



[논문 리뷰] MultiResUNet [2019]

☀ 본 논문은 SOTA를 달성한, **U-Net**을 개선한, **MultiResUNet**을 제안하였습니다.

Introduction

일반적인 Semantic Segmentation은 non-overlapping regions의 이미지들이 분할될 때, 비슷한 Pixel값을 가진 Pixel들이 함께 Clustering되어, 겹쳐지는 문제가 있습니다.

논문에서는 Network의 개선을 위해, potential scopes를 발견하여, Architecture를 개선하였습니다. MultiResUNet을 3D image에 적용하였을 때도, outperforms를 보였습니다.

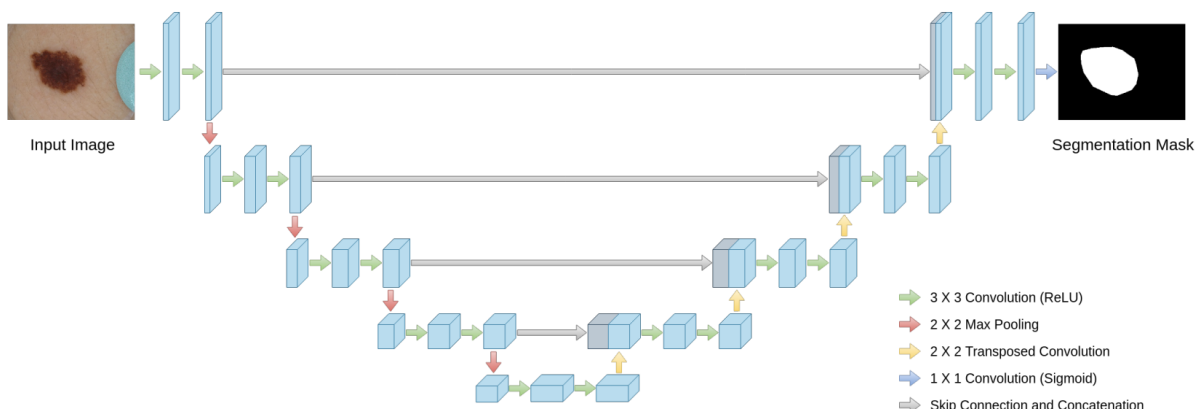
MultiResUNet은 일반적인 U-Net 모델이 비하여, 적은 parameter를 가지고, 좋은 성능을 나타내었습니다.

본 논문이 기여한 내용들

1. U-Net의 Architecture의 Depth를 분석한 후, enhancement할 수 있는 Potential opportunities와 병합하였습니다.
2. MultiResUNet은 standard U-Net을 enhanced한 Architecture입니다.
3. Medical image datasets에서, 다양한 modalities를 하고, MultiResUNet에 적용하였을 때, 상당한 performance를 나타내었음을 알 수 있었습니다. (Challenge Dataset 마찬가지.)
4. 3D U-Net과 비교하였을 때도, 3D MultiResUNet이 성능이 좋음을 알 수 있었습니다.

Overview of the UNet Architecture

☀ U-Net의 일반적인 구조에 관하여, 언급하고 있습니다.
자세한 내용은 U-Net 논문을 참고하는 것이 좋습니다.



- Semantic Segmentation에 주로 활용되는 Encoder-Decoder Structure에 Skip-Connection을 적용합니다.
- Pooling layer에 소실되는 Spatial info를, Residual Block을 활용해서, retrieve해줍니다.

Variation of Scale in Medical Images

Medical Image들은, 대부분 irregular하며, 불규칙한 크기를 가지고 있습니다.
 그러므로, Network들은 different scales에 관하여, 충분히 robust를 가지고 있어야 합니다.

여러가지 실험을 하였을 때, Inception을 U-Net Architecture에 적용하였을 때, Image내에서 학습한 것들을 서로 다른 위치에서, 정보를 가져와 활용할 수 있습니다.

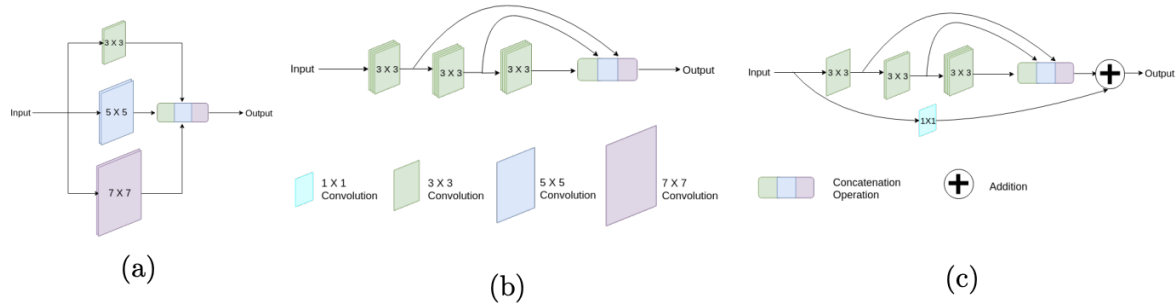


Figure 3: Developing the Proposed *MultiRes* block. We start with a simple Inception like block by using 3×3 , 5×5 and 7×7 convolutional filters in parallel and concatenating the generated feature maps (Fig. 3a). This allows us to reconcile spatial features from different context size. Instead of using the 3×3 , 5×5 and 7×7 filters in parallel, we factorize the bigger and more expensive 5×5 and 7×7 filters as a succession of 3×3 filters (Fig. 3b). Fig 3c illustrates the *MultiRes* block, where we have increased the number of filters in the successive three layers gradually and added a residual connection (along with 1×1 filter for conserving dimensions).

Inception Network의 접근법을 활용하여, U-Net 내에, 3×3 , 5×5 , 7×7 를 활용하여, multi-resolutional 분석을 활용하는 것이다. (a)

그렇게 됨으로서, Different Scale의 Image로부터 learn한 feature들을 reconcile하는 것이 가능합니다.

이렇게되면,

Memory를 상당히 많이 소모하므로, (b)와 같이 layer들을, 3×3 smaller and lightweight로 변경합니다.

Probable Semantic Gap between the Corresponding Levels of Encoder-Decoder

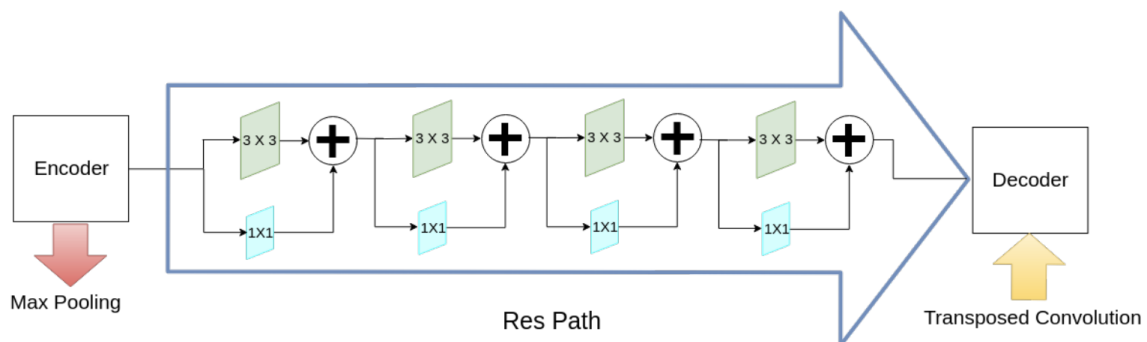


Figure 4: Proposed *Res* path. Instead of combining the the encoder feature maps with the decoder feature in a straight-forward manner, we pass the encoder features through a sequence of convolutional layers. These additional non-linear operations are expected to reduce the semantic gap between encoder and decoder features. Furthermore, residual connections are also introduced as they make the learning easier and are very useful in deep convolutional networks.

U-Net에서의 Encoder의 처음 layer정보는 decoder의 마지막 layer로 전달됩니다.

이러한 과정에서 발생할 수 있는 **Spatial information loss**를 줄이기 위해, 추가적인 **non-linear transformation**이 필요하며, **Residual Connection**이 포함된 **Conv Layer**를 추가하여, 학습을 할 수 있습니다.

Proposed Architecture

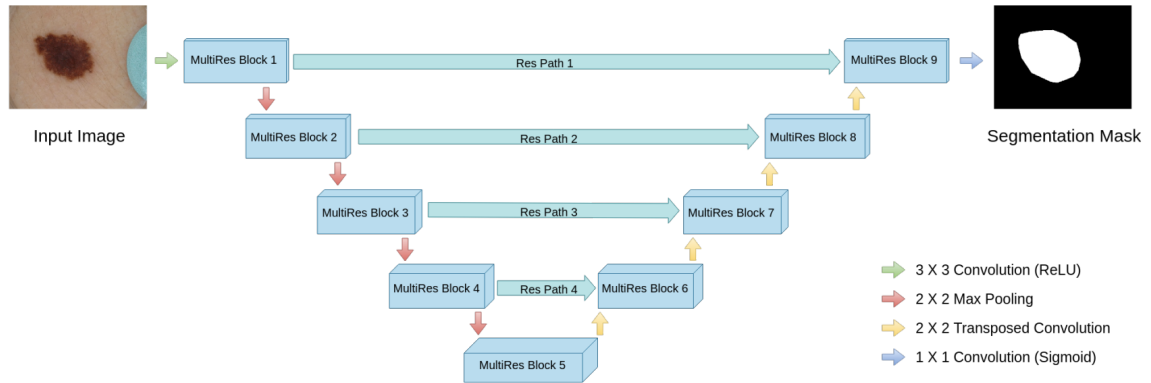


Figure 5: Proposed MultiResUNet architecture. We replace the sequences of two convolutional layers in the U-Net architectures with the proposed *MultiRes* block. Furthermore, instead of using plain shortcut connections, we use the proposed *Res* paths.

각 **Block** 내에 **Convolutional layer filter** 수를 기존 **U-Net**과 유사하게, **Control**하기 위해서, **W parameter**를 도입하였다.

또한, 논문을 보면, **U-Net**의 **parameter**보다, 적은 개수로, 좋은 **performance**를 나타내는 것을 알 수 있습니다.

Experiments

Modality	MultiResUNet (%)	U-Net (%)	Relative Improvement (%)
Dermoscopy	80.2988 \pm 0.3717	76.4277 \pm 4.5183	5.065
Endoscopy	82.0574 \pm 1.5953	74.4984 \pm 1.4704	10.1465
Fluorescence Microscopy	91.6537 \pm 0.9563	89.3027 \pm 2.1950	2.6326
Electron Microscopy	87.9477 \pm 0.7741	87.4092 \pm 0.7071	0.6161
MRI	78.1936 \pm 0.7868	77.1061 \pm 0.7768	1.4104

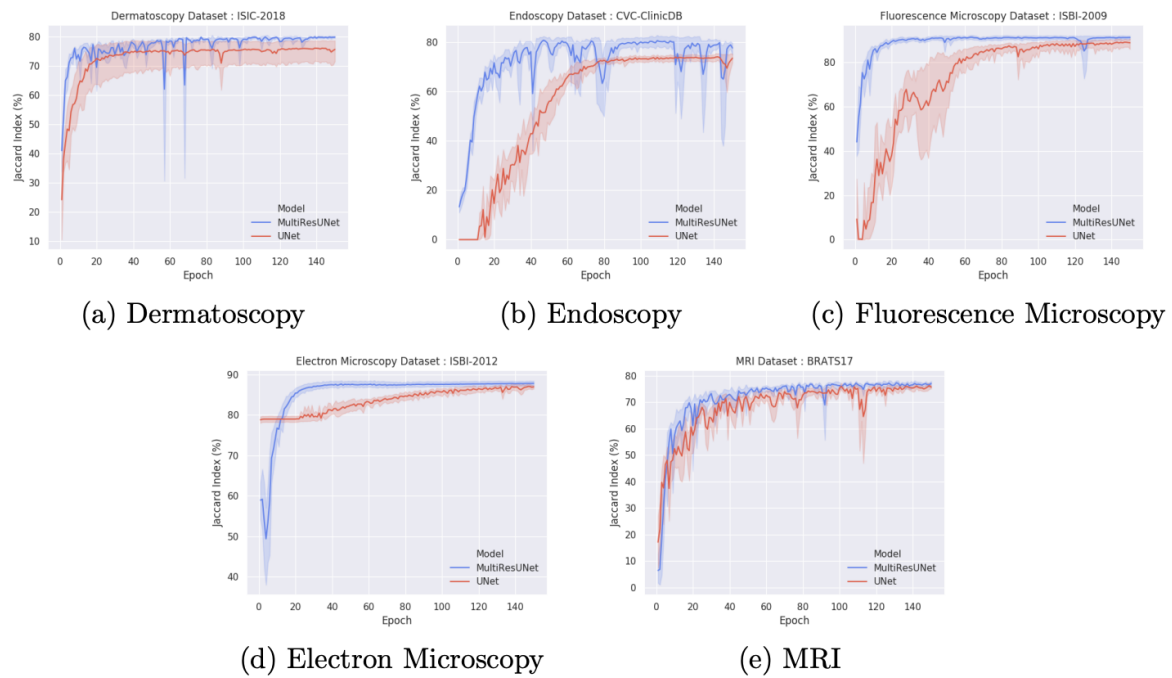


Figure 6: Progress of the validation performance with the number of epochs. We record the value of Jaccard Index on validation data after each epoch. It can be observed that not only MultiResUNet outperforms the U-Net model, but also the standard deviation of MultiResUNet is much smaller.

Conclusion

본 논문에서는 U-Net의 Architecture를 분석하여, 개선하였습니다. U-Net의 Encoder와 Decoder 내에는 training하는 과정에서 discrepancy가 발생하지만, Res-Block을 도입하여, reconcile되게끔 조정하였다.

또한, Multi-Resolution Analysis를 보강하기 위해 U-Net에 Res-Block을 도입하였습니다.

Inception Block으로부터 영감을 얻고, lightweight하고, demand less memory를 공식화하였습니다.

이러한 과정을 통하여, MultiResUNet을 고안하였습니다.

MultiResUNet : Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation


In recent years Deep Learning has brought about a breakthrough in Medical Image Segmentation. U-Net is the most prominent deep network in this regard, which has been the most popular architecture in the medical imaging community.

 <https://arxiv.org/abs/1902.04049>

arXiv

MultiResUNet - arxiv 2019

Information Title: MultiResUNet : Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation, arxiv 2019 Reference Review By: Taeyup Song Edited by: Taeyup Song Last updated on Jan. 16, 2022 Resource medical imaging 분야에서 잘 쓰이던 U-Net을 최신 architecture를 적용해서 개선 Multi-resolution analysis (figure a) → inception-like block을 도입 → stride conv.를 적용할 수 있으나 U-Net archi에 의해 효과가 없었음.

 https://pseudo-lab.github.io/SegCrew-Book/docs/ch1/01_02_01_multiresunet.html