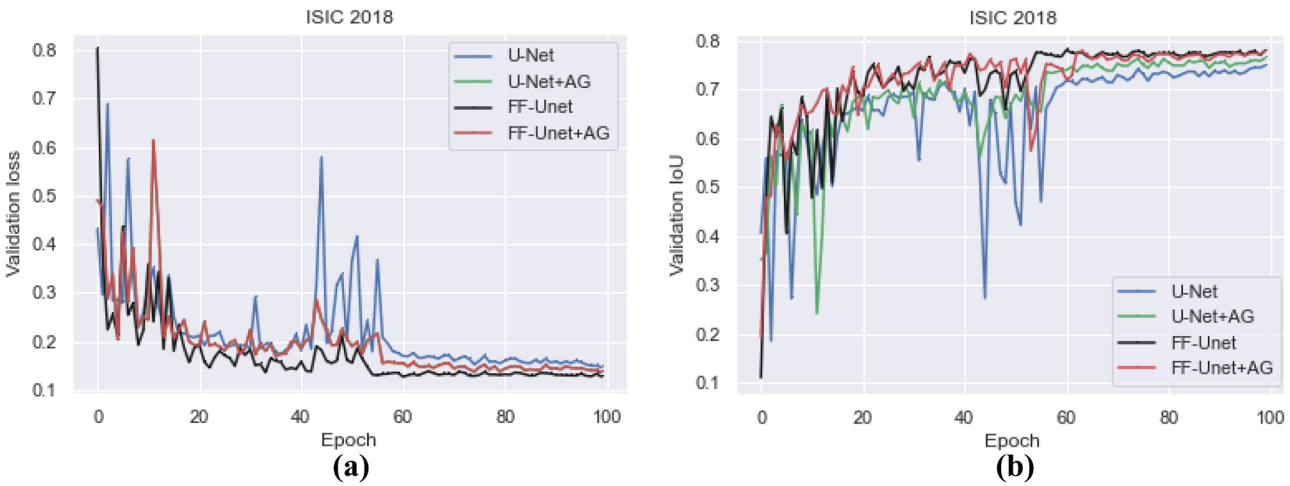

Fig. 10 Performance of validation loss (**a**) and IoU (**b**) of various models throughout training epochs are shown for three ISIC 2018 datasets

Table 4 The performance evaluation on 2018 Data Science Bowl Challenge (mean ± standard deviation)

Method	Precision	Recall	IoU	Dice	Accuracy
UNet	0.947 (±0.018)	0.871 (±0.010)	0.855 (±0.021)	0.922 (±0.012)	0.950 (±0.940)
UNet + AG	0.948 (±0.014)	0.875 (±0.011)	0.859 (±0.018)	0.924 (±0.010)	0.950 (±0.970)
FF-UNet	0.937 (±0.019)	0.887 (±0.007)	0.860 (±0.018)	0.924 (±0.010)	0.950 (±0.940)
FF-UNet + AG	0.944 (±0.018)	0.880 (±0.010)	0.861 (±0.021)	0.925 (±0.012)	0.951 (±0.960)

[19], and FU-Net [49], respectively. The proposed model usefulness is further validated on the ISIC 2018 dataset; results are compared with classical UNet [18], UNet + AG [18], FCN8s [38], DeepLab [39], BCDU-Net [43], BCDU-Net [43], and CKDNet [25], respectively. Comparative results indicate that our proposed gain a dice of 0.894, which is superior to all other models. In addition, the total numbers of parameters of various networks are also reflected in the second column of Table 6. The proposed FF-UNet and FF-UNet + AG have the lowest number of parameters containing only about 3.94 million parameters, which is significantly lesser than existing models. And the number parameters of FF-UNet (3.94 million) is only 51% of classical UNet architecture (7.75 million).

Discussion

A detailed discussion of the proposed contrast enhancement and medical image segmentation deep networks is presented in this section. For the contrast enhancement of medical images, tri-threshold fuzzy intensification-based technique is applied on five datasets. After using the technique for contrast enhancement, the black top-hat filtering-based technique is employed to remove hair-like artifacts from ISIC 2018 skin lesion dataset. Numerous deep learning-based networks have been presented by the medical imaging research community. Finding a robust and effective deep learning model to perform biomedical image segmentation isanother challenge for the research community. A novel U-shaped convolutional neural network is proposed, called FF-UNet, which incorporates feature-fused module into classical UNet architecture, i.e., one of the well-recognized networks in the medical image segmentation field. We also introduced an attention gate (AG) mechanism to rapidly reduce the feature response and suppress irrelevant background areas. A regular UNet model was also developed, and the attention gate mechanism is stacked to compare our proposed FF-UNet + AG model results with the baseline models. The colonoscopy polyp segmentation is performed, and their quantitative results are demonstrated tabularly in Tables 1, 2 and 3. Then, microscopy segmentation results are reflected in Table 4. Similarly, skin lesion segmentation obtained results are reflected in Table 5. The comparative results demonstrate that FF-UNet obtained the highest results as compared with UNet, UNet + AG, and FF-UNet + AG models. Similarly, FF-UNet also achieved the highest results in the ISIC 2018 challenge dataset in IoU, dice. However, FF-UNet + AG achieved the best results in terms of precision and accuracy measures, as depicted in Table 5. The results reflect that the attention gate mechanism does not share much significance in FF-UNet architecture and without attention gate mechanism also contribute supervisor results in most datasets. Figures 6, 8 and 10 illustrate the three random sample images predicted by our proposed models, and visual results are compared with baseline UNet models. It shows our performance results are better than bassline models, and images are segmented accurately. The resulting performance of validation loss and validation IoU of various models throughout training epochs is shown in Figs. 7, 9 and 10. Results demonstrate that our models converge quicker and reach closer to best performance in fewer epochs. A comparative analysis based on dice results is conducted with other previously proposed models to evaluate our work effectiveness. Hence, Table 6 reports that the proposed model outperforms most of the datasets with the highest dice score.

Table 7 shows training time, inference time, and speed comparison between our proposed FF-UNet with classical U-Net models. Our FF-UNet requires only 11.26 min for training the network on ETIS-Laribpolypdb dataset. However, FF-UNet inference time (ms) is slightly greater than UNet, UNet + AG, and FF-UNet + AG, MultiResUnet [17]. We also verified the clinical applicability of proposed

Table 5 The performance evaluation on the ISIC 2018 challenge (mean ± standard deviation)

Method	Precision	Recall	IoU	Dice	Accuracy
UNet	0.898 (±0.023)	0.870 (±0.028)	0.792 (±0.038)	0.883 (±0.024)	0.964 (±1.030)
UNet + AG	0.847 (±0.051)	**0.908 (±0.038)**	0.779 (±0.046)	0.875 (±0.029)	0.958 (±1.560)
FF-UNet	0.890 (±0.046)	0.904 (±0.045)	**0.809 (±0.029)**	**0.894 (±0.018)**	0.964 (±1.590)
FF-UNet + AG	**0.894 (±0.077)**	0.882 (±0.038)	0.802 (±0.085)	0.887 (±0.055)	**0.965 (±1.760)**

Table 6 The dice coefficient comparison of proposed models with other states of the art models

Method	Year	Parameters	ETIS-LaribPolypDB	CVC-ColonDB	CVC-ClinicDB	DSB 2018	ISIC 2018
U-Net [18]	2015	7.75 million	0.807	0.919	0.884	0.922	0.883
U-Net + AG [18]	2015	7.75 million	0.761	0.915	0.884	0.925	0.875
FCN8s [38]	2015	138 million	-	-	-	-	0.848
DeepLab [39]	2015	-	-	-	-	-	0.858
Zhang et al. [40]	2017	-	-	0.701	-	-	-
Akbari et al. [41]	2018	-	-	0.810	-	-	-
R2U-Net [42]	2018	-	-	-	-	-	0.861
BCDU-Net [43]	2019	-	-	-	-	-	0.838
Unet + + [44]	2019	8.80 million	0.653	0.758	0.874	-	-
ResUNet + + [45]	2019	16.2 million	0.609	0.732	0.794	-	-
MRCNN [46]	2020	-	-	-	-	0.888	-
LinkNet [47]	2020	20.3 million	-	-	-	0.890	-
FCANet [22]	2020	-	-	-	-	0.898	-
Banik et al. [48]	2020	-	-	-	0.813	-	-
Nguyen et al. [26]	2020	-	-	0.908	0.913	-	-
Double-Unet [19]	2020	29.3 million	0.762	0.858	0.923	0.913	-
MultiResUnet [17]	2020	7.26 million	0.691	0.778	0.885	-	-
PolypSegNet [24]	2020	5.50 million	0.848	0.928	0.915	-	-
CKDNet [25]	2020	-	-	-	-	-	0.877
FU-Net [49]	2021	4.39 million	-	-	-	0.904	-
FF-UNet	2021	3.94 million	**0.860**	**0.932**	**0.932**	0.924	**0.894**
FF-UNet + AG	2021	3.94 million	0.844	0.923	0.917	**0.925**	0.887

Dashes (-) shows results are not reported in cited papers, and the highest results are marked as bold

model by computing frame per second (FPs). The proposed FF-UNet achieved 27.3 FPS, which is slightly better than Double-Unet [19] and LinkNet [47]. In addition, FF-UNet pretrained weights only require 46-MB size, which can easily be deployed in mobile devices. Although proposed models indicate strong performance and accelerate the medical planning process in various medical domains, it is still far from perfection. Our model results are more accurate on large organs, but few issues still existed for small organs such as nuclei cells.

Table 7 The proposed method training time, inference time, and speed comparison between other state-of-the-art U-Net models

Method	Training time	Inference time (ms)	Speed (FPS)
UNet	12.10 m	34.50	31.88
UNet + AG	14.50 m	33.70	30.98
MultiResUnet [17]	30.90 m	33.21	30.01
Double-Unet [19]	25.80 m	47.12	22.31
LinkNet [47]	26.70 m	50.12	20.00
FF-UNet	11.26 m	36.99	27.03
FF-UNet + AG	12.34 m	40.86	24.46

Conclusion

The highlighted contribution of our research work is two-fold. In this work, the tri-threshold fuzzy intensification-based contrast enhancement technique improves the contrast of biomedical datasets. The black top-hat filtering-based method is adopted to remove hair-like artifacts from the ISIC 2018 skin lesion dataset. Furthermore, we proposed FF-UNet models for accurate segmentation of colorectal polyps, skin lesions, and nuclei images. A feature-fused module and attention gate mechanism are introduced to address the fixed receptive field challenges in U-shaped network convolution operation. And fair-result comparison analysis is conducted with classical UNet and previously proposed methods. The proposed model is tested on five freely available datasets (i.e., ETIS-LaribPolypDB, CVC-ColonDB, CVC-ClinicDB, DSB 2018, and ISIC 2018 datasets). The experimental results across five datasets show that the proposed methodology is proven more generalized. The research has the potential to further extend to other domains of biomedical imaging. Few challenges should be addressed in feature research. In the Kaggle 2018 data science bowl, visual results demonstrate that the decoder module is simpler to retrain boundary information in

up-sampling path. So, transfer learning is another option to upgrade our encoder receptive field to retrain more prominent information of input image. Because various transfer learning models have a substantial trainable parameter size, it is another challenge to reduce computational complexity and improve our proposed model robustness.

Declarations

Ethical Approval Not applicable.

Conflict of Interest The authors declareno competing interests.

References

1. Tashk A, Herp J, Nadimi E. Fully automatic polyp detection based on a novel U-Net architecture and morphological post-process. Proc - 2019 3rd Int Conf Control Artif Intell Robot Optim ICCAIRO 2019. 2019;37–41.
2. American Cancer Society, Atlanta G. American Cancer Society: cancer facts &figures 2021. [Internet]. 2021. Available from: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2021.html
3. Park CH, Kim JO, Choi MG, Kim KJ, Kim YH, Kim YS, et al. Utility of capsule endoscopy for the classification of Crohn's disease: A multicenter study in Korea. Dig Dis Sci. 2007;52:1405–9.
4. Choi HN, Kim HH, Oh JS, Jang HS, Hwang HS, Kim EY, et al. Factors influencing the miss rate of polyps in a tandem colonoscopy study. Korean J Gastroenterol. 2014;64:24–30.
5. Rabeneck L, Souchek J, El-Serag HB. Survival of colorectal cancer patients hospitalized in the Veterans Affairs Health Care System. Am J Gastroenterol [Internet]. 2003;98:1186–92. Available from: https://journals.lww.com/00000434-200305000-00038
6. Rigel DS, Friedman RJ, Kopf AW. The incidence of malignant melanoma in the United States: issues as we approach the 21st century. J Am Acad Dermatol. 1996;34:839–47.
7. Khan MA, Akram T, Sharif M, Shahzad A, Aurangzeb K, Alhussein M, et al. An implementation of normal distribution based segmentation and entropy controlled features selection for skin lesion detection and classification. BMC Cancer BMC Cancer. 2018;18:1–20.
8. Barata C, Ruela M, Francisco M, Mendonca T, Marques JS. Two systems for the detection of melanomas in dermoscopy images using texture and color features. IEEE Syst J [Internet]. 2014;8:965–79. Available from: http://ieeexplore.ieee.org/document/6570764/
9. Losina E, Walensky RP, Geller A, Beddingfield FC, Wolf LL, Gilchrest BA, et al. Visual screening for malignant melanoma. Arch Dermatol [Internet]. 2007;143:21–8. Available from: http://archderm.jamanetwork.com/article.aspx?doi=10.1001/archderm.143.1.21
10. Caicedo JC, Goodman A, Karhohs KW, Cimini BA, Ackerman J, Haghighi M, et al. Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. Nat Methods [Internet]. Springer US; 2019;16:1247–53. Available from: https://doi.org/10.1038/s41592-019-0612-7
11. Bennai MT, Guessoum Z, Mazouzi S, Cormier S, Mezghiche M. A stochastic multi-agent approach for medical-image segmentation: application to tumor segmentation in brain MR images. Artif Intell Med. 2020;110.
12. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. Z Med Phys [Internet]. Elsevier B.V.; 2019;29:102–27. Available from: https://doi.org/10.1016/j.zemedi.2018.11.002
13. De Brabandere B, Jia X, Tuytelaars T, Van Gool L. Dynamic filter networks. Adv Neural Inf Process Syst [Internet]. 2016;667–75. Available from: http://arxiv.org/abs/1605.09673
14. Qin X, Wu C, Chang H, Lu H, Zhang X. Match Feature U-Net: dynamic receptive field networks for biomedical image segmentation. symmetry (Basel) [Internet]. 2020;12:1230. Available from: https://www.mdpi.com/2073-8994/12/8/1230
15. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. 2015. p. 234–41. Available from: http://link.springer.com/10.1007/978-3-319-24574-4_28
16. Milletari F, Navab N, Ahmadi SA. V-Net: fully convolutional neural networks for volumetric medical image segmentation. Proc - 2016 4th Int Conf 3D Vision, 3DV 2016. 2016;565–71.
17. Ibtehaz N, Rahman MS. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation. Neural Networks [Internet]. Elsevier Ltd; 2020;121:74–87. Available from: https://doi.org/10.1016/j.neunet.2019.08.025
18. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) [Internet]. 2015;9351:234–41. Available from: http://arxiv.org/abs/1505.04597
19. Jha D, Riegler MA, Johansen D, Halvorsen P, Johansen HD. DoubleU-Net: a deep convolutional neural network for medical image segmentation. Proc - IEEE Symp Comput Med Syst. 2020;2020-July:558–64.
20. Byra M, Jarosik P, Szubert A, Galperin M, Ojeda-Fournier H, Olson L, et al. Breast mass segmentation in ultrasound with selective kernel U-Net convolutional neural network. Biomed Signal Process Control [Internet]. 2020;61:102027. Available from: https://linkinghub.elsevier.com/retrieve/pii/S174680942030183X
21. Iqbal A, Sharif M. MDA-Net: Multiscale dual attention-based network for breast lesion segmentation using ultrasound images. J King Saud Univ - Comput Inf Sci [Internet]. 2021; Available from: https://linkinghub.elsevier.com/retrieve/pii/S1319157821002895
22. Cheng J, Tian S, Yu L, Lu H, Lv X. Fully convolutional attention network for biomedical image segmentation. Artif Intell Med [Internet]. Elsevier; 2020;107:101899. Available from: https://doi.org/10.1016/j.artmed.2020.101899
23. Hasan MK, Dahal L, Samarakoon PN, Tushar FI, Martí R. DSNet: Automatic dermoscopic skin lesion segmentation. Comput Biol Med [Internet]. Elsevier Ltd; 2020;120:103738. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0010482520301190
24. Mahmud T, Paul B, Fattah SA. PolypSegNet: A modified encoder-decoder architecture for automated polyp segmentation from colonoscopy images. Comput Biol Med [Internet]. Elsevier Ltd; 2021;128:104119. Available from: https://doi.org/10.1016/j.compbiomed.2020.104119
25. Jin Q, Cui H, Sun C, Meng Z, Su R. Cascade knowledge diffusion network for skin lesion diagnosis and segmentation. Appl Soft Comput [Internet]. Elsevier B.V.; 2021;99:106881. Available from: https://doi.org/10.1016/j.asoc.2020.106881
26. Nguyen NQ, Vo DM, Lee SW. Contour-aware polyp segmentation in colonoscopy images using detailed upsamling encoder-decoder networks. IEEE Access. 2020;8:99495–508.
27. Lei B, Xia Z, Jiang F, Jiang X, Ge Z, Xu Y, et al. Skin lesion segmentation via generative adversarial networks with dual discriminators. Med Image Anal [Internet]. Elsevier B.V.; 2020;64:101716. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1361841520300803
28. Qamar S, Ahmad P, Shen L. Dense encoder-decoder–based architecture for skin lesion segmentation. Cognit Comput [Internet]. 2021;13:583–94. Available from: http://link.springer.com/10.1007/s12559-020-09805-6

29. Silva J, Histace A, Romain O, Dray X, Granado B. Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer. Int J Comput Assist Radiol Surg. 2014;9:283–93.
30. Bernal J, Sánchez J, Vilariño F. Towards automatic polyp detection with a polyp appearance model. Pattern Recognit. 2012;45:3166–82.
31. Bernal J, Sánchez FJ, Fernández-Esparrach G, Gil D, Rodríguez C, Vilariño F. WM-DOVA maps for accurate polyp highlighting in colonoscopy: validation vs saliency maps from physicians. Comput Med Imaging Graph. 2015;43:99–111.
32. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data [Internet]. Springer International Publishing; 2019;6. Available from: https://doi.org/10.1186/s40537-019-0197-0
33. Al-Ameen Z. Visibility enhancement for images captured in dusty weather via tuned tri-threshold fuzzy intensification operators. Int J Intell Syst Appl. 2016;8:10–7.
34. Morphological image analysis: principles and applications. Sens Rev [Internet]. 2000;20. Available from: https://www.emerald.com/insight/content/doi/10.1108/sr.2000.08720cae.001/full/html
35. Li X, Wang W, Hu X, Yang J. Selective kernel networks. 2019 IEEE/CVF Conf Comput Vis Pattern Recognit [Internet]. IEEE; 2019. p. 510–9. Available from: https://ieeexplore.ieee.org/document/8954149/
36. Oktay O, Schlemper J, Folgoc L Le, Lee M, Heinrich M, Misawa K, et al. Attention U-Net: learning where to look for the pancreas. arXiv [Internet]. 2018; Available from: http://arxiv.org/abs/1804.03999
37. Schlemper J, Oktay O, Schaap M, Heinrich M, Kainz B, Glocker B, et al. Attention gated networks: learning to leverage salient regions in medical images. Med Image Anal [Internet]. 2019;53:197–207. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1361841518306133
38. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. IEEE Trans Pattern Anal Mach Intell [Internet]. 2017;39:640–51. Available from: http://ieeexplore.ieee.org/document/7478072/
39. Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Trans Pattern Anal Mach Intell. 2018;40:834–48.
40. Zhang L, Dolwani S, Ye X. Automated polyp segmentation in colonoscopy frames using fully convolutional neural network and textons. Commun Comput Inf Sci. 2017;723:707–17.
41. Akbari M, Mohrekesh M, Nasr-Esfahani E, Soroushmehr SMR, Karimi N, Samavi S, et al. Polyp segmentation in colonoscopy images using fully convolutional network. 2018 40th Annu Int Conf IEEE Eng Med Biol Soc [Internet]. IEEE; 2018. p. 69–72. Available from: https://ieeexplore.ieee.org/document/8512197/
42. Alom MZ, Yakopcic C, Hasan M, Taha TM, Asari VK. Recurrent residual U-Net for medical image segmentation. J Med Imaging [Internet]. 2019;6:1. Available from: 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
43. Azad R, Asadi-Aghbolaghi M, Fathy M, Bi-directional ES, U-net ConvLSTM, with densley connected convolutions. Proc -,. Int Conf Comput Vis Work ICCVW 2019. IEEE. 2019;2019:406–15.
44. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. IEEE Trans Med Imaging. 2020;39:1856–67.
45. Jha D, Smedsrud PH, Riegler MA, Johansen D, De Lange T, Halvorsen P, ResUNet++: an advanced architecture for medical image segmentation. Proc -, et al. IEEE Int Symp Multimedia. ISM. 2019;2019(2019):225–30.
46. Fujita S, Han XH. Cell detection and segmentation in microscopy images with improved mask R-CNN. Proc Asian Conf … [Internet]. 2020;1–13. Available from: https://openaccess.thecvf.com/content/ACCV2020W/MLCSA/papers/Fujita_Cell_Detection_and_Segmentation_in_Microscopy_Images_with_Improved_Mask_ACCVW_2020_paper.pdf
47. Natarajan VA, Sunil Kumar M, Patan R, Kallam S, Noor Mohamed MY. Segmentation of nuclei in histopathology images using fully convolutional deep neural architecture. 2020 Int Conf Comput Inf Technol ICCIT. 2020;1:319–25.
48. Banik D, Bhattacharjee D, Nasipuri M. A multi-scale patch-based deep learning system for polyp segmentation. Adv Intell Syst Comput. 2020;1136:109–19.
49. Olimov B, Sanjar K, Din S, Ahmad A, Paul A, Kim J. FU-Net: fast biomedical image segmentation model based on bottleneck convolution layers. Multimed Syst [Internet]. Springer Berlin Heidelberg; 2021;27:637–50. Available from: https://doi.org/10.1007/s00530-020-00726-w