

U-Net

Convolutional Networks for Biomedical Image Segmentation

INDEX

전처리

1. Data Exploratory Analysis
2. Image Resize
3. Augmentation Method



모델

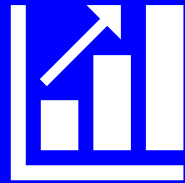
1. Baseline
2. Validation
3. Optimizer
4. Scheduler
5. Metric (Custom loss)
6. Apex

후처리

1. Pesudo labeling
2. TTA
3. Ensemble



전처리



1. Data Exploratory Analysis
 - 1.1. 이미지를 살펴보는 이유
2. Image Resize
3. Augmentation Method
 - 3.1. Augmentation을 하는 이유
 - 3.3. Albumentations 패키지를 이용한 데이터 증강 방법
 - 3.4. 최신 Augmentation 기법
 - 3.5. 그 외

0. Abstract

• (1) Abstract

✓ Unet[1]의 개요

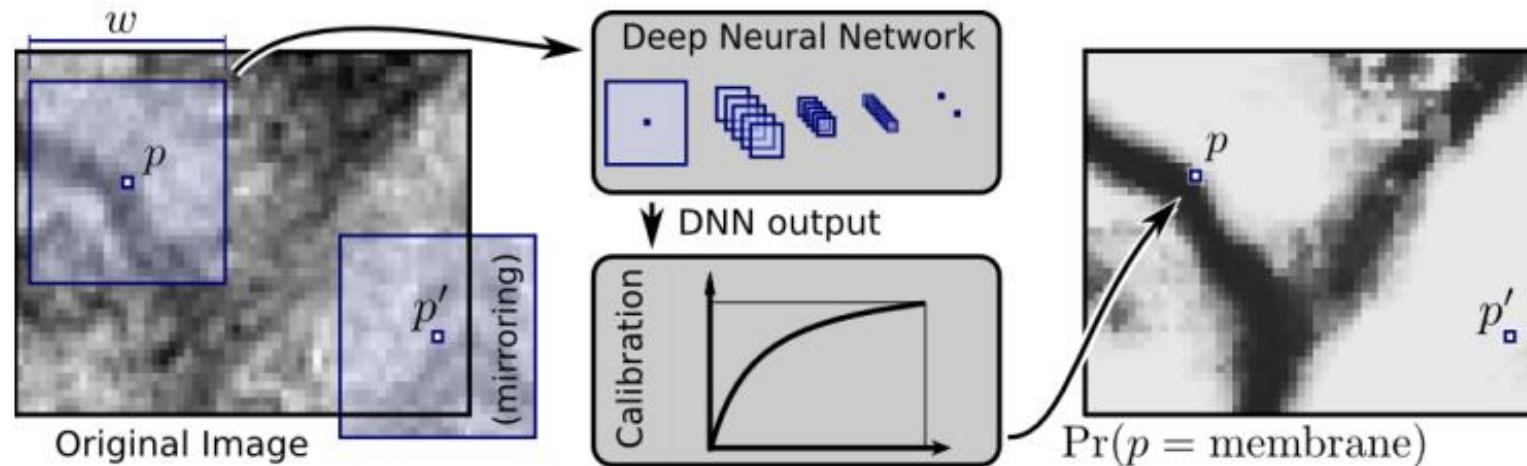
1. 딥러닝 네트워크의 경우 라벨을 가진 많은 학습 데이터가 필요
 - Unet에서는 네트워크 / 학습 전략 소개? 학습데이터를 효과적으로 사용하도록 하는 augmentation
2. 네트워크의 구성은 context를 잡는 contracting path와 localization을 가능하게하는 symmetric expanding path으로 구성
3. Unet을 통해 적은 이미지로 end-to-end 학습을 가능하게하고 sota 성능을 ISBI 대회에서 달성
4. 게다가 모델은 빠르다.

1. Introduction

• (0) Overview

✓ Biomedical 분야에서의 Image Segmentation

1. 기본적인 딥러닝은 파라미터가 많고 네트워크가 깊어서 학습데이터가 많이 필요하고 보통의 ImageNet dataset을 이용해서 학습한 모델은 1개의 Output을 필요로하는 Classification 테스트임
2. Biomedical 분야에서는 데이터의 수가 적고 localization 정보 뿐만아니라 모든 픽셀에 대해 class label을 부여해야함
 - Ciresan[2]의 경우 이를 해결하기 위해서 Sliding Window 방식을 통해 학습을 진행
 - Localize가 가능함
 - patch 단위의 학습을 할 경우 데이터의 양이 많아지는 효과가 있음
 - EM segmentation challenge에서 좋은 결과를 가져옴

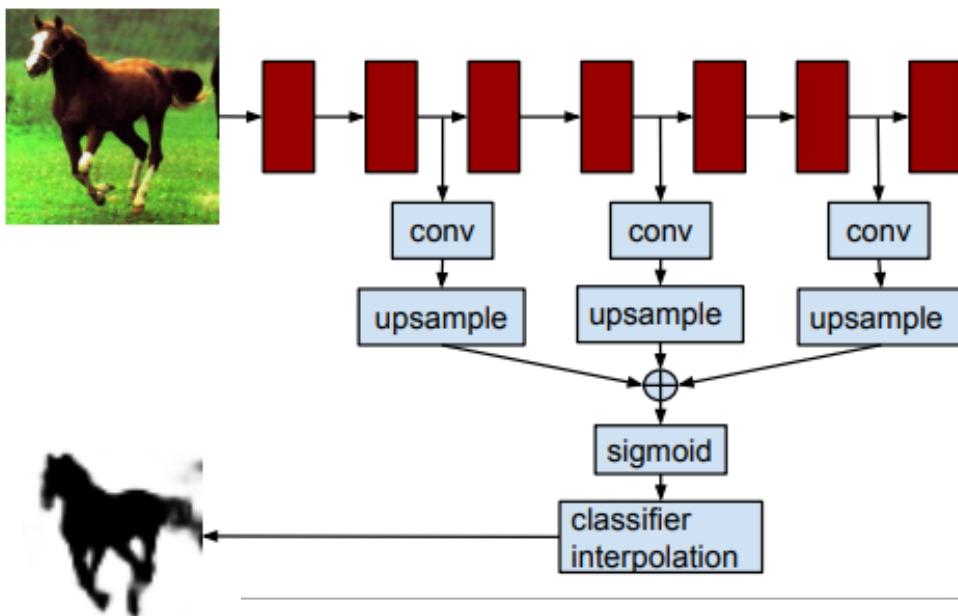


1. Introduction

• (0) Overview

✓ Biomedical 분야에서의 Image Segmentation

3. 하지만 Ciresan[2]의 Patch 단위의 학습은 2가지의 문제점이 있음
- 각각에 Patch에 대해 학습을 하기에 속도가 느리고 패치마다 많은 영역이 겹치는 부분이 생김
 - Localization 정확도와 context의 사용간에 trade-off가 존재
 - Patch가 크면 max-pooling을 많이 필요해서 localization accuracy가 떨어짐
 - Patch가 작으면 조그만 context만 보는 문제가 있음
 - 위의 문제를 해결하기위해서 [3]에서는 classifier의 output을 multiple layer의 features로 고려하는 방법이 제안되었고, Good Localization과 use of context가 가능해짐

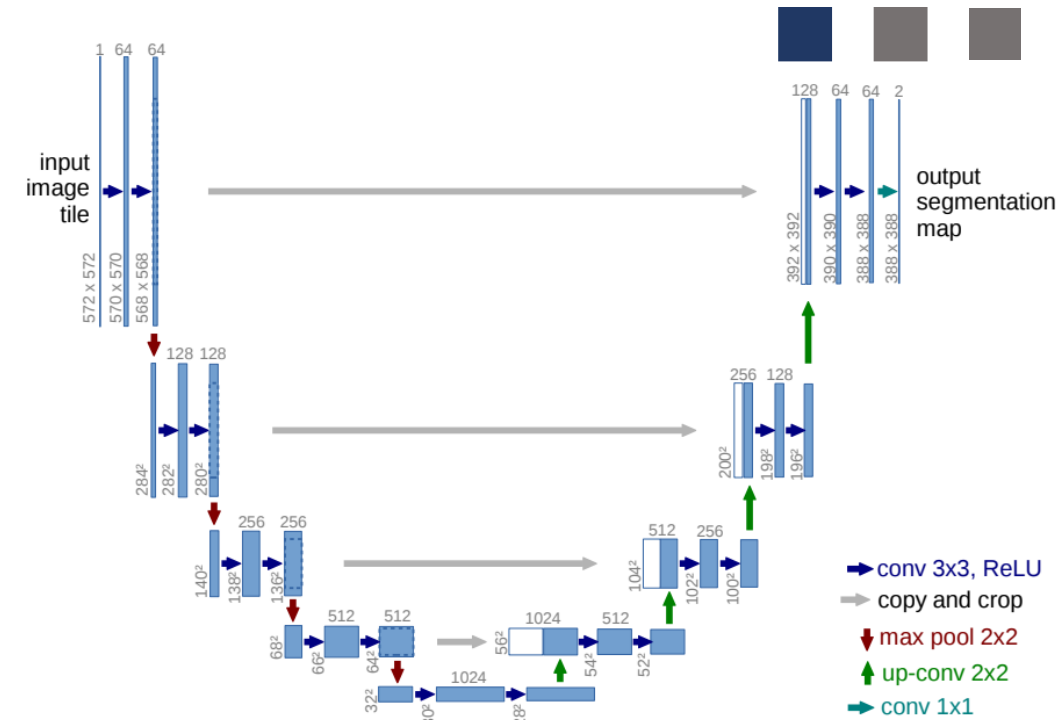


1 Introduction

• (0) Overview



Biomedical 분야에서의 Image Segmentation



4. Unet의 경우 FCN을 확장한 U 모양의 네트워크를 제안

- 적은 학습 데이터를 가지고도 높은 성능을 가짐
- [9]의 주요 아이디어는 풀링 운영자가 업샘플링 운영자로 대체되는 일반적인 계층 네트워크를 연속적인 계층으로 보완하는 것이다. 따라서, 이러한 계층들은 출력의 해상도를 증가시킨다. 로컬리제이션하려면 높은 계층 경로의 해상도 기능이 업샘플링된 출력과 결합됩니다. 그런 다음 연속 컨볼루션 레이어는 이 정보를 기반으로 더 정밀한 출력을 결합하는 방법을 배울 수 있다.
- Upsampling 채널의 수를 크게 함으로서 context information을 높은 resolution layer에 전파
- 결론적으로 expansive path는 contracting path와 대칭이고 U-shape의 모양을 가짐
- fully connected layers를 사용하지 않고 오직 convolution만 사용
 - overlap-tile 전략에 의해 만들어진 임의의 이미지를 입력으로 받아도 문제가 없게함
- 이미지의 테두리 영역에서 픽셀을 예측하기 위해 입력 이미지를 미러링하여 누락된 컨텍스트를 추정합니다.

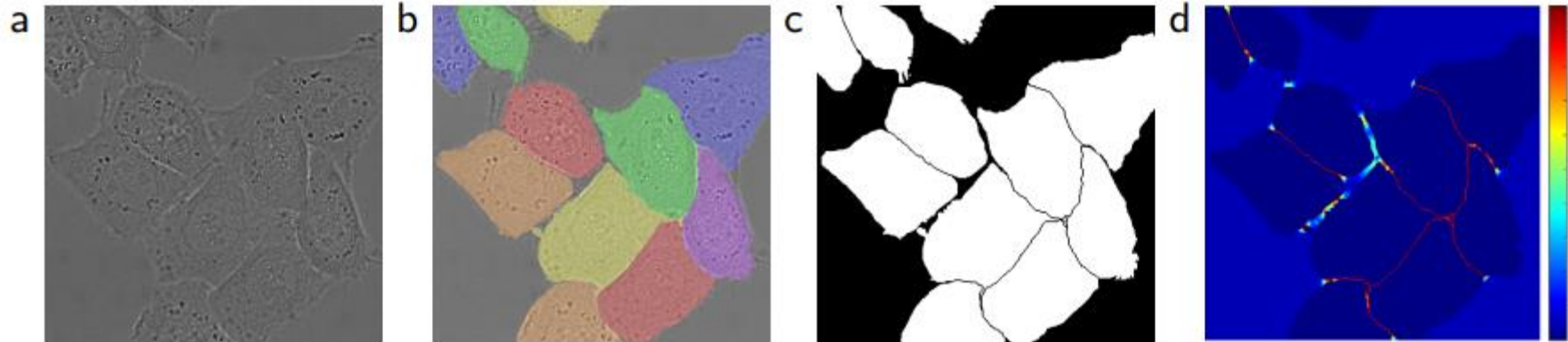
1. Introduction

- (0) Overview

- ✓ Biomedical 분야에서의 Image Segmentation

- 5. 학습 데이터의 부족 해결하기 위해 data augmentation을 사용

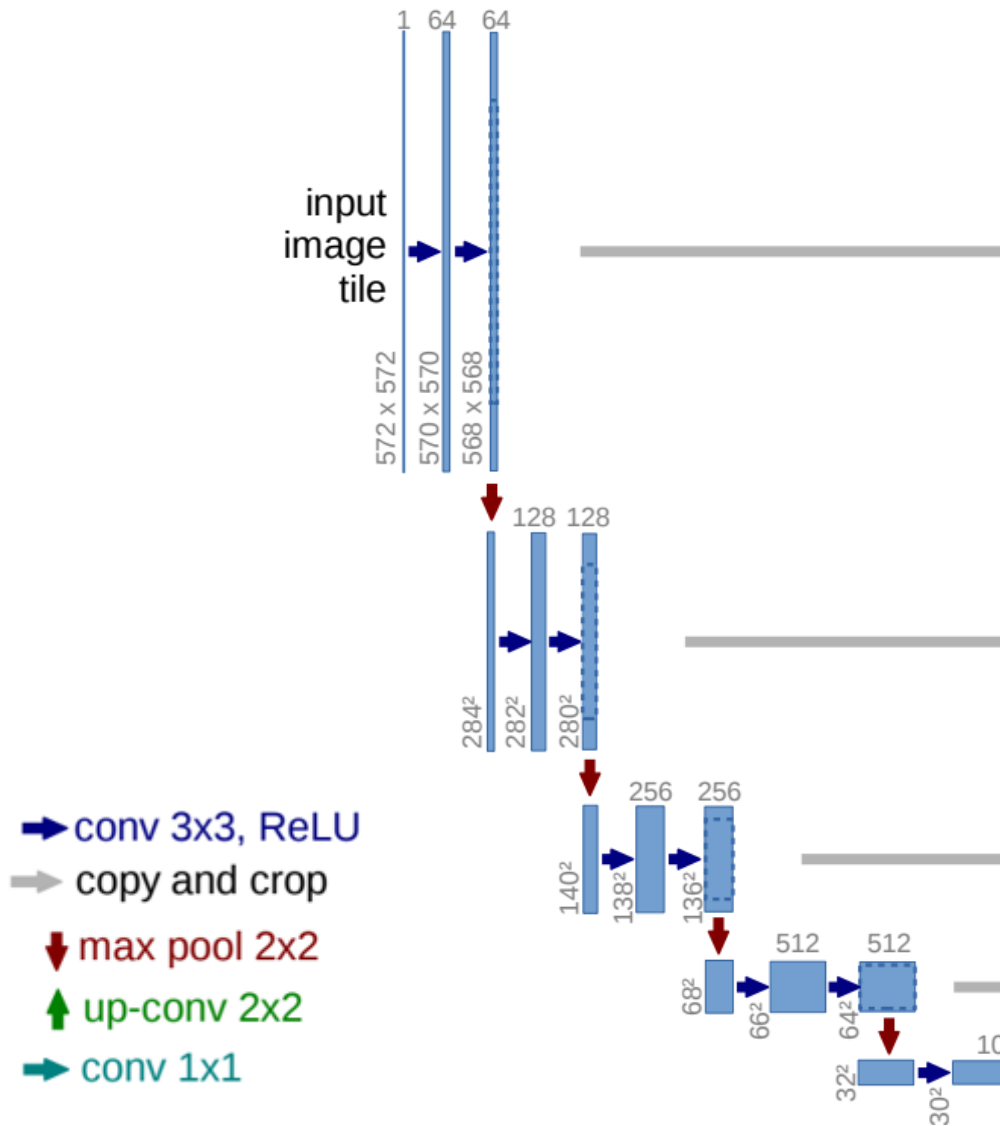
- 6. Cell segmentation에서 동일한 클래스의 겹치는 물체를 분리하는 작업



- 7. 결론적으로 2012에서 wg은 tjdsmb보임

2. Network Architecture

(1) Contracting Path



Contracting Path

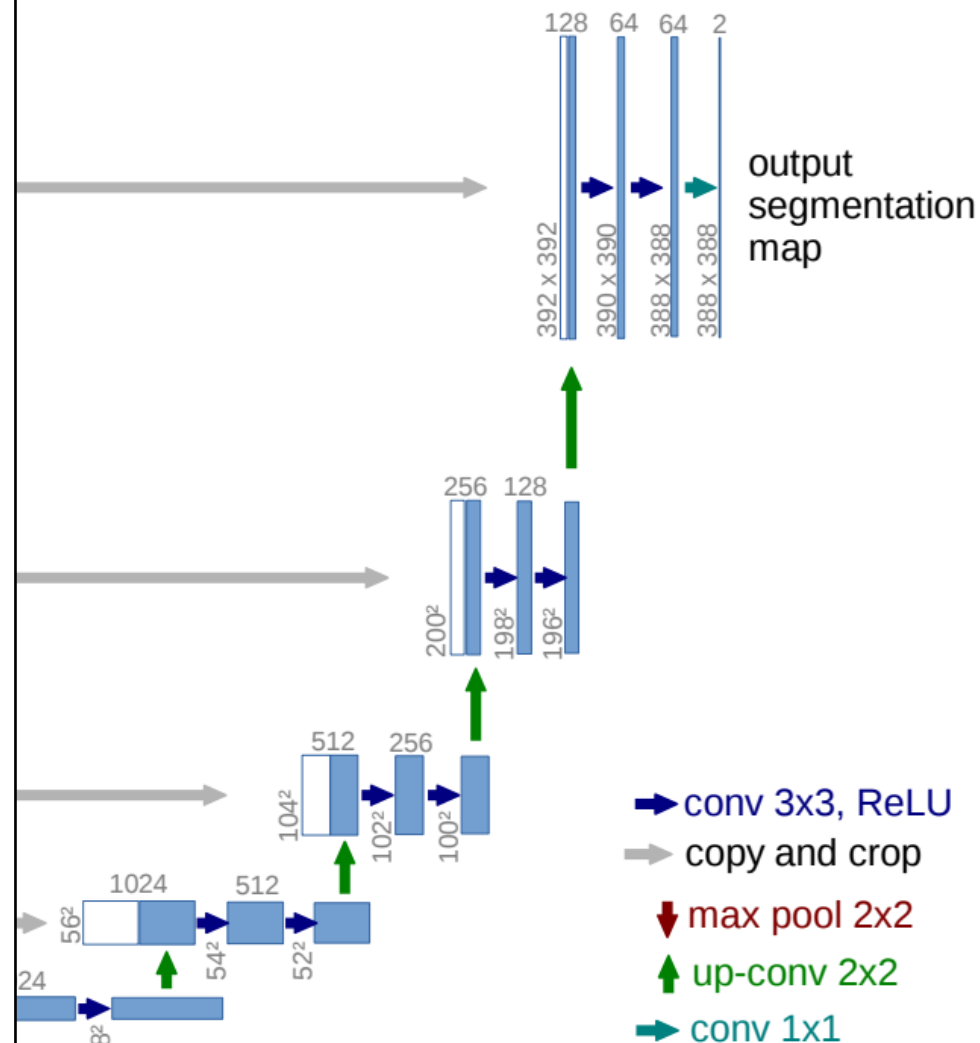
1. 기본적인 딥러닝은 파라미터가 많고 네트워크가 깊어서 학습데이터가 많이 필요하고 보통의 ImageNet dataset을 이용해서 학습한 모델은 1개의 Output을 필요로하는 Classfication 테스트임
2. Biomedical 분야에서는 데이터의 수가 적고 localization 정보 뿐만아니라 모든 픽셀에 대해 class label을 부여해야함

2. Network Architecture

(2) Expanding Path

Expanding Path

1. 기본적인 딥러닝은 파라미터가 많고 네트워크가 깊어서 학습데이터가 많이 필요하고 보통의 ImageNet dataset을 이용해서 학습한 모델은 1개의 Output을 필요로하는 Classification 테스트임
2. Biomedical 분야에서는 데이터의 수가 적고 localization 정보 뿐만아니라 모든 픽셀에 대해 class label을 부여해야함





✓ Biomedical 분야에서의 Image Segmentation

1. 기본적인 딥러닝은 파라미터가 많고 네트워크가 깊어서 학습데이터가 많이 필요하고 보통의 ImageNet dataset을 이용해서 학습한 모델은 1개의 Output을 필요로 하는 Classification 테스트임
2. Biomedical 분야에서는 데이터의 수가 적고 localization 정보 뿐만아니라 모든 픽셀에 대해 class label을 부여해야함
 - Ciresan[2]의 경우 이를 해결하기 위해서 Sliding Window 방식을 통해 학습을 진행
 - Localize가 가능함
 - patch 단위의 학습을 할 경우 데이터의 양이 많아지는 효과가 있음
 - EM segmentation challenge에서 좋은 결과를 가져옴

3. Training

(1) Data Augmentation

✓ Biomedical 분야에서의 Image Segmentation

1. 기본적인 딥러닝은 파라미터가 많고 네트워크가 깊어서 학습데이터가 많이 필요하고 보통의 ImageNet dataset을 이용해서 학습한 모델은 1개의 Output을 필요로하는 Classfication 테스트임
2. Biomedical 분야에서는 데이터의 수가 적고 localization 정보 뿐만아니라 모든 픽셀에 대해 class label을 부여해야함
 - Ciresan[2]의 경우 이를 해결하기 위해서 Sliding Window 방식을 통해 학습을 진행
 - Localize가 가능함
 - patch 단위의 학습을 할 경우 데이터의 양이 많아지는 효과가 있음
 - EM segmentation challenge에서 좋은 결과를 가져옴

4. Experiments

• (0) Overview



✓ Biomedical 분야에서의 Image Segmentation

1. 기본적인 딥러닝은 파라미터가 많고 네트워크가 깊어서 학습데이터가 많이 필요하고 보통의 ImageNet dataset을 이용해서 학습한 모델은 1개의 Output을 필요로하는 Classfication 테스트임
2. Biomedical 분야에서는 데이터의 수가 적고 localization 정보 뿐만아니라 모든 픽셀에 대해 class label을 부여해야함
 - Ciresan[2]의 경우 이를 해결하기 위해서 Sliding Window 방식을 통해 학습을 진행
 - Localize가 가능함
 - patch 단위의 학습을 할 경우 데이터의 양이 많아지는 효과가 있음
 - EM segmentation challenge에서 좋은 결과를 가져옴

5. Conclusion

(0) Overview



✓ Biomedical 분야에서의 Image Segmentation

1. 기본적인 딥러닝은 파라미터가 많고 네트워크가 깊어서 학습데이터가 많이 필요하고 보통의 ImageNet dataset을 이용해서 학습한 모델은 1개의 Output을 필요로 하는 Classification 테스트임
2. Biomedical 분야에서는 데이터의 수가 적고 localization 정보 뿐만아니라 모든 픽셀에 대해 class label을 부여해야함
 - Ciresan[2]의 경우 이를 해결하기 위해서 Sliding Window 방식을 통해 학습을 진행
 - Localize가 가능함
 - patch 단위의 학습을 할 경우 데이터의 양이 많아지는 효과가 있음
 - EM segmentation challenge에서 좋은 결과를 가져옴

5. Conclusion

(1) Advantages

✓ Biomedical 분야에서의 Image Segmentation

1. 기본적인 딥러닝은 파라미터가 많고 네트워크가 깊어서 학습데이터가 많이 필요하고 보통의 ImageNet dataset을 이용해서 학습한 모델은 1개의 Output을 필요로 하는 Classification 테스트임
2. Biomedical 분야에서는 데이터의 수가 적고 localization 정보 뿐만아니라 모든 픽셀에 대해 class label을 부여해야함
 - Ciresan[2]의 경우 이를 해결하기 위해서 Sliding Window 방식을 통해 학습을 진행
 - Localize가 가능함
 - patch 단위의 학습을 할 경우 데이터의 양이 많아지는 효과가 있음
 - EM segmentation challenge에서 좋은 결과를 가져옴

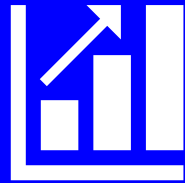
5. Conclusion

(2) Disadvantages

✓ Biomedical 분야에서의 Image Segmentation

1. 기본적인 딥러닝은 파라미터가 많고 네트워크가 깊어서 학습데이터가 많이 필요하고 보통의 ImageNet dataset을 이용해서 학습한 모델은 1개의 Output을 필요로하는 Classficiation 테스트임
2. Biomedical 분야에서는 데이터의 수가 적고 localization 정보 뿐만아니라 모든 픽셀에 대해 class label을 부여해야함
 - Ciresan[2]의 경우 이를 해결하기 위해서 Sliding Window 방식을 통해 학습을 진행
 - Localize가 가능함
 - patch 단위의 학습을 할 경우 데이터의 양이 많아지는 효과가 있음
 - EM segmentation challenge에서 좋은 결과를 가져옴

전처리



1. Data Exploratory Analysis
 - 1.1. 이미지를 살펴보는 이유
2. Image Resize
3. Augmentation Method
 - 3.1. Augmentation을 하는 이유
 - 3.3. Albumentations 패키지를 이용한 데이터 증강 방법
 - 3.4. 최신 Augmentation 기법
 - 3.5. 그 외

1. 공금증



• (2) PASCAL VOC 2012

1. Many of these architectures have used a host of supporting techniques and multi-stage training recipes to arrive at high accuracies on datasets but this makes it difficult to gather evidence about their true performance under time and memory constraints. Instead we chose to perform a controlled benchmarking where we used batch normalization to enable end-to-end training with the same solver (SGD). However, we note that this approach cannot entirely disentangle the effects of model versus solver (optimization) in achieving a particular result. This is mainly due to the fact that training these networks involves gradient back-propagation which is imperfect and the optimization is a non-convex problem in extremely large dimensions. Acknowledging these shortcomings, our hope is that this controlled analysis complements other benchmarks [62] and reveals the practical trade-offs involved in different well known architectures. (p12)
2. Note that the decoder corresponding to the first encoder (closest to the input image) produces a multi-channel feature map, although its encoder input has 3 channels (RGB). This is unlike the other decoders in the network which produce feature maps with the same number of size and channels as their encoder inputs. (p5)

2. 참고자료

(2) PASCAL VOC 2012

1. O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In MICCAI, pages 234–241. Springer, 2015
2. Ciresan, D.C., Gambardella, L.M., Giusti, A., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. In: NIPS. pp. 2852–2860 (2012)
3. Hariharan, B., Arbellez, P., Girshick, R., Malik, J.: Hypercolumns for object segmentation and fine-grained localization (2014), arXiv:1411.5752 [cs.CV]

감사합니다

2. 이해안가는 목록



1. Secondly, there is a trade-off between localization accuracy and the use of context. Larger patches require more max-pooling layers that reduce the localization accuracy, while small patches allow the network to see only little context
2. The main idea in [9] is to supplement a usual contracting network by successive layers, where pooling operators are replaced by upsampling operators. Hence, these layers increase the resolution of the output. In order to localize, high resolution features from the contracting path are combined with the upsampled output. A successive convolution layer can then learn to assemble a more precise output based on this information
3. To predict the pixels in the border region of the image, the missing context is extrapolated by mirroring the input image