

A NEW MULTIREOLUTION DEEP CONVOLUTIONAL NEURAL NETWORK WORKFLOW FOR GLOMERULI SEGMENTATION USING AN ITERATIVE ANNOTATION STRATEGY

Master Thesis

Systems biology master program

Vilnius university

STUDENT NAME: OLUWASEUN EZEKIEL TAIWO

STUDENT NUMBER: 1930424

SUPERVISOR: ALLAN RASMUSSEN, PHD.

SUPERVISOR DECISION:

FINAL GRADE

DATE OF SUBMISSION: 13 05 2022

Table of Contents

LIST OF ABBREVIATIONS	4
INTRODUCTION	5
AIM AND TASKS	7
LITERATURE REVIEW	8
Function of the kidney	8
Regulatory function of the kidneys	8
The secretory function of the kidneys	9
Morphological changes in glomeruli diseases	9
Worldwide concern	10
Digital Pathology	11
Whole slide Image (WSI)	12
Role of the pathologist in diagnosis of renal disorders	13
Image analysis	14
Image Segmentation	15
Approaches to Image Segmentation	15
Conventional Image Analysis	15
Machine learning	16
Methods of ML	16
Convolutional Neural Networks	18
Convolution layer	18
CNN Pooling Layer	20
Activation function	20
Batch Normalization	22
Dropout	22
Fully connected layer	23
Loss function	23
Optimizer	23
Basic U-NET Architecture	26
Innovations of deep CNN	27
Residual block	27
Attention block	28

Multiscale VS Multiresolution inputs	29
Image interpolation	31
Innovation of U-net Architecture	32
Kidney glomeruli segmentation review	33
The concept of Human in the loop (HIL).....	35
Evaluation of HIL performance	36
DATA AND METHODS	38
Data Materials.....	38
Dataset1	38
Dataset2	40
Patch extraction.....	41
Preprocessing.....	43
Thresholding	43
Stain color normalization	45
Network topology.....	46
Experimental setup	49
RESULTS	50
DISCUSSION	57
CONCLUSIONS	60
RECOMMENDATION	61
ACKNOWLEDGEMENTS	62
REFERENCES	63
SUMMARY	76
APPENDICES	78

LIST OF ABBREVIATIONS

- CIA** Conventional Image Analysis
- DCNN** Deep Convolutional Neural Networks
- ELU** Exponential Linear Unit
- GAN** Generative Adversarial Networks
- GPU** Graphical Processing Units
- HIL** Human in the Loop
- HOG** Histogram of Oriented Gradients
- HTM** Hierarchical Temporal Memory
- ING** Idiopathic Nodular Glomerulosclerosis
- MAR-UNET** Multi-Resolution Attention Residual UNET
- MCD** Minimal Change Disease
- MCNN** Multiscale Convolutional Neural Networks
- ML** Machine Learning
- OATP** Online Advertising, Traffic Prediction
- PH** Potential Hydrogen
- PRELU** Parametric Relu
- RELU** Rectified Linear Unit
- S-HOG** Segmental Histogram of Oriented Gradients
- TANH** Hyperbolic Tangent Function
- WSI** Whole Slide Image

INTRODUCTION

Just like every other organ of the body, the kidney is also prone to developing pathological changes known as diseases of the kidney. Kidney diseases are usually associated with loss of glomerular function thus homeostatic control of the body's internal environment by the kidney is lost (Pesce and Schena, 2010). Glomerular diseases can destroy the kidneys and may eventually end in kidney failure. During the investigation of every kidney disease, kidney pathologists always want to know the morphological states of the kidney including the structure of glomeruli, because the glomeruli filtration rate can be used for kidney disease diagnosis (Ebert and Schaeffner, 2018).

The use of whole slide image (WSI) has improved clinical diagnosis and really helped kidney pathologists. Now, pathologists can have their scanned glass slide as an image at different magnification levels, stored and shared among other pathologists. WSI is used in image analysis (Hanna et al., 2020). Image analysis is the way of extracting essential information from an image. Semantic segmentation is one of the techniques of image analysis used to classify each pixel in each image independently (Alom et al., 2019). Semantic segmentation of glomeruli is a way of distinguishing pixels of glomeruli from other classes of pixel in each image that consist of glomeruli (Zhang and Ji, 2011).

Generally, in image analysis, the use of the traditional method of conventional image analysis is challenging when it comes to analyzing medical images. This is because medical images are high content images, the contents have different structures and different magnification (Wetteland et al., 2020). Also, conventional image analysis itself involves customization and adjustment of parameters as it involves many independent steps and algorithms, and high human error trying to implement this method (Godinez et al., 2017).

However, a better approach of medical images analysis is the use of machine learning algorithms where image analysis is performed by machines automatically (He et al., 2015). Traditional machine learning workflow is used in classification, segmentation, regression, and anomaly detection in image analysis. (Komura and Ishikawa, 2018). Feature extraction stage in machine learning workflow is also challenging because filters (such as Gabor, Sobel, edge detection etc.) must be handcrafted on images before ML algorithm like random forest or support vector machine is used (Komura and Ishikawa, 2018).

The most recent method of machine learning is deep convolutional neural networks (DCNN) (Gibson et al., 2018; Wulczyn et al., 2021). Deep convolutional neural networks are better because a lot of parameters are generated from images and trained automatically (Aatresh et al., 2021; Roy et al., 2021a, 2021b).

In recent years, there have been advancements through the innovation of neural network architectures for segmentation of high content images in medical image analysis. These advancements came through the introduction of inception block (Szegedy et al., 2015), skip connections ResNet (He et al., 2015), visualization of CNN layer (Simonyan and Zisserman, 2015), the use of residual blocks for vanishing gradient problems (He et al., 2016), cardinality (He et al., 2020; Zagoruyko and Komodakis, 2016), multiscale convolutional neural networks (MCNN) (Godinez et al., 2017), channel boosting (Liang et al., 2019; Wang et al., 2017) and ensemble of deep convolutional neural network models (Gu et al., 2022).

In many current applications DCNN workflow has been engaged in the field of medicine for segmentation of medical image, but few innovations of CNN architecture have been used for kidney glomeruli segmentation. One of the most recently used architecture for kidney glomeruli semantic segmentation is basic UNET architecture but no innovation of processing unit was used in this research (Gallego et al., 2021). Most recently, Moradi and his research team conducted a feasibility study on tomographic kidney images using a customized UNET architecture but whole slide image fully adopted by pathologist was not used (Moradi et al., 2022).

More importantly, building a model with deep convolution neural networks for glomeruli segmentation involves bulk annotated data, drawing annotations on high content cellular images are challenging due to human error, to draw a quality annotation takes a lot of time (Lutnick et al., 2019). Thus, it is difficult for pathologists to get efficient annotated data to train a good deep convolutional neural network algorithms for glomeruli segmentation in whole slide images. And because the structure of kidney glomeruli varies in sizes in whole slide images, the prediction accuracy by trained neural network model for kidney glomeruli segmentation is affected by image resolution (Wetteland et al., 2020).

In this current research, a new multiresolution of deep convolutional neural network workflow is examined. It applies an iterative annotation strategy which involves the customization processing units of DCNN architecture. It optimizes kidney glomeruli semantic segmentation accuracy in the whole slide image.

AIM AND TASKS

Aim

Develop a new multiresolution deep convolutional based workflow for glomeruli segmentation using an iterative annotation strategy.

Tasks

Develop a workflow for multiresolution image patch extraction

Develop and validate a CNN based algorithm that would be able to generate more annotations using an iterative strategy.

Develop and validate a multiresolution UNET model with attention blocks and residual blocks that would be able to perform glomeruli segmentation task.

Test and evaluate the performance of built models using kidney glomeruli datasets from ZENODO

LITERATURE REVIEW

Function of the kidney

The kidney is an important organ of the human urinary system. There are two bean shaped kidneys located at the posterior ends of the abdomen below the rib cage. Blood enters the kidney directly from an abdominal aorta (artery). In the kidney, there are tiny blood vessels called glomeruli. On daily basis of a human's life, about 135 litres of the blood are being filtered by the glomeruli and because of this, about 2 litres of urine is being produced every day. Each glomerulus is attached to small structure called tubule. The main function of the kidneys is excretion of foreign substances and metabolic waste from the blood via the urine, but the kidney also has other regulatory or secretory functions (Kaartinen et al., 2019). A unit of a glomerulus and a tubule is called nephron, Nephron excrete foreign substances and metabolic waste in the urine (Kaartinen et al., 2019).

Regulatory function of the kidneys

Kidney regulates the internal environment of the body by constantly monitoring the quality and quantity of blood that flows through it per unit time and with the information obtained it makes decisions on how to prevent deviation of blood composition away from the expected normal (Hoenig & Zeidel, 2014).

The regulates the internal environment by secreting renin an enzyme which potentiates the Renin Angiotensin Aldosterone System (RAAS) and eventually leads to an increase in blood pressure (Pluznick and Caplan, 2015); Regulation water loss in the urine (Danziger and Zeidel, 2015); Regulation of sodium concentrations by controlling its transportation across the tubular wall (Palmer and Schnermann, 2015); Maintaining Potassium balance (Subramanya and Ellison, 2014); Maintaining potential hydrogen (PH) balance (Curthoys and Moe, 2014); Maintaining the Balance of magnesium, Phosphate and Calcium (Blaine et al., 2015).

The secretory function of the kidneys

Secretion of Calcitriol the active form of Vitamin D. Also, secretion of Erythropoietin: an important hormone found in the bone exactly bone marrow, and erythropoietin mainly stimulates red blood cells' production resulting to a low Oxygen concentration of the tissue (Wang and Kestenbaum, 2018).

Morphological changes in glomeruli diseases

Glomerular cytotoxicity is characterised by reduction of the glomerular cells and monolayer (Haltia et al., 1998). Changes in the thickness of the basement membrane of the glomeruli was observed in examination of glomerulonephritis (Bohle et al., 1976). Alteration of ultra-structures of the glomeruli could be a biomarker for pathophysiology of kidney disease and clinical diagnosis. (Kannan et al., 2019).

Minimal Change disease: or MCD as an acronym which is one of leading causes of nephrotic syndrome, whereby protein in large amounts is deposited in the urine. When the filters in the kidney are destroyed, protein could escape into the urine to form what is called proteinuria (Feehally, 2010).

Inability of the glomeruli to filter blood could result to immunoglobulin A (antibody) to bind up in the kidney, because of this disease such as Berger's disease emanated (Swaminathan et al., 2006; Pesce and Schena, 2010).

In Membranous nephropathy, the glomerular capillary walls thicken homogenously in the later stages, and this is due to the formation of immune complex on the basement membrane outer parts (Nortier et al., 2021). Idiopathic nodular glomerulosclerosis (ING) has been identified by the expansion of the mesangial cells with nodularity of the glomerular highlighted in renal biopsies (Chandragiri et al., 2016).

Worldwide concern

Renal disease can be either acute or chronic with the latter having a poor prognosis and higher prevalence. Due to the increasing amount of cases, high cost of treatment and high mortality rate, renal disease has become a disease of global public health concern (Levey et al., 2011). According to Bornstein, 10% of the world population suffer from renal disease (Bornstein et al., 2014). Amount of renal disease diagnosed in Lithuania has risen dramatically (Vincerževskienė et al., 2014). \$ US 23 million was budgeted for the treatment and diagnosis of renal diseases in Uruguay. An article written by Znaor shows that renal cancer is the most predominant in Europe (Znaor et al., 2015). \$ US 48 billion is budgeted yearly for the diagnosis and management of Renal Disorders. In sub-Saharan Africa, the risk diseases of the kidney are high among the young adults (Arogundade et al., 2020) .

Focal and segmental glomerulosclerosis is the most common glomerular diseases among African Americans, and this is a concern today (Pesce and Schena, 2010). In Latin American country like Brazil, Lupus nephritis is prominent and most common glomerular disease (Kanjanaabuch, Kittikowit and Eiam-Ong, 2009).

According to Briganti et al. in 2001 in Pesca and Schena, 2010, on the prevalence of glomerulonephritis in Australia, they observed that “Immunoglobulin A neuropathy, focal and segmental glomerulosclerosis, Lupus nephritis and vasculitis were the most common renal diseases in adults with a male predominance for all glomerulonephritides except Lupus nephritis” (Pesca and Schena, 2010).

Digital Pathology

Digital pathology is one of the fields of pathology in which with technologies pathologists acquire, interpret, manage, and share pathology information (Keim et al., 2008; Corvò et al., 2019). Specimen slides are converted from glass slides and then studied and analyzed on a computer screen.

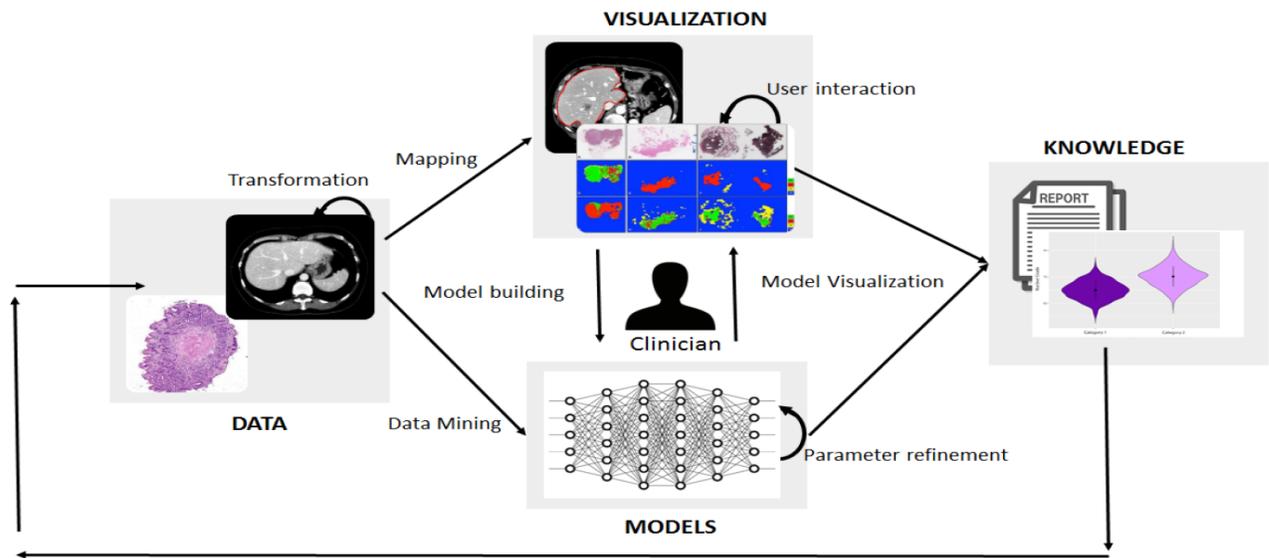


Figure 1. Schematical representation showing the interaction between the three pillars (Data, models, and visualization) (Keim et al., 2008; Corvò et al., 2019).

Figure 1 above shows a general step taken by pathologists to investigate the content of glass slides. The content of these slides is easily converted to images when they are being scanned by digital cameras. These image data could be transformed using some image processing algorithm. Visualization is done through mapping and models could also be built by the pathologist using some computer algorithms during investigation.

Whole slide Image (WSI)

Whole slide image aka WSI is a specialized type of images, it's like a pyramid of images as shown in figure 2 below (Khened et al., 2021; van Rijthoven et al., 2021; Roy et al., 2021a; Wetteland et al., 2020). Normally, Brightfield illumination is used to capture most WSIs. Stained slides such as haematoxylin and eosin (H and E), periodic acid Schiff etc. Deep learning applications is made more attractive by a wider accessibility of H and E-stained slides compared to custom-tailored labeling reagents. Hence in clinical practice, for the characterization of medical tissue samples and identification of morphology in tissue samples, H and E-stained tissues are an excellent pick (Dimitriou et al., 2019).

The resolution of the Whole slide image is measured in micrometer per pixel. The pixel size of a whole slide image can be influenced by resolution of camera, length of light tube, objective magnification. Generally, when 10X scanner is used, images are saved at about $1.0\mu\text{m}$ per pixel, when 20X scanner is used images are saved at about $0.5\mu\text{m}$ per pixel and when 40X scanner is used images would be saved at $0.25\mu\text{m}$ per pixel (Marini et al., 2021).

However, the use of stained slides with a cell for in situ molecular data and in cancer research is really challenging (Wong et al., 2019). As opposed to numerous other fields where deep learning techniques are supervised and adopted, obtaining labeled data is much more difficult and this poses a problem on practicing supervised approaches.

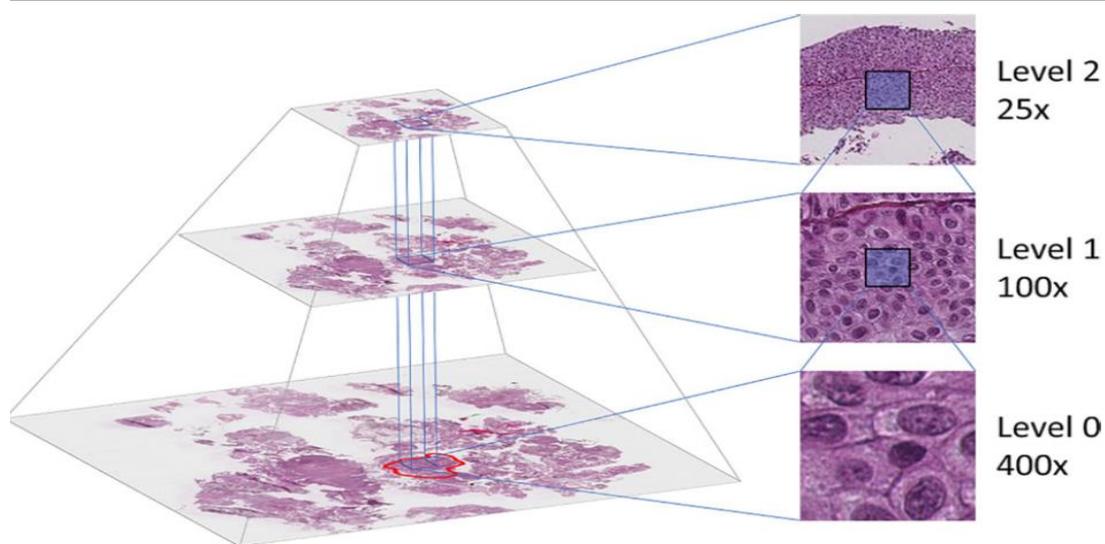


Figure 2. Schematical representation of Whole slide image in a pyramidal file format, showing different stage of down-sampling versions of the base image (Wetteland et al., 2020).

Generally, there are two approaches to be followed to facilitate a deep learning model outcome. The first approach involves color variability normalization (Magee et al., 2009) and the second approach is through artifact removal (for example, using image filters). The second approach, however, is less direct. It makes learning an important part of the training process by developing data with synthetically generated data that takes note of a representative variability in staining and in artifacts. Although, this finding has not been made universal (Liu et al., 2017), these two approaches have been used to adjust the variation from clinical samples that have been archived in different clinics and from batch effect (Dimitriou et al., 2019).

In some real-life applications, the image patches gotten from whole slide images analysis are enhanced by image filter-based preprocessing. Color Normalization is used to unify the color variation in RGB patches before the training of the network (Anghel et al., 2019). There are various other methods of normalization which are also used to improve the quality of patch images such as contrast normalization (Komura and Ishikawa, 2018).

Because a huge amount of training data is required to train a good neural network, data augmentation is required to reduce overfitting when the training set is small. Data augmentation algorithm is required in most cases to increase the amount of training data. The common algorithms for image data augmentation are metal learning, neural time transfer, kernel filters, color space augmentation, erasing etc. but the most recently used algorithm is generative adversarial networks (GAN) (Shorten and Khoshgoftaar, 2019).

Role of the pathologist in diagnosis of renal disorders

In the laboratory, renal function can be assessed by various sample analytic methods which includes serum creatinine, blood urea nitrogen, urinalysis, renal biopsy to mention just a few (Jialal et al., 2019). Serum creatinine is the gold standard investigation in the diagnosis of renal pathology (Koyner, 2012). Creatinine is an endogenous biomarker that is filtered constantly by the kidney thus a good biomarker of renal function. It can be used to estimate renal glomerular filtration rate. In situations of renal disorders, renal function can be monitored by investigating the serum Creatinine level. An increased serum Creatinine level point towards poor renal filtration rate and subsequent accumulation of Creatinine in blood (Koyner, 2012).

It is important to note that the serum creatinine test cannot detect early stages of renal impairment. This is because Serum Creatinine level can only be detected in the later part of the disease when about half renal function has been lost (Jialal., 2019). This calls for a need to find another

diagnostic approach to detecting early stages of renal deterioration and alleviating the deadly effect of late diagnosis (Jiang et al., 2020).

Pathologists make use of appropriate image analysis methods to extract relevant information from whole slide images (Aeffner et al., 2016). Method of analysing medical images started at an infant stage where a lot of human input was required to arrive at a final diagnosis (Conventional image analysis) and is developing towards a more automated form which works on principles of convolutional neural network.

In image analysis with convolutional neural network, pathologist can identify the early stages of renal impairment in patients with high risk of developing renal failure (Jiang et al., 2020). For example, streams program developed by google could predict the early stages of acute kidney injury. It also takes it further by informing the doctor about this new development, thus medical intervention is initiated early (Powles and Hodson, 2017). In the kidney, kidney diseases are explained by tissue changes and the relationship between different morphology (cells, arteries, capsules, etc.) changes in the kidney (Farnik and Zeuzem, 2012).

Image analysis

The term "Image analysis" is used to describe the process of retrieving usable information from images (Lantuejoul, 2019). The use of images that are produced within the clinical settings (e.g., histopathology slide, x-ray film) to solve medical related problems is known as Medical Image Analysis).

Traditionally, medical image analysis involves the use of tools like the microscope to view and measure parameters in the laboratory (Aeffner et al., 2019). The pathologist usually identifies glomerular diseases by changes in glomerular cells (Farnik and Zeuzem, 2012). In recent years, medical image analysis has evolved into a digitized and computerized form. And this has brought about the use of a new set of tools like digital slide scanners and Image storage device (Aeffner et al., 2019).

Digital image analysis makes the measurements of cellular components easier. More details of the image on the slide can be appreciated better when viewed on a screen as compared to viewing under the microscope. Digitized slides can be archived and organized. There is reduction of individual bias when reading the image thus teaching and analysis can be standardized. Specialized software can be used for extracting complex data in a manner that is reproducible (Aeffner et al., 2016; Godinez et al., 2017).

Image Segmentation

There is need for digital images to be broken down into various segments, regions, or a set of pixels to be able to locate lines, curves, and objects in images (Linda G. Shapiro, George C. Stockman, 2001), this process is called image segmentation. With the use of any image segmentation method, pathologists can extract important information needed for diagnosis and then do analysis (Zhang and Ji, 2011). Image contrast, texture, colour etc are criterions utilised to achieve successful segmentation of an image (Zhang and Ji, 2011). The segmented image is identified and classified by the clinical expert thereby detecting abnormalities in tissue.

Approaches to Image Segmentation

There are different approaches to image segmentation. These approaches include a discontinuity approach and similarity approach. In discontinuity approach, there is a sudden change in the intensity level of the images and the partitions observed in image is based on this sudden change. What is of interest in this approach is the identification of isolated points, edges and lines in the image and a 3 x 3 mask operation is used in the identification (Kumar, Arthanari and Sivakumar, 2012). However, in similarity approach images are segmented based on similarity. These similarities are detected between image pixels by Machine Learning (ML) algorithms based on a threshold (Komura and Ishikawa, 2018).

Conventional Image Analysis

Conventional image analysis (CIA) is the first set of technique used to analyse digitized medical image. Convectional Image analysis workflow pipeline consists of various independent data analytic steps, which includes Segmentation, Classification, Detection of abnormalities (Anwar et al., 2018). According to Godinez, "Conventional image analysis approaches essentially transform the image data into different levels of abstraction, starting from the pixel intensities and ending in the higher-level semantics describing the data" (Godinez et al., 2017).

The disadvantage of CIA is that the steps are independent of each other and thus requires massive human participation (Anwar et al., 2018). It is a time-consuming process (Anwar et al., 2018). The amount of data necessary is massive and analysing it as a human can be quiet challenging (Jiang et al., 2020) . There is risk of human error occurring (Jiang et al., 2020).

In-dept analysis of information is necessary and this might be difficult for humans to achieve (Jiang et al., 2020). In a bid to solve this problem research has been ongoing on how to integrate all the steps in conventional image analysis and automate it.

“Conventional image analysis approaches have certain limitations. For instance, several steps along the analysis pipeline, such as object segmentation, dimension reduction, and phenotype classification, typically require customization to each specific assay using a priori knowledge, such as the geometric properties of the expected phenotypes” (Godinez et al., 2017).

Machine learning

Machine learning is a technique of data analysis that automates the creation of analytical models. ML uses algorithms capable of detecting patterns in existing data, then makes predictions on new data after training itself (Komura and Ishikawa, 2018; Sultan et al., 2020). Machine and deep learning are both subsets of artificial intelligence of which deep learning is the most recently used in research (Lutnick et al., 2019). In traditional machine learning for image segmentation, features from the input image are hand crafted for independent pixel classification (Alom et al., 2019). It has a training phase where good values are learned for bias and weight from labelled data. The model accuracy based on predictions of the learned model goes through a testing phase on blind data to determine the model having the lowest error on generalization (Wulczyn et al., 2021).

Several applications of machine learning in modern day exist in various forms such as online advertising, traffic prediction (OATP) for global positioning system, email spam filtering, storm model weather predictions and recommendation systems used for books or movies online (Sultan et al., 2020).

Methods of ML

Supervised and unsupervised learning are the major two techniques in machine learning mostly used in image analysis under digital pathology (Guégan and Hassani, 2018). For supervised learning, several algorithms are used which include random forest algorithm, support vector machines, and convolutional neural networks (Huang et al., 2018). Conversely, deducing a function which, from unlabeled images can describe hidden structures, is the main purpose of unsupervised learning. This task involves anomaly detection, clustering and reduction of dimensionality and the algorithms involved include autoencoders, k-means and principal component analysis (Sultan et al., 2020).

Other techniques of ML are semi-supervised learning and multiple instance learning (figure 3). The derivative of unsupervised and supervised learning is semi-supervised. The idea of semi-supervised learning is that few data are labeled while in multiple instant learning groups of data are labelled instead of individual labelling (Komura and Ishikawa, 2018).

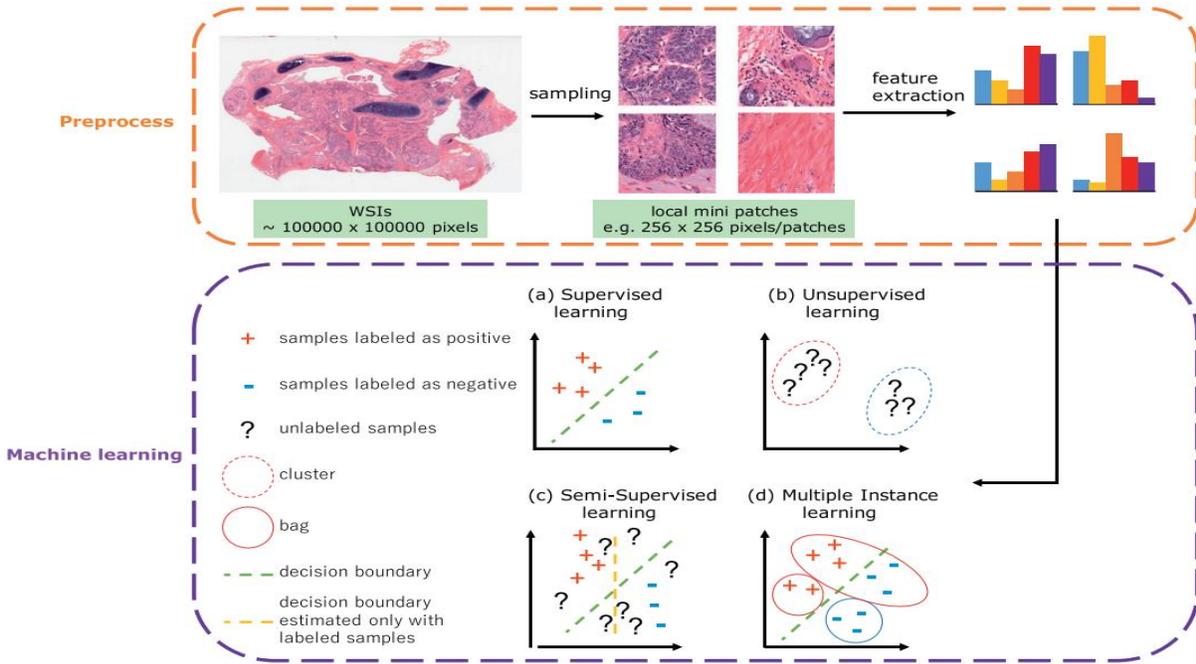


Figure 3. Schematic diagram explaining the techniques of ML in histopathology (Komura and Ishikawa, 2018).

Convolutional Neural Networks

CNN is an algorithm that is used to process data that is in grid form (such as image) (Woo et al 2018). Convolutional neural network is recently used in image segmentation (Kayalibay et al., 2017), image classification (Recht et al., 2019), lesion detection (Jiang et al., 2020), image reconstruction (Huang and Cao, 2020) etc. In contrast to traditional machine learning, convolutional neural networks are designed in such a way that, they can automatically learn features from images and use these features for prediction, there is no hand engineering of features in convolutional neural networks (Zhou et al., 2014).

Convolution, pooling, and fully connected layers are the major components of most CNN architectures (Lutnick et al., 2019; Yamashita et al., 2018). Sometimes, an average global pooling layer replaces a fully connected layer. Each of these layers performs its functions independently. Aside from mapping functions in these layers, dropout and normalization of batches which are regulatory units are also included to increase the performance of CNNs (Bouvier, 2006). In attaining enhanced performance and in new architectural designs, arrangement of components (layers) in the CNN plays an important role (Khan et al., 2020a).

Convolution layer

The first layer of the convolutional neural network is the convolution layer; therefore, it is being referred to as the basic unit of CNN, this layer plays a leading role in CNN (Yamashita et al., 2018). A convolutional neural network layer consists of convolution operations (equation 1 equation 2) and activation function. In convolution layer, mapping of features is done by mathematical operations called convolution (matrix multiplication) (Qu et al., 2021). Input image is also called tensor and it just an array of numbers in form of grid. Kernel is a small array of numbers (grid of parameters) which could be 3x3 or 5x5 or 7x7 matrix (Yamashita et al., 2018). During convolution, there is binary operation (matrix multiplication) between the elements of the kernel and the element of the tensor of a given stride. This product is summed to form an element of feature map in equivalent position of output tensor (He et al., 2021). This operation is repeatedly done at a given stride resulting in feature maps which signify the input image characteristics. Then nonlinear operation (such as sigmoid or rectified linear unit) is applied on feature maps.

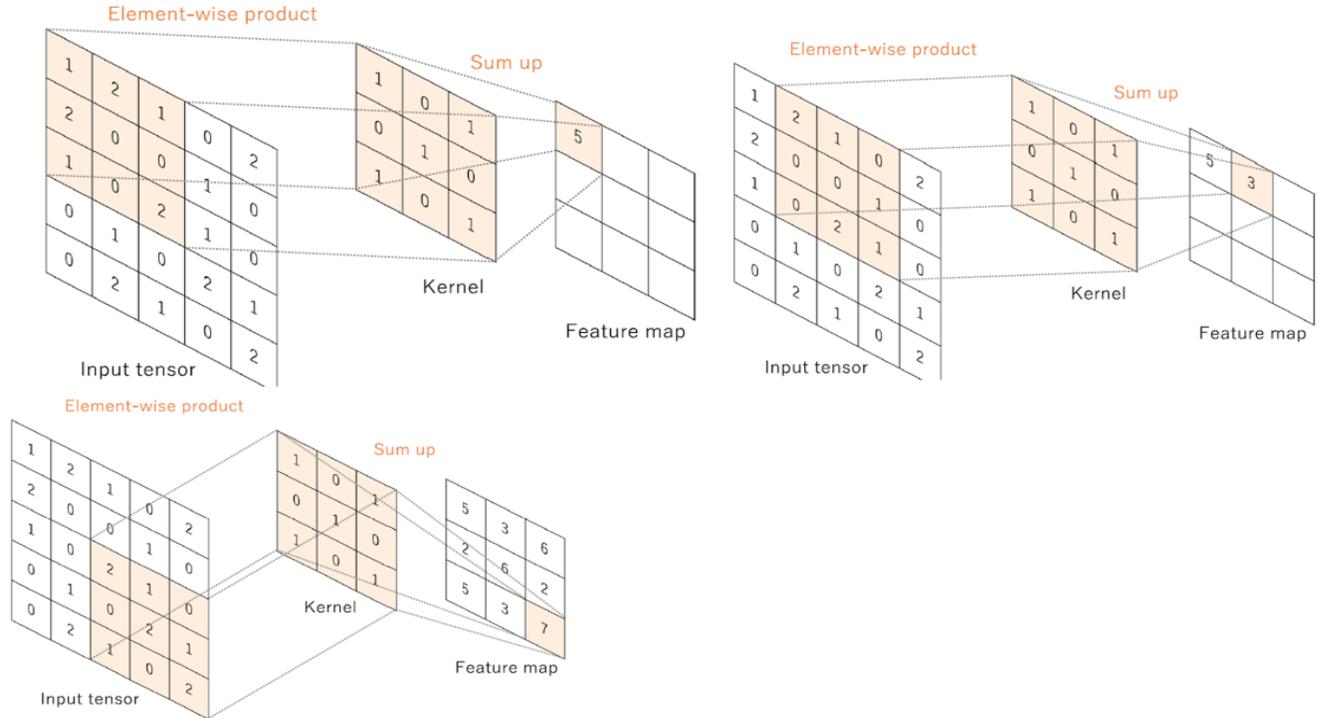


Figure 4. This is an example of how each element of a feature map is calculated during convolution. Kernel serves as the feature extractor (Yamashita et al., 2018).

Image (input) is the tensor while the output is the feature map. In figure 4 above, 3 x 3 matrix is used as the kernel to multiply the input tensor resulting to feature map.

$$f_l^k(p, q) = \sum_c \sum_{x, y} i_c(x, y) \cdot e_l^k(u, v)$$

(1)

Equation 1. General mathematical formular of convolutional operation.

From equation 1 above, $f_l^k(p, q)$ the product of matrix multiplication of $i_c(x, y)$ and $e_l^k(u, v)$. i_c represents the tensor (input image) and e_l is the kernel (filter), x and y are the elements of input image i_c , u and v are the elements of kernel e_l , k is the index of convolutional kernel, f_l is the feature map, p and q are the elements of the feature map, c is the image channel index (Litjens et al., 2017).

As a result of convolutional operation weight sharing ability, various types of image features can be drawn out by sliding kernels with the same weight set on the image. As opposed to fully connected networks, this makes CNN parameter efficient. Categorically, convolution operations differ by direction of convolution, size and type of filters and type of padding (Lecun et al., 2015).

In some multiscale convolutional neural networks, it is possible to concatenate multiple features. Concatenation of features means summing of two or more feature maps (Roy et al., 2021a) or using a function that would take the union of two or more feature maps (Ma et al., 2019).

CNN Pooling Layer

In image analysis, a result of convolution operation is feature motifs which can occur at various locations. The precise location of a feature, once extracted, becomes unimportant so long the estimated relative position to others remains unchanged. An interesting local operation is pooling. It collates information on receptive fields within the local region and produces the dominant response (maximum number is selected) (Lee et al., 2016).

$$\mathbf{Z}_l^k = g_p(\mathbf{F}_l^k) \tag{2}$$

Equation 2: Mathematical expression explaining the pooling operation.

\mathbf{Z}^k represents the feature map that had been pooled, \mathbf{F}_l^k is the feature map to be pooled, $g_p(\cdot)$ is the pooling operation. (Litjens et al., 2017).

Pooling operations (equation 2) help when extracting a blend of features that usually don't change in the presence of little distortions and translational shifts (Scherer et al., 2010). Generalization made by overfitting reduction increases and network complexity is regulated when a feature-map size is reduced relative to a non-changing feature set. In previous research, pooling methods such as average spatial pyramid, max pooling, L2, overlapping, max pooling were used while max pooling is the most popular method (Bouvier, 2006; Wang et al., 2017).

Activation function

Activation function helps in making decisions and learning patterns that are intricate. Selecting the right activation function helps to fasten up the learning process.

$$\mathbf{T}_l^k = g_a(\mathbf{F}_l^k) \tag{3}$$

Equation 3. Mathematical expression of activation function, \mathbf{F}_l^k denotes the output of convolution passed to the activation function $g_a(\cdot)$ (Litjens et al., 2017).

According to Litjens et al in 2017, Activation functions such as tanh, sigmoid, RELU, parametric ReLU (PReLU) are the most common activation functions in CNN (Litjens et al., 2017).

$$g_a(\mathbf{F}_l^k) = \max(0, \mathbf{F}_l^k) \tag{4}$$

Equation 4: Simple formular of ReLU activation function. This function would return zero (0) if the value of \mathbf{F}_l^k (input) is negative or returns \mathbf{F}_l^k if the input is positive. \mathbf{F}_l^k is the output of convolution layer passed to the activation function ReLU ($g_a(\cdot)$). (Litjens et al., 2017).

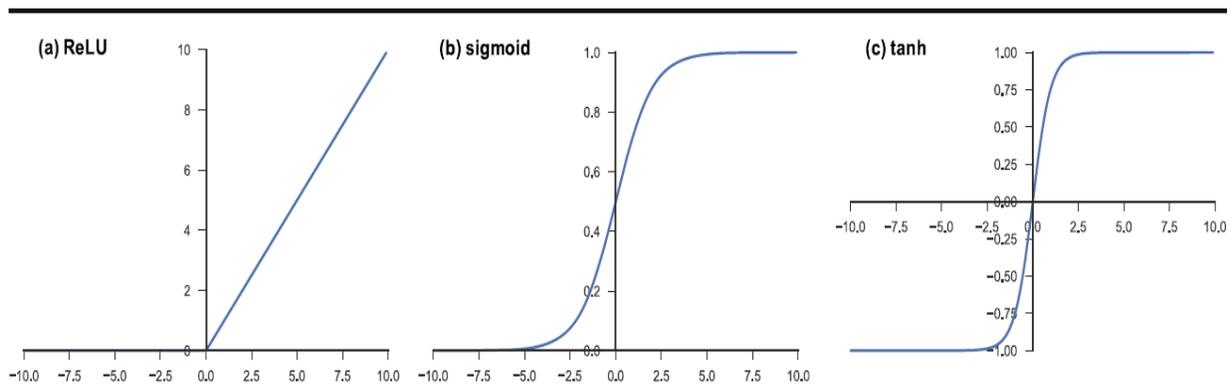


Figure 5. Showing the curve of the most used activation function in neural networks: figure 5a is ReLU, figure 5b is sigmoid function and figure 5c is tanh (Yamashita et al., 2018).

Batch Normalization

Batch normalization takes care of problems in feature maps involving internal covariance shift. This internal covariance shift causes hidden unit value distribution to change which in turn slows convergence down and demands careful parameter initialization.

$$\mathbf{N}_l^k = \frac{\mathbf{F}_l^k - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (5)$$

Equation 5. Mathematical expression of batch normalization operation, \mathbf{N}_l^k is the normalized feature map, μ_B is the mean, while σ_B^2 is the variance with batch normalization there is unification of the feature map's distribution, \mathbf{F}_l^k is the output of convolution, ϵ is the number being added for numerical stability (Litjens et al., 2017).

Dropout

Dropout improves generalization when it is introduced within the network by dropping some connections randomly or some units with a certain probability, network with low weight are dropped out of the connection to avoid overfitting of the network. Overfitting occurs in deep convolutional neural network because a non-linear relation that has been learnt by the multiple connections co-adapt. Some units or connections are randomly dropped, and they produce respective thinned architectural networks and then, small weights are picked with a representative network (Litjens et al., 2017).

Placing dropouts in between convolutions is now outdated as new architectures are no more using, instead of average pooling is calculated in some architecture to improve generalization, some architectures are making use of residual blocks (He et al., 2016).

Fully connected layer

This layer is majorly used for classifying a network at the end. It is a global operation unlike convolution and pooling. It surveys the output of all preceding layers globally when it receives input from the stages in feature extraction. It helps in classifying data by making selected features of a non-linear combination (Litjens et al., 2017). This layer could also be referred to as the dense layer, in this layer, there is transformation of the last pooling or convolution layer (2-dimensional array of number) into 1-dimensional vector. Each fully connected layer is associated with weight and bias.

Loss function

Loss function is the error function or can be called the cost function. It is one of the hyperparameters that needs to be defined in DCNN. To measure the similarities (quantifies error) between labels of a ground truth and a network output, a particular loss function is required in CNN workflow (Yamashita et al., 2018). Binary cross entropy is used in classification task while mean squared error is used in regression task. Nevertheless, sometimes it is required to combine two or more loss functions for sanity (Gu et al., 2022).

$$\mathcal{L}_{hybrid}(p, y_1, y_2) = -w \sum_i^n p^i * \log(y_1^i) - (1 - w) \sum_i^n p^i * \log(y_2^i) \quad (6)$$

Equation 6. Showing the mathematical expression of the cross-entropy loss function, p denotes the network prediction, y_1 and y_2 are the annotations in the input image and w is the weight, n is the pixel number of input image (Zhang et al., 2021; Gu et al., 2022).

Optimizer

The major interest of optimization is to find the minimum value of the loss function (equation 6)(Lee and Myung, 2017). Gradient descent is a highly used iterative optimization algorithm. In gradient descent, the parameters (weight and bias) are initialized to random values and move in slope's direction in small step at each iteration.

Gradient descent repeatedly updates the weight and bias (learnable parameters) (Yamashita et al., 2018). Practically, because of insufficient computation memory, mini batch (subset of training dataset) is widely used to compute the loss function gradient irrespective of the parameters involved, this process is called stochastic gradient descent. (Yamashita et al., 2018).

Most recently used first order optimization algorithm is adaptive moment estimation aka ADAM. Adam is a method for stochastic gradient descent that incorporates root mean square propagation and momentum (Lee and Myung, 2017).

However, one of the major problems neural networks faced when using any gradient descent-based algorithms for training is called vanishing gradient problem. When the layers of neural network increase and more activations occur, loss function gradient approaches zero and this makes the network stop learning. (Alom et al., 2019). The output of loss function activation like sigmoid function (figure 5) is small, therefore its derivative becomes smaller. To tackle this problem some researchers used rectified linear units instead of loss function. However, adding residual blocks (figure 9) neural networks can also help to prevent vanishing gradient problem when sigmoid function must be used (He et al., 2016).

Learnable parameters are generated in each layer of CNN before training, for example learnable parameters weights and bias. Hyperparameters are not trainable but set before training any CNN algorithm. For example, in convolution layers hyperparameter are size of the kernel, the value of stride, padding and the type of activation function to be executed (equation 4 equation 5).

In pooling layers methods of pooling, the size of filter, value of stride, padding must also be set before the training of network (Yamashita et al., 2018).

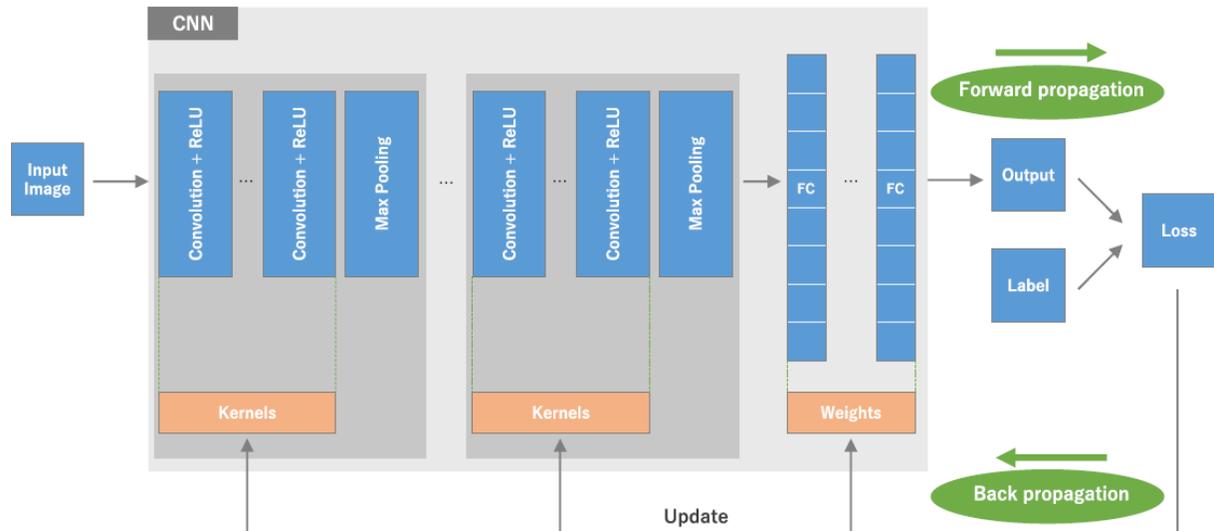


Figure 6. General overview of CNN architecture (Yamashita et al., 2018).

Figure 6 shows the basic fully convolutional neural network, kernels are associated with convolution blocks while the weights and bias are associated with the dense layer (fully connected layer).

The way to move from input layer to output is called forward propagation while the opposite movement is back propagation. Back propagation

Basic U-NET Architecture

U-NET is a CNN architecture to segment medical images (Alom et al., 2019; Cireşan et al., 2012; Ronneberger et al., 2015). This architecture has a right part (encoder) which is an expansive path and a left part (decoder) which is a contracting path as shown in figure 7. The right part generates automatic features from an input image while the left part captures context and allows learning from the image masks (ben Hamida et al., 2021).

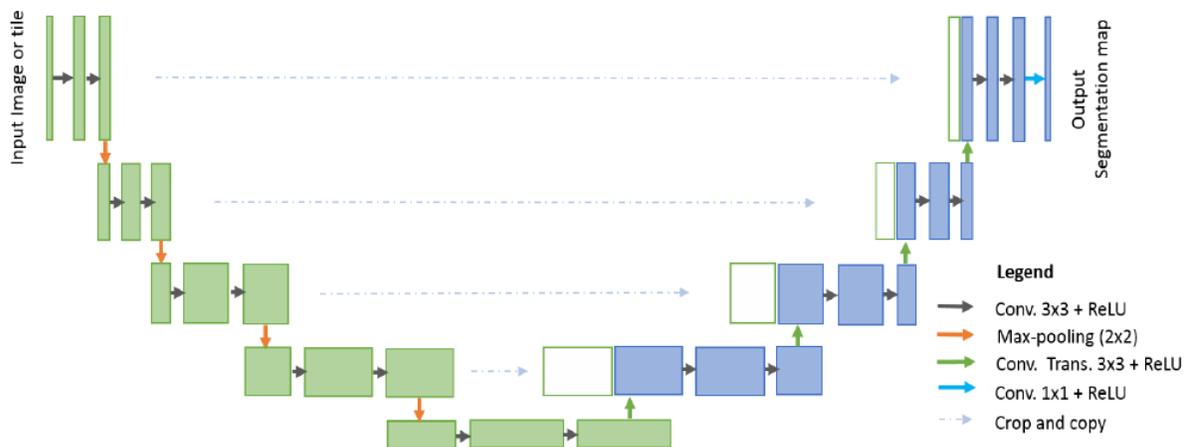


Figure 7. Showing the structure of a U-Net architecture; encoding unit and the decoding unit, the green blocks are contracting path while the blue blocks are expanding path, the layers of the architecture from the convolution layers are connected to the output layer formed a U structure. (Alom et al., 2019).

Furthermore, one of the major reasons for using U-NET architecture in research is that U-NET models can perform segmentation task correctly with small amount of training and validation data better than many other known architectures such as Mask R-CNN and SegNet (Alom et al., 2019).

Innovations of deep CNN

All deep learning are neural networks, but not all neural networks are deep learning. Example of neural networks deep that are not learning are self-organising maps (Qu et al., 2021), radial basis function networks, spiking neural networks, physical neural networks using physical “synapse”, hierarchical temporal memory (HTM) (Zyarah and Kudithipudi, 2019). However, CNN could be explored considering distinct aspects like loss functions, different activation, regularisation, optimization of parameters and dynamic innovations of CNN architecture (Khan et al., 2020a; Litjens et al., 2017). In recent years the great advancement is accomplished via the dynamic innovations of CNN architecture. Information on channels and spaces, architectural depth and width, processing multipath information are being exploited garnering attention.

Block layered structural units have likewise amassed attention. CNN innovations involve various features such as pattern design, layer connectivity, processing unit modifications, parameter and hyperparameter), etc. Notably, redesigning of new blocks and processing units have been suggested and accredited to CNN innovations (Khan et al., 2020).

Residual block

Residual blocks were introduced to DCNN to overcome the problem called vanishing gradient problem (Alom et al., 2019). Residual networks come with the idea of skip connections (He et al., 2016).

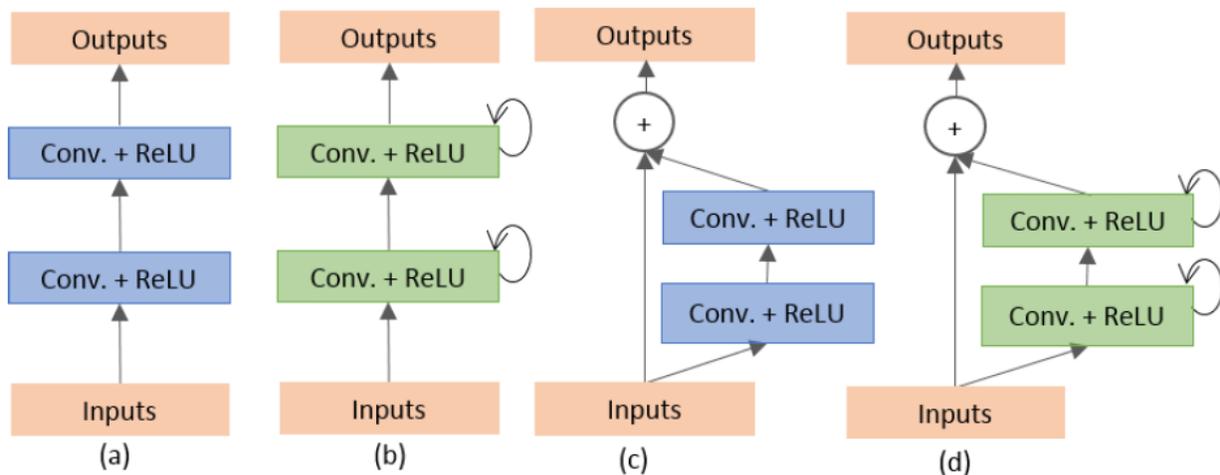


Figure 8. Showing different processing units in CNN (Alom et al., 2019).

In figure 8 above, 8a depicts an example of forward convolutional units (Gadermayr et al., 2019), recurrent convolution block is found in 8b (Alom et al., 2019), 8c is residual convolutional connection unit (He et al., 2016) and 8d is showing an example of recurrent residual convolutional units (Alom et al., 2019).

$$y = \mathcal{F}(x, \{W_i\}) + x \quad (7)$$

Equation 7. Showing the mathematical representation of convolution blocks with residual block added. y is the output vector; x is the input vector, W_i is the weight, " $\mathcal{F}(x, \{W_i\})$ " is the feature map, i is the index of the weight (He et al., 2016).

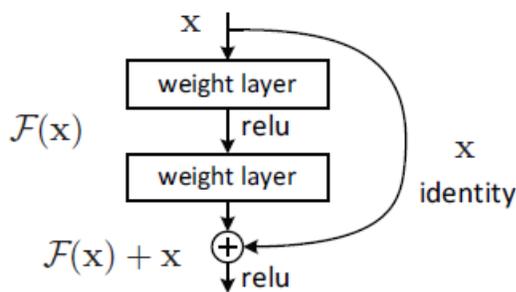


Figure 9. A residual building block (He et al., 2016).

From figure 9, a residual block is a combination of normal convolution block (convolutional operation and activation function) and features from lower layer (equation 7). Sometimes residual blocks are called residual convolution blocks.

Attention block

During the training process, only the relevant activations are highlighted by attention method. Computational resources wasted on irrelevant activations is minimized and provides better generalization of the network (Oktay et al., 2018). There are two types of attention: hard and soft attention (Luong et al., 2015; Oktay et al., 2018). Hard attention highlights relevant regions by cropping. In hard attention, just one image region is considered per time and its either network pay attention or not. There is no backpropagation in hard attention (Luong et al., 2015).

Soft attention adds weight to pixels of the image based on relevance. The relevant parts gain weight and less relevant parts gain a small weight (Khan et al., 2020b). This can be trained with backpropagation (figure 6). During training in soft attention, the weight also gets trained making the model pay more attention to relevant regions (Luong et al., 2015).

The process called skip connection combines spatial information from the down-sampling path with the up-sampling path to retain good spatial information. But the problem of skip connection is that it brings along the poor feature representation from the initial layers (Oktay et al., 2018). Therefore, soft attention suppresses activations at irrelevant regions when its being implemented at the skip connections.

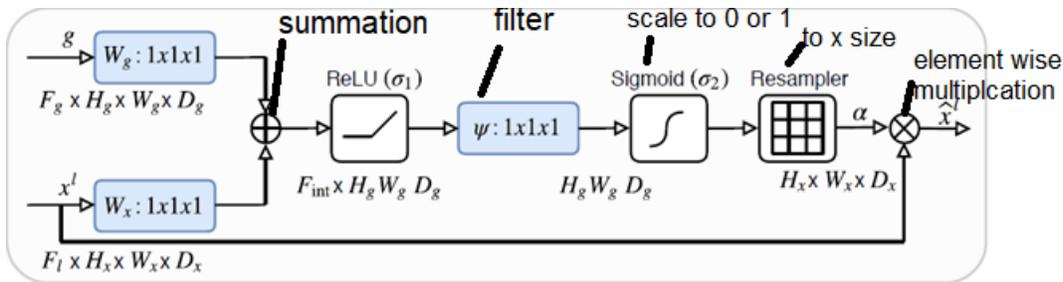


Figure 10. Attention gate (Oktay et al., 2018). There are two inputs g and x^l , g is the gating signal which comes from the next step lower layer of the network, and it has better feature representation since it comes from deeper layer. x^l comes from skip connections, since it comes from early layers, it has better spatial information but not more feature information. “ $F_g \times H_g \times W_g \times D_g$ ” is the dimension of g while “ $F_l \times H_x \times W_x \times D_x$ ” is the dimension of input image x^l . F , H , W and D . are features, height, width, and channel, respectively.

Attention block checks what comes from skip connection using a particular algorithm (figure 10).

Multiscale VS Multiresolution inputs

The input of convolution neural network cannot be underestimated because the performance of the network relies on the strategies and the methods adopted to build the network. One must be sure of the input to trust the network result accuracy. Nevertheless, let us consider what researchers have called multiscale patch images in their research works.

Because biological samples (high content images) have different shapes, texture and levels and numerous dimensions, they are difficult to analyze (Godinez et al., 2017). Godinez et al., was able to capture a spatial scale of image samples by developing an architecture he called Multiscale convolutional Neural network (MCNN) for classification problem. He fabricated a single network that handles seven different scales of image inputs simultaneously.

MCNN takes seven different image scale inputs through two channels. He therefore uses the trained network for both binary and multiclass labelled prediction. Notably in this research work scaling dimension is referred to as (width/scale X height/scale) while the value of scale is 1,2,4,8,16,32 or 64 (figure 11). Given an Image I with dimension of 1000×1000 , the dimension of image I_2 when scale = 4 would be 250×250 while the aspect ratio is being kept constant. In view of this research work this scaling method only affects the physical size of the input images while the magnification of the image object is kept constant. A similar method of scaling was published in 2019 by “Institute of Electrical and Electronics Engineers” (IEEE Circuits and Systems Society et al., 2019), also most recently in Oakland University (Oakland University et al., 2021).

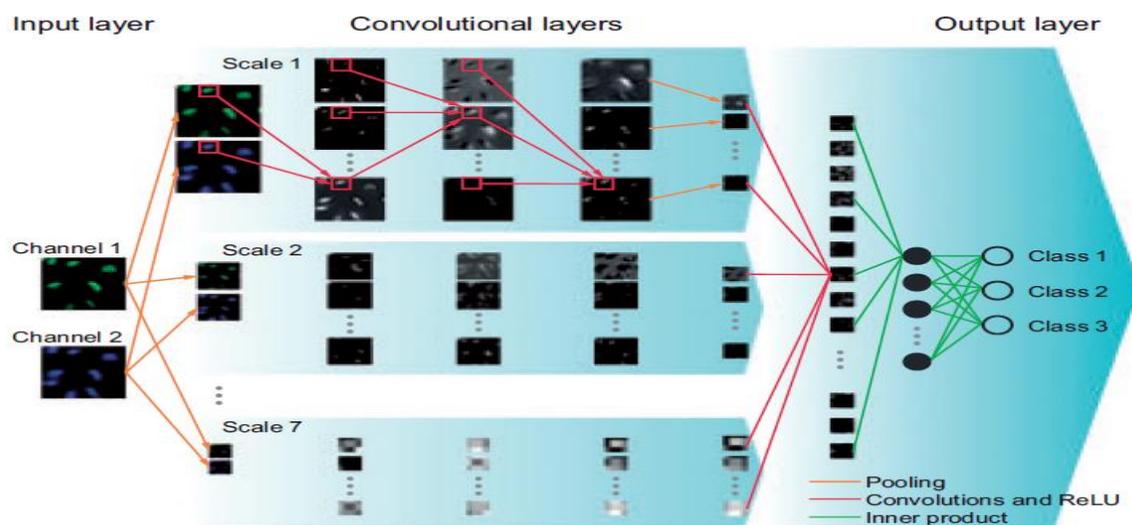


Figure 11. Showing multiscale convolutional neural network algorithm developed by Godinez et al (Godinez et al., 2017).

However, another method of scaling a network input (Wetteland et al., 2020) could be considering the pixel size of an image. In this case there would be input with same physical dimension (e.g., 256×256) but different zoom level (e.g., 10X, 20X, 40X as shown in figure 12). In this case the objects in the images appeared to be in varied sizes or levels of magnification, this means object are at different resolution. The higher the magnification level the bigger the object in the image. Wetteland et al., eventually adopted the use of VGG16 (CNN with 16 layers) as the base architecture and three different input scales.

A trained VGG16 network was used for both single and multiclass image segmentation. Roy et al., 2021 adopted the style of image scaling but different CNN architecture. Most recently, Rijthoven and his colleagues used the same scaling method to innovate an architecture called Hooknet (van Rijthoven et al., 2021).

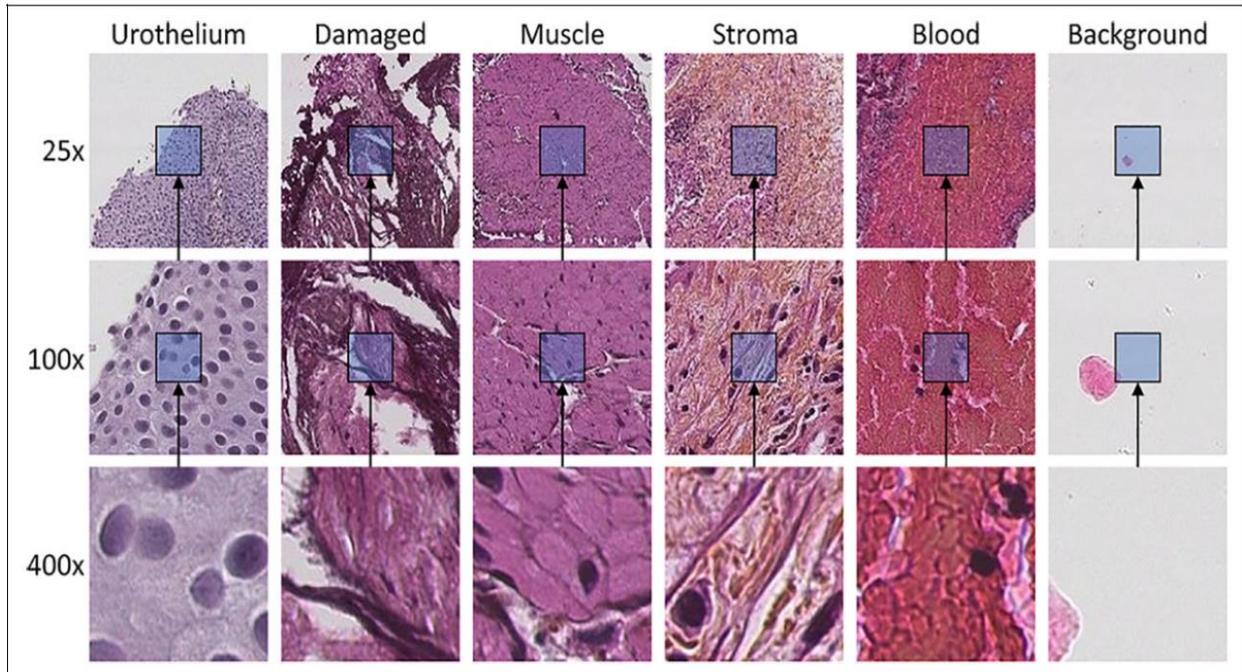


Figure 12. Showing various levels of resolution of multiscale convolutional neural network adopted by Wetteland (Wetteland et al 2020).

Image interpolation

The word interpolation simply means, calculation a value from two given values (Roszkowiak et al., 2017). In image interpolation, there is transformation of image from discrete to continuous image sample. It could also mean changing a 1-dimensional interpolation signal to 2-dimensional and 3-dimensional image stacks. Therefore, interpolation is used when we need to get a number from two discrete values. In image processing, interpolation could be used for resizing, image registration, image compression, image decomposition, spatial distortion correction, geometric operation like scaling, translation, and rotation of image (Sellaro et al., 2013).

The major reason why interpolation is required in research while Whole slide images are being used is that immunohistochemical tissue samples vary in shapes, pixel size and resolution. Methods of image interpolation are Nearest neighbor, Linear/bilinear, bicubic (Huang and Cao, 2020), Cosine, b-splines, Hermite, interarea interpolation, Lagrange polynomial interpolation and quadratic polynomial equation (Roszkowiak et al., 2017). According to (Roszkowiak et al., 2017) in their recent.

Innovation of U-net Architecture

Figure 7 above shows the basic configuration and structure of U-net architecture. This architecture is used enormously in much medical research to demonstrate segmentation tasks (Alom et al., 2019). Nevertheless, with U-Net architecture feature maps can easily be copied from the encoder unit to any decoding unit (Ronneberger et al., 2015). Basic U-net architecture consists of pipeline process from end-to-end and protects the integrity of the image patches (Türk et al., 2020). Nabla-Net was proposed in 2016 for multiple sclerosis segmentation (McKinley et al., 2016).

Modification of the processing units in U-net framework had resulted in innovation of architecture Attention U-Net etc. (Zhang et al., 2018). No tangible significant result was observed when compared the performance of basic U-net model and U-net model with multiple feature maps summation (Kayalibay et al., 2017), V-net is a type of 3-D U-net for multiple sclerosis segmentation (Milletari et al., 2016), Recurrent residual U-net (Alom et al., 2019), Recurrent U-net (Alom et al., 2019).

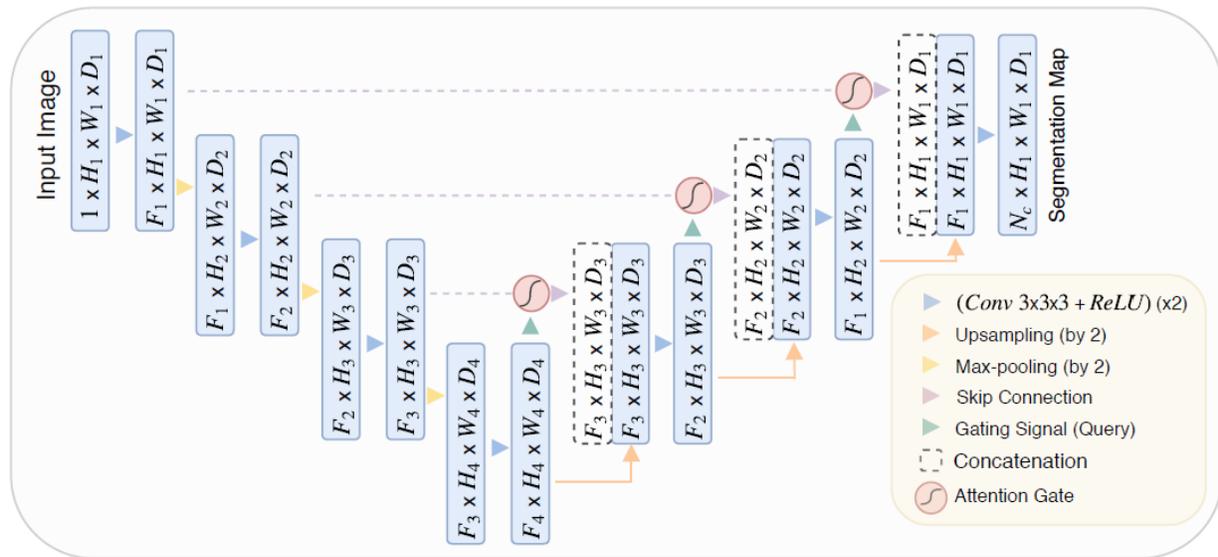


Figure 13. Showing attention- UNET architecture (Oktay et al., 2018).

This is how attention gates can be added to the decoding unit of basic UNET architecture (figure 13). In most deep CNN, vanishing gradient problem is the major issue where the weight and bias are unable to be updated, therefore loss function is increased. This prevents the network from learning and in turn decreases the performance of the network.

Kidney glomeruli segmentation review

In recent years different approaches were used to segment glomeruli from other kidney structures. Temerinac-Ott used a method called mutual information and convolutional neural network for segmentation, detection, and evaluation of glomeruli structures in whole slide images, Temerinac-Ott obtained 76% F1-Score (Temerinac-Ott et al., 2017). Kakimoto used “histogram of oriented gradients (HOG) and support vector machine” for glomeruli segmentation and obtained 82.1% F1 score (Kakimoto et al., 2014). Kato used “Segmental histogram of oriented gradients (S-HOG) and support vector machine” for glomeruli segmentation and obtained 86.6% F1-Score (Kato et al., 2015) Gagermayr et al used U-net to obtain 90% accuracy of Dice Coefficient performance metrics (Gadermayr et al., 2019). Bueno recently used “SEGNET and VGG19” for glomeruli identification and characterization, he obtained 81.9% F1-Score (Bueno, Gonzalez-Lopez, et al., 2020). Gallego et al. obtained 86.6% F1-score using basic UNET (Gallego et al., 2021).

Dice coefficient or F1-score (table 1) is mostly used to measure the performance of DCNN model for glomeruli segmentation when image ground truth is given. Some researchers use the meaning of two intersections over union or mean of F1-score.

During segmentation of glomeruli in whole slide images, one of the most recent approaches that could be used to optimizing the segmentation accuracy is ensemble modelling (Gu et al., 2022). In ensemble models, more than one model is combined during prediction process to obtain more predictive accuracy. There are a lot of challenges when a single estimator is built, accuracy might be very low, there might be high variance and bias from image feature noise. To this effect ensemble models are used during prediction.

Performance Metric	Formular
Accuracy	$\frac{TP + TN}{TP + FP + TN + FN}$
Recall	$\frac{TP}{TP+FN}$
precision	$\frac{TP}{TP+FP}$
F1-Score	$\frac{2 * Precision * Recall}{Precision + Recall}$
Mean of F1-score	$\frac{\text{Total of F1 – Scores from prediction}}{\text{Total number of predictions}}$

Table 1. Showing the performance metrics that could be used to evaluate model for glomeruli segmentation (Bueno et al., 2020; Gu et al., 2022).

The concept of Human in the loop (HIL)

For supervised deep learning methods, well-annotated samples are an important prerequisite because large amount of training parameters are involved (Zhou et al., 2014; Siméoni et al., 2020). And due to the bulky amount of effort and time required to generate these samples, annotations could become a very daunting task (Lutnick et al., 2019; Wu et al., 2021). The expert annotation that is required to produce training sets is less practical as compared to the traditional data sets due to the medical dataset's complex nature, imagine glomeruli existing with varied sizes in the same cell (Meng et al., 2021). This therefore becomes a problem when CNNs are applied to medical imaging databases which require an adept knowledge of image annotation to perform (Bobes-Bascarán et al., 2021).

Human in the loop is a concept where human subject is actively involved in machine learning modelling process (Wu et al., 2021). Nevertheless, with humans in the loop interface in any DCNN research, false positive predictions are being deleted in a particular loop of modelling which in turn increases the chance of true positive predictions at the final stage of modelling. (Lutnick et al., 2019).

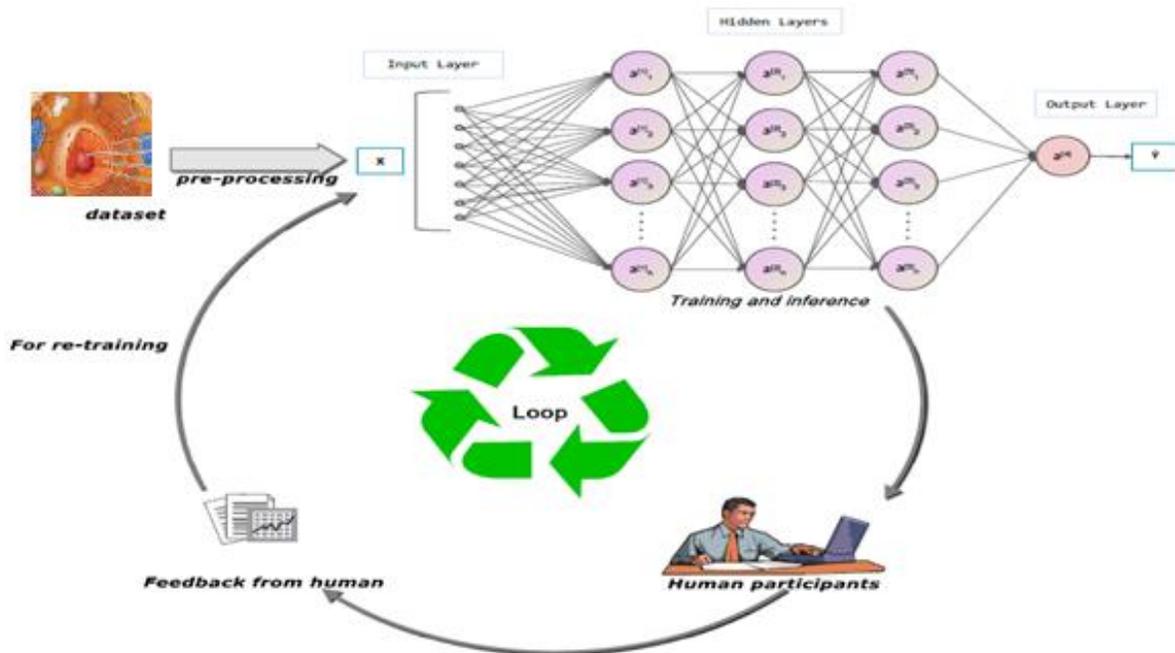


Figure 14. Showing an example of human in the loop interface usage in deep convolutional neural Network (Wu et al., 2021)

The main stages of the workflow in figure 14 above involve preprocessing, training of the network, human participant, human feedback and retraining of the network. After the collection of data, data is preprocessed and serves as input layer into the network, the first prediction from this process is the beginning of the first loop. Human participants look at the result gotten from the first output of the network and correct the errors, then retrain the network for better performance. This process continues until the perfect output is gotten. the gaps in network performances can easily be identified and the predictions of a network are corrected by users using this method Chai and Li, 2020; Gnjatović et al., 2020; Lutnick et al., 2019; Wang et al., 2019; Wu et al., 2021; Yeruva et al., 2020).

This concept of human in the loop had been used in different research work. In security systems, it had been used to check the effect of warning on behavior (Wogalter, 2019). Human in the loop is used in producing software inform of integrated developed environments for developers (Taleb et al., 2019). Most common applications of human in the loop in real world are multimedia search engine, robotics, optimization of health care systems (Wu et al., 2021).

Evaluation of HIL performance

HIL performance is estimated based on the different evaluation criteria and this is challenging in most applications (Recht et al., 2019; Wu et al., 2021). However, it's important to observe the progress of the model performance in each loop although there are no universal evaluation criteria. HIL performance can be determined by comparing the rate of false positives predictions in each loop. The decrease the error of prediction the better the model performance. In other words, the lower the error of prediction the higher the model predicting accuracy (Lutnick et al., 2019). Also, the duration of HIL method can be measured by estimating time per every annotated region and the number of annotations.

$$H = \int_0^R A(r)dr = \tau \left[1 - e^{-\frac{R}{\tau}} \right] \quad (8)$$

Equation 8. Showing the mathematical expression of HIL (Lutnick et al., 2019).

R denoted the overall number annotated region while the number annotated regions is r, τ is the constant of exponential time. H is the normalized time for annotation, A is annotation time per region.

$$P = \left(1 - \frac{H}{B}\right) \times 100 = \left(1 + \frac{\tau}{R} \left[e^{\frac{-R}{\tau}} - 1 \right]\right) \times 100 \quad (9)$$

Equation 9. Formular to measure the time saving performance in HIL (Lutnick et al., 2019).

R denoted the overall number annotated region while the number annotated regions is r, τ is the constant (exponential time). H is the normalized time for annotation, P is the time saving performance, B denotes the baseline annotation time in normalized form.

However, the above estimation method (equation 8 and equation 9) was not used in this research, instead patch counts from good predictions per iteration was used.

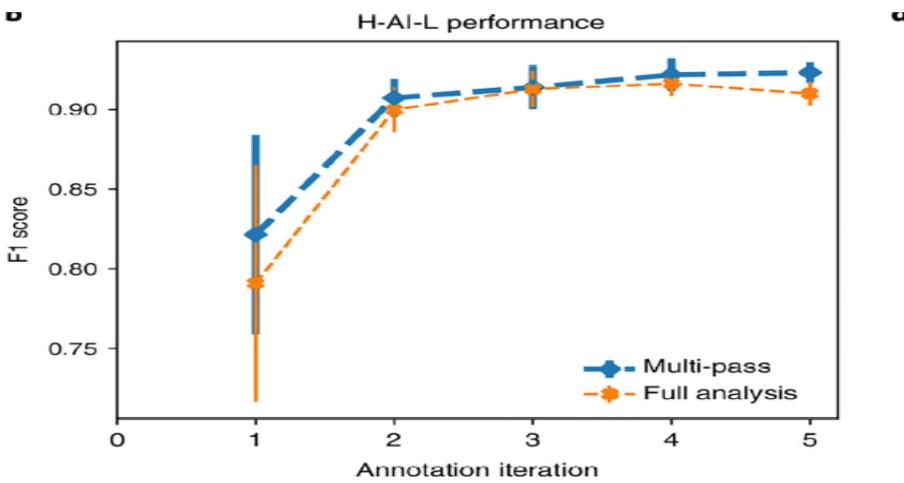


Figure 15. Showing the F1 score of HIL performance in segmentation of glomeruli (Lutnick et al., 2019).

Figure 15 shows how F1 score is used to evaluate HIL performance when the golden standard annotation is given. This shows the results of F1 score in 5 iterations. The F1 increases in each iteration increase.

DATA AND METHODS

In this chapter, the data materials used for this research work are introduced and the methods used are explained in detail.

Data Materials

In this study two data sources were used, and the datasets were named dataset1 and dataset2. The major dataset for this research is dataset1 but dataset2 was introduced to build the control models for the experiment and ZENODO kidney dataset was chosen because it has ground truth labels.

Dataset1

The first datasets used in this research work were obtained from Vilnius national center of pathology at Santariskes. The data sets consist of about 250 whole slide images in sv5 extension stained with hematoxylin and pico-Sirius and their corresponding segmented markup file in tiff format. The markup images consist of segmented glomeruli, background, and other structures. There were manual annotations drawn by kidney pathologist, these annotations were used to produce given masks using halo software. The round green structures (in figure 17) are the segmented glomeruli, and these are the region we are interested in. In this dataset, annotations were converted to these given masks

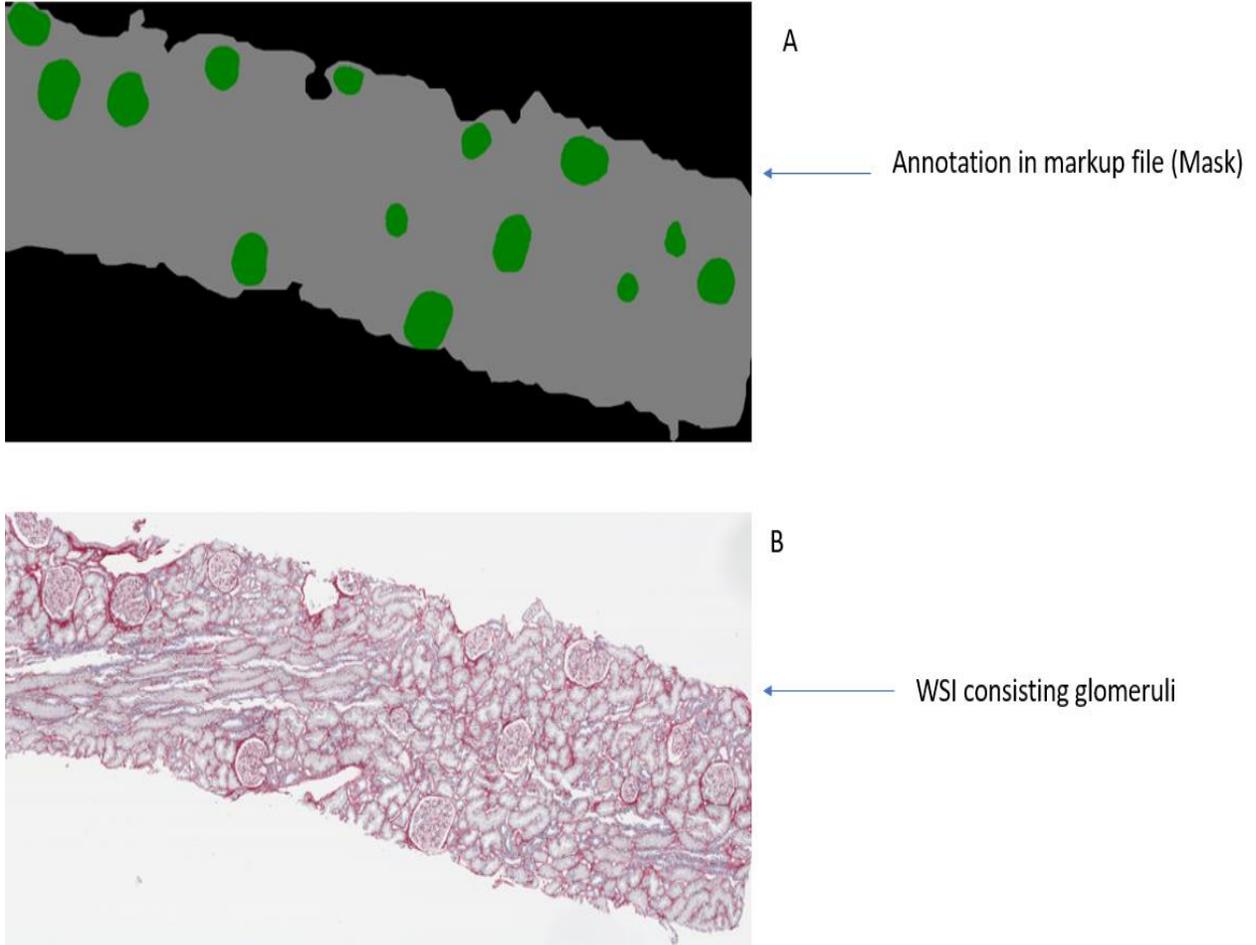


Figure 16. Showing whole slide image and its corresponding markup file.

Figure 16A is one of the markup images, the green circles are showing the segmented glomeruli, while figure 16B is the corresponding whole slide image.

Dataset2

The second dataset used in this research is gotten from an opensource database of ZENEDO. This dataset is used as a control set. ZENEDO kidney dataset is mostly used in research that involves glomeruli segmentation (Gallego et al., 2021). This dataset consists of 47 masks and 47 RGB whole slide images of kidney showing glomeruli stained with H and E, captured with x10 scanner, and stored at 1 μ m.

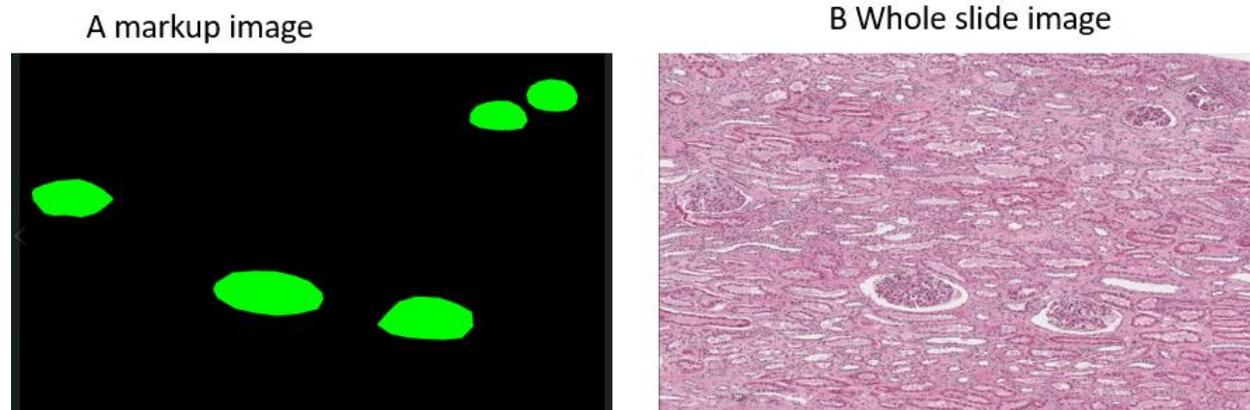


Figure 17. Showing one mask image and WSI from ZENODO kidney glomeruli dataset. (Appendix B)

The mask shows segmented glomeruli annotations in green color and background in black color (figure 17A). Figure 17B is the corresponding RGB WSI. This dataset2 is useful in this research because its annotations are used as ground truth for performance metrics (figure 7). Also, dataset2 was captured at x10 magnification and this is relevant to testing the performance of MAR-UNET.

In this research, this dataset is used in the following investigation.

1. To train, validate and test basic UNET model.
2. To train, validate and test Attention-Residual-U-NET model.
3. To test models from proposed architecture.

Patch extraction

Training, testing, and evaluation of the networks in this research is done at the patch level. Image patches are small portions of whole slide images. The networks input layers take cropped image patches which is better than using complete whole slide image (Hou et al., 2016).

Dataset1:

Normally glomeruli are spherical in shape, and in normal kidney (before kidney section) the diameter of glomeruli ranges from 100 to 350 μm (Gallego et al., 2018).

During image patch extraction from dataset1, mask is being converted to border using “find_contours” function from “skimage.measure” library package in python, this function returns borders of the glomeruli in the masks and then the borders were used to draw a boundary box on the on glomeruli in whole slide image using “regionprops” from “skimage.measure” library package in python.

These boundary boxes range from 21 x 21 (width (pixels) x length (pixels)) to 381 x 366 (width (pixels) x length (pixels)) in size. In most medical research 256 x 256 (width (pixels) x length (pixels)) image patch is being used. Therefore, all these drawn boundary boxes were adjusted to 256 x 256 (width (pixels) x length (pixels)) on all the located glomeruli and cropped out, this makes the first image patch size used, then 176 x 176 (width (pixels) x length (pixels)), 336 x 336 (width (pixels) x length (pixels)) and 416 x 416 (width (pixels) x length (pixels)).

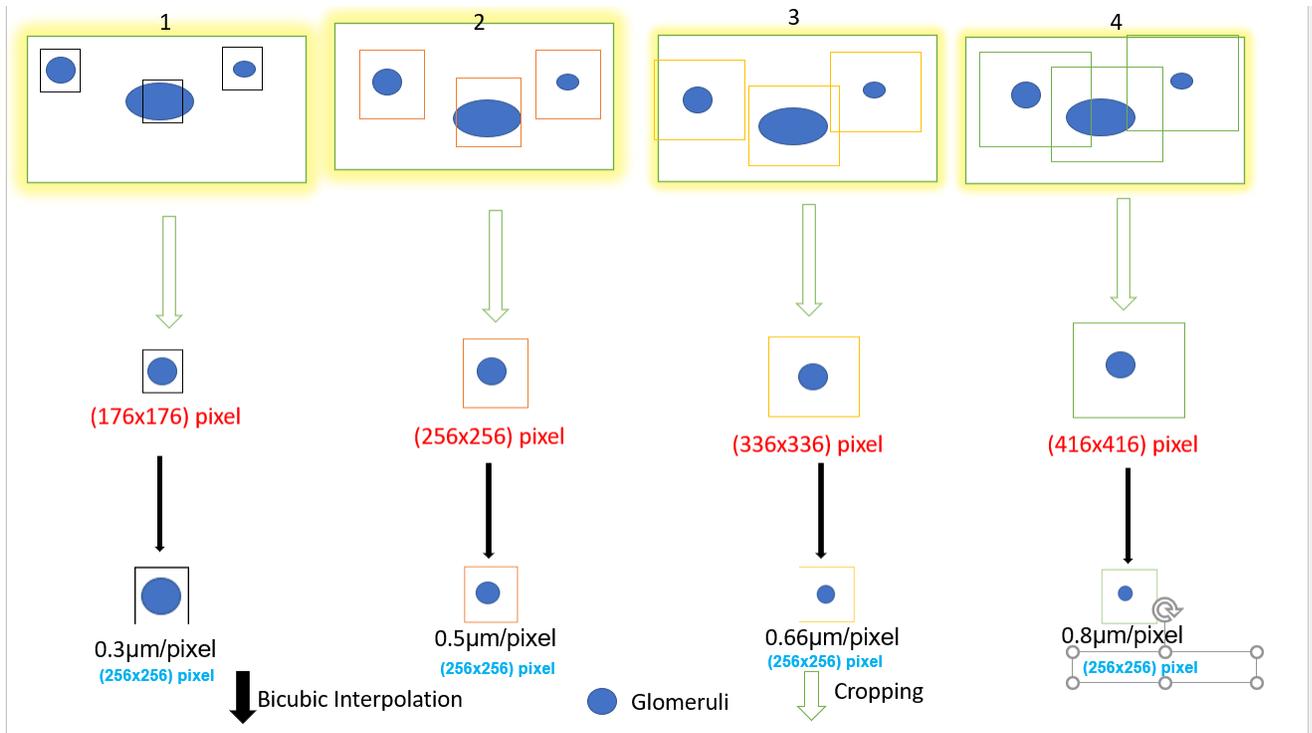


Figure 18. Showing 4 different resolutions were extracted (Appendix A, file 'patch_extraction.py').

All the patches were interpolated using the bicubic interpolation method figure 18, bicubic interpolation was chosen because it is very fast and more accurate than nearest neighbor or bilinear (Li et al., 2019).

$$\text{Size} = (n \cdot \text{width}, n \cdot \text{height}) \quad (10)$$

Equation 10. Showing how the size of pixels in the whole slide image is calculated. n is the value of apparent magnification (resolution in $\mu\text{m}/\text{pixel}$) of the WSI.

In most cases resolution of WSI at 20x is usually $0.5\mu\text{m}$ per pixel (Marini et al., 2021).

98.8% of the dataset1 was captured at 20x, and the value of n is $0.4925 \mu\text{m}$ per pixel (approximately $0.5 \mu\text{m}$ per pixel).

Dataset 2:

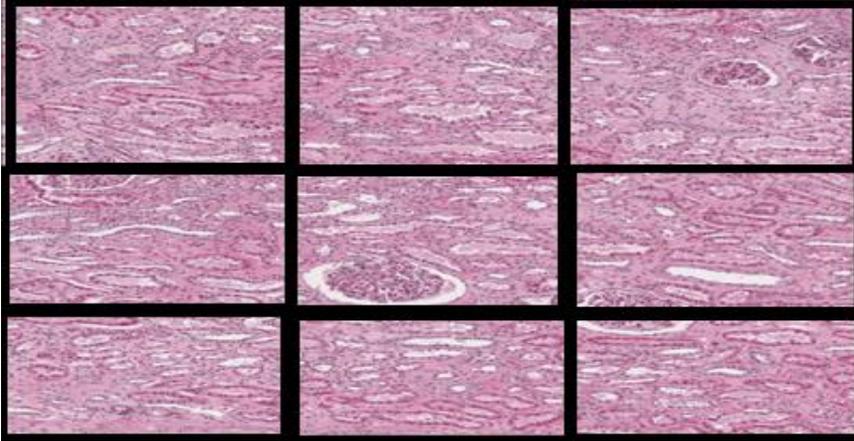


Figure 18. Showing how patches in dataset 2 were extracted.

From figure 19, one single image patch resolution was cropped out from whole slide image both the widths and heights were 256 pixels (1 $\mu\text{m}/\text{pixel}$) (Appendix A, file 'svs_slice.py').

.

Preprocessing

Thresholding

Thresholding is a type of image segmentation; Thresholding method is required on the markup patches to convert the multiclass mask image to single class (binary) since the major task is to segment glomeruli. Using mode attribute from pillow library in python to checking the color channel of masks in dataset1, the color mode is 'palettized', which means there are more than 3 colors [0, 64, 128, 192], the threshold of glomeruli pixels consists of 64 and 128 and were selected and set as 1 while 0 and 64 were set to 0. Therefore, the threshold of pixel intensities of masks was converted to 0 and 1. In each mask image **I**, the glomeruli pixels are set to 1 and all other pixels are set to 0 resulting to the new mask up image **IB**.

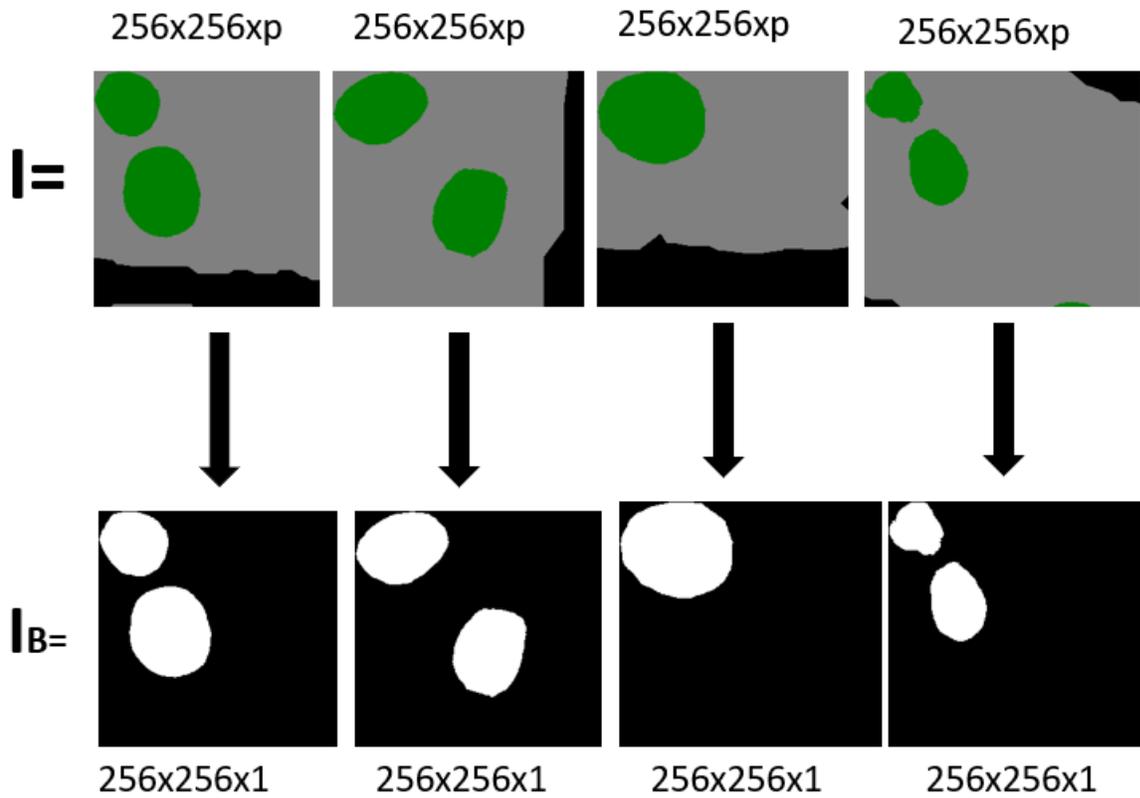


Figure 19. Showing image thresholding method (Kotte et al., 2018) (Appendix A, file 'multiclas_to_binary.py').

Figure 20 above shows how the multiclass marks images were converted to single class (binary)

Stain color normalization

Color variations influence model performance (Tellez et al., 2019). Therefore, in this current study it is necessary to regulate color variations in image patches. Dataset2 has two different types of color. Dataset1 was produced from hematoxylin and pico-Sirius stain while dataset2 was produced from H and E stain. To unify these color variations, a normalization algorithm is applied.

Procedure according to (Macenko et al., 2009), the results are shown in figure 21. "Step1: import of RGB Image patch (P) using 'IMREAD' function from cv2 library in python. Step2: Conversion of image P RGB color space to optical density (D_p)

$$D_p = -\log_{10}P$$

11

From equation 11 (Macenko et al., 2009), D_p is the optical density space, P is the RGB input image.

Step3: Filter D_p using 0.15 threshold for transparent pixels (Default threshold value is 0.15), denotes filtered D_p as F. Step4: Compute the eigen values and vector of F. Step5: data is projected and normalized to unit length. Step 6: "Calculation of angle of each point wrt the first single value decomposition direction". Step 7: normalize the concentration of stain by calculating the percentile of the angle. Step 8: convert back to OD and save to disk.

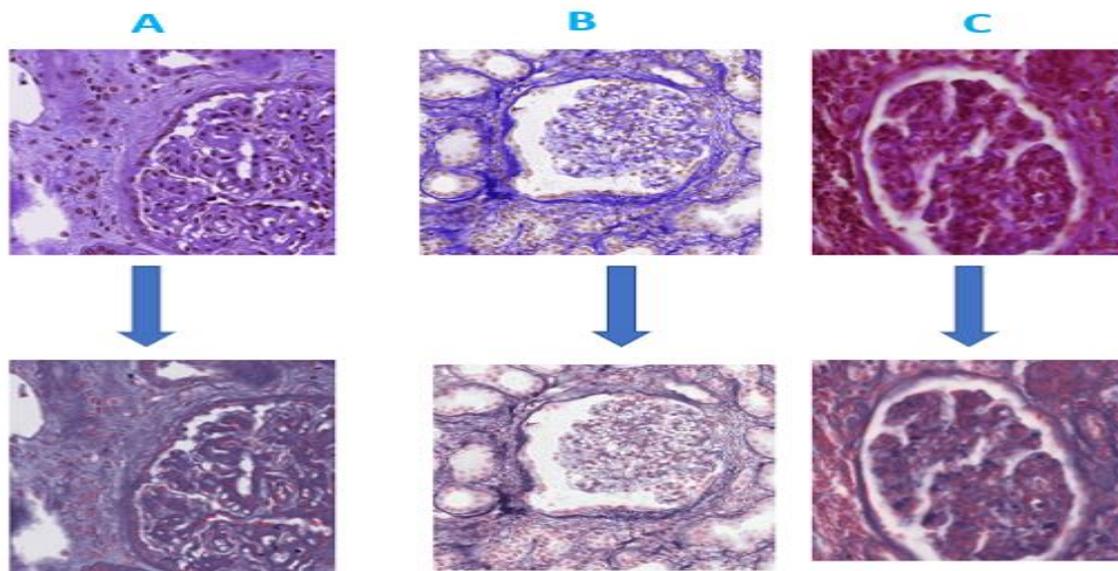


Figure 20. showing the result of color normalization in dataset1 and dataset2. (Code: Appendix A, file "color_normalisation.py").

Network topology

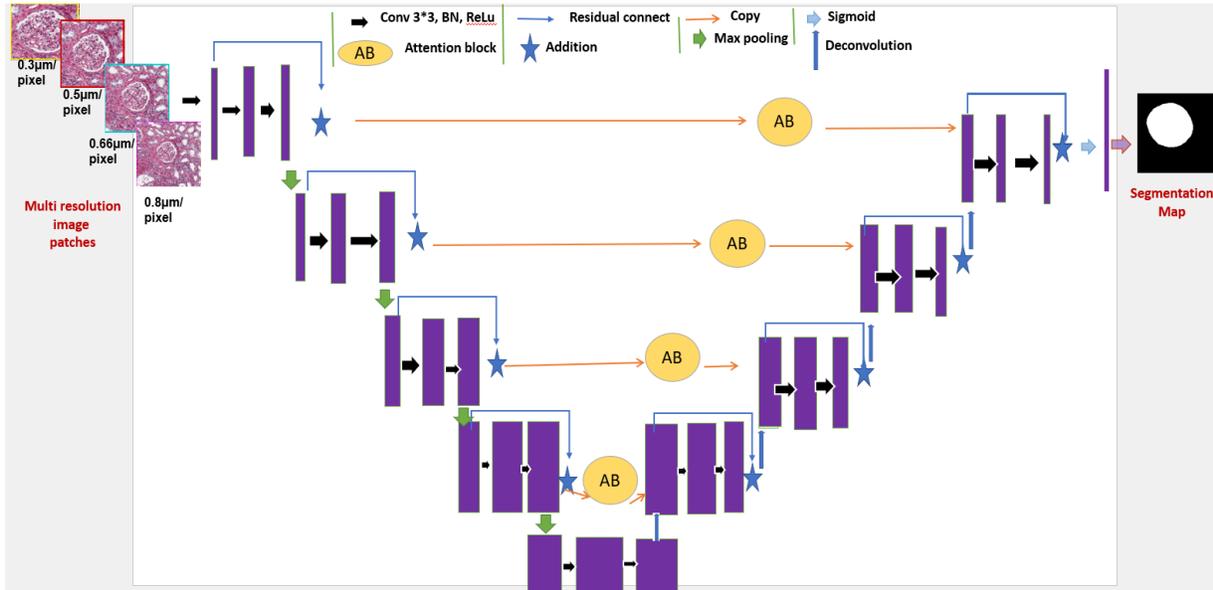


Figure 21. Showing the schematic diagram of the proposed architecture multi-resolution-attention-residual-UNET (MAR-UNET) (Appendix A, file 'all_model.py').

Figure 22 depicts the proposed architecture called MAR-UNET. This innovation combines a basic UNET architecture using multiple resolution image patches with attention blocks between the encoding and the decoding units, and residual connection (figure 8c). This architecture consists of the expansive path called the encoder and the contractive path called the decoder. The encoder consists of an input layer which takes the multiresolution image patches, In between each residual connection is the attention block (figure 10). The decoder involves the deconvolutional operations and segmentation map, and each decoding unit receives a signal from the encoding units through residual connection. The input layer has one channel that takes 4 different resolutions of image patch, images pass through the input layer individually at random.

Network hyperparameters	Set values
Convolutional kernel size	3 x 3
Up sampling filter size	2 x 2
Polling size	2 x 2
Activation function	Sigmoid (figure 5b, equation 4)
Optimizer	Adam
Loss function:	Binary Focal Loss (gamma=2)
Batch normalization	True (equation 5)
Dropout rate	0.0
Batch size	1

Table 4. Showing the parameters set before training of the networks.

For the evaluation of the model performance, the Mean of the F1-score (Dice metrics in table 1) was used when prediction was made on 500 image patches from ZENODO kidney dataset.

To optimize glomeruli segmentation in dataset 2, ensemble models were used.

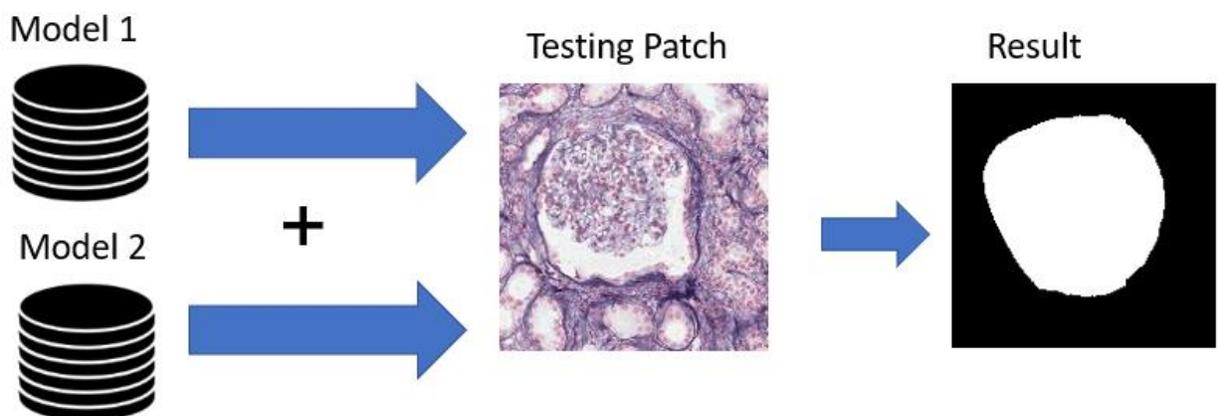


Figure 22. Showing ensemble model (Appendix A, file 'Em_model.py').

Proposed workflow

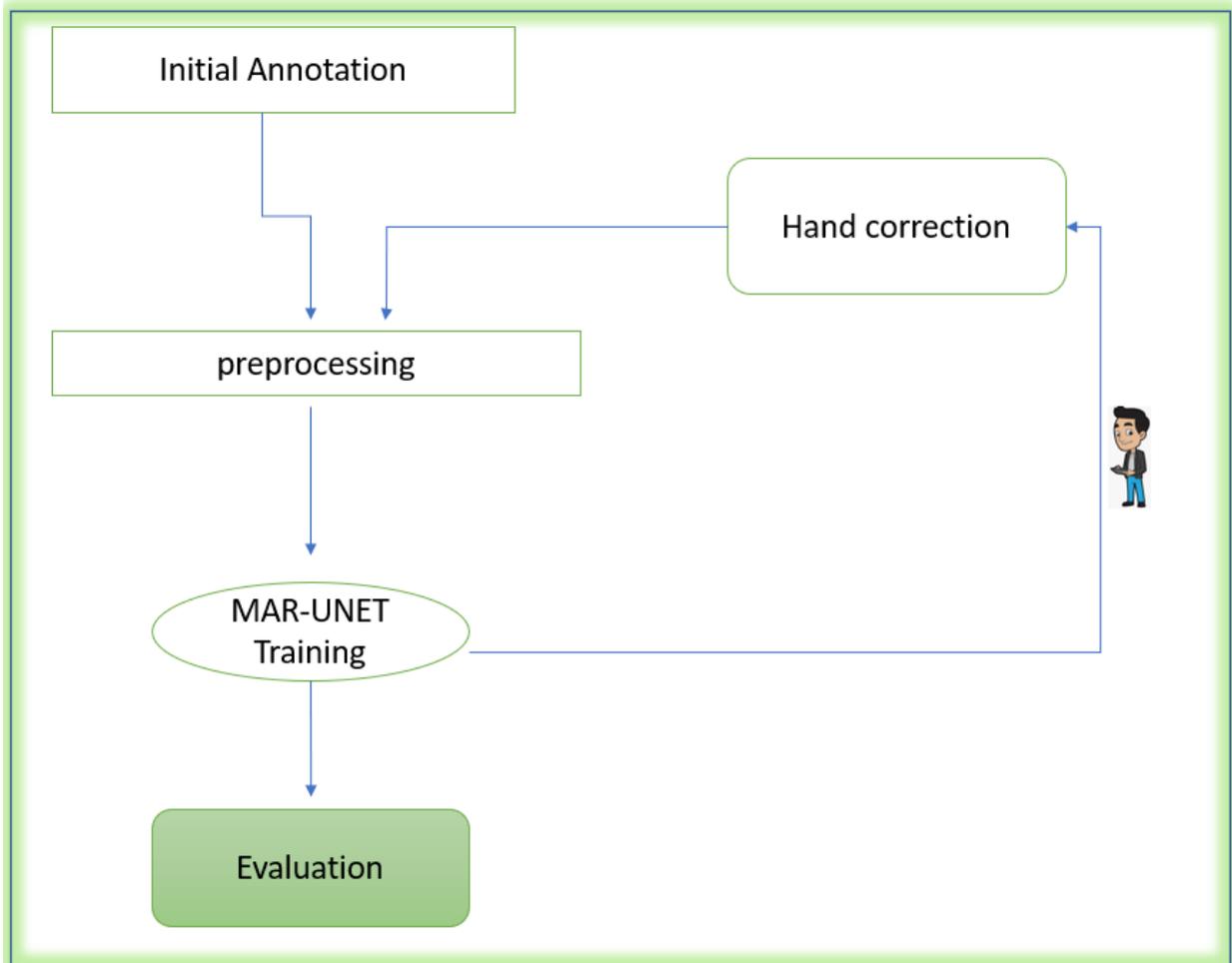


Figure 23. Showing the flow diagram of the proposed workflow.

The above workflow shows how MAR-UNET was used to generate more annotations while humans would do some corrections, this creates an iterative patch extraction workflow.

The pathologist at Vilnius pathology center annotates some glomeruli in whole slide Images, then trains a network with halo software, then tests this network/model on some whole slide images. The results obtained are segmented multiclass markup files or masks (Dataset 1).

The dataset1 used in this research is a product of this built network (M0) from halo software. Multi resolution image patches were extracted from these given samples using the proposed patch extraction method, then preprocessed. This gives a set of training and validation image patches and their respective binary masks. These samples are used to train a MAR-UNET network (figure 22) which produced another model. This model is then used to segment glomeruli in a new set of

preprocessed multiresolution image patches from another set of images from 250 WSI that were not initially involved MAR-UNET network. Then the accuracy of MAR-UNET was tested on 500 image patches from ZENODO kidney glomeruli datasets. These steps were repeated several times until the best model was produced based. The mean of F1- score was calculated in each iteration (figure 15)

Experimental setup

The following UNET architectures were configured in this investigation.

Basic U-net architecture was setup and configured (figure 7 and table 4). This is a basic U-net architecture, this architecture uses single resolution image patches extracted from ZENODO dataset (dataset 2) at the input layer. This is set up to compare its performance with MAR-UNET performance.

Attention-Res-U-net architecture was setup and configured (table 4). This is a basic U-net architecture with residual block and attention block, this architecture uses a single resolution image patches extracted from ZENODO dataset (dataset 2) at the input layer. To investigate the role of attention blocks and residual blocks in glomeruli segmentation using deep learning.

MAR-U-NET architecture was setup and configured (figure 22 and table 4). This is a basic U-net architecture with residual block and attention block and uses 4 different resolutions of image patches extracted from dataset1 at the input layer.

Datasets	Basic U-net	Attention-Res-U-net	MAR-U-NET
Training	0	0	16580
Validation	0	0	2072
Test	0	0	0

Table 5. Showing the number of image patches used from dataset1

Datasets	Basic U-net	Attention-Res-U-net	MAR-U-NET
Training	6966	6966	0
Validation	775	775	0
Test	500	500	500

Table 6. Showing the number of image patches from ZENODO dataset used by the network

RESULTS

DCNN models for kidney glomeruli segmentation were trained on Nvidia RTX3050 GPU (graphical processing units), 24G RAM using keras machine learning libraries on TensorFlow framework. All training in this research was stopped at 100 epochs. No early stopping or call back was used.

Using dataset1, the first MAR-UNET model was built and used to generate more annotations on the image patches. M1 prediction accuracy for glomeruli segmentation was tested on 500 patches from ZENODO dataset (dataset2) and the mean of F1-score was calculated, to know if there is any tangible relationship between these two datasets.

This same procedure continues until the best model was chosen after minimal correction of prediction and high accuracy on ZENODO kidney dataset.

Model	Number of Training patches	Number of Validation patches	Number of Additional Patches	Mean F1-Score Dataset 2 (%)
M1	3713	413	1094	45.13
M2	4698	522	1891	51.30
M3	6399	712	2963	59.09
Best MAR-UNET(M5)	16580	1658	0	61.38

Table 7. Summarize the results using human in the loop annotation strategy.

Table 7 explained how new curated annotations in image patches were added to training and validation set after prediction my built model and careful observations by the pathologist. The accuracy of these models was checked on newly introduced ZENODO dataset patch images.

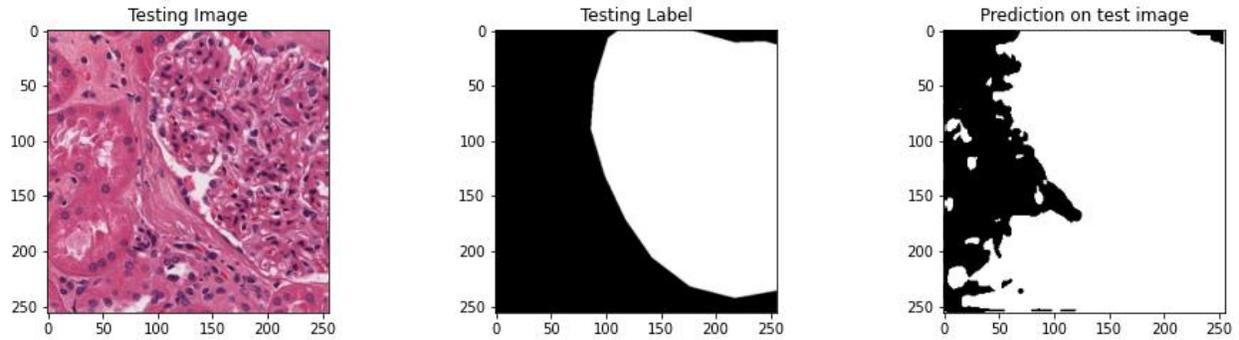


Figure 24. Showing M1 prediction. 58.09% F1-score, 45.13% mean of F1-score using 500 image patches from dataset 2 (1 μm per pixel) (Appendix A, file 'fig.pptm').

This is the lowest F1-score observed in this research, at this stage it was difficult for the network to classify some pixels correctly.

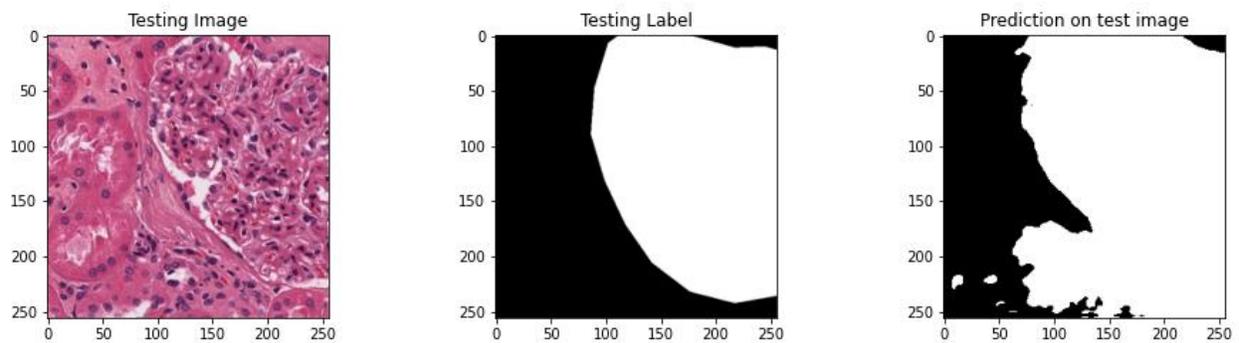


Figure 25. Showing M2 prediction. 71.67% F1-score, 51.30% mean of F1-score using 500 image patches from dataset 2 (1 μm per pixel) (Appendix A, file 'fig.pptm').

The F1- score became better, when more patches were added, from predictions of model M1.

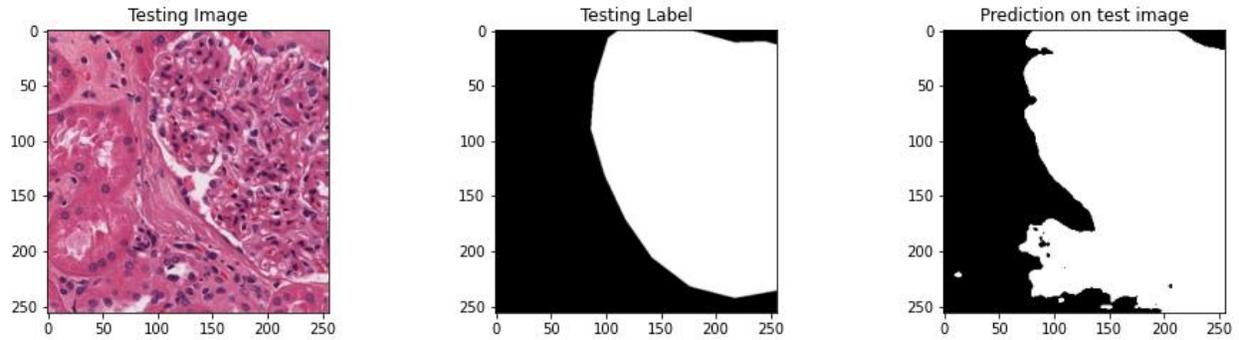


Figure 26. Showing M3 prediction. 76.78% F1-score, 59.09 mean of F1-score using 500 image patches from dataset 2 (1 μm per pixel) (Appendix A, file 'fig.pptm').

The average F1-score (mean IOU) of 500 prediction was increased using M3, as more curated annotations were added by pathologists.

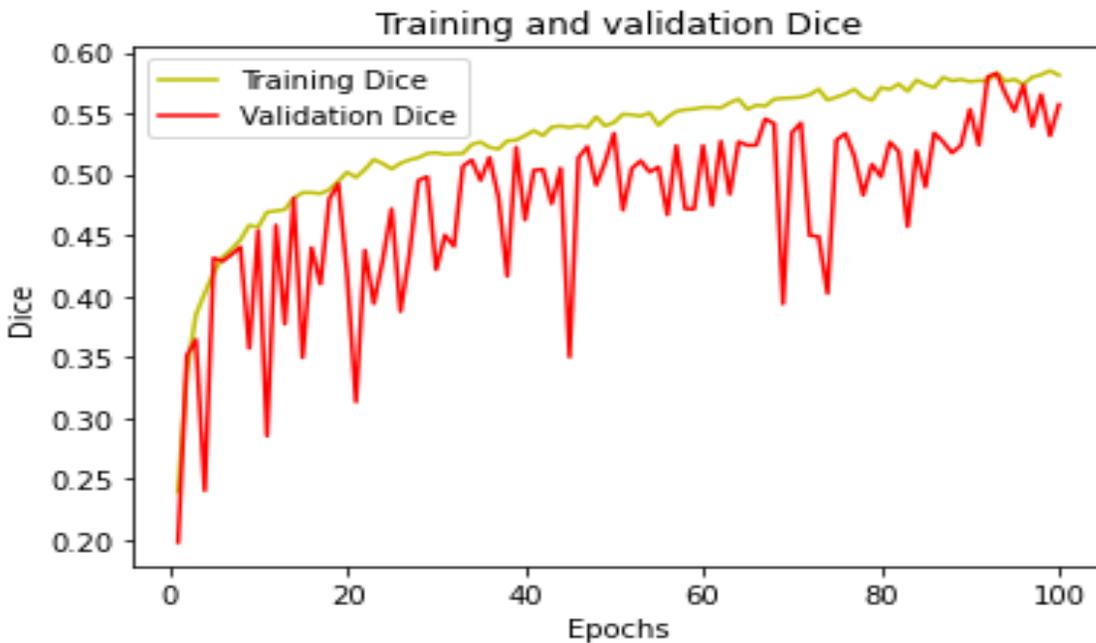


Figure 27. Training and validation dice of the Best MAR-UNET (loop5) chosen after several iterations from and predictions using more curated image patches (Appendix A, file 'fig.pptm').

The validation dice score could be better if more epochs were accomplished.

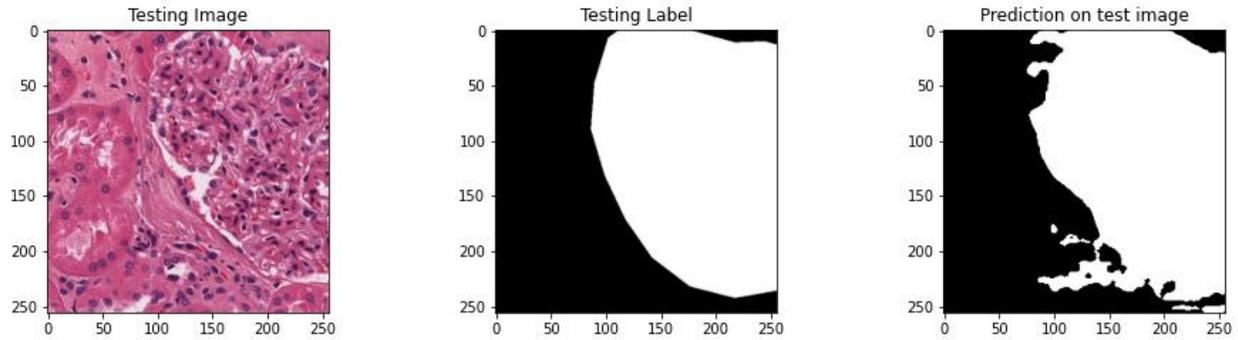


Figure 28. Best MAR-UNET prediction. 82.75% F1-score, 61.38% mean of F1-score using 500 image patches from dataset 2 (1 μm per pixel) (Appendix A, file 'fig.pptm').

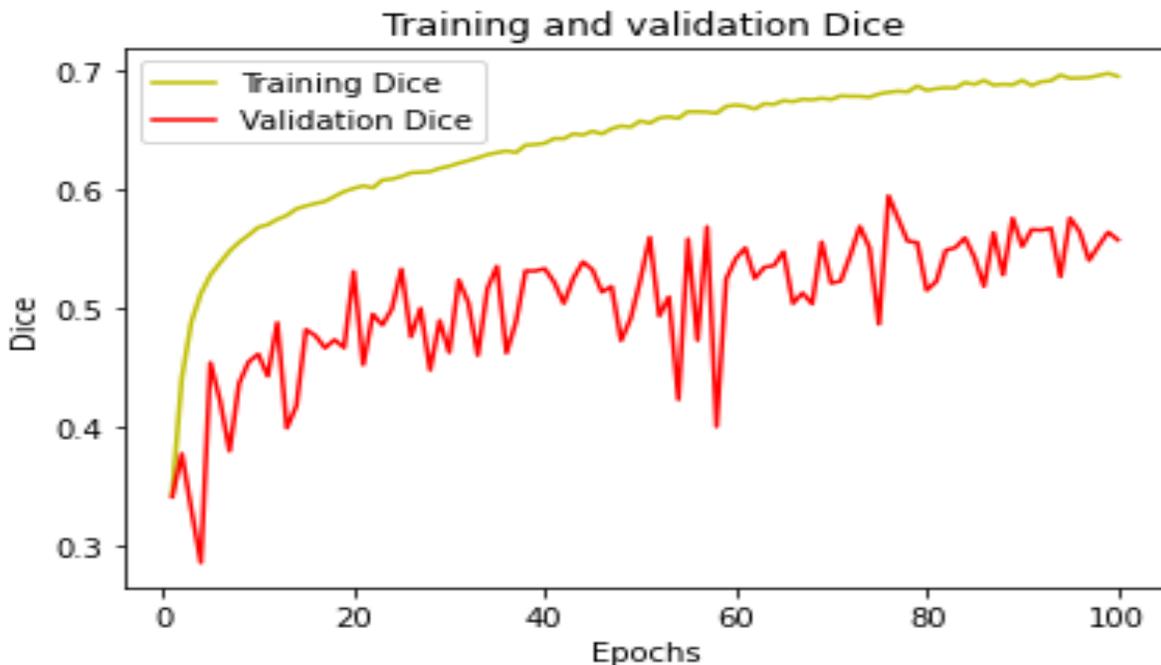


Figure 29. Training and validation dice of Basic U-net architecture using ZENODO kidney dataset (Dataset2), patches were extracted at x10 magnification (1 μm per pixel) (Appendix A, file 'fig.pptm').

The much difference in training and validation dice scores might be due to zero dropout hyperparameter used in building the network. Most recent architectures are not using dropout.

It might also be because of the small batch size used in training the network.

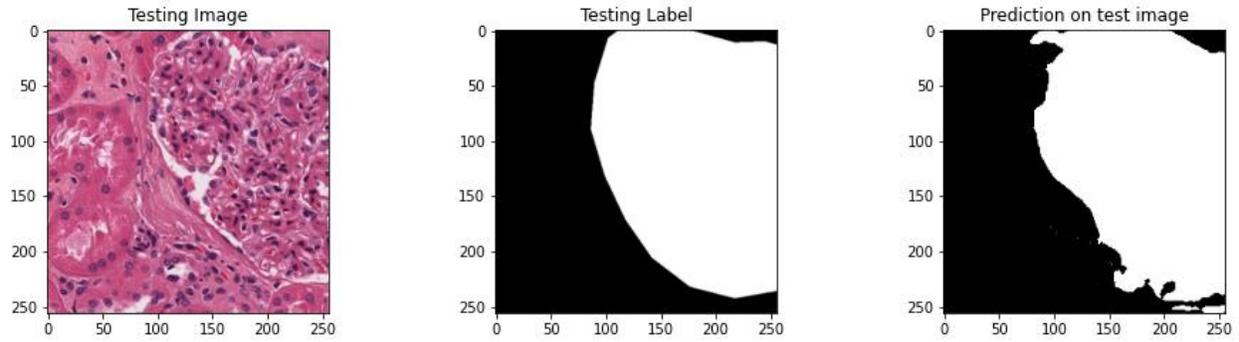


Figure 30. Basic UNET prediction on dataset 2 (1 μm per pixel). 87.05% F1-score, 76.61% mean F1-score using 500 image patches from dataset 2 (1 μm per pixel) (Appendix A, file 'fig.pptm').

The mean of F1-score obtained from using Basic UNET for glomeruli segmentation in 500 image patches from dataset 2 (1 μm per pixel) is more than the result obtained from the best model of MARUNET.

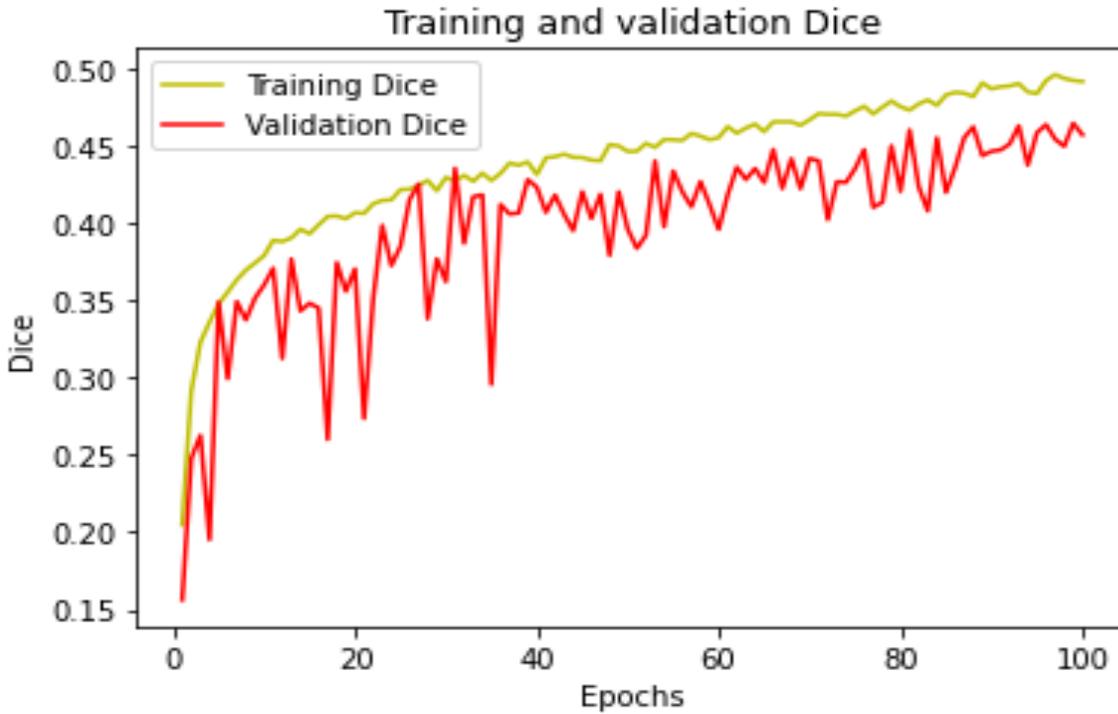


Figure 31. Training and validation dice of Attention-res-UNET network using ZENODO kidney dataset (Dataset2), patches were extracted at x10 magnification (1 μm per pixel) (Appendix A, file 'fig.pptm').

Attention-res-UNET network has a better fit than basic UNET on ZENODO kidney dataset (Dataset2) even with zero dropout. This might be due to the ability of the residual blocks to regularize vanishing gradient problems in the deep layers of CNN architecture.

SIG

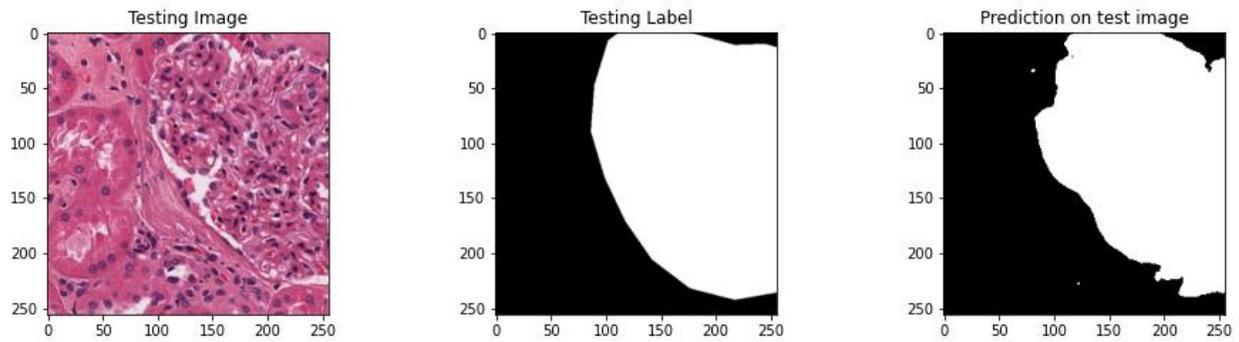


Figure 32. Attention-Res-UNET's prediction. 88.91% F1-score, 78.4% mean of F1-score using 500 image patches from dataset 2 (1 μm per pixel) (Appendix A, file 'fig.pptm').

Attention-Res-UNET's has the highest mean of F1-score.

DISCUSSION

The word multiscale is used interchangeably with multiresolution DCNN research. Some research considered a neural network to be multiscale neural network when different resolutions of an input image are given to a network through multiple channels or many input layers (Wetteland et al., 2020). Mostly in this type of MCNN architecture, concatenation of feature maps is used. However, in some research, a neural network is multiscale neural network when different physical dimensions of inputs are being used in the network input layer through multiple input channel(Godinez et al., 2017) and in some MCNN architectures, different up sampling filter size are used but one input layer (Aatresh et al., 2021). In contrast to these, a single input channel was used in this research but taking more than one resolution of an input image.

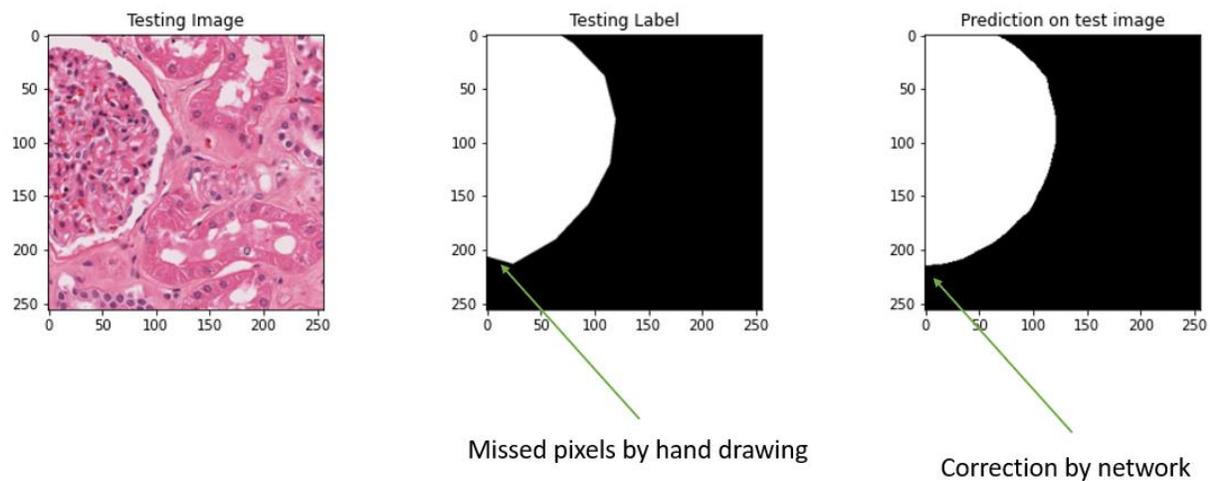


Figure 33. Showing the effect of using neural networks to draw annotations

Using an iterative patch annotation strategy is better than relying on hand drawing annotation.

Glomeruli is said to be round, when drawing annotations with hand, there might be some pixels left out. These pixels might also be important to differentiate a diseases glomerulus from normal glomeruli in kidney diagnosis.

Gallego et al (Gallego et al., 2021) obtained 86.6% F1-score for glomeruli segmentation using basic UNET architecture on ZENODO kidney dataset (H& E), this is close to the result obtained from this research when a basic UNET architecture was use on ZENODO kidney dataset. If we are to consider a single patch prediction (F1-score of 87.05%).

The 61.38% mean of F1-score obtained from glomeruli segmentation in 500 image patches from dataset2 (1 μm per pixel) using the best MAR-UNET model is low to result obtained using basic UNET or attention-res-UNET. This might be because the resolution of test image patches is more than all the resolutions used to build MAR-UNET. MAR-UNET could not classify some pixels correctly in ZENODO kidney dataset.

A single attention residual UNET model produced from using ZENODO glomeruli image patches for training and validation has the highest mean of F1-scores on 500 image patches from dataset2 (1 μm per pixel). Attention residual UNET performs better than basic UNET for glomeruli segmentation.

Using an attention residual UNET model and the best model of MAR-UNET model by model ensemble (figure 23) gives a better performance of 89.89% F1-sore and 79.38% mean of F1-score on 500 image patches from dataset 2 (1 μm per pixel).

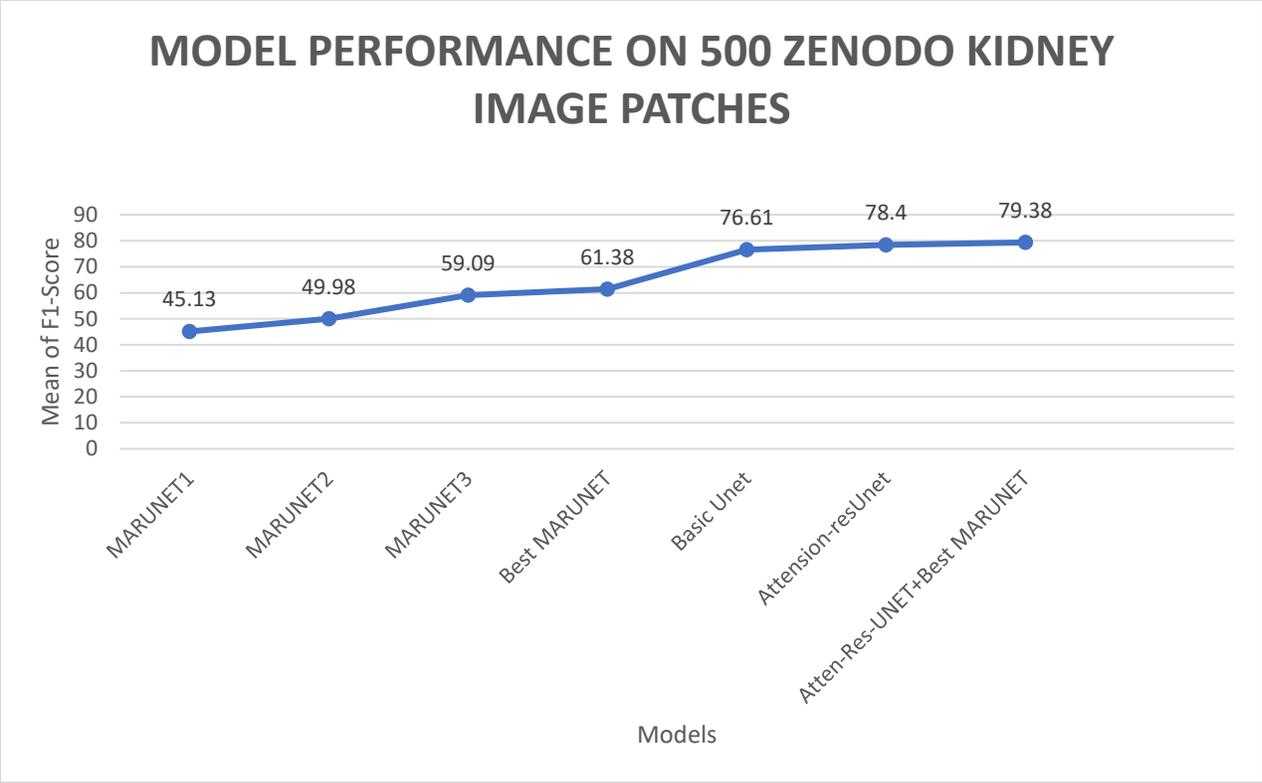


Figure 34. Showing the performance of all the models for glomeruli segmentation
Resolution of neural network input can affect the performance of neural network.

CONCLUSIONS

A new method of extracting multiresolution image patches is by capturing the region of interest at the center of the patch image. A multiresolution deep convolutional neural network can optimize glomeruli segmentation in a whole slide image using ensemble models. A new way of evaluating human in the loop annotation strategy instead of using time per annotation area in whole slide image is counting the number of segmented images patches that are correctly predicted.

RECOMMENDATION

The training of Attention-residual-UNET networks is very slow, it is highly recommended to use high GPU workstation when attention-residual-UNET architecture or customizing CNN processing unit in the future.

ACKNOWLEDGEMENTS

My acknowledgement goes to the Almighty God who has given me the physical, psychological, and intellectual acumen to complete this research successfully. Also, to my supervisor Dr. ALLAN RASMUSSEN for his motivation and encouragement during this research work. Moreso, to Vilnius kidney pathology research team headed by Prof. Dr ARVYDAS LAURINAVICIUS.

To my academic coordinator Prof. AUDRONĖ JAKAITIENĖ, who has coordinated this master's thesis and given updates on various submission deadlines at the appropriate time.

I dedicate this research to my grandmother Mrs. S.O. Odeku for sponsoring my master's program in the Republic of Lithuania.

REFERENCES

1. Aatresh, A.A., Alabhya, K., Lal, S., Kini, J. and Saxena, P.P. (2021), "LiverNet: efficient and robust deep learning model for automatic diagnosis of sub-types of liver hepatocellular carcinoma cancer from H&E stained liver histopathology images", *International Journal of Computer Assisted Radiology and Surgery*, Springer Science and Business Media Deutschland GmbH, Vol. 16 No. 9, pp. 1549–1563.
2. Aeffner, F., Zarella, M.D., Buchbinder, N., Bui, M.M., Goodman, M.R., Hartman, D.J., Lujan, G.M., et al. (2019), "Introduction to digital image analysis in whole-slide imaging: A white paper from the digital pathology association", *Journal of Pathology Informatics*, Vol. 10 No. 1, available at:https://doi.org/10.4103/jpi.jpi_82_18.
3. Alom, M.Z., Yakopcic, C., Hasan, M., Taha, T.M. and Asari, V.K. (2019), "Recurrent residual U-Net for medical image segmentation", *Journal of Medical Imaging*, Vol. 6 No. 01, available at:<https://doi.org/10.1117/1.jmi.6.1.014006>.
4. Anghel, A., Stanisavljevic, M., Andani, S., Papandreou, N., Rüschoff, J.H., Wild, P., Gabrani, M., et al. (2019), "A High-Performance System for Robust Stain Normalization of Whole-Slide Images in Histopathology", *Frontiers in Medicine*, Vol. 6, available at:<https://doi.org/10.3389/fmed.2019.00193>.
5. Anwar, S.M., Majid, M., Qayyum, A., Awais, M., Alnowami, M. and Khan, M.K. (2018), "Medical Image Analysis using Convolutional Neural Networks: A Review", *Journal of Medical Systems*, Springer New York LLC, 1 November, available at:<https://doi.org/10.1007/s10916-018-1088-1>.
6. Arogundade, F.A., Omotoso, B.A., Adelakun, A., Bamikefa, T., Ezeugonwa, R., Omosule, B., Sanusi, A.A., et al. (2020), "Burden of end-stage renal disease in sub-Saharan Africa", *Clinical Nephrology*, Vol. 93 No. 1, available at:<https://doi.org/10.5414/CNP92S101>.
7. Blaine, J., Chonchol, M. and Levi, M. (2015), "Renal control of calcium, phosphate, and magnesium homeostasis", *Clinical Journal of the American Society of Nephrology*, American Society of Nephrology, Vol. 10 No. 7, pp. 1257–1272.
8. Bobes-Bascarán, J., Mosqueira-Rey, E. and Alonso-Ríos, D. (2021), "Improving Medical Data Annotation Including Humans in the Machine Learning Loop", available at:<https://doi.org/10.3390/engproc2021007039>.

9. Bohle, A., Eichenseher, N., Fischbach, H., Neild, G.H., Wehner, H., Edel, H.H., Losse, H., et al. (1976), "The different forms of glomerulonephritis morphological and clinical aspects, analyzed in 2500 patients", *Klinische Wochenschrift*, Vol. 54 No. 2, available at:<https://doi.org/10.1007/BF01468771>.
10. Bornstein, J., McCullough, K., Combe, C., Bieber, B., Jadoul, M., Pisoni, R., Mariani, L., et al. (2014), "DIALYSIS. EPIDEMIOLOGY, OUTCOME RESEARCH, HEALTH SERVICES 2", *Nephrology Dialysis Transplantation*, Vol. 29 No. suppl 3, pp. iii516–iii527.
11. Bouvrie, J. (2006), "Notes on convolutional neural networks", *In Practice*, available at:<https://doi.org/http://dx.doi.org/10.1016/j.protcy.2014.09.007>.
12. Bueno, G., Fernandez-Carrobles, M.M., Gonzalez-Lopez, L. and Deniz, O. (2020), "Glomerulosclerosis identification in whole slide images using semantic segmentation", *Computer Methods and Programs in Biomedicine*, Vol. 184, available at:<https://doi.org/10.1016/j.cmpb.2019.105273>.
13. Bueno, G., Gonzalez-Lopez, L., Garcia-Rojo, M., Laurinavicius, A. and Deniz, O. (2020), "Data for glomeruli characterization in histopathological images", *Data in Brief*, Vol. 29, available at:<https://doi.org/10.1016/j.dib.2020.105314>.
14. Chai, C. and Li, G. (2020), "Human-in-the-loop Techniques in Machine Learning", *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*.
15. Chandragiri, S., Raju, S., Mukku, K.K., Babu, S. and Uppin, M.S. (2016), "Idiopathic nodular glomerulosclerosis: Report of two cases and review of literature", *Indian Journal of Nephrology*, available at:<https://doi.org/10.4103/0971-4065.164233>.
16. Cireşan, D.C., Giusti, A., Gambardella, L.M. and Schmidhuber, J. (2012), "Deep neural networks segment neuronal membranes in electron microscopy images", *Advances in Neural Information Processing Systems*, Vol. 4.
17. Corvò, A., Westenberg, M.A., Wimberger-Friedl, R., Fromme, S., Peeters, M.M.R., van Driel, M.A. and van Wijk, J.J. (2019), "Visual analytics in digital pathology: Challenges and opportunities", *Eurographics Workshop on Visual Computing for Biology and Medicine, VCBM 2019*, available at:<https://doi.org/10.2312/vcbm.20191240>.
18. Curthoys, N.P. and Moe, O.W. (2014), "Proximal tubule function and response to acidosis", *Clinical Journal of the American Society of Nephrology, American Society of Nephrology*, Vol. 9 No. 9, pp. 1627–1638.
19. Danziger, J. and Zeidel, M.L. (2015), "Osmotic homeostasis", *Clinical Journal of the American Society of Nephrology, American Society of Nephrology*, Vol. 10 No. 5, pp. 852–862.

20. Dimitriou, N., Arandjelović, O. and Caie, P.D. (2019), “Deep Learning for Whole Slide Image Analysis: An Overview”, *Frontiers in Medicine*, Frontiers Media S.A., 22 November, available at:<https://doi.org/10.3389/fmed.2019.00264>.
21. Ebert, N. and Schaeffner, E. (2018), “New biomarkers for estimating glomerular filtration rate”, *Journal of Laboratory and Precision Medicine*, Vol. 3, available at:<https://doi.org/10.21037/jlpm.2018.08.07>.
22. Farnik, H. and Zeuzem, S. (2012), “New antiviral therapies in the management of HCV infection”, *Antiviral Therapy*, available at:<https://doi.org/10.3851/IMP2127>.
23. Feehally, J. (2010), *Comprehensive Clinical Nephrology*, *Comprehensive Clinical Nephrology*, available at:<https://doi.org/10.1016/C2009-0-46539-5>.
24. Gadermayr, M., Dombrowski, A.K., Klinkhammer, B.M., Boor, P. and Merhof, D. (2019), “CNN cascades for segmenting sparse objects in gigapixel whole slide images”, *Computerized Medical Imaging and Graphics*, Vol. 71, available at:<https://doi.org/10.1016/j.compmedimag.2018.11.002>.
25. Gallego, J., Pedraza, A., Lopez, S., Steiner, G., Gonzalez, L., Laurinavicius, A. and Bueno, G. (2018), “Glomerulus classification and detection based on convolutional neural networks”, *Journal of Imaging*, Vol. 4 No. 1, available at:<https://doi.org/10.3390/jimaging4010020>.
26. Gallego, J., Swiderska-Chadaj, Z., Markiewicz, T., Yamashita, M., Gabaldon, M.A. and Gertych, A. (2021), “A U-Net based framework to quantify glomerulosclerosis in digitized PAS and H&E stained human tissues”, *Computerized Medical Imaging and Graphics*, Elsevier Ltd, Vol. 89, available at:<https://doi.org/10.1016/j.compmedimag.2021.101865>.
27. Gibson, E., Li, W., Sudre, C., Fidon, L., Shaker, D.I., Wang, G., Eaton-Rosen, Z., et al. (2018), “NiftyNet: a deep-learning platform for medical imaging”, *Computer Methods and Programs in Biomedicine*, Elsevier Ireland Ltd, Vol. 158, pp. 113–122.
28. Gnjatović, M., Maček, N. and Adamović, S. (2020), “Putting humans back in the loop: A study in human-machine cooperative learning”, *Acta Polytechnica Hungarica*, Vol. 17 No. 2, available at:<https://doi.org/10.12700/APH.17.2.2020.2.11>.
29. Godinez, W.J., Hossain, I., Lazic, S.E., Davies, J.W. and Zhang, X. (2017), “A multi-scale convolutional neural network for phenotyping high-content cellular images”, *Bioinformatics*, Oxford University Press, Vol. 33 No. 13, pp. 2010–2019.

30. Gu, Y., Ruan, R., Yan, Y., Zhao, J., Sheng, W., Liang, L. and Huang, B. (2022), "Glomerulus Semantic Segmentation Using Ensemble of Deep Learning Models", *Arabian Journal for Science and Engineering*, Springer Science and Business Media Deutschland GmbH, available at:<https://doi.org/10.1007/s13369-022-06608-9>.
31. Guégan, D. and Hassani, B. (2018), "Regulatory learning: How to supervise machine learning models? An application to credit scoring", *Journal of Finance and Data Science*, Vol. 4 No. 3, available at:<https://doi.org/10.1016/j.jfds.2018.04.001>.
32. Haltia, A., Solin, M.L., Holmberg, C., Reivinen, J., Miettinen, A. and Holthöfer, H. (1998), "Morphologic changes suggesting abnormal renal differentiation in congenital nephrotic syndrome", *Pediatric Research*, Vol. 43 No. 3, available at:<https://doi.org/10.1203/00006450-199803000-00017>.
33. ben Hamida, A., Devanne, M., Weber, J., Truntzer, C., Derangère, V., Ghiringhelli, F., Forestier, G., et al. (2021), "Deep learning for colon cancer histopathological images analysis", *Computers in Biology and Medicine*, Vol. 136, available at:<https://doi.org/10.1016/j.compbiomed.2021.104730>.
34. Hanna, M.G., Parwani, A. and Sirintrapun, S.J. (2020), *Whole Slide Imaging: Technology and Applications*, available at: <http://journals.lww.com/anatomicpathology>.
35. He, K., Gkioxari, G., Dollár, P. and Girshick, R. (2020), "Mask R-CNN", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 42 No. 2, available at:<https://doi.org/10.1109/TPAMI.2018.2844175>.
36. He, K., Zhang, X., Ren, S. and Sun, J. (2015), "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 37 No. 9, available at:<https://doi.org/10.1109/TPAMI.2015.2389824>.
37. He, K., Zhang, X., Ren, S. and Sun, J. (2016), "Deep residual learning for image recognition", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2016-December, available at:<https://doi.org/10.1109/CVPR.2016.90>.
38. He, X., Xu, W., Yang, J., Mao, J., Chen, S. and Wang, Z. (2021), "Deep Convolutional Neural Network With a Multi-Scale Attention Feature Fusion Module for Segmentation of Multimodal Brain Tumor", *Frontiers in Neuroscience*, Vol. 15, available at:<https://doi.org/10.3389/fnins.2021.782968>.

39. Hoenig, M.P. and Zeidel, M.L. (2014), "Homeostasis, the milieu inte´rieur, and the wisdom of the nephron", *Clinical Journal of the American Society of Nephrology, American Society of Nephrology*, Vol. 9 No. 7, pp. 1272–1281.
40. Hou, L., Samaras, D., Kurc, T.M., Gao, Y., Davis, J.E. and Saltz, J.H. (2016), "Patch-Based Convolutional Neural Network for Whole Slide Tissue Image Classification", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2016-December, available at:<https://doi.org/10.1109/CVPR.2016.266>.
41. Huang, S., Nianguang, C.A.I., Penzuti Pacheco, P., Narandes, S., Wang, Y. and Wayne, X.U. (2018), "Applications of support vector machine (SVM) learning in cancer genomics", *Cancer Genomics and Proteomics*, available at:<https://doi.org/10.21873/cgp.20063>.
42. Huang, Z. and Cao, L. (2020), "Bicubic interpolation and extrapolation iteration method for high resolution digital holographic reconstruction", *Optics and Lasers in Engineering*, Vol. 130, available at:<https://doi.org/10.1016/j.optlaseng.2020.106090>.
43. Jialal., V.G.I. (2019), "StatPearls [Internet]", *StatPearls [Internet]*.
44. Jiang, H., Li, S., Liu, W., Zheng, H., Liu, J. and Zhang, Y. (2020), "Geometry-Aware Cell Detection with Deep Learning", *MSystems*, Vol. 5 No. 1, available at:<https://doi.org/10.1128/msystems.00840-19>.
45. Kaartinen, K., Safa, A., Kotha, S., Ratti, G. and Meri, S. (2019), "Complement dysregulation in glomerulonephritis", *Seminars in Immunology*, available at:<https://doi.org/10.1016/j.smim.2019.101331>.
46. Kakimoto, T., Okada, K., Hirohashi, Y., Relator, R., Kawai, M., Iguchi, T., Fujitaka, K., et al. (2014), "Automated image analysis of a glomerular injury marker desmin in spontaneously diabetic Torii rats treated with losartan", *Journal of Endocrinology*, Vol. 222 No. 1, available at:<https://doi.org/10.1530/JOE-14-0164>.
47. Kanjanabuch, T., Kittikowit, W. and Eiam-Ong, S. (2009), "An update on acute postinfectious glomerulonephritis worldwide", *Nature Reviews Nephrology*, available at:<https://doi.org/10.1038/nrneph.2009.44>.
48. Kannan, S., Morgan, L.A., Liang, B., Cheung, M.G., Lin, C.Q., Mun, D., Nader, R.G., et al. (2019), "Segmentation of Glomeruli Within Trichrome Images Using Deep Learning", *Kidney International Reports*, Vol. 4 No. 7, available at:<https://doi.org/10.1016/j.ekir.2019.04.008>.

49. Kato, T., Relator, R., Ngouv, H., Hirohashi, Y., Takaki, O., Kakimoto, T. and Okada, K. (2015), "Segmental HOG: New descriptor for glomerulus detection in kidney microscopy image", *BMC Bioinformatics*, Vol. 16 No. 1, available at:<https://doi.org/10.1186/s12859-015-0739-1>.
50. Kayalibay, B., Jensen, G. and van der Smagt, P. (2017), "CNN-based Segmentation of Medical Imaging Data", available at: <http://arxiv.org/abs/1701.03056>.
51. Keim, D., Andrienko, G., Fekete, J.D., Görg, C., Kohlhammer, J. and Melançon, G. (2008), "Visual analytics: Definition, process, and challenges", *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 4950 LNCS, available at:https://doi.org/10.1007/978-3-540-70956-5_7.
52. Khan, A., Sohail, A., Zahoora, U. and Qureshi, A.S. (2020a), "A survey of the recent architectures of deep convolutional neural networks", *Artificial Intelligence Review*, Springer Science and Business Media B.V., Vol. 53 No. 8, pp. 5455–5516.
53. Khan, A., Sohail, A., Zahoora, U. and Qureshi, A.S. (2020b), "A survey of the recent architectures of deep convolutional neural networks", *Artificial Intelligence Review*, Vol. 53 No. 8, available at:<https://doi.org/10.1007/s10462-020-09825-6>.
54. Khened, M., Kori, A., Rajkumar, H., Krishnamurthi, G. and Srinivasan, B. (2021), "A generalized deep learning framework for whole-slide image segmentation and analysis", *Scientific Reports, Nature Research*, Vol. 11 No. 1, available at:<https://doi.org/10.1038/s41598-021-90444-8>.
55. Komura, D. and Ishikawa, S. (2018), "Machine Learning Methods for Histopathological Image Analysis", *Computational and Structural Biotechnology Journal*, Elsevier B.V.
56. Koyner, J.L. (2012), "Assessment and diagnosis of renal dysfunction in the ICU", *Chest*, available at:<https://doi.org/10.1378/chest.11-1513>.
57. Krizhevsky, A., Sutskever, I. and Hinton, G.E. (n.d.). *ImageNet Classification with Deep Convolutional Neural Networks*, available at: <http://code.google.com/p/cuda-convnet/>.
58. Kumar, R., Arthanari, M. and Sivakumar, M. (2012), "Image Segmentation using Discontinuity-Based Approach", *International Journal of Multimedia and Image Processing*, Vol. 2 No. 1/2, available at:<https://doi.org/10.20533/ijmip.2042.4647.2012.0009>.
59. Lantuejoul, S. (2019), "MS17.03 Quantitative Image Analysis, Image-Based Profiling Including 3D Digital Printing and AI", *Journal of Thoracic Oncology*, Vol. 14 No. 10, available at:<https://doi.org/10.1016/j.jtho.2019.08.379>.

60. Lecun, Y., Bengio, Y. and Hinton, G. (2015), "Deep learning", *Nature*, available at:<https://doi.org/10.1038/nature14539>.
61. Lee, C.Y., Gallagher, P.W. and Tu, Z. (2016), "Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree", *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics, AISTATS 2016*.
62. Lee, D. and Myung, K. (2017), "ADAM: Method for Stochastic Optimization", *2017 IEEE International Conference on Consumer Electronics, ICCE 2017*.
63. Levey, A.S., de Jong, P.E., Coresh, J., Nahas, M. el, Astor, B.C., Matsushita, K., Gansevoort, R.T., et al. (2011), "The definition, classification, and prognosis of chronic kidney disease: A KDIGO Controversies Conference report", *Kidney International, Nature Publishing Group*, Vol. 80 No. 1, pp. 17–28.
64. Li, Y., Qi, F. and Wan, Y. (2019), "Improvements on Bicubic Image Interpolation", *Proceedings of 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference, IAEAC 2019*, available at:<https://doi.org/10.1109/IAEAC47372.2019.8997600>.
65. Liang, L., Cao, J., Li, X. and You, J. (2019), "Improvement of Residual Attention Network for Image Classification", *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 11935 LNCS, available at:https://doi.org/10.1007/978-3-030-36189-1_44.
66. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., et al. (2017), "A survey on deep learning in medical image analysis", *Medical Image Analysis, Elsevier B.V.*, 1 December.
67. Liu, Y., Gadepalli, K., Norouzi, M., Dahl, G.E., Kohlberger, T., Boyko, A., Venugopalan, S., et al. (2017), "Detecting Cancer Metastases on Gigapixel Pathology Images", available at: <http://arxiv.org/abs/1703.02442>.
68. Luong, M.T., Pham, H. and Manning, C.D. (2015), "Effective approaches to attention-based neural machine translation", *Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, available at:<https://doi.org/10.18653/v1/d15-1166>.
69. Lutnick, B., Ginley, B., Govind, D., McGarry, S.D., LaViolette, P.S., Yacoub, R., Jain, S., et al. (2019), "An integrated iterative annotation technique for easing neural network training in medical image analysis", *Nature Machine Intelligence, Nature Research*, Vol. 1 No. 2, pp. 112–119.

70. Ma, C., Mu, X. and Sha, D. (2019), "Multi-Layers Feature Fusion of Convolutional Neural Network for Scene Classification of Remote Sensing", IEEE Access, Vol. 7, available at:<https://doi.org/10.1109/ACCESS.2019.2936215>.
71. Macenko, M., Niethammer, M., Marron, J.S., Borland, D., Woosley, J.T., Guan, X., Schmitt, C., et al. (2009), "A method for normalizing histology slides for quantitative analysis", Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009, available at:<https://doi.org/10.1109/ISBI.2009.5193250>.
72. Magee, D., Quirke, P., Treanor, D., Crellin, D., Shires, M., Smith, K. and Mohee, K. (2009), Colour Normalisation in Digital Histopathology Images, available at: www.aperio.com.
73. Marini, N., Otálora, S., Podareanu, D., van Rijthoven, M., van der Laak, J., Ciompi, F., Müller, H., et al. (2021), "Multi_Scale_Tools: A Python Library to Exploit Multi-Scale Whole Slide Images", Frontiers in Computer Science, Frontiers Media S.A., Vol. 3, available at:<https://doi.org/10.3389/fcomp.2021.684521>.
74. McKinley, R., Wepfer, R., Gundersen, T., Wagner, F., Chan, A., Wiest, R. and Reyes, M. (2016), "Nabla-net: A deep dag-like convolutional architecture for biomedical image segmentation", Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 10154 LNCS, available at:https://doi.org/10.1007/978-3-319-55524-9_12.
75. Meng, H., Li, Q., Liu, X., Wang, Y. and Niu, J. (2021), "Multi-scale view-based convolutional neural network for breast cancer classification in ultrasound images", available at:<https://doi.org/10.1117/12.2581918>.
76. Milletari, F., Navab, N. and Ahmadi, S.A. (2016), "V-Net: Fully convolutional neural networks for volumetric medical image segmentation", Proceedings - 2016 4th International Conference on 3D Vision, 3DV 2016, available at:<https://doi.org/10.1109/3DV.2016.79>.
77. Moradi, M., Du, X., Huan, T. and Chen, Y. (2022), "Feasibility of the soft attention-based models for automatic segmentation of OCT kidney images", Biomedical Optics Express, Vol. 13 No. 5, available at:<https://doi.org/10.1364/boe.449942>.
78. Niazi, K.K., Gurcan, M.N., Khalid, M., Niazi, K., Parwani, A. v and Gurcan, M.N. (2019), Digital Oncology 1 Digital Pathology and Artificial Intelligence, Series Lancet Oncol, Vol. 20, available at: www.thelancet.com/oncology.

79. Nortier, J.L., Remiche, G., Delrée, P., Nauta, J., Notermans, N.C., Vivarelli, M., Diodato, D., et al. (2021), "Antenatal Membranous Nephropathy and Type 2 (Axonal) Charcot-Marie-Tooth With Mutations in the Metallo-Membrane Endopeptidase Gene: A Call for Family Screening and Pharmacovigilance", *Kidney International Reports*, Vol. 6 No. 7, available at:<https://doi.org/10.1016/j.ekir.2021.04.034>.
80. Oktay, O., Schlemper, J., Folgoc, L. le, Lee, M., Heinrich, M., Misawa, K., Mori, K., et al. (2018), "Attention U-Net: Learning Where to Look for the Pancreas", available at: <http://arxiv.org/abs/1804.03999>.
81. Palmer, L.G. and Schnermann, J. (2015), "Integrated control of na transport along the nephron", *Clinical Journal of the American Society of Nephrology*, American Society of Nephrology, Vol. 10 No. 4, pp. 676–687.
82. Pesce, F. and Schena, F.P. (2010), "Worldwide distribution of glomerular diseases: The role of renal biopsy registries", *Nephrology Dialysis Transplantation*, available at:<https://doi.org/10.1093/ndt/gfp620>.
83. Pluznick, J.L. and Caplan, M.J. (2015), "Chemical and physical sensors in the regulation of renal function", *Clinical Journal of the American Society of Nephrology*, American Society of Nephrology, Vol. 10 No. 9, pp. 1626–1635.
84. Powles, J. and Hodson, H. (2017), "Google DeepMind and healthcare in an age of algorithms", *Health and Technology*, Springer Verlag, Vol. 7 No. 4, pp. 351–367.
85. Qu, X., Yang, L., Guo, K., Ma, L., Sun, M., Ke, M. and Li, M. (2021), "A Survey on the Development of Self-Organizing Maps for Unsupervised Intrusion Detection", *Mobile Networks and Applications*, Vol. 26 No. 2, available at:<https://doi.org/10.1007/s11036-019-01353-0>.
86. Recht, B., Roelofs, R., Schmidt, L. and Shankar, V. (2019), "Do ImageNet classifiers generalize to ImageNet?", *36th International Conference on Machine Learning, ICML 2019*, Vol. 2019-June.
87. van Rijthoven, M., Balkenhol, M., Siliņa, K., van der Laak, J. and Ciompi, F. (2021), "HookNet: Multi-resolution convolutional neural networks for semantic segmentation in histopathology whole-slide images", *Medical Image Analysis*, Elsevier B.V., Vol. 68, available at:<https://doi.org/10.1016/j.media.2020.101890>.
88. Ronneberger, O., Fischer, P. and Brox, T. (2015), "U-Net: Convolutional Networks for Biomedical Image Segmentation", available at: <http://arxiv.org/abs/1505.04597>.

89. Roszkowiak, L., Korzynska, A., Zak, J., Pijanowska, D., Swiderska-Chadaj, Z. and Markiewicz, T. (2017), "Survey: interpolation methods for whole slide image processing", *Journal of Microscopy*, available at:<https://doi.org/10.1111/jmi.12477>.
90. Roy, M., Kong, J., Kashyap, S., Pastore, V.P., Wang, F., Wong, K.C.L. and Mukherjee, V. (2021a), "Convolutional autoencoder based model HistoCAE for segmentation of viable tumor regions in liver whole-slide images", *Scientific Reports, Nature Research*, Vol. 11 No. 1, available at:<https://doi.org/10.1038/s41598-020-80610-9>.
91. Roy, M., Kong, J., Kashyap, S., Pastore, V.P., Wang, F., Wong, K.C.L. and Mukherjee, V. (2021b), "Convolutional autoencoder based model HistoCAE for segmentation of viable tumor regions in liver whole-slide images", *Scientific Reports, Nature Research*, Vol. 11 No. 1, available at:<https://doi.org/10.1038/s41598-020-80610-9>.
92. Scherer, D., Müller, A. and Behnke, S. (2010), "Evaluation of pooling operations in convolutional architectures for object recognition", *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 6354 LNCS, available at:https://doi.org/10.1007/978-3-642-15825-4_10.
93. Sellaro, T., Filkins, R., Hoffman, C., Fine, J., Ho, J., Parwani, A., Pantanowitz, L., et al. (2013), "Relationship between magnification and resolution in digital pathology systems", *Journal of Pathology Informatics*, Vol. 4 No. 1, available at:<https://doi.org/10.4103/2153-3539.116866>.
94. Shorten, C. and Khoshgoftaar, T.M. (2019), "A survey on Image Data Augmentation for Deep Learning", *Journal of Big Data*, Vol. 6 No. 1, available at:<https://doi.org/10.1186/s40537-019-0197-0>.
95. Siméoni, O., Budnik, M., Avrithis, Y. and Gravier, G. (2020), "Rethinking deep active learning: Using unlabeled data at model training", *Proceedings - International Conference on Pattern Recognition*, available at:<https://doi.org/10.1109/ICPR48806.2021.9412716>.
96. Simonyan, K. and Zisserman, A. (2015), "Very deep convolutional networks for large-scale image recognition", *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*.
97. Subramanya, A.R. and Ellison, D.H. (2014), "Distal convoluted tubule", *Clinical Journal of the American Society of Nephrology, American Society of Nephrology*, Vol. 9 No. 12, pp. 2147–2163.

98. Sultan, A.S., Elgharib, M.A., Tavares, T., Jessri, M. and Basile, J.R. (2020), "The use of artificial intelligence, machine learning and deep learning in oncologic histopathology", *Journal of Oral Pathology and Medicine*, Blackwell Publishing Ltd, 1 October.
99. Swaminathan, S., Leung, N., Lager, D.J., Melton, L.J., Bergstralh, E.J., Rohlinger, A. and Fervenza, F.C. (2006), "Changing incidence of glomerular disease in Olmsted County, Minnesota: a 30-year renal biopsy study.", *Clinical Journal of the American Society of Nephrology : CJASN*, Vol. 1 No. 3, available at:<https://doi.org/10.2215/CJN.00710805>.
100. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., et al. (2015), "Going deeper with convolutions", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 07-12-June-2015, available at:<https://doi.org/10.1109/CVPR.2015.7298594>.
101. Taleb Zadeh Kasgari, A., Saad, W. and Debbah, M. (2019), "Human-in-the-Loop Wireless Communications: Machine Learning and Brain-Aware Resource Management", *IEEE Transactions on Communications*, Vol. 67 No. 11, available at:<https://doi.org/10.1109/TCOMM.2019.2930275>.
102. Tellez, D., Litjens, G., Bándi, P., Bulten, W., Bokhorst, J.M., Ciompi, F. and van der Laak, J. (2019), "Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology", *Medical Image Analysis*, Vol. 58, available at:<https://doi.org/10.1016/j.media.2019.101544>.
103. Temerinac-Ott, M., Forestier, G., Schmitz, J., Hermsen, M., Braseni, J.H., Feuerhake, F. and Wemmert, C. (2017), "Detection of glomeruli in renal pathology by mutual comparison of multiple staining modalities", *International Symposium on Image and Signal Processing and Analysis, ISPA*, available at:<https://doi.org/10.1109/ISPA.2017.8073562>.
104. Türk, F., Lüy, M. and Barışçı, N. (2020), "Kidney and renal tumor segmentation using a hybrid v-net-based model", *Mathematics*, Vol. 8 No. 10, available at:<https://doi.org/10.3390/math8101772>.
105. Vincerževskienė, I., Krilavičiūtė, A. and Smailytė, G. (2014), "Trends in cancer incidence in Lithuania between 1991 and 2010", *Acta Medica Lituanica*, Vol. 20 No. 4, available at:<https://doi.org/10.6001/actamedica.v20i4.2811>.
106. Wang, F., Jiang, M., Qian, C., Yang, S., Li, C., Zhang, H., Wang, X., et al. (2017), "Residual attention network for image classification", *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Vol. 2017-January, available at:<https://doi.org/10.1109/CVPR.2017.683>.

107. Wang, K. and Kestenbaum, B. (2018), "Proximal tubular secretory clearance: a neglected partner of kidney function", *Clinical Journal of the American Society of Nephrology*, available at:<https://doi.org/10.2215/CJN.12001017>.
108. Wang, P., Peng, D., Li, L., Chen, L., Wu, C., Wang, X., Childs, P., et al. (2019), "Human-in-the-loop design with machine learning", *Proceedings of the International Conference on Engineering Design, ICED*, Vol. 2019-August, available at:<https://doi.org/10.1017/dsi.2019.264>.
109. Wetteland, R., Engan, K., Eftestøl, T., Kvikstad, V. and Janssen, E.A.M. (2020), "A Multiscale Approach for Whole-Slide Image Segmentation of five Tissue Classes in Urothelial Carcinoma Slides", *Technology in Cancer Research & Treatment*, SAGE Publications, Vol. 19, p. 153303382094678.
110. Wogalter, M.S. (2019), "Communication-Human Information Processing (C-HIP) Model", *Forensic Human Factors and Ergonomics*, available at:<https://doi.org/10.1201/9780429462269-3>.
111. Wong, P.F., Wei, W., Smithy, J.W., Acs, B., Toki, M.I., Blenman, K.R.M., Zelterman, D., et al. (2019), "Multiplex quantitative analysis of tumor-infiltrating lymphocytes and immunotherapy outcome in metastatic melanoma", *Clinical Cancer Research*, Vol. 25 No. 8, available at:<https://doi.org/10.1158/1078-0432.CCR-18-2652>.
112. Woo, S., Park, J., Lee, J.Y. and Kweon, I.S. (2018), "CBAM: Convolutional block attention module", *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 11211 LNCS, available at:https://doi.org/10.1007/978-3-030-01234-2_1.
113. Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T. and He, L. (2021), "A Survey of Human-in-the-loop for Machine Learning", available at: <http://arxiv.org/abs/2108.00941>.
114. Wulczyn, E., Steiner, D.F., Moran, M., Plass, M., Reihls, R., Tan, F., Flament-Auvigne, I., et al. (2021), "Interpretable survival prediction for colorectal cancer using deep learning", *Npj Digital Medicine*, Nature Research, Vol. 4 No. 1, available at:<https://doi.org/10.1038/s41746-021-00427-2>.
115. Yamashita, R., Nishio, M., Do, R.K.G. and Togashi, K. (2018), "Convolutional neural networks: an overview and application in radiology", *Insights into Imaging*, available at:<https://doi.org/10.1007/s13244-018-0639-9>.

116. Yeruva, V.K., Chandrashekar, M., Lee, Y., Rydberg-Cox, J., Blanton, V. and Oyler, N.A. (2020), "Interpretation of Sentiment Analysis with Human-in-the-Loop", Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020, available at:<https://doi.org/10.1109/BigData50022.2020.9378221>.
117. Zagoruyko, S. and Komodakis, N. (2016), "Wide Residual Networks", British Machine Vision Conference 2016, BMVC 2016, Vol. 2016-September, available at:<https://doi.org/10.5244/C.30.87>.
118. Zhang, L. and Ji, Q. (2011), "A Bayesian Network Model for Automatic and Interactive Image Segmentation", IEEE Transactions on Image Processing, Institute of Electrical and Electronics Engineers (IEEE), Vol. 20 No. 9, pp. 2582–2593.
119. Zhang, Y., Kers, J., Cassol, C.A., Roelofs, J.J., Idrees, N., Farber, A., Haroon, S., et al. (n.d.). U-Net-and-a-Half: Convolutional Network for Biomedical Image Segmentation Using Multiple Expert-Driven Annotations, available at: <https://orcid.org/0000-0002-5312-8644>.
120. Zhou, B., Lapedriza, A., Xiao, J., Torralba, A. and Oliva, A. (2014), "Learning deep features for scene recognition using places database", Advances in Neural Information Processing Systems, Vol. 1.
121. Znaor, A., Lortet-Tieulent, J., Laversanne, M., Jemal, A. and Bray, F. (2015), "Reply from Authors re: Mehrad Adibi, Jose A. Karam, Christopher G. Wood. Reporting geographic and temporal trends in renal cell carcinoma: Why is this important? Eur Urol 2015;67:531-2: Tackling inequalities in renal cell carcinoma", European Urology, available at:<https://doi.org/10.1016/j.eururo.2014.11.008>.
122. Ziyarah, A.M. and Kudithipudi, D. (2019), "Neuromorphic Architecture for the Hierarchical Temporal Memory", IEEE Transactions on Emerging Topics in Computational Intelligence, Vol. 3 No. 1, available at:<https://doi.org/10.1109/TETCI.2018.2850314>.

SUMMARY

Institution: Vilnius university, **Study program:** Systems Biology, **Name, surname:** Oluwaseun Ezekiel Taiwo.

One of the major approaches engaged by kidney pathologist to knowing the morphological states of the kidney glomeruli in any glomeruli disease diagnosis is image analysis using image segmentation techniques. There were attempts using various deep convolutional neural network workflow to automate glomeruli segmentation in kidney images but not still achieve the most accurate results. In this current research, a new multiresolution deep convolutional neural network workflow is developed. It applies an iterative annotation strategy which involves the customization processing units of deep convolutional neural network architecture. It optimizes kidney glomeruli semantic segmentation accuracy in whole slide images. The result from this study showed that a multiresolution attention residual UNET model achieves the best segmentation accuracy of 89.89% with ensemble models.

SUMMARY IN LITHUANIAN

Iestūde: Viļņas universitāte, studiju programma: Sistēmas bioloģija, vārds, uzvārds: Oluwaseun Ežekīls Taiwo.

Viena no galvenajām pieejām, ko nieru patoloģi izmanto, lai noskaidrotu nieru glomeruli morfoloģiskos stāvokļus jebkuras glomeruli slimības diagnozes gadījumā, ir attēlu analīze, izmantojot attēlu segmentācijas metodes. Tika veikti mēģinājumi, izmantojot dažādu konvolūciju nervu tīkla darbplūsmu, lai automatizētu glomeruli segmentāciju nieru attēlos, bet joprojām nerasniegtu visprecīzākos rezultātus. Šajā pašreizējā izpētē tiek izstrādāta jauna multiresolution dziļo konvolucionālo nervu tīkla darbplūsma. Tas izmanto iteratīvu anotāciju stratēģiju, kas ietver dziļās konvolucionālās nervu tīkla arhitektūras pielāgošanas vienības. Tas optimizē nieru glomeruli semantiskās segmentācijas precizitāti veselos slaidos attēlos. Šā pētījuma rezultāti parādīja, ka multiresolution uzmanības atlikušais UNET modelis sasniedz vislabāko segmentācijas precizitāti 89,89% kopā ar komplektiem modeļiem.

APPENDICES

Appendix A

Scripts written for this research work and figures used in this manuscript are available on GitHub "<https://github.com/Ezekiel2020/Thesis.git>".

Appendix B

ZENODO dataset "<https://zenodo.org/record/4299694#.Ym02dtpBy3D>"