

**REPUBLIC OF TURKEY
AKDENİZ UNIVERSITY**



KI67 LABELING IN GLIOMAS WITH ARTIFICIAL INTELLIGENCE

Caner SONGÜL

INSTITUTE OF NATURAL AND APPLIED SCIENCES

DEPARTMENT OF COMPUTER ENGINEERING

MASTER THESIS

MAY 2021

ANTALYA

**REPUBLIC OF TURKEY
AKDENİZ UNIVERSITY**



KI67 LABELING IN GLIOMAS WITH ARTIFICIAL INTELLIGENCE

Caner SONGÜL

INSTITUTE OF NATURAL AND APPLIED SCIENCES

DEPARTMENT OF COMPUTER ENGINEERING

MASTER THESIS

MAY 2021

ANTALYA

REPUBLIC OF TURKEY
AKDENİZ UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

KI67 LABELING IN GLIOMAS WITH ARTIFICIAL INTELLIGENCE

Caner SONGÜL

DEPARTMENT OF COMPUTER ENGINEERING

MASTER THESIS

This thesis unanimously accepted by the jury on 10/05/2021.

Assist. Prof. Dr.Hüseyin Gökhan AKÇAY

Assoc. Prof. Dr. Havva Serap TORU

Assist. Prof. Dr.Emre GÜNGÖR

ÖZET

YAPAY ZEKA İLE GLİOMLARDA KI67 İŞARETLEME

Caner SONGÜL

Yüksek Lisans Tezi, Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Hüseyin Gökhan AKÇAY

Mayıs 2021 ; 31 sayfa

Kanserin evresinin belirlenmesi, tedaviler için en önemli etmenlerden biridir. Hücre- sel çoğalma indeksi, kanserin evresini belirleyen faktörlerden biridir. Bu nedenle hücre- sel proliferasyon indeksinin belirlenmesi tedavi için hayati öneme sahiptir. Patologlar, hücre- sel çoğalmayı belirlemek için KI67 nükleer proteinlerinin prognostik özelliğini kullanır. KI67 nükleer protein uygulandığında yüksek çoğalma hızına sahip hücreler kahveren- gimsi bir renge sahip olurlar. Bu hücrelerin miktarı, proliferasyon indeksinin hesaplanma- sında kullanılır. Bu tez kapsamında hücrelerin KI67 pozitif ve negatif olarak etiketlenmesi için yapay zeka destekli görüntü işleme teknikleri kullanılmıştır.

Bu çalışmada Faster R-CNN obje tanıma modeli kullanılmıştır. Faster R-CNN obje tanıma modelinin başarı oranları, kullanılan CNN modeline bağlıdır. Günümüzde Faster R-CNN obje tanıma algoritması için kullanılabilecek birden çok CNN modeli bulunmak- tadır. Bu tez kapsamında ResNet50, ResNet18, VGG16, VGG19 CNN modelleri kullanıl- mıştır ve bu modellerin Faster R-CNN obje tanıma algoritmasına entegre edilerek KI67 hücre etiketlemedeki başarımları test edilmiştir.

Derin öğrenme teknikleri ile obje tespiti veri setine de bağlıdır. Veri seti görüntüleri, eğitim sürecinden önce etiketlenmelidir. Bu tezde iki farklı görüntü etiketleme tekniği ile veri seti oluşturularak tam denetimli ve yarı denetimli öğrenme yöntemleri uygulan- mıştır. Tam denetimli öğrenme yönteminde tüm görüntüler manuel olarak etiketlenmiştir. Yarı denetimli öğrenme yönteminde ise varolan obje tanıma modeli etiketlemede kullanıl- mıştır. Yarı denetimli öğrenme yöntemi Faster R-CNN obje tanıma algoritmasına entegre edilerek daha başarılı sonuçlar alınmıştır. Bu sayede en uygun şekilde KI67 nükleer eksp- rasyonu gösteren gliom hücrelerinin tanınması sağlanmıştır.

ANAHTAR KELİMELEER: Derin Öğrenme, Biyomedikal, Görüntü İşleme, Nesne Tanıma, Yapay Sinir Ağları, Dijital Patoloji.

JÜRİ: Dr. Öğr. Üyesi Hüseyin Gökhan AKÇAY

Doç. Dr. Havva Serap TORU

Dr. Öğr. Üyesi Emre GÜNGÖR



ABSTRACT

KI67 LABELING GLIOMAS WITH ARTIFICIAL INTELLIGENCE

Caner SONGÜL

MSc Thesis in Computer Engineering

Supervisor: Assist. Prof. Dr.Hüseyin Gökhan AKÇAY

May 2021; 31 pages

Determining the stage of cancer is one of the essential steps for treatments. The cellular proliferation index is one of the factors determining the stage of cancer. Therefore, determining the cellular proliferation index is a vital part of treatment. Pathologists use KI67 nuclear proteins' prognostic feature to determine cellular proliferation. High proliferation rate cells have a brownish color when KI67 nuclear protein is applied. Therefore, the quantity of these cells is a crucial part of calculating the proliferation index. In the scope of this thesis, image processing with artificial intelligence techniques is used for labeling cells as KI67 positive and negative.

In this study, the Faster R-CNN algorithm was selected for object detection. Faster R-CNN object detectors' accuracy depends on CNN-based feature extractor model. Today, many different CNN models are available for this purpose. In this study, ResNet50, ResNet18, VGG16, VGG19 CNN models were used for feature extractors, and their KI67 cell-labeling performance was tested.

Object detection with deep learning techniques depends on the dataset. Dataset images should be labeled before the training process. In this study, two dataset image labeling techniques were implemented: Supervised and Semi-Supervised learning methods. For the Supervised learning method, all dataset images were labeled manually. For the Semi-Supervised learning method dataset, images were labeled with a previously trained Faster R-CNN object detector. Integrating the Semi-Supervised learning system into the Faster R-CNN object detection model, the object-labeling process has achieved more accurate results. In that way, nuclear KI67 expressing glioma cells and non-expressing cells were identified most appropriately.

KEYWORDS: Deep Learning, Biomedical Image Processing, Object Detection, Neural Networks, Digital Pathology

COMMITTEE: Assist. Prof. Dr. Hüseyin Gökhan AKÇAY

Assoc. Prof. Dr. Havva Serap TORU

Assist. Prof. Dr. Emre GÜNGÖR



ACKNOWLEDGEMENTS

I would like to thank my supervisors Assist. Prof. Dr. Hüseyin Gökhan AKÇAY and Assoc. Prof. Dr. Havva Serap TORU who supervised, guided and motivated me during this study. I would like to thank our chairman Prof. Dr. Melih GÜNAY for valuable contribution to my academic knowledge. I am also grateful to my dear colleagues and friends Melih ÖZ and Erdiñ TÜRÖ for their support during this study. Finally, I would like to thank my family for supporting me throughout my education life.



LIST OF CONTENTS

ÖZET	i
ABSTRACT	iii
ACKNOWLEDGEMENTS	v
TEXT OF OATH	viii
SYMBOLS AND ABBREVIATIONS	ix
LIST OF FIGURES	x
LIST OF TABLES	xi
1. INTRODUCTION	1
2. LITERATURE REVIEW	3
2.1. KI67 Marker	3
2.2. Convolutional Neural Network	3
2.3. Hand-Crafted Feature Extraction	4
2.4. CNN based Feature Extraction Methods	5
2.5. Object Detection in Medical Area	6
3. MATERIAL AND METHOD	8
3.1. Dataset	8
3.2. Pre-Processing	8
3.2.1. Image Resizing	9
3.2.2. Image Dividing	9
3.2.3. Image Labeling	9
3.2.4. Supervised Image Labeling	10
3.2.5. Semi-Supervised Image Labeling	10
3.3. Faster R-CNN	12
3.4. Convolutional Neural Network Models For Faster R-CNN	14
3.4.1. ResNet50	14
3.4.2. ResNet18	15
3.4.3. VGG16	16
3.4.4. VGG19	17
3.5. Anchor Boxes	17
3.6. Augmentation	18
3.7. Evaluation	20
4. RESULTS AND DISCUSSION	21
4.1. Batch Size	21

4.2. Anchor Numbers	22
4.3. Different Network Models	22
4.4. Supervised Learning and Semi-Supervised Learning	24
5. CONCLUSION.	28
6. REFERENCES	29
CURRICULUM VITAE	



TEXT OF OATH

I declare that this study "KI67 Labeling Gliomas with Artificial Intelligence", which I present as master thesis, is in accordance with the academic rules and ethical conduct. I also declare that i cited and referenced all material and results that are not original to this work.

10/05/2021

Caner SONGÜL



ABBREVIATIONS

CNN	: Convolutional Neural Network
ResNet	: Residual Neural Network
VGG	: Visual Geometry Group
mIoU	: Mean Intersection Over Union
VGG	: Visual Geometry Group
R-CNN	: Region Based Convolutional Neural Networks
Fast-RCNN	: Fast Region Based Convolutional Neural Networks
Faster-RCNN	:Faster Region Based Convolutional Neural Networks
GPU	:Graphics Processing Unit
CPU	:Central Processing Unit
RAM	:Random Access Memory
ReLU	:Rectified Linear Unit

LIST OF FIGURES

Figure 3.1.	Sample of Two Datasets	8
Figure 3.2.	Divided Dataset Image	10
Figure 3.3.	Supervised Image Labeling	11
Figure 3.4.	Semi-Supervised Image Labeling	13
Figure 3.5.	Faster R-CNN Architecture	13
Figure 3.6.	Region Proposal Network	14
Figure 3.7.	a)ResNet18 with Faster R-CNN b)ResNet50 with Faster R-CNN	16
Figure 3.8.	a)VGG16 with Faster R-CNN b)VGG19 with Faster R-CNN . .	18
Figure 3.9.	Ground Truth of Dataset Aspect Ratio Pixel Size Distrubution . .	19
Figure 3.10.	AnchorBoxes Number Effect on Mean IoU	19
Figure 3.11.	Augmented Data	20
Figure 4.12.	VGG16 with Faster R-CNN a) Supervised b) Semi-Supervised .	26
Figure 4.13.	VGG19 with Faster R-CNN a) Supervised b) Semi-Supervised .	27
Figure 4.14.	ResNet18 with Faster R-CNN a) Supervised b) Semi-Supervised	27
Figure 4.15.	ResNet50 with Faster R-CNN a) Supervised b) Semi-Supervised	27

LIST OF TABLES

Table 4.1.	Batch Size Results	22
Table 4.2.	Anchor Box Number	23
Table 4.3.	Different Network Model Comparison	24
Table 4.4.	Supervised and Semi Supervised Learning Comparison	25



1. INTRODUCTION

Cancer is a leading cause of death(Twombly et al. 2005). Many people around the world suffer from that disease and their life expectancy depends on treatment and diagnosis time (Syriopoulou et al. 2017). Many of the cancer cells are growing rapidly and patient's treatment can be changed during the treatment time. In the medical area, pathologists focus on these subjects to determine the growth rate of these malignant cells. KI67 is an excellent marker to determine the growth fraction of a given cell population. The nuclear protein KI67 (pKI67) is an established prognostic and predictive indicator. The fraction of KI67-positive tumor cells (the KI67 labeling index) is often correlated with the clinical course of cancer(Kontzoglou et al. 2013). KI67 is an important prognostic and predictive marker for many cancers. The KI67 Labelling Index is used as an ancillary tool in glioma diagnostics.

One of the main problems in the medical area is many diagnostics are done manually, so its diagnostics become time-consuming and interobserver-intraobserver variability has been reported and no precise guidelines are available. Using novel digital approaches would be a promising technique. For pathologists, KI67 labeling is important for the brain tumor process. However, many times pathologists need to scan many areas to find the hot point of KI67 expression and count the KI67 expressing tumor cells. After that pathologist must count every tumor cell and normal cells manually in that area to decide the patient's condition. For this problem, automated computer-aided KI67 labeling model was proposed to support pathologists during the diagnostic process. Object detection methods are very suitable for this problem so deep learning based object detection method is used for cell labeling.

Nowadays, progress in artificial intelligence and deep learning algorithms give better solutions for object detection. In addition, the rise of the Region-Based Convolutional Neural Network has created a new era for object detection.

One of the usage areas of object detection with deep learning methods is in the medical area. For example, Faster R-CNN object detection methods are used for blood cell detection(Xia et al. 2019).

Deep learning methods highly depend on training data quantity and quality, and these

should be high enough to give satisfying results for object detection. So pathologists should label every cell on microscopic images before training the Faster R-CNN object detection method. Because these problems remain in the object detection area, a newer approach should solve these problems.

In this study, the Semi-Supervised learning method is adapted to the Faster R-CNN object detector method to lower the time-consuming image labeling process. KI67 nuclear protein marker applied tissue images are taken from 40X magnified microscopy. Only ten percent of dataset images were used, and brown stained cells labeled as KI67 positive and non-stained cells were labeled as negatives. A Faster R-CNN system was used for training this dataset. The remaining images were applied to the Faster R-CNN object detector. After that, the detection results of these images were automatically labeled to generate a more extensive dataset for another training process. Generating a bigger dataset for the training process gives opportunity to learn from different cells and gives better results for KI67 labeling in gliomas.

2. LITERATURE REVIEW

2.1. KI67 Marker

Antigen KI67 is a nuclear protein that is associated with cellular proliferation. KI67 protein is present during all active phases of the cell cycle (G1, S, G2, and mitosis) but is absent in resting (quiescent) cells (G0). The cellular content of KI67 protein markedly increases during cell progression through the S phase of the cell cycle. Thus it is used for determining the proliferation index of tumor cells. One of the essential things is KI67 protein stains the proliferating cells, and cells become brownish to differentiate from the other cells (Li et al. 2015). KI67 primarily stains tumor cells which are highly associated with proliferation and growth. It is a prevalent tool in the pathology area to detect malign cells and many pathologists see the KI67 proliferation marker. KI67 marker is applied on biopsies of suspicious tissue. Also, cells with KI67 marker give strong evidence for metastasis and stage of the tumors. The prognosis feature of the KI67 marker can be used for breast, cervix, lung, brain tissue. Glioma is a type of tumor that occurs in the brain and spinal cord. Three types of glial cells can produce tumors. Gliomas are classified according to the type of glial cell involved in the tumor and the tumor's genetic features, which can help predict how the tumor will behave over time and the treatments most likely to work McKinsey et al. 2012. KI67 labeling is used for grading gliomas which is an essential step for the clinical approach.

2.2. Convolutional Neural Network

One of the first attempts on Convolutional Neural Networks was about gradient-based learning applied to document recognition (LeCun et al. 1998). In that study, a gradient-based learning method based on a backpropagation algorithm. According to the study, the gradient descent algorithm is used for minimizing loss on multilayer neural networks. This process made neural networks more successful for classification. Thus, LeNet-5 architecture is created. As a dataset, the system used 32×32 pixels images which were MNIST datasets, to classify hand-written characters.

Until the 2000s, research on the CNN area was limited. After (K. Chellapilla et al. 2006) make an implementation CNN network to run on GPUs, CNN models run much

faster than the model that runs on CPUs. In their paper, they got four times less processing time than older CPUs-based methods. After that development in the CNN area, it allows to develop more complex deep neural networks. Around 2010s, CNNs have become the main tools for image recognition as the paper issued by Ciregan et al. 2012 . CNN models became deeper by containing more layers, and the term of deep learning is derived from these improvements. Also on Krizhevsky et al. 2012 proposed a network called Alexnet for the classification problem on the large dataset, which is called ImageNet. This model won image recognition content, and this event shows that Convolutional Neural Networks are a fundamental part of the computer vision area. Because of less pre-processing, sufficient accuracy, and improvement of computer hardware leads researchers began to use CNNs for computer vision area.

2.3. Hand-Crafted Feature Extraction

Before the improvement of deep learning methods, hand crafted feature extraction techniques were used for object detection problems. Examples about Hand-Crafted feature extraction methods given below(O'Mahony et al. 2019):

- Scale Invariant Feature Transform (SIFT)
- Speeded Up Robust Features (SURF)
- Features from Accelerated Segment Test (FAST)
- Histogram Oriented Gradients (HOG)
- Geometric hashing

Hand-Crafted feature extraction methods require less processing power for object detection than deep learning based object detection methods. Histogram oriented Gradients (HOG) feature extraction method is used for human detection (Zhu. et. al. 2006). In that work Histogram Oriented Gradients (HOG) and Support Vector Machine (SVM) feature extraction methods are used for human detection on the INRIA data set.

2.4. CNN based Feature Extraction Methods

Recent advances in technology allow doing complex tasks on computers. Therefore, deep learning methods on object detection area rise. First implementations of deep neural networks are generally used for classification. On the other hand, object detection needs a localization parameter, so modification of deep neural networks is needed for object detection problems. Deep neural networks, the major success in object detection area is feature extraction. Convolution operations do feature extractions, and using different convolution operations gives feature maps that contain valuable information for object detection problems. Today different object detection methods become available researchers, and many different deep neural networks can be implemented to solve object detection problems (Goodfellow et al. 2016).

One of the most popular object detection methods is region-based convolutional neural networks (R-CNN) family-type object detection. The first research about R-CNN was published by Girshick et al. 2014, became a research milestone for object detection with deep neural networks. In that work, region proposals are created with a selective search algorithm instead of using only sliding windows, and nearly 2000 region proposals are created with the selective search algorithm. Then region proposals are fed into the CNN network, which in that research Alexnet was selected for backbone CNN network for R-CNN. After every region proposals feature, a vector is created with the AlexNet CNN module. After that, these region proposals feature vectors are classified with SVM to determine which class they are related to (Noble et al. 2006). On the 200-class ILSVRC2013 detection dataset, R-CNN's mAP is 31.4%, a significant improvement over OverFeat, which had the previous best result at 24.3% (Russakovsky et al. 2015).

Although the R-CNN object detection method achieved many object detection goals, the algorithm was still slow for training and testing. Solutions for these drawbacks of R-CNN object detection Fast R-CNN method is proposed (Girshick and Ross et al. 2015). The Fast R-CNN method has significant improvements in training time, testing time, and accuracy. On official paper Fast R-CNN with VGG16 network data training nine times faster than R-CNN with VGG16 network. Object detection time is 213 times faster than R-CNN. Also, the Fast R-CNN method lowers the feature cache, so a large

storage medium is not needed. This advancement in speed was achieved with some changes to the algorithm. On the R-CNN algorithm, approximately 2000 region proposals are done. These proposed regions are fed into CNN, but in the Faster R-CNN object detection algorithm, dataset images first go into CNN, then extracted feature maps are used for region proposals. Again selective search algorithm was used to find the region proposals. Each region proposal is made RoI operation then RoI vector created with Fully Connected Layer. These vectors are used for two processes. First, to find out which class is object-related, another usage is to find bounding boxes. One of the other differences is on the R-CNN algorithm class prediction done by SVM, but the Fast R-CNN method class prediction is made by the Softmax classifier. In the paper, the Softmax classifier gives a small amount of accuracy gain on object detection. Another important update for R-CNN based object detector family is the Faster R-CNN algorithm. Faster R-CNN greatly improved training and test time. The Faster R-CNN algorithm also uses the CNN backbone for feature extraction. However, this time instead of using a selective search algorithm to find region proposals, a region proposal network is used (Ren et al. 2016). Therefore, that way, the system learns region proposals.

2.5. Object Detection in Medical Area

Before deep learning algorithms, there were implementations of traditional object detection methods in the medical area. Traditional computer vision methods are used for detecting abnormal cervical cells (Sophea et al. 2018) . Feature extraction is done by the Histogram of Oriented Gradients method. Cell detection and classification are done by Support Vector Machines(SVM). The proposed method is tested with the Harlev dataset. Despite using only traditional methods, their proposed method reach 94.7% accuracy for detecting normal and abnormal cells. Although getting a high accuracy score, the dataset used in that method contains a single cell per image. Thus, using these methods for more complex detection problems that contain multiple objects for detection can lead to less accurate results.

Other examples of object detection in the medical area contain deep neural network implementation. One of the examples of tumor cells being used is CNN and Single Shot multi-box detector for detecting KI67 stained breast cancer cells. For feature extraction,

the VGG net is used. For detecting cells SSD method proposed by Zhang et al. 2018, is used in their work CNN, and SSD is used for KI67 stained breast cancer cells to detecting affected cells. Also, for blood cell detection, work has been done by the Faster R-CNN algorithm to detect live White Blood Cells (Xia et al. 2019).



3. MATERIAL AND METHOD

3.1. Dataset

The image dataset of this study consists of 40 times magnified microscopy images of brain tissue. First, all tissue samples are stained with a KI67 protein marker. All visual data was provided by Akdeniz University School of Medicine, Department of Pathology. To acquire and visibly enlarged samples, Olympus BX53 light microscope was used for this study. Microscopic images were taken by light microscope and transferred to the digital platform using the "CellSens Entry" program. The first dataset consists of 25 patients, and all patients had three different images from brain tissue, a total of making 75 images. The second dataset contains 21 patients, and all patients again had three different images from brain tissue, a total of making 63 images. Hot point selecting was provided by pathologists conventionally for taking microscopic images for the data set. Using only one source of image dataset can lead to overfitting the model so, to overcome this problem, two different datasets were used for training the model. In that way, the dataset was prepared with different images with different brightness values.

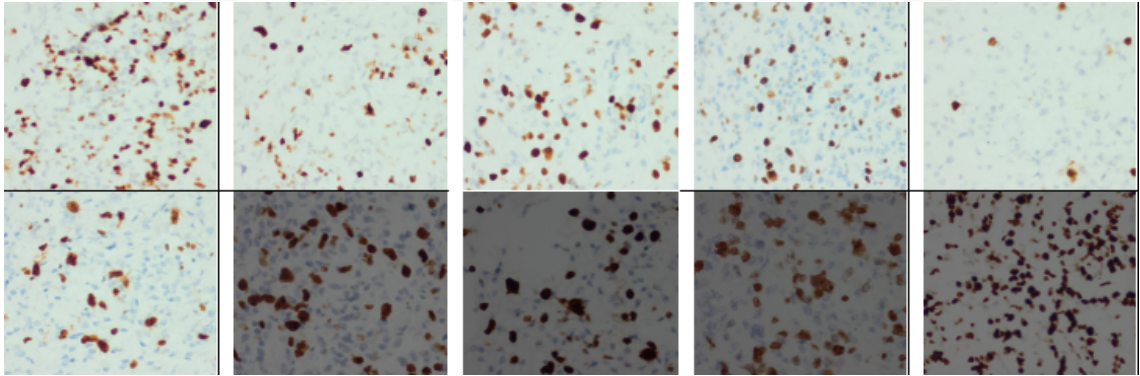


Figure 3.1. Sample of Two Datasets

3.2. Pre-Processing

In this study, two different datasets were used. Images in each dataset have diverse brightness contrast values. Also, these images have 2560×1920 resolution, which was too high for the training process. Therefore dataset images resolution should be lowered

to 1280×960 . By reducing image resolution and dividing sample images into equal sub-images, training time for neural network and GPU memory optimization was satisfied.

3.2.1. Image Resizing

Every image should have the same resolution before training and labeling and also for reducing computational time. High-resolution images, which in this case, 2560×1920 resolution, should be lowered in the resolution, which is not suitable for GPU memory. Thus every dataset images were first reduced to 1280×960 resolution. This resolution was selected because the aspect ratio of the images is 4:3. Thus, the aspect ratio and the cell characteristics do not change.

3.2.2. Image Dividing

Object detection method with deep neural networks is an intensive task. High resolution images can easily overflow the GPU memory. To cope with that problem, generally reducing the size of images to use for training the neural network but in this case, reducing resolution leads to another problem that labeled cell data on images becomes very small or even vanishes. Therefore optimizing dataset images with lowering the resolution is not an option for KI67 labeling on images. Instead of reducing image resolution, all images on the dataset will be divided according to the minimum accepted resolution for CNN networks. Many CNN networks used for this thesis accept 240×240 pixels or higher, so without changing the aspect ratio of 4:3, the best resolution for training these images will be 320×240 pixels. Every image from the dataset was divided, and every image had a resolution of 320×240 pixels. The single image was divided equally into 16 images. In that way, GPU memory optimization was satisfied without overflowing GPU memory during of training process.

3.2.3. Image Labeling

Image labeling one of the essential things in deep neural networks. Image labeling is used for creating ground truth for a deep neural network. Using the ground truth values convolution neural network is used for back-propagation to make estimations closer to ground truth values. In that way, weight and bias are adjusted on deep neural networks.

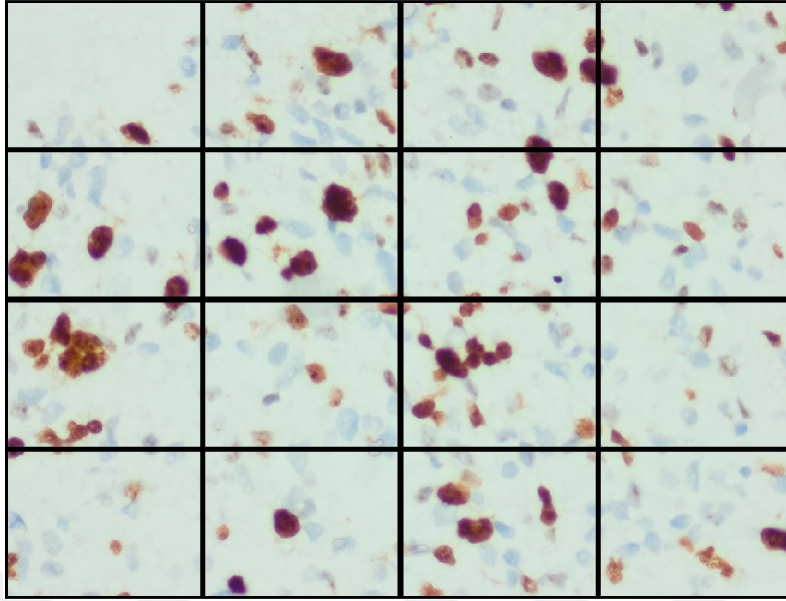


Figure 3.2. Divided Dataset Image

So using properly labeled data is a vital part of the deep neural network. In the thesis, two types of image labeling processes were used. Those were being supervised and semi-supervised labeling methods.

3.2.4. Supervised Image Labeling

Supervised Image labeling relies on a person manually annotates images for ground truth values. In this case, with a consultation with Akdeniz University School of Medicine, Department of Pathology, ten percent of the dataset images were annotated with MATLAB image labeling toolbox. Every image on the dataset KI67 labeled positive and negative cells are annotated with a rectangular annotator with x and y-axis coordinates. Even though manual segmentation is a good way to annotate images, it is time-consuming. Instead of annotating the whole dataset of images with supervised image labeling, only a sufficient part of 10 percent of the dataset was annotated with the supervised method. Flowchart of Supervised based learning method is shown in Fig 3.3.

3.2.5. Semi-Supervised Image Labeling

The Semi-Supervised learning method is used for different purposes. Creating labeled data for deep neural networks highly time-consuming job. Also, deep neural networks'

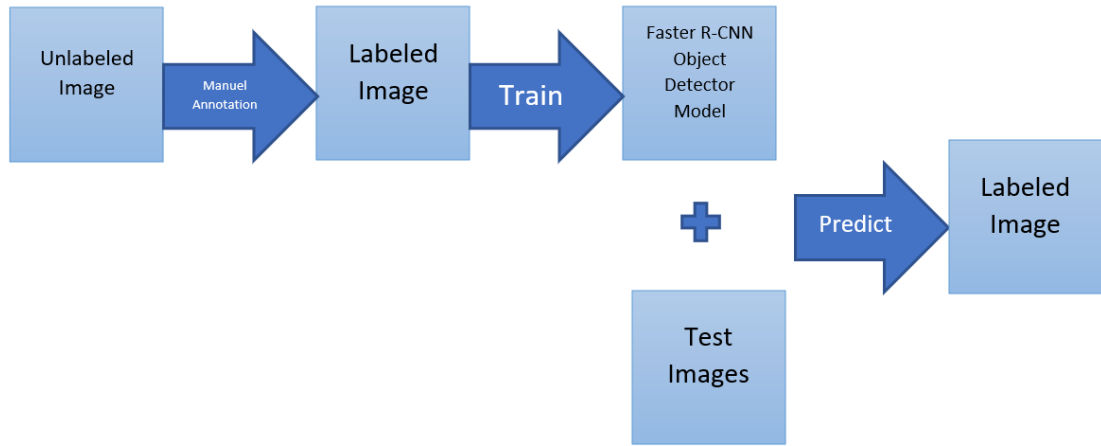


Figure 3.3. Supervised Image Labeling

success rate depends on the labeled dataset. To cope with that problem semi-supervised learning method is used for solutions to these problems. The semi-supervised learning method is used for three different situations. The first is for a labeling process where the supervisor cannot decide the classes of the objects. The second usage is for the inexact labeling process when the supervisor is the falsely classified object. The last usage is for semi-supervised learning is where the lower quantity of objects in the dataset are labeled by the supervisor, and other remaining data is remained unlabelled (Zhou et al. 2018). In this study, Semi-Supervised learning methods were applied used for the image labeling process. For this thesis, a dataset was incompletely labeled by the supervisor, so a Semi-Supervised learning technique is implemented for object detection problems. The Semi-supervised method for image labeling significantly increased efficiency. After the first training process was done with a supervised labeled image dataset, a new object detector was created. A newly created object detector was used for labeling objects on the remaining dataset images. Because of the low quantity of dataset images of the first training object detector, only some labeled images were used for the second training process. At that point, automated labeled images with high accuracy were selected and create a bigger dataset for another training process. Using the semi-supervised labeling method significantly reduced the time-consuming image labeling process and created a bigger labeled dataset. By using the semi-supervised learning method iteratively, more images with the high accurately labeled dataset are created. Therefore, two iterations of

the Semi-Supervised learning method were used for the image labeling process in this study. The first iteration of semi-supervised learning was done by an object detector created by the Supervised learning method. Object detector labels the unlabeled data, and newly labeled data was used for creating a better object detector. In this process, the previous object detector was used to continue the training process to get better results, but training parameters were changed for the new training process. Since new training images were different from the first training process, the number of epoch and learning rates were changed for better training. Learning rate is a parameter used for updating the network's weights, and therefore it should be tuned correctly for the Semi-Supervised learning process. The learning rate was selected 10^{-4} for the first training process to update the weights of the network quickly. However, for the second training process, the previously used detector was used to continue training. Using a high learning rate which was 10^{-4} in this case, was not fine-tune model, increases error rate, and even led to a lower success rate. Hence using a lower learning rate for the late training section is a better solution for neural networks. Every iteration with a Semi-Supervised learning rate is selected ten times lower than the previous training process (Bengio et al. 2012). After the training process was done, the first iteration of semi-supervised learning was finished. Then, for the second iteration for semi-supervised learning, the newly created detector was used for another image labeling process for the remaining dataset. Hence using semi-supervised iteratively make automated image labeling more accurate. The process of the Semi-Supervised learning method is shown in Fig 3.4.

3.3. Faster R-CNN

The Faster R-CNN object detection method is a robust method for object detection, especially for its lower training time and better object detection than R-CNN and Faster R-CNN object detection methods. For this study, the Faster R-CNN object detection method is a fundamental part. The Faster R-CNN method contains several neural network blocks. Faster R-CNN is a combination of Fast R-CNN as detector and Region Proposal Network. The first whole image is fed into CNN. After convolution operations on layers, image size is reduced to 15×20 resolution, and feature maps for input images are created. One of the crucial things using Faster R-CNN is selecting the right CNN model. That way, accurate

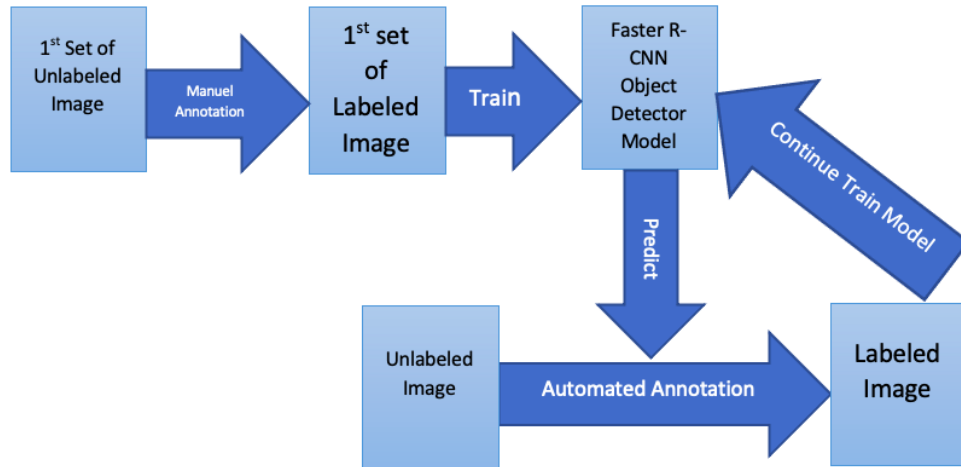


Figure 3.4. Semi-Supervised Image Labeling

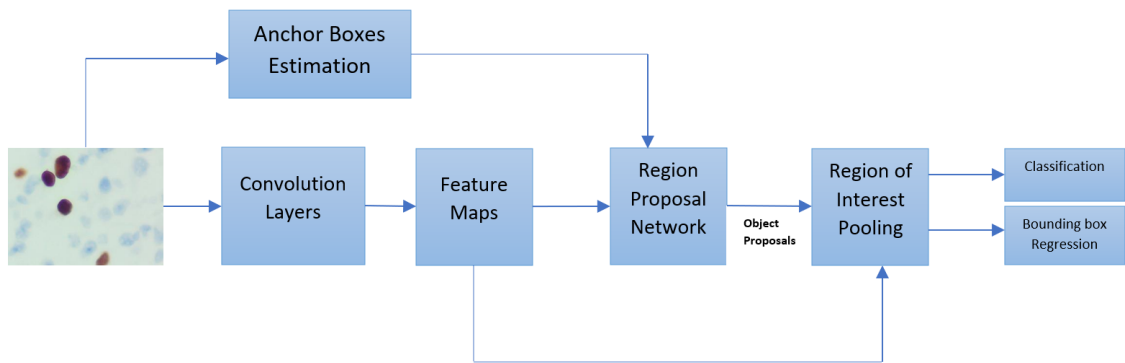


Figure 3.5. Faster R-CNN Architecture

feature maps are created. Feature maps are used for different purposes. One of them is creating region proposals.

Region proposal network is itself a neural network for region proposals. Region proposal network gets feature maps from CNN and does sliding windows operation on feature maps with anchor boxes. Anchor boxes are the set of rectangular windows that are determined in this study with ground truth labels of training data. Their aspect ratio and resolution important for sliding windows operation in the Region proposal network. Because the anchor boxes give estimated dimensions of the object, selecting the correct number of anchor boxes is essential. Different anchor boxes can overlap a single object, but only the

highest Intersection over Union value which overlaps most of the ground truth object, is selected. In this study, the IoU value must be higher than 0.5 value for a label classify as an object. After sliding windows operation with anchor boxes is completed, object proposals reconnect to the main network with the RoI pooling layer. After that operation, the network diverges two ways. One is for classification operation and regression operation. Regression operation is used to find bounding boxes coordinates, and it is done by smooth L1 loss regression operation. Classification operation is used for the decision of the class of the proposed image. For the Faster R-CNN method softmax classifier is selected for classification. Regression operation is estimating the coordinates of selected object coordinates. After that, objects are detected on the images(Ren et al. 2016) .

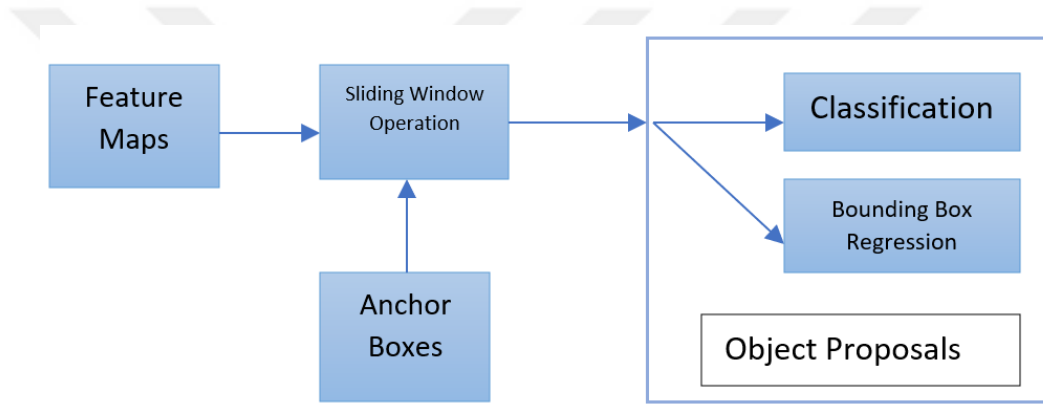


Figure 3.6. Region Proposal Network

3.4. Convolutional Neural Network Models For Faster R-CNN

In this section, the CNN models used for the thesis are inspected. Faster R-CNN object detection model highly relies on CNN for feature extraction. Different CNN models are reconstructed for implementation of Faster R-CNN.

3.4.1. ResNet50

ResNet50 contains 50 layers. It is a residual learning-based Convolution Neural Network (He et al. 2016). It differentiates from the other network is implementing residual learning. Typical ResNet models are implemented with double- or triple-layer skips that contain nonlinearities (ReLU) and batch normalization in between. ResNet50 convolutional network makes another connection by skipping layers to cope with the vanishing

gradients problem. Vanishing gradients problem occurs during the training process. According to error functions partial derivatives, the weight of the neural network updates, but if the value is too small, it is not changed and leads to the network not updating its weights. Therefore, ResNet50 has speed advantages over speed from the plain network. The speed advantage is achieved by reducing the importance of vanishing gradients. In that way, fewer layers are used for propagation. After that network restores skipped layers, then the network learns feature space. For classification, the ResNet50 model uses softmax classification, which gives probability scores for vector value. The softmax function maps the vector values to 0 to 1. The vector is classified to the highest probability class by the softmax classifier. In this study, ResNet50 was used for creating feature maps for Faster R-CNN. ResNet50 network's ActivationRelu40 layer was used for feature extraction and concatenates the network with Faster R-CNN, which this layer has $15 \times 20 \times 1024$ feature maps. Object detection was done by using Resnet50's $15 \times 20 \times 1024$ feature map. This node was selected for feature extraction because it feeds the object detector with a high enough feature map for a single image. However, using a small number of feature maps can lead to a low success rate because of limited feature quantity. Resnet50 with Faster R-CNN network structure is shown in Fig 3.7.

3.4.2. ResNet18

In this study, another residual learning-based convolutional network with smaller layers was inspected. ResNet18 has significantly lower layers which have 18 compared to ResNet50 networks 50 layers.

To test object detector performance and accuracy of object detector, ResNet18 was also used for this thesis. The main reason for testing ResNet18 is more shallow network can achieve similar results for object detection. Moreover, ResNet18 is used for feature extraction for the Faster R-CNN model. In this study, the res4bRelu layer was used as a feature extractor for Faster R-CNN. The res4bRelu layer has $15 \times 20 \times 256$ feature map, and its connection with Faster R-CNN is shown in Fig 3.7.

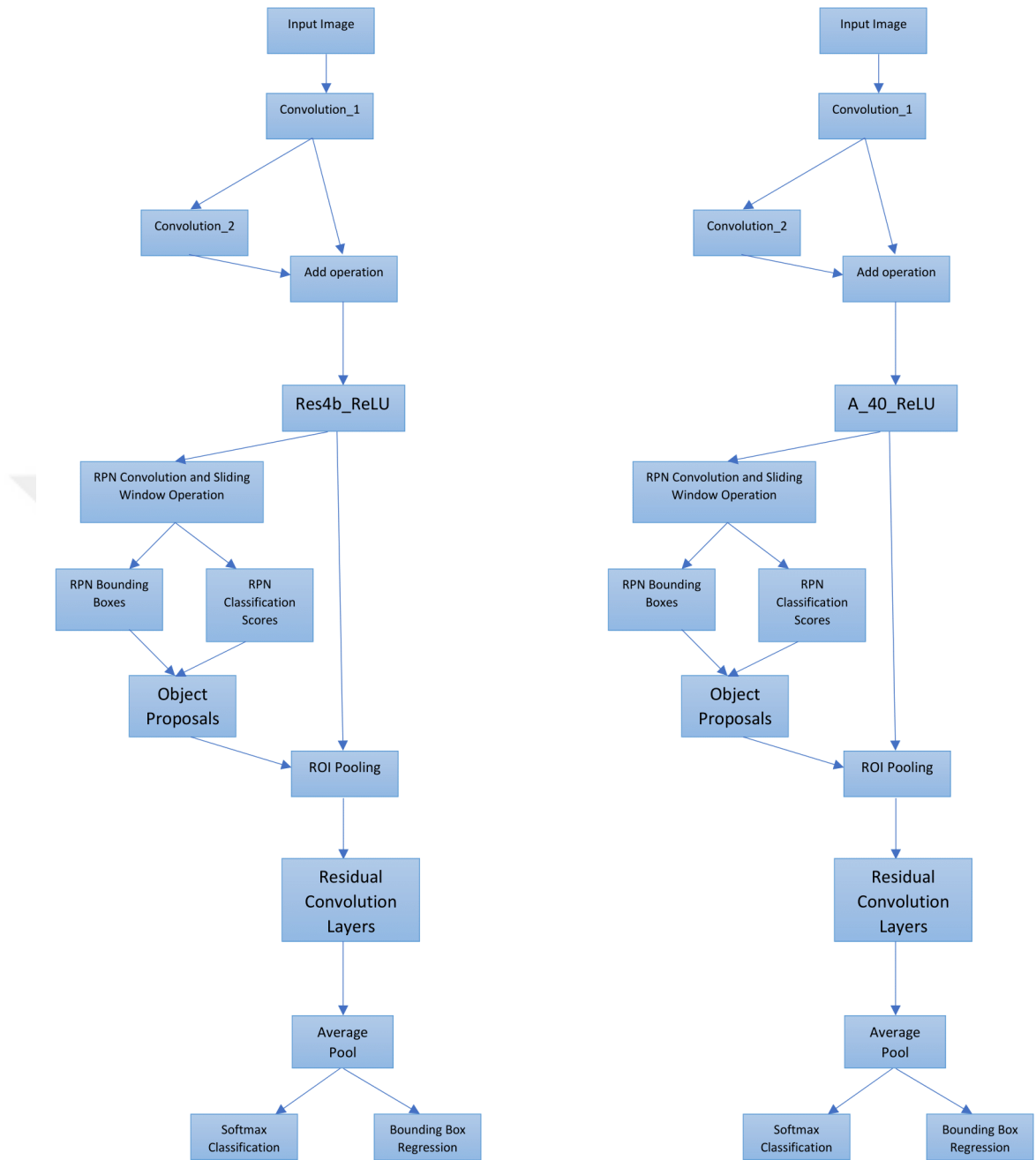


Figure 3.7. a) ResNet18 with Faster R-CNN b) ResNet50 with Faster R-CNN

3.4.3. VGG16

VGG16 contains 16 layers, and it proposed with a paper named Very Deep Convolutional Networks for Large-Scale Image Recognition. Differences from the residual convolutional network, there is no skip connection between layers as it is identified as plain network architecture. For this study, the VGG16 network was used as a feature ext-

ractor for the Faster R-CNN object detector algorithm. The relu5-3 layer was selected as a feature extraction layer for Faster R-CNN. Feature maps created with the VGG16 network are $15 \times 20 \times 512$. This feature map is used, and Faster R-CNN layers are connected network from this layer. Faster R-CNN with VGG16 connection is shown in the Fig 3.8.

3.4.4. VGG19

VGG19 contains 19 layers. VGG19 has a more deep layer for convolution operations. For this study, the VGG19 network was used as a feature extractor for the Faster R-CNN object detector algorithm. The relu5-4 layer was selected as a feature extraction layer for Faster R-CNN. Feature maps created with the VGG19 network are $15 \times 20 \times 512$. This feature map was used, and Faster R-CNN layers were connected network from this layer. Faster R-CNN with VGG19 connection is shown in Fig 3.8.

3.5. Anchor Boxes

Anchor boxes are an essential part of the Faster R-CNN method. Anchor boxes are a predefined set of windows with a fixed size of height and width values to do sliding windows operation on the Region Proposal Network. Anchor boxes are potential object locations, so that is why it is used in sliding windows operation. The number of anchor boxes dramatically affects the success of the Faster R-CNN object detection method. The small or large number of anchor boxes can lead to undesirable consequences. For this study, anchor boxes were defined by training datas' ground truth boxes. Training data objects' ground truth pixel areas and aspect ratios are shown in Fig 3.9.

K-means clustering algorithm is used to determine Anchor Boxes values. The number of the desired anchor boxes are given class number for the K-means clustering algorithm. It creates anchor box coordinates depending on the number of classes. Selecting the number of anchor boxes are another important work, and it depends on training datas' ground truth values. A low amount of anchor boxes leads to low intersection over union value, which means proposed anchor boxes do not overlap the real ground truth boxes. Increasing too much anchor box number leads to waste of computational power for slight increase for MeanIoU value (Redmond et al. 2007).

In this study, the graphic that is shown as Fig 3.10 is constructed with training dataset

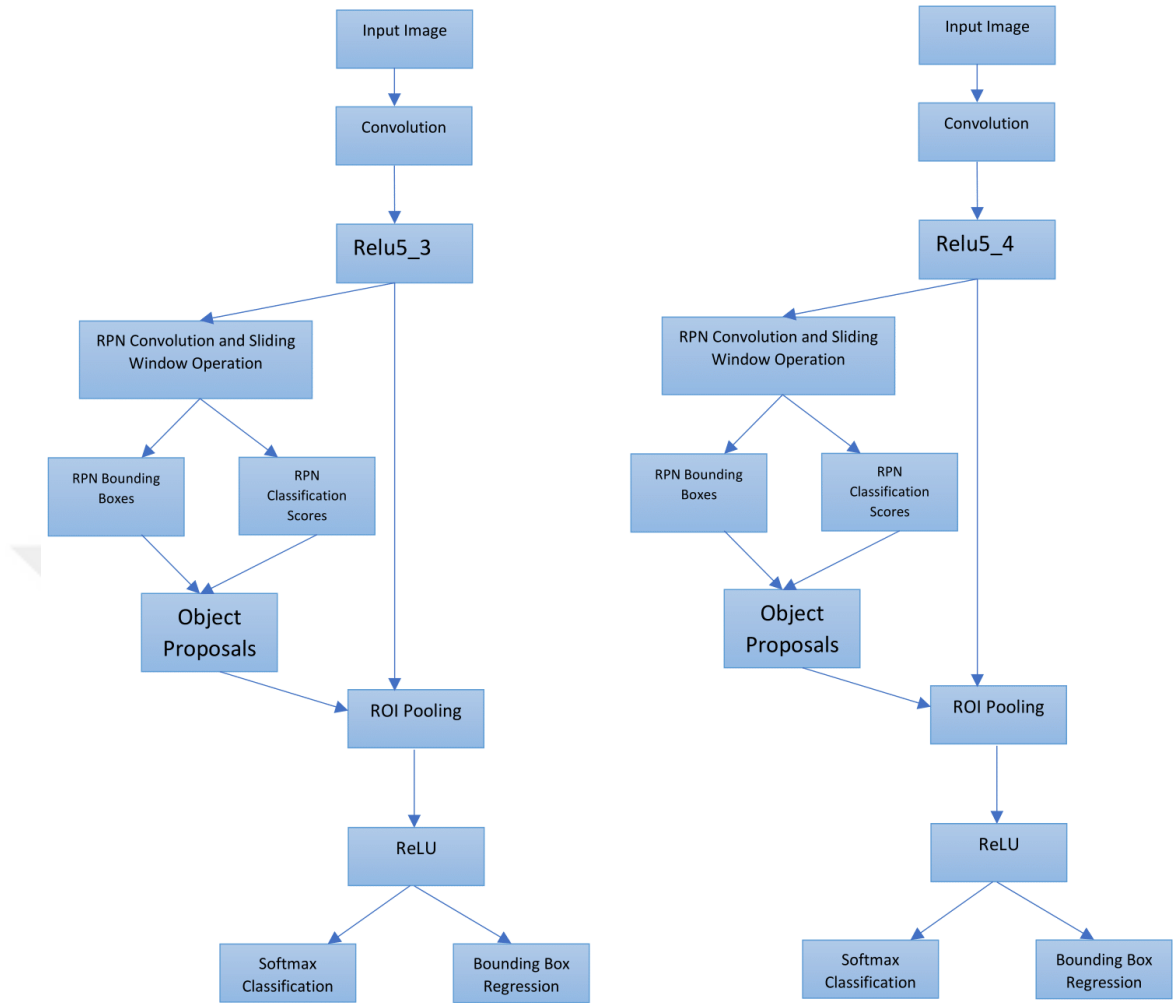


Figure 3.8. a)VGG16 with Faster R-CNN b)VGG19 with Faster R-CNN

images' ground truth values. Using only one anchor box leads to unsuccessful IoU with ground truth values. However, simply increasing the anchor box number also increases the IoU value. Increasing the number of anchor boxes more than five is not dramatically affect the mean IoU value. Because of that, the number of anchor boxes is selected between 2-5.

3.6. Augmentation

Data augmentation is used to increase the dataset and prevent overfitting the model. For this study, training images were mirrored Y-axis to create more images for the training dataset. Also, ground truth values were also rotated with the augmentation process.

Rotation is a common image augmentation technique and new location of the rotated pixel can be represented as:

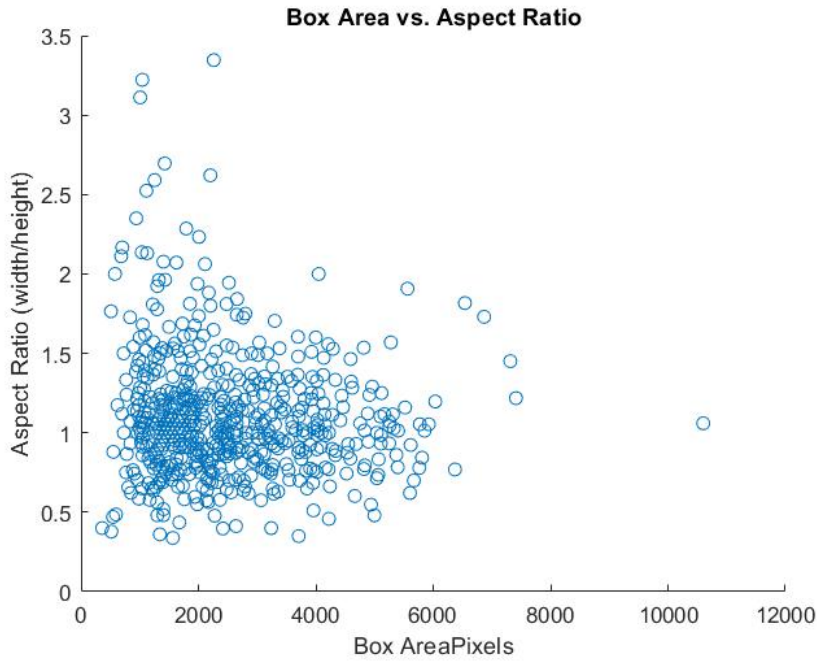


Figure 3.9. Ground Truth of Dataset Aspect Ratio Pixel Size Distrubution

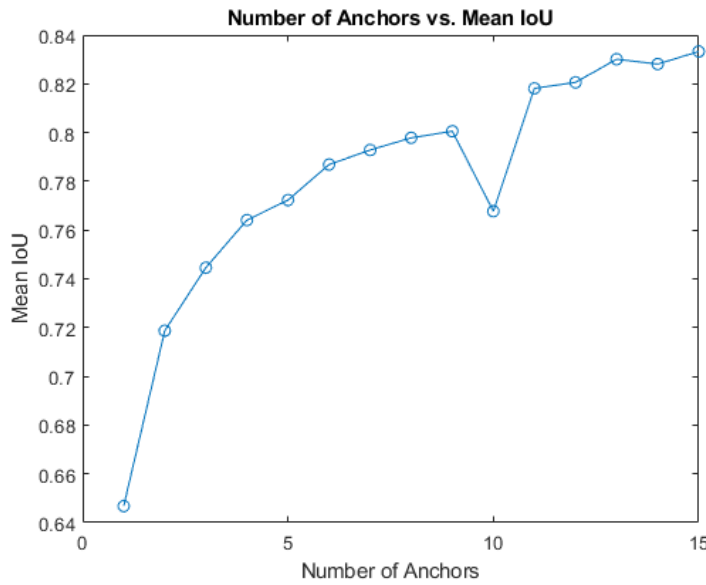


Figure 3.10. AnchorBoxes Number Effect on Mean IoU

$$x_n = \cos(\theta) \times (X_{n-1} - X_0) + \sin(\theta) \times (Y_{n-1} - Y_0) \quad (3.1)$$

$$Y_n = \sin(\theta) \times (X_{n-1} - X_0) + \cos(\theta) \times (Y_{n-1} - Y_0)$$

X_0 and Y_0 represents selected fixed point of rotation X_{n-1} and Y_{n-1} old location of the pixel value and X_n and Y_n new location of the pixel. Example of augmented dataset

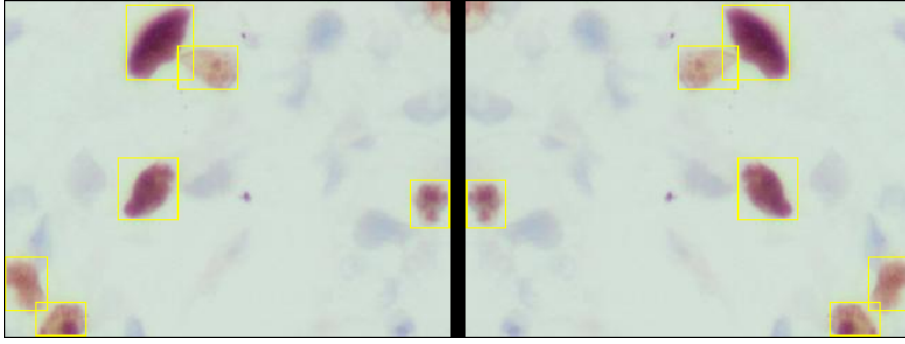


Figure 3.11. Augmented Data

image is shown in Fig 3.11.

3.7. Evaluation

For the evaluation of this thesis, several methods were used. Five evaluation criteria were used to measure the accuracy of constructed object detection model. These evaluation criteria are, Recall, Precision, Average Precision, and F1 scores were used. For evaluations, if detected object locations' IoU value is higher than 0.5 model classifies as successful labeling. If detected object locations' IoU value is lower than 0.5 model does not classify as an object. Precision is used to calculate to decide how many positive selected objects are positive objects. It is calculated with equation (3.2),

$$Precision = (TruePositive) / (TruePositive + FalsePositive) \quad (3.2)$$

The recall is used to measure how many real positives are selected as positive on the system. Recall is calculated with equation (3.3),

$$Recall = (TruePositive) / (TruePositive + FalseNegative) \quad (3.3)$$

F1 score is harmonic mean of precision and recall values and F1 score is calculated with equation (3.4), .

$$F1Score = 2 \times (Precision \times Recall) / (Precision + Recall) \quad (3.4)$$

4. RESULTS AND DISCUSSION

In this thesis, different CNN models with Faster R-CNN object detection models with different training parameters' performance were assessed. MATLAB 2020b program was used for experiments. All networks training process were performed on a computer equipped with Intel Core I5 9400F CPU, Nvidia RTX 2080 Super 8 GB GPU and 24 GB of RAM running on Windows 10 64 bit operating system. In order to find the optimal network parameters, several experiments were performed.

4.1. Batch Size

Batch size one of the important part of the Convolutional Neural Networks' training process. One of the effects of batch sizes are Convolutional Neural Networks' weights on layers. If the batch size is equal to the training samples, it is called batch gradient descent. If the batch size is equal to one, only one training sample updates the weights at a single iteration. Selecting batch size one is called stochastic gradient descent. If the batch size is between one and training dataset examples, it is called mini-batch size. Selecting too small batch size changes the weights of convolutional neural network layers with only one image, leading to unsuccessful results. On the other hand, selecting a mini-batch size too high can consume much more memory. In this study, different batch sizes effects on the object detection model were tested. Especially with ResNet50 with Faster R-CNN model only tested with selecting mini-batch size one. If mini-batch size increased more than one, GPU memory overflows and leads to the failed training process. In the scope of this thesis, VGG19 with Faster R-CNN model was tested with different batch sizes to test the mini-batch effect on the object detection model. Choosing a batch size more than 1 prevents the model from only updating weights per image for one iteration; however, increasing too much batch size lowers the system's accuracy, precision and leads to higher GPU memory usage. On Table 4.1 VGG19 networks performance per class was tested by different batch sizes.

Table 4.1. Batch Size Results

CNN Model	Batch Size	Object Class	Average Precision
VGG19	1	Positive	0.8016
VGG19	1	Negative	0.7298
VGG19	2	Positive	0.7776
VGG19	2	Negative	0.7260
VGG19	4	Positive	0.8382
VGG19	4	Negatives	0.7605
VGG19	8	Positives	0.7506
VGG19	8	Negatives	0.6909

4.2. Anchor Numbers

One of the essential parameter selections is anchor boxes which are determined by ground truth values. Choosing the correct anchor box number increases the success rate of the object detection model. Chosen anchor box values should have a higher overlap value with ground truth boxes. More anchor boxes mean more anchor boxes similar to ground truth labels which increases IoU value. However, in some cases, more anchor boxes can lead to unsuccessful results. On Table 4.2 different CNN models with different anchor box number performance assessed. For the ResNet18 model, selecting anchor box number three was a sweet spot for getting good average precision. ResNet50 model, on the other hand, behaves differently; increasing anchor box number badly affects average precision. For VGG type network models, selecting Anchor Box number four was a sweet spot for networks' success. VGG19 networks' success rate increased with anchor boxes number increases. For the anchor box number comparison test, anchor box estimations and network performances were done only for KI67 positive classes.

4.3. Different Network Models

A critical part of object detection is selecting the correct feature extractor network. The Faster R-CNN model uses feature maps for object detection. Because of that, the Faster

Table 4.2. Anchor Box Number

CNN Model	Batch Size	Anchor Box Number	Average Precision
ResNet18	2	2	0.69394
ResNet18	2	3	0.7484
ResNet18	2	4	0.71349
ResNet18	2	5	0.78110
ResNet50	1	2	0.7521
ResNet50	1	3	0.7345
ResNet50	1	4	0.6981
VGG16	3	2	0.7581
VGG16	3	3	0.7185
VGG16	3	4	0.7649
VGG16	4	5	0.7691
VGG16	4	6	0.6917
VGG19	3	2	0.7148
VGG19	3	3	0.7251
VGG19	3	4	0.7944
VGG19	3	5	0.7578

R-CNN object detector model highly depends on the backbone Convolutional Neural Network type. In this Table 4.3 the success rate of CNN models are compared. All networks' best parameters were selected for comparison and their performances. For this test, all classes were included to measure the performance of the networks. Also, the anchor box number was selected two for all networks. For batch sizes, the only difference was that ResNet50 back-boned models' mini-batch size which was selected one for high memory consumption problems. Other models' mini-batch size is selected two for testing network performances. For KI67 positive object class detection, the ResNet50 back-boned Faster R-CNN model have achieved the highest value. For KI67 negative object class detection, VGG19 back-boned network have achieved the highest value.

Table 4.3. Different Network Model Comparison

CNN Model	Object Class	Average Precision	Recall	Precision	F1 Score
ResNet18	Positives	0.7857	0.8462	0.9112	0.8775
ResNet18	Negatives	0.7102	0.7897	0.8035	0.7965
ResNet50	Positives	0.8393	0.8407	0.9217	0.8793
ResNet50	Negatives	0.7025	0.7725	0.8126	0.7921
VGG16	Positives	0.8055	0.8132	0.8268	0.8199
VGG16	Negatives	0.7082	0.7554	0.7928	0.7736
VGG19	Positives	0.7776	0.7582	0.8903	0.8190
VGG19	Negatives	0.7356	0.7725	0.7982	0.7852

4.4. Supervised Learning and Semi-Supervised Learning

In the scope of this study, two different learning systems were adapted to the object detection model. All CNN models were tested with these learning methods. Test results are given in Table 4.4. Evaluation criteria for different learning methods with CNN models are Average Precision, Recall, F1 Score values. Recall and Precision values on Table 4.4 were selected when F1 Score had the highest value. First, all CNN models with Faster R-CNN models were trained with the Supervised learning method. For the Supervised learning system, all ground truth labeling processes were done manually with human interaction. The Semi-Supervised learning method allows for the continuation of the parameter update of the already existing Faster R-CNN model with automatically detected images by the same model. Therefore, only an automated labeled dataset was used. The Semi-Supervised learning method was also used for two iterations. For the first iteration of the Semi-Supervised method, a detector created with the Supervised learning method was used to label images that were not present in training, validation, and test datasets. Then only the automated labeled dataset images were used to continue training the supervised detector. Therefore training process did not start from the beginning, continued from the previous state. After the first iteration of the training process of the detector was done, that detector was used for the new dataset labeling process for the second iteration

Table 4.4. Supervised and Semi Supervised Learning Comparison

CNN Model	Learning Type	Object Class	Average Precision	Recall	Precision	F1 Score
ResNet18	Supervised	Positives	0.7857	0.8462	0.9112	0.8775
ResNet18	Supervised	Negatives	0.7102	0.7897	0.8035	0.7965
ResNet18	1st Semi-Supervised	Positives	0.7749	0.8571	0.8478	0.8525
ResNet18	1st Semi-Supervised	Negatives	0.7331	0.7747	0.7917	0.7831
ResNet18	2nd Semi-Supervised	Positives	0.8338	0.8462	0.8750	0.8603
ResNet18	2nd Semi-Supervised	Negatives	0.7412	0.8047	0.7813	0.7928
ResNet50	Supervised	Positives	0.8393	0.8407	0.9217	0.8793
ResNet50	Supervised	Negatives	0.7025	0.7725	0.8126	0.7921
ResNet50	1st Semi-Supervised	Positives	0.8475	0.8407	0.9273	0.8818
ResNet50	1st Semi-Supervised	Negatives	0.7546	0.8133	0.7929	0.8030
ResNet50	2nd Semi-Supervised	Positives	0.8613	0.8681	0.8977	0.8827
ResNet50	2nd Semi-Supervised	Negatives	0.7619	0.7618	0.8353	0.7969
VGG16	Supervised	Positives	0.8055	0.8132	0.8268	0.8199
VGG16	Supervised	Negatives	0.7082	0.7554	0.7928	0.7736
VGG16	1st Semi-Supervised	Positives	0.8071	0.8187	0.8663	0.8418
VGG16	1st Semi-Supervised	Negatives	0.6661	0.7682	0.7427	0.7553
VGG16	2nd Semi-Supervised	Positives	0.8108	0.7857	0.8773	0.8290
VGG16	2nd Semi-Supervised	Negatives	0.7700	0.7704	0.8215	0.7951
VGG19	Supervised	Positives	0.7776	0.7582	0.8903	0.8190
VGG19	Supervised	Negatives	0.7260	0.8133	0.7610	0.7863
VGG19	1st Semi-Supervised	Positives	0.7824	0.7857	0.8462	0.8148
VGG19	1st Semi-Supervised	Negatives	0.7541	0.7918	0.7785	0.7851
VGG19	2nd Semi-Supervised	Positives	0.8164	0.8187	0.8371	0.8278
VGG19	2nd Semi-Supervised	Negatives	0.7356	0.7725	0.7982	0.7852

with the Semi-Supervised learning method. Then the second set of the automated labeled dataset was used for the second iteration of the object detector training process.

For the Semi-Supervised method, an automated labeling process with human intervention was used for the object labeling process. The supervisor was asked to add labeled

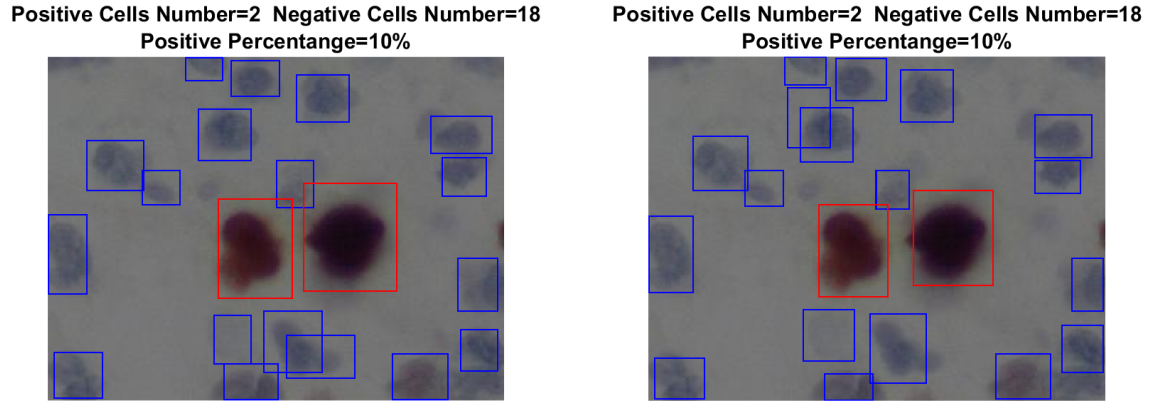


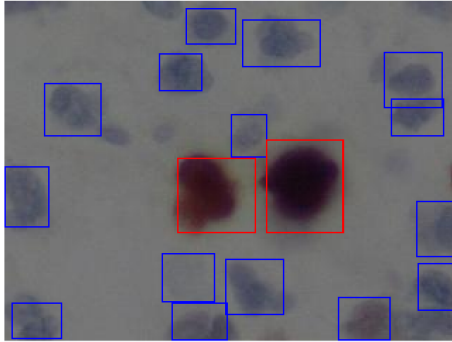
Figure 4.12. VGG16 with Faster R-CNN a) Supervised b) Semi-Supervised

objects as ground truth objects.

For these tests, different CNN models with Faster R-CNN model was used. ResNet18 backbone Faster R-CNN model with Semi-Supervised learning method, 17 new images were labeled in the first, and 29 new images were labeled in the second iteration of the Semi-Supervised learning process. ResNet50 backbone Faster R-CNN model with Semi-Supervised learning method, 12 new images were labeled for the first and second iteration of Semi-Supervised learning process. VGG16 backbone Faster R-CNN model with Semi-Supervised learning method, 8 new images are labeled for the first and second iteration of Semi-Supervised learning process. VGG19 backbone Faster R-CNN model with Semi-Supervised learning method, 9 new images were labeled for the first and 12 new images were labeled for the second iteration of the Semi-Supervised learning process. For Faster R-CNN models with the Supervised Learning method, the learning rate was selected as 10^{-4} for the training process. For Faster R-CNN models with the Semi-Supervised Learning method, the learning rate was selected 10^{-5} for the first iteration of the training process, and 10^{-6} was selected for the second iteration of the training process.

As shown in Table 4.4 by using Semi-Supervised learning methods, the performance of average precision value for object detectors was significantly increased for both classes. For example, one of the test images randomly selected for testing the networks' object detector performance for positive and negative classes and proliferation index for images is also shown in figures. VGG16 based test is shown in Fig 4.12 . VGG19 based test is shown in Fig 4.13 . ResNet18 based test is shown in Fig 4.14. ResNet50 based test is shown in Fig 4.15.

Positive Cells Number=2 Negative Cells Number=15
Positive Percentange=12%



Positive Cells Number=2 Negative Cells Number=17
Positive Percentange=11%

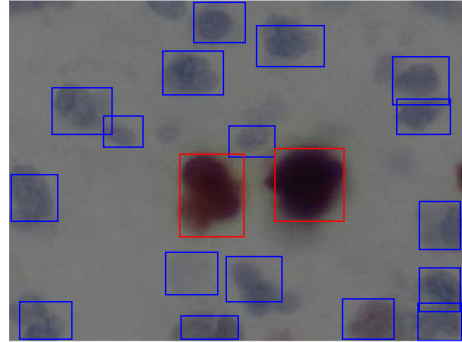
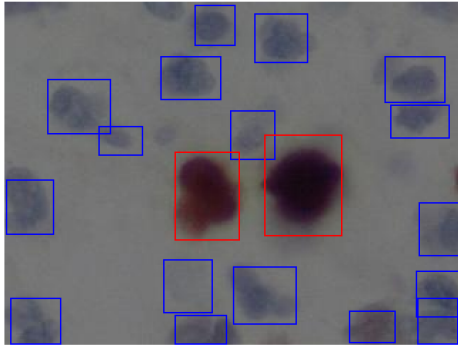


Figure 4.13. VGG19 with Faster R-CNN a) Supervised b) Semi-Supervised

Positive Cells Number=2 Negative Cells Number=17
Positive Percentange=11%



Positive Cells Number=2 Negative Cells Number=17
Positive Percentange=11%

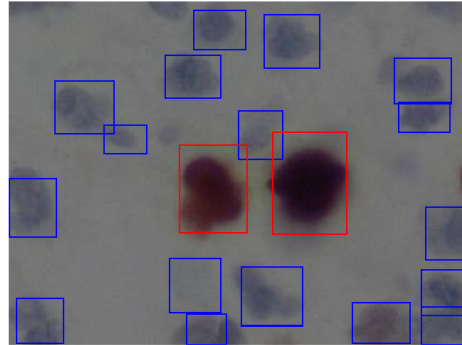
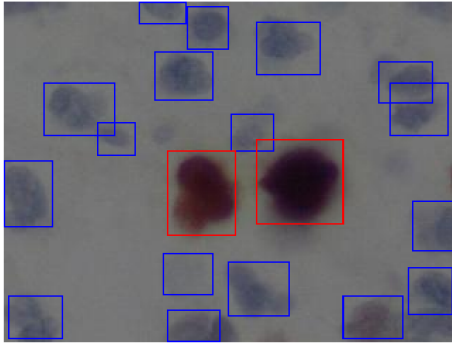


Figure 4.14. ResNet18 with Faster R-CNN a) Supervised b) Semi-Supervised

Positive Cells Number=2 Negative Cells Number=17
Positive Percentange=11%



Positive Cells Number=2 Negative Cells Number=18
Positive Percentange=10%

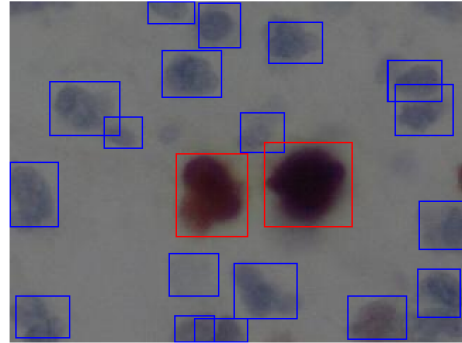


Figure 4.15. ResNet50 with Faster R-CNN a) Supervised b) Semi-Supervised

5. CONCLUSION

For this thesis, KI67 nuclear protein applied cell labeling performed by deep learning algorithms. Different network structures and convolutional network parameters were used to find the optimal model and parameters for that object detection problem. Also, different learning methods were tested to increase the efficiency of the KI67 labeling process. The iteratively Semi-Supervised method was tested for validation of the method's success rate. Advantages of using these methods are ;

- High Precision value even with an insufficient labeled dataset.
- Model can be implemented to Semi-supervised learning method.
- an automated image labeling process can replace manual image labeling process, but only with human observation.
- Automated image labeling process reduces the time for image labeling process and increases dataset quantity.
- Iteratively using the Semi-Supervised learning method significantly increase the accuracy of the labeled KI67s.

Some shortcomings are

- Even in small increases of batch size quickly overflow memory can lead to training failure. Especially for ResNet 50 back-boned system
- Wrong human intervention on semi-supervised learning can lead to significant success loss.
- Models are too sensitive to training parameters even slight change can lead to a decrease of success of object detection

6. REFERENCES

- Bengio, Y.: Practical recommendations for gradient-based training of deep architectures. In: Neural networks: Tricks of the trade, pp. 437–478. Springer, 2012.
- Chellapilla, K., Puri, S. and Simard, P. 2006. High Performance Convolutional Neural Networks for Document Processing. In: Lorette, G. (ed.) Tenth International Workshop on Frontiers in Handwriting Recognition. Suvisoft.
- Ciregan, D., Meier, U. and Schmidhuber, J. 2012. Multi-column deep neural networks for image classification. In: 2012 IEEE conference on computer vision and pattern recognition, pp. 3642–3649. IEEE.
- Girshick, R. 2015. Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp. 1440–1448.
- Girshick, R., Donahue, J., Darrell, T. and Malik, J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580–587.
- Goodenberger, M.L. and Jenkins, R.B. 2012. Genetics of adult glioma. *Cancer Genetics*, 205(12):613–621. <https://doi.org/https://doi.org/10.1016/j.cancergen.2012.10.009>, <https://www.sciencedirect.com/science/article/pii/S2210776212002608>.
- Goodfellow, I., Bengio, Y. and Courville, A. 2016. Deep Learning. MIT Press, <http://www.deeplearningbook.org>.
- He, K., Zhang, X., Ren, S. and Sun, J. 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2016-December, pp. 770–778. IEEE Computer Society.
- Kontzoglou, K., Palla, V., Karaolanis, G., Karaikos, I., Alexiou, I., Pateras, I., Konstantoudakis, K. and Stamatakis, M. 2013. Correlation between ki67 and breast cancer prognosis. *Oncology*, 84(4):219–225.

- Krizhevsky, A., Sutskever, I. and Hinton, G.E. 2017. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90.
- LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.
- Li, L.T., Jiang, G., Chen, Q. and Zheng, J.N. 2015. Ki67 is a promising molecular target in the diagnosis of cancer. *Molecular medicine reports*, 11(3):1566–1572.
- Noble, W.S. 2006. What is a support vector machine? *Nature biotechnology*, 24(12):1565–1567.
- O’Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G.V., Krpal-kova, L., Riordan, D. and Walsh, J. 2019. Deep learning vs. traditional computer vision. In: Science and Information Conference. pp. 128–144. Springer.
- Redmond, S.J. and Heneghan, C. 2007. A method for initialising the k-means clustering algorithm using kd-trees. *Pattern recognition letters*, 28(8):965–973.
- Ren, S., He, K., Girshick, R. and Sun, J. 2016. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6):1137–1149.
- 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.
- Simonyan, K. and Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition.
- Sophea, P., Handayani, D.O.D. and Boursier, P. 2018. Abnormal cervical cell detection using hog descriptor and svm classifier. In: 2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA). pp. 1–6. IEEE.

- Syriopoulou, E., Bower, H., Andersson, T.M., Lambert, P.C. and Rutherford, M.J. 2017. Estimating the impact of a cancer diagnosis on life expectancy by socio-economic group for a range of cancer types in england. *British journal of cancer*, 117(9):1419–1426.
- Twombly, R. 2005. Cancer surpasses heart disease as leading cause of death for all but the very elderly. *Journal of the National Cancer Institute*, 97(5):330–331.
- Xia, T., Jiang, R., Fu, Y.Q. and Jin, N. 2019. Automated blood cell detection and counting via deep learning for microfluidic point-of-care medical devices. In: IOP Conference Series: Materials Science and Engineering. vol. 646, p. 012048. IOP Publishing.
- Zhang, R., Yang, J. and Chen, C. 2018. Tumor cell identification in ki-67 images on deep learning. *Molecular & Cellular Biomechanics*, 15(3):177.
- Zhou, Z.H. 2018. A brief introduction to weakly supervised learning. *National Science Review*, 5(1):44–53.

CURRICULUM VITAE

Caner SONGÜL

EDUCATION

Master of Science 2017-2021	Akdeniz University Institute of Natural and Applied Sciences, Computer Engineering, Antalya, Turkey
Bachelor of Science 2010-2014	Akdeniz University Akdeniz University Faculty of Engineering, Electrical Electronics Engineering, Antalya, Turkey

WORK EXPERIENCE

Research Assistant 04/2019-Present	Alanya HEP University Faculty of Engineering, Computer Engineering, Antalya, Turkey
---------------------------------------	---