Regular paper | Journal of the KIECS. pp. 529-536, vol. 17, no. 3, Jun. 30. 2022, t. 113, pISSN 1975-8170 | eISSN 2288-2189
http://dx.doi.org/10.13067/JKIECS.2022.17.3.529

COVID-19 폐 CT 이미지 인식

수징제* · 김강철**

COVID-19 Lung CT Image Recognition

Jingjie Su* · Kang-Chul Kim**

요 약

지난 2년 동안 중증급성호흡기증후군 코로나바이러스-2(SARS-CoV-2)는 점점 더 많은 사람들에게 영향을 미치고 있다. 본 논문에서는 COVID-19 폐 CT 이미지를 분할하고 분류하기 위해서 서브코딩블록(SCB), 확장 공간파라미드풀링(ASSP)와 어텐션게이트(AG)로 구성된 혼합 모드 특징 추출 방식의 새로운 U-Net 컨볼루션 신경망을 제안한다. 그리고 제안된 모델과 비교하기 위하여 FCN, U-Net, U-Net-SCB 모델을 설계한다. 제안된 U-Net-MMFE 는 COVID-19 CT 스캔 디지털 이미지 데이터에 대하여 atrous rate가 12이고, Adam 최적화 알고리즘을 사용할 때 다른 분할 모델에 비하여 94.79%의 우수한 주사위 분할 점수를 얻었다.

ABSTRACT

In the past two years, Severe Acute Respiratory Syndrome Coronavirus-2(SARS-CoV-2) has been hitting more and more to people. This paper proposes a novel U-Net Convolutional Neural Network to classify and segment COVID-19 lung CT images, which contains Sub Coding Block (SCB), Atrous Spatial Pyramid Pooling(ASPP) and Attention Gate(AG). Three different models such as FCN, U-Net and U-Net-SCB are designed to compare the proposed model and the best optimizer and atrous rate are chosen for the proposed model. The simulation results show that the proposed U-Net-MMFE has the best Dice segmentation coefficient of 94.79% for the COVID-19 CT scan digital image dataset compared with other segmentation models when atrous rate is 12 and the optimizer is Adam.

키워드

COVID-19, Lung CT Images, Lesion Segmentation, Convolution Neural Networks
COVID-19, 폐 CT 이미지, 병변, 컨볼루션 신경망

* 전남대학교 대학원 컴퓨터공학과
(sujingjie0330@gmail.com)
** 교신저자 : 전남대학교 공학대학 전기컴퓨터공학부
• 접 수 일 : 2022. 05. 18
• 수정완료일 : 2022. 06. 02
• 게재확정일 : 2022. 06. 17
• Received : May. 18, 2022, Revised : Jun. 02, 2022, Accepted : Jun. 17, 2022
• Corresponding Author : Kang-Chul Kim
School of Electricity and Computer Engineering, Chonnam National University,
Email : kkc@jnu.ac.kr

Ⅰ. Introduction

Since December 2019, COVID-19 has been sweeping the world, and more than 514 million people worldwide have been confirmed infected, resulting in about 6.24 million deaths[1]. COVID-19 is a serious global epidemic[1]. In diagnosing the COVID-19 lesions, doctors always use CT scans. However, CT images often suffer from poor image quality, interference from noise, and unclear edge boundaries of tissues in the images. It is necessary to achieve accurate recognition by semantic segmentation[2] with the help of computers. Semantic segmentation relies on deep learning[3], and the Convolutional Neural Network(CNN)[4] is a classic deep learning model.

We propose a novel U-Net[5] Convolutional Neural Network with Mixed Mode Feature Extraction (U-Net-MMFE) to classify and segment COVID-19 lung CT images. The three modules of Sub Coding Block (SCB)[6], Attention Gate (AG)[7], and Atrous Spatial Pyramid Pooling (ASPP)[8] are embedded into the U-Net. The information transfer between the feature maps of each layer is enhanced by Sub Coding Block (SCB) during the same-layer coding process. The addition of the Attention Gate (AG) in the same-layer encoder-to-decoder process solves the problem that the semantic vector does not fully represent the information of the whole sequence. Using Atrous Spatial Pyramid Pooling (ASPP) in the back-propagation process at the bottom of the model preserves more image features. All modules we have embedded are lightweight without more extra parameters and complex operations, and our proposed model not only suppresses feature information loss effectively but also does not burden the computer. U-Net-MMFE obtains a higher Dice coefficient.

We design four different models such as FCN, U-Net and U-Net-SCB to compare the proposed model to them and choose the best optimizer for our model by using Adam, AdaDelta, AdaGrad, and SGDM[9~12].

The rest of this paper is organized as follows. In section Ⅱ, the theoretical foundations of semantic segmentation models and related work are introduced. In section Ⅲ, we describe the dataset, the structure of the proposed model, and how to evaluate the model. Experimental results and analysis are shown in section Ⅳ. Section Ⅴ concludes this paper.

Ⅱ. Related Work

Fully Convolutional Network (FCN)[13] is the originator of semantic segmentation and proposes to replace the fully connection layer with a convolutional layer to obtain a two-dimensional feature map. U-Net[5] is proposed based on FCN, and composed of the left encoder's contracting path and the right decoder's expansive path. U-Net adds an up-sampling operation after the fully connection layer. Up-sampling corresponds to down-sampling and is an operation process to increase the image resolution. The encoder-decoder architecture is a simple and effective structure in semantic segmentation.

Amine Amyar et al.[14] proposed a novel multi-task deep learning model and utilized helpful information contained in multiple related tasks. Their proposed model is based on U-Net with two group of decoders for image reconstruction and infection segmentation. The output feature information of the encoder is directly used as the input feature information of the decoder without any other operation. Although they also add a fully connected layer for classification, it does not solve the problem of feature information loss fundamentally. Yu-Huan Wu et al.[15] developed a

1) https://coronavirus.jhu.edu/map.html (accessed on 12 May 2022).

novel Joint Classification and Segmentation (JCS) system to perform real-time and explainable COVID-19 chest CT diagnosis. In their proposed model, VGG-16 Backbone Block is used as the encoder, Attentive Feature Fusion (AFF) is used as the decoder, and Enhanced Feature Module (EFM) is used to derive a more abundant feature map with various receptive fields. In this model, the last two fully connected layers of the VGG-16 Backbone Block in the encoder are deleted, but the model structure is more complex, increasing the burden on the computer. Aryan Mobiny et al.[16] proposed a novel learning architecture, called Detail-Oriented Capsule Networks (DECAPS), to automatically diagnose COVID-19. DECAPS used an Inverted Dynamic Routing mechanism which prevented the passage of information from non-descriptive regions and employed a Peekaboo training procedure to encourage the network to generate activation maps for every target concept. However, this model still cannot segment the lung CT lesion region of COVID-19 very clearly.

III. Materials and Methods

3.1 Dataset

The COVID-19 CT scan digital image dataset[17] used in this paper was created by Zhao et al. Considering the limited amount of original data, we increased the number of images by rotating images and adding noise on top of the original 397 COVID-19 negative CT images and 349 COVID-19 positive CT images. Our dataset is divided into training set, validation set, and test set. Each folder in turn contains two subfolders (NonCOVID-19 and COVID-19). The number of images for each type is shown in Table 1.

표 1. 코로나19 CT 데이터세트
Table 1. COVID-19 CT dataset

	NonCOVID-19	COVID-19
Training set	3851	3384
Validation set	1284	1128
Test set	1280	1120

Patients with COVID-19 show small slices on early CT images of the lungs and glassy structures[18] (circled in yellow in Fig. 1). The comparison of CT image of NonCOVID-19 and COVID-19 is shown in Fig. 1.

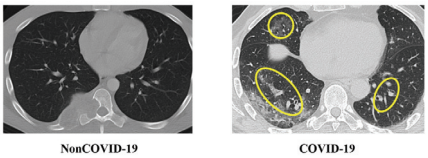


그림 1. NonCOVID-19 및 COVID-19 폐 CT 이미지
Fig. 1 NonCOVID-19 and COVID-19 Lung CT images

3.2 Image Pre-processing

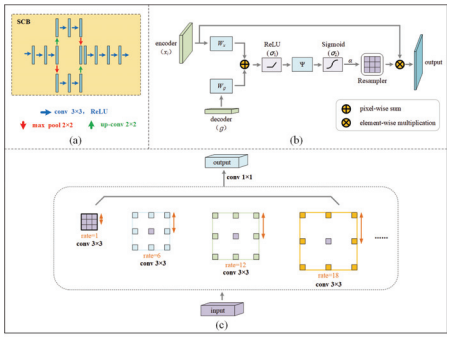
Histogram equalizations and filtering operations are used to pre-process the CT images. External interference may lead to large differences in the grayscale values of CT images. Both too large or too small differences in grayscale values can decrease the contrast of lung organ tissues. The histogram equalization operation is performed on the original image to make the gray values of the histogram evenly distributed in the range of 0-255 to increase the image's contrast.

After the histogram equalization, Gaussian filtering, Mean filtering, Median filtering, and Bilateral filtering[19~22] are used to remove noise interference. The filtering operation causes image blurring and distortion to a certain extent, and different filters have different blurring effects on the image.

3.3 Basic Module

The Sub Coding Block(SCB)[6] is used to

capture more useful information in the encoder. The SCB with simultaneous up-and down-sampling in each feature extraction module of the encoder stage is shown in Fig. 2(a). In the SCB, the feature map starts to go through three convolutional layers and then divides into two branches: up branch and down branch, the up branch by up-sampling operation and two convolutional operations followed by down-sampling operation, and the down branch by down-sampling operation and two convolutional operations followed by up-sampling operation. The features from two branches are added, and then three convolution operations are performed.



(a)SCB (b)AG (c)ASPP

그림 2. 기본 모듈
Fig. 2 Basic Module

In the encoder-decoder structure of the same layer of U-Net, the output of the encoder is used as input in the decoder. In the up-sampling process, the output of the lower layer is also used as input in the decoder of the upper layer. However, the feature information may be lost during the above operation. As shown in Fig. 2(b), the encoder output of the same layer and the decoder output of the lower layer are added in Attention Gate(AG)[7]. And the output of AG becomes the input of the decoder. There are two types of AG, additive and multiplicative[23,24]. Multiplicative attention module is faster and more space-efficient than the additive attention module in practice. By multiplying the attention coefficients α by the feature maps, this module focuses attention on the features that are useful for the final prediction.

AG provides better attention to feature regions (such as the glass structure of the COVID-19 CT image shown in Fig. 1) and prevents feature information loss.

As the depth of the U-Net increases, it loses a large amount of feature information during down-sampling, and the lost feature information cannot be restored during up-sampling, which affects the accuracy of image segmentation by convolutional neural networks.

To overcome the above problem, Atrous Spatial Pyramid Pooling(ASPP)[8] is used to restore feature information by filling the "0" between kernel values without increasing the parameters in the up-sampling operation, as shown in Fig. 2(c). Atrous rate represents the interval between kernel values. As the atrous rate becomes larger, the number of valid weights(the weights that are applied to the valid feature region) becomes smaller[25].

3.4 Proposed Model U-Net-MMFE

The three modules of SCB, AG, and ASPP are embedded into the U-Net structure to cooperate to achieve more accurate recognition and segmentation of the COVID-19 lesion region. Fig. 3 shows the proposed network structure, U-Net-MMFE.

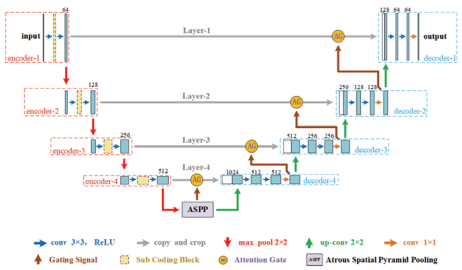


그림 3. U-Net-MMFE의 구조
Fig. 3 The Structure of U-Net-MMFE

The U-Net-MMFE model consists of encoders and decoders. The middle convolutional layer of each encoder in U-Net is replaced by an SCB, which deepens the network depth in the encoder. The SCB mixes the features in the image by up- and down-sampling. The output features of encoder-i are used as the input of AG in the same layer and next encoder-i+1. One decoder has three parts: add operation, two 3×3 convolutions operations, and one 1×1 convolution operation.

The input image size is 480×480. After the SCB operation and convolution operation, the image size remains unchanged, channels are 64. After the down-sampling, the image size is half of the original size (240×240) , and the channels are doubled (128) . After three times down-sampling operations, the image size becomes a 30×30 feature map, and the convolution operation in the bottom layer of U-Net is replaced by ASPP in order to restore the lost feature information during down-sampling. During up-sampling, the image size is enlarged to double size of the previous input (60×60) . After three times up-sampling operations, the image size is restored to 480×480. The decoder-1 outputs the segmentation result.

3.5 Evaluation Methods

To evaluate the segmentation performance of the algorithm, three metrics, C_{Dice} is Dice coefficient, P Precision, and R Recall are used in the experiments[26]. The calculation formula for the indicators is respectively, as follows :

$$C_{Dice}(A,B) = \frac{2\,TP}{2\,TP+FP+FN} = \frac{2|A \cap B|}{|A|+|B|} \quad (1)$$

$$P(Precision) = \frac{TP}{TP+FP} \quad (2)$$

$$R(Recall) = \frac{TP}{TP+FN} \quad (3)$$

Where A is the set of the marked COVID-19 lesion region. B is the set of the COVID-19 lesion regions obtained by segmentation. TP represents the number of correctly detected as positive samples, FP represents the set of incorrectly detected as positive samples, and FN represents the number of falsely detected as negative samples. The value range of the Dice coefficient is $[0,1]$. As Dice value gets larger, the closer the result of the algorithm segmentation is to the result of the label, and the better the segmentation effect. If the R is too low, it means that more COVID-19 lesion region have not been segmented, and if the P is too low, it means that the segmentation is not accurate enough.

Ⅳ. Experimental Results and Analysis

4.1 Pre-processing Results

The histograms for the original and the equalized CT image are shown in Fig. 4, and the red line is the CDF(Cumulative Distribution Function).

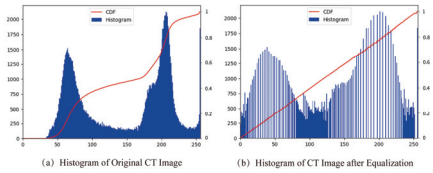


그림 4. 원본과 평활화 이미지 히스토그램
Fig. 4 The histograms for original and equalized image

The results show that the histograms of the equalized CT images have a relatively uniform gray distribution, and the CDF curve indicates that the bright and dark parts of the equalized CT images have a uniform distribution, and the image quality is high.

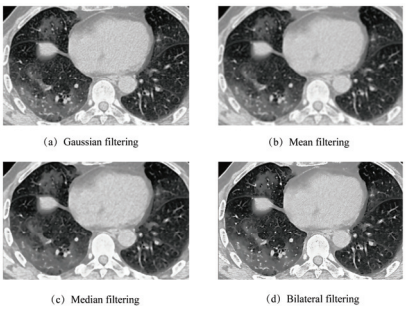


그림 5. 필터 효과 비교
Fig. 5 Comparison of filtering operation

Fig. 5 shows that the noise of the image is significantly reduced and the image is smoothed after the Gaussian filtering operation. Although some slight blurring is obtained, the image features are wholly preserved.

4.2 Optimizer

Table 2 shows the performance results for four optimizers after 50 epochs of training test. The Adam optimizer has the best results for Dice coefficient and is used in U-Net-MMFE.

표 2. 최적화 알고리즘 결과 비교
Table 2. Comparison of Optimizers

Optimizer	P %	R %	C_{Dice} %
Adam[9]	89.43	88.13	87.31
AdaDelta[10]	70.55	68.01	69.26
AdaGrad[11]	71.07	65.68	68.27
SGDM[12]	74.36	67.29	70.65

4.3 Determination of atrous rate

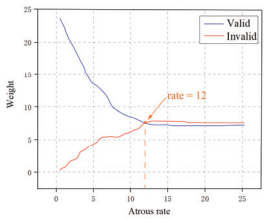


그림 6. atrous rate 에 따른 가중치의 변화곡선
Fig. 6 Variation curve of Weights with atrous rate

The training is set with batch_size=32, epoch=100, iteration=220, and learning rate=0.0001. In Fig. 6, both valid and invalid weights start to reach the saturation state when atrous rate=12, and this atrous rate is used in U-Net-MMFE.

4.4 Results and Comparison

Four models FCN, U-Net, and U-Net-SCB and U-Net-MMFE are designed to measure the effects of the mixed mode feature extraction in U-Net-MMFE.

In Fig. 7(a)(c), Dice coefficient of U-Net-MMFE is higher than the other models, and it has saturated at around 30 epochs. In Fig. 7(b)(d), the loss curves of the U-Net-MMFE are fewer than the other models.

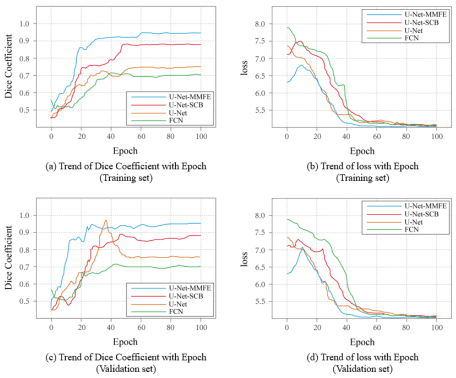


그림 7. 4 모델에 대한 주사위 계수 및 손실 곡선
Fig. 7 Dice Coefficient and loss curves for 4 models

표 3. 4 모델에 대한 COVID-19 초점 영역 분할 비교
Table 3. Comparison of COVID-19 focal region segmentation for 4 models

Model	P %	R %	C_{Dice} %
FCN	74.85	66.36	70.35
U-Net	80.16	71.12	75.37
U-Net-SCB	90.01	87.63	88.36
U-Net-MMFE	95.36	94.22	94.79

The results of 4 models for COVID-19 focal region segmentation performance metrics on the test set are shown in Table 3. And U-Net-MMFE model has better precision, recall, and Dice coefficient for COVID-19 focal regions segmentation performance than 3 other models.

표 4. U-Net-MMFE와 다른 논문의 결과 비교
Table 4. Comparison of U-Net-MMFE to other papers

Model	C_{Dice} %
Amine Amyar et al.'s Model[14]	88.00
Yu-Huan Wu et al. 's Model JCS[15]	78.50
Our's Model U-Net-MMFE	94.79

Table 4 shows that our proposed model has the best Dice coefficient of 94.79% compared to other recent papers.

Ⅴ. Conclusion

Accurate segmentation of lesion regions of COVID-19 lung CT images is an important application of artificial intelligence technology in the medical field. In this paper, SCB, AG, and ASPP are embedded in U-Net model. Compared with the recent COVID-19 lung CT images segmentation models, U-Net-MMFE has advantages in the segmentation of COVID-19 lung CT images, more lightweight, and the Dice coefficient reached 94.79%. The proposed U-Net-MMFE model can assist physicians in diagnosing COVID-19 patients.

References

[1] Y. Sun and Y. Lee, "A Semantic Diagnosis and Tracking System to Prevent the Spread of COVID-19," *J. of The Korea Institute of Electronic Communication Sciences,* vol. 15, no. 3, 2020, pp. 611-616.

[2] H. Seo, V. Vasudevan, H. Ren, R. Xiao, X. Jia and L. Xing, "Machine learning techniques for biomedical image segmentation An overview of technical aspects and introduction to state-of-art applications," *Int. J. of Medical Physics Research and Practice*, vol. 47, no. 5, 2020, pp. e148-e167.

[3] B. R. Kim, S. Shin, N. K. Kim, M. S. Park and H. J. Yoon, "Analysis of Floating Population in Schools Using Open Source Hardware and Deep Learning-Based Object Detection Algorithm," *J. of The Korea Institute of Electronic Communication Sciences,* vol. 17, no. 1, 2022, pp. 91-98.

[4] Y. Ma and K. C. Kim, "CNN Based 2D and 2.5D Face Recognition For Home Security System," *J. of The Korea Institute of Electronic Communication Sciences,* vol. 14, no. 6, 2019, pp. 1207-1214.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, October 2015, pp. 234-241.

[6] G. Haijun, Z. Xiangyin, P. Dazhi, and Z. Bochuan, "Rectal tumor segmentation method based on improved U-Net model," *J. of Computer Applications,* vol. 40, no. 8, 2020, pp. 2392-2397.

[7] O. Oktay, J. Schlemper, M. Lee, M. Heinrich, N. Y. Hammerla, and D. Rueckert, "Attention U-Net: Learning Where to Look for the Pancreas," *Conf. on Medical Imaging with Deep Learning,* Amsterdam, Netherlands, May 2018, pp. 1-10.

[8] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE Trans. Pattern Anal Mach Intell*, vol. 40, no. 4, 2018, pp. 834-848.

[9] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *Int. Conf. for Learning Representations,* San Diego, USA, January 2017, pp. 1-15.

[10] M. D. Zeiler, "ADADELTA: An Adaptive Learning Rate Method," *J. of Computer Science*, vol. 12, no. 4, 2012, pp. 1-6.

[11] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *J. of Machine Learning Research*, vol. 12, no. 61, 2011, pp. 2121-2159.

[12] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning," *Int. Conf. on Machine Learning*, Atlanta, Georgia, USA, May 2013, pp. 1139-1147.

[13] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Conf. on CV and Pattern Recognition,* Boston, MA, USA, June 2015, pp. 3431-3440.

[14] A. Amyar, R. Modzelewski, H. Li, and S. Ruan, "Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia Classification and segmentation," *J. of Computers in Biology and*

Medicine, vol. 126, no. C, 2020, pp. 1-10.

[15] Y. Wu, S. Gao, J. Mei, R. G. Zhang and M. M. Cheng, "Jcs: An explainable covid-19 diagnosis system by joint classification and segmentation," *J. of IEEE Transactions on Image Processing*, vol. 30, no. 99, 2021, pp. 3113-3126.

[16] A. Mobiny, P. A. Cicalese, S. Zare, P. Yuan, M. Abavisani, C. C. Wu, J. Ahuja, P. M. Groot and H. V. Nguyen, "Radiologist-Level COVID-19 Detection Using CT Scans with Detail-Oriented Capsule Networks," *J. of Electrical Engineering and Systems Science,* vol. 20, no. 4, 2020, pp. 1-11.

[17] X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang, and P. Xie, "COVID-CT-Dataset: A CT Image Dataset about COVID-19," *J. of Computer Vision and Pattern Recognition*, vol. 20, no. 11, 2020, pp. 1-14.

[18] K. Li, W. Li, C. Pan, Y. Zhong, X. Liu, Y. Liao and S. Li, "CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19)," *J. of European radiology,* vol. 30, no. 8, 2020, pp. 4407-4416.

[19] E. S. Gedraite and M. Hadad, "Investigation on the Effect of a Gaussian Blur in Image Filtering and Segmentation," *Conf. of Int. Symposium ELMAR-2011,* Zadar, Croatia, September 2011, pp. 393-396.

[20] U. Erkan, D. N. Thanh, L. M. Hieu and S. Engínoğlu, "An Iterative Mean Filter for Image Denoising," *J. of IEEE Access*, vol. 7, no. 19, 2019, pp. 167847-167859.

[21] P. A. Lyakhov, A. R. Orazaev, N. I. Chervyakov and D. I. Kaplun, "A New Method for Adaptive Median Filtering of Images," *Conf. of Russian Young Researchers in Electrical and Electronic Engineering*, Saint Petersburg and Moscow, Russia, January 2019, pp. 1197-1201.

[22] P. D. Patil and A.D. Kumbhar, "Bilateral Filter for Image Denoising," *2015 Int. Conf. on Green Computing and Internet of Things,* Greater Noida, India, October 2015, pp. 299-302.

[23] M. Luong, H. Pham, and C. D. Manning, "Effective Approaches to Attention-based Neural Machine Translation," *Proc. of the 2015 Conf. on Empirical Methods in Natural Language Processing,* Lisbon, Portugal, September 2015, pp. 1412-1421.

[24] D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," *3rd Int. Conf. on Learning Representations,* San Diego, CA, USA, May 2015, pp. 1-15.

[25] L. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation," *J. of Computer Vision and Pattern Recognition*, vol. 19, no. 2, 2017, pp. 1-14.

[26] M. Heeswijk, C. A. Graaff, B. T. Regina and S. X. Rao, "Automated and Semiautomated Segmentation of Rectal Tumor Volumes on Diffusion-Weighted MRI Can It Replace Manual Volumetry," *Int. J. of Radiation Oncology* Biology* Physics*, vol. 94, no. 4, 2016, pp. 824-831.

저자 소개

수징제(Jingjie Su)

2019년 Beijing Institute of Petrochemical Technology, Communication Engineering 졸업(공학사)

2020년 ~현재 전남대학교 대학원 컴퓨터공학과 재학(공학석사)

※ 관심분야 : Image Recognition, IoT

김강철(Kang-Chul Kim)

1981년 서강대학교 전자공학과 학사
1983년 서강대학교 전자공학과 석사
1996년 경상대학교 전자공학과 박사
현재 전남대학교 전기전자통신컴퓨터공학부 교수

※ 관심분야 : 임베디드시스템, NoC, IoT Pattern Recognition