ISTANBUL TECHNICAL UNIVERSITY ★ INFORMATICS INSTITUTE

DEEP LEARNING BASED SEGMENTATION PIPELINE FOR LABEL-FREE PHASE-CONTRAST MICROSCOPY IMAGES

M.Sc. THESIS

Aydin AYANZADEH

Department of Applied Informatics

Applied Informatics Programme

JULY 2020

ISTANBUL TECHNICAL UNIVERSITY ★ INFORMATICS INSTITUTE

DEEP LEARNING BASED SEGMENTATION PIPELINE FOR LABEL-FREE PHASE-CONTRAST MICROSCOPY IMAGES

M.Sc. THESIS

Aydin AYANZADEH
(708181005)

Department of Applied Informatics

Applied Informatics Programme

Thesis Advisor: Assoc. Prof. Dr. Behçet Uğur TÖREYİN
Co-advisor: Assoc. Prof. Dr. Devrim ÜNAY

JULY 2020

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ BİLİŞİM ENSTİTÜSÜ

ETİKETSİZ FAZ-KONTRAST MİKROSKOPİ GÖRÜNTÜLERİ İÇİN DERİN ÖĞRENME TABANLI SEGMENTASYON BORU HATTI

YÜKSEK LİSANS TEZİ

Aydin AYANZADEH
(708181005)

Bilişim Uygulamaları Anabilim Dalı

Bilişim Uygulamaları

Tez Danışmanı: Assoc. Prof. Dr. Behçet Uğur TÖREYİN
Eş Danışman: Assoc. Prof. Dr. Devrim ÜNAY

TEMMUZ 2020

Aydin AYANZADEH, a M.Sc. student of ITU Informatics Institute student ID 708181005 successfully defended the thesis entitled "DEEP LEARNING BASED SEGMENTATION PIPELINE FOR LABEL-FREE PHASE-CONTRAST MICROSCOPY IMAGES ", which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Assoc. Prof. Dr. Behçet Uğur TÖREYİN** Istanbul Technical University	.....
Co-advisor :	**Assoc. Prof. Dr. Devrim ÜNAY** İzmir Democracy University	.....
Jury Members :	**Prof. Dr. Hazım Kemal EKENEL** Istanbul Technical University	.....
	Assoc. Prof. Dr. Nazım Kemal ÜRE Istanbul Technical University	.....
	Prof. Dr. Devrim Pesen OKVUR İzmir Institute of Technology	.....

Date of Submission : **15 June 2020**
Date of Defense : **21 July 2020**

To my dear family,

FOREWORD

I do appreciate all the supervision which I received from my advisor, Assoc. Prof. Dr. Behçet Uğur TÖREYİN during the process of my Master's degree in general and thesis in particular. Moreover, I extend my best appreciation to my co-advisor, Assoc. Prof. Dr. Devrim ÜNAY for his advance help in the process of my master thesis. I also do appreciate all the support which I received from my family during my academic career.

Moreover, I should also add my appreciation of the motivational support which I received from Öznur DURUKAN during the thesis process.

This thesis is supported in part by Istanbul Technical University(ITU) Vodafone Future Lab under Project No.ITUVF20180901P04.
The data used in this study is collected under the Marie Curie IRG grant(no: FP7 PIRG08-GA-2010-27697).
This work is in part funded by ITU BAP MGA-2017-40964. This work is partially supported by the Tubitak ARDEB 1001 grant 119E578.

JULY 2020

Aydin AYANZADEH
(Student)

TABLE OF CONTENTS

Page

ABBREVIATIONS

App : Appendix
DCNN : Deep Convolutional Neural Network
ResNet : Residual Networks
ReLU : Rectified Linear Unit Parametric
FCN : Fully Convolutional Network
TP : True Positive
FP : False Positive
FN : False Negative
TN : True Negative
IoU : Intersection over Union
DC : Dice Coefficient
GAN : Generative Adversarial Network
FPN : Feature Pyramid Network
LSTM : Long Short Term Memory
ResNet : Residual Neural Networks
GT : Ground Truth
ResNet : Residual Neural Networks
DL : Dice Loss
ResNet : Residual Neural Networks
JI : Jaccard Index
RP : Residual Path

SYMBOLS

DL	**:** Dice Loss
BCE	**:** Binary Cross Entropy loss
L	**:** Loss function
p_i	**:** Probability map
g_i	**:** Ground truth binary volume
n_{fp}	**:** Number of false positive
n_{tp}	**:** Number of true positive
n_{fn}	**:** Number of false negative

LIST OF TABLES

Page

LIST OF FIGURES

Page

DEEP LEARNING BASED SEGMENTATION PIPELINE FOR LABEL-FREE PHASE-CONTRAST MICROSCOPY IMAGES

SUMMARY

Since the discovery of the microscope, analysis towards cell biology is one of the integral parts for the understanding of cell behaviour. These analyses of the cells biology have two main perspectives comprising of quantitative and qualitative factors. Cell segmentation is one of the critical tasks which make ease the process of cell analysis. Phase-contrast Microscopy(PCM) images have an integral role in a detailed analysis of the cell structures as it allows a wide range of possibilities in studying living cells from different perspectives in a label-free manner. Segmentation of the PCM images has been challenging for the biologists due to problems specific to the examined specimen or the microscopy technique used, such as perturbations in the cell shapes, overlapping and transparent appearance of the cells in PCM images which complicates the cell segmentation with conventional computer vision approaches, effective extraction of the Region of Interest (RoI) from images. Therefore, most of the segmentation and tracking solutions available in the literature do not qualify as satisfactory in terms of robustness, completeness, and accuracy.

The recently popular Deep Neural Networks(DNNs) have a significant effect on the improvement of segmentation accuracy from different perspectives, including robustness and completeness compared to conventional methods; moreover, deep learning-based approaches are also unbeaten in classification, detection and tracking applications. However, deep learning bases approaches require large numbers of annotated datasets to converge at high performances. By the way, applying the specific architectures which are well-known in the field of medical imaging segmentation can be successful even in low annotated images efficiently.

We proposed an alternative feature extractor by applying the modified ResNet18 in the encoder of U-Net and replacing the plain blocks with residual blocks in the decoder of U-Net architecture. As an extension of the ResNe18-U-Net, we employed residual-pathway as an alternative of the skip-connection of the model which made the model effective in taking low and high-level semantics and the detail of the images into account. Furthermore, we utilized ResNet18 in the backbone of Feature Pyramid Network(FPN) which has higher performance in comparison to conventional methods.

In our experimental results, we employ our proposed models in comparison to the base-line models on MDA-MB-231 and DSB2018 dataset. Our proposed model achieved Jaccard Index of 85.4% $\pm$ 2.7, 87.9% $\pm$ 1.7, 87.1% $\pm$ 2.3 and 89.2% $\pm$ 1.3 with TipNet, ResNet18-U-Net, ResNet18-FPN and ResNet18-U-Net+RP respectively on MDA-MB231 dataset and achieve the Jaccard index of 83.1% $\pm$ 1.8, 81.2% $\pm$ 2.2 and 85%$\pm$ 2.1 in the DSB2018 dataset which is slightly less than BCDUNet in the Jaccard Index metric. All of the determined results are come out after applying the

TTA as post processing at the proposed method on the DSB2018 and MDA-MB231 datasets. However, our proposed method passes the BCDUNet on MDA-MB-231 dataset. Moreover, the proposed approach is provided satisfactory segmentation results in terms of robustness and completeness on the MDA-MB231 dataset and DSB2018 dataset. Moreover, our proposed methods are also robust in the segmentation of the boundary regions of the cells in the presented dataset; from the other perspective, these methods are reliable on the detection of the outlier and it causes to reduce the amounts of false prediction in comparison of baseline methods. Among the proposed methods ResNet18-U-Net+RP is outperforming in the separation of the boundaries with faint characteristics.

ETİKETSİZ FAZ-KONTRAST MİKROSKOPİ GÖRÜNTÜLERİ İÇİN DERİN ÖĞRENME TABANLI SEGMENTASYON BORU HATTI

ÖZET

Mikroskop keşfinden bu yana, hücre biyolojisinin analizi, hücre davranışlarının analizinde ayrılmaz parçalardan biridir. Hücre kültürünün kantitatif ve kalitatif tespiti, hücre yapısı analizinin güçlü tespitinin ayrılmaz bir parçasıdır. Biyolojik araştırma süreci alanlarında mikroskopi hücre analizi ve hücre kültürlerindeki hücrelerin epitet yapıları çok önemlidir. Hücre biyolojisinin bu analizi, nicel ve nitel faktörlerden oluşan iki ana perspektife sahiptir. Hücre segmentasyonu, hücre analizi sürecini kolaylaştıran önemli görevlerden biridir.

Faz-kontrast Mikroskopi optik(PCM) görüntüleri, hücre yapılarının detaylı analizinde ayrılmaz bir role sahiptir. Çünkü PCM, canlı hücreleri farklı perspektiflerden etiketsiz bir şekilde incelemek için çok çeşitli olanaklar sunar. PCM görüntülerinin bölümlere ayrılması, incelenen numuneye özgü problemler veya kullanılan hücre mikroskoplarındaki bozulmalar, PCM görüntülerindeki hücrelerin üst üste binme ve şeffaf hücre görünümü ile karmaşıklaşan karmaşık görünümü gibi biyologlar için zorlayıcı olmuştur. Bilgisayarlı görme yaklaşımları, ilgi alanlarının görüntülerden etkili bir şekilde çıkarılması gerekmektedir Bu nedenle, literatürde mevcut olan segmentasyon ve izleme çözümlerinin çoğu sağlamlık, bütünlük ve doğruluk açısından tatmin edici değildir. Ayrıca canlı hücre mikroskopi görüntülemesinin uygun görüntü analiz araçlarıyla birlikte biyologlara biyolojik olayları ölçmede yardımcı olma potansiyeli vardır. Örneğin, hücrelerin segmentasyon görevleri için uygun araçlar olmadan, hızlandırılmış mikroskopide hücrelerin özelliklerinin incelenmesi biyologlar için zaman alıcı ve sıkıcı bir görev olabilir.

Son zamanlarda popüler olan Derin Sinir Ağları, geleneksel yöntemlere kıyasla sağlamlık ve bütünlük dahil olmak üzere farklı yönlerden segmentasyon doğruluğunun geliştirilmesi üzerinde önemli bir etkiye sahiptir; ayrıca derin öğrenme tabanlı yaklaşımlar sınıflandırma, tespit ve izleme uygulamalarında da başarılıdır. Bu arada, tıbbi görüntüleme alanında iyi bilinen spesifik mimarileri uygulayarak, düşük açıklamalı veri kümesine sahip veri kümesinin segmentasyonu düşük açıklamalı görüntülerde bile başarılı olabilir. Son zamanlarda popüler olan Derin Sinir Ağları, geleneksel yöntemlere kıyasla sağlamlık ve bütünlük dahil olmak üzere farklı yönlerden segmentasyon doğruluğunun geliştirilmesi üzerinde önemli bir etkiye sahiptir; ayrıca derin öğrenme tabanlı yaklaşımlar sınıflandırma, tespit ve izleme uygulamalarında da başarılıdır. Bununla birlikte, derin öğrenme tabanı yaklaşımı çok sayıda açıklamalı veri kümesine açlıktan ölmektedir. Bu arada, tıbbi görüntüleme alanında iyi bilinen spesifik mimarileri uygulayarak, düşük açıklamalı veri kümesine sahip veri kümesinin segmentasyonu düşük açıklamalı görüntülerde bile başarılı olabilir.

Bu tezin amacı, özellikle faz-kontrast mikroskopi görüntülerinde tıbbi görüntülemenin segmentasyonunda derin sinir ağının etkinliğini ve sağlamlığını araştırmaktır. Canlı

hücre mikroskopisi görüntüleme ile birlikte uygun imageanalysis araçları biyologların biyolojik fenomenleri ölçmede yardımcı olma potansiyeline sahiptir.

Veri kümesi hakkında, faz-kontrast mikroskopi görüntülerinde belirli bir veri kümesi oluşturduk. MDA-MB-231 mezenkimal morfolojisi olan ve bir Olympus IX71 mikroskobu kullanılarak yakalanan invaziv meme kanseri hücreleridir. Hücrenin canlılık, enfeksiyon ve proliferasyon gibi spesifik özelliklerinin incelenmesi için, hücreler ters ışık mikroskobu ile günlük olarak takip edilir. Optimize edilmiş sayılarla 6 oyuklu petri kaplarında hücreler görüntülendikten sonra, 24 saat boyunca yapışmaları beklenecektir. Bireysel hücre hareketlerini görselleştirmek için farklı yoğunluklardaki hücreler test edilecektir. Yara kapatma görüntüleri için, hücreler 100% dolgunluğa ulaşacak bir sayıya ekilecek ve yara, görüntüleme başlamadan hemen önce sarı uçlu bir çizgide yüzeyi kazıyarak oluşturulacaktır. Görüntüleme için 6-çukurlu petri kapları, 37°C sıcaklık ve Leica SP8 mikroskop sisteminin 5% C02'si sağlayacak şekilde inkübasyon odasına yerleştirilecek ve 48 veya daha fazla süre için belirlenen alanlardan faz kontrastlı görüntüler alınacaktır. Her 30 veya 60 dakikada bir 72 saat. Hücrelerin sürekliliğini analiz etmek için düzenli olarak stok çorabı uygulanır. Son olarak, hücreler FBS ve DMSO içeren spesifik bir çözelti içinde dondurulmuştur.

Bu tezde, PCM görüntülerinde hücre segmentasyonu için U-Net mimari yönteminden esinlenen ve geleneksel yaklaşımlara kıyasla üstün performansını gösteren hibrid derin sinir ağlarını önerdik. Kodlayıcı özelliğindeki eşitsizliği ve U-Net mimarisinin kod çözücüsünde yayılan özellikleri azaltmak için. ResNet18'i Net kodlayıcısına uygulayarak alternatif bir özellik çıkarıcı önerdik ve düz blokları U-Net mimarisinin kod çözücüsünde kalan bloklarla değiştirdik. Ayrıca, geleneksel yöntemlere göre daha yüksek performansa sahip FPN'nin omurgasında ResNet18'i kullandık. Ayrıca, U-Net benzeri mimariye uygulanan atlama bağlantısına alternatif olarak değiştirilmiş residual yolu önerdik. Bu değişiklik, görüntülerin detaylarını dikkate alarak etkili hale getirdi. Ayrıca sunulan yaklaşımda transfer öğreniminin kullanılması eğitim yakınsamasını artırmıştır. Tahmin adımındaki sonuçların esnekliğini geliştirdi, eğitim süresini azalttı ve modelin aşırı uymasını önlemeye de katkıda bulundu.

Makine öğrenimi uygulamaları tıbbi görüntü analizi alanı çoğunlukla probleme bağlıdır ve açıklamalı görüntülerin sayısı çok düşüktür. Özellikle tıbbi görüntü segmentasyonu alanında, büyük miktarda açıklamalı veri setinin hazırlanması, uzmanlar için sıkıcı ve zaman alıcı bir prosedürdür. Ayrımcı özelliklerin çıkarılmasında derin öğrenme uygulamasının gücünden esinlenerek, hücre segmentasyonu için U-Net [4] benzeri mimarilerden esinlenen otomatik kodlayıcı ağlarını varsayıyoruz. Araştırmamızda, sorularımız için MDA-MB-231 adlı veri kümesini kullandık. İlk olarak, ResNet18(Conv1-Conv5) 'in omurgası ile otomatik kodlayıcı yapılarına sahip mimariyi, ilk ImageNet ağırlığına sahip olan mimariyi önermiştik. Ayrıca, FPN'nin omurgasındaki ResNe18'in (Conv1-Conv5) yoğun katmanlarını hariç tutarak Özellik Piramidi Ağı(FPN) için bu işlemi tekrarlıyoruz. İkinci olarak, enkoder ve dekoder yolu arasındaki atlama bağlantısını yeniden tasarlıyoruz, bu da özel residual blokları uygulayarak enkoder ve dekoder arasındaki anlam boşluğunu azaltır. Atlama bağlantısının bu değişimi, modelin daha ayrımcı bilgi yakalamasını sağlar ve bu da kaldıraçlı veri kümesinde bölümleme sonuçlarında daha fazla sağlamlığa neden olur. Ayrıca, işlem sonrası olarak, mimarilerin belirsizlik tahmin düzeyini ölçmek için yararlı hileler olan Test Süresi Arttırımı(TTA) kullandık, fakat aynı zamanda kullanılan veri setimizde segmentasyon sonuçlarının kesinliğini arttırdı.

Önerilen yaklaşımımızın performansını EGT [8] ,BCDUNet [6], TipNet [9] ve PHANTAST [10] gibi çeşitli taban çizgileriyle karşılaştırıyoruz. Karşılaştırmamızda uygulanan U-Net'in her aşamasında (32, 64, 128, 256, 512) filtreler vardır. PHANTAST ve EGT gibi geleneksel yöntemler, hücre sınırı ve ROI'nin çıkarılmasında tatmin edici değildir. Önerilen yöntem, en gelişmiş yöntemlerden daha iyi performans gösterir ve hücre sınırlarının çıkarılmasında belirgin bir şekilde güçlü bir performansa sahiptir. Ayrıca, spesifik hücre yapısının öngörülmesinde de esnektir. ResNet18-U-Net, ResNet18-FPN ve ResNet18-Unet+Rp ile sırasıyla **87.9%**$\pm$ 1.7, **87.1%**$\pm$ 2.3 ve **89.2%**$\pm$ 1.3 Jaccard dizinlerini elde ettik. Önerilen yaklaşımımızın sadece segmentasyon doğruluğunu arttırmakla kalmayıp aynı zamanda üst üste binen hücrelerin herhangi bir son işlem prosedürü kullanılmadan ayrılmasında da etkili olduğunu kanıtlamaktadır. Ayrıca, önerdiğimiz yaklaşım, MDA-MB-231 ve DSB2018 veri kümesindeki segmentasyon görevindeki U-Net ve diğer temel çizgilerin eksiksizliğini, sağlamlığını ve diğer verimlilik faktörlerini aşmıştır.

DSB2018 veri seti ile ilgili önerilen yaklaşımımızı U-Net , SegNet gibi çeşitli taban çizgilerini analiz eder ve karşılaştırırız. Önerilen uygulama en gelişmiş yöntemlere ulaşmıştır ve hücre sınırlarının çıkarılmasında gözle görülür derecede sağlam bir performansa sahiptir. Ayrıca, önerilen yaklaşım özellikle ResNet18-U-Net + Residual path değerlere, hücre sınırlarına karşı güçlü bir performansa sahiptir. ResNet18-FPN, ResNet18-U-Net ve ResNet18-U-Net+RP ile sırasıyla **81.2%**$\pm$ 2.2, **83.1%**$\pm$ 1.8 ve **85%**$\pm$ 2.1 Jaccard dizinlerini elde ettik. Önerilen yaklaşımımızın sadece segmentasyon doğruluğunu geliştirmekle kalmayıp aynı zamanda bitişik veya üst üste binen bazı hücrelerde üst üste binen hücrelerin ayrılmasında da etkili olduğunu kanıtlamaktadır. DSB2018 veri kümesinin bazı çerçevelerinde yüksek bozulma nedeniyle. önerdiğimiz yaklaşım, DSB2018 veri kümesindeki segmentasyon görevindeki U-Net ve diğer temel çizgilerin eksiksizliğini, sağlamlığını ve diğer verimlilik faktörlerini aşmıştır.

Bu araştırmanın gelecekteki yönünün birkaç dalı vardır. Gelecek çalışmada, bölümleme görevleri gibi yöntemlerimizi konuşlandırma ve izleme aşaması için sağlam hale getirme planımız var. Bununla birlikte, gelecekte model için en iyi hiper parametre setini daha ayrıntılı olarak belirlemek için deneyler yapmak istiyoruz. Ayrıca, tıbbi yöntem veri kümelerinde farklı yöntemlere ilişkin önerilen yöntemlerin performansını değerlendirmek istiyoruz. Daha önce de belirtildiği gibi, zaman serileri için yöntemlerimizi hücre davranışı hakkında ayrıntılı bilgi veren soy ilişkileri kurarak uygulamak istiyoruz.

1. INTRODUCTION

Phase-contrast Optical Microscopy (PCM) images play an indispensable role in detailed analysis of cell structures because they offer a variety of possibilities to examine living cells from different perspectives in an unlabeled manner. Visual analysis of such time-series data is essential but challenging for biologists due to the time consuming and tiring nature of the process as well as factors such transparent and mixed light cell appearance. For this reason, most of the segmentation and monitoring solutions for PCM images which are available in the literature are not satisfactory in terms of robustness, integrity, and accuracy. Additionally, live cell microscopy imaging has the potential to assist biologists in measuring biological events, along with appropriate image analysis tools. For example, it may be a time consuming and tedious task for biologists to examine the properties without the appropriate tools for the segmentation of cells. Deep learning approaches has significant impact on the process of automated analysis of biomedical images and ease the process of automatic analysis which includes classification, detection, segmentation, tracking and quantitative analysis of the cells.

1.1 Purpose of Thesis

The purpose of this thesis is to investigate the effectiveness and robustness of deep neural network(DNN) in segmentation of biomedical images specifically phase-contrast optical microscopy time-series images. Live-cell microscopy imaging together with appropriate image analysis tools have the potential to aid biologists in quantifying biological phenomena. For instance, without proper tools for segmentation tasks of cells, scrutinizing the properties of the cells in time-lapse microscopy could be a time consuming and expensive task for biologists. For instance, to have efficient preparation of the for annotation purpose for segmentation tasks, you have to pass the determined workflow which is shown in Figure 1.1.

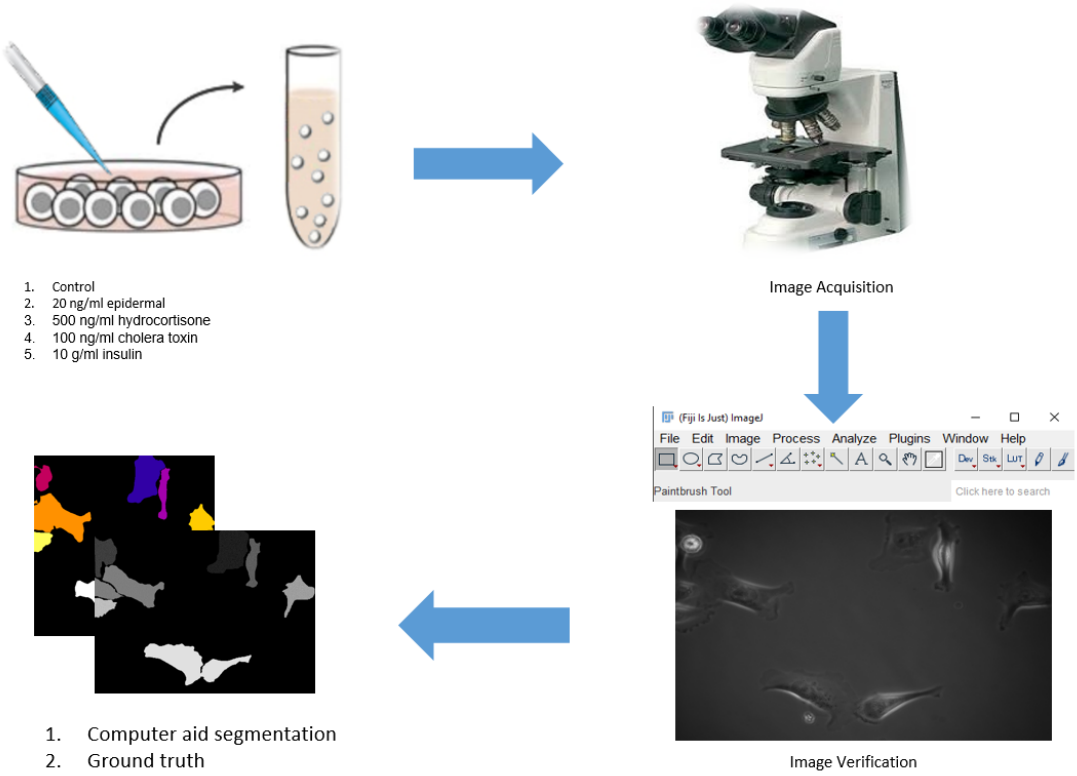

Figure 1.1 : Workflow of the preparation of the dataset for segmentation task.

1.2 Literature Review

We review the literature of segmentation of medical images in general, and particularly focus on the literature of cell segmentation. Finally, we present previous and current works which are related to the result of segmentation in this thesis.

Literature among the cell segmentation application is vast and rich. Conventional method is containing of image processing approach such as automatic segmentation approaches such as thresholding, active contours [11] and watershed algorithms [12]. In particular, morphological approach for detection of Phase-Contrast Microscopy images and it extended with Hierarchical Mergence which introduced in. [13].

Recently, Deep learning has significant effect on the performance of segmentation task drastically. In [4] a encoder-decoder segmentation architecture is introduced which is called U-Net due to its similarity to U like shape. U-Net and U-Net like models have been successfully used in variant modality microscopy images such as fluorescence [14], nuclei [15], electron, endoscopy, dermoscopy, etc. The proposed method, employing

a symmetric contracting and expanding path to capture context and path for precise localization respectively. U-Net added up the skip connections for encoder-decoder segmentation networks which transmit the features which learned from encoder to the decoder which were effective in decreasing of the disparity between the encoder and decoder learned features.

Among the deep learning approaches microscopy images in PCM and nuclei segmentation, U-Net [4] [16]and U-Net based methods has robust accuracy in compare of are the most of popular deep learning models and achieve the state-of-the-art segmentation in robustness, completeness and other qualification factors of segmentation. In [17], Mask R-CNN architecture is leveraged as a solver for the segmentation of PCM images. The proposed method is powerful on the extraction of the cell shapes on adjacent and touching cells, however it is not easily applicable to PCM image datasets corresponding to different biological domains. Moreover, the proposed approach has a high level of complexity in computation to reach the determined results.

Recently generative deep learning models has achieved remarkable attention among the deep learning models. In [18], a Generative Adversarial Network (GAN) is proposed in the segmentation of microscopy images. The proposed method is shown its robustness in cell segmentation even with the low number of training example and achieve high performance and achieved the accuracy of 80.6% on segmentation of H1299 data set [19]. In the same manner, [20] addressed the problem of cell segmentation by leveraging of GAN. The proposed method utilized the GAN is an the extractor of cells object and transfer the learned representation to the CNN network which pre-trained on the segmentation of the cells which is called iRPE.

In [21], key point graph based bounding boxes is presented to a deep learning framework for multi-resolution cell instance segmentation which was successful in separation of the overlapping cells and adjacent cells. This methods applied the Residual Neural Networks (ResNet-50) in the backbones of the model encoder which has the U-Net like architecture.

In [22], an architecture is propose which was inspired from the of U-Net model. This model replaces Convolutional Long Short Term Memory (C-LSTM) with the common convolution layers of the U-Net to encoder and decoder of the architecture. The proposed

model earn the mean Jaccard Index of **81.1%** and **79.3%**on Fluo-N2DH-SIM+ and DIC-C2DL-HeLa dataset [23] respectively.

1.3 Hypothesis

Machine learning applications in the field of medical image analysis is mostly problem dependent, and the number of annotated images is very low. Especially in the area of medical image segmentation, the preparation of a vast amount of the annotated datasets is a tedious and time-consuming procedure for the experts. Inspired by the power of the deep learning application in the extraction of the discriminative features, we hypothesis the auto-encoder networks, which are inspired by U-Net [4] like architectures for the cell segmentation. In our investigation, we employed the on the MDA-MB-231 [9]. Firstly, we proposed the architecture which has the auto-encoder structures with the backbones of ResNet18(Conv1-Conv5) [7], which has the initial weights of the ImageNet [24]. Moreover, we repeat this process for Feature Pyramid Network(FPN) by excluding the dense layers of the ResNe18 (Conv1-Conv5) in the backbone of the FPN. Secondly, we redesign the skip connection between the encoder and decoder path, which reduces the semantic gap between the encoder and decoder by applying the specific residual blocks. This alternation of the skip connection makes the model capture more discriminative information, which causes more robustness in segmentation results in the leveraged dataset. Furthermore, as post-processing, we employed Test Time Augmentation(TTA) in which are useful tricks for measuring the level of uncertainty estimation of the architectures, but also it increases the precision of the segmentation results in our utilized dataset.

1.4 Contribution

The contribution of this thesis has two main perspectives: the first one is preparation of the specific phase contrast microscopy dataset with the help of expert for pixel classification of the cell boundaries. The second part is applying the method which increase the accuracy of the model in the segmentation task. We extended the concept of network which is inspired by U-Net with effective modification which increase the accuracy of the model in comparison of U-Net and other state-of-the-art methods for segmentation of cells in two-photon microscopy images.

We proposed a model which is inspired with U-Net with applying the ResNet18 backbone as feature extractor of encoder in our proposed model. So, instead of the sequential convolution which is applied on the original U-Net, we have the ResNet architecture (Conv1-5) with elimination of the last dense layer. Furthermore, we applied the same procedure on Feature Pyramid Network(FPN); In FPN, we just replace the backbones of the encoder to ResNet-18 which is outperform in compare of the other variants of the residual network on the MDA-MB-231 datasets.

One of the significant improvement of the thesis is applying the residual blocks between the skip-connection which propagate the input features from encoder to the decoders. This residual blocks is applied to alleviate the trace of the gap between the semantic gap between the encoder and decoder parts of the network. This utilization efficiently increase the accuracy of the res18-UNet in both of MDA-MB-231 and DSB2018 dataset.

The interesting contribution of the thesis is applying the Test Time Augmentation(TTA) which is commonly employed on classification application. We applied the specific type of TTA which is effective on improving the accuracy of the the model in segmentation tasks. Moreover, TTA is robust measurement in prediction of the uncertainty estimation of the model, specially on the proposed model.

The experimental results has applied on MDA-MB-231 dataset which is prepared with the help of experts and DSB2018 [25] dataset which is a public benchmark dataset. The proposed method showed high gain in semantic segmentation in compare to state-of-the-art alternatives and also achieve high accuracy in segmentation even with the low number of training set. Therefore, this achievement significantly reduce the level of workload in the process of annotation step.

2. CONVOLUTIONAL NEURAL NETWORKS

2.1 Component

In this thesis, we leveraged the Convolutional Neural Networks(CNN) to segmentation of medical images. CNN has different layers that are complicated in Deep Neural Network(DNN) with more layers of representation power. In this section, we define the Deep Neural Network different segments and elaborates it on different parts of deep learning architecture.

2.1.1 Activation Functions

In Neural Network, for breaking up of the linear functions, activation functions are utilized to enumerate the non-linearity to the neural network. Activation Function is used active neurons whenever the each neuron input are related to model prediction. The common activation functions which is used in deep learning is comprising of Sigmoid, Tanh, Rectified Linear Units(ReLU) and Leaky Rectified Linear Units(LReLU) which is formulate in Equation2.1, 2.2, 2.3 and 2.4 respectively. Moreover, Visualization of the determined activation function is shown in Figure2.1.

Sigmoid

$$S(a) = \frac{1}{1+e^{-a}} \tag{2.1}$$

Tanh

$$tanh(a) = \frac{e^{a}-e^{-a}}{e^{a}+e^{-a}} \tag{2.2}$$

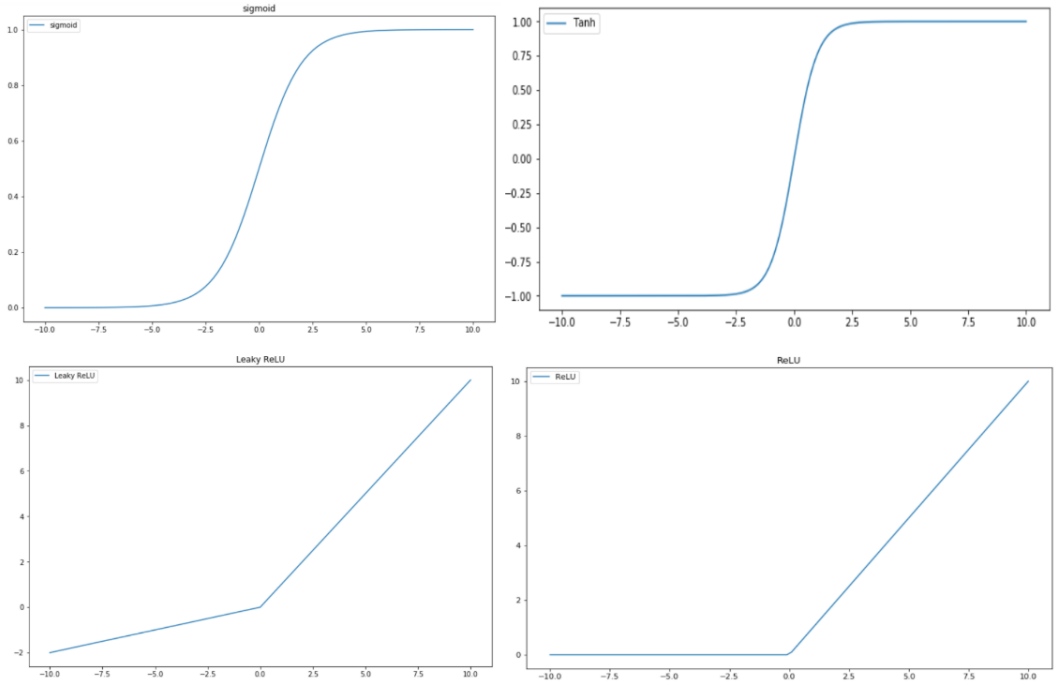

Figure 2.1 : Visualization of Activation function. The represented activation fucntions are Sigmoid, Tanh, Leaky ReLU and ReLU respectively.

ReLU

$$R(a) = \begin{Bmatrix} a & a > 0 \\ 0 & a <= 0 \end{Bmatrix} \tag{2.3}$$

Leaky ReLU

$$R(a) = \begin{Bmatrix} a & a > 0 \\ \alpha a & a <= 0 \end{Bmatrix} \tag{2.4}$$

2.1.2 Pooling

The pooling layer is a sub-sampling layer that reduces the spatial size of the input features. The down-sampling can happen with applying the stride in convolution; But the more common and well-known technique is using the pooling function. This input feature contains is also contain the crucial elements of the previous input features. Among the deep learning literature, Max-pooling and average pooling are the most addressed pooling functions. The max-pooling calculates the max value of each patch on the feature map, and average pooling makes the average of each patch in the feature map.

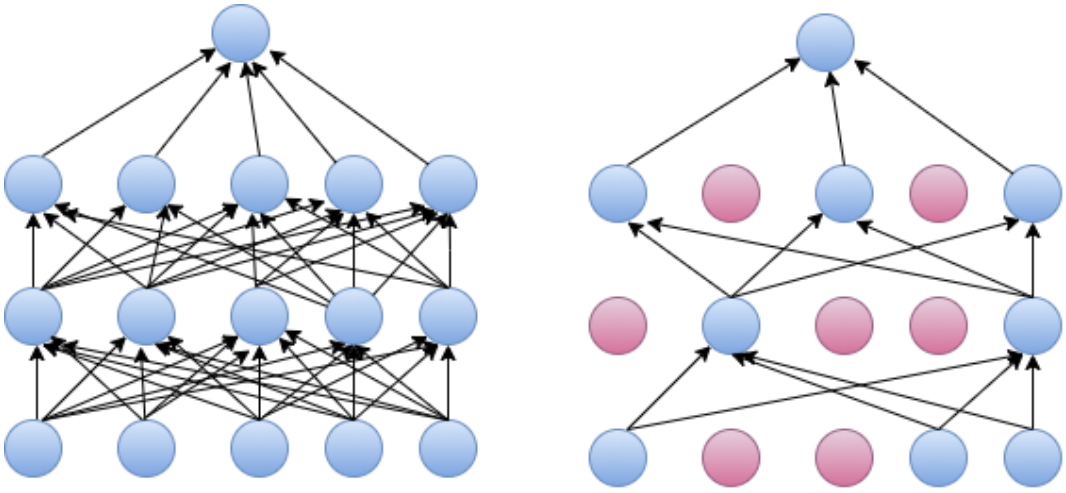

Figure 2.2 : Representation of the model before and after applying the dropout [1].

2.1.3 Dropout

In deep learning, regularization is one of the significant technique which is applied to alleviate the trace of the over-fitting during the training phase. Technically, in the dropout the nodes of model in training section are either kick off with the probability of p or kept with the probability of $1 - p$. Figure 2.2 shows a neural network before and after applying Dropout [1]. Moreover, this utilization force the neurons to learn more robust and reduce the rate of error.

2.1.4 Batch Normalization

Normalization is a necessary layer of the deep learning models to speed up the training process. There are different kinds of normalization contains of Batch normalization(BN) [26], group normalization [27], spectral normalization [28], instance normalization [29] and local response normalization [30] which is common in the literature of the deep learning applications. among the discussed normalization method, Batch Normalization is an optimization method that is widely leverage for the training of deep learning architectures. Batch Normalization has the behavior of regularization effect on the deep learning models.

2.2 Loss Functions

In this section, we illustrate the losses which are utilized for training of the models in this thesis and explain the other important losses which are used in the image

segmentation, particularly medical image segmentation.

2.2.1 Cross Entropy

Cross Entropy loss is mainly employed on multi-class classification applications. This loss is employed to improve the level of the divergence in label prediction in classification tasks. The probability value is located between the range of 0 and 1. Calculation of the probability for each class happens after applying the softmax function. The loss function is calculated separately for each class and summing up the loss for each class. The Equation2.5 is the formula of the cross-entropy whenever C>2 where C is the number of the class and p and y is probability prediction of observation and a binary indicator for each class in the observation of o. The Equation2.6 is the formula of the Binary Cross Entropy, which y is the generated label, and p is the predicted probability of the point for all points.

$$L(y,\hat{y}) = -\sum_{k=1}^{C} \log \hat{y}, c \tag{2.5}$$

$$L = -\frac{1}{N}\sum_{i=1}^{N} (y\log(p) + (1-y)\log(1-p)) \tag{2.6}$$

2.2.2 Dice Loss

Milletari et al. [91] proposed dice loss for the segmentation of volumetric data. Dice loss is inspired by the Dice metric, which is not differentiable. Let p_i and p_j represent the value of the corresponding pixel in prediction and ground truth, respectively. Therefore, the equation of Dice Loss is written as in Equation 2.7. Dice loss is a better option for overlap measures, and also there is no need to weight processing to create the balance between the background and foreground pixels.

$$\mathrm{DL}\left(p_i, p_j\right) = 1 - \frac{2\sum p_i p_j}{\sum p_i + \sum p_j} \tag{2.7}$$

3. DEEP LEARNING MODEL

In this chapter, we are scrutinizing the details of the state-of-the-art method besides the conventional methods, which we have employed some of them in our experimental results. Moreover, at the end of this chapter, evaluation metrics which we have employed in our experiments for measuring the performance of the methods are illustrated.

3.1 Fully Convolutional Networks

Fully Convolutional Networks(FCN) is proposed by Long et al. [2]. This model employed an encoder-decoder architecture, which is used VGG layers to the backbone of the encoder and decode them with up-sampling layers for the reconstruction of the learned features. This model eliminates the trace of a fully connected layer and replaces it with the convolution layer in the last layers of FCN. Instead of using fully connected layers, they only use convolutional layers. Pixel-wise loss functions are used sparsely for pixel-wise prediction of semantic segmentation map. Instead of the train a segmentation network from scratch, they fine-tune classification models such as AlexNet, VGG16, and GoogLeNet.

3.1.1 SegNet

SegNet [3] is one of the encoder-decoder based models which is applied on road segmentation applications. This method contains 13 convolution layers like VGG-16 with excluding of the fully connected layer, which makes the model faster and reduces the number of parameters from 147M to 14.7M. In each stage, feature map using the memorized max-pooling indices which store the corresponding max-pooling location in each operation. In the end, the pixel-wise classification layer is applied, which is used to predict the class for the corresponding pixel.

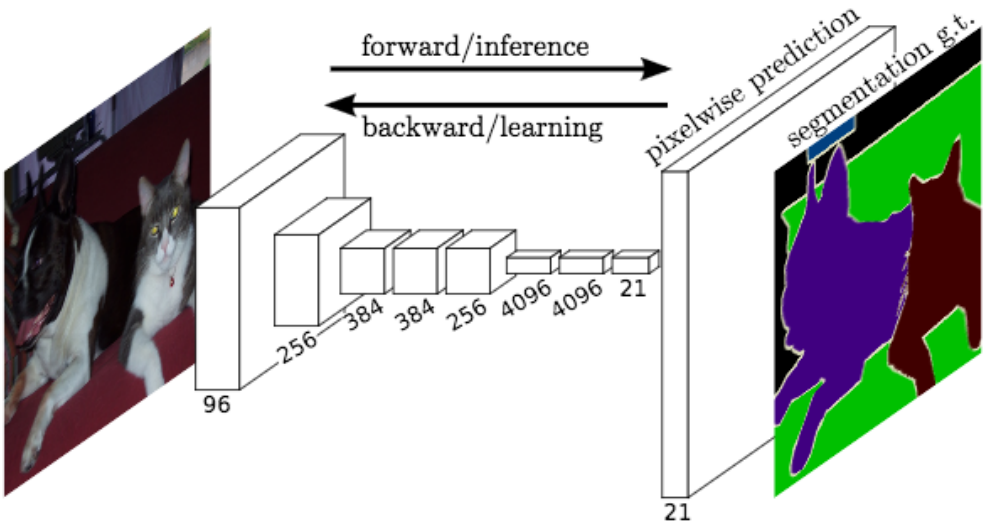

Figure 3.1 : Fully convolutional neural network [2].

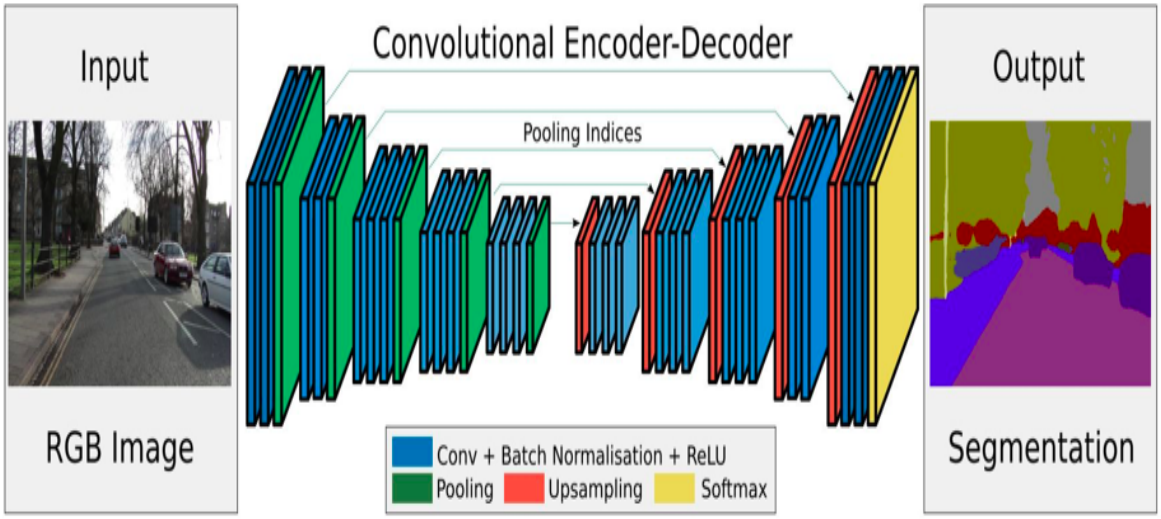

Figure 3.2 : The Schematic representation of SegNet [3].

3.1.2 U-Net

U-Net is one of the successful methods in the field of medical image segmentation, which is proposed by Ronneberger et al [4]. Due to the time-consuming of the annotation of medical images, especially for the segmentation task, U-Net with its encoder-decoder architecture has proven its high accuracy even with a low number of annotated images. Therefore, U-Net carries high importance in the field of medical image segmentation. This architecture is comprising of three main segments: contraction, bottleneck, and the expansion section. The contraction path is containing the pair of 3×3 in each stage, which is followed by 2×2 max-pooling for reduction of the spatial dimension in

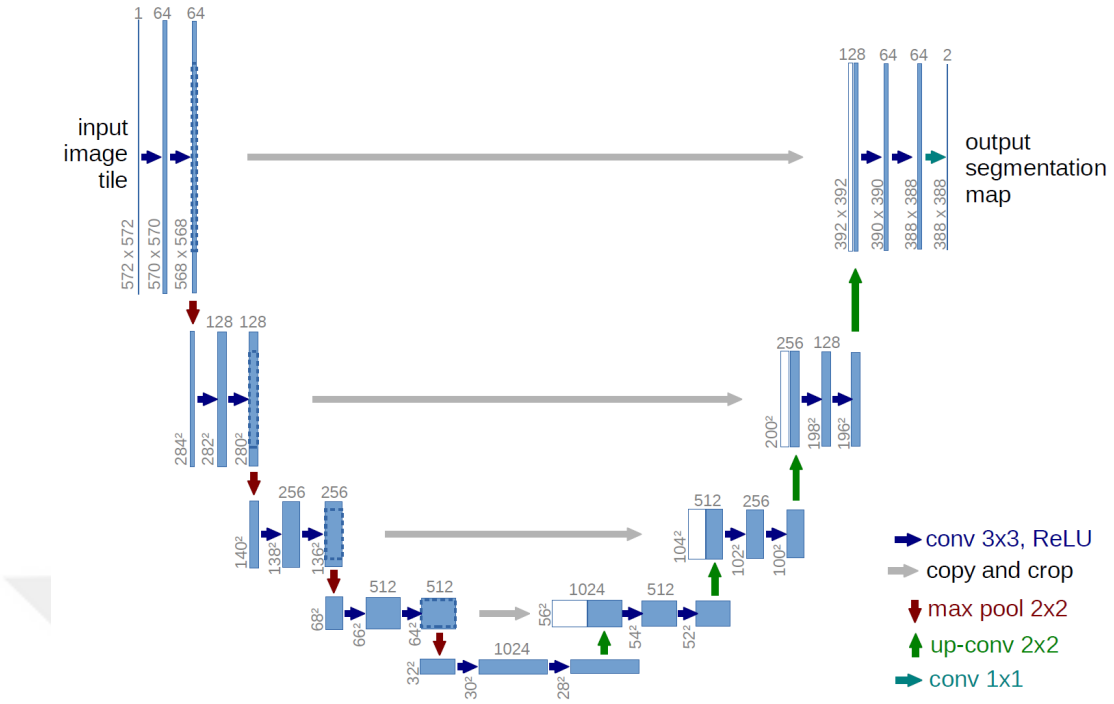

Figure 3.3 : The representation of U-Net Architecture [4].

each stage; Moreover, the number of the kernel in each feature maps is regularly added up to increase the learning capacity. In the Expansive path, there is the same number of up-sampling to save the symmetric operation in the proposed network. This path is leveraged reconstruct learned feature in contracting path. This patch is contained of 3×3 convolution, which is followed by 2×2 up-sampling layers with transposed convolutions. On the contrary to the contracting path, the number of activation maps is decreased while increasing the spatial dimensions of features in the Expansive path.

Due to robust performance of U-Net and U-Net like architecture in the literature of medical image segmentation, U-Net has the variant type of architecture which it is express in the following. To start with, Zhou.et. al [31] replace the skip connection of the encoder and decoder with dense block pathway which is inspired from DenseNet architecture. R2U-Net is is inspired from U-Net architecture with replacing of residual recurrent conv-relu blocks in the network.

As an extension of U-Net architecture, Khened et al [5] integrate the attention gate on the head of U-Net architecture. The attention gate is applied beyond the concatenation operation to have the specific activation to the given task. In the backward pass, the gradients are down-weighted from the background regions, which make the models parameters free in prior layers to be updated upon to the spacial region for the domain

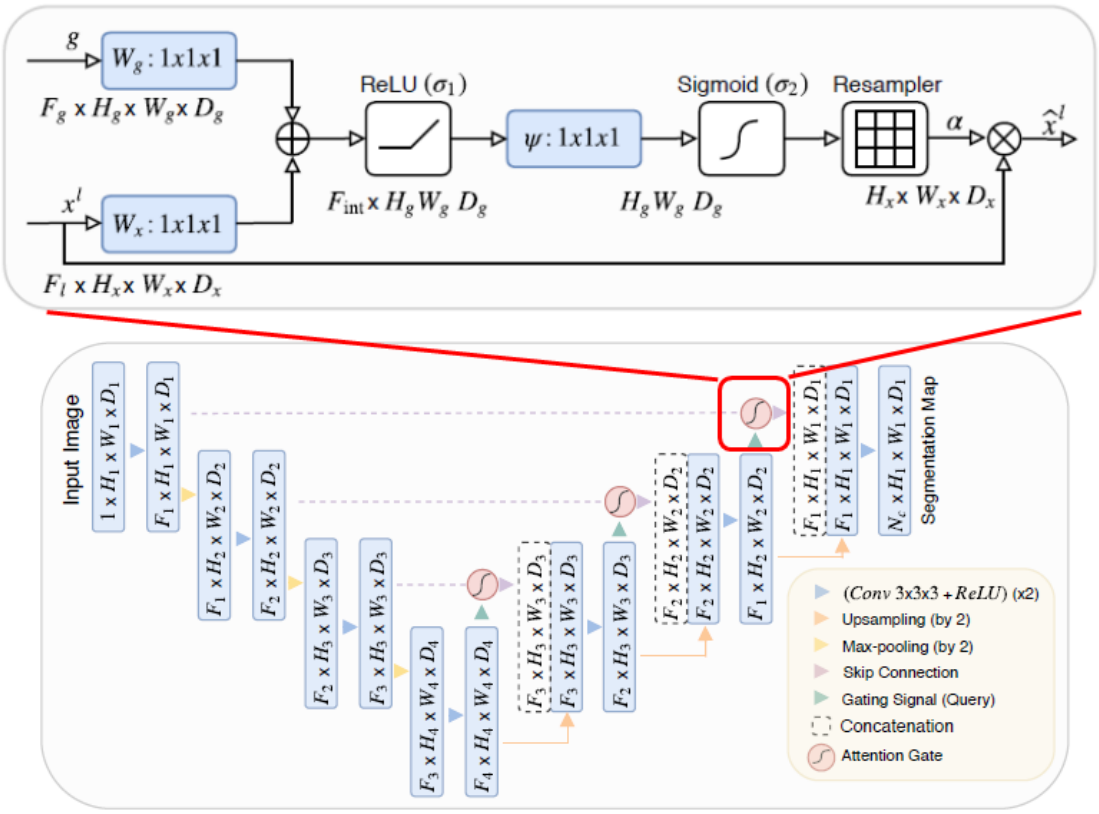

Figure 3.4 : Schematic diagram of Attention U-Net [5].

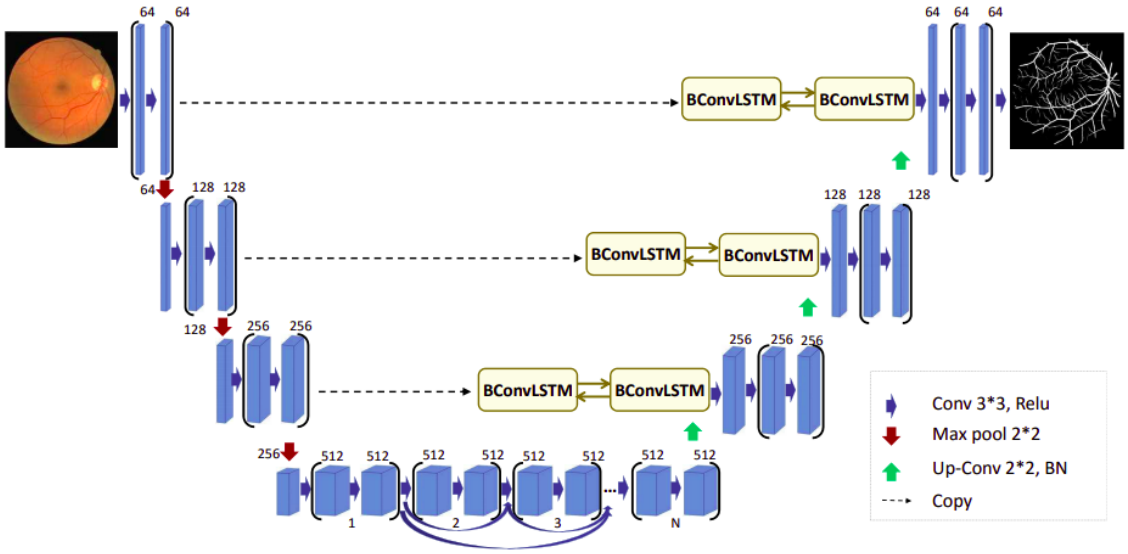

Figure 3.5 : Representation of the architecture of BCDU-Net [6].

of the target task. The schematic of attention U-Net is shown in 3.4. In the recent works in the field of medical image segmentation, an extension of U-Net is introduced by Azad, Reza, et al. [6] which integrated the bi-directional ConvLSTM(BConvLSTM) densely connected layer into U-Net instead of applying the plain skip-connection which concatenate the features from encoder to the decoder and is know as one of the important state-of the art method in the field of medical segmentation. The schematic architecture of the BCDUNet is shown in Figure3.5.

3.1.3 Residual Neural Network

Residual Neural Networks(ResNet) [7] is employed to add skip connection, which allows information to flow from input to the output. In this approach, the underlying mapping is learned by the residual mapping in comparison to underlying mapping, which makes the model learn faster. Let suppose that we make the model learn an $H(X)$ function, which is mapping the x from input to the output after the specific convolution layers. If we add the X from input to the output, we make the model to estimate $F(X) = H(X) + X$, which is easier to learn of the $F(X)$ in comparison of the $H(X)$. The different versions of ResNet is shown in Figure. We applied the ResNet18 in our experiments as the backbone of our proposed network.

3.1.4 Transfer Learning

The deep learning method is starving for a large amount of dataset, and gathering a large amount of dataset, which annotated is a tedious task, especially in the medical images. So, training an architecture that has a high number of parameters needs a vast number of annotated datasets. Transfer learning is a technique which transfers the knowledge of pre-trained-model which is trained on large scale annotated dataset like the ImageNet to the magnitude of our target domain in a specific problem. Transfer learning is applicable for fine-tuning and feature extraction, and we can deploy it in many applications of deep learning, especially classification and segmentation tasks.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Figure 3.6 : The architecture detail of ResNet with different residual layer [7].

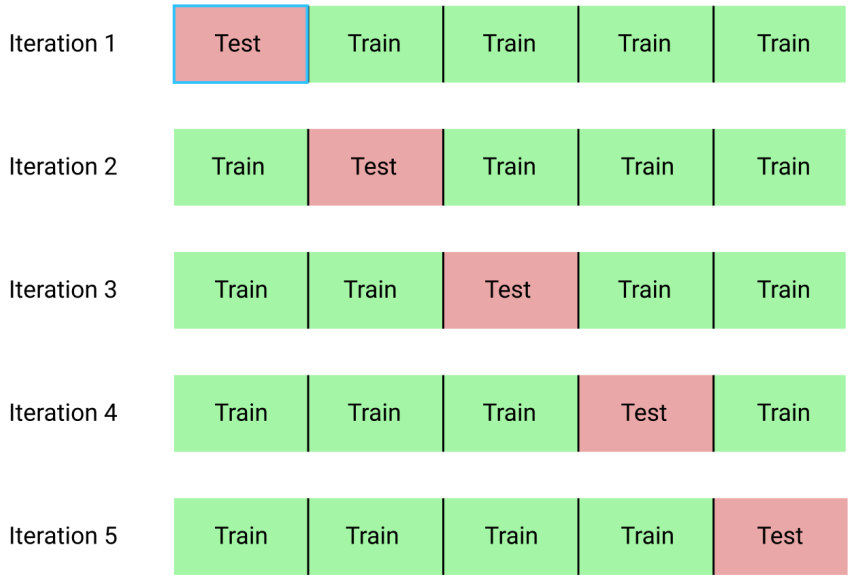

Figure 3.7 : Schematic representation of K-Fold Cross Validation when k=5.

3.1.5 Fine-Tuning

Fine-tuning is a specific type of transfer learning which is only applied on top or more layers of a pre-trained model. Fine-tuning is employed to improve the accuracy of our target domain. If there are enough datasets in our target domain, we can fine-tune the classification layer plus some of the top layers to extract more representative features specific to our target domain.

3.1.6 K-Fold Cross Validation

Evaluation of the performance of the machine learning approach can be challenging whenever we have a low number of the annotated dataset for our test set in the prediction step. Cross-Validation is a technique that measures the efficiency of the algorithm. In k-Fold Cross validation [32], the method traversed randomly among the dataset and split it into training and test set in the independent dataset. Moreover, This technique is applied for ensuring the balance between bias and variance. The schematic representation of the K-Fold Cross Validation is shown in 3.7 when k is assigned 5. In K-Fold Cross Validation, Dataset is to k equal folds and each dataset is The algorithm is run k times subsequently, each time taking one of the k splits as the validation set and the rest as the training set.

3.2 Evaluation Metrics

In order to show the efficacy of the algorithm, evaluation of the images from different perspectives is necessary. Moreover, some of the standard evaluation metrics, in some cases, can not show the real performance of the algorithm from a different perspective. Particularly, in the field of cell segmentation, the RoI has contained the small parts of the images. Hence, employing one metric for evaluation of the method leads to a wrong understanding of the real performance of the algorithm. In the following, we cover some of the well-known evaluation metrics [33], which is common in the field of segmentation, especially in cell segmentation applications.

3.2.1 Precision-Recall

Precision represents the accuracy of the model, which is predicted positive among the number of actual positives. Applying the precision is necessary whenever the cost of false positive is high among the actual true positive. Therefore, Precision measures how much of the predicted area is correctly overlapped with the RoI in cell instance. The precision and Recall formula is shown in 3.1 and 3.2 respectively.

Recall is the represents the fraction of the tp among the actual positive class. Moreover, Recall is applied to show the relevancy of the point of interest in the datasets. In simple words, the formulation of the Recall is the number of true positive over the false positive and true positive, which is shown in3.2.

$$Precision = \frac{n_{tp}}{n_{fp} + n_{tp}} \tag{3.1}$$

$$Recall = \frac{n_{tp}}{n_{fn} + n_{tp}} \tag{3.2}$$

3.2.2 Jaccard Index

Jaccard Index, which is also called Intersection over Union(IoU), is one of the useful metrics on the evaluation of the algorithm in the fields of computer vision and pattern recognition methods. This metric is applied to measure how the predicted bounding box in the prediction step is overlapped to the ground truth. The IoU formula is shown

in 3.3.

$$Jaccard(X,Y) = \frac{|X \bigcap Y|}{|X \bigcup Y|} = \frac{|X \bigcap Y|}{|X| + |Y| - |X \bigcap Y|} \tag{3.3}$$

3.2.3 Dice Coefficient

Dice Coefficient is metrics that are necessary for measuring the balance between Precision and Recall and represent the harmonical average between the Precision and Recall. Dice Coefficient is effective whenever the cost of the FN and FP is different. The formula of Dice is shown in 3.4.

$$Dice(X,Y) = \frac{2|X \cdot Y|}{|X| + |Y|} \tag{3.4}$$

4. METHODOLOGY

In this chapter, we introduce the dataset which we used in our experiments besides the proposed method. In section 4.1, we explain the details of the dataset structures. The dataset which we used is the MDA-MB-231 [9] and DSB2018 [25]. In section 4.2 and the following subsections, we introduced the proposed method, which we have applied in the datasets. The subsections encompass the models and post-processing trick, such as Test Time Augmentation(TTA), which we applied for increasing the accuracy of the method.

4.1 Dataset

4.1.1 MDA-MB-231

MDA-MB-231 [9] is the invasive breast cancer cells that which has mesenchymal morphology and captured using an Olympus IX71 microscope. Some frames of the MDA-MB-231 dataset has been shown in Figure4.1. To scrutinizing of the cell's specific characteristics such as viability, infection and proliferation, the cells are tracked daily by inverse light microscope. After imaging the cells in 6-well petri dishes in optimized numbers, they will be expected to adhere for 24 hours. Cells of different densities will be tested to visualize individual cell movements. For wound closure images, cells will be planted in a number that will reach 100% fullness, and the wound will be created by scraping the surface in a line with the yellow tip just before imaging begins. 6-well petri dishes for imaging will be placed in the incubation chamber providing 37°C temperature and 5% CO2 of the Leica SP8 microscope system and phase contrast images will be taken from the designated areas for 48 or 72 hours every 30 or 60 minutes. For analyzing the continuity of the cells period stocking is applied regularly. Finally, the cells were frozen in a specific solution includes FBS and DMSO. As pre-processing of the dataset, we adjusted the level of contrast and normalized the pixel values to the [0,1] range. The final dataset contains 600 frames of PCM images, each with a dimension of 2568×1912

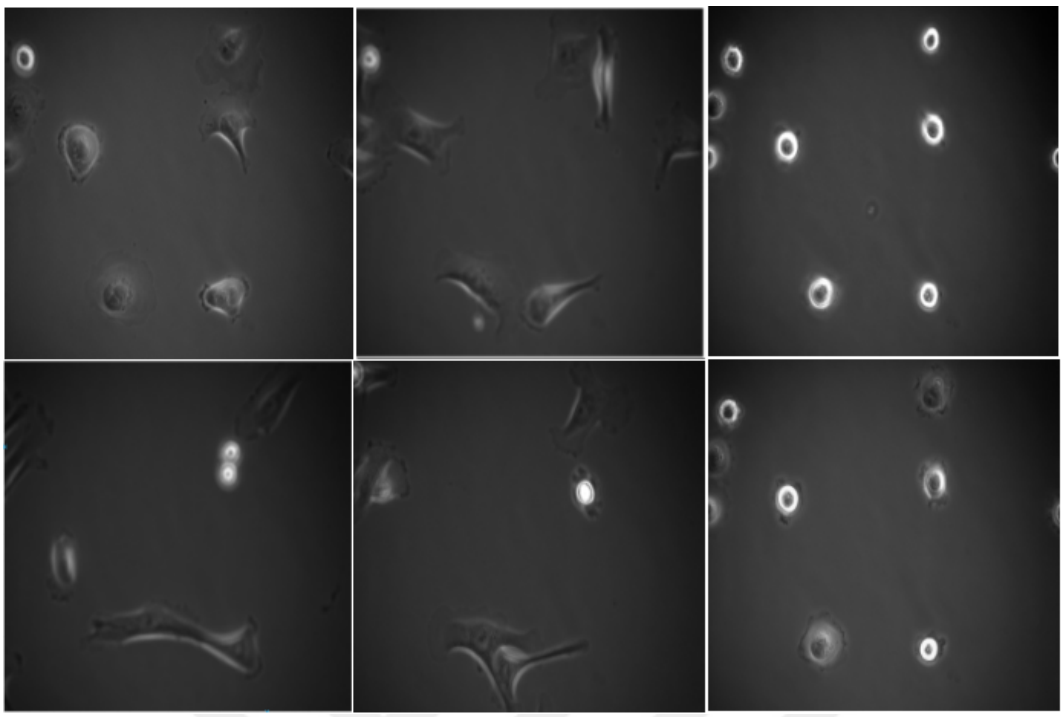

Figure 4.1 : Representation of breast cancer MDA-MB-231 dataset on Phase-Contrast Optical Microscopy images.

pixels. For manual annotation of the cell boundary on the frames, the Fiji distribution of ImageJ [34] [35] is leveraged. We annotated manually 45 of the frames, where 30 of them are used for training while the rest is utilized for the test set.

4.1.2 DSB2018

Data Science Bowl 2018 [25] is one of the challenging datasets which is containing the nuclei cells with different illumination and appearance. Moreover, the size and shape of the nuclei in different frames vary, and the nucleus are overlapped in some frames in the crowded frames. The dataset has a different modality of fluorescence microscopy, such as bright and dark field fluorescence microscopy. Some frames of the DSB2018 is shown in Figure4.2. We employ our experiments only on the grey-scale images and exclude the pathological frames from the dataset. The dataset contains 600 frames which We split the dataset into training and validation set and employed the 527 of the dataset for training set and applied the rest of them for validation set.

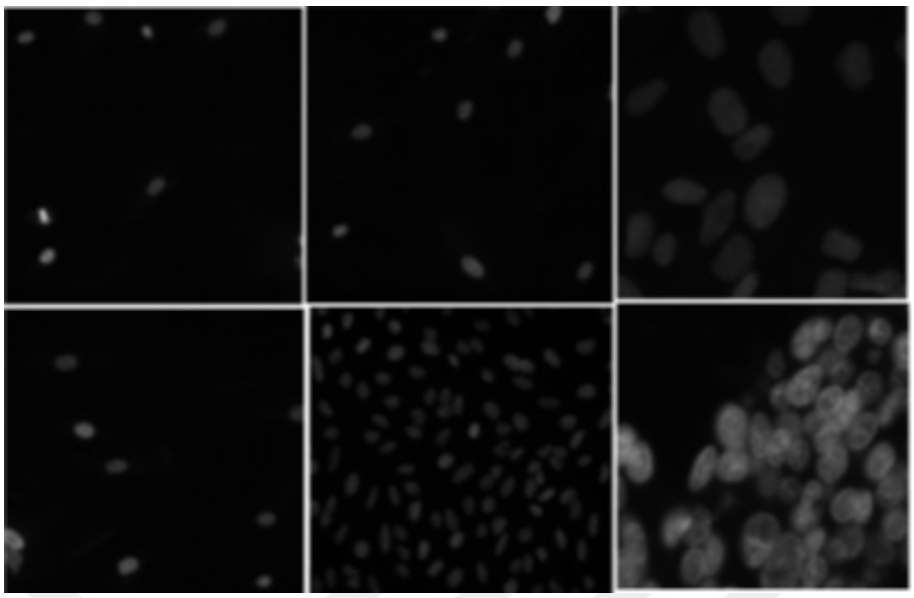

Figure 4.2 : Representation of breast cancer DSB2018 dataset images.

4.2 Proposed Method

In this section, The proposed method which we employed in our experimental results is explained. In section 4.2.1, we explained the Feature Pyramid Network; section 4.2.2 introduced a model which is inspired U-Net. In section 4.2.3 and 4.2.4, the structure of ResNet18-U-Net and ResNet18+Residual path is explained. Moreover, Test-Time Augmentation(TTA) is introduced, which we applied as post-processing in our experimental results.

4.2.1 Feature Pyramid Network

Feature Pyramid Network(FPN) [36]comprises of bottom-up and top-down pathways as seen in Figure4.3. In the bottom-up pathway, we employed ResNet18(Conv1-5) as the backbone of the model's feature encoder, and in each block, the stride is doubled in each stage to reduce the spatial dimension of the pyramid level. The detail structure of the ResNet18 represents in Figure in details. To reduce the dimension of the last block in FPN, 1×1 convolution with the channel depth of 256 is utilized. In the top-down pathway, as we move-up, the spatial dimension of layers is increased by 2 with nearest-neighbor interpolation. Therefore, two sequential 3×3 convolutions to reach module T in each stage of the pyramid. In the end, feature maps are upsampled to the same dimension and concatenated.

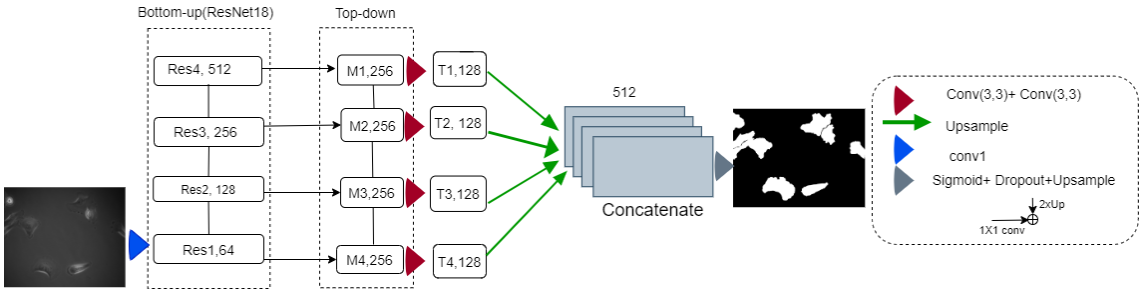

Figure 4.3 : Schematic structure of FPN with the backbone as modified ResNet18 pre-trained on ImageNet. Conv stands for Convolution in the above figure. The number in each block represents the number of channel which is employed in each stage.

In the second section of the method, all of the T modules which in the previous steps. These modules are concatenated and downsampled with the factors of 4. Then, the modules have the 512 channels, which concatenated from previous layers, which each of them has 128 channels. Afterward, 3×3 convolution filters, batch normalization, and ReLU activation function is utilized in the in this step. Finally, the number of output channels is reduced by 1×1 convolutions based on the feature map. In the upsampling path, the original image size is reconstructed with bi-linear interpolation.

4.2.2 Multi-Resolution Network

The first network we trained entitled TipNet, was a multi-resolution CNN architecture which inspired from U-Net. The architecture eliminates the contracting path and replaces it with sequential convolution 3×3, which downscale with the factors of 4. The convolution is followed with Batch Normalization(BN), which is applied to speed up the training process and increase the stability of the neural network. The network contains three main parts, each of the convolution is the 3×3 convolution filters, which is applied in downsampling and upsampling path with leaky ReLU activation function. Firstly, it operates to the lowest resolution, which is downscaled from the input image with the factors of 4; The lowest stage has the resolution of 128×128 pixel with 15 convolution layers and number of feature map sets 64. For maintaining the spatial dimensions of input features, the padding layer is applied. Afterward, this segment is followed by a 2×2 upsampling layer which followed by the segment by factor two. In this stage, there are three 3×3 convolution layers in which the channel size is determined 96 in this step. There is another 2×2 up-convolution, which double the dimensions of input

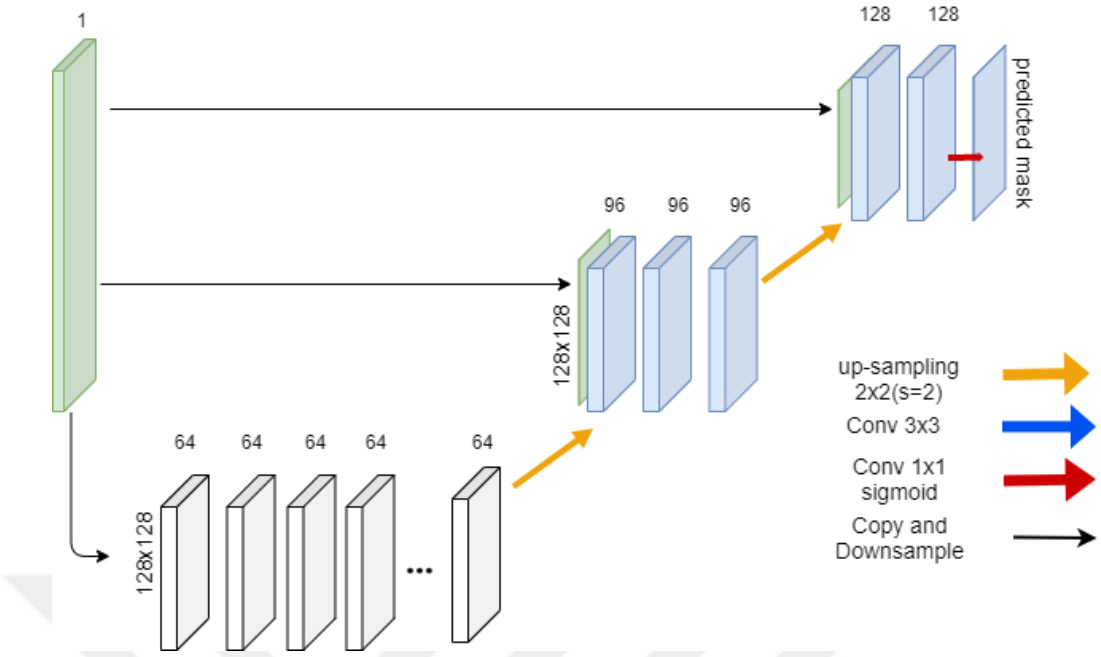

Figure 4.4 : Representation of the diagram of TipNet Architecture.

size to the original input size of the image, which is considered 512×512 pixel in this architecture. Finally, the last layer of this model is followed by a sigmoid activation function for binary segmentation task. The diagram of the TipNet is shown in Figure4.4

4.2.3 ResNet18-U-Net

We replace the backbones of U-Net with ResNet18(Conv 1-5) [7] with excluding of the dense layers from ResNet architecture. The schematic structure of ResNet18-U-Net is shown in Figure.4.5. One of the modifications which we have applied in ResNet18 architecture is the reduce the stride in Conv1 from 2 to 1. Applying the 7×7 layer in the ResNet leads to loss of important information during the downsampling operation, especially for medical datasets. Therefore, after applying of Conv1 in ResNet18, the dimension of the network is not halved. In the backbones of ResNet18, we have the two residual blocks in each resolution stage, which we have the residual block it symmetrically to the decoding path with one residual block for each dimension level. The residual blocks contain layers with 3×3 kernels where each convolutional layer is followed by a Batch Normalization(BN) layer and a ReLU activation function. The detail seting of Resblock is represented in Table4.1. In addition, the number of channels in the first block is 64, and the number of feature map is regularly doubled after doubling the stride in each resolution stage. In the decoder section, to have symmetric operation,

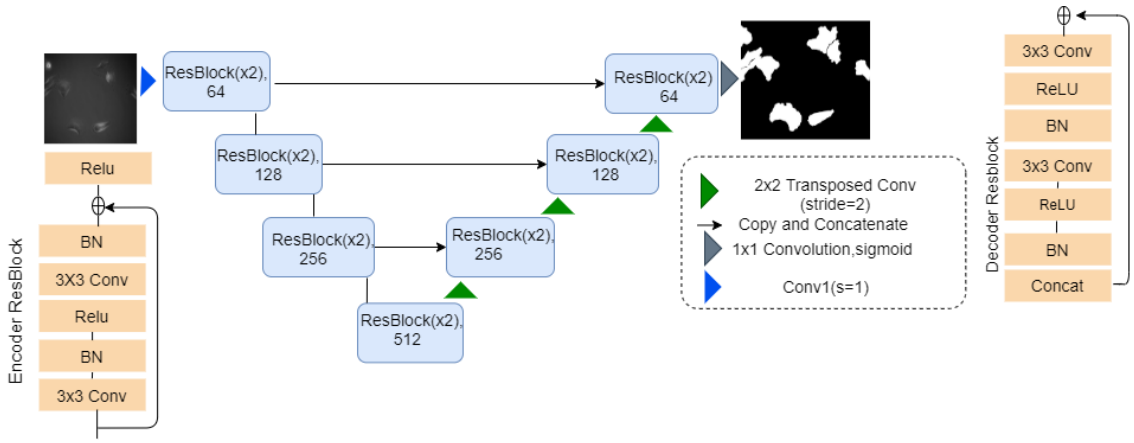

Figure 4.5 : Schematic structure of Res18-U-Net. Each number in the block indicates the number of filters in each stage.

we apply upsampling by 3 which is equal to the number of dimension reduction in the encoder part for reproducing the input images that consist of residual blocks in each resolution stage. The Schematic structure of ResUNet is shown in Figure4.5. Moreover, as an extension, we applied the residual blocks as an alternative of plain skip-connection which increase the performance of the network on MDA-MB231 and DSB2018 dataset.

4.2.4 Residual Pathway

In the encoder-decoder networks such as U-Net, there is a meaningful gap between the model's encoder and decoder features and to overcome this, some papers [31] has comes up with effective structure for this pathway. The encoder features contain high resolution and have the local information of the input images. Conversely, the encoder comprises semantic information of the input images. To overcome this shortage, we proposed a residual skip-connection, which is a robust alternative for the skip-connection of the U-Net, which concatenates the encoder features to the decoder features. Therefore, we also gradually reduce the number of convolutional blocks used along the residual pathway. About the number of Residual blocks along the skip-connection, we gradually reduce the number of residual blocks 4 to 2 from high-resolution stages to the low-resolution stage in the architecture, which we applied this on ResUNet architecture as an extension. Moreover, about the number of feature maps in the residual path(RP), we employ the set of (64,128,256) filters in the blocks of the four residual-Pathway, respectively. The schematic structure of the proposed residual pathway is shown in 4.6. The blocks contain of 3×3 filters which followed

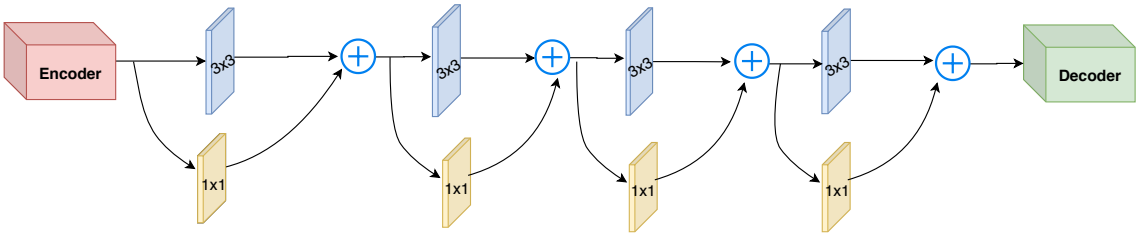

Figure 4.6 : Schematic structure of Residual pathway.

Table 4.1 : Detailed structure of Residual Pathway in ResNet18-U-Net model.

No	Layer	Detail
1	Conv(1×1)	Activation: No
2	Conv(3×3)	Activation: ReLU, Padding: Same
3	Add	—
4	ReLU(Activation)	—
5	Batch Normalization(BN)	—

1×1 filters for generating of residual connections. The 1×1 filters is kept here if we wished to apply the residual block with different size; However, the size of filter is kept same in each block and it has applied for extension purpose. Moreover, this pathway is employed on ResNet18-U-Net, which is entitled ResNet18-U-Net+RP and efficiently increase the performance of the architecture in comparison to applying the plain skip-connection in ResNet18-U-Net in MDA-MB-231 and DSB2018 dataset. The detail structure of residual pathway has been shown in Table 4.1.

4.3 Test-Time Augmentation

Deep learning architectures contains massive numbers of parameters which make them greedy for the high number of the annotated dataset. In the field of biomedical imaging, the lack of a comprehensive bank of annotated dataset makes them challenging in the prediction phase. To solve the lack of dataset in training steps, data augmentation is an effective alternation for solving the lack of annotated dataset in the training step. However, this shortage has shown its impact in the prediction step.

Test-Time Augmentation(TTA) is a technique that redesigns the data augmentation in the prediction step. This technique is efficiently applied to the classification task for the uncertainty estimation of the model and recently has been applied on the segmentation application [37]. TTA applies the prediction both on the original and on the augmented versions of the image and starts to make the decision based on the specific methods

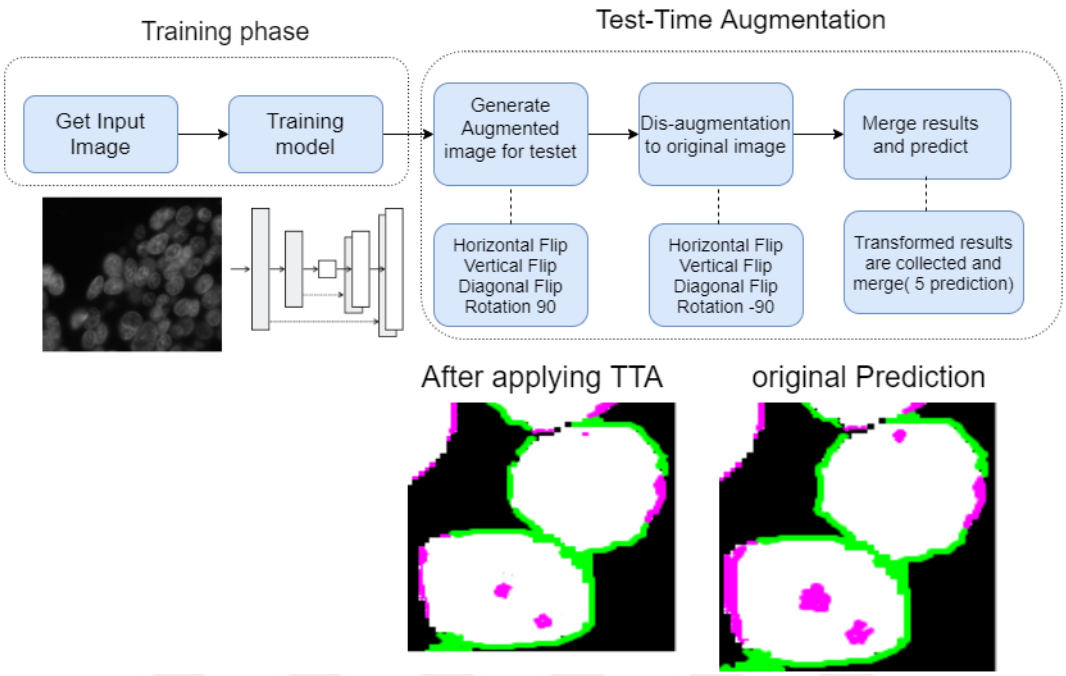

Figure 4.7 : Diagram of TTA.

which can satisfy the results in the prediction step. TTA increases the inference time on prediction steps due to the addition prediction of the augmented image in the prediction step. TTA contains the main four-step which we redesigned it for our application:

1) Augmentation: in this procedure, like the augmentation in the training phase, we employ the augmentation in the test set.

2) Prediction: The model not only predicts the original image in testset, but also predict the image which is generated in the augmentation step.

3) Re-augmentation: in this step, all of the augmented images are transformed into the original image by reversing the augmentation for each of the augmented images.

4) Merging: In this step, we have different versions of prediction from the original image, which is predicted with the different geometrical views. The rules of the merging based on the number of voting for each image by summing and averaging of all predicted image with probability maps. The schematic diagram of the TTA is represented in Figure4.7. According to the determined figure, after applying the TTA on the method the level of errors is decreased in compare of original prediction. Therefore, prediction is efficiently increased in especially in the cell borders and where has not predicted due to unbalance illumination of the RoI in the DSB2018 dataset.

5. EXPERIMENTAL RESULTS

In this chapter, we analyze the experimental results on MDA-231 datasets. The dataset encompasses PCM images. Thus, these datasets characterize the pathological dataset, which can be utilized to study the robustness of the image segmentation from deep learning models employed on medical images datasets. In section 5.1, we introduce the general implementation details of the models which applied to our investigation. In section 5.2, explains the specific details of the applied model in reported on the utilized datasets. In Section 5.3, the quantitative and qualitative results of the model are shown in MDA-231 and DSB2018 datasets.

5.1 Implementation Details

We applied data augmentation to the annotated training set, where the main transformations utilized comprised of random elastic transformation, rotation, and horizontal and vertical flip. In the training phase. To have a robust prediction, we applied(TTA) involving horizontal, vertical and diagonal flip and the rotation at 90 degree. In the prediction step. Moreover, spatial Dropout is utilized with the rate of 0.5 at the last layer of the ResNet18-FPN network. Batch size is considered 8 for the network, and weights of the backbone in the encoder are pre-trained in the ImageNet [24] in ResNe18-UNet and ResNet18-U-Net+RP. Adam [38], is considered as the optimization function and loss function is controlled in the validation set to be stopped for avoiding the method from over-fitting. All of the experiments have been realized on a single NVIDIA TitanX graphics card with 12 GB memory. The training set was resized to 518×518 px and 262×262 px in ResNet18-U-Net and ResNet18-FPN to reduce memory consumption in MDA-MB213 dataset. The experiments are implemented in Keras [39] with Tensorflow backend. We consider the different evaluation metrics to perform the experimental comparative, including Precision, Recall, Jaccard Index and Dice Coefficient. For the Loss function, we integrate a hybrid loss function, which is comprised of binary cross-entropy and dice loss. Binary cross-entropy is special type

cross-entropy which is applied whenever the target is either 0 or 1 and this specific modification is achievable in prediction step with applying the sigmoid activation. Moreover, Dice Loss is also sum up to the loss function which is powerful loss function in overlapping measures. The applied loss function is shown in 5.1 to 5.3.

$$BCE\,(p,\hat{g}) = -\left(p\log\left(\hat{g}\right) + (1-p)\log\left(1-\hat{g}\right)\right) \tag{5.1}$$

$$\mathrm{DL}\,(p,\hat{g}) = 1 - \frac{2\sum_i^N p_i\hat{g}_i}{\sum_i^N p_i + \sum_i^N \hat{g}_i} \tag{5.2}$$

$$L(p,\hat{g}) = DL(p,\hat{g}) + BCE(p,\hat{g}) \tag{5.3}$$

where p_i is predicted binary segmentation volume and g_i the ground truth binary volume which is part of $\hat{g}$.

5.1.1 Experimental Results on MDA-MB231

We compare the performance of our proposed approach with various baselines such as Empirical gradient threshold(EGT) [8], U-Net [4], TipNet [9], and PHANTAST [10]. The U-Net, which has applied in our comparison have (32, 64, 128, 256, 512) filters in each stage. As can be seen from Figure 5.1. According to the qualitative results, classical methods like PHANTAST and EGT are not satisfactory in the extraction of the cell border and the RoI. The proposed method outperforms the state-of-the-art methods and has noticeably robust performance in the extraction of cell boundaries. Moreover, it is resilient in the prediction of the specific cell structure. The qualitative results of the utilized methods have shown in Figure 5.2. In the visualization, magenta marks represent the FN, and green marks illustrate the FP in our proposed dataset. According to the quantitative results which has reported in Table5.1, We achieved Jaccard indexes of **85.4% ± 2.7**, **87.9% ± 1.7**, **87.1% ± 2.3** and **89.2% ± 1.3** with TipNet, ResNet18-UNet, ResNet18-FPN and ResNet18-U-Net+RP respectively. Moreover, to have better analysis on the reliability of the applied methods, we add up the standard deviation to the each evaluation metrics which delegate the average scores among the testset in the applied dataset. Therefore, the result is applied similar to (Mean±STD) on the experimental results section. The aforementioned table, represents the quantitative results of MDA-MB231 on Fold k=2 after applying the k-fold cross validation, Furthermore, there is a comprehensive table which show the IoU of each models in 3-folds. The rest of the quantitative results on MDA-MB231 on 3-folds

Table 5.1 : Qualitative results of the different architecture which applied on the MDA-MB-231 dataset. TipNet, ResNet18-UNet, ResNet18-U-Net+RP is our suggested models for k=2. The quantitave results of EGT and PHANTAST is not passed among k-fold cross validation.

Method	JI(%)	Dice(%)	Precision(%)	Recall(%)
EGT [8]	42.2 ±3.8	59.4±3.1	45.7±3.8	84.7±2.4
PHANTAST [10]	54.6±4.4	70.6±3.5	68.2±4.1	73.3±3.6
U-Net+TTA [4]	86.4±1.9	92.7±1.7	93.3±1.9	92.4±1.8
BCDUNet+TTA [6]	82.2±2.4	90.2±2.1	89.2±2.8	91.2 ±1.6
Tip-Net+TTA [9]	85.4±2.7	92.1±2.2	93.8±2.9	90.5 ±1.9
ResNet18-U-Net+TTA	87.9±1.7	93.5±1.3	95.0±1.8	92.1±1.5
ResNet18-U-Net+Rp+TTA	89.2±1.3	94.3±1.1	95.8±1.3	92.8±1
ResNet18-FPN+TTA	87.1±2.3	93.1±2	93.0±2.6	94.1±1.9
U-Net [4]	84.43±2	90.8±1.9	92.5±2.3	89.1 ±1.8
BCDUNet [6]	81.33 ±2.6	89.7±2.2	88.6±2.9	90.7±2
Tip-Net [9]	83.26±3.1	90.8±2.5	91.6 ±3.3	90.1±2.1
ResNet18-UNet	86.69±2.2	92.8±1.6	93.8±2.1	91.8 ±1.7
ResNet18-U-Net+Rp	87.26 ±1.8	93.1±1.5	94.5±1.7	91.8±1.6
ResNet18-FPN	85.25±2.6	92.0±2.3	91.9 ±2.7	92.1±2.2

with the rest of evaluation metrics in Table A.2 and Table A.3 in the appendix section. According to qualitative and quantitative results, they prove that our proposed approach not only improves segmentation accuracy but also it is effective in the separation of overlapping cells in compare of baseline and even state-of-the art method such as BCDUNet [6]. Moreover, our proposed approach has surpassed the results of U-Net and other baselines in completeness, robustness, and other factors of efficiency in the segmentation task on the MDA-MB-231 dataset.

5.1.2 Experimental Results on DSB2018

We analyse and make comparision between our proposed approach various baselines such as U-Net [4], SegNet [3]. As can be seen from qualitative results of DSB2018 [25] which has been shown in Figure5.2, the proposed method has achieve the state-of-the-art methods and has noticeably robust performance in the extraction of cell boundaries. Moreover, the proposed approach especially ResNet18-U-Net+RP has robust performance against the outliers, cell boundaries. The qualitative results of the utilized methods have shown in Figure5.2. We achieved Jaccard indexes of **81.1% ± 2.2**, **83.1% ± 1.8** and **85.0% ± 2.1** with ResNet18-FPN, ResNet18-UNet and ResNet18-U-Net+RP, respectively. Moreover, to have better analysis on the reliability

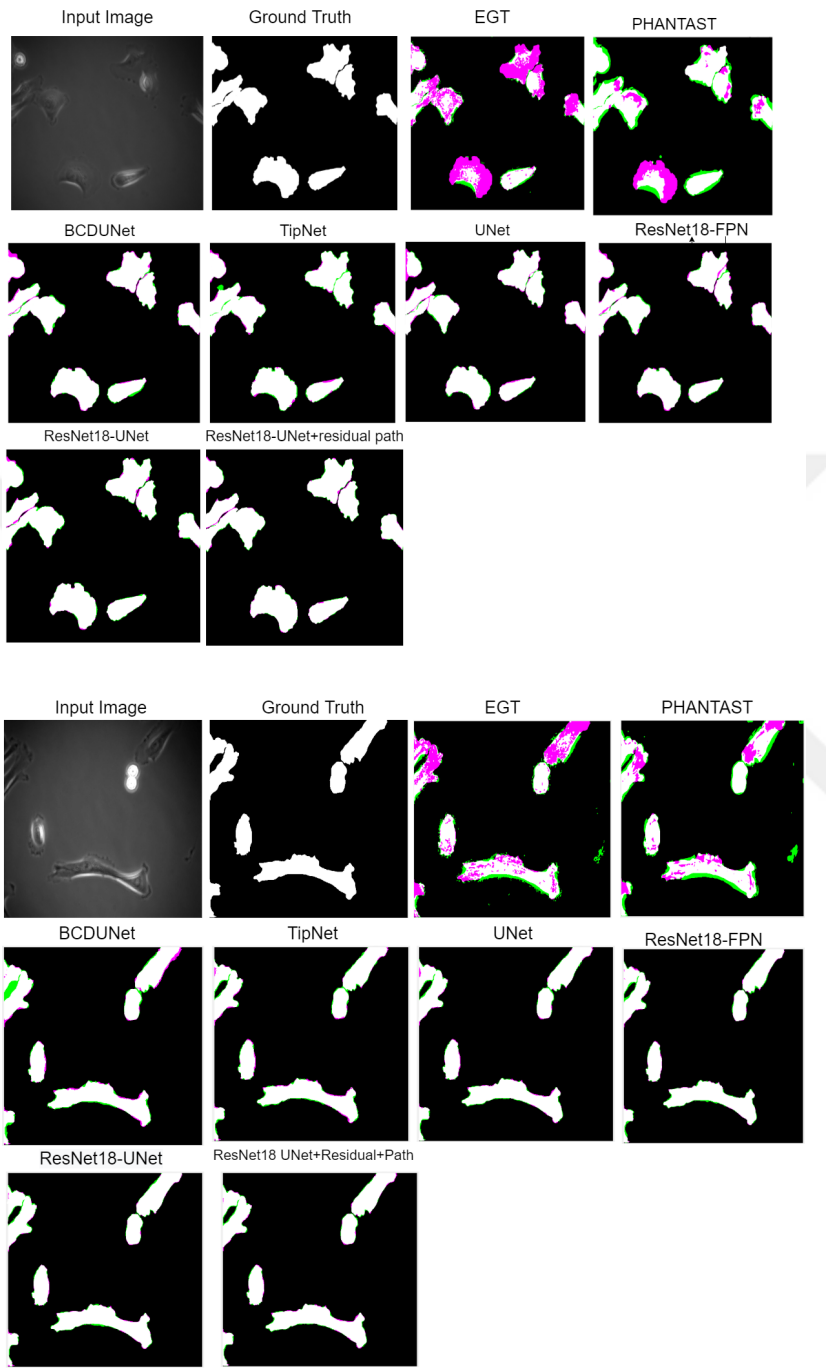

Figure 5.1 : Segmentation results for the MDA-MB-231 dataset. Magenta marks represent the false negatives and green marks indicate the false positives. The presented methods are the EGT, PHANTAST, BCDUNet, U-Net, TipNet, ResNet18-FPN, ResNet18-UNet, ResNet18-U-Net+RP respectively.This results come out after applying the TTA on the methods.This results come out after applying the TTA on the methods.

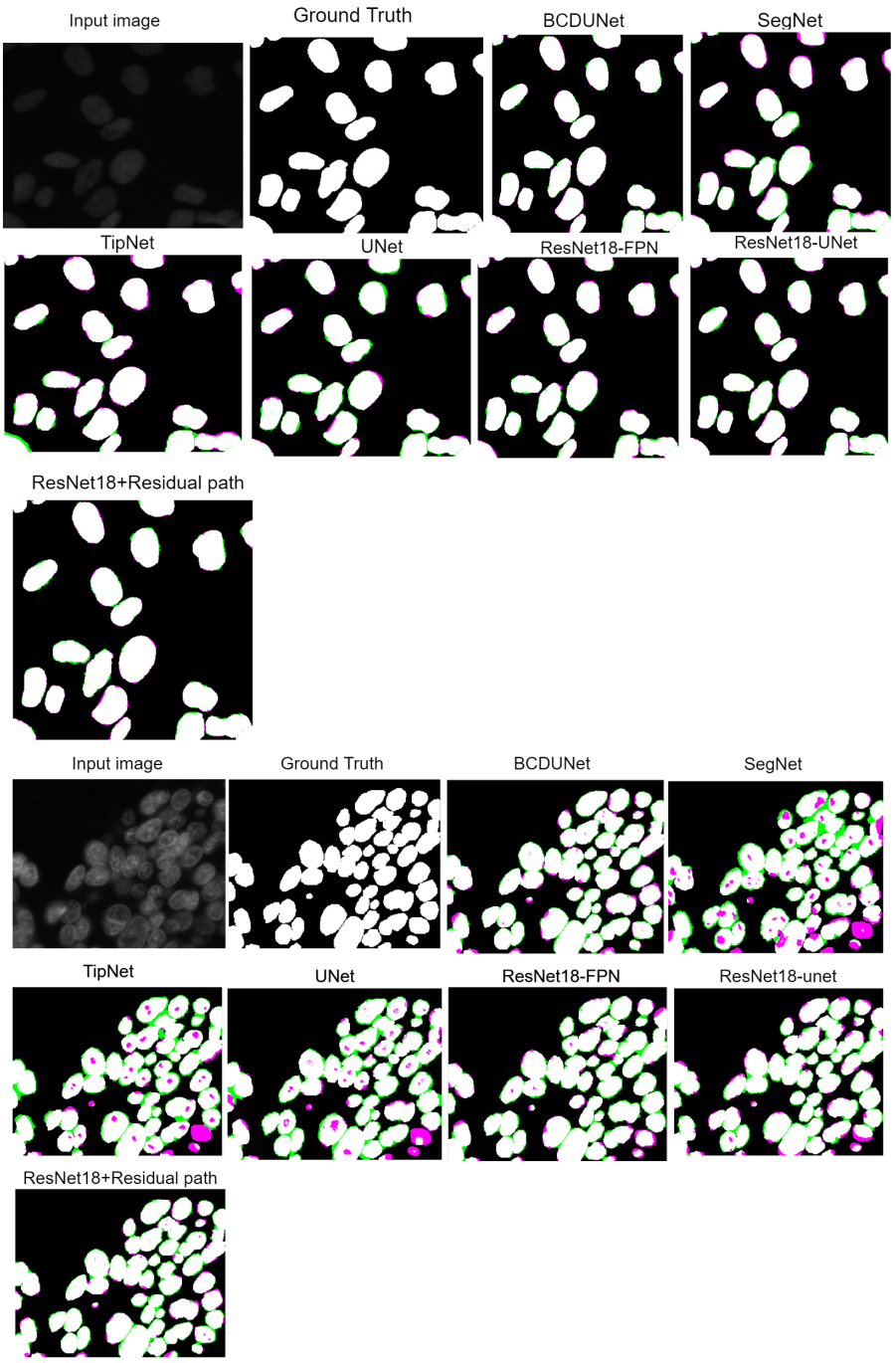

Figure 5.2 : Segmentation results for the DSB2018 dataset. Magenta marks represent the false negatives and green marks indicate the false positives. The presented methods are the BCDUNet, SegNet, UNet, TipNet, ResNet18-FPN, ResNet18-UNet, ResNet18-U-Net+RP respectively.This results come out after applying the TTA on the methods.

Table 5.2 : Quantitative results of DSB2018.

Method	JI(%)	Dice(%)	Precision(%)	Recall(%)
SegNet+TTA	72.3 ±2.9	83.9±3.8	81.5±4.5	86.5±3.3
BCDUNet+TTA	**86.7**±1.6	**92.9**±1.9	**93.9**±3.5	**91.8**±2.2
Tip-Net+TTA	74.1±2.8	85.1±2.5	87.5 ±2.9	86.9 ±2.4
U-Net+TTA	79.2 ±2.2	88.4 ±2.1	87.3 ±2.5	89.5 ±1.9
ResNet18-U-Net+TTA	83.1±1.8	90.7 ±2.2	92.3 ±2.3	90.5 ±1.6
ResNet18-FPN+TTA	81.1 ±2.2	89.6±2.4	91.0±2.7	88.2±1.9
ResNet18-U-Net+RP+TTA	85.0 ±2.1	91.9 ±1.9	93.0 ±2.4	90.8 ±1.6
SegNet [4]	69.2 ±3.2	81.8 ±4	80.1 ±4.7	83.5 ±3.6
BCDUNet [6]	**85.5** ±1.7	**92.1** ±2	92.5 ±2.7	**91.8** ±2.5
Tip-Net [9]	73.5±2.9	84.7 ±2.7	85.9±3	83.6 ±2.6
U-Net [4]	76.4±2.3	86.6±2.3	87.0±2.6	86.2±2.1
ResNet18-UNet	82.1±2	90.1±2.3	91.5 ±2.5	88.8±1.7
ResNet18-FPN	79.6±2.6	88.6±2.5	89.7 ±2.8	87.5 ±2.2
ResNet18-U-Net+RP	84.8±2	91.7 ±2.1	**92.7**±2.5	90.8±1.8

of the applied methods, we add up the standard deviation to the each evaluation metrics which delegate the average scores among the testset in the applied dataset. Therefore, the result is applied similar to (Mean±STD) on the experimental results section. It proves that our proposed approach not only improves segmentation accuracy but also it is effective in the separation of overlapping cells in some of the adjacent or overlapping cells.Due to high perturbation in some frames of DSB2018 dataset. our proposed approach has surpassed the results of U-Net and other baselines in completeness, robustness, and other factors of efficiency in the segmentation task on the DSB2018 dataset.

5.1.3 Model Parameters

The number of parameter of the method is one of the important characteristics of the model complexity which has key role on determination of the inference time in the training and prediction steps. In experimental results we have applied different type of methods which include with different level of complexity. In our proposed approach we applied the ResNet18 in the backbone of our architecture. The ResNet18 contains the 11M parameters which is partially reduce after excluding the dense layer from the discussed model. Table 5.3 represents the number of parameters in the deep learning models which has applied in our study.

Table 5.3 : The number of parameters on of the models which is used in this study.The number of parameters is based on the million which **M** is stand for it.

Method	# of params(M)
SegNet [3]	17.5
U-Net [4]	7.76
ResNet18-UNet	14.3
BCDUNet [6]	20.66
Tip-Net [9]	6.69
ResNet18-U-Net+Rp	16.31
ResNet18-FPN	13.81

5.1.4 Discussion

According to aforementioned sections, analyzing the details of the experimental results seems necessary. We have utilized our proposed methods on two different dataset which is entitled MDA-MB231 and DSB2018 respectively. For scrutinizing the performance of the methods we apply the similar configuration on the datasets to have better evaluations. For instance, we employed the TTA as an post processing method on evaluate the improvement on the performance of the methods. TTA is increased the Jaccard index between **1%** to **3%** in comparison of the proposed method without utilizing the presented methods. For instance, Jaccard index of ResNet18-U-Net+Rp is increased from **87.26%** to **89.24%**. Moreover, TTA is also effective in reduction of the STD of the results in comparison of the No-TTA methods. Qualitatively, TTA is also successful in excluding of the outlier in compare of No-TTA version of the proposed method specially on the DSB2018 dataset which is represented visually in Figure4.2. About the TTA time complexity, due to composing of multiple prediction on each original prediction, computational times of the applied method is multiply increased. By the way, due to importance of the medical application, utilizing the TTA can be effective in increasing the performance of the method from qualitative and quantitative perspectives. For instance, in our application TTA compose $4x$ computational time in comparison of the No-TTA version of the applied methods. About the loss function, we applied the Dice + Binary Cross Entropy which was success full in comparison of the single loss function whenever applying either dice or cross entropy function.

6. CONCLUSIONS AND RECOMMENDATIONS

In this thesis, we trained different convolutional neural network architecture for the segmentation of Phase-Contrast Optical Microscopy images. The second proposed method is applying the well-known classification architecture in the backbone of the auto-encoder architectures like U-Net. We integrate ResNet18 pre-trained model as an encoder of the proposed network, which has the initial weights of the ImageNet. As the Third model, we apply an extension in the architecture of ResNet18-U-Net with applying a specific convolution path between the encoder and decoder skip-connection to reduce the discrepancy between the encoder and decoder feature map information.

Our empirical results are compared based on state-of-the-art on MDA-MB-231 [40] dataset. Our proposed model achieved Jaccard Index of **85.4% ± 2.7**, **87.9% ± 1.7**, **87.1% ± 2.3** and **89.2% ± 1.3** with TipNet, ResNet18-UNet, ResNet18-FPN and ResNet18-U-Net+RP respectively on MDA-MB231 dataset and outperforms in comparison to the state-of-the-art methods and satisfies the robustness, completeness factors in the segmentation task. Moreover, discussed methods are also achieves Jaccard index of **74.1% ± 2.8**, **81.1% ± 2.2**, **83.1% ± 1.8** and **85.0% ± 2.1** on DSB2018 dataset. By the way our best method can achieve the near to the state of the art method which entitled BCDUNet which achieve the From other perspectives, the proposed methods are also robust against the outliers and reliable on the detection of the outliers which causes to reduce the amounts of false prediction in compare of baseline methods.

The future direction of this research has several branches. In Future study, we have the plan to deploy our methods, for instance, segmentation tasks and make it robust for the tracking stage. However, in the future, we wish to conduct experiments to determine the best set of hyper-parameters for the model more exhaustively. We want to evaluate the performance of the proposed methods on the different modality in medical image datasets. Besides points as mentioned earlier, we would like to deploy our methods

for time-series by the construction of lineage relationships, which provide detailed information about cell behavior.

REFERENCES

[1] **Srivastava, N.**, **Hinton, G.**, **Krizhevsky, A.**, **Sutskever, I. and Salakhutdinov, R.** (2014). Dropout: a simple way to prevent neural networks from overfitting, *The journal of machine learning research*, *15*(1), 1929–1958.

[2] **Long, J.**, **Shelhamer, E. and Darrell, T.** (2015). Fully convolutional networks for semantic segmentation, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.3431–3440.

[3] **Badrinarayanan, V.**, **Kendall, A. and Cipolla, R.** (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation, *IEEE transactions on pattern analysis and machine intelligence*, *39*(12), 2481–2495.

[4] **Ronneberger, O.**, **Fischer, P. and Brox, T.** (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation, ***N. Navab**, **J. Hornegger**, **W.M. Wells and A.F. Frangi**, editors, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Springer International Publishing, Cham, pp.234–241.

[5] **Oktay, O.**, **Schlemper, J.**, **Folgoc, L.L.**, **Lee, M.**, **Heinrich, M.**, **Misawa, K.**, **Mori, K.**, **McDonagh, S.**, **Hammerla, N.Y.**, **Kainz, B.** *et al.* (2018). Attention u-net: Learning where to look for the pancreas, *arXiv preprint arXiv:1804.03999*.

[6] **Azad, R.**, **Asadi-Aghbolaghi, M.**, **Fathy, M. and Escalera, S.** (2019). Bi-Directional ConvLSTM U-Net with Densley Connected Convolutions, *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp.0–0.

[7] **He, K.**, **Zhang, X.**, **Ren, S. and Sun, J.** (2016). Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.770–778.

[8] **Chalfoun, J.**, **Majurski, M.**, **Peskin, A.**, **Breen, C.**, **Bajcsy, P. and Brady, M.** (2015). Empirical gradient threshold technique for automated segmentation across image modalities and cell lines, *Journal of microscopy*, *260*(1), 86–99.

[9] **Ayanzadeh, A.**, **Yağar, H.**, **Özuysal, **, **Okvur, D.P.**, **Töreyin, B.U.**, **Ünay, D. and Önal, S.** (2019). Cell Segmentation of 2D Phase-Contrast Microscopy Images with Deep Learning Method, *2019 Medical Technologies Congress (TIPTEKNO)*, pp.1–4.

[10] **Jaccard, N.**, **Griffin, L.D.**, **Keser, A.**, **Macown, R.J.**, **Super, A.**, **Veraitch, F.S. and Szita, N.** (2014). Automated method for the rapid and precise estimation of adherent cell culture characteristics from phase contrast microscopy images, *Biotechnology and bioengineering*, *111*(3), 504–517.

[11] **Bamford, P. and Lovell, B.** (1998). Unsupervised cell nucleus segmentation with active contours, *Signal processing*, *71*(2), 203–213.

[12] **Vincent, L. and Soille, P.** (1991). Watersheds in digital spaces: an efficient algorithm based on immersion simulations, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (6), 583–598.

[13] **Chen, L.**, **Zhang, J.**, **Chen, S.**, **Lin, Y.**, **Yao, C. and Zhang, J.** (2014). Hierarchical mergence approach to cell detection in phase contrast microscopy images, *Computational and mathematical methods in medicine*, *2014*.

[14] **Vuola, A.O.**, **Akram, S.U. and Kannala, J.** (2019). Mask-RCNN and U-net ensembled for nuclei segmentation, *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, IEEE, pp.208–212.

[15] **Long, F.** (2020). Microscopy cell nuclei segmentation with enhanced U-Net, *BMC bioinformatics*, *21*(1), 1–12.

[16] **Falk, T.**, **Mai, D.**, **Bensch, R.**, **Çiçek, Ö.**, **Abdulkadir, A.**, **Marrakchi, Y.**, **Böhm, A.**, **Deubner, J.**, **Jäckel, Z.**, **Seiwald, K.** *et al.* (2019). U-Net: deep learning for cell counting, detection, and morphometry, *Nature methods*, *16*(1), 67.

[17] **Tsai, H.F.**, **Gajda, J.**, **Sloan, T.F.**, **Rares, A. and Shen, A.Q.** (2019). Usiigaci: Instance-aware cell tracking in stain-free phase contrast microscopy enabled by machine learning, *bioRxiv*, 524041.

[18] **Arbelle, A. and Raviv, T.R.** (2018). Microscopy cell segmentation via adversarial neural networks, *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, IEEE, pp.645–648.

[19] **Cohen, A.A.**, **Geva-Zatorsky, N.**, **Eden, E.**, **Frenkel-Morgenstern, M.**, **Issaeva, I.**, **Sigal, A.**, **Milo, R.**, **Cohen-Saidon, C.**, **Liron, Y.**, **Kam, Z.** *et al.* (2008). Dynamic proteomics of individual cancer cells in response to a drug, *science*, *322*(5907), 1511–1516.

[20] **Payer, C.**, **Štern, D.**, **Neff, T.**, **Bischof, H. and Urschler, M.** (2018). Instance segmentation and tracking with cosine embeddings and recurrent hourglass networks, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, pp.3–11.

[21] **Yi, J.**, **Wu, P.**, **Huang, Q.**, **Qu, H.**, **Liu, B.**, **Hoeppner, D.J. and Metaxas, D.N.** (2019). Multi-scale Cell Instance Segmentation with Keypoint Graph Based Bounding Boxes, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, pp.369–377.

[22] **Arbelle, A. and Raviv, T.R.** (2019). Microscopy cell segmentation via convolutional LSTM networks, *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, IEEE, pp.1008–1012.

[23] **Ulman, V.**, **Maška, M.**, **Magnusson, K.E.**, **Ronneberger, O.**, **Haubold, C.**, **Harder, N.**, **Matula, P.**, **Matula, P.**, **Svoboda, D.**, **Radojevic, M.** *et al.* (2017). An objective comparison of cell-tracking algorithms, *Nature methods*, *14*(12), 1141.

[24] **Russakovsky, O.**, **Deng, J.**, **Su, H.**, **Krause, J.**, **Satheesh, S.**, **Ma, S.**, **Huang, Z.**, **Karpathy, A.**, **Khosla, A.**, **Bernstein, M.** *et al.* (2015). Imagenet large scale visual recognition challenge, *International journal of computer vision*, *115*(3), 211–252.

[25] **Caicedo, J.C.**, **Goodman, A.**, **Karhohs, K.W.**, **Cimini, B.A.**, **Ackerman, J.**, **Haghighi, M.**, **Heng, C.**, **Becker, T.**, **Doan, M.**, **McQuin, C.** *et al.* (2019). Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl, *Nature methods*, *16*(12), 1247–1253.

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

[27] **Wu, Y. and He, K.** (2018). Group normalization, *Proceedings of the European Conference on Computer Vision (ECCV)*, pp.3–19.

[28] **Miyato, T.**, **Kataoka, T.**, **Koyama, M. and Yoshida, Y.** (2018). Spectral normalization for generative adversarial networks, *arXiv preprint arXiv:1802.05957*.

[29] **Ulyanov, D.**, **Vedaldi, A. and Lempitsky, V.** (2016). Instance normalization: The missing ingredient for fast stylization, *arXiv preprint arXiv:1607.08022*.

[30] **Szegedy, C.**, **Liu, W.**, **Jia, Y.**, **Sermanet, P.**, **Reed, S.**, **Anguelov, D.**, **Erhan, D.**, **Vanhoucke, V. and Rabinovich, A.** (2015). Going deeper with convolutions, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.1–9.

[31] **Zhou, Z.**, **Siddiquee, M.M.R.**, **Tajbakhsh, N. and Liang, J.**, (2018). Unet++: A nested u-net architecture for medical image segmentation, Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer, pp.3–11.

[32] **Kohavi, R.** *et al.* (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection, *Ijcai*, volume 14, Montreal, Canada, pp.1137–1145.

[33] **McGuinness, K. and O'connor, N.E.** (2010). A comparative evaluation of interactive segmentation algorithms, *Pattern Recognition*, *43*(2), 434–444.

[34] **Schneider, C.A.**, **Rasband, W.S. and Eliceiri, K.W.** (2012). NIH Image to ImageJ: 25 years of image analysis, *Nature methods*, *9*(7), 671–675.

[35] **Schindelin, J.**, **Arganda-Carreras, I.**, **Frise, E.**, **Kaynig, V.**, **Longair, M.**, **Pietzsch, T.**, **Preibisch, S.**, **Rueden, C.**, **Saalfeld, S.**, **Schmid, B.** *et al.* (2012). Fiji: an open-source platform for biological-image analysis, *Nature methods*, *9*(7), 676–682.

[36] **Seferbekov, S.S.**, **Iglovikov, V.**, **Buslaev, A. and Shvets, A.** (2018). Feature Pyramid Network for Multi-Class Land Segmentation., *CVPR Workshops*, pp.272–275.

[37] **Wang, G.**, **Li, W.**, **Aertsen, M.**, **Deprest, J.**, **Ourselin, S. and Vercauteren, T.** (2018). Test-time augmentation with uncertainty estimation for deep learning-based medical image segmentation.

[38] **Kingma, D.P. and Ba, J.** (2014). Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*.

[39] **Chollet, F.**, (2015), keras, https://github.com/fchollet/keras.

[40] **Binici, R.C.**, **Şahin, U.**, **Ayanzadeh, A.**, **Töreyin, B.U.**, **Önal, S.**, **Okvur, D.P.**, **Özuysal, Ö.Y. and Ünay, D.** (2019). Automated Segmentation of Cells in Phase Contrast Optical Microscopy Time Series Images, *2019 Medical Technologies Congress (TIPTEKNO)*, IEEE, pp.1–4.

APPENDICES

APPENDIX A.1 : Experiment Details

Table A.1 : The detail structure of the applied method on discussed dataset.

Configuration	FPN	ResNet18UNet
Activation(last layer)	Sigmoid	Sigmoid
Input shape	518×518	518×518
Encoder weights	ImageNet	ImageNet
Decoder block	(256,128,64,32)	(256,128,64,32)
Decoder type	Upsampling 2D	bi-linear interpolation
block filter(pyramid)	128	—
Encoder Freeze	False	False
pyramid Aggregation	Concat	—-
pyramid dropout	0.5	—-

Table A.2 : Quantitative results of the different architecture which applied on the MDA-MB-231 dataset. TipNet, ResNet18-UNet, ResUNet-Respath is our suggested models for k=1.

Method	JI(%)	Dice(%)	Precision(%)	Recall(%)
U-Net+TTA	83.51	91.20	91.98	90.08
BCDUNet+TTA [6]	81.68	89.91	89.58	90.25
Tip-Net+TTA [9]	82.50	90.41	91.56	89.28
ResNet18-U-Net+Rp+TTA	86.09	92.52	94.02	91.08
ResNet18-UNet+TTA	85.50	92.18	93.14	91.25
ResNet18-FPN+TTA	85.53	92.20	91.63	92.78
U-Net [4]	83.20	90.83	90.75	90.91
BCDUNet [6]	78.54	88.21	86.85	89.64
Tip-Net [9]	80.19	89.07	90.01	88.02
ResNet18-UNet	83.52	91.02	91.78	90.28
ResNet18-U-Net+Rp	85.39	92.12	93.0	91.25
ResNet18-FPN	82.96	90.68	90.02	91.37

APPENDIX A.1

1.1 Detailed Experiments

The detail experimental results of the models after applying the k-fold cross validation is shown in TableA.2 andA.3 respectively. Moreover, the detail configuration which is applied on ResNet18-UNet and ResNet18-FPN is also show in TableA.1.

Table A.3 : Quantitative results of the different architecture which applied on the MDA-MB-231 dataset. TipNet, ResNet18-UNet, ResUNet-Respath is our suggested models for k=3.

Method	JI(%)	Dice(%)	Precision(%)	Recall(%)
U-Net+TTA [4]	82.23	90.25	91.07	89.45
BCDUNet+TTA	81.06	89.54	89.02	90.07
Tip-Net+TTA	81.20	89.62	90.75	88.53
ResNet18-U-Net+TTA	86.67	92.91	93.87	91.98
ResNet18-U-Net+RP+TTA	86.02	92.48	93.11	91.87
ResNet18-FPN+TTA	83.61	91.07	90.08	92.17
U-Net [4]	81.15	89.59	90.07	89.12
BCDUNet [6]	79.10	83.33	88.41	88.25
Tip-Net [9]	79.12	88.34	88.79	87.91
ResNet18-UNet	84.10	91.36	92.22	90.53
ResNet18-U-Net+RP	85.02	91.90	92.81	91.02
ResNet18-FPN	81.86	90.02	88.25	91.88

CURRICULUM VITAE

Name Surname: Aydin Ayanzadeh

Place and Date of Birth: Tabriz, Iran - 04/09/1993

E-Mail: ayanzadeh17@itu.edu.tr

EDUCATION:

- **B.Sc.:** 2016, University of Tabriz, Department of Mathmatic, Computer Science
- **M.Sc.:** 2020, Istanbul Technical University, Department of Applied Informatics

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2019-2020 Research Fellowship in Vodafone FutureLab.

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- A. Ayanzadeh, O.Y. Ozuysal, D.P Okvur, S. Onal, D. Unay, B.U.Toreyin " Deep Learning based Segmentation Pipeline for Label-Free Phase-Contrast Microscopy Images", *THE 28th IEEE CONFERENCE ON SIGNAL PROCESSING AND COMMUNICATIONS APPLICATIONS, SIU2020, Gaziantep,Turkey*