REPUBLIC OF TURKEY

AKDENIZ UNIVERSITY

AUTOMATIC RECOGNITION OF LYMPH NODE METASTASIS OF BLADDER CANCER WITH ARTIFICIAL INTELLIGENCE

Muhammet Fatih ÇAKMAKÇI

INSTITUTE OF NATURAL AND APPLIED

SCIENCES DEPARTMENT OF COMPUTER

ENGINEERING MASTER THESIS

MAY 2021

ANTALYA

REPUBLIC OF TURKEY

AKDENIZ UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

AUTOMATIC RECOGNITION OF LYMPH NODE METASTASIS OF BLADDER CANCER WITH ARTIFICIAL INTELLIGENCE

MUHAMMET FATİH ÇAKMAKÇI

DEPARTMENT OF COMPUTER ENGINEERING

MASTER THESIS

This thesis was unanimously accepted by the jury on 07/05/2021.

Asst. Prof. Hüseyin Gökhan AKÇAY

Assoc. Prof. Dr. Havva Serap TORU

Asst. Prof. Shahram TAHERI

ÖZET

YAPAY ZEKÂ İLE MESANE KANSERİNİN LENF NODU METASTASLARININ OTOMATİK TANINMASI

Muhammet Fatih ÇAKMAKÇI

Yüksek Lisans Tezi, Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Hüseyin Gökhan AKÇAY

Mayıs 2021; 41 sayfa

Yapay zekânın önemli bir alt başlığı olan derin öğrenme, sağlık hizmetlerinde önemli gelişmelere ve kolaylaştırmalara olanak tanımaktadır. Bu bağlamda derin öğrenme yöntemleri kullanılarak geliştirilen sağlıkla alakalı uygulamalar gün geçtikçe artmaktadır. Kanserli hücrelerin tespiti, tıbbi görüntüleme, hastalıkları öngörmek ve hastalıkların bulut sistemlerinde kayıtlarının alınması gibi çalışmalar buna örnek olarak verilebilir. Özellikle hastalıklı ve kanserli hücrelerin tespiti büyük öneme sahiptir. Çünkü tıbbın en önemli kuralı olan "erken teşhis hayat kurtarır" dan yola çıkıldığında, kanserli hücrelerin yapay zekâyla erken teşhis edilmesi hastaların hayata döndürülmesine büyük katkı sağlamaktadır. Çok çeşitli kanser türü olmakla beraber mesane kanseri en sık görülen kanser türlerinden biridir. Daha geniş bir doku alanına yayılan kanserli hücrelerin yanında küçük boyutlu tümör hücrelerini de içeren mesane kanseri teşhis edilmesi zor bir kanser türüdür.

Bu araştırmada, mesane kanseri tümörlerinin lenf nodu metastazlarının derin öğrenme kullanılarak tespit edilmesi amaçlanmıştır. Yapılan deney ve eğitimlerde büyük alanlara yayılmış kanser hücreleri derin öğrenme modeliyle büyük ölçüde tespit edilebildiği halde küçük alana yayılan tümörlerde bu başarı düşüktür. Bu çalışmanın asıl amacı, Mask R-CNN algoritması kullanılarak büyük tümörlü alanlarda elde edilen başarılı tanıma sonuçlarının, modelin tanıma oranının düşük olduğu küçük tümörlü alanlarda da aynı ya da yakın bir oranla elde edilmesidir. Böylece uzun, zorlu çalışma saatleri ve oldukça yüksek hasta sayısına sahip olan doktorların, gözden kaçırabileceği küçük tümör alanlarının daha yüksek oranda tanınmasıyla doktorlara büyük kolaylık sağlanacaktır. Çalışma için özel oluşturulan veri seti, ResNet modelini kullanan Mask R-CNN algoritması ile test edilmiştir. Test sonuçlarında büyük alana yayılmış tümörlü hücrelerde kayda değer başarı sağlanırken küçük alanlarda modelin aynı başarıyı elde edemediği gözlenmiştir. Bu amaçla modelin aşama ve katman sayılarında değişiklikler yapılarak ve küçük bir veri setine rağmen küçük tümör alanlarında önemli bir tanıma artışı elde edilmiştir.

ANAHTAR KELİMELER: Yapay zekâ, Derin öğrenme, Mesane kanseri, Metastaz, E-Sağlık, Mask R-CNN

JÜRİ: Dr. Öğr. Üyesi Hüseyin Gökhan AKÇAY
Doç. Dr. Havva Serap TORU
Dr. Öğr. Üyesi Shahram TAHERI

ABSTRACT

AUTOMATIC RECOGNITION OF LYMPH NODE METASTASIS OF BLADDER CANCER WITH ARTIFICIAL INTELLIGENCE

Muhammet Fatih ÇAKMAKÇI

MSc Thesis in Computer Engineering

Supervisor: Asst. Prof. Hüseyin Gökhan AKÇAY

May 2021; 41 pages

Deep learning, which is an important subheading of artificial intelligence, enables significant improvements and facilitations in healthcare services. In this scope, health-related applications developed by deep learning methods are increasing day by day. Studies such as the detection of cancer cells, medical imaging, predicting diseases and recording diseases in cloud systems can be given as examples. Detection of diseased and cancerous cells is of great importance. Since "early diagnosis saves lives" is the most important medicine rule, early diagnosis of cancer cells through artificial intelligence makes a significant contribution to getting patients back to life. There are several different types of cancer, and bladder cancer is one of the most common types of cancer. Bladder cancer, which involves small tumor cells as well as cancer cells that spread over a wider area of tissue, is a difficult form of cancer to diagnose.

The aim of this study is to detect the lymph node metastasis of bladder cancer using artificial neural networks. In experiments, while cancer cells that spread across large areas can be identified to a large extent by a deep learning model, this performance is lower in tumors that spread to small areas. The main purpose of this study is to achieve successful diagnosis accuracy in small cancer areas as well as in large cancer areas. Thus, significant convenience will be achieved with the recognition of small tumor areas in which doctors who have long, challenging working hours and quite high number of patients can overlook. Specially created dataset for the study was tested with the Mask R-CNN algorithm using the ResNet model. In the test results, it was observed that while significant success was achieved in tumor cells which covers large areas, the model did not achieve the same success on the small areas. For this purpose, changes and additions were made in the number of layers and stages of the model and a significant improvement in small tumor area detection was obtained.

KEYWORDS: Artificial intelligence, Deep learning, Bladder cancer, Metastasis, E-Health, Mask R-CNN

COMMITTEE: Asst. Prof. Hüseyin Gökhan AKÇAY

Assoc. Prof. Dr. Havva Serap TORU

Asst. Prof. Shahram TAHERI

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my advisor, Asst. Prof. Hüseyin Gökhan AKÇAY, who supervised, guided, and motivated me during this study. He integrated me into this field and steered me in the right direction with his invaluable knowledge and experience.

I would like to thank Assoc. Prof. Dr. Havva Serap TORU who is a lecturer at Akdeniz University School of Medicine, Department of Pathology, and Elif OCAK GEDİK, MD who is a research assistant at the same department for their warm approach, support and providing microscopic images.

I would like to thank Prof. Dr. Melih GÜNAY and all our Computer Engineering Department Teaching Members who support and guide.

I thank to appreciate my sincere gratitude to my dear friend MSc Egehan ÇETİN for his support during the thesis writing process and on other subjects.

Finally, I thank to express my gratitude to my mother and father who have always supported me with their knowledge and experience.

LIST OF CONTENTS

TEXT OF OATH

I declare that this study "**Automatic Recognition of Lymph Node Metastasis of Bladder Cancer 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.

07/05/2021

Muhammet Fatih ÇAKMAKÇI

ABBREVIATIONS

Abbreviations

AI	: Artificial Intelligence
CNN	: Convolutional Neural network
CSV	: Comma Separated Values
CT	: Computed Tomography
DCNN	: Deep Convolutional Neural Networks
DL	: Deep Learning
EMR	: Electronic Health Record
FPN	: Feature Pyramid Network
GAN	: Generative Adversarial Network
ILSVRC	: ImageNet Large Scale Visual Recognition Competition
LDA	: Latent Dirichlet Allocation
mAP	: Mean Average Precision
ML	: Machine Learning
MRI	: Magnetic Resonance Imaging
OAR	: Organs At Risk
ReLU	: Rectified Linear Unit
RNN	: Recurrent Neural Network
ROI	: Region Of Interest
RPN	: Region Proposal Network
SAF	: Semantic Analysis Framework
SVM	: Support Vector Machine
TNM	: Staging of Malignant Tumors
VGG	: Visual Geometry Group
WHO	: World Health Organization

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Every year, more than 18 million new instances of disease are analysed around the world. Universally, one in each five individuals builds up the sickness before the age of 75 years. However, this ratio varies from 1 in 10 in South-Central Asia and Central Africa to 1 in 2.5 in Australia and New Zealand. At the present, there are around 44 million malignant tumor growth patients alive who were analysed under 5 years back. Right around 10 million individuals pass away from cancer growth yearly (Bray et al. 2018).

Around 550.000 individuals were determined to have bladder cancer in 2018 (Richters et al. 2020). This rate constitutes to about 3% of all new cancer findings and Bladder cancer is the eleventh most popular cancer type around the world. This type of cancer is normally first suspected because of hematuria and afterward diagnosed to a cystoscopy or an adjustable endoscopy of the bladder or transabdominal ultrasound, or/and PC tomography (CT) urography or MRI. Upwards of 7 out of each 10 instances of bladder cancer are recognized in beginning phases, so this early recognition rate provides higher endurance and to stay alive. While durability rates have improved with before conclusion, robotic surgical strategies, and presentation of immunotherapy, bladder cancer growth continues to be a large and growing supporter of disease problems worldwide, especially in large and developed countries (Saginala et al. 2020).

World Health Organisation (WHO) and International Society of Urological Pathology (ISUP) have divided urothelial tumors into two major groups: (1) Infiltrating urothelial carcinoma (2) Non-invasive urothelial Lesions. Non-invasive urothelial Lesions are subclassified to (a) Urothelial carcinoma in situ; (b) Non-invasive papillary urothelial carcinoma, low-grade; (c) Non-invasive papillary urothelial carcinoma, high-grade; (d) Papillary urothelial neoplasm of low malignant potential; (e) Urothelial papilloma; (f) Inverted urothelial papilloma; (g)Urothelial proliferation of uncertain malignant potential; (h) Urothelial dysplasia (Warren and Harrison 2018). Infiltrating tumors tend to spread to lymph nodes. TNM staging is used for bladder. In TNM staging includes tumor size, lymph node metastasis and solid organ metastasis. So lymph node metastasis is an important and necessary component for staging of bladder cancer (Teoh et al. 2020).

Over the past few years, researchers have shown a growing interest in deep learning models that automate tasks with accuracy equivalent to those of experts and demonstrate superior performance and powerful knowledge discovery capabilities in various fields. In this process, the imaging and visual functions of deep convolutional neural network models (DCNN) are widely used. It has been revealed that it can be used to solve many computer vision problems (Dong et al. 2020; Liu et al. 2020), including medical image analysis. However, there are not many DCNN applications used in medical imaging because of lack of sufficient amount of annotated images (Mechria et al. 2019). Deep learning methods typically require too many samples to train the network, and there are also memory and computational power limitations (Yogananda et al. 2020, Sharma and Mehra 2020). There are a lot of algorithms which based on deep learning. Mask R-CNN is one of deep learning-based algorithms that a theoretically basic, flexible, and generic structure for instance segmentation. It has been shown that Mask R-CNN can successfully segment the gross tumor volume and organs at risk (OAR), and it has been

proposed that deep learning through convolutional neural networks (CNN) can create correct outline in every stage of treatment (Breto et al. 2019). Deep learning used in grading bladder cancer in previous studies has produced encouraging results (Jansen et al. 2020). Similarly, Woerl et al. (2020) reported that the deep learning method performed similar to or better than experts in predicting Muscle-invasive bladder cancer from hematoxylin and eosin slides in patients. While some studies have focused on using Machine learning (ML) algorithms to aid in the detection, diagnosis and staging of bladder cancer, and especially deep learning (Suarez-Ibarrola et al. 2019), their use in bladder cancer prognostication and outcome prediction is relatively limited.

Within the scope of the thesis, lymph nodes with urothelial carcinoma metastasis are selected for automate image analysis. It was aimed to detect the lymph node metastasis of urothelial carcinomas with deep learning algorithms to improve the accuracy and shorten the evaluation time. Digital images are taken from metastatic and non-metastatic lymph nodes with the CellSense computer program. Afterwards, a CNN-based Mask R-CNN algorithm was trained in this direction and a recognition model was developed using 365 metastatic and non-metastatic lymph node images obtained from 11 patients. The modified model has achieved significant success on large and small tumor areas.

The Literature Section of the thesis contains reviews articles that are similar to the subject of the study and also basic information is given about the physiology of bladder cancer. Diagnostic technique, cystoscopy process, images and database formation, artificial intelligence and Mask R-CNN algorithm and proposed modification are listed in the section Materials and Methods. In the Discussion and Results section, comparison between original model and proposed modification to the original model is done. Conclusion section includes relevant information and summary of this study.

2. LITERATURE REVIEW

2.1. Literature Summary

Deep learning models that can process large amounts of data and automatically extract features from input are based on Neural Networks that mimic the work of the brain (Krishnakumar et al. 2019). Deep neural networks are a state-of-the-art concept in which neural network techniques are used. The best-known deep learning model is the convolutional neural network (Gao et al. 2017). Deep neural networks simulate visual system of human and can achieve in a human-like equality in image classification, object sensation, and segmentation (Yun et al. 2018). In fact, deep learning neural networks can learn and predict unknown values of data once known data is obtained. Also, this method always requires prior training on a specific data form to provide the best predictive results. Research has shown that deep learning techniques are consistently applied to image-based medical diagnosis and outperform classic machine learning methods and exceed human performance in visual tasks (Kamnitsas et al. 2017; Khosravi et al. 2018).

Investigation of the cell classification methods such as blood image segmentation (Shirazi et al. 2016), and an overview of medical image and deep learning (Razzak et al. 2018), nuclei segmentation (Kurnianingsih et al. 2019), organ and structure segmentation (Dong et al. 2020) and leukocyte segmentation (Shirazi et al. 2016). It has been reported that the deep learning method, which is used in medical injury research (Yu 2020), medical diagnosis (Gao et al. 2017; Iqbal et al. 2019), classification of diseased and normal cells (Iqbal et al. 2020), and other fields such as Internet of things (IoT) technology (Zhao et al. 2019; Kumar et al. 2020), overperforms classic learning methods in high dimensional image recognition problems (Iqbal et al. 2019), multi-classification and detection of cancer cells (Sharma and Mehra 2020), analysis of histopathological images (Wang et al. 2018), and classification of apical echocardiograms (Khamis et al. 2017).

In previous research, the deep learning methods have been used in analyzing huge amounts of data (Angermueller et al. 2016), classifying cancer cells (He et al. 2018; Iqbal et al. 2019; Yang et al. 2020), counting cells and mapping cell spatial density between images (Xie et al. 2018), extracting characteristics of a single cell and extracting and capturing image phases and intensity and classifying diseased and normal cells (Iqbal et al. 2020). On the other hand, the deep learning model was used in the prediction and detection of lung nodules (Wang et al. 2017), breast cancer (Krishnakumar et al. 2019; Mechria et al. 2019; Sharma and Mehra 2020), cervical cancer (Liu et al. 2018), skin cancer (Yu et al. 2017) and disease such as medical image analysis (Hao et al. 2019).

In a study conducted by Cha et al. (2017a), who evaluated the possibility of distinguishing bladder cancers with predictive models based on radiomics using pre- and post-treatment computed tomography. They used 3 different radiomic predictive methods such as DL-CNN, RF-SL and RF-ROI predictive model. RF-SL extracted radiomics features from segmented lesions. RF-ROI extracted radiomics features from pre-post treatment paired ROI's. In this experiment conducted with samples taken from 67 male and 15 female patients which are their age range between 37 and 84, the CNN algorithm was trained with 6700 ROI. The used CNN algorithm modified as two convolution layers,

two local layers and a full linked layer. Test Accuracy of this study are range of 0.69 to 0.77.

Cha et. al (2017b) tried to modify CNN model features such as changing network sizes and then trained it with transfer learning or not using ROI's or natural scene images. After that they expected resulted as inside or outside of bladder. The 87 lesions of 82 patients produced 104 pairs of temporal lesions and 6,700 paired ROIs prior to treatment. CNN without transfer learning test result is 0.77 since using natural scene images and transfer learning test result is 0.73 since using bladder images. Thus, they showed that Transfer learning did not improve the estimate of response to treatment. According to the results of this study, the most appropriate and accurate result was obtained with DL-CNN model when transference learning with bladder images used.

Garapati et al. (2017) used machine learning algorithms for bladder cancer stagging. They used two stages depending on pathological stage greater than or equal to T2 or less than equal to T2. It is the decision threshold for a clinically new adjuvant chemotherapy. 84 lesions which were auto segmented with their old and enhanced level sets algorithm were obtained 43 cancer lesions which is from stage $\geq$ T2 and 41 cancer lesions from stage $\leq$T2. Their new method is morphological and textural features were extracted as first. Then, it was divided into 3 different ways as morphological, texture features, and both morphological and texture features. Also, cross validation method used for augmented dataset. For that purpose, dataset was divided as Set1 and Set2. Feature progressive determination was used to determine which features were most effective. 3 different methods are LDA, RAF, SVM and NN combined and obtained one single accuracy rate. NN's accuracy is 0.95 on test set for combined features and 0.90, 0.89 and 0.96 for LDA, SVM and RAF, respectively. Sum up, developed model is promising for recognition of classification of two stage of bladder cancer.

Dolz et al. (2018) investigated a network contains extended convolutions to increase the field of reception without incurring additional costs or affecting its performance. By adding progressive expansions in each convolutional block, they offered incremental expansions and created wide receptive fields without the need for great expansion rates. Actually, its traditional CNN algorithm based on U-net and used dilated convolutions in addition.53 Male and 7 Female patients' tumors are collected for the dataset. Their age average is 67 years old. Dataset were divided by training/validation and test sets and their amounts are 40,5 and 15 patients, respectively. This proposed method has achieved 98%, 84% and 69% for inner wall, outer wall, and tumor area segmentation accuracy rate. The results showed their algorithm had high level accuracy than state of the art algorithms.

Liu et al. (2018) suggested a method of cervical core seperation that uses pixel-level to train regional convolutional neural network (Mask R-CNN). It has been emphasized that the structure of a Mask-RCNN is basically on a convolutional neural network, which is a deep network consisting of various layer types and is mainly composed of three phases. Furthermore, the researchers ultimately suggested that this method is a powerful technique and will play an important role in automated cytology screening and cancer diagnosis, particularly nuclear segmentation.

Kurnianingsih et al. (2019) used Mask R-CNN for segmenting cervical cells and VGG-like Net for classification. According to their reports, at the segmentation stage, when Mask R-CNN was applied the whole cell, precision was 0.92, recall 0.91 and ZSI 0.91.

Sompawong et al. (2019) used it for cervical cancer detection and classification by applying a pre-trained Mask R-CNN to liquid-based histological slides. In this study, they reported that Mask R-CNN was used for the first time for cervical cell nucleus detection, analysis and nuclear properties, and the accuracy of the results was 89.8%, sensitivity 72.5% and specificity 94.3%.

Shkolyar et al. (2019) emphasized that bladder tumors' 20% has lost under white light cystoscopy. Cystoscopy enhanced with deep learning may improve the white light problem and bladder cancer's tumor localization, intraoperative mobility, and surgical resection. Due to such kind of problems, their goal was to develop a deep learning model for increased cystoscopic detection of bladder cancer. They built their own CysoNet which is platform depending on CNN for the automatic detection of bladder tumors through a development.100 patients' samples were used for this study. 95 samples were arranged for training and 5 samples for testing. Also 54 additional patients added to dataset for CystoNet's diagnostic performance. In the validation data set, the sensitivity and specificity per frame were 90.9% and 98.6, respectively. Every single tumor sensitivity was 90.9%. Moreover, CystoNet performance was very good that 39 of 41 papillary tumors were detected and 3 of 3 surface bladder cancers. To sum up created CystoNet can improved recognition of cystoscopy and efficiency of TURBT with high value of sensitivity and specificity.

Ma et al. (2019) developed a deep learning U-Net-based approach to bladder segmentation for urological tomography as part of a pipeline to assess response to treatment and find computer-aided bladder cancer. They already built a bladder segmentation algorithm which using neural network level sets inclusive of user bounding box. However, they faced some problems that failure recognition results cause of low image quality or extended bladder cancer outspread to other near body organs. So, they had to develop a new U-net deep learning method for estimating a bladder probability map in CTU.U-net has some good advantages such as no need user-input box or level sets for post processing. They compared some models for finding the best model for this study. For the test set Initial accuracy rate for base model was 81.9 % (DCNN-LS). 3D U-DL got the highest accuracy number is 90.6% (test set) from the compared models. They used 256 × 256 input size for the model. 173 cases are used as dataset. 81 cases for training and remaining cases for test sets. Consequently, new U-DL model accuracy performance was much better than DCNN-LS.

Xu et al. (2019) developed a deep learning algorithm to diagnose lung cancer with CT images of patients. They created the model that uses transfer learning from convolutional to recurring neural networks using tumor localization with a single seed point. Time series functions greatly improved the performance of the deep learning model that is ResNet CNN model, especially for disease progression, regional recurrence, and distant metastases. So, model performance turned into better with every extra observe-up scan into the CNN version. The algorithm tested by obtaining 739 CT scans from 268 patients which are stage III. Patient's age average was 63 and 52.8% were women. Model

accuracy rate is 0.74 (AUC). They showed that time series and transfer learning combination is improving CNN model accuracy for lung cancer detection accuracy.

Du et al. (2020), in order to determine bladder cancer by applying CNN to cystoscopy images, 1002 normal tissue photographs from 103 patients and 734 bladder tumors from 72 patients were taken under cystoscopy. The research has shown that bladder cancer can be correctly recognized with the Caffe deep learning.

Gordon et al. (2017) proposed a model that to detect bladder cancer in CT urography. While they are developing this system, they faced some hard and challenging issues such as wall differences between contrasting and non-contrasting regions, bladder appearance and wall, and size and shape differences. The CNN model is trained to predict the probability of a specific pixel being present within the bladder wall using neighbourhood information Segmentation precision was assessed by comparing segmented wall diagrams to manual outlines.79 training cases and 15 test cases are used for the accuracy calculation. The calculation has done by average volume intersection % as the metric. They achieved results 90 % and 93.7 % for inner and outer wall training sets respectively. 87.6 % and 87.2 % obtained for inner and outer wall test sets. The results showed that DL-CNN with the level groups was effective in segmentation of the inner and outer bladder walls.

Iqbal et al. (2020) researched the performance of deep learning and traditional machine learning classification methods in order to categorize normal and diseased cells in their studies. They used support vector machines and logistic regression as machine learning method and CNN as deep learning method for to classify normal and diseased cells. It is detected normal and healthy cells 98% with the CNN algorithms which are ResNet-50 and ResNet-152 models. In the data set, 60 training images, 16 test images were used in half for normal and diseased cells.

In a study conducted on 328 non-invasive bladder cancer (NMIBC) samples on 232 patients from three different medical centres, Jansen et al. (2020) proposed a fully automated detection and grading network based on deep learning. As a result, the researchers suggested that fully automatic detection and grading of common NMIBC is possible with deep learning and that this method can be used.

Ünaldı (2020) proposed classification with support vector machines (DVM) was performed using different CNN models such as VGG16, VGG19, Xception, InceptionV3, ResNet50 and Inception-ResNet-V2 for feature extraction and compared the performance of these methods. BreakHis dataset are used for this study. The BreakHis database consists of microscopic biopsy photographs of benign and malignant breast tumors. photos received during a scientific study performed between January 2014 and December 2014. 7909 breast tumor micrographs with various ratios (40X, 100X, 200X, 400X) obtained from 82 patients with clinical signs of breast cancer were used. According to test accuracy results, InceptionV3 was the best model with a test accuracy of 98.2% and a magnification factor of 400X. ResNet-50 was second best model with 98 % accuracy rate. This study shows that CNN models especially ResNet-50 and InceptionV3 are very successful for breast cancer tumor detection and another clinical research.

Yang et al. (2020) trained and compared 3 different deep learning models which are GoogleNet, LeNet and AlexNet and EasyDL platform to recognize bladder cancer on the images. Also, diagnosis efficiency of urology doctors was contrasted with those deep learning models. General model features were not changed or modified. Sigmoid and Relu were used for loss and activation function, respectively. They used Adam for optimization method. Learning rate was 0.0001 and dropout was 0.4. Epoch size was 150 and batch size was 20 1,200 cystoscopic images of cancer were gathered from 224 patients with bladder cancer and 1,150 cystoscopic images from 221 patients without bladder cancer. EasyDL's recognition rate was highest, and accuracy was 96.9. GoogLeNet's recognition rate was the second highest, and the accuracy was 92.54. They also compared accuracy of neutral network model and urology doctors. Out of 33 bladder cancer nodes and 11 no bladder cancer nodes, the neural network accuracy was 83.36% and medical expert's accuracy was 84.09%. They showed that a neutral network model was at least as much as good medical experts.

Yogananda et al. (2020) used to simplify the complex multi-layer fragmentation problem to individual binary hash problems for each subcomponent with 3 independent 3D Dense U-Nets. In this study, 285 cases were used as the data set, and the validation scores for segmentations of the whole tumor, nucleus and the augmenting tumor were 0.92, 0.84 and 0.80, respectively. Then, they retrained each binary segmentation network using 265 of the 285 cases, with 20 cases kept out for testing. The best accuracy result is obtained 90% for whole tumor. The best accuracy result was achieved with 90% for whole tumors, 84 % and 80 % for tumor nucleus and the augmenting tumor, respectively. In addition, they tested this model for some other datasets which are taken 46 cases from BraTS2017, 66 cases from BraTS2018 and 52 cases from an exclusive clinical dataset. Results for BraTS2017 are 0.90, 0.80, and 0.78 and 0.85, 0.80, and 0.77 for exclusive clinical cases. Thus, model was promising for brain tumor detecting for clinics and doctors.

2.2. Bladder Cancer

In the lower pelvis, the bladder is a small organ. It has elastic, muscular walls that can extend and retain the urine and squeeze it out of the body. The main task of the bladder is to collect urine. Two kidneys make up urine, which is a liquid waste. The urine is then transferred to the bladder through 2 tubes called ureters. During urination, the muscles of the bladder are stretched, and urine is pushed out of the bladder into a channel called the urethra (Anonymous 1). Details of the bladder are shown in Figure 2.1.

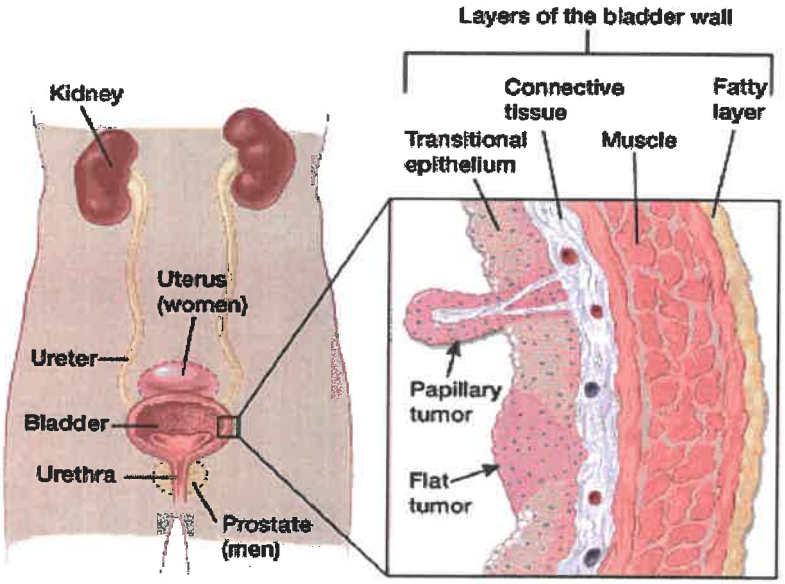

Figure 2.1. Details of bladder (Anonymous 1)

Bladder cancer is a common malignant tumor most of the tumors originate from urothelial cells that are lining inner part of bladder. The tumor expands to the bladder muscle in some circumstances. The most prominent symptom of bladder cancer is hematuria, that is typically painless (Anonymous 2). Other symptoms are dysuri, urgency and low back pain (Anonymous 3).

Overall, the incidence of bladder cancer is four times higher in men than women. The cancer is usually diagnosed in elder ages; 90 percent of the patients are over age of 55 and 80 percent of them are over the age of 65. Smoking is most important risk for bladder cancer and accounts for about 50-65 percent of new cases each year. Smoking increases the risk of the cancer 3 to 4 times (Saginala et al.2020).

Lymph node metastasis is important for tumor staging. Scanning lymph nodes is time consuming and detecting small metastatic tumor foci called micrometastasis is challenging (Lee et al. 2015).

3. MATERIAL AND METHOD

In this section, used image dataset features, their gathering process will be explained. After giving general information about deep learning and some important algorithms, proposed modification to the model for small tumor metastasis areas will be discussed in detail.

3.1. Data Collection

First, patients who has bladder cancer and suitable for this study were selected. A total of 365 samples were obtained from 11 different patients. After collecting cancer samples, Olympus BX53 light microscope (shown in Figure 3.1) was used for acquiring and visibly enlarging samples.

Olympus BX53 is a well-equipped microscope for cell studies. It has an advanced LED light source that provides exceptional clear and lifelike images. True Color LED generates a consistent output greater than the reference of 100-watt halogen and reflects the sample in real way by displaying every colour cast (Anonymous 4).

This microscope transfers the microscopic images to a computer application called 'cellSens Entry'. CellSens Entry is very detailed computer software which gives, transferred cell images can be enlarged, cropped, analysed and many more image processing process. 365 microscopic images were taken and enlarged 400 times with this program. Image's resolution, DPI and image bit depth is 2560 × 1920, 96 and 24-bit, respectively.

3.2. Study Protocol

There are many process sequences from patient selection to test results of cancer cells. All these procedures have been carried out in collaboration with the School of Medicine Department of Pathology of the Akdeniz University. Process scheme of this study is shown in Figure 3.1.

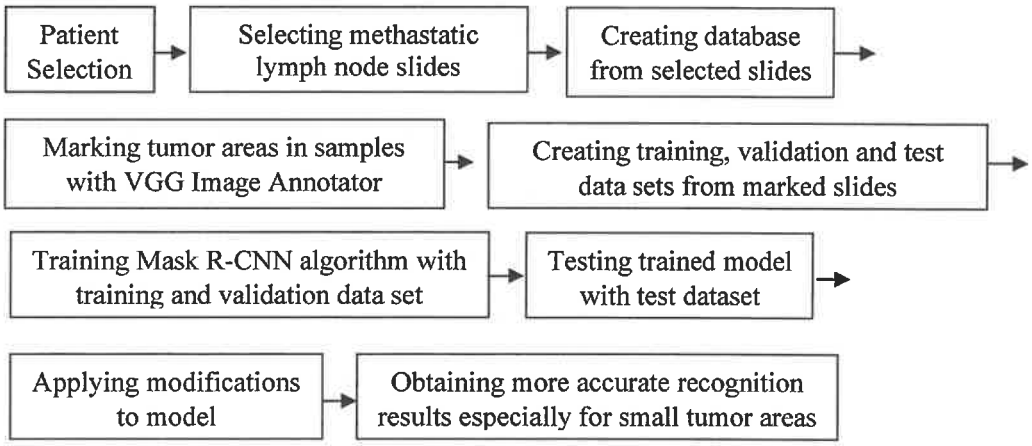

Figure 3.1. Process scheme

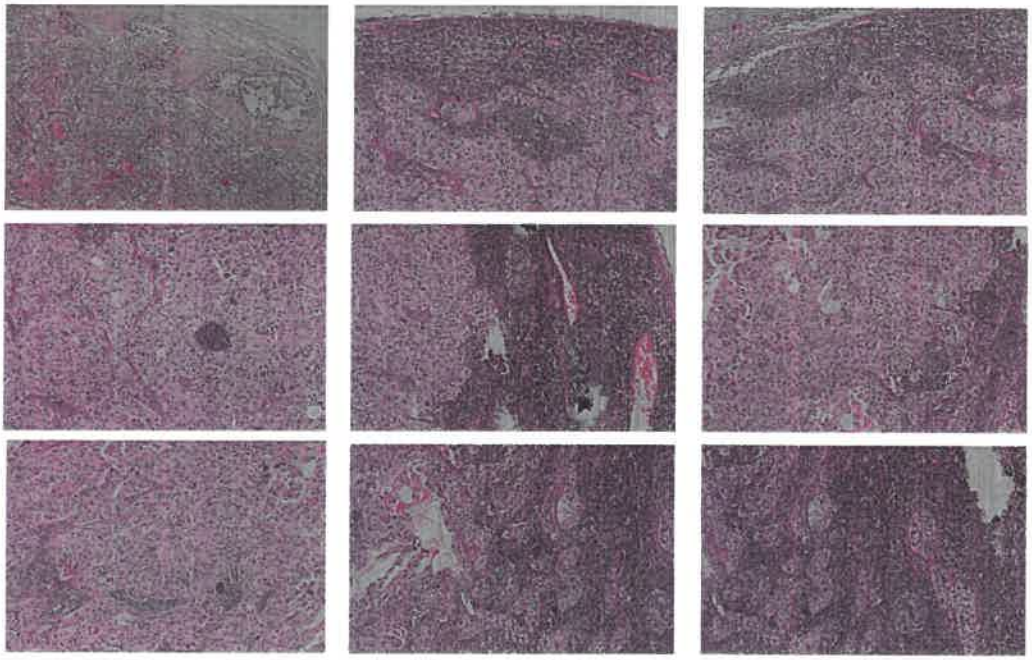

Figure 3.2. Examples of lymph node metastasis images

After obtaining bladder cancer tumor cells and groups (discussed on Section 3.1), a database was created by gathering all images. Some examples of dataset images are given in Figure 3.2. Then, tumor areas in microscopic images were marked with the web based VGG Image Annotator program (Dutta and Zisserman 2019).

After annotation, all microscopic images were divided into train, validation, and test sets. According to 60:20:20 percentage ratio rule, 225, 70 and 70 samples were used for train, validation, and test sets, respectively.

All annotated sample images imported to Mask R-CNN source code (that will be given more details in Section 3.3.3). After importing images and annotations to source code, Mask R-CNN was trained with training and validation data sets. After the training phase is complete, testing and recognition phase were done with 70 test images. Diagnosed bounding boxes, their accuracy values were given on the test images as recognition result. Examples of recognition results on test images are shown in Figure 3.3.

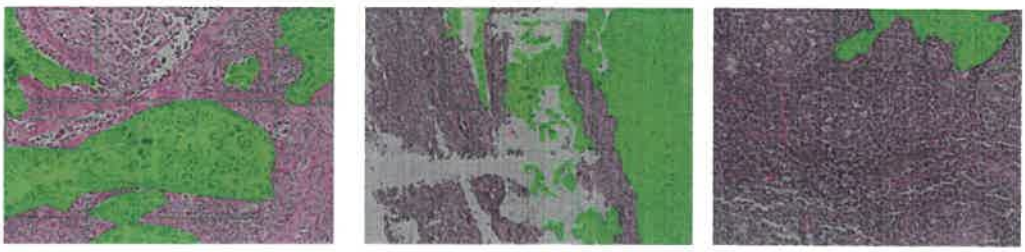

Figure 3.3. Examples of target metastasis regions (shown with green color)

After the testing phase, although the overall recognition results were quite satisfactory, the results were not good enough for small tumor areas. Therefore, the main purpose of this study was determined to be not only increasing the overall recognition accuracy but also increasing the recognition accuracy particularly in small metastatic tumor cell groups (that will be shown in detail in the Section 4.2).

3.3. Automated Metastasis Detection in Lymph Nodes

Artificial Intelligence (AI) is an overall expression that includes using a computer to model intellectual conduct with minimum human involvement. The invention of AI is considered to have begun robots. It is a fact that in this area, Da Vinci's robots' sketches helped lay the groundwork for this invention. His sketches inspired robotic surgery for complex procedures in urology and gynecology. The term of AI was officially born in 1956 by John McCarthy during a conference held on this subject. and defined as the science and engineering of smart machinery (Hamet and Tremblay 2017). Actually, possibility of machines simulating human behaviour and actually thinking was proposed earlier by Alan Turing, who developed the Turing test to separate humans from machines (Mintz and Brodie 2019). After that computing power came to the point of instantaneous calculations and the ability to evaluate new data in real time based on previously evaluated data. AI is a machine intelligence tool that provides many possibilities such as data collection, identifying and selecting alternatives, making it easier to take action, make decisions and review and predict. Recently, AI has begun to be introduced into medicine to improve patient care by speeding up processes and achieving greater accuracy. Radiological images, pathology slides and electronic medical records (EMR) of patients are evaluated with machine learning, helping diagnosis and treatment method of patients, and improving the skills and shorten evaluation time of doctors. In this scope, this study was developed in cooperation with the Department of Computer Engineering and School of Medicine, Department of Pathology.

Next, general information of deep learning is given in section 3.3.1. After mentioned about deep learning, Convolutional Neural Network that used for this study is described in section 3.3.2. Then, main used algorithm that Mask R-CNN is explained in section 3.3.3. ResNet that Mask R-CNN uses for feature extraction is described in section 3.3.4. Finally, two most important contributions of Mask R-CNN for object detection are given in section 3.3.5.

After all these explanations, proposed modification method for small metastasis areas is given section 3.4.

3.3.1. Deep learning algorithms

Deep learning is a machine learning method that was developed from the concept of simulating the human brain (Zou et al. 2008). An example of neural network architecture is shown in Figure 3.4.

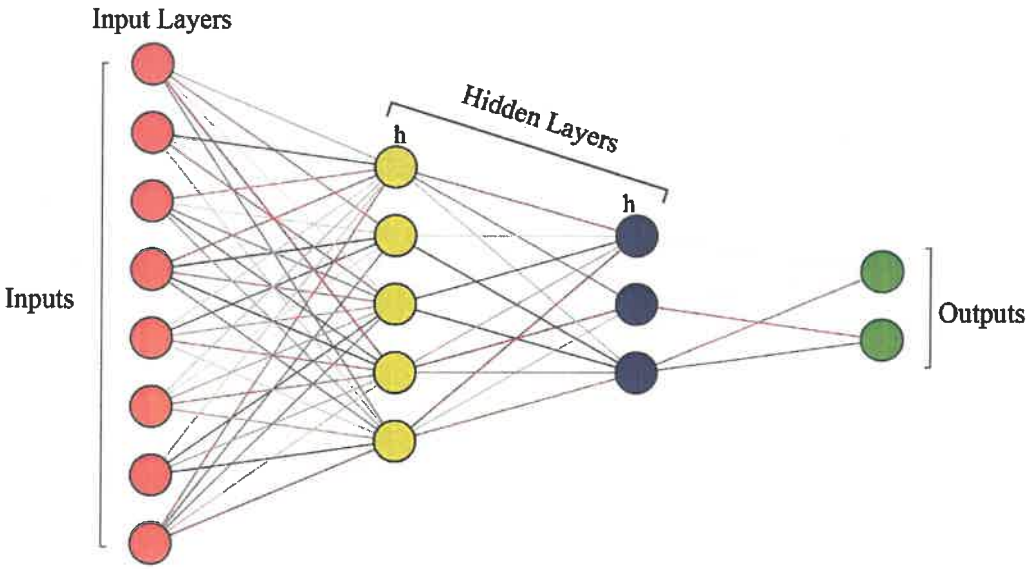

Figure 3.4. An example of neural network architecture

A neural network has well-organized processing units which are input, hidden and output layers.

- The units or nodes in each layer are linked to nodes in the adjacent layers (shown in 3.8).
- There is a weight value for each connection (shown on 3.9).
- The inputs are multiplied by the weights and added together for each layer (shown in 3.9).
- The sums value often undergoes a transformation with ReLU that is a sigmoid function. After sigmoid function applied, some classification results are obtained (shown in 3.5).
- Finally, deep learning algorithm gives a probability ratio of what the object is (shown in 3.5).

There are some popular deep learning algorithms such as CNNs, RNNs, GANs and some others. (Shrestha and Mahmood 2019). More details about studied deep learning algorithm will be given in section 3.4.

3.3.2. Convolutional neural networks

CNN is a deep learning algorithm that uses convolutional layers. This algorithm processes the image using some functional layers. Details of CNN architecture is given in Figure 3.6.

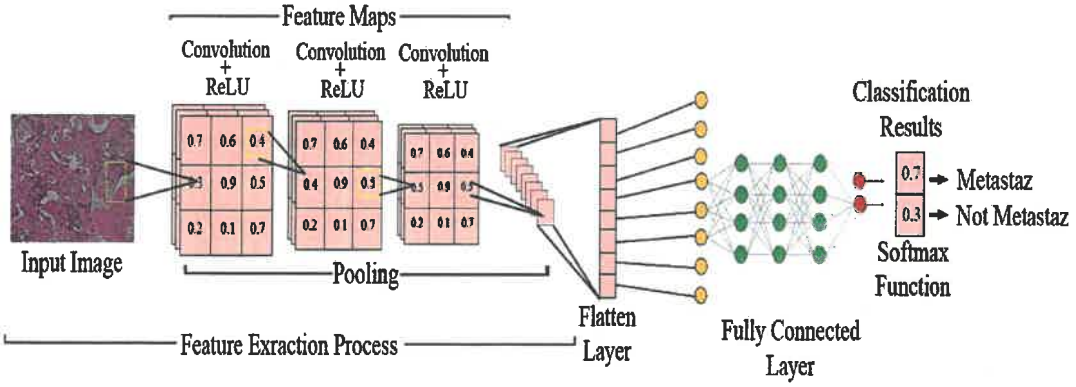

Figure 3.5. Detailed CNN architecture

Convolution layers are used for feature extraction. This layer selects and eliminates the best features and patterns from the input picture and retains information in a matrix. The input image matrix is multiplied elementally with filters and summed to create a characteristic feature map similar to that of the point product between vector combinations.

After all convolutional layers, the non-linearity layer typically is used. This layer is most widely used to enable the neural network to take into account nonlinear relationships. Non-linearity layer is also called as activation layer because of using an activation function. There are some well-known activation functions such as Sigmoid, TanH, ReLU, Leaky ReLU and Softmax. Among these functions, ReLU is the most widely used activation function. A representation of applying non-linear function to feature map is shown in Figure 3.6. ReLU sets all negative values to zero in a given matrix x, and all other values remain unchanged. It is expressed mathematically by (3.1).

$$f(x) = \max(0, x) \qquad (3.1)$$

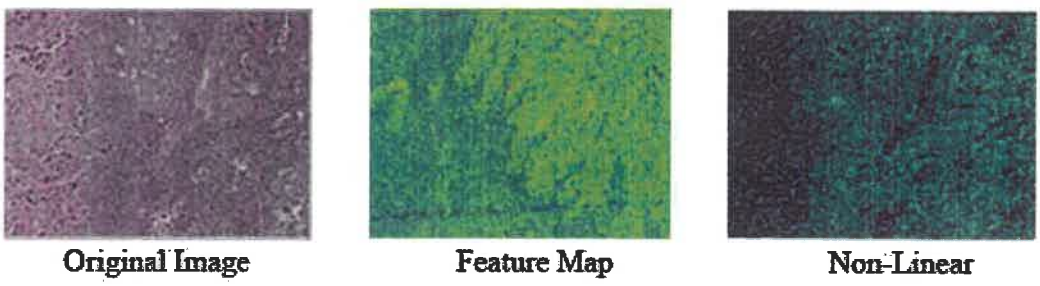

Figure 3.6. A representation of applying Non-linear (ReLU) function to feature map

Pooling layer is frequently added between successive convolutional layers. This layer uses to reduce feature map size, number of measurements, parameters, and calculations in network. Thus, it reduces transaction complexity on the network and prevents wasted time. There are some types of pooling process such as max pooling, min pooling, average pooling. Max pooling is the most widely used pooling method. Shown in Figure 3.7.

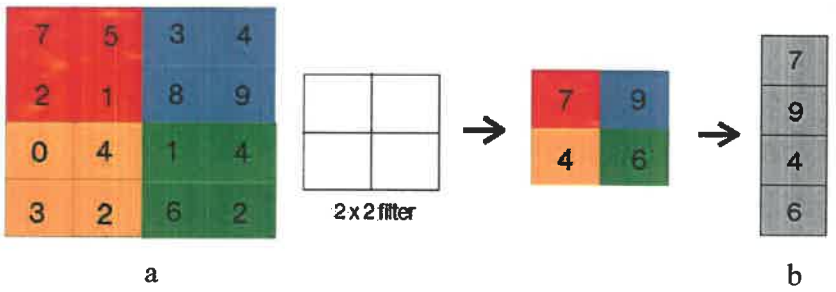

Figure 3.7. Max pooling (a) and Flatting (b) process

The 2 × 2 filter generated is passed over the convolution layer and the largest value in the past region is transferred to the pooling layer (Figure 3.7.a). Flattening layer basically prepares the data to the entrance of the fully connected layer. The data in this neural network is that the matrixes coming from the convolutional and pooling layers are transformed into a one-dimensional array (Figure 3.7.b).

Fully connected layer is the last and the most important layer of ConvNet. This layer takes the data from the flattening process and performs the learning process through neural network. Also, there are some weights and bias values between neurons. An example of neural network architecture is given in Figure 3.8.

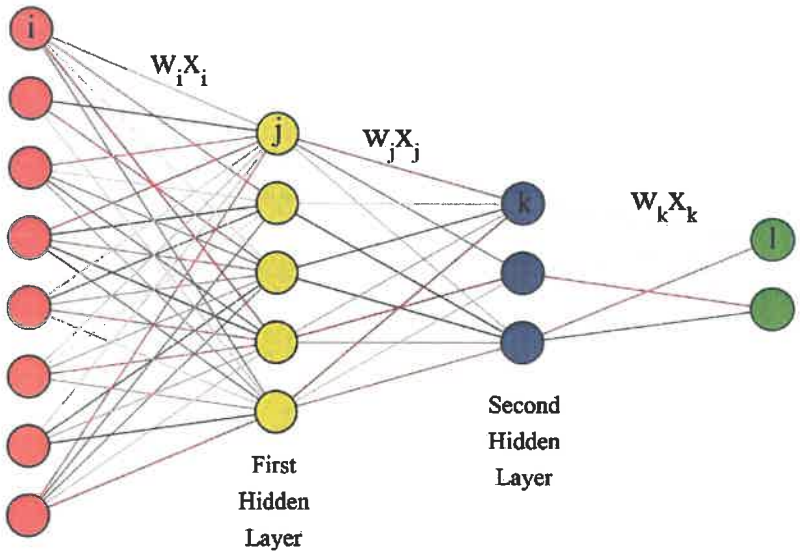

Figure 3.8. An example of fully connected layer (Neural network) architecture

These layers and neurons created by inspiring human nervous system. The mathematical representation of a human nerve cell can be seen in Figure 3.10.

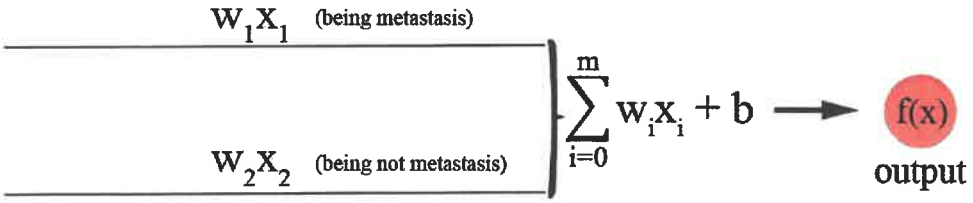

Figure 3.9. Representation of formula of neutral network. x is initial data for the neural network. There are 2 input values that being metastasis or being not metastasis for this study. W(weight) represents how important a node is for the next layer. It can be called as transformation tool for inputs. Because if a input value (x) has higher possibility than other inputs, change of weights take an important role for increasing being possibility of x as final output. b(bias) is an extra parameter in the neural network, along with weight, that is used to change the output as well as the weighted sum of the neuron's inputs. As a result, bias is a constant that makes the model suit better with the given data. $f(x)$ is final output score. Basic operation in a deep learning model as shown on the Formula 3.1 to calculate the w and b parameters that the model will give the best output score.

3.3.3. Object detection algorithms based on CNN

There are some basic and commonly used object detection algorithms which are R-CNN, Fast R-CNN, and Faster R-CNN. The Mask R-CNN algorithm used in this study is based on these algorithms. Basic information about these algorithms is given below.

R-CNN: Input image is first divided into about 2000 zone suggestions and then CNN (ConvNet) is applied for each zone, respectively. The sizes of the regions are determined, and the correct region is placed in the neural network. The biggest problem of this method is time consuming. Since CNN is applied separately to each region in the picture, training time is approximately 84 hours, and the estimation time is around 47 seconds.

Fast R-CNN: Main difference between this algorithm and R-CNN is that Fast R-CNN makes regional recommendations on the covn5 feature map after the CNN process, instead of dividing the picture into regional recommendations as the first process. Also, Fast R-CNN uses softmax classification instead of SVM. Due to using CNN once, training time is decreased significantly. Training time is about 8.75 hours, and the estimated time is approximately 2.3 seconds.

Faster R-CNN: Unlike the Fast R-CNN algorithm, Faster R-CNN algorithm uses the "Region Proposal Network (RPN)" to generate the recommended regions. RPN creates a "sliding window" and moves over the feature map created in the convolution layer. The sliding window could be called as a filter. This filter assumes an object in each zone and assigns neutral scores to the zone. For that purpose, sliding window looks adjacent pixels, colour, density, etc. Then, a new feature map is created using the "ReLU" activation function. Other remaining processes are similar with Fast-RCNN. Faster R-CNN algorithm is shown in Figure 3.10.

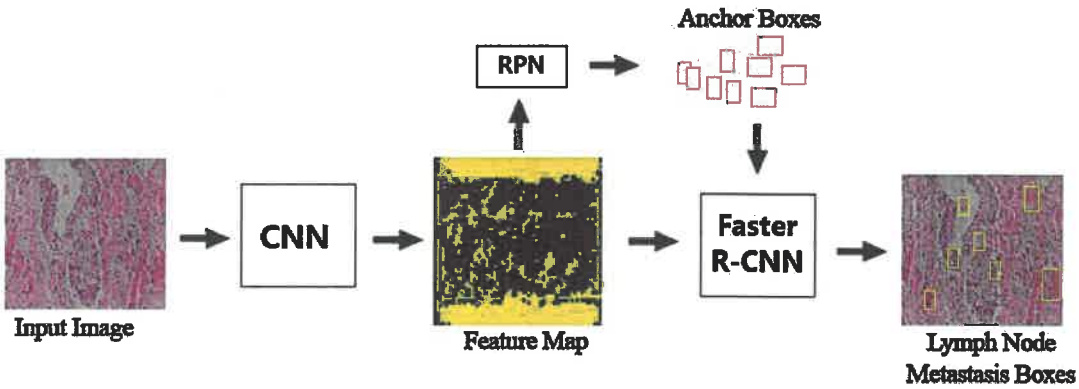

Figure 3.10. Faster R-CNN architecture

Mask R-CNN: This algorithm is extended version of Faster R-CNN which has two outputs that class label and bounding box for each candidate object. The third feature that distinguishes Mask R-CNN from Faster R-CNN is object mask output (He et al. 2017). More details about Mask R-CNN and proposed modification to the model will be given Sections 3.3.5 and 3.4.

3.3.4. ResNet

There are some feature extraction algorithms that are based on CNN. Hyperparameters, layer features and other some features fine-tune modified and summed

with a Net model. There are some well-known models such as AlexNet, VGG, Inception and ResNet. Also, Mask R-CNN algorithm uses ResNet-50 architecture.

ResNet is created by Microsoft engineers and obtained 3.6 % error rate. Considering that the error rate for human beings is around 5%, this is be an impressive rate. In architectures before ResNet, increased success was often attributed to an increase in the number of layers. This idea can be criticized due to the fact that in deep networks, especially during backpropagation, the multipliers decreased too much until they came to the first layers that the effect of these first layers on last layers decreased to almost 0. This idea directly is affected recognition rate. The problem is called as "Vanishing Gradient". ResNet offered a "Residual Network" for solving vanishing gradient (He et al. 2016). Representation of residual block is shown in Figure 3.11.

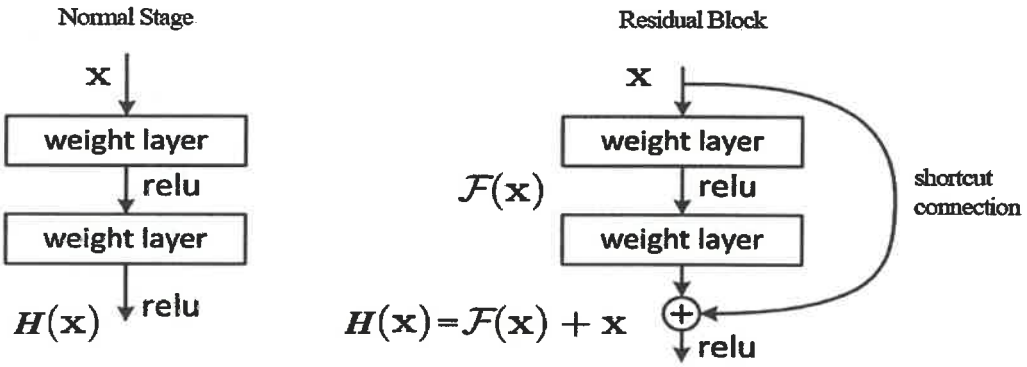

Figure 3.11. Representation of residual block (Adapted from He et al. 2016)

According to Residual Network idea, normally, a layer has an input. This input value leaves the layer and adds into the next layer as a new value, and so on, one after another. The problem is that if the weights are reset to any layer, that layer will no longer have an impact on the result. At this point ResNet defined a shortcut connection (shown in Figure 3.11). Since the input value will be added as input again after 2 layers, the input will not be lost even if it is reset in the next layer, will be added again as input after every 2 layers. Thus, even if the number of layers increases, the initial input will be preserved. Hence, adding more layers to model will not have a negative effect on the result.

The detailed comparison of ResNet and VGG architecture is shown in Figure 3.13. Firstly, the image size initialized as 224 × 224 was first reduced to 112 × 112 with 64 depth size filters (stride: 2), then the image size was reduced to half again with pooling (stride 2), and the structure was obtained 56 × 56 size 64-depth layer. 64 depth layers come from the number of filters in the most recent convolution process. Also, the filter number of one layer indicates the channel number of the next layer. If it is noticed that some skips in the residual blocks are shown with a dashed line (shown in Figure 3.13), this is because the additions here are not made directly as they are in the others. In inserts with dashed lines, the depth between where the input value is and the point to be inserted is different. Depth differences between blocks is shown in Figure 3.12.

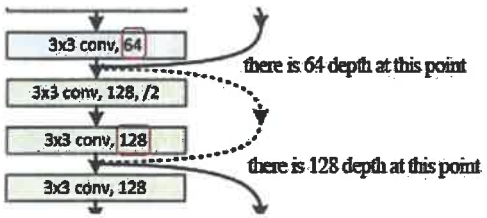

Figure 3.12. Depth differences between blocks

Input that has 64 depth size cannot be directly added to next layer which has 128 depth size. For this to happen, firstly the depth sizes must be equalized. Equalization process has been done using a 1 × 1 filter. When 1 × 1 filters which has 128 depth size were used, the input value can be added to the next point. Thus, the depth size of the input and depth size of the next added point are 128. Comparison between VGG and ResNet models are shown in Figure 3.13.

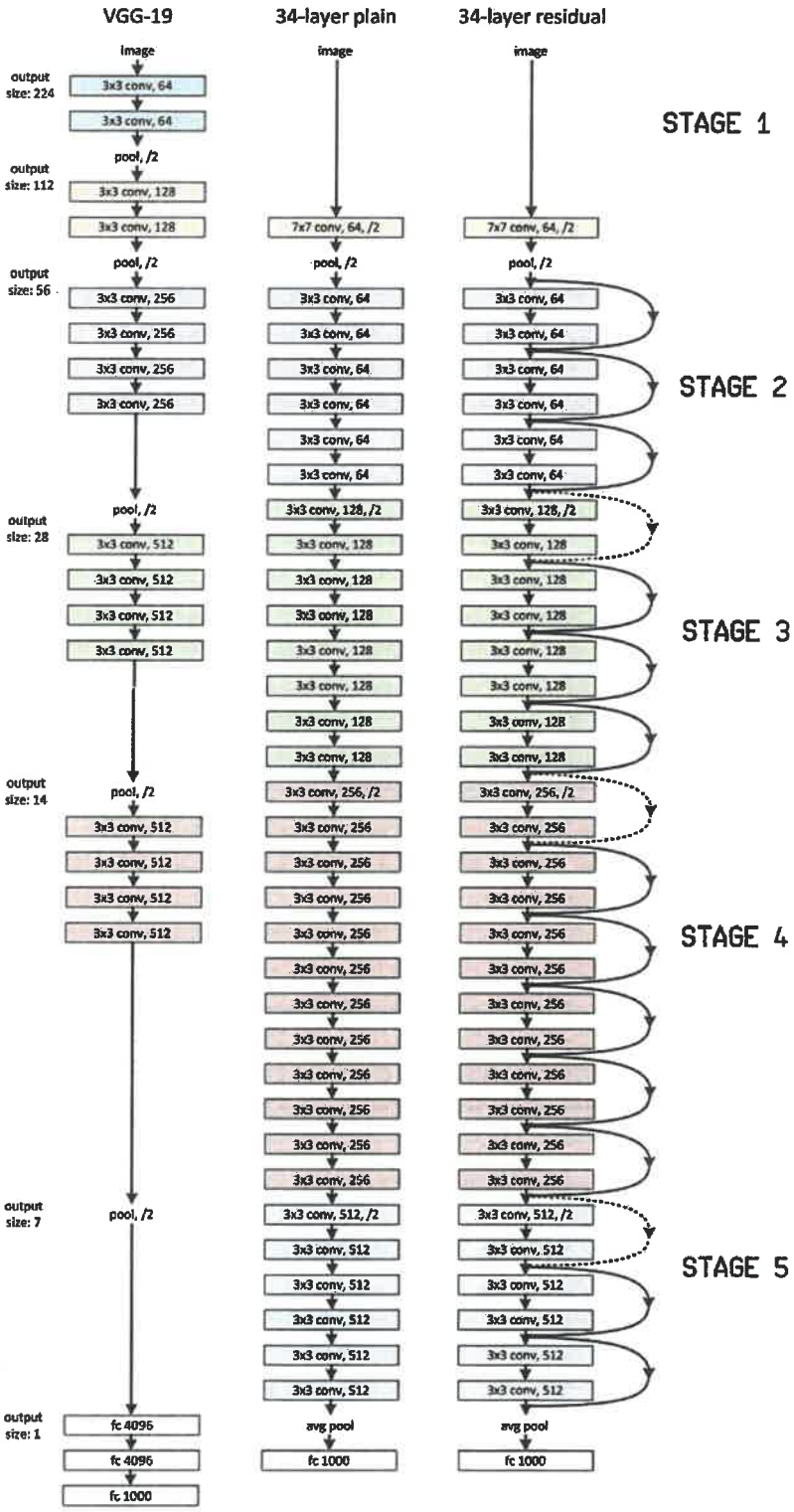

Figure 3.13. Comparison between VGG and Resnet

As a result of the above-mentioned improvement, ResNet achieved very successful result from ILSVRC competitions. Comparison between models on ILSVRC competitions is shown in 3.14.

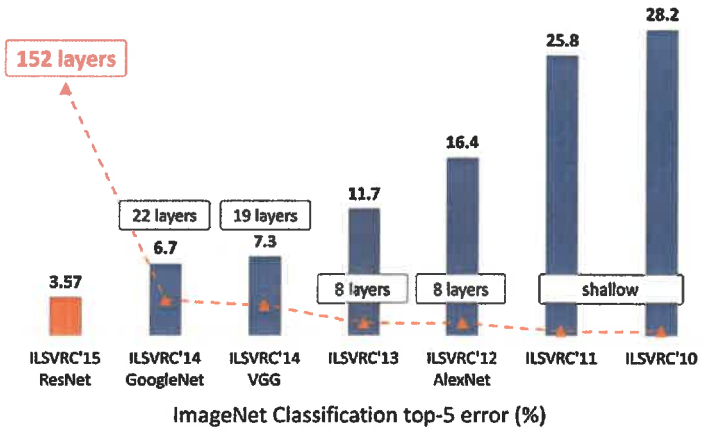

Figure 3.14. Comparison between models on ILSVRC competitions (Hu et al. 2020)

3.3.5. Contribution to the Mask R-CNN

Mask R-CNN algorithm uses Faster R-CNN basically (discussed in 3.3.3). However, Mask R-CNN offers a new method for ROI pooling phase. After obtaining all ROI's, they have a special size (width × height × channel size) for each single of them. ResNet-50 is used model for Mask R-CNN. This model has 224 × 224 × 3 input size. Therefore, ROI's size should be arranged according to the model input size. For instance, ResNet-50 is mapping it into a 16 × 16 × 512 feature map. A scale factor is 32. So, the original image size should be 32 times smaller. While this make smaller process, ROI's size should be 32 times smaller as well for keeping localization.

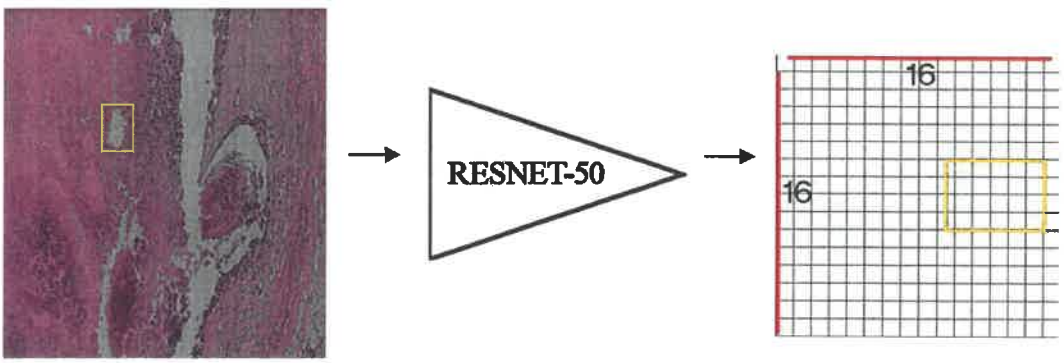

Figure 3.15. ROI make smaller process

Then, original ROI input size is 145 × 200 (shown in Figure 3.15 a). This box size should be divided by 32. ROI replacement and ROI replacement after rounding is shown in Figure 3.16 below.

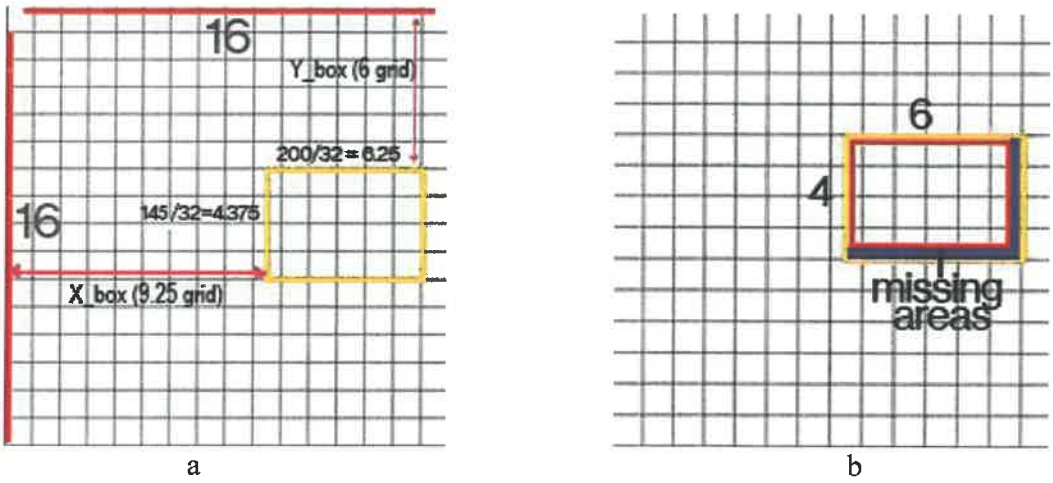

Figure 3.16. ROI replacement **(a)** and ROI replacement after rounding **(b)**

These values (6.25, 4.375) are rounded to 6 and 4, respectively. After rounding, small piece of ROI areas is missed (shown in Figure 3.16 b). So, it effects model performance in bad way because of features loss. Mask R-CNN provides a solution for that problem. ROI box (yellow square on Figure 3.16 a) divided into 9 equal size boxes of original ROI and applied bilinear interpolations within each box. ROI box size and division into boxes are given in Figure 3.17.

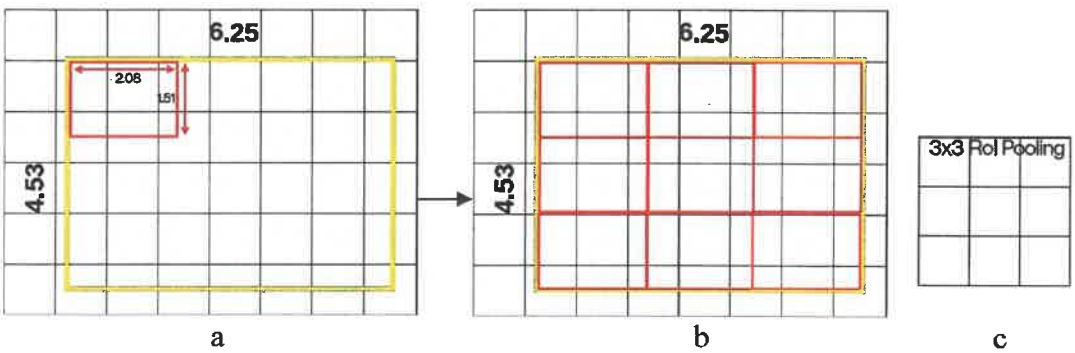

Figure 3.17. ROI box size and dividing into boxes

If top left box (b) is checked, it can be seen that it covers 6 different grids. Extracting value for the pooling layer, it must sample some data from each of 6 grids. So "four sampling points" is used for this purpose as shown in Figure 3.18.

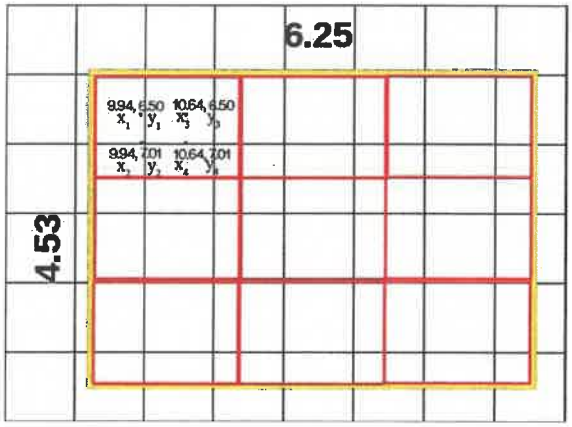

Figure 3.18. Four sampling points

Those 4 sample points are calculated where each of those points should be by diving the height and width of the box by 3 as follows:

$$x_1 = x_{box} + \left(\frac{width}{3}\right) \times 1 = 9.94 \quad (3.2)$$

$$y_1 = y_{box} + \left(\frac{height}{3}\right) \times 1 = 6.50 \quad (3.3)$$

$$x_2 = x_{box} + \left(\frac{width}{3}\right) \times 1 = 9.94 \quad (3.4)$$

$$y_2 = y_{box} + \left(\frac{height}{3}\right) \times 2 = 7.01 \quad (3.5)$$

where x is x coordinate values of four sampling points for each grid, y is y coordinate values of four sampling points for each grid, x_{box} is initial x coordinate value of feature box (shown on 3.18. a), y_{box} is initial y coordinate value of feature box (shown on 3.18. a), *height* is height of feature box, *width* is width of feature box.

After all the calculations made for all points using the formulas (3.3), (3.4), (3.5), (3.6) shown above, bilinear interpolation can apply to sample data for this box. Bilinear Interpolation for the first point is shown in Figure 3.19. Bilinear interpolation is popular method for image processing to sample colours.

$$P \approx \frac{y_2 - y}{y_2 - y_1}\left(\frac{x_2 - x}{x_2 - x_1}Q_{11} + \frac{x - x_1}{x_2 - x_1}Q_{21}\right) + \frac{y - y_1}{y_2 - y_1}\left(\frac{x_2 - x}{x_2 - x_1}Q_{12} + \frac{x - x_1}{x_2 - x_1}Q_{22}\right) \quad (3.6)$$

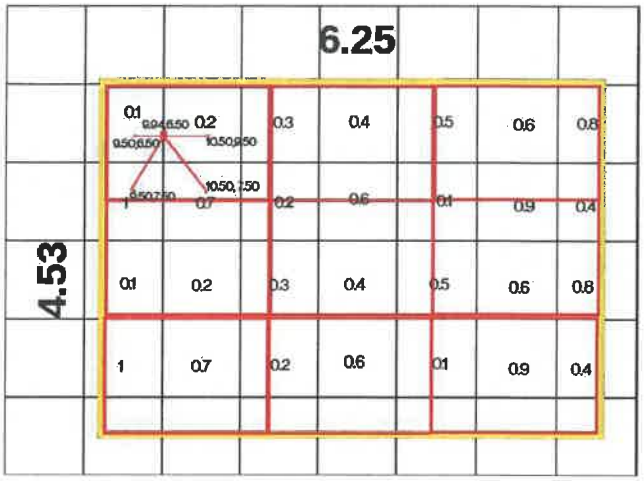

Figure 3.19. Bilinear Interpolation for the first point

Formula for the first point is calculated by the formula (3.8).

$$P \approx \frac{7.5-6.5}{7.5-6.5}\left(\frac{10.5-9.94}{10.5-9.5}0.1+\frac{9.94-9.5}{10.5-9.5}0.2\right)+\frac{6.5-6.5}{7.5-6.5}\left(\frac{10.5-9.94}{10.5-9.5}1+\frac{9.94-9.5}{10.5-9.5}0.7\right) \quad (3.7)$$

This calculation is done for other 8 ROI boxes as well and obtained weights shown in Figure 3.20.

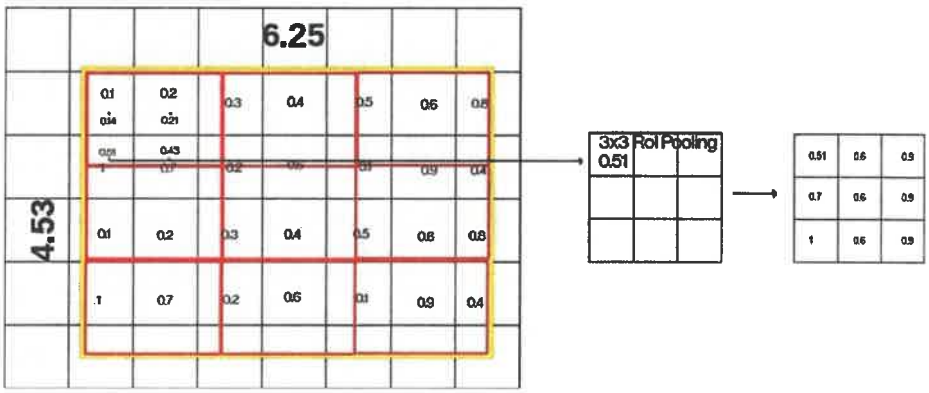

Figure 3.20. Max pooling processing for four sample points

Max-pooling applies and selects the highest weights between those 4 sampling dots (shown in Figure 3.20). It is done for all 8 remaining boxes and fills 3 × 3 ROI pooling. Unlike rounding, maximum performance is achieved by processing the values to the right of the comma. (6.25, 4.53). This method is most important contribution of Mask R-CNN for object detection. General concept of Mask-RCNN is shown in Figure 3.21.

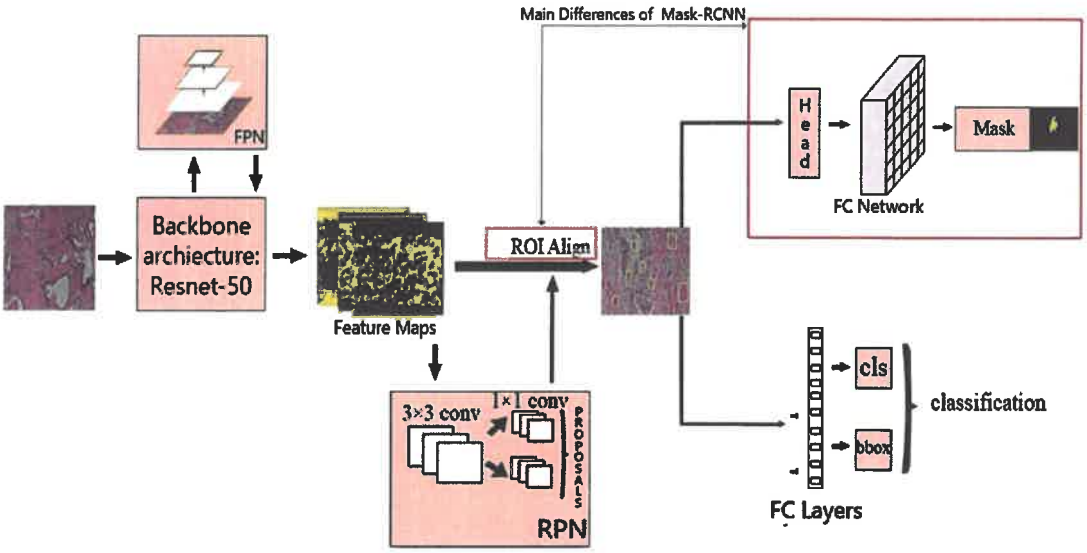

Figure 3.21. General concept of Mask R-CNN

After ROI align and pooling process, collected pools goes to FC layers and gives some classification result with masks (Instance segmentation). Besides of ROI alignment there is another contribution of Mask R-CNN which is head and mask branch. It creates mask boxes for diagnosed areas (shown in Figure 3.21).

3.4. Improvement on ResNet for better detection of small tumor areas

ResNet is main model behind of Mask R-CNN algorithm. If the ResNet architecture is examined carefully, it can be seen that it consists of 5 different stages with different layer depths. It produces results by shrinking ROI regions in multiples of each stage 2 (stride 2^n). FPN architecture on ResNet is shown in Figure 3.22.a and b.

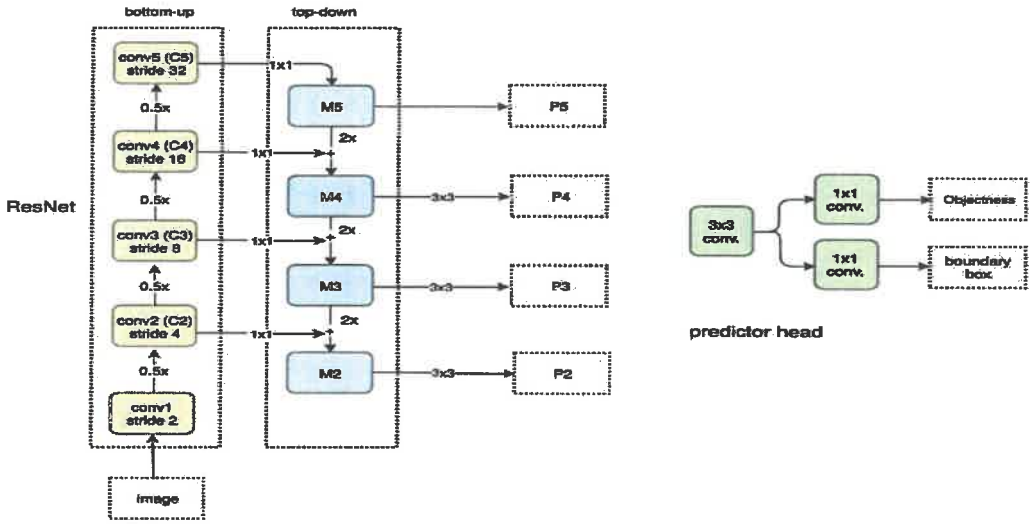

Figure 3.22. a. ResNet Stages

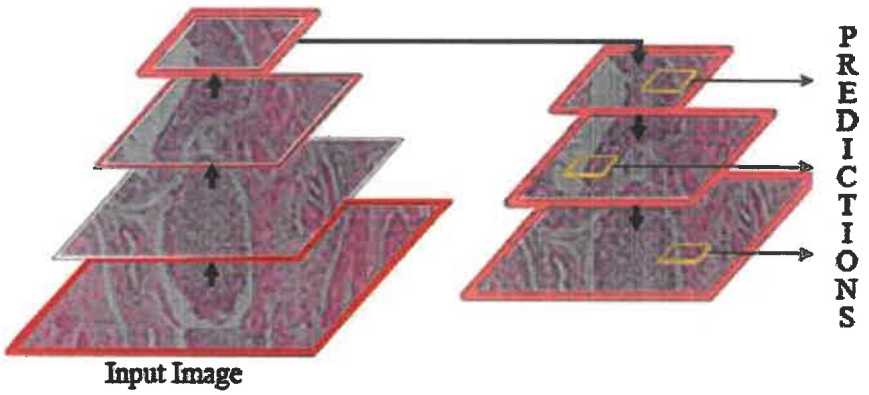

Figure 3.22. b. FPN Architecture

In this study, foci of lymphnode metastasis of urothelial carcinoma were detected with a significantly higher accuracy by the original Mask R-CNN algorithm. Some test results for cancer tumors are as shown in Figure 3.23 below.

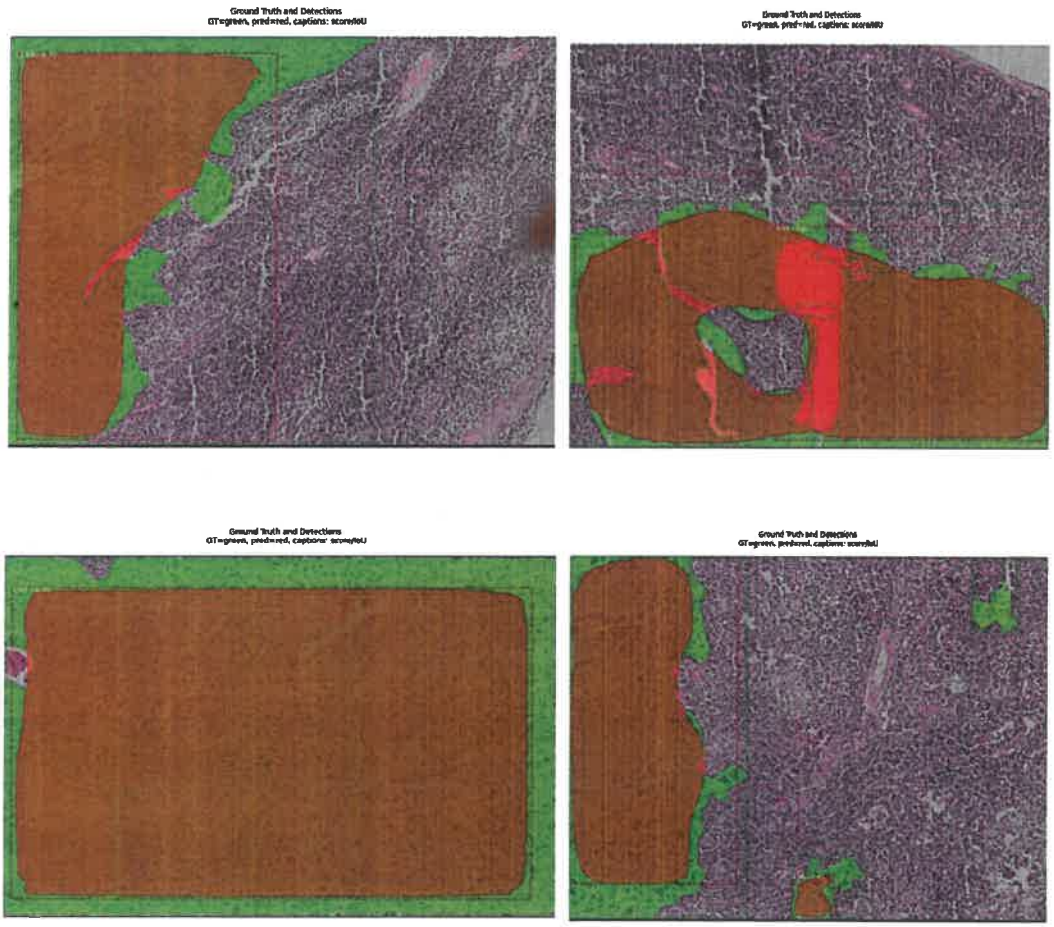

Figure 3.23. Some test results on the bladder cancer tumors (Green area: Ground Truth Dark Orange area: Predictions made by model)

If we evaluate all figures given in Figure 3.23, the model appears to work well in large tumor areas, but has difficulty especially recognizing small areas. This study proposes a method modification for solving the recognition problem in micro metastasis foci.

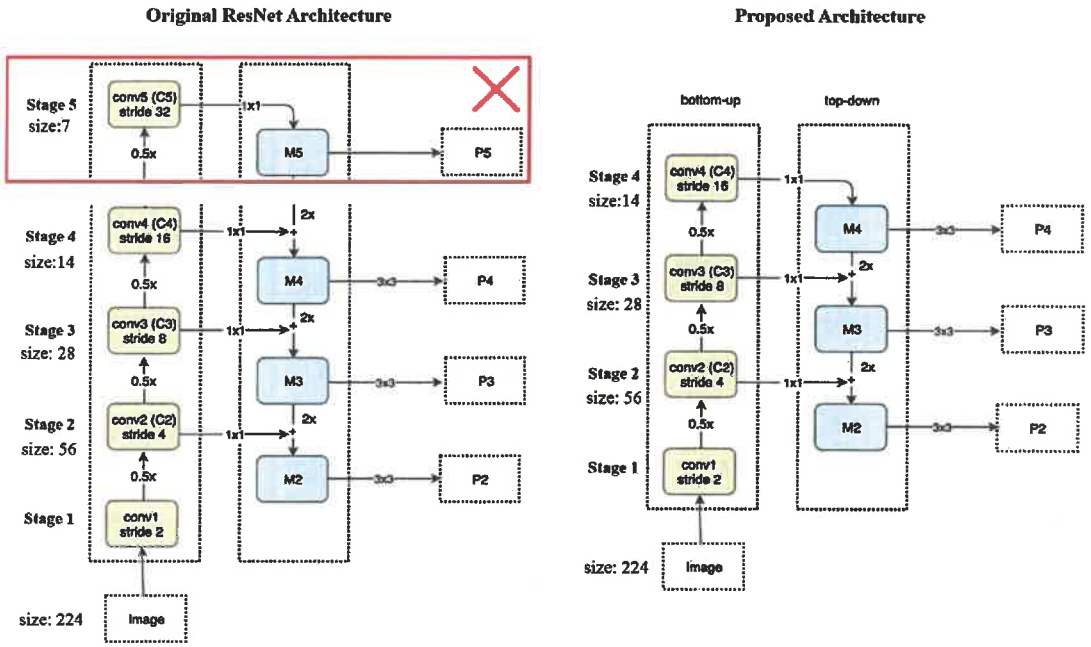

Figure 3.24. Removing stage 5 from the model.

According to ResNet discussed in Section 3.3.4, the size of the input image starts with 224 × 224 and down sampled to 7 × 7 at the last stage by striding (shown in Figure 3.24). Since the main purpose of this study is the better detection of small tumor areas, if the input image size shrinks too much, this process causes to lose features for small target regions and decreases detection performance of such small tumor areas. Therefore, the fifth stage was removed from the model. Thus, final image size became 2 times larger than original model. Also, each of the stages consists of different number of feature extraction map layers. While the fifth stage was removed, its layer blocks were removed as well. This removing process caused to decrease performance of diagnosing because of feature maps loss. Therefore, same number of layer blocks of stage 5 were added to stage 4. Thus, the model's layer depth was preserved. An important point is that feature map depth of the added extra layer blocks are same as Stage 4 (3 × 3 conv, 256) and also not applied any extra striding (shown in Figure 3.25). In other training trials, in addition to fifth stage, fourth stage was removed to observe the effect of removing more stages on the test results. These testing results were not satisfactory (will be shown on Section 4.2).

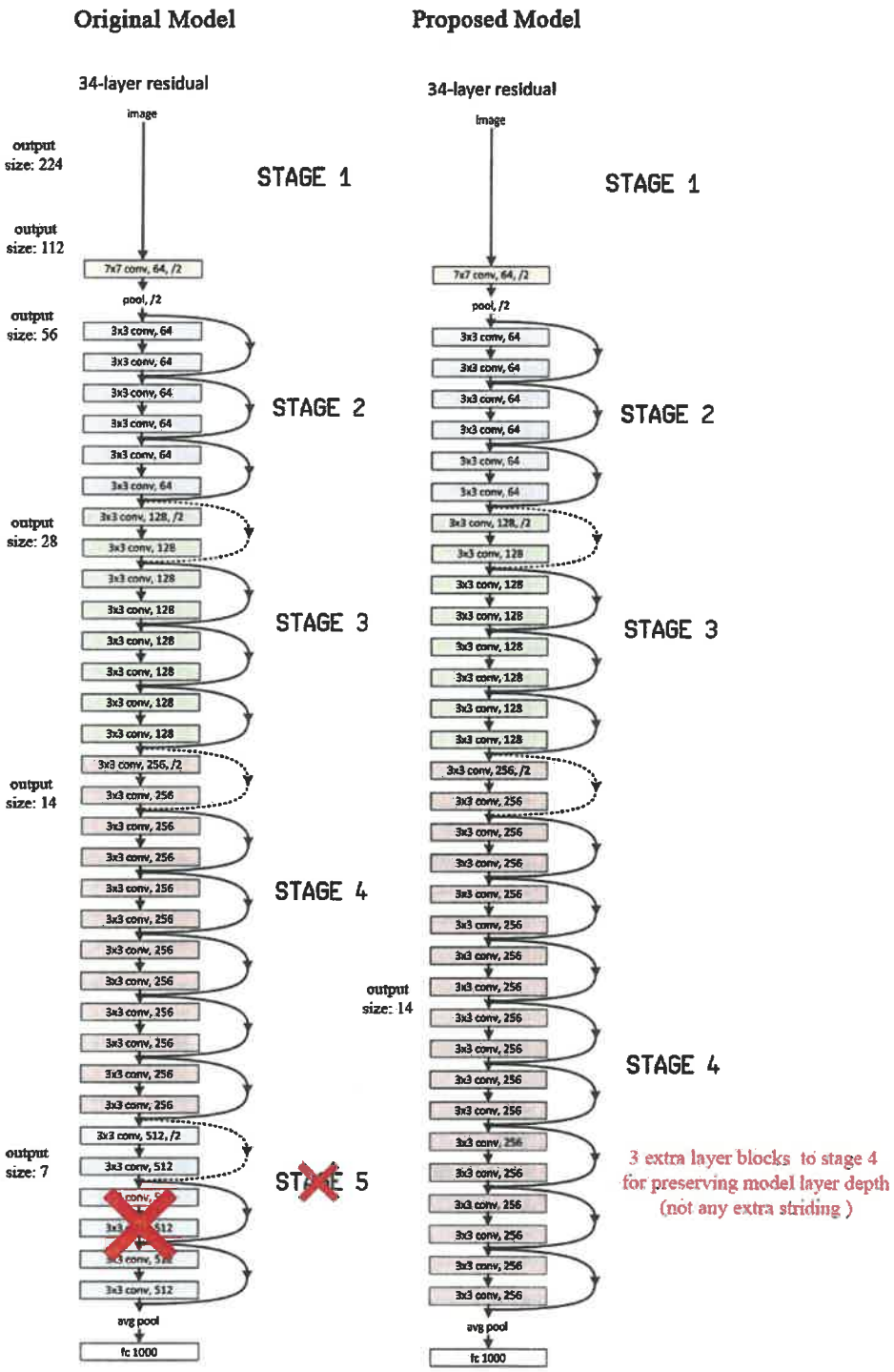

Figure 3.25. Proposed model architecture for better detection of small tumor areas

4. RESULTS AND DISCUSSION

4.1. Experimental Details

Developing artificial intelligence applications needs powerful hardware. Training a network may take hours and even days as the dataset grows. However, these training times can be reduced considerably with a powerful GPU. In this study, Google's Cloud Artificial Intelligence Development System Google Colab was used. Google Colab gives great number of GPU and cores to developers. Thus, training periods of models are shortened significantly. Table 4.1 shows the System features provided by Google Colab.

Table 4.1. System features provided by Google Colab (Anonymous 5)

GPU Name	Nvidia Tesla V100 GPU
GPU Memory	16 GB RAM
Clock Rate	1530 MHz
Tensor Core Size	640

In this study, Tensorflow and Keras open-source artificial intelligence libraries were used for creating and developing for deep learning model which is ResNet. These libraries provide ready-to-use codes for building deep learning models. This makes it easier to create and modify a deep learning model (Anonymous 6, 7).

While images labelling (discussed on Section 3), VGG Image Annotator program was used for this purpose. It is a fast and independent picture, audio and video manual annotation program developed by Department of Engineering Science, University of Oxford. It works in a web browser without downloading or installing. VGG Image Annotator is also program that open source and enables both the academic and business application to be useful. This program gives annotations as a CSV or JSON file format. JSON file format was used for this study.

All training trials were done with 50 epoch and 0.9 learning momentum. Learning rate and batch size were 1.10^{-3} and 2, respectively. ReLU was used as the activation function while performing convolution process. In addition, another function used in classification process is Softmax.

4.2. Experiment Results

Firstly, deep learning algorithms need a huge data set because of it cannot do feature extraction itself. The data set used in this study was not taken from huge size databases involved of thousands of pictures such as Kaggle, MNIST etc. A relatively small data set consisting of 365 samples created by ourselves was used. Therefore, even with the best model optimization, it is not possible to achieve very high success rates as in other studies. As a result, the accuracy values found were evaluated between the original model and the modified model.

There are some metrics for calculating accuracy ratios on deep learning models. Also ground-truth definition should be known. The targeted values in the test data set are called ground truth. Figure 4.1 shows the presentation of ground truth and predicted area by the model.

Figure 4.1. Presentation of ground truth and predicted area by the model

In this study, there are two classes that are either being a cancerous cell region or not. While image annotating, unmarked areas were defined as tumor free cells. Also, following 4 metrics were used to calculate confusion matrix

TP: who is sick is called sick.

FP: who is not sick is called sick.

TN: who is not sick is called not sick.

FN: who is sick is called not sick.

In this study, the Mask R-CNN algorithm yielded satisfactory results for large tumor areas. The focus of this study was better detection of small tumor areas. There has been an improvement in the diagnosis of small tumor areas with the model modification described in the method section.

According to the testing results, the proposed method modification increased accuracy of small TP and TN tumor areas. Also, FP tumor areas were decreased. Thus, false diagnosed tumor areas were reduced as well. The comparison of the modified model results with the original model are shown in Figure 4.2.a and b.

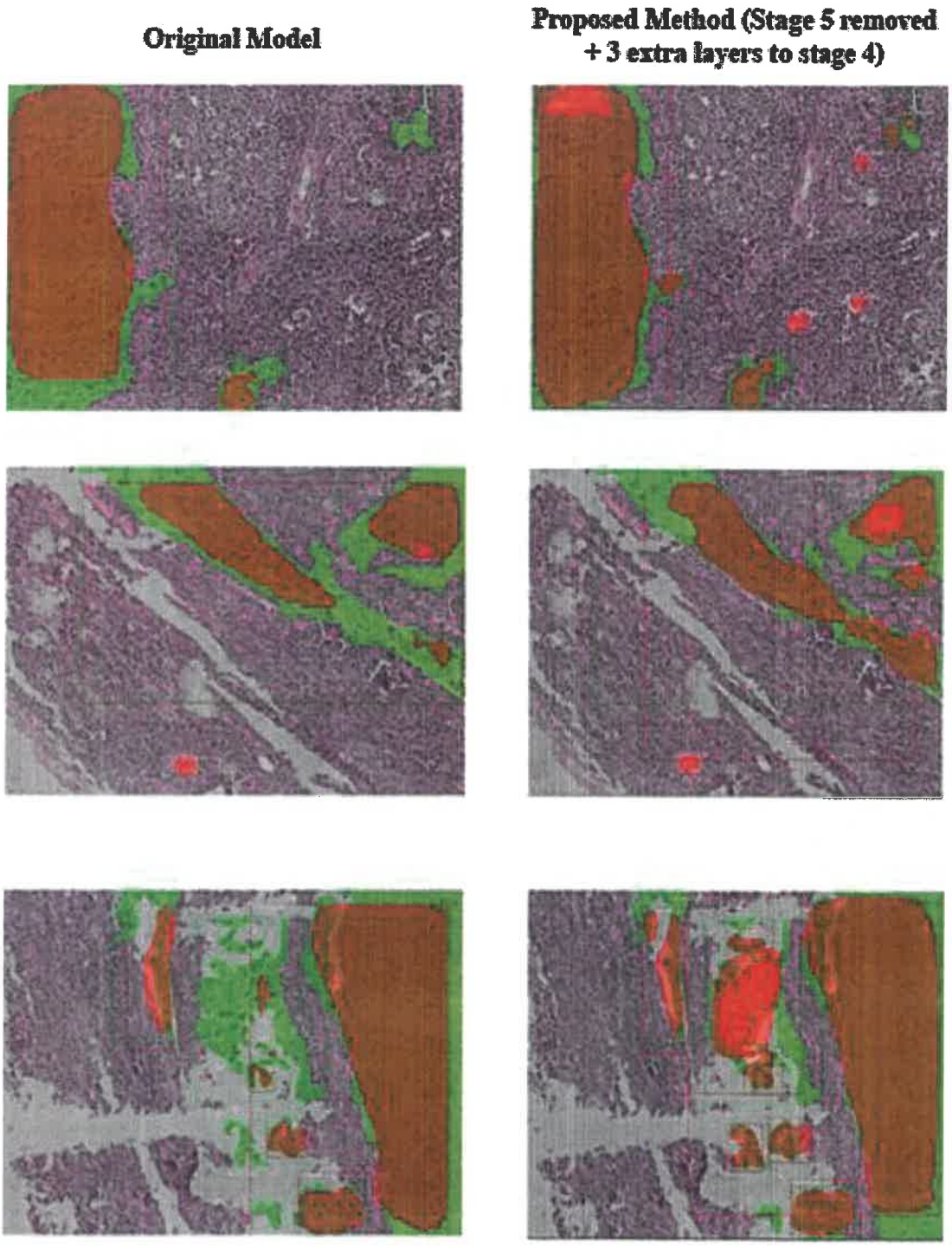

Figure 4.2 a. Test results comparison

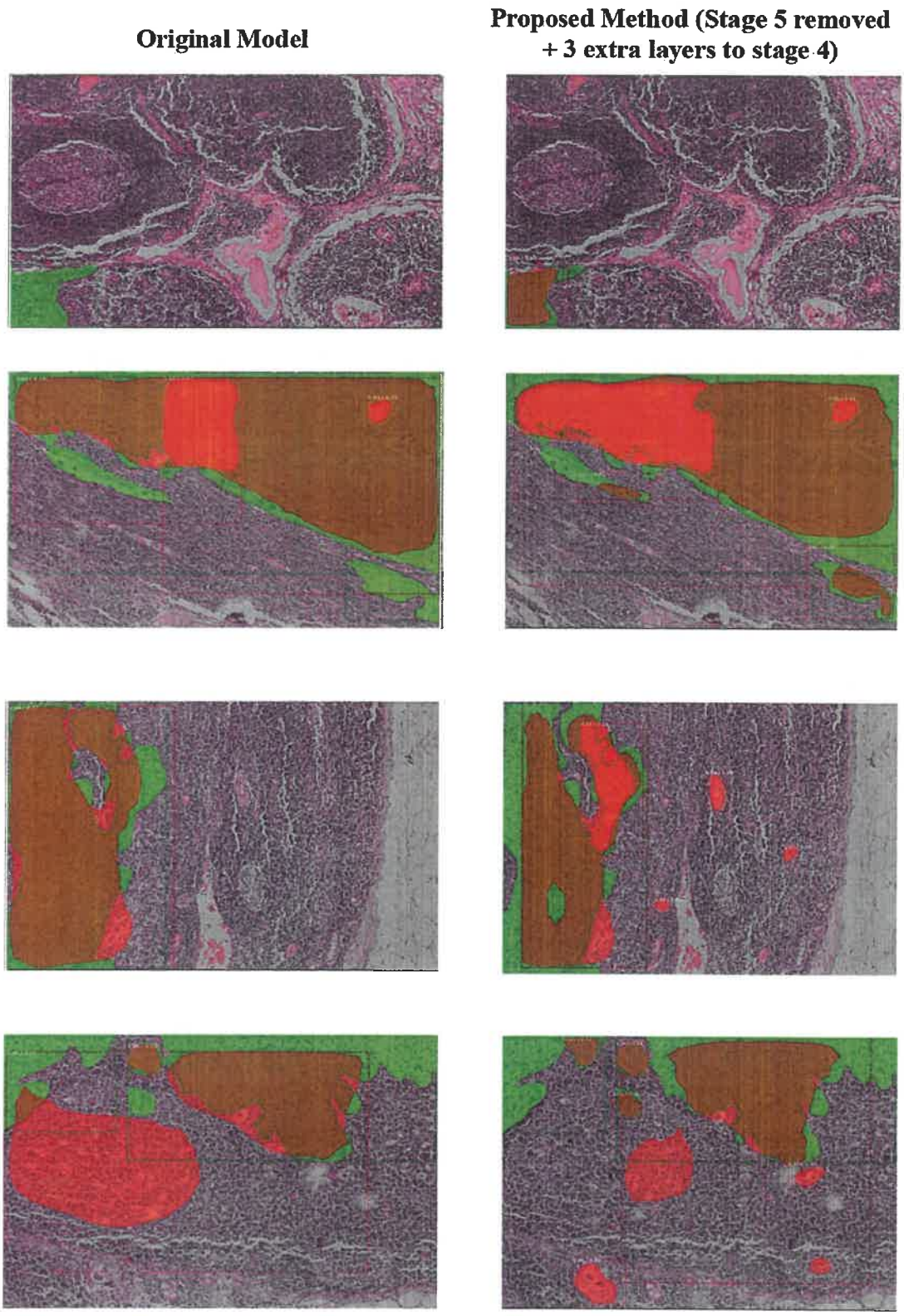

Figure 4.2 b. Test result comparison

According to testing results (shown in Figure 4.2 a and b), annotated small tumor areas detection rate increased. Also, there are some detected tumor areas which were not annotated by pathologists (shown with red color in Figure 4.2 a and b). Due to challenging and intensive working conditions, pathologists may overlook small tumor areas that require more attention. Proposed modification will help to doctors for this problem. Finally diagnostic time of pathologist will be reduced and also, they will be able to start to treat their patients faster.

In this study, main purposed idea is removing Stage 5 and adding 3 extra layer blocks to Stage 4 for obtaining better detection results on small tumor areas (discussed on Section 3). In addition to the fifth stage, fourth stage was deleted to monitor the effect on the test results of removing other stages. Recognition of small and large tumor areas rate significantly decreased because of reducing of model layers depth, stage, and feature map size. Testing results are shown in Figure 4.3.

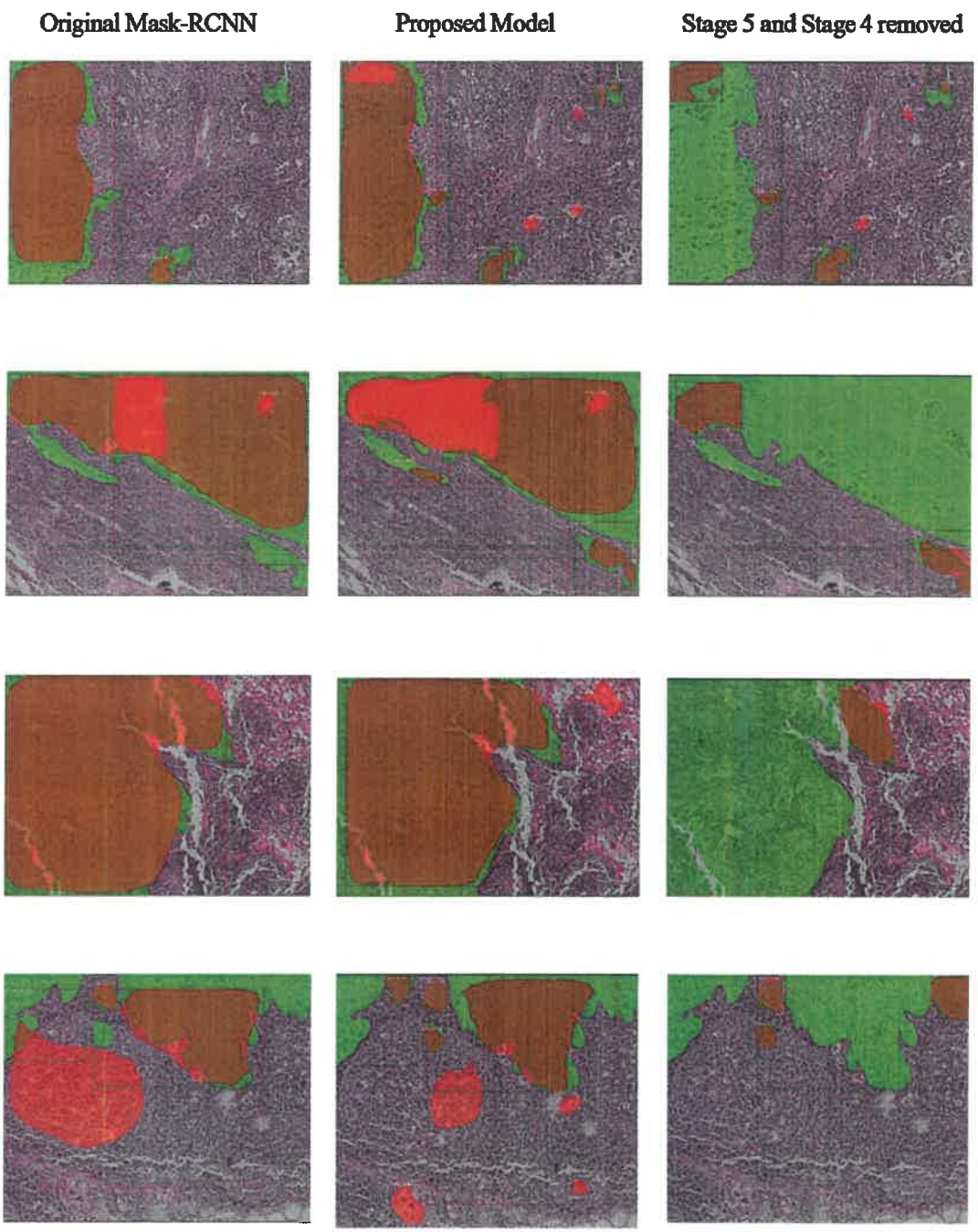

Figure 4.3. Test results comparison between main model, proposed model modification and stage 5 and 4 removed model

5. SUMMARY AND CONCLUSIONS

Today, deep learning has commonly started to use in the health field as in other fields. Dramatic rise in patients and diseases has negative impacts on doctors' workforce. Cancer is an important diagnostic area in medicine. Bladder cancer is a type of cancer which has an increasing incidence day by day. In this study, we have done a model modification to already existing a deep learning model which is able to detect lymph node metastasis of bladder cancer.

Mask R-CNN algorithm was used and tested for diagnosis of bladder cancer tumor areas. The main contribution was to increase the detection rate of small tumor areas by modifying the ResNet model used inside Mask R-CNN. There were 5 stages in the convolutional process of the ResNet. Feature extraction is performed by reducing the size of the detected ROIs at each stage. Since the size and resolution of the ROIs decreases with each stage, it caused to lose small tumor areas on extracted feature maps. Hence last stage of the original model which is Stage 5 was removed from the model. Aim of the removing stage 5 is to keep image size bigger and reduce the loss of feature maps in small areas. While removing last stage, model's layers depth was decreased. Due to preserving the model layer depth, extra 3-layer blocks (same as Stage 5-layer depth size) are added to Stage 4. However, depth of feature maps of this extra 3-layer blocks have not been changed and also has not been applied any extra striding. Finally, small tumor area detections were improved comparing to original model.

To sum up, the results obtained with this study have been promising for doctors and their clinical studies. Due to overworking hours and heavy working conditions and large number of patients, pathologists can miss small tumor areas even if they easily can recognize large tumors. Proposed deep learning model modifications were improved for that purpose. Thus, great convenience will be provided to doctors with detecting especially small lymph node metastasis areas of bladder cancer using artificial intelligence and deep learning. In future studies, more satisfactory and striking results will be obtained with a model produced only for the detection of small tumor areas with a larger dataset.

6. REFERENCES

Angermueller, C. Pärnamaa, T., Parts, L., and Stegle, O. 2016. Deep learning for computational biology. *Molec. Syst. Biol.* 12(7): 878.

Anonymous 1: About Bladder Cancer. https://www.cancer.org/cancer/bladder-cancer/about/ what-is-bladder-cancer [Last access date: 06.01.2021].

Anonymous 2: Overview Bladder Cancer. https://www.nhs.uk/conditions/bladder-cancer/ [Last access date: 06.01.2021].

Anonymous 3: Bladder Cancer: Symptoms and Signs. https://www.cancer.net/ cancer-types/ bladder-cancer/symptoms-and-signs [Last access date: 06.01.2021].

Anonymous 4: Trust the Color – Olympus BX53 microscope. https://www.olympus-europa.com/company/en/news/press-releases/olympus-bx53-microscope.html [Last access date: 28.12.2020].

Anonymous 5: Cloud GPU's. https://cloud.google.com/gpu [Last access date: 07.1.2021].

Anonymous 6: TensorFlow. https://www.tensorflow.org/. [Last access date: 07.1.2021].

Anonymous 7: Keras, simple, flexible, powerful. https://keras.io. [Last access date: 07.1.2021].

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J. Clin*, *68*(6), 394–424.

Breto, A.L., Zavala-Romero, O., Asher, D., Baikovitz, J.B., Ford, J., Stoyanova, R. and Portelance, L. 2019. A deep learning pipeline for per-fraction automatic segmentation of GTV and OAR in cervical cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 105: S202.

Cha, K.H., Hadjiiski, L., Chan, H.P., Weizer, A.Z., Alva, A., Cohan, R.H., Caoili, E.M., paramagul, C. and Samala, R.K. 2017a. Bladder cancer treatment response assessment in CT using radiomics with deep-learning. *Sci. Rep.* 7(1): 8738 1-12.

Cha, K.H., Hadjiiski, L.M., Chan, H.P., Samala, R., Cohan, R.H., Caoili, E.M., Paramagul, C., Alva, A. and Weizer, A.Z. 2017b. Bladder cancer treatment response assessment using deep learning in CT with transfer learning. *Proceedings SPIE Vol. 10134, Medical Imaging 2017: Computer-Aided Diagnosis.* International Society for Optics and Photonics, 2017, Orlando, Florida, United States.

Dolz, J., Xu, X., Rony, J., Yuan, J., Liu, Y., Granger, E., Desrosiers, C., Zhang, X., Ayed, I.B. and Lu, H. 2018. Multi-region segmentation of bladder cancer structures in MRI with progressive dilated convolutional networks. *Med. Phys.* 45 (12): 5482-5493.

Dong, Z., Du, X. and Liu Y. 2020. Automatic segmentation of left ventricle using parallel end-end deep convolutional neural networks framework. *Knowl.-Based Syst.* 204: 106210.

Du, Y., Yang, R., Chen, Z., Wang, L., Weng, X. and Liu, X. 2020. A deep learning network-assisted bladder tumor recognition under cystoscopy based on Caffe deep learning framework and EasyDL platform. *Int. J. Med. Robot.* 2020: e2169.

Dutta, A. and Zisserman, A. 2019. The VIA annotation software for images, audio and video. In Proceedings of the 27th ACM International Conference on Multimedia (MM '19), October 21–25, 2019, Nice, France. ACM, New York, NY, USA, 2276-2279.

Gao, X.W., Hui, R. and Tian Z. 2017. Classification of CT brain images based on deep learning networks. *Comput. Methods Progr. Biomed.* 138: 49-56.

Garapati, S.S., Hadjiiski, L., Cha, K.H, Chan, H.P., Caoili, E.M., Cohan, R.H., Weizer, A., Alva, A., Paramagul, C., Wei, J. and Zhou, C. 2017. Urinary bladder cancer staging in CT urography using machine learning. *Med. Phys.* 44 (11): 5814-5823.

Gordon, M., Hadjiiski, L., Cha, K., Chan, H.P., Samala, R., Cohan, R.H. and Caoili, E.M. 2017. Segmentation of inner and outer bladder wall using deep-learning convolutional neural network in CT urography. *Proceedings SPIE Vol. 10134, Medical Imaging 2017: Computer-Aided Diagnosis.* International Society for Optics and Photonics, Orlando, Florida, United States.

Hamet, P. and Tremblay, J. 2017. Artificial intelligence in medicine. *Metabolism* 69: S36-S40.

Hao, Y., Usame, M., Yang, J., Hossain, M.S. and Ghoneim, A. 2019. Recurrent convolutional neural network based multimodal disease risk prediction. *Future Gener. Comput. Syst.* 92: 76–83.

He, K., Gkioxari, G., Dollár, P. and Girshick, R 2017. Mask R-CNN. *Proceedings of the IEEE İnternational Conference on Computer Vision (ICCV)*. Venice, Italy, 22-29 Oct 2017, 1: pp. 2980-2988.

He, Y., Wang, G., Zhu, S., Chen, H., Zhang, A. and Xu, Z. 2018. Ghost imaging based on deep learning. *Sci. Rep.* 8: 6469.

He, Z., Zhang, X., Ren, S. and Sun J. 2016. Deep residual learning for image recognition. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), Las Vegas, NV, USA, 27-30 June 2016, pp.770-778., DOI Bookmark: 10.1109/CVPR.2016.90.

Hu, H., Lyu, J., and Yin, X. 2020. Research and prospect of image recognition based on convolutional neural network. *Journal of Physics: Conference Series.* 1574:012161, IOP Publishing. doi:10.1088/1742-6596/1574/1/012161.

Iqbal, M.S., Ahmad, I., Bin, L., Khan, S. and Rodrigues, J.J. 2020. Deep learning recognition of diseased and normal cell representation. *Trans. Emerg. Telecommun. Technol.* e4017.

Iqbal, M.S., el-Ashram, S., Hussain S., Khan, T., Huang, S., Mehmood, R. and Luo, B. 2019. Efficient cell classification of mitochondrial images by using deep learning. *J. Opt.* 48 (1): 113-122.

Jansen, I., Lucas, M., Bosschieter, J., de Boer, O.J., Meijer, S.L., van Leeuwen, T.G., Marquering, H.A., Nieuwenhuijzen, J.A., de Bruin, D.M., Savci-Heijink, C.D.

2020. Automated detection and grading of non–muscle-invasive urothelial cell carcinoma of the bladder. *Am. J. Pathol.* 190: 1483-1490.

Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D. and Glocker, B. 2017. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* 36: 61-78.

Khamis, H., Zurakhov, G., Azar, V., Raz, A., Friedman, Z. and Adam, D. 2017. Automatic apical view classification of echocardiograms using a discriminative learning dictionary. *Med Image Anal.*, 36: 15-21.

Khosravi, P., Kazemi, E., Imielinski, M., Elemento, O. and Hajirasouliha, I. 2018. Deep convolutional neural networks enable discrimination of heterogeneous digital pathology images. *EBioMedicine,* 27: 317-328.

Krishnakumar, B., Kousalya, K., Mohana, R.S., Dinesh, K. and Santhiya, S. 2019. Classification of Breast Cancer using Deep Learning Architecture. *Int. J. Recent Technol. Eng.* 8 (4): 7451-7454.

Kumar, S., Asthana, R., Upadhyay, S., Upreti, N. and Akbar, M. 2020. Fake news detection using deep learning models: A novel approach. *Trans. Emerg. Telecommun. Technol.* 31: e3767: 1–23.

Kurnianingsih, Allehaibi, K.H.S., Nugroho, L.E., Widyawan, Lazuardi, L., Prabuwono, A.S. and Mantoro, T. 2019. Segmentation and classification of cervical cells using deep learning. *IEEE Access* 2019, 7: 116925–116941.

Lee, C.M., Park, S.S. and Kim, J.H. 2015. Current status and scope of lymph node micrometastasis in gastric cancer. *J. Gastric Cancer* 15 (1): 1–9.

Liu, Y., Zhang, P., Song, Q., Li, A., Zhang, P. and Gui, Z. 2018. Automatic segmentation of cervical nuclei based on deep learning and a conditional random field. *IEEE Access,* 6: 53709–53721.

Liu, M., Zhou, Z., Shang, P. and Xu, D. 2020. Fuzzified image enhancement for deep learning in iris recognition. *IEEE Trans. Fuzzy Syst.* 28: 92-99.

Ma, X., Hadjiiski, L.M., Wei, J., Chan, H.P., Cha, K.H., Cohan, R.H., Caoili, E.M., Samala, R., Zhou, C. and Lu, Y. 2019. U-Net based deep learning bladder segmentation in CT urography. *Med. Phys.* 46(4): 1752-1765.

Mechria, H., Gouider, M.S. and Hassine, K. 2019. Breast cancer detection using deep convolutional neural network. *Proceedings of the 11th International Conference on Agents and Artificial Intelligence*, Vol 2: ICAART, Prague, Czech Republic, pp. 655-660.

Mintz, Y. and Brodie, R. 2019. Introduction to artificial intelligence in medicine. *Minim. Invasive Ther. Allied Technol.* 28(2): 73-81.

Razzak M.I., Naz, S. and Zaib, A. 2018. Deep learning for medical image processing: Overview, challenges and the future. In: Dey, N., Ashour, A. and Borra, S. (eds.), Classification in BioApps. 26. Springer, pp. 323-350.

Richters, A., Aben, K.K.H. and Kiemeney, L.A.L.M. 2020. The global burden of urinary bladder cancer: An update. *World J. Urol.* 38: 1895–1904.

Saginala, K., Barsouk, A., Aluru, J.S., Rawla, P., Padala, S.A., and Barsouk, A. 2020. Review: Epidemiology of Bladder Cancer. *Med. Sci.* 8:15.

Sharma, S. and Mehra, R. 2020. Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images-a comparative insight. *J. Digit. Imaging*, 33: 632–654.

Shirazi, S.H., Umar, AI., Naz, S. and Razzak, M.I. 2016. Efficient leukocyte segmentation and recognition in peripheral blood image. *Technol. Health Care.* 24(3): 335-347.

Shkolyar, E., Jia, X., Chang, T.C., Trivedi, D., Mach, K.E., Meng, M.Q.H., Xing, L. and Liao, J.C. 2019. Augmented bladder tumor detection using deep learning. *Eur. Urol.* 76 (6): 714-718.

Shrestha, A., and Mahmood, A. 2019. Review of deep learning algorithms and architectures. *IEEE Access* 7: 53040-53065.

Sompawong, N., Mopan, J., Pooprasert, P., Himakhun, W., Suwannarurk, K., Ngamvirojcharoen, J., Vachiramon, T. and Tantibundhit, C. 2019. Automated pap smear cervical cancer screening using deep learning. In Proceedings of the 2019 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Berlin, Germany, 23–27 July 2019, pp. 7044-7048.

Suarez-Ibarrola, R., Hein, S., Reis, G., Gratzke, C. and Miernik, A. 2019. Current and future applications of machine and deep learning in urology: a review of the literature on urolithiasis, renal cell carcinoma, and bladder and prostate cancer. *World J. Urol.,* 38: 2329-2347.

Teoh, J.Y-C., Huang, J., Ko, W.Y-K., Lok, V., Choi, P., Ng, C-F., Sengupta, S., Mostafid, H., Kamat, A.M., Black P.C., Shariat, S., Babjuk, M., Wong., M.C-S. 2020. Global trends of bladder cancer incidence and mortality, and their associations with tobacco use and gross domestic product per capita. *Eur. Urol.,* 78 (6): 893-906.

Ünaldı, I. 2020. Breast Cancer Histopathological Image Classification with Convolutional Neural Networks Models. Master's thesis, Ondokuz Mayıs University, Samsun, 86 s.

Wang, C., Elazab, A., Wu, J. and Hu, Q. 2017. Lung nodule classification using deep feature fusion in chest radiography. *Comput. Med. Imaging Graph.* 57: 10–18.

Wang, P., Xu, S., Li, Y., Wang, L. and Song, Q. 2018. Feature-based analysis of cell nuclei structure for classification of histopathological images. *Digit Signal Process.* 78: 152–162.

Warren, A.Y. and Harrison, D. 2018. WHO/ISUP classification, grading and pathological staging of renal cell carcinoma: standards and controversies. *World J. Urol.* 36(12): 1913-1926.

Woerl, A.C., Eckstein, M., Geiger, J., Wagner, D.C., Daher, T., Stenzel, P., Fernandez, A., Hartmann, A., Wand, M., Roth, W. and Foersch, S. 2020. Deep learning predicts molecular subtype of muscle-invasive bladder cancer from conventional histopathological slides. *Eur Urol.* 78: 256-264.

Xie, W., Noble, J.A. and Zisserman, A. 2018. Microscopy cell counting and detection with fully convolutional regression networks. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* 6(3): 283-292.

Xu, Y., Hosny, A., Zeleznik, R., Parmar, C., Coroller, T., Franco, I., Mak, R.H. and Aerts Hugo J.W.L. 2019. Deep learning predicts lung cancer treatment response from serial medical imaging. *Clin. Cancer Res.* 25(11): 3266-3275.

Yang, R., Du, Y., Weng, X., Chen, Z., Wang, S. and Liu, X. 2020. Automatic recognition of bladder tumors using deep learning technology and its clinical application. *Int. J. Med. Robot.*: e2194.

Yogananda, C.G.B., Shah, B.R., Jahromi, M.V., Nalawade, S.S., Murugesan, G.K., Yu, F.F., Pinho, M.C., Wagner, B.C., Emblem, K.E., Bjørnerud, A., Fei, B., Madhuranthakam, A.J. and Maldijian, J.A. 2020. A Fully automated deep learning network for brain tumor segmentation. *Tomography* 6 (2): 186-193.

Yu, H. 2020. Research and optimization of sports injury medical system under the background of Internet of Things. *Trans. Emerg. Telecommun. Technol.* 31: e3929.

Yu, L., Chen, H., Dou, Q., Qin, J., and Heng, P.A. 2017. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Trans. Med. Imaging* 36: 994–1004.

Yun, K., Huyen, A. and Lu, T. 2018. Deep Neural Networks for Pattern Recognition. https://arxiv.org/abs/1809.09645, pp. 1-31.

Zhao, J., Qu, H., Zhao, J. and Jiang, D. 2019. Spatiotemporal traffic matrix prediction: a deep learning approach with wavelet multiscale analysis. *Trans. Emerg. Telecommun. Technol.* 30(12): e3640.

Zou, J., Han, Y. and So, S-S. 2008. Overview of artificial neural networks. *Methods Mol Biol.* 458:15-23.

RESUME

Muhammet Fatih ÇAKMAKÇI

fcakmakci07@gmail.com

EDUCATION

Master of Science 2018-2021	Akdeniz University Institute of Natural and Applied Sciences, Computer Engineering, Antalya, Turkey
Bachelor of Science 2010-2016	Eastern Mediterrian University Faculty of Engineering, Computer Engineering, Gazimagusa, K.K.T.C