

# Project Report

**Project Title:** U-Net: Convolutional Networks for Biomedical Image Segmentation

**Paper URL:** <https://arxiv.org/pdf/1505.04597.pdf>

**Code:** <https://github.com/S498/u-net>

## Problem Definition

In the field of computer vision, the typical use of convolutional networks is on classification tasks, where the output to an image is a single class label. However, in many visual tasks, especially in biomedical image processing, the desired output should include localization, i.e., a class label is supposed to be assigned to each pixel. Moreover, thousands of training images are usually beyond reach in biomedical tasks. Though convolutional networks existed, its success was limited due to the lack of sufficient training data. To get accurate image segmentation, and localization of different classes of the image, a large amount of image data is required [1], but in the biomedical field most of the image data present is private and not accessible for training purposes.

In digital image processing and computer vision, image segmentation is the process of partitioning a digital image into multiple image segments, also known as image regions or image objects (sets of pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

So, there is an ultimate need of a model which could be used to get accurate localization of different classes of the image, and could provide accurate image segmentation, when the amount of data present was limited, and could use it efficiently. With the improved models, that could provide accurate segmentation of different classes of the image compartments would help in development of better models to predict rare diseases in the biomedical field where the amount of the images present is very less [2].

## Project Objectives

The main objective of this paper was to present a network and training strategy that relies on the strong use of data augmentation to use the available annotated samples more efficiently. The main idea is to supplement a usual contracting network by successive layers, where pooling operations are replaced by upsampling operators. Hence these layers increase the resolution of the output. A successive convolutional layer can then learn to assemble a precise output based on this information.

The system mainly focuses on the following objective:

- Collect the image data, in this experiment the original data has been taken from the ISBI Challenge.
- Preprocess the image data and perform Data Augmentation, as the data consists of less number of images, we performed data augmentation using the ImageDataGenerator module from Keras.
- Use the pretrained model, as mentioned in the research paper that is built using the U-Net architecture for the Image Segmentation Challenge, to classify different cell compartments.

So, with the model trained using U-Net Architecture, we will be able to predict accurate results, even though there are very few training samples, with low error. This will be helpful for scenarios like biomedical image segmentation and classification of tumours which are rare, and when there is no prior research done in that field.

## **Multidisciplinary Analysis**

### **Data Analysis:**

- Total Images: 30

### **Image Analysis:**

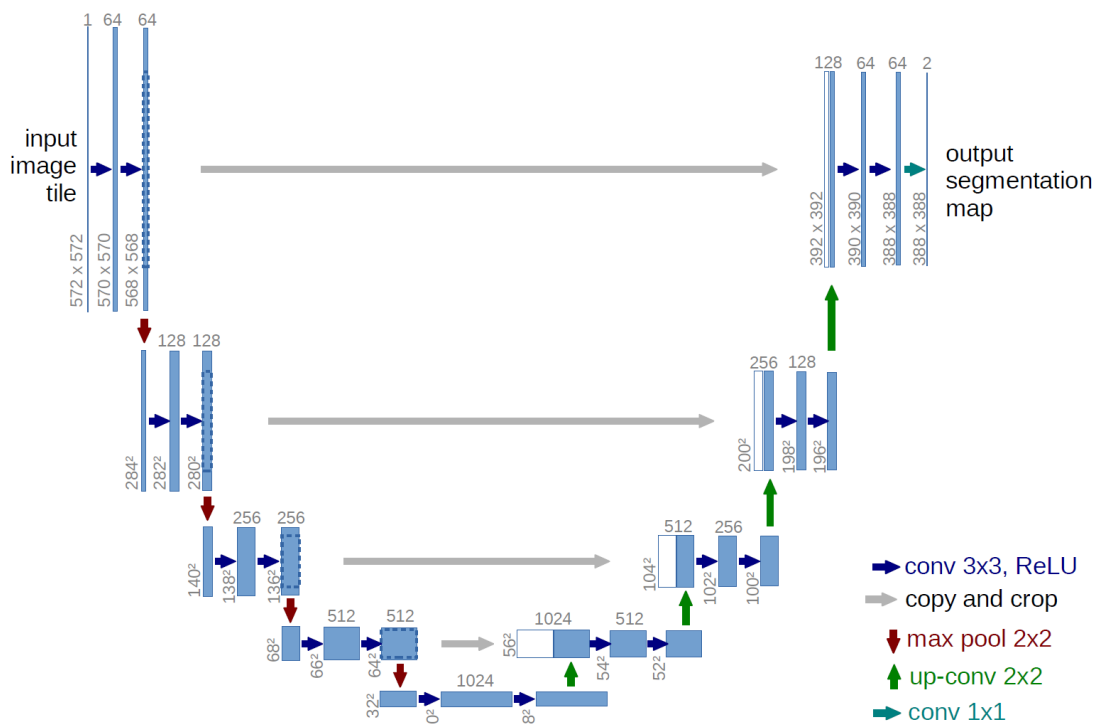
- Remove any personal information related to the patients and preprocess the data.
- Image prediction rate per image: 1s

### **U-Net Model Analysis:**

- One important modification in U-Net is that there are a large number of feature channels in the upsampling part, which allow the network to propagate context information to higher resolution layers. As a consequence, the expansive path is more or less symmetric to the contracting part, and yields a u-shaped architecture. The network only uses the valid part of each convolution without any fully connected layers. To predict the pixels in the border region of the image, the missing context is extrapolated by mirroring the input image. This tiling strategy is important to apply the network to large images, since otherwise the resolution would be limited by the GPU memory.
- The u-net architecture achieves very good performance on very different biomedical segmentation applications. Thanks to data augmentation with elastic deformations, it needs very few annotated images and has a very reasonable training time of only 10 hours on a NVidia Titan GPU (6 GB).
- U-Net is a convolutional neural network which takes as input an image and outputs a label for each pixel. U-Net initially was developed to detect cell boundaries in biomedical images. U-Net follows classical autoencoder architecture, as such it contains two sub-structures. The encoder structure follows the traditional stack of convolutional and max pooling layers to increase the receptive field as it goes through the layers. It is used to capture the context in the image. The decoder structure utilizes transposed convolution layers for upsampling so that the end dimensions are close to that of the input image. Skip connections are placed between convolution

and transposed convolution layers of the same shape in order to preserve details that would have been lost otherwise.

- The U-Net architecture consists of 2 paths:
  - **Contraction Path (Downsampling):** Look like a typical CNN architecture, by consecutive two 3x3 convolutions (blue arrow) followed by 2x2 Max Pooling for downsampling. At each downsampling step, the number of steps is doubled.
  - **Expansion Path (Up Convolution):** A 2x2 up-convolution (green arrow) for upsampling and two 3x3 convolutions (blue arrow). At each upsampling step, the number of channels is halved.
- After each 2x2 up-convolution, a concatenation of feature maps with correspondingly layer from the contracting path (grey arrows), to provide localization information from contraction path to expansion path, due to the loss of border pixels in every convolution.

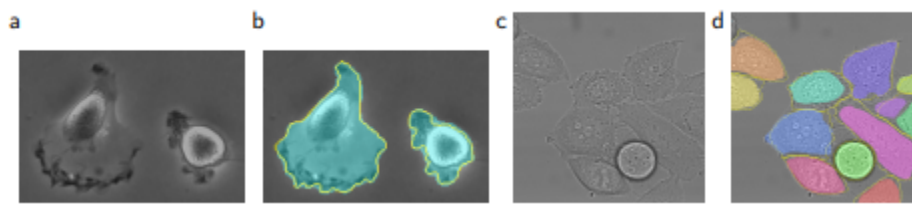


## Results

The results of the U-Net architecture is better than the previously better performed sliding-window technique using convolutional network. When compared with the Electron Microscope Segmentation challenge, U-Net architecture had the least warping error when compared with the other techniques.

Rank	Group name	Warping Error	Rand Error	Pixel Error
	<b>** human values **</b>	0.000005	0.0021	0.0010
1.	u-net	<b>0.000353</b>	0.0382	0.0611
2.	DIVE-SCI	0.000355	0.0305	0.0584
3.	IDSIA [1]	0.000420	0.0504	0.0613
4.	DIVE	0.000430	0.0545	<b>0.0582</b>
⋮				
10.	IDSIA-SCI	0.000653	<b>0.0189</b>	0.1027

Apart from the EM segmentation challenge, the author's also participated in the ISBI Cell Tracking challenge, where This architecture even did well for segmenting cells that had strong shape variations, weak outer borders, similar structures achieving an Intersection over Union(IoU) of 92%, and the evaluation metric used to measure the accuracy of an object detector on a particular dataset, which was much higher than the second-best method's IoU of 83%.



Name	PhC-U373	DIC-HeLa
IMCB-SG (2014)	0.2669	0.2935
KTH-SE (2014)	0.7953	0.4607
HOUS-US (2014)	0.5323	-
second-best 2015	0.83	0.46
u-net (2015)	<b>0.9203</b>	<b>0.7756</b>

## Discussion

The Image Segmentation accuracy generated by the Model Created using U-Net architecture is very good compared to the sliding window technique that was previously used. This model achieved very low warping error compared to its competitive models.

As computer vision and image segmentation are growing very fast, many recent technologies are becoming obsolete very fast — the hottest field for image segmentation application in biomedical imaging. As much as this technology can be helpful in the early detection and mitigation of diseases, it has high stakes as well. Biomedical applications are susceptible fields that demand rigorously tested, efficient, safe, and accurate models for implementation. When these factors are considered, they can immensely improve the quality of human life.

## Evaluation and Reflection

There is some incompatibility associated with the propagation of features from the contraction phase of the network to the expansion phase. Also, U-Net has discrepancies in handling multi-modal classes and multiresolution analysis. Architectures like MultiResUNet[3] have been developed by improvising the U-Net, which tackles the incompatibility issues of feature maps and aids in multi-resolution analysis. There has also been a re-invention on the skip connections of the old U-Net architecture into a new improvised architecture, UNet++. The semantic gap is reduced very much, resulting in faster and much more efficient learning.

Apart from the above mentioned results, the U-Net architecture has been used for the Drishti-GS Dataset, which is different from what Ronnerberger have used in their paper. This dataset contained 101 retina images, and annotated mask of the optical image disc and optical cup, for detecting Glaucoma, one of the major cause of blindness in the world. In this experiment 50 images were used for training and 51 were used for validation [4].

Since 2015, many improvised architectures have been proposed and built on top of the U-Net to improvise it in its weak areas. There have been few publications in which the U-Net architecture has been extended for 3D images using semi-automated and fully-automated segmentation [5].

## References

- [1]:<https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/35179.pdf>
- [2]:<https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e#:~:text=Studying%20biological%20vision,way%20to%20go.>
- [3]:<https://doi.org/10.1016/j.neunet.2019.08.025>
- [4]:<https://towardsdatascience.com/biomedical-image-segmentation-u-net-a787741837fa>
- [5]:<https://arxiv.org/abs/1606.06650>