Received: 9 December 2024 | Revised: 20 January 2025 | Accepted: 4 July 2025 | Published online: 23 September 2025

RESEARCH ARTICLE

Journal of Data Science and Intelligent Systems
2025, Vol. 00(00) 1-11
DOI: 10.47852/bonviewJDSIS52024977

BON VIEW PUBLISHING

DCRU-Net: Dynamic Contextual Residual U-Net for Medical Image Segmentation

Manoj Kumar Singh[1,*], Satish Chand[1] and Devender Kumar[2]

[1] *School of Computer and Systems Sciences, Jawaharlal Nehru University, India*

[2] *Department of Information Technology, Netaji Subhas University of Technology, India*

Abstract: Accurate medical image segmentation (MIS) is crucial for computer-assisted diagnosis and treatment planning. This research proposes a deep learning (DL) architecture for accurate and efficient MIS, named Dynamic Contextual Residual U-Net (DCRU-Net). This design is a variation of the conventional U-Net that combines dynamic contextual residual block (DCRB) with a squeeze-and-excitation (SE) block. DCRU-Net combines the strengths of the DCRB and SE block. The SE block improves the feature-capture performance of the model by retuning the channel-specific feature responses. The DCRB adaptively modifies feature representations, selecting and adding significantly relevant contextual features at each network step, making DCRU-Net adaptable across multiple medical imaging modalities and to the difficulties of segmentation tasks. Experimental tests on six medical image collections show that DCRU-Net is superior to state-of-the-art (SOTA) methods in terms of Dice similarity coefficients (DSC) and intersection over union (IoU). Consistent performance is achieved across different medical imaging datasets with less annotated data because of the resilience and generalizability of the architecture. DCRU-Net is a new approach to accurate and automated MIS that can transform healthcare by improving segmentation accuracy and flexibility and becoming an invaluable instrument in computer-aided diagnosis.

Keywords: deep learning, medical image segmentation, DCRU-Net, computer-aided diagnosis

1. Introduction

Segmenting medical images is a vital part of many diagnostic procedures, making it an important task in medical imaging. Organs, tumors, lesions, and other anatomical features are just a few examples of regions of interest that need to be carefully delineated and extracted from medical images [1]. Clinicians can benefit from precise diagnosis, treatment planning, disease monitoring, and research through the segmentation of medical images. Several approaches have been developed for medical image segmentation (MIS) due to the widespread use of recent breakthroughs in computer vision, machine learning (ML), and deep learning (DL) [2].

Medical images are notoriously difficult to accurately segment due to their complexity and variability. Because anatomy, pathology, imaging technique, and image quality of each patient vary, these images can appear significantly different from one another in terms of size, form, and intensity [3]. Artefacts, noise, and partial volume effects are also common in medical images and can make segmentation more difficult. Although thresholding, region-growing, and active contours were promising early approaches to automated segmentation, they ultimately fell short because of their inability to adequately handle the intricacies inherent in medical images [4].

The introduction of DL, and in particular convolutional neural networks, or CNNs, has transformed the field of medical imaging by facilitating automated learning of features. CNNs are useful for gathering both coarse-grained features (such edges and textures) and fine-grained semantic meaning because to their capacity to learn high-dimensional input [5]. CNN models built on massive image datasets are modified for MIS in a process called transfer learning, which allows researchers to overcome the difficulty of sparsely annotated medical image datasets [6]. Fully convolutional networks (FCNs), U-Net, SegNet, and DeepLab are just a few examples of state-of-the-art (SOTA) approaches to medical image segmentation [7]. When it comes to pixel-wise segmentation, FCNs [8] were among the earliest designs, while U-Net skip connections facilitated better data transfer. To improve computational performance, SegNet used an encoder-decoder design with skip links, whereas DeepLab incorporated dilated convolutions to improve multi-scale segmentation [9–11].

Although much progress has been made with DL-based methods, there are still many obstacles to overcome. Clinicians need to have clear knowledge of the decision-making process of the model, which makes the interpretability of DL models a major challenge [12]. Models must maintain consistency in their performance across a variety of datasets and clinical contexts, making robustness and generalizability essential. Multimodal imaging [13] combines data from several imaging modalities and is an intriguing path toward improving segmentation accuracy. The problem of insufficient annotated data can be mitigated with the use of semi-supervised or unsupervised learning approaches, as well as efficient augmentation techniques [14].

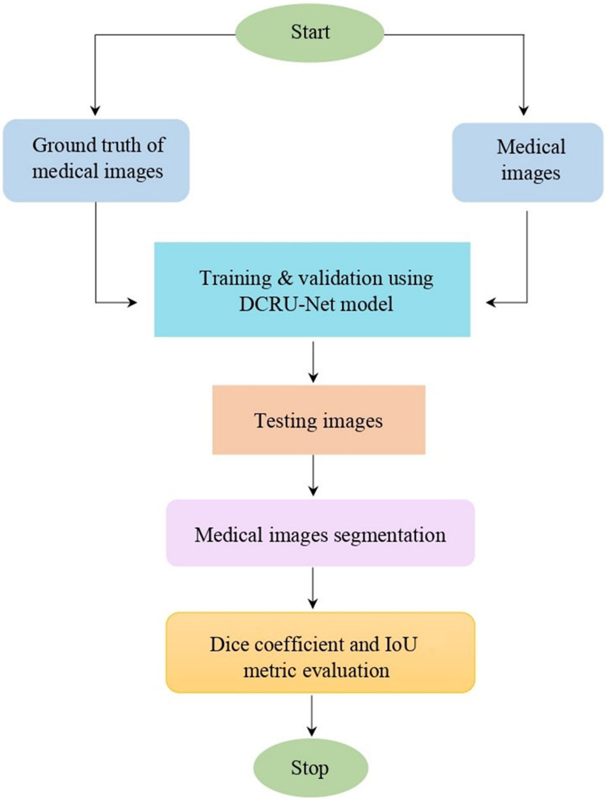
Improvements in MIS are expected to have a significant impact on clinical decision-making and patient care [15]. Researchers are making great strides toward a more promising future in MIS by exploring novel structures, optimization strategies, and interpretability approaches. Collaboration in the fields of medical imaging, computer vision, and ML will define future advancements in this crucial area of research and the future of healthcare in general [16].

***Corresponding author:** Manoj Kumar Singh, School of Computer and Systems Sciences, Jawaharlal Nehru University, India. Email: manojk42_scs@jnu.ac.in

The remainder of this paper is outlined as follows: In Section 2, we present the relevant work, and in Section 3, we present the proposed methodology; the flowchart is shown in Figure 1. In Section 4 we present the experimental data and methodology. In Section 5, we discuss the results of our experiments and the analysis of the collected data. Additionally, we will evaluate the proposed model against competing SOTA models. In the last section, we conclude the work with future directions. Below are the key contributions of this research:

- **DL Architecture:** Dynamic contextual residual U-Net (DCRU-Net) builds on the well-established U-Net framework, enhancing performance through the integration of two critical components: the Dynamic Contextual Residual Block (DCRB) and the Squeeze-and-Excitation (SE) block.
- **Feature-Capture Efficiency:** The SE block improves feature-capture performance by recalibrating channel-specific feature responses, ensuring that the model prioritizes the most relevant information.
- **Adaptive Feature Modification:** The DCRB introduces adaptive modifications to feature representations, dynamically adjusting the relevance of features at each network stage. This process enables DCRU-Net to capture both local and global contextual information, ensuring robust segmentation across diverse medical imaging modalities and challenges.
- The extensive testing on various medical image datasets shows that DCRU-Net outperforms SOTA methods in terms of evaluation metrics such as Dice similarity coefficients (DSC) and IoU. This superior performance highlights the effectiveness of the proposed architecture in accurately segmenting medical images.

Figure 1
Designed procedure flow diagram


Start
Ground truth of medical images
Medical images
Training & validation using DCRU-Net model
Testing images
Medical images segmentation
Dice coefficient and IoU metric evaluation
Stop

2. Literature Review

Disease diagnosis, treatment planning, and patient care all rely heavily on medical imaging, making MIS a vital role in this profession. Researchers have explored several approaches over the years to improve the efficiency in the MIS process [17]. The complexity of medical images was too great for early efforts, which relied on proven techniques such as thresholding [18], region expanding [19], and edge detection [20]. However, the introduction of DL has greatly improved segmentation outcomes by using big annotated datasets with models such as U-Net [9], SegNet [10], and FCNs [8]. Undersupervised and semi-supervised techniques, in addition to unsupervised and self-supervised methods, have emerged as feasible options in situations with minimal annotated data [21]. Multimodal and multitask segmentation algorithms have been developed due to the promising results achieved by integrating information from multiple imaging modalities or tackling multiple tasks concurrently [22].

Some of the challenges in MIS have been addressed through recent advances. To help segmentation models generalize to other clinical situations, for instance, experts have begun using complex data augmentation approaches [23]. Overall segmentation accuracy has also improved because of the use of adversarial learning, which has shown progress in the development of realistic and high-quality segmented outputs. The use of explainable artificial intelligence methodologies has resulted to more interpretable models [24], leading to increased confidence and acceptability among medical professionals. To better assess and compare segmentation algorithms [25], the medical and computer vision communities have worked together to create specialized medical image datasets, customized to particular clinical applications [26]. As MIS continues to advance, it is vital that researchers focus on real-world deployment and validation of these algorithms to guarantee their smooth incorporation into clinical processes, and eventually help patients and advance medical knowledge [27].

Ronneberger et al. [8] developed U-Net in 2015, and since then it has become a popular CNN architecture for MIS. There are mainly two parts to this design: the encoder [28] and the decoder networks [29]. The encoder network takes an input image and extracts its hierarchical features; the decoder network then takes those feature maps [21] and upsamples them to recreate the segmentation mask. Because of the inclusion of skip connections in the encoder and decoder networks [30], the model can learn both the global and local context [31] of the input image.

Khan et al. [7] proposed a new framework for endoscopy image classification that consists of three essential modules: Local-Global Convolutional Neural Network, Endoscopy-Lesion Attention Module, and Gastrointestinal Endoscopy CNN. The performance of the framework is evaluated on two publicly available datasets, Kvasir and HyperKvasir, demonstrating its efficacy in effectively classifying endoscopy images. DoubleU-Net, suggested by Jha et al. [2] to enhance performance of U-Net on different segmentation tasks, is a unique DL architecture that consists of two stacked U-Net topologies. For medical imaging, Oktay et al. [32] created an attention gate model called Attention U-Net, which is able to train itself to concentrate on target structures without the assistance of any external tissue localization components [33]. TransUNet, suggested by Chen et al. [34], is a robust alternative to MIS that combines Transformers [35] and U-Net. To improve features, it recovers localized spatial information from CNN feature maps and encodes it into tokenized image patches. Improved segmentation accuracy and real-time efficiency were the motivation for the proposal of the parallel reverse attention network (PraNet) by Fan et al. [36]. In their work, Lin et al. [37] proposed the DS-TransUNet framework for deep MIS, which uses a hierarchical Swin transformer in place of the usual U-shaped encoder-decoder architecture [38].

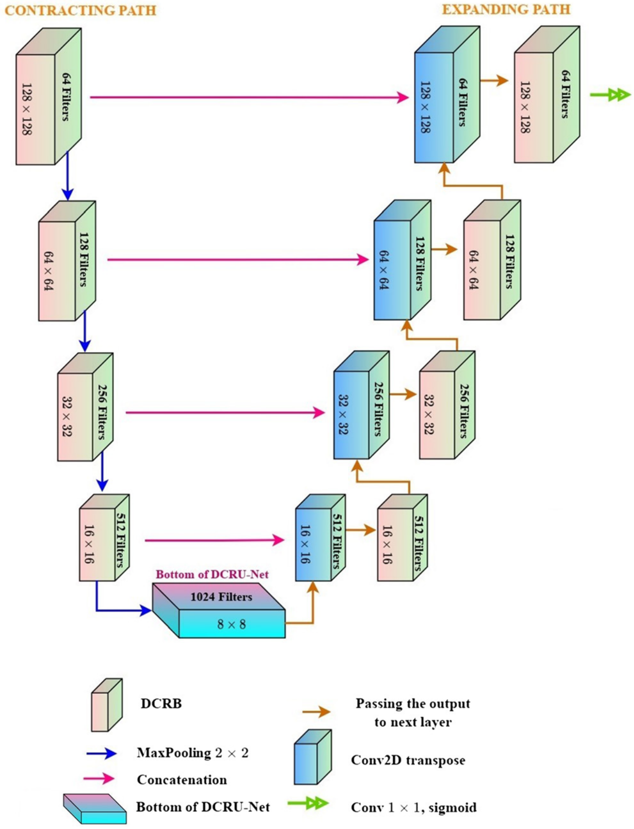
3. Research Methodology

This section discusses about our proposed DCRU-Net in detail. We discuss the SE block adopted in this network. Next, the proposed DCRB is illustrated, then the complete architecture is introduced which includes the encoder and decoder blocks.

The network design of DCRU-Net consists of the contracting path (encoder), the bottom of the DCRU-Net, and expanding path (decoder). The contracting path extracts the contextual and high-level characteristics from the input image. There are four encoder layers of increasing depth, all of which use DCRB and max pooling to downsample the feature maps. The integration of DCRB and SE blocks in the U-Net framework is achieved by embedding DCRBs in both the encoder and decoder paths, where they process the feature maps at each stage. SE blocks are applied after the convolutional layers (CLs) to recalibrate channel-wise feature responses, and skip connections merge the encoder and decoder outputs, ensuring that the network effectively combines spatial details with high-level features for accurate segmentation. The introduction of a central bottleneck layer with enhanced capacity allows the acquisition of more abstract features.

Each of the four decoder layers uses transposed convolution and concatenation to merge features from the encoding layers. The expanding path begins with the transposed CLs, which increases the spatial dimensions of the feature image while reducing the number of channels. After each transposed CL, the feature maps are concatenated via skip connections into the appropriate layer in the contracted route. Therefore, the network employs both coarse- and fine-grained properties to aid the segmentation process. After the first concatenation, its DCRB further concatenates with transposed convolutions. The final output layer typically consists of a single CL with one channel using a sigmoid activation function (AF). The complete network design with its input size is given in Figure 2.

Figure 2
DCRU-Net architecture



The computational efficiency of DCRU-Net is maintained through the optimized design of the DCRB and SE blocks, which enhance feature extraction and recalibration without significantly increasing the computational overhead. Residual connections in the DCRB facilitate gradient flow, thereby reducing the training time, while SE blocks focus on relevant features, thereby minimizing redundant computations. This design ensures robust performance on diverse datasets even with limited computational resources.,

3.1. Squeeze-and-excitation block

By improving feature recalibration of the CNN, the SE block contributes to making them more representative. To rebalance the feature maps, the block learns channel-wise scaling factors to enhance relevant features and minimize irrelevant ones. Figure 3 outlines the structure of the SE block and the five steps involved in the block are

Step 1: Global Average Pooling (GAP)

The input tensor is initially processed using GAP in the SE block. GAP calculates the average value of each feature channel across the height and width spatial dimensions, so each channel has only one value. Assume X is an input feature map with height (H), width (W), and channels (C). For each channel $c \in C$, the average value across all spatial locations is calculated as

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_{ijc} \quad (1)$$

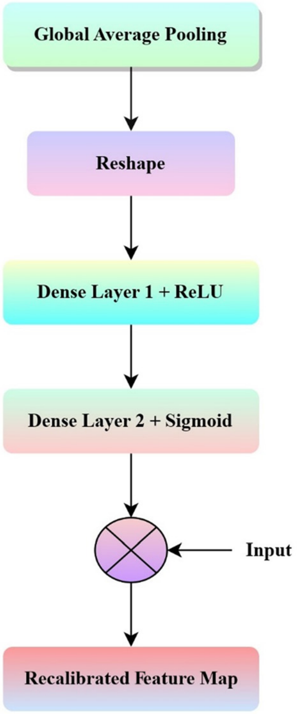
where z_c is the average value for channel c and X_{ijc} represents the pixel value located at coordinates (i, j) in channel c.

By performing this action, the dimensionality of each channel is reduced to 1×1.

Step 2: Reshape

After GAP, the output is reshaped to have the dimension 1×1×*number of channels*. This reshaping is done to prepare the data for the subsequent fully connected (FC) layers. This can be represented as

Figure 3
Feature recalibration architecture



$$Y = Reshape(z, [1, 1, C]) \tag{2}$$

where Y is the reshaped tensor, GAP procedure yields z as its result, and [1, 1, C] denotes the desired shape of the reshaped tensor.

Step 3: Squeeze or Feature Reduction

After Y is transformed, it is sent into an FC layer that uses a rectified linear unit (ReLU) AF. This layer reduces the number of channels by dividing the number of channels by a reduction ratio *r*. The aim of this step is to make the calibration process nonlinear and reduce its dimensionality. This operation can be defined as

$$S = ReLU\left(Dense\left(Y, \frac{C}{r}\right)\right)$$

where S represents the squeezed tensor and $Dense\left(Y, \frac{C}{r}\right)$ is the FC layer that reduces the dimension to C/r.

Step 4: Excitation or Feature Activation

After the squeeze step, the excitation tensor E is generated by applying another FC layer with a sigmoid AF. This tensor contains per-channel scaling factors that determine how to emphasize the information of each channel:

$$E = \sigma(Dense(S, C))$$

where E is the excitation tensor, σ denotes the sigmoid AF, and $Dense(S, C)$ is an FC layer that maps the squeezed tensor S to a tensor of the same dimension as the input C.

Step 5: Feature Recalibration

The last step involves recalibrating the input tensor X with the help of the excitation tensor E. This is achieved by element-wise multiplication between E and X:

$$Z = X \cdot E \tag{5}$$

where Z represents the recalibrated feature map and · denotes element-wise multiplication.

3.2. Dynamic contextual residual block

DCRB is a robust building block for deep neural networks that integrates convolutional feature extraction with block normalization (BN), AFs, an SE block, and residual connections as shown in Figure 4. Training problems, such as vanishing gradients, are reduced, and the network is better able to absorb contextual information both locally and globally. Each block starts with an input feature map, which is then passed through a sequence of CLs with normalization and activation, recalibrated with an SE block, combined with the original input through a residual connection, and then activated.

The DCRB combines multiple CLs with contextual information and residual connections to boost feature extraction in the CNN. The steps for the DCRB are as follows:

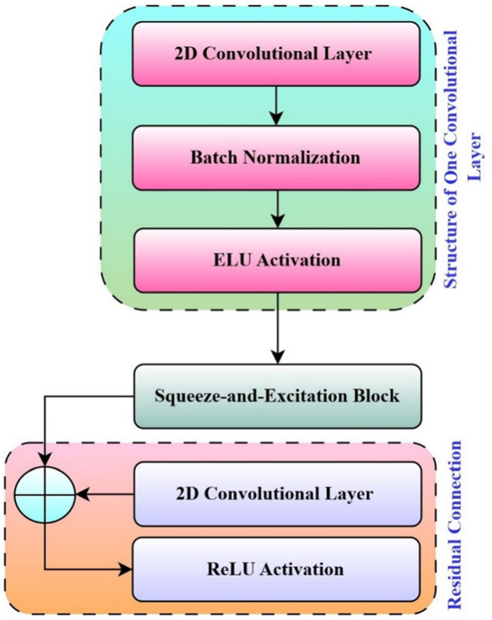
Step 1: Convolutional Layers

First, the block uses a sequence of CLs because we have a specified number of CLs to adjust the depth of the network. Features are extracted from the input tensor by applying a set of learnable filters to each CL. For the *i^th^* CL in the block, where *i* lies between 1 and the number of CLs

$$X_i = Conv2D(X_{i-1}, W_i) + b_i \tag{6}$$

where the output feature map of the i^{th} CL is denoted as X_i, the input feature map for the i^{th} CL is represented as X_{i-1}, W_i represents the trainable weights for the i^{th}CL, and the bias for the i^{th} CL is denoted by b_i.

Figure 4
Dynamic Contextual Residual Block Architecture



(3)

(4)

Step 2: Batch Normalization (BN) and Activation

After each CL, BN is applied to normalize the feature maps:

$$Y_i = BatchNormalization(X_i) \tag{7}$$

BN normalizes the activations of each layer, which helps stabilize and accelerate training. Then nonlinearity is added incrementally with the use of exponential linear unit (ELU) AFs:

$$Z_i = ELU(Y_i) \tag{8}$$

where Y_i is the batch-normalized feature map and Z_i is the activation output by the ELU.

Step 3: Squeeze-and-Excitation Block

Following CLs, feature maps can be flexibly recalibrated using the SE block. It has the same steps as mentioned above. The SE improves feature recalibration by exploring channel-wise scaling factors *E* for the feature map Z_i.

Step 4: Residual Connection

The output of the SE block is connected to the input tensor through a residual connection. To maintain the residual connection, the input feature map X_{i-1} is combined with the recalibrated feature map Z_i:

$$R_i = X_{i-1} + Z_i \tag{9}$$

where R_i is the residual output of the i^{th} convolutional block.

This connection eliminates the vanishing gradient problem by allowing the gradient to pass directly through the block during training.

Step 5: ReLU

The output of the SE block is combined with a residual connection, and then an element-wise ReLU activation is performed on the output tensor:

$$O_i = ReLU(R_i) \tag{10}$$

4. Datasets and Experimental Design

This section details the datasets used in this work and the intricacies of the experimental design. It provides insights into the data source and its characteristics. A detailed description of the experimental setup is also given.

4.1. Datasets

Polyp segmentation and nuclei segmentation datasets are used to generalize the DCRU-Net. Polyp datasets consist of five endoscopy datasets: ETIS [40], ClinicDB (CVC-ClinicDB) [39], EndoScene [41], ColonDB (CVC-ColonDB) [42], and Kvasir [43] datasets. Nuclei segmentation is generalized on Data Science Bowl (DSB) 2018 dataset [38]. The majority (80%) of datasets are used for training, while 10% are utilized for model validation and the rest are used for testing. Below is a complete description of all the available datasets:

4.1.1. ETIS

This dataset contains colorectal cancer images at early diagnosis. ETIS has 196 polyp images of 1225×966 resolution taken from 34 colonoscopy recordings. This dataset presents a greater challenge because of the variation in size and form of the polyp items in ETIS.

4.1.2. ClinicDB

From 31 colonoscopy videos, this dataset includes 612 high-resolution images of polyps with 384×288 pixels. This dataset helps in the field of medical imaging in detecting polyps from colonoscopy videos.

4.1.3. EndoScene

ClinicDB and CVC-300 are combined to create this dataset, which includes 912 photos collected from 36 patients using 44 colonoscopy sequences. For our experiments, only 60 images with 574×500 resolution of CVC-300 are used for testing.

4.1.4. ColonDB

ColonDB is a collection of colonoscopy video sequences with annotations. It has 15 distinct colonoscopy clips covering 15 studies. Only 380 frames at 574×500 resolution from the entire sequence are used for the polyp segmentation and annotated with high-quality labels for the entire region containing the polyps.

4.1.5. Kvasir

The information was gathered by endoscopic methods by Vestre Viken Health Trust in Norway. The Kvasir dataset has 1000 annotated polyp images from gastrointestinal endoscopy operations. The resolution of each polyp image is 626×547.

4.1.6. DSB 2018

Images of cell nuclei have been segmented and are included in this dataset in large numbers. Different cell types, magnification levels, and imaging techniques (brightfield vs. fluorescence) were used to acquire these images. This dataset contains 670 images of human nuclei.

Both the DSB 2018 dataset and the polyp segmentation datasets have been downsampled to 128×128. Data augmentation is used for some datasets. It artificially increases the size of data by using several transformations. In addition to preventing the overfitting problem, this also boosts the performance of the model. Table 1 lists all datasets, whether or not they have undergone data augmentation, and also gives the number of samples in both the original and augmented datasets. The horizontal flip, 180˚ rotation, 90˚ counterclockwise rotation, 90˚ clockwise rotation, translation between (-20, 20), and rotation between (-30, 30) are the techniques used to augment the data.

Table 1
Augmented data with original count

Dataset	Augmentation	Count before	Count after
ETIS	Yes	196	1372
ClinicDB	Yes	612	3060
EndoScene	Yes	912	3360
ColonDB	Yes	380	2660
Kvasir	Yes	1000	7000
DSB 2018	No	670	670

4.2. Evaluation metrics

The metrics Precision (Pre), recall (Rec), intersection over union (IoU), and Dice coefficient are used to evaluate the performance of the proposed segmentation model. A brief description of these evaluation metrics is given below.

4.2.1. Precision

Precision evaluates the accuracy of a positive prediction of a model. It is calculated as the ratio of accurate predictions (true positives, TP) relative to all predictions (correct and wrong, false positives, FP).

$$Precision = \frac{TP}{TP+FP} \quad (11)$$

4.2.2. Recall

Recall is the percentage of relevant outcomes that are correctly identified. It is the ratio of true positives (TP) over the sum of TP and false negatives (FN), the missed positive outcomes.

$$Recall = \frac{TP}{TP+FN} \quad (12)$$

4.2.3. Dice Coefficient

Harmonic mean of Precision and Recall is used to determine the Dice coefficient.

$$Dice\ Coefficient = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (13)$$

4.2.4. Intersection of Union (IoU)

The IoU assesses the overlap between the predicted and actual regions in object segmentation tasks. It is the quotient of the expected and actual regions over the union of the two.

$$IoU = \frac{Intersection}{Union} \quad (14)$$

4.3. Experimental design

Designing experiments for DL requires careful planning to ensure meaningful conclusions about the performance of the model. So hyperparameter setting is very important for the DL model. Our learning rate for the DCRU-Net model during training is 1e-3. The learning rate decreasing factor is 0.90 for EndoScene, DSB 2018, and ClinicDB datasets; 0.98 for Kvasir dataset; 0.94 for ETIS; and 0.95 for ColonDB. Adam is used to optimize the results with batch sizes 16 for EndoScene and ClinicDB, 8 for DSB 2018, 64 for Kvasir, 24 for ETIS, and 32 for ColonDB. For ClinicDB, EndoScene, Kvasir, and DSB 2018, 80% is used for training and 10% each for validation and testing. For ETIS and ColonDB, 60% is used for training and 20% each for validation and testing. DCRU-Net uses three CLs for ETIS, ColonDB, EndoScene, DSB 2018, ClinicDB, and Kvasir. EarlyStopping callback method is used to avoid the overfit issue of the model. DCRU-Net is trained using the binary cross-entropy loss. The experiment is run on TensorFlow version 2.12 and the Kaggle platform. Kaggle provides Nvidia Tesla P100 GPU. TensorFlow is used with Python 3.10.

5. Results and Output Analysis

The proposed DCRU-Net is tested on six different datasets. We start by testing polyp segmentation on two separate datasets, ClinicDB and Kvasir, so that we can make a fair comparison. We conduct a cross-study using all five datasets to ensure the effectiveness of the proposed DCRU-Net. The segmentation models have improved and achieved varied degrees of success in successfully segmenting images compared with SOTA methods on the Kvasir and ClinicDB datasets, as can be seen in Table 2 and Table 3. Table 4 displays the mIoU and mDice scores of several different image segmentation models on many different medical image datasets. Our DCRU-Net achieves an average mIoU of 80.9% and an average mDice of 89.2%, making it a good choice for MIS tasks.

It can be seen from Table 2 and Table 3 that U-Net and its variants have achieved distinct degrees of success. The other FCN variant of U-Net improves the mDice and mIoU for better segmentation of the conventional encoder-decoder architecture as the DoubleU-Net model is achieving 0.813 mDice and 0.733 mIoU on the Kvasir dataset. Transformer-based architectures are segmenting images with better performance metrics. For example, DS-TransUNet achieves 0.913, 0.942 mDice scores and 0.859, 0.894 mIoU scores on the Kvasir and ClinicDB datasets, respectively. The proposed DCRU-Net performs better on each independent dataset in terms of the evaluation metrics. DCRU-Net achieves 0.924 and 0.944 mDice scores on Kvasir and ClinicDB datasets, respectively. The proposed model outperforms the SOTA methods such as HarDNet-MSEG with 0.904 mDice and FANet with 0.936 mDice. The quantitative results are taken from Siddique et al. [38] for the comparative study.

Table 2
Quantitative evaluation of polyp segmentation on Kvasir dataset

		Pre	Rec	mIoU	mDice
Model	**Year**	**(%)**	**(%)**	**(%)**	**(%)**
U-Net [8]	2015	82.8	80.8	68.4	78.3
DoubleU-Net [2]	2020	86.1	84.0	73.3	81.3
Attention U-Net [32]	2018	85.2	79.3	68.6	78.7
UNet++ [33]	2018	82.0	81.7	67.8	78.4
TransUNet [34]	2021	91.3	91.2	83.3	89.6
MCTrans [3]	2021	-	-	-	86.2
Swin-Unet [30]	2021	90.6	90.6	82.5	89.0
HarDNet-MSEG [39]	2021	90.7	92.3	84.8	90.4
FANet [29]	2021	90.1	90.6	81.0	88.0
SegFormer [35]	2021	90.4	93.5	84.8	90.9
DS-TransUNet-B [38]	2022	91.4	93.5	85.6	91.1
DS-TransUNet-L [38]	2022	91.6	**93.6**	85.9	91.3
ConvMLPSeg [40]	2023	-	-	86.9	92.1
BLE-Net [41]	2023	-	-	85.4	90.5
HarDNet-CPS [42]	2023	-	-	85.6	91.1
CTNet	2024	-	-	**86.3**	91.7
DCRU-Net (Ours)	-	**93.3**	92.5	86.0	**92.4**

"-" means that the result is not available in that respective research paper, and the most promising results are highlighted in bold for each column.

Table 3
Quantitative evaluation of polyp segmentation on ClinicDB dataset

		Pre	Rec	mIoU	mDice
Model	**Year**	**(%)**	**(%)**	**(%)**	**(%)**
U-Net [8]	2015	91.7	86.8	80.4	87.2
DoubleU-Net [2]	2020	**95.9**	84.6	86.1	92.4
Attention U-Net [32]	2018	90.9	88.7	82.7	89.0
HarDNet-MSEG [39]	2021	94.5	91.2	86.4	91.8
FANet [29]	2021	94.0	93.4	89.4	93.6
SegFormer [35]	2021	91.1	94.2	86.0	91.1
UNet++ [33]	2018	88.5	91.0	81.9	88.1
TransUNet [34]	2021	91.7	94.2	86.9	92.3
MCTrans [3]	2021	-	-	-	92.3
Swin-Unet [30]	2021	90.7	91.8	84.9	90.6
DS-TransUNet-B [38]	2022	93.1	94.6	88.5	93.5
DS-TransUNet-L [38]	2022	93.7	**95.0**	**89.4**	94.2
ConvMLPSeg [40]	2023	-	-	87.1	92.4
BLE-Net [41]	2023	-	-	87.8	92.6
HarDNet-CPS [42]	2023	-	-	88.7	91.7
CTNet	2024	-	-	88.7	93.6
DCRU-Net (Ours)	-	95.7	89.4	**89.4**	**94.4**

"-" means that the result is not available in that respective research paper, and the most promising results are highlighted in bold for each column.

DCRU-Net exhibited significant improvements in medical imaging modalities, particularly in endoscopic and microscopic imaging, which highlights its versatility and effectiveness. In endoscopic imaging, the model excelled in polyp segmentation tasks and achieved remarkable results on datasets such as Kvasir and ClinicDB. On the Kvasir dataset, DCRU-Net achieved a DSC of 92.4% and an IoU of 86.0%, demonstrating precise segmentation of gastrointestinal polyps despite challenges such as anatomical variations and image noise. Similarly, on the ClinicDB dataset, it recorded a DSC of 94.4% and an IoU of 89.4%, surpassing SOTA methods in accurately delineating polyp boundaries in high-resolution colonoscopy images. Additionally, on the more challenging ETIS dataset, which features colorectal cancer images with significant variations in polyp size and shape, DCRU-Net achieved a DSC of 84.9% and an IoU of 74.1%, showcasing its robustness in handling complex cases.

In microscopic imaging, DCRU-Net also demonstrated exceptional performance in nuclei segmentation, a critical task for analyzing cellular structures and identifying pathological changes. On the DSB 2018 dataset, it achieved a DSC of 92.4% and an IoU of 86.7%, outperforming competing methods across diverse imaging techniques, including brightfield and fluorescence microscopy. The ability of the model to accurately segment nuclei, despite variations in cell shapes, sizes, and imaging conditions, highlights its adaptability to cellular-level segmentation tasks.

The mDice and IoU metrics for polyp datasets demonstrate that DCRU-Net consistently outperforms SOTA methods, achieving significant improvements across Kvasir, ClinicDB, and ETIS datasets. The superior performance highlights the effectiveness of the DCRB and SE blocks in recalibrating features and capturing contextual information critical for accurate polyp segmentation.

DCRU-Net achieves an mDice of 94.4% and an IoU of 89.4% on the ClinicDB dataset, marking a significant improvement over SOTA methods. This superior performance highlights the capacity of

AQ9

Table 4
Cross-study experimental results on the polyp segmentation task. The most promising results are highlighted in bold for each column

		ClinicDB		Kvasir		ColonDB		ETIS		EndoScene		Average	
Model	**Year**	**mIoU (%)**	**mDice (%)**	**mIoU (%)**	**mDice (%)**	**mIoU (%)**	**mDice (%)**	**mIoU (%)**	**mDice (%)**	**mIoU (%)**	**mDice (%)**	**mIoU (%)**	**mDice (%)**
U-Net [8]	2015	75.5	82.3	74.6	81.8	44.4	51.2	62.6	71.0	33.5	39.8	58.1	65.2
UNet++ [33]	2018	72.9	79.4	74.3	82.1	41.0	48.3	62.4	70.7	34.4	40.1	57.0	64.1
Attention U-Net [32]	2018	78.9	85.0	73.0	81.4	48.4	56.1	30.5	37.1	68.2	77.3	59.8	67.4
TransUNet [34]	2021	85.6	91.0	86.0	91.2	71.5	79.7	67.1	75.4	81.5	88.7	78.3	85.2
PraNet [36]	2020	84.9	89.9	84.0	89.8	64.0	70.9	56.7	62.8	79.7	87.1	73.9	80.0
HarDNet-MSEG [39]	2021	88.2	93.2	85.7	91.2	66.0	73.1	61.3	67.7	82.1	88.7	76.7	82.8
Swin-Unet [30]	2021	83.6	89.9	83.5	89.6	66.6	75.9	58.6	68.1	76.4	85.0	73.7	81.7
SETR-PUP [6]	2021	88.5	93.4	85.4	91.1	69.0	77.3	64.6	72.6	81.4	88.9	77.8	84.7
SegFormer [35]	2021	82.6	89.1	84.4	90.4	67.4	76.2	65.8	74.8	78.0	85.6	75.6	83.2
TransFuse-S [31]	2021	86.8	91.8	86.8	91.8	69.6	77.3	65.9	73.3	83.3	90.2	78.5	84.9
TransFuse-L [31]	2021	88.6	93.4	86.8	91.8	67.6	74.4	66.1	73.7	83.8	90.4	78.6	84.7
DS-TransUNet-B [38]	2022	89.1	93.8	86.8	93.4	71.7	79.8	69.8	77.2	81.0	88.2	80.1	86.5
DS-TransUNet-L [38]	2022	88.7	93.6	**88.9**	**93.5**	72.2	79.8	68.7	76.1	**84.6**	91.1	80.6	86.8
ConvMLPSeg [40]	2023	87.1	92.4	86.9	92.1	71.8	79.3	67.6	75.3	82.2	89.3	79.1	85.7
BLE-Net [41]	2023	87.8	92.6	85.4	90.5	65.8	73.1	59.4	67.3	80.5	87.9	75.8	82.3
HarDNet-CPS [42]	2023	88.7	91.7	85.6	91.1	65.8	72.9	61.9	69.0	82.6	89.1	77.0	82.8
CTNet [43]	2024	88.7	93.6	86.3	91.7	**74.4**	81.3	**73.4**	81.0	84.4	90.8	**81.4**	82.7
DCRU-Net (Ours)	-	**89.4**	**94.4**	86.0	92.4	74.1	**84.9**	70.6	**82.6**	**84.6**	**91.6**	80.9	**89.2**

the model to adapt to high-resolution datasets and effectively segment polyps, even in challenging cases with subtle boundary variations.

The training and validation graphs for loss and accuracy show consistent convergence, with both loss decreasing and accuracy increasing steadily over epochs as represented in Figure 5. The close alignment between the training and validation curves indicates a well-generalized model with no signs of underfitting or overfitting. The qualitative results shown in Figure 5 demonstrate the effectiveness of the model by comparing the predicted outputs with the ground truth. The visual comparisons highlight the ability of the model to accurately segment, showing clear boundaries and minimal errors.

The training and validation loss curves for polyp segmentation datasets demonstrate steady convergence, indicating effective learning by DCRU-Net.

The qualitative results in Figure 6 of polyp segmentation demonstrate precise delineation of polyp boundaries, even in challenging cases with small or irregularly shaped polyps.

Further, the DCRU-Net is evaluated on the DSB 2018 dataset for the nuclei segmentation task. The quantitative results are given in Table 5 with the comparison of SOTA methods. Table 5 shows the performance metrics for various semantic segmentation models on the DSB 2018 dataset. The included models are FANet with high precision and recall, UNet++ with high mIoU, and Swin-Unet with competitive scores. The proposed DCRU-Net outperforms others by achieving a superior precision of 95.6%, recall of 92.5%, mIoU of 86.7%, and mDice of 92.4%, demonstrating its effectiveness in accurately segmenting objects in images. The training and validation trends of loss and accuracy are depicted in Figure 7, showing the effectiveness of the model. The visual segmentation images results are shown in Figure 8.

The tabulated metrics for nuclei segmentation show that DCRU-Net achieves higher mDice and mIoU scores than SOTA methods on the DSB 2018 dataset. This superior performance highlights the robustness and adaptability of the model for cellular-level segmentation tasks, reinforcing the contributions of its novel architectural components.

The training and validation curves for the DSB 2018 dataset show smooth convergence, with minimal gaps between the training and validation loss. This demonstrates that DCRU-Net effectively learns generalized features for nuclei segmentation tasks without overfitting, even in the presence of diverse cell types and imaging conditions.

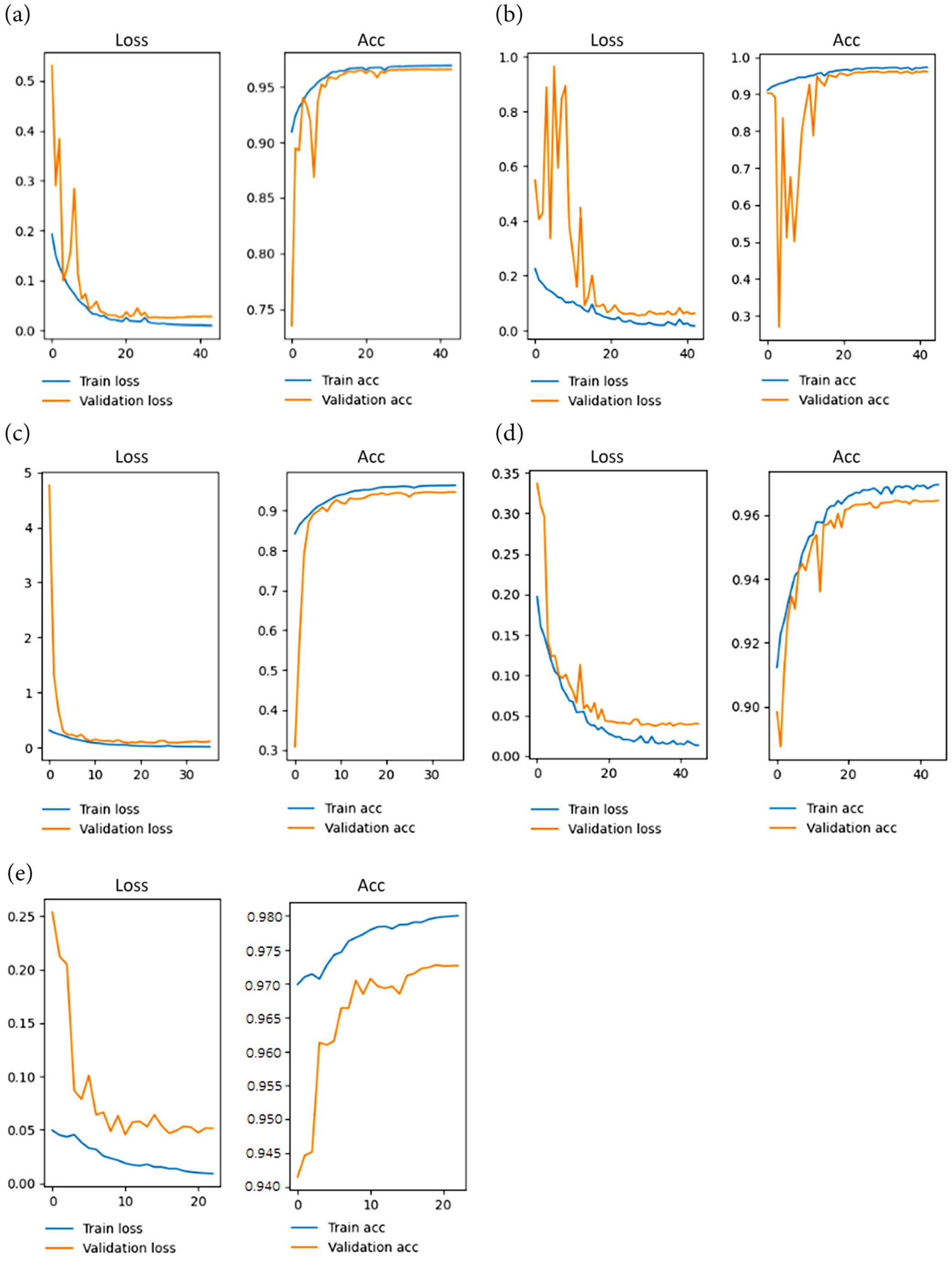
The qualitative visualization of nuclei segmentation reveals the ability of DCRU-Net to accurately segment nuclei, including those with irregular shapes or overlapping structures.

To validate the performance improvements achieved by DCRU-Net over SOTA models, a paired t-test is conducted using mDice and mIoU metrics across multiple datasets. The paired t-test evaluates whether the observed differences in performance are statistically significant by analyzing the mean difference between paired measurements, the variability of these differences, and the standard error.

For the mDice metric, the differences between the DCRU-Net and SOTA models across datasets are calculated as $D_i = DCRU - Net_i - SOTA_i$. The calculated mean difference $(\overline{D})$ is 1.3, with a standard deviation (SD) of 0.6164. Using the standard error $\left(SE = \frac{SD}{\sqrt{n}} = 0.2757\right)$, the derived t-value is $t = \frac{\overline{D}}{SE} = 4.716$. Similarly, for the mIoU metric, the mean difference is 1.06 with a standard deviation of 0.4336 and a standard error of 0.1939, resulting in a t-value of t=5.467.

Both t-values exceed the critical threshold for statistical significance at a 95% confidence level (p<0.05) confirming that the observed improvements in mDice and mIoU by DCRU-Net are statistically significant. These results highlights the robustness of the DCRU-Net architecture and demonstrate its superior performance in

Figure 5
Training and validation performance of loss and accuracy over epochs: (a) ClinicDB, (b) ColonDB, (c) Kvasir, (d) EndoScene, and (e) ETIS



MIS tasks across diverse datasets. This statistical validation provides strong evidence of the efficacy of DCRU-Net compared to existing SOTA methods.

Table 6 compares the parameter counts (in millions) for various segmentation models and emphasizes the relative complexity and design considerations. Traditional architectures, including U-Net (24.56 million [M]), UNet++ (25.09M), Attention U-Net (25.09M), and DoubleU-Net (29.30M), have relatively low parameter counts, demonstrating their lightweight nature. Transformer-based models, such as MCTrans (23.79M), SegFormer (84.59M), TransUNet (105.28M), TransFuse (115.59M), and Swin-Unet (149.22M), show a significant increase in parameter counts, reflecting the additional computational complexity introduced by transformer components. The DS-TransUNet variants, DS-TransUNet-B (171.44M), and DS-TransUNet-L (287.75M), increase the parameter count to accommodate more advanced features and capabilities. In contrast, the proposed

Figure 6
Qualitative results of the polyp segmentation datasets. Ground truths and predicted masks are for better visualization

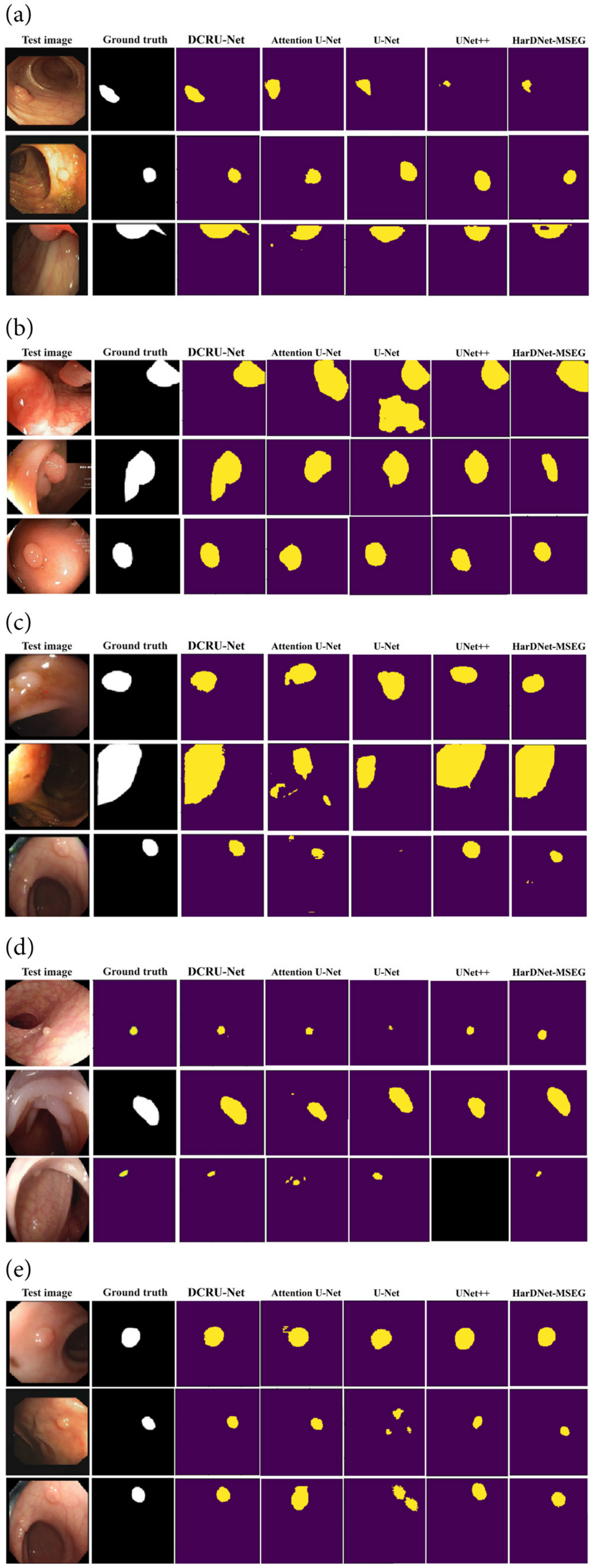


Table 5
Quantitative segmentation results on DSB 2018 dataset

Model	Year	Pre (%)	Rec (%)	mIoU (%)	mDice (%)
U-Net [8]	2015	-	-	91.0	75.7
FANet [29]	2021	91.9	92.2	85.7	91.8
DoubleU-Net [2]	2020	95.0	64.1	84.1	91.3
Attention UNet [32]	2018	91.6	-	91.0	90.8
UNet++ [33]	2018	-	-	**92.6**	89.7
TransUNet [34]	2021	89.7	92.3	83.6	90.7
SegFormer [35]	2021	91.2	93.1	85.5	91.9
Swin-Unet [30]	2021	91.5	92.4	85.1	91.6
DS-TransUNet-B [37]	2022	90.6	**94.3**	85.9	92.0
DS-TransUNet-L [37]	2022	91.2	93.8	86.1	92.2
DCRU-Net (Ours)	-	**95.6**	92.5	86.7	**92.4**

The most promising outcomes are highlighted in **bold** for each column. "-" means that the result is not available in that respective research paper.

Figure 7
Training and validation trends of loss and accuracy over epochs on DSB 2018

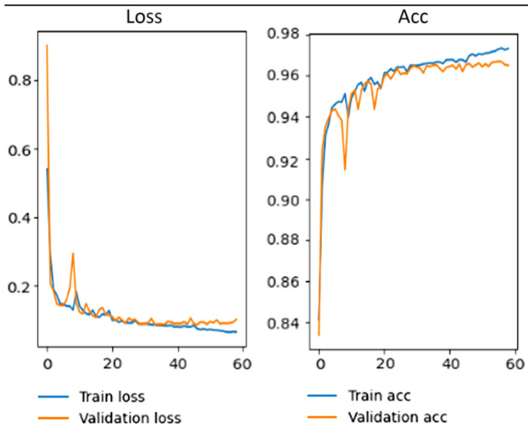


Figure 8
Qualitative segmentation results of DCRU-Net on DSB 2018 dataset

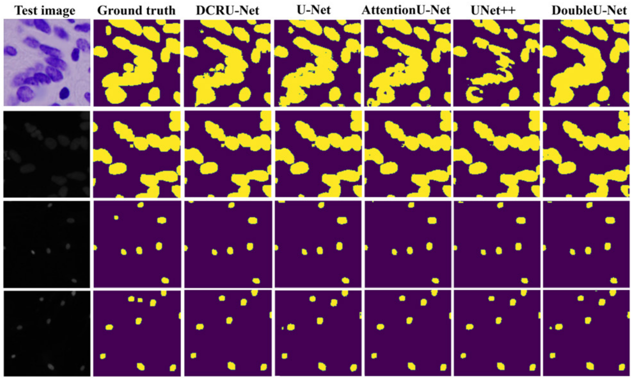


Table 6
Parameters comparison of the models

Model	Parameters (in millions)
U-Net [8]	24.56
UNet++ [33]	25.09
Attention U-Net [32]	25.09
Double U-Net [2]	29.30
MCTrans [3]	23.79
SegFormer [35]	84.59
TransUNet [34]	105.28
TransFuse [31]	115.59
Swin-Unet [30]	149.22
DS-TransUNet-B [37]	171.44
DS-TransUNet-L [37]	287.75
DCRU-Net (Ours)	48.30

DCRU-Net has 48.30M parameters, which strikes a balance between traditional architectures and transformer-based models, resulting in an efficient yet powerful solution for segmentation tasks. This comparison demonstrates the trade-offs between model complexity and potential performance gains in segmentation research.

The parameter count of DCRU-Net (48.30M) strikes a balance between traditional U-Net-based architectures and computationally intensive transformer-based models. For example, DCRU-Net is significantly lighter than models such as TransFuse (115.59M parameters) and DS-TransUNet-L (287.75M parameters), ensuring its computational feasibility for practical use in clinical settings.

6. Conclusion

In this article, we introduce DCRU-Net to improve the segmentation quality of medical images. By incorporating the DCRB and SE block, DCRU-Net leverages the strengths of both components. The SE block enhances feature-capture performance by channel specific feature response adjustment, while DCRB ensures adaptability across various medical imaging modalities and segmentation challenges. Extensive testing on diverse medical image datasets shows that DCRU-Net has superior performance in terms of Dice coefficients and IoU than existing methods. Moreover, its consistent performance across different datasets, even with limited annotated data, underscores its resilience and generalizability. Despite its efficiency in 2D segmentation, DCRU-Net requires high-performance hardware and relies heavily on data augmentation, which increases the computational costs. Its emphasis on 2D limits its applicability to volumetric medical imaging applications. Future work includes optimizing DCRU-Net for real-time 2D segmentation on standard hardware and integrating semi-supervised learning to minimize the reliance on annotated datasets.

Ethical Statement

The polyp segmentation images presented in Figure 6 of this article are sourced from a public dataset available on Google Drive at https://drive.google.com/drive/folders/10QXjxBJqCf7PAXqbDvoceWmZ-qF07tFi?usp=share_link. The authors of this article did not directly collect these images.

Conflicts of Interest

The authors declare that they have no conflict of interest in this work.

Data Availability Statement

The data that support the findings of this study are openly available in Kaggle at https://www.kaggle.com/competitions/data-science-bowl-2018. The data that support the findings of this study are openly available in Google Drive at https://drive.google.com/drive/folders/10QXjxBJqCf7PAXqbDvoceWmZ-qF07tFi?usp=share_link. The data that support the findings of this study are openly available in Polyp DataSet at https://doi.org/10.6084/m9.figshare.21221579.v2.

Author Contribution Statement

Manoj Kumar Singh: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review and editing, Visualization. **Satish Chand:** Validation, Formal analysis, Writing – review and editing, Visualization. **Devender Kumar:** Formal analysis, Writing – review and editing, Visualization, Supervision, Project administration.

References

[1] Moftah, H. M., Azar, A. T., Al-Shammari, E. T., Ghali, N. I., Hassanien, A. E., & Shoman, M. (2014). Adaptive k-means clustering algorithm for MR breast image segmentation. *Neural Computing and Applications*, *24*, 1917–1928

[2] Jha, D., Riegler, M. A., Johansen, D., Halvorsen, P., & Dagenborg,, H. J. (2020, July). Doubleu-net: A deep convolutional neural network for medical image segmentation. In *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, 558–564

[3] Yuanfeng, J. et al. (2021). Multi-compound transformer for accurate biomedical image segmentation. *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24*. Springer International Publishing.

[4] Jiang, F., Grigorev, A., Rho, S., Tian, Z., Fu, Y., Jifara, W., Adil, K., & Liu, S. (2017). Medical image semantic segmentation based on deep learning. *Neural Computing and Applications*, *29*, 1257–1265.

[5] Azad, R., Asadi-Aghbolaghi, M., Fathy, M., & Escalera, S. (2019). Bi-directional ConvLSTM U-Net with densely connected convolutions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops* (pp. 0–0).

[6] Zheng, S., Lu, J., Zhao, H., Zhu, X., Luo, Z., Wang, Y., ... & Zhang, L. (2021). Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 6881–6890).

[7] Khan, S. D., Basalamah, S., & Lbath, A. (2024). Multi-module attention-guided deep learning framework for precise gastrointestinal disease identification in endoscopic imagery. *Biomedical Signal Processing and Control*, *95*, 106396

[8] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3431–3440).

[9] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, part III 18*. Springer International Publishing.

[10] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(12), 2481–2495

[11] Chen, L.-C. et al. (2017). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(4), 834–848

[12] Gite, S., Mishra, A., & Kotecha, K. (2023). Enhanced lung image segmentation using deep learning. *Neural Computing and Applications*, *35*(31), 22839–22853

[13] Zhang, Y. et al. (2021). Deep multimodal fusion for semantic image segmentation: A survey. *Image and Vision Computing*, *105*, 104042

[14] Nalepa, J., Marcinkiewicz, M., & Kawulok, K. (2019). Data augmentation for brain-tumor segmentation: A review. *Frontiers in Computational Neuroscience*, *13*, 83

[15] Murugappan, M. et al. (2023). Automated semantic lung segmentation in chest CT images using deep neural network. *Neural Computing and Applications*, *35*(21), 15343–15364

[16] Jha, D. et al. (2019). ResUNet++: An advanced architecture for medical image segmentation. *2019 IEEE International Symposium on Multimedia (ISM)*. IEEE

[17] Wang, R. et al. (2022). Medical image segmentation using deep learning: A survey. *IET Image Processing*, *16*(5), 1243–1267

[18] Pare, S. et al. (2020). Image segmentation using multilevel thresholding: A research review. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, *44*(1), 1–29

[19] Zhong, Q. et al. (2023). Joint image and feature adaptative attention-aware networks for cross-modality semantic segmentation. *Neural Computing and Applications*, *35*(5), 3665–3676

[20] Aquino, A., Gegúndez-Arias, M. E., & Marín, D. (2010). Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques. *IEEE Transactions on Medical Imaging*, *29*(11), 1860–1869

[21] Alom, M. Z. et al. (2018). Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net). *NAECON 2018-IEEE National Aerospace and Electronics Conference*. IEEE.

[22] Kaur, A., Kaur, L., & Singh, A. (2021). GA-Unet: UNet-based framework for segmentation of 2D and 3D medical images applicable on heterogeneous datasets. *Neural Computing and Applications*, *33*(21), 14991–15025

[23] Chaitanya, K. et al. (2021), Semi-supervised task-driven data augmentation for medical image sgmentation. *Medical Image Analysis*, *68*, 101934

[24] Shi, F. et al. (2020). Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19. *IEEE Reviews in Biomedical Engineering*, *14*, 4–15

[25] Hesamian, M. H. et al. (2019). Deep learning techniques for medical image segmentation: Achievements and challenges. *Journal of Digital Imaging*, *32*, 582–596

[26] Peng, D. et al. (2021). DGFAU-Net: Global feature attention upsampling network for medical image segmentation. *Neural Computing and Applications*, *33*, 12023–12037

[27] Avalos, O. et al. (2021). An accurate cluster chaotic optimization approach for digital medical image segmentation. *Neural Computing and Applications*, *33*, 10057–10091

[28] Wu, Y. et al. (2023). D-former: A U-shaped dilated transformer for 3D medical image segmentation. *Neural Computing and Applications*, *35*(2), 1931-1944

[29] Tomar, N. K. et al. (2022). FANet: A feedback attention network for improved biomedical image segmentation. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(11), 9375–9388

[30] Cao, H. et al. (2022). Swin-UNet: UNet-like pure transformer for medical image segmentation. *European Conference on Computer Vision*. Cham: Springer Nature Switzerland

[31] Zhang, Y., Liu, H., & Hu, Q. (2021). TransFuse: Fusing transformers and CNNs for medical image segmentation. *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24*. Springer International Publishing

[32] Oktay, O. et al. (2018). Attention u-net: Learning where to look for the pancreas. *arXiv preprint* arXiv:1804.03999.

[33] Zhou, Z. et al. (2018). UNet++: A nested U-net architecture for medical image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*. Springer International Publishing

[34] Chen, J. et al. (2021). Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint* arXiv:2102.04306.

[35] Xie, E. et al. (2021). SegFormer: Simple and efficient design for semantic segmentation with transformers. *Advances in Neural Information Processing Systems*, *34*, 12077–12090

[36] Fan, D.-P. et al. (2020). PraNet: Parallel reverse attention network for polyp segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer International Publishing

[37] Lin, A. et al. (2022). DS-TransUNet: Dual Swin transformer U-net for medical image segmentation. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1–15

[38] Azad, R. et al. (2024). Medical image segmentation review: The success of U-net. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*

[39] Huang, C.-H., Wu, H.-Y., & 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 preprint* arXiv:2101.07172

[40] Jin, Y. et al. (2023). Polyp segmentation with convolutional MLP. *The Visual Computer*, *39*(10), 4819-4837

[41] Ta, N. et al. (2023). BLE-net: Boundary learning and enhancement network for polyp segmentation. *Multimedia Systems*, *29*(5), 3041-3054

[42] Yu, T. & Wu, Q. (2023). HarDNet-CPS: Colorectal polyp segmentation based on harmonic densely united network. *Biomedical Signal Processing and Control*, *85*, 104953

[43] Xiao, B et al. (2024). CTNet: Contrastive transformer network for polyp segmentation. *IEEE Transactions on Cybernetics.*

How to Cite: Singh, M. K., Chand S., & Kumar D. (2025). DCRU-Net: Dynamic Contextual Residual U-Net for Medical Image Segmentation. *Journal of Data Science and Intelligent Systems, 00*(00), 1-11. https://doi.org/10.47852/bonviewJDSIS52024977