Dense Dilated Inception Network for Medical Image Segmentation

Surayya Ado Bala[1]

Department of Computer Science and Engineering
Sharda, University, Greater Noida, India

Shri Kant[2]

Research and Technology Development Centre
Sharda University, Greater Noida, India

Abstract—In recent years, various encoder-decoder-based U-Net architecture has shown remarkable performance in medical image segmentation. However, these encoder-decoder U-Net has a drawback in learning multi-scale features in complex segmentation tasks and weak ability to generalize to other tasks. This paper proposed a generalize encoder-decoder model called dense dilated inception network (DDI-Net) for medical image segmentation by modifying U-Net architecture. We utilize three steps; firstly, we propose a dense path to replace the skip connection in the middle of the encoder and decoder to make the model deeper. Secondly, we replace the U-Net's basic convolution blocks with a modified inception module called multi-scale dilated inception module (MDI) to make the model wider without gradient vanish and with fewer parameters. Thirdly, data augmentation and normalization are applied to the training data to improve the model generalization. We evaluated the proposed model on three subtasks of the medical segmentation decathlon challenge. The experiment results prove that DDI-Net achieves superior performance than the compared methods with a Dice score of 0.82, 0.68, and 0.79 in brain tumor segmentation for edema, non-enhancing, and enhancing tumor. For the hippocampus segmentation, the result achieves 0.92 and 0.90 for anterior and posterior, respectively. For the heart segmentation, the method achieves 0.95 for the left atrial.

Keywords—Deep learning; Dense-Net; inception network; medical image segmentation; U-Net

I. INTRODUCTION

Accurate and automated segmentation of anatomical structures is the most critical and challenging task in analyzing medical images. Medical image segmentation extracts the region of interest for the diagnosis and treatment of various diseases [1], including brain cancer [2], cardiovascular diseases [3], liver cancer [4], pulmonary disease [5], etc., and the list goes on. Accurate and automatic segmentation of anatomical structures is the most important and demanding activity of medical imaging Medical image analysis aims to provide radiologists and clinicians with an efficient, accurate, and precise interpretation of medical images, reducing the time, cost, and error for effective diagnosis. Medical images such as magnetic resonance images (MRI) provides a variety of information (i.e., shape, size, and position) for a diagnostic which achieves multiple anatomical tomographic imaging by setting different parameters [6].

Deep learning (DL) models recently achieved huge success in segmenting medical images [7] because of their great ability to learn critical data features automatically [8][9].

Compared to traditional approaches, multi-layered DL has become the preferred solution for various complicated tasks. Motivated by its performance, multiple types of medical image segmentation research were conducted, notably using a convolutional neural network (CNN) such as brain tumor segmentation [10], heart segmentation [11], and hippocampus segmentation [12].

Over the years, many sophisticated CNN models have been proposed such as Alex Net [13], VGG [14], Google Net [15], Dense Net [16], ResNet [17], Deeplab [18], fully convolution network (FCN) [19] and U-Net [20]. Among these CNN networks, U-Net, an encoder-decoder based model, makes an outstanding achievement and becomes the most famous model in medical image segmentation tasks and computer vision at large that outperformed the existing approaches [21]. The encoder extracts the features while the decoder performs the segmentation based on the extracted features, which results in a remarkable performance on medical images. However, these encoder-decoder architecture has a drawback in learning multi-scale features in complex segmentation tasks and a weak ability to generalize to other tasks. The network structure needs to be optimized to be robust enough to make the parameter space wider and deeper to solve the problem. Although network widening and deepening increase network parameters and computational cost, which causes difficulty while training, causing the gradient to vanish during training [22]. Therefore, the challenge is to make the network wider and deeper without gradient vanishing and fewer parameters.

To overcome the above-aforementioned challenges, we propose a generalized encoder-decoder model called dense dilated inception network (DDI-Net) for medical image segmentation by modifying U-Net architecture. More specifically, we utilize three steps; firstly, we propose a dense path to replace the skip connection between the encoder and decoder to make the model deeper. Secondly, we replace the U- Net's basic convolution blocks with a modified inception module called multi-scale dilated inception module (MDI) to make the model wider without gradient vanishes and with fewer parameters. Thirdly, data augmentation and normalization was applied to the training data to improve the model generalization. We evaluated our DDI-Net on three subtasks of medical segmentation decathlon challenge (MSD) datasets [23]. The experimental results show that our proposed method outperformed the existing ones in each task. Our contribution to this paper is as follows:

- A generalized network named Dense Dilated Inception Network (DDI-Net) is proposed for medical image segmentation.

- We proposed a simple and efficient Preprocessing pipeline, which uses data normalization and augmentation for training and testing DDI-Net to improved segmentation generalization and accuracy.

- We conduct experiments with three different medical segmentation tasks to verify integrated components' performance and the overall model's generalization. The results show that our model outperforms other state-of-the-art models with fewer parameters.

The remaining part of this paper is as follows; we review the related work in Section II. In Section III represents our proposed DDI-Net. The experimental setup, including dataset preprocessing, implementation details, and evaluation, are describes in Section IV. Section V discusses the experiments to evaluate the effectiveness of our DDI-Net. Finally, we conclude in Section VI.

II. RELATED WORK

Nowadays, many encoder-decoder based architectures have been proposed for medical image segmentation. Based on recent studies, the encoder-decoder architecture, such as U-Net, has shown excellent performance due to its flexibility and extensible structure. Several extensions of U-Net have been proposed by integrating sophisticated network blocks such as residual network [24], dense network [25], inception module [26], and dilated convolution [27] for improving segmentation accuracy. Li et al. [25] proposed a hybrid densely U-Net (H-DenseU-Net) for 3D liver and tumor segmentation. H-DenseUNet combines densely connected paths and U-Net to improve performance. Alternatively, Yang et al. [28] propose

a U-Net with dilated convolution, and they called their structure DCU-Net for brain tumor segmentation.

Similarly, Chen et al. [29] embedded dense and residual blocks into a U-Net segmentation network. Ibtehaz and Rahman [30] combine a U-Net with residual inception modules for multi- scale feature extraction and perform segmentation on different modalities. Also, Wang et al. [31] integrate the inception module in U-Net architecture for segmentation of left atrial. Li and Tso [32] in cooperated inception modules and dilated inception modules in U-Net architecture for liver and tumor segmentation. Furthermore, Zang Z.et al [33] integrates the inception module with a dense connection into U-Net architecture. Jingcong L. et al. [34] replace the basic convolution block of U-Net architecture with a dilated inception block for multi-scale feature aggregation for cardiac right ventricle segmentation. Moreover, Bala S.B. and Kant S. [35] proposed a hybrid network. They combined CNN and Gated Recurrent Unit (GRU) using the U-Net structure to perform segmentation of cardiac MRI.

III. PROPOSED METHOD

In this study, inspired by U-Net, Dense-Net, Inception module, and Dilation convolution, we proposed a generalized medical segmentation model. The model was built upon U-Net based encoder-decoder architecture by integrating dense path and MDI blocks into U-Net. We modify U-Net by replacing the skip connection with the proposed dense path between the encoder and decoder and in cooperating MDI block to replace the basic convolutional block to improve the model's accuracy. Fig. 1 illustrate the proposed DDI-Net architecture. The DDI-Net comprises four dense paths, nine MDI blocks, four down sampling layers, four up sampling layers, and one output layer.
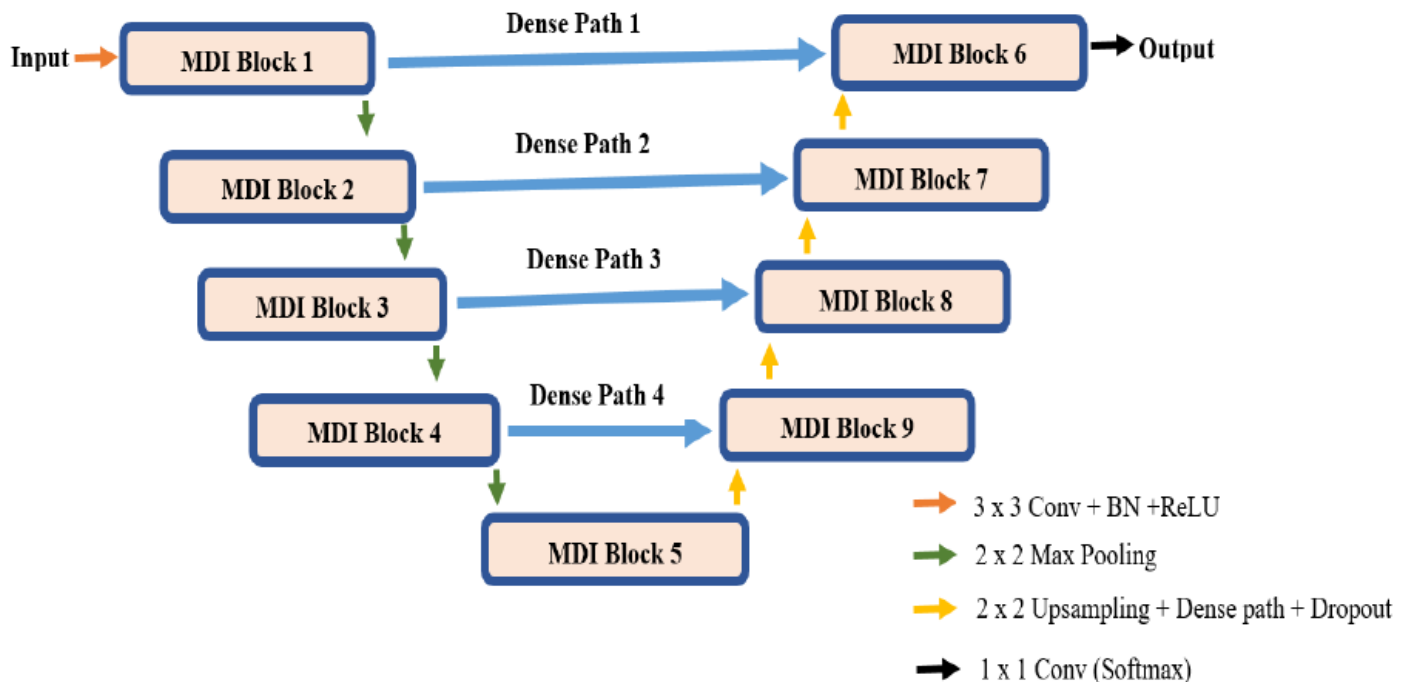

Fig. 1. Overall Architecture of the DDI-Net Model.

A. Dense Path

In U-Net, features extracted at the encoding path are pass using skip connections to their corresponding decoder path, which results in the passing of extra features forward, leading to the mortifying the exactitude of segmentation [36][37]. Also, we observe a large semantic gap between the encoder and the decoder feature map. Thus concatenation (feature fusion) of the feature maps from the encoder and the decoder will cause disparity during learning, thereby affecting segmentation prediction. Therefore to alleviate these challenges, we proposed to replace the skip connection with convolutional layers densely connected, which we referred to as dense path. Rather than merely concatenating the feature. As illustrated in Fig. 2, the dense path comprises densely connected convolution layers with 3x3 filters and a bottleneck layer. The dense path allows in-depth supervision to make the model deeper to allow the encoder to extract low-level features, thus helping the decoder recover the lost spatial information. The dense path also improved the flow of information and the gradient all over the network. This aids in alleviating the difficulty in training the network and hence reduces overfitting with its regularizing effect. Moreover, the dense path performs feature reuse to utilize the network's potential, with a resilient condensed model that is easy to train and highly parameter efficient.

B. Multi-Scale Dilated Inception Block

There are usually different scales of interest in medical image segmentation, such as tumors, lesions, and organs. Therefore, we need a network that can learn and extract multi-scale features with fewer parameter Networks models like googleNet [12] propose the inception module. The inception module consists of multiple convolutional layers with kernels of different sizes that learn multi-scale features. In each convolutional layer, the receptive field size is determined by the kernel size [38]. These kernel sizes include both small and large sizes. The small kernels are used to learn small scale features such as 1 x 1, 3 x 3, while the large scale kernel is used to learn large scale features such as 7 x 7 and 13 x 13[38]. According to [39] [40], multi-scale features improved the performance of the network model. However, large convolutional kernels used in obtaining large scale features increase the parameters and computational cost. To overcome this challenge, [39] apply dilated convolution. Dilated convolution is a convolution type that expands the receptive field to obtain large scale features using different dilation rates without increasing the parameters and computational cost. Inspired by the inception module [12] and dilated convolution [39], we propose a modified inception module by incorporating dilated convolutions called multi-scale dilated inception module (MDI). MDI module is developed to be used in the encoder as well as the decoder path to extract and aggregate the multi-scale feature maps. These feature maps are aggregated from kernels of different sizes with different dilation rates to widen the network to learn multi-scale features to improve the segmentation performance [41]. As depicted in Fig. 3, three convolutional layers with 3x3 kernels with four different dilation rates are used in the MDI module. The dilation rates are 1, 2, 4, and 6. Each convolutional kernel's feature scale is $(2l+1)^2$, where l is the kernel's dilation

rate. Features extracted from the dilated convolution result produce a different scale of 3 x 3, 5 x 5, 9 x 9, and 13 x 13, as illustrated in Fig. 4. The output of the four dilated convolution layers is concatenated. Batch normalization [42] is applied to accelerate the training and enhance the model's stability, followed by a 1x1 convolution to reduce the dimension and ReLU is used as the activation function for each convolutional layer [43].
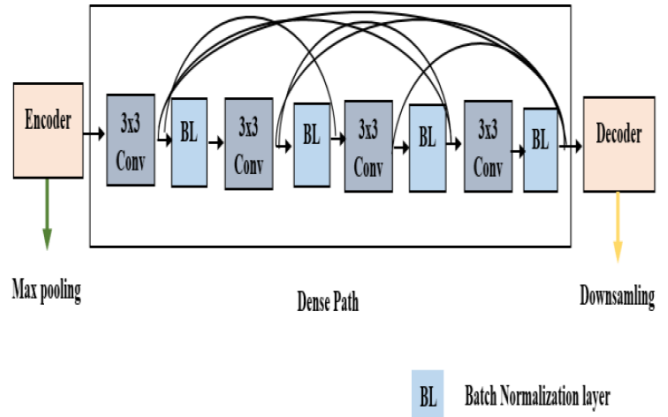

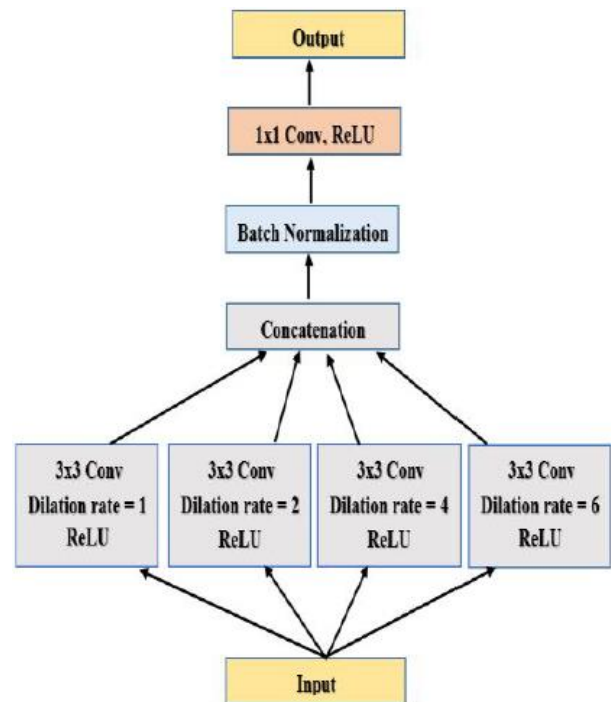Fig. 2. The Architecture of Dense Path.

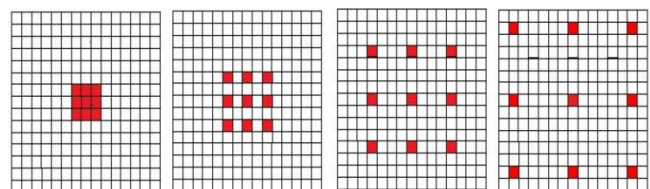Fig. 3. The Architecture of Multi-Scale Dilated Inception (MDI) Block.

Fig. 4. Dilated Convolution with 3x3 Kernel with a Dilation Rate of 1, 2, 4, and 6.

We modified the U-Net architecture by replacing the convolution block with the proposed MDI module. Experiments verified that our proposed MDI enhanced the segmentation performance by learning more multi-scale features without any free blow up in computational complexity [37] with fewer parameters than the original inception module.

IV. EXPERIMENTAL SETUP

The experimental setup, including the preprocessing of the dataset, implementation description, and evaluation metrics, is discussed in this section.

A. Datasets

We use three subtasks from the decathlon challenge dataset for medical segmentation. There are 484, 260, and 20 image data for brain tumors, hippocampus, and heart tasks. In Table I, the dataset is briefly outlined.

1) Preprocessing of training and testing data: Various scanners, institutions, and anatomical structures with different pixel spacing were used in collecting training and testing data used. Hence, these differences make it very important to preprocess the training and testing data before feeding our model. Fig. 5 shows the overview of the preprocessing steps followed during training and testing. Precisely, we performed image resampling to make the pixel spacing of all the images the same, and then we normalized the images. Lastly, data augmentation is applied during the training and testing process to improved generalization.

a) Image Resampling: Since the dataset used for both training and testing, the experiments are from three different datasets with pixel spacing ranging from 1mm to 1.25mm.We do image resampling to eliminate the difference. For brain MRI the pixel spacing is 1mm x 1mm x 1mm, while the hippocampus 1mm x 1mm x 1mm and heart is 1.25mm x 1.25mm x 2.70mm. Therefore, we resample the heart images to 1mm to make the spatial resolution the same as the brain and hippocampus images. After image resampling, we applied intensity normalization to the three datasets' images to normalize the image.

b) Data Normalization: We normalize the images using intensity normalization by subtracting the volume's mean and dividing by the volume standard deviation to the range of 0, 1.

After normalization, we applied augmentation to increase the training data to improve model generalization and avoid overfitting.

c) Data Augmentation: Data augmentation increases the training data by artificially generating more training data to generalize the model. The training data is augmented by;

- Random rotation of angle between -5 and 5 degrees.
- Vertical flipping with a probability of 0.2 for increasing the orientation variety.
- Random image scaling with a scale factor s: s E [0.2, 0.6] to maximize the images' variance.

B. Implementation Details

The network model has been implemented using keras [44] with tensorflow [45] backend using python 3 programming languages. Our network was trained and tested on a desktop computer with NVIDIA GeForce RTX 2080Ti with 11 GB of memory and 2 graphics card. During the training, the network was initialized with the normal weight [46], 0 bias, 0.0001 learning rate, and cross-entropy as our loss function. We optimize our network with Adam optimizer [47] with Beta-1=0.90, Beta-2 = 0.99 and epsilon = 0.000001. We executed 5-fold cross-validation and trained the model for 100 epochs. After every epoch, we evaluate the model using the validation data, and then the best model is selected for evaluating the test data. For the training and validation, we use a batch size of 4. In each epoch, 4 data is transposed to the model as input. All layers use a Rectified Linear Unit (ReLU) as an activation function except the output layer that uses softmax. We use batch normalization to normalize the feature maps and stabilize the network while training.

C. Evaluation Metric

The performance of our model is to assess using the Dice score. It is evaluated as;

$$Dice\ (GT, SR) = \frac{2|GT \wedge SR|}{(|GT| + |SR|)} \tag{1}$$

GT and SR are ground truth and segmentation results, respectively. Ground truth is the segmented region extracted by experienced experts manually using standard annotation protocol. In contrast, the segmentation result is the segmented region from the evaluated method.

TABLE I.   GENERAL DESCRIPTION OF THE DATASET

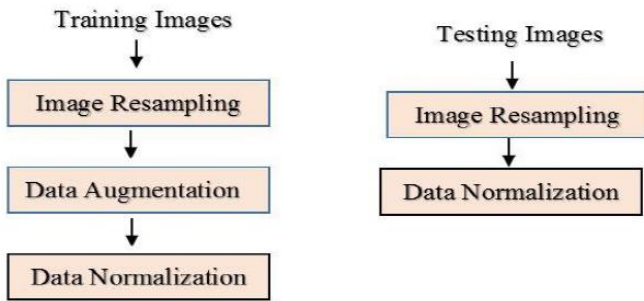| Name	Modality	Number of subjects	Scanners	Source	Image Spatial Resolution	Target
Brats 2016 & 2017	Multimodal Multisite MRI Data FLAIR, T1w, T1gd, T2w	484	1T to 3T	Brats 2016 & 2017 Datasets	Gliomas segmentation necrotic, active tumor and oedema	$1 \times 1 \times 1mm^3$
LASC	Mono-modal MRI	20	1.5T Achieva Scanner Philips	Kings College London	Left Atrium	$1.25 \times 1.25 \times 1.25mm^3$
Hippocampus	Mono-modal MRI	260	1.5T Achieva Scanner Philips	Vanderbilt University Medical Centre	Hippocampus Head and Body	$1 \times 1 \times 1mm^3$

Fig. 5.   Overview of Preprocessing of Images Steps during Training and Testing.

V.   EXPERIMENTAL RESULT

This section evaluates our proposed model's effectiveness and generalizability on three separate segmentation tasks, including multimodal MRI segmentation of brain tumors, mono-modal MRI segmentation of the hippocampus, and MRI segmentation of the heart.

A.   Brain Tumour Segmentation

We experiment with brain MRI images for brain tumor diagnosis of glioma to test our model's efficacy. The most common brain tumor found in the brain and spinal cord is a glioma. Due to the diverse and heterogeneously positioned targets shown in Fig. 6, glioma segmentation is a difficult task. This segmentation is targeted at segmenting glioma into edema, tumor non-enhancement, and tumor enhancement. 484 multi-parametric magnetic resonance imaging (MRI) scans from patients diagnosed with glioblastoma or lower grade glioma with the same number of ground-truth images are included in the brain dataset given. The proposed method uses all four sequences to segment brain MRI images, comprising volumes of Native T1-weighted (T1), Post-contrast T1-weighted (T1-Gd), T2-weighted (T2), and T2-fluid attenuation inversion recovery (FLAIR). 70 % of the data in this experiment is used for training, 15 % for validation, and 15% for testing. To get an accurate and stable model, we performed a 5-cross validation. DDI-Net results were contrasted with two recently published state-of-the-art models, and the outcome is shown in Table II. The results of the dice score obtained from DDI-Net demonstrated superior performance over the existing models.

B.   Hippocampus Segmentation

The hippocampus is a complex organ of the brain embedded deep in the temporal lobe. In learning and memory, it has the most responsible function. For Alzheimer's disease (AD) diagnosis, hippocampus segmentation is essential. As shown in Fig. 6, a complicated task is hippocampus segmentation. It has two adjacent tiny structures with high precision. The data set consisted of 260 stable adults and adults with non-affective psychotic illness, taken from the Vanderbilt University Medical Center phenotype data repository. 70% of the data in this experiment is used for training, 15% for validation, and 15% for testing.

To get an accurate and stable model, we performed a 5-cross validation. The hippocampus's entire MRI is used as the input to the network, as shown in Table III. Compared to the

other two art method states, our proposed method gets the highest result.

C.   Heart Segmentation

The heart is one of the human body's vital organs that pump blood throughout the body. Segmentation of the Left atrial from the heart plays a vital role in diagnosing atrial fibrillation (AF).

Segmentation of the left atrial from the heart is challenging because of the small training dataset with considerable variability, as shown in Fig. 6. The provided dataset consists of 20 MRI images from the left atrial segmentation challenge (LASC), Kings College Kingdom, London, United Kingdom. We use the whole MRI of the heart as input to the network. As shown in Table IV, the best result compared to other method states is obtained by our proposed method.

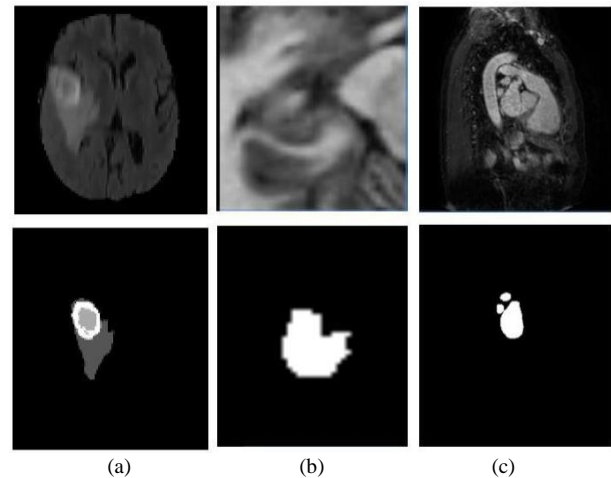(a)   (b)   (c)

Fig. 6.   A Sample of the Three Tasks of Segmentation. The First Row Demonstrates the Original Images and the Ground Truths was Shown in the Second Row. (a) Brain Tumor. (b) Hippocampus. (c) Heart.

TABLE II.   COMPARISONS RESULT ON BRAIN TUMOR SEGMENTATION

Networks	Edema	Non-enhancing	Enhancing
NDN	0.71	0.60	0.72
nnU-Net	0.68	0.48	0.68
DDI-Net (ours)	**0.82**	**0.68**	**0.79**

TABLE III.   COMPARISONS RESULT ON HIPPOCAMPUS SEGMENTATION

Networks	Anterior	Posterior
NDN	0.88	0.89
nnU-Net	0.90	0.89
DDI-Net (ours)	**0.92**	**0.90**

TABLE IV.   COMPARISONS RESULT ON HEART SEGMENTATION

Networks	Left Atrial
NDN	0.85
nnU-Net	0.93
DDI-Net (ours)	**0.95**

D. Ablation Studies

We propose and introduce dense paths and the MDI blocks to improve the baseline encoder-decoder-based U-Net model's segmentation accuracy in the proposed method.

To verify these introduced modules' effectiveness, we conduct the following ablation studies to investigate their contributions to the overall DDI-Net performance. We use the heart dataset for the ablation studies because it is the most challenging dataset used in our experiment. Hence, we make a comparison among the U-Net, U-net with dense paths (U-Net + dense path), and U-net with MDI blocks (U-Net +MDI) and the DDI-Net (U-Net+ Dense path +MDI). We initially start with the baseline U-Net and then assess the dense path and MDI block's effect on the results.

1) Ablation study for replacing the skip connection with the dense path: To verify the dense path's effectiveness, we replaced the skip connection with the proposed dense path. Table V illustrates the segmentation result. The results show that we achieved 0.9 on the dice score compare to 0.89 in the original U-Net. Our result signifies that the dense path proposed has improved the segmentation accuracy, making the network deeper and without a vanishing gradient. The dense path also alleviates the semantic gap between the encoder and the decoder by adding more blocks of convolutional operation and dense connection, which aids in a proper fusion of the feature maps.

2) Ablation study for replacing the convolutional layer with MDI blocks: To verify MDI blocks' effectiveness, we replaced the basic convolutional blocks with MDI blocks. Table VI depicted the segmentation results. The results illustrate that we achieved 0.93 on the dice score compare to 0.89 in the original U-Net. We observed that MDI blocks make the network wider; this aid in extracting multi-scale features from different scales. This indicates that using a filter of different sizes allowed the network to capture multi-scale features and improved the segmentation result.

3) Ablation study for the proposed DDI-Net: To verify the effectiveness of DDI-Net, We experimented with dense path and MDI blocks together. The results of the comparison are depicted in Table VII. Our results show that we achieved 0.95 on the dice score compare to 0.89 in the original U-Net. Table VII shows that the DDI-Net contributes to improving medical image segmentation's performance and accuracy. The cooperation between these two proposed modules into U-Net has yielded the best result from the segmentation results.

E. Evaluating the Effect of Data Normalization and Data Augmentation on DDI-Net Generalization

Using two data normalization and three data augmentation techniques, including image resampling, intensity normalization, rotation, flipping, and scaling, this section verifies the efficacy of data normalization and data augmentation on DDI-Net generalizability. We trained DDI-Net using all three datasets with the same setting to analyze the impact of data normalization and augmentation in model generalization. Firstly, we experiment with data normalization only. Secondly, we experimented with data augmentation and experimented with normalization and augmentation of data, as seen in Table VIII. From Table VIII, it indicates that the data normalization and augmentation increase Dice score result. By integrating data normalization and augmentation operations, the best segmentation efficiency is obtained for all three datasets.

F. Algorithm Run-Time

Table IX shows the training and testing time for all the models in each experiment. It can be found that in both segmentation tasks, the proposed model requires less time for training and testing compared to nnU-Net and NDN. Besides, brain data requires more time than the hippocampus and heart dataset for training and research.

G. Comparison with State-of-the-Art Methods

To verify the effectiveness of our proposed improvements with the state of the art methods. We compare our method with two proposed methods by Wang L. et al. [48] and Isensee F.et al. [49].For the brain and hippocampus dataset, the result is from the papers. For the heart dataset, Wang L.et al. do not perform implementation with the heart dataset. We obtained the result using Wang L.et al; implementation details and Isensee F. result from their paper. Tables II, III, and IV show the two methods' dice score and the DDI-Net on the three datasets. As observed visually from the tables, the proposed DDI-Net improves the segmentation's accuracy and generalizes all three datasets. Fig. 7 visually illustrates the output results of the DDI-Net proposed.

TABLE V.   ABLATION STUDY FOR REPLACING THE SKIP CONNECTION WITH DENSE PATH

Networks	Heart Segmentation Left Atrial	Parameters
U-Net	0.89	2.81M
U-Net + Dense path	**0.90**	**1.75M**

TABLE VI.   ABLATION STUDY FOR REPLACING THE CONVOLUTION BLOCK WITH MDI BLOCKS

Networks	Heart Segmentation Left Atrial	Parameters
U-Net	0.89	2.81M
U-Net + MDI	**0.93**	**1.10M**

TABLE VII.   ABLATION STUDY FOR REPLACING THE SKIP CONNECTION AND CONVOLUTION BLOCK WITH DENSE PATH AND MDI BLOCKS

Networks	Heart Segmentation Left Atrial	Parameters
U-Net	0.89	2.81M
DDI-Net	**0.95**	**0.90M**

TABLE VIII.   PERFORMANCE OF DDI-NET WITH DIFFERENT TRAINING CONFIGURATION

Configuration Data Norm	Data Aug	Brain Segmentation Edema	Non-enhancing	Enhancing	Hippocampus Segmentation Anterior	Posterior	Heart Segmentation Left Atrial
X	√	0.73	0.50	0.71	0.85	0.81	0.82
√	X	0.79	0.62	0.75	0.89	0.86	0.89
√	√	**0.82**	**0.68**	**0.79**	**0.92**	**0.90**	**0.93**

TABLE IX.   MODEL TRAINING AND TESTING TIME

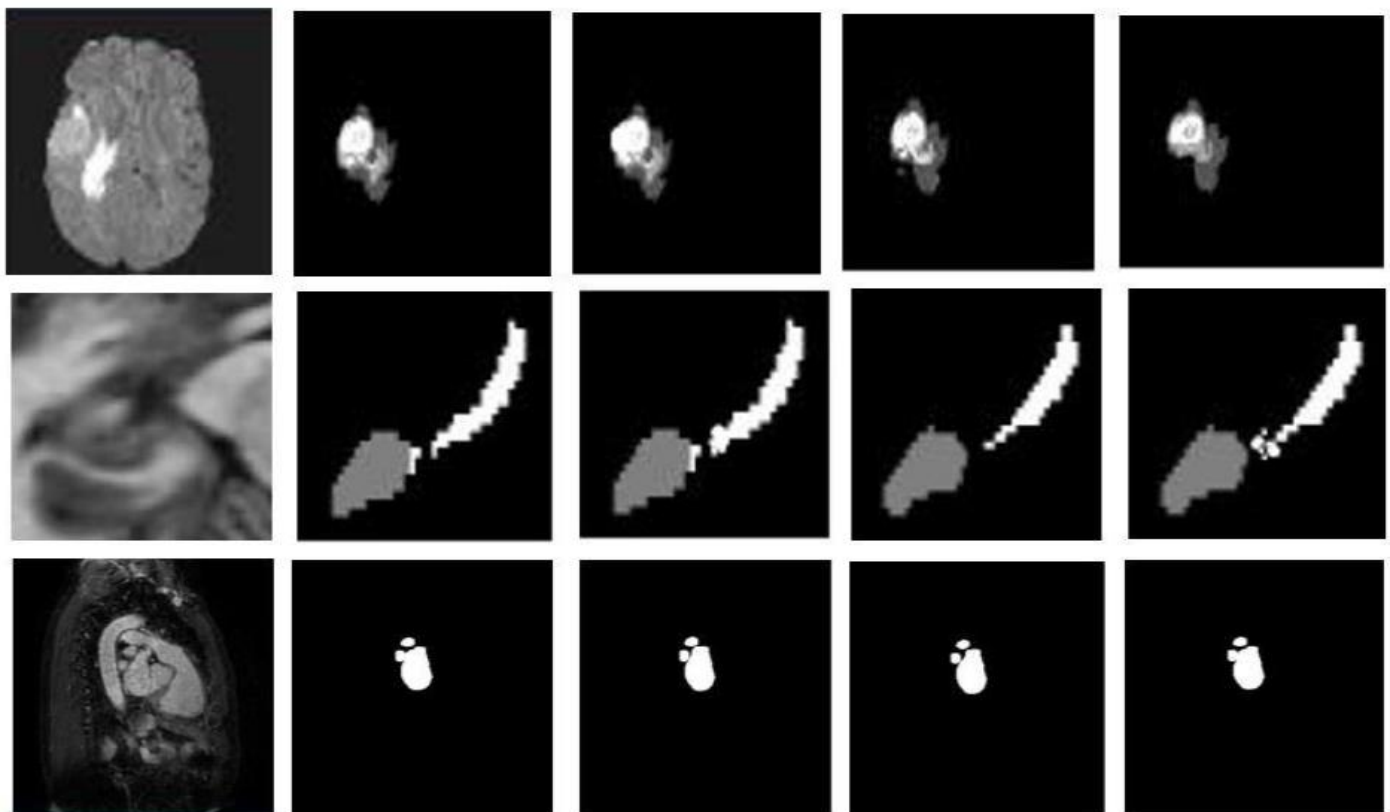| Networks	Brain Segmentation Training Time	Testing Time	Hippocampus Segmentation Training Time	Testing Time	Heart Segmentation Training Time	Testing Time
NDN	8.2h	1.5sec	7.4h	0.91sec	4.7h	0.8sec
nnU-Net	7.6h	1.8sec	8.1h	1.32sec	5.2h	1.3sec
DDI-Net (ours)	**7.5h**	**0.9sec**	**6.5h**	**0.68sec**	**4.3h**	**0.7sec**

Fig. 7.   Visual Illustration of Segmentation Output on Brain, Hippocampus and Heart Dataset. from Left to Right: Original Image, Ground truth, DDI-Net, NDN and nnU-Net.

VI. CONCLUSION

In this paper, by modifying the U-Net architecture using Dense-Net, Dilated Convolution, and Inception network, we propose a new encoder-decoder network called DDI-Net. There are two features on the DDI-Net, namely dense paths and MDI blocks. The dense path enables in-depth supervision to deepen the model. Low-level features can be extracted by the encoder, thus helping the decoder recover the missing spatial information. This helps to facilitate the reuse of features with a resilient simplified training path and highly efficient parameters.

The MDI block, meanwhile, makes the model wider without the gradient vanishing but with fewer parameters. Besides, using data normalization and augmentation, we propose a general training and testing process. The experiments conducted show that they play an essential role in generalizing the model across images from various tasks. To prove the DDI-Net generalization, the model is tested using

three distinct medical image segmentation tasks. The result shows that our DDI-Net performs better than the state of the art method, including nnU-Net and NDN. However, we have limited the tasks only to MRI images. Therefore, to make the network more general and efficient for all medical problems in the future, we plan to experiment with other modalities.

REFERENCES

[1] S.A.Bala, S.Kant. " Deep learning based model architectures for cardiac MRI segmentation: A Survey," International Journal of Innovative Science, Engineering & Technology, pp.129-135, 2020.

[2] N.M.Aboelenein, S.Piao, K. Anis, N.Alam, and A.Ahmed. "HTTU-Net: Hybrid Two Track U-Net for automatic brain tumor segmentation," IEEE Access, 2020.

[3] Lindsey, Tony, and J.J.Lee." Automated Cardiovascular Pathology Assessment Using Semantic Segmentation and Ensemble Learning," Journal of Digital Imaging, pp. 1-6, 2020.

[4] S.Almotairi, G. Kareem, M. Aouf, B. Almutairi, and M.A.M. Salem. "Liver tumor segmentation in CT scans using modified segnet," Sensors,5 (20), pp. 1516, 2020.

[5] C.Liu, R. Zhao, W.Xie, and M.Pang." Pathological lung segmentation based on random forest combined with deep model and multi-scale superpixels." Neural Processing Letters, pp.1-19, 2020.

[6] Lundervold, S.Alexander, and L.Arvid. "An overview of deep learning in medical imaging focusing on MRI," Zeitschrift für Medizinische Physik, 2(29), pp. 102-127, 2019.

[7] G.Litjens, et al. "A survey on deep learning in medical image analysis," Medical Image Analysis, 42, pp. 60-88, 2017.

[8] L.Yann, B.Yoshua, and H.Geoffrey. "Deep learning," Nature, 521(7553), pp. 436-444, 2015.

[9] S.A. Bala, S.Kant and K. Kumar." Impact of deep learning in medical imaging: a systematic new proposed model," International journal of recent technology and engineering, pp.112-118,2019.

[10] D.Daimary, M.B.Bora, K.Amitab, and D.Kandar. "Brain Tumor Segmentation from MRI Images using Hybrid Convolutional Neural Networks," Procedia Computer Science, 167, pp. 2419-2428,2020.

[11] S.M.Ibrahim, M.S.Ibrahim, Shakeel Muhammad, M.Usman, I.Naseem, and M.Moinuddin. "Heart Segmentation From MRI Scans Using Convolutional Neural Network," arXiv preprint arXiv:1911.09332, 2019.

[12] M.Liu, L.Fan, Y.Hao, W.Kundong, M.Yixin, and X.Mingqing. "A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease." NeuroImage 208, 2020.

[13] A. Krizhevsky, I.Sutskever, and G.E Hinton. "ImageNet classification with deep convolutional neural networks," Advances in neural information processing systems, pp.1097-1105,2012.

[14] K.Simonyan, A. Zisserman." Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556,2014.

[15] Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition, 2015.

[16] G.Huang, L.Zhung, M. Laurens, and K.Q Weinberger Gao. "Densely connected convolutional networks," Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700-4708 2017.

[17] K.He, Z.Xiangyu, R.Shaoqing, and S.Jian. "Deep residual learning for image recognition," Proceedings of the IEEE conference on computer vision and pattern recognition, pp.770-778, 2016.

[18] L.C.Chen, G.Papandreou, I.Kokkinos, K.Murphy, and A.L Yullie. "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," IEEE transactions on pattern analysis and machine intelligence, 40 (4), pp.834-848, 2017.

[19] J.Long, E.Shelhamer, and T. Darrell. "Fully convolutional networks for semantic segmentation," Proceedings of the IEEE conference on computer vision and pattern recognition, pp.3431-3440, 2015.

[20] O.Ronneberger, P. Fischer, and T.Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, pp. 234-241 2015.

[21] L.Liu, J.Cheng, Q.Quan, F.Wu, Y.P.Wang, and J.A Wang." A survey on U-shaped networks in medical image segmentation," Neurocomputing, 409, pp.244-258, 2020.

[22] L.Bi, J.Kim, A.Kumar, M.Fulham, D.Feng." Stacked fully convolutional networks with multi-channel learning: application to medical image segmentation." The Visual Computer 33, pp.1061-1071 2017.

[23] Simpson, Amber L., et al. "A large annotated medical image dataset for the development and evaluation of segmentation algorithms," arXiv preprint arXiv:1902.09063, 2019.

[24] M.Z.Alom, M.Hassan, C.Yakopcic, T.M.Taha,V.K.Asari. "Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation," arXiv preprint arXiv:1802.06955 , 2018.

[25] X.Li, H.Chen, X.Qi, Q.Dou, C.W.Fu, P.A Heng. "H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes," IEEE transactions on medical imaging 37(12), pp.2663-2674, 2018.

[26] D.E.Cahall, G.Rasool, N.C.Bouaynaya, H.M.Fathallah-Shaykh. "Inception modules enhance brain tumor segmentation," Frontiers in computational neuroscience,13, pp.44, 2019.

[27] A.Tureckova and A. J. Rodríguez-Sánchez. "ISLES challenge: U-shaped convolution neural network with dilated convolution for 3D stroke lesion segmentation." International MICCAI Brainlesion Workshop. Springer, Cham, pp.319-327, 2018.

[28] T.Yang, Y.Zhou, L.Li, and C.Zhu." DCU-Net: Multi-scale U-Net for brain tumor segmentation." Journal of X-Ray Science and Technology Preprint, pp. 1-18, 2020.

[29] L.Chen, P.Bentley, K.Mori, K.Misawa, M.Fujiwara, and D.Rueckert "DRINet for medical image segmentation," IEEE transactions on Medical Imaging, 37 (11),pp.2453-2462, 2018.

[30] N.Ibtehaz and M.S. Rahman. "MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation," Neural Networks 121,pp.74-87,2020.

[31] C.Wang, M.Rajchi, A.D Chan, and E. Ukwatta. "An ensemble of U-Net architecture variants for left atrial segmentation," Medical Imaging 2019: Computer-Aided Diagnosis. Vol. 10950. International Society for Optics and Photonics, 2019.

[32] S.Li and G.K. Tso. "Bottleneck supervised u-net for pixel-wise liver and tumor segmentation," arXiv:1810.10331, 2018.

[33] Z.Zhang, C.Wu, S.Coleman, D.Kerr." DENSE-INception U-net for medical image segmentation," Computer Methods and Programs in Biomedicine (192),2020.

[34] J.Li, Z.L.Yu, Z.Gu, H.Liu, and Y.Li. "Dilated-inception net: multi-scale feature aggregation for cardiac right ventricle segmentation." IEEE Transactions on Biomedical Engineering 66(12), pp. 3499-3508, 2019.

[35] S.A Bala and S. Kant S. "A hybrid spatio-temporal network for cardiac MRI segmentation," International journal of advanced trends in computer science and engineering, pp.9089-9097,2020.

[36] Q.Tong, et al. "RIANet: Recurrent interleaved attention network for cardiac MRI segmentation," Computers in biology and medicine, 109,pp.290-302, 2019.

[37] Z.Zhang, X.Zhang, C.Peng, X.Xue and J.Sun." Exfuse: Enhancing feature fusion for semantic segmentation." Proceedings of the European Conference on Computer Vision (ECCV),pp.269-284,2018.

[38] W.Luo, Y.Li, R.Urtasun, and R.Zemel. "Understanding the effective receptive field in deep convolutional neural networks," In Advances in Neural Information Processing Sytems, pp. 4898–4906, 2016.

[39] A. Krizhevsky, I. Sutskeyer, and G.E Hinton. "Imagenet classification with deep convolutional neural networks," In Advances in Neural Processing Systems,pp.1097-1105,2012.

[40] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," arXiv preprint arXiv:1511.07122, 2015.

[41] C. Peng, X.Zhang, G.Yu, G.Luo, and J. Sun. "Large kernel matters-improve semantic segmentation by the global convolutional network," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.4353-4361,2017.

[42] S.Ioffe, and C. Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." arXiv preprint arXiv:1502.03167,2015.

[43] A.F.Agarap. "Deep learning using rectified linear units (ReLU)." arXiv preprint arXiv:1803.08375, 2018.

[44] F.Chollet. "Keras: The python deep learning library," [online] Available: https://keras.io/.

[45] M.Abadi et al., "Tensorflow: Large-scale machine learning on heterogeneous systems," arXiv:1603.04467, 2015.

[46] D. P. Kingma, J. Ba, "Adam: A method for stochastic optimization," Int. Conf. Learn. Represent, pp. 1-13, 2015.

[47] K. He, X. Zhang, S. Ren, J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," Proceeding of IEEE International Conference on Computer Vision, pp. 1026-1034, 2015.

[48] L.Wang, R. Chen, S.Wang, N.Zeng, X.Huang, C.Liu. "Nested Dilation Network (NDN) for Multi-Task Medical Image Segmentation." IEEE Access. 1(7) ,pp.44676-44685,2019.

[49] F. Isensee, et al. "nnU-net: Self-adapting framework for u-net-based medical image segmentation." arXiv preprint arXiv:1809.10486 ,2018.