

Deep Learning Techniques for Quantification of Tumour Necrosis in Post-neoadjuvant Chemotherapy Osteosarcoma Resection Specimens for Effective Treatment Planning

T. S. Saleena ¹ , P. Muhamed Ilyas ¹, V. M. Kutty Sajna ², A. K. M. Bahalul Haque ³ 

¹ Sullamussalam Science College, Areekode, India

² MVR Cancer Centre and Research Institute, Calicut, India

³ School of Engineering Science, LUT University, Finland

Corresponding author: T. S. Saleena (tssaleena@gmail.com)

Abstract

Osteosarcoma is a high-grade malignant bone tumour for which neoadjuvant chemotherapy is a vital component of the treatment plan. Chemotherapy brings about the death of tumour tissues, and the rate of their death is an essential factor in deciding on further treatment. The necrosis quantification is now done manually by visualizing tissue sections through the microscope. This is a crude method that can cause significant inter-observer bias. The suggested system is an AI-based therapeutic decision-making tool that can automatically calculate the quantity of such dead tissue present in a tissue specimen. We employ U-Net++ and DeepLabv3+, pre-trained deep learning algorithms for the segmentation purpose. ResNet50 and ResNet101 are used as encoder parts of U-Net++ and DeepLabv3+, respectively. Also, we synthesize a dataset of 555 patches from 37 images captured and manually annotated by experienced pathologists. Dice loss and Intersection over Union (IoU) are used as the performance metrics. The training and testing IoU of U-Net++ are 91.78% and 82.64%, and its loss is 4.4% and 17.77%, respectively. The IoU and loss of DeepLabv3+ are 91.09%, 81.50%, 4.77%, and 17.8%, respectively. The results show that both models perform almost similarly. With the help of this tool, necrosis segmentation can be done more accurately while requiring less work and time. The percentage of segmented regions can be used as the decision-making factor in the further treatment plans.

Keywords

Necrosis; U-Net++; DeepLabv3+; ResNet; Osteosarcoma; Tissue volume calculation.

Citation: Saleena, T. S., Ilyas, P. M., Sajna, V. M. K., & Haque, A. K. M. B. (2023). Deep Learning Techniques for Quantification of Tumour Necrosis in Post-neoadjuvant Chemotherapy Osteosarcoma Resection Specimens for Effective Treatment Planning. *Acta Informatica Pragensia*, 12(1), 87–103. <https://doi.org/10.18267/j.aip.207>

Special Issue Editors: Mazin Abed Mohammed, University of Anbar, Iraq
Seifedine Kadry, Noroff University College, Norway
Oana Geman, Ștefan cel Mare University of Suceava, Romania

Academic Editor: Zdenek Smutny, Prague University of Economics and Business, Czech Republic

Copyright: © 2023 by the author(s). Licensee Prague University of Economics and Business, Czech Republic.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (CC BY 4.0).

1 Introduction

Ferlay et al. (2015) made a survey of the mortality caused by 27 major cancers based on evidence given in GLOBOCAN 2012, a data series of the International Agency for Research on Cancer. The study showed that more than 30,000 new cases are diagnosed per year. All cancer patients should not be treated the same. The treatment plan should be based on the patient's body condition, their body's response to the therapeutics, and multiple prognostic factors. Prognostic scores are the key elements in the selection of treatment plans in clinical practice (ESMO, 2014; Picci et al., 1994; Davis et al., 1994). In the case of osteosarcoma, chemotherapy-induced tumour necrosis is one of the most important prognostic factors.

Necrosis is a kind of cell death. Among the several types of cell necrosis scenarios, we focus only on the tumour cell necrosis that arises due to neoadjuvant treatment (NAT). Such treatment can be radiation therapy, chemotherapy, or hormone therapy, which is an initial step to diminish the cancerous region before the main treatment. In our study, we deal with the case of neoadjuvant chemotherapy or NACT. Its primary purpose is to assess the clinical response of the patient's body to such therapeutics. It can reduce the volume of tumours, reduce risk stratification, and also improves the success rate of surgical resection (Caparica et al., 2019; Wang et al., 2011). The advantages and disadvantages of NAT in advanced-stage breast cancer patients are described by Ikeda et al. (2002). Initially, neoadjuvant chemotherapy (NAC) was introduced to reduce the volume of the cancerous region, where the cancer is almost in an unresectable stage. Nowadays, however, it is proclaimed by experts that NAC followed by de-bulking surgery, is the best treatment for advanced cancer (Elies et al., 2018). Several studies are reported to evaluate the clinical and pathological response after NAC on various cancers (Alawad et al., 2014; Burcombe et al., 2002). Nevertheless, all these are done either via laboratory tests or medical imaging modalities such as PET.

Artificial intelligence has introduced tremendous changes in the healthcare field, especially in oncology. The journey of AI from the beginning to the current scenarios and future perspectives has been clearly explained by Luchini et al. (2022). It touches every stage of cancer treatment, including initial screening, diagnosis, detection, classification, segmentation, bio-marker detection, strategies for follow-up, and up to the drug discovery (Luchini et al., 2022; Kann et al., 2019). In the past few years, digital pathology has received a lot of attention in the above-mentioned stages. The application of AI in digital pathology helps the entire medical team reduce their workload. The proposed system can act as an automated tool to guide *clinical decision-making*, using digital histopathology images.

1.1 Literature review

Image segmentation is a process of grouping voxels that belong to the same tissue structures according to their morphology (Estienne et al., 2019). The traditional clustering and methods are outperformed by the introduction of neural networks, which enable the model to predict a particular region of interest from a given image. Zhang et al. (2017) pioneered segmentation using U-Net and ResNet by highlighting the concept of pyramid-dilated convolution. The pooling layer is replaced here with a convolution layer, which can reduce information loss. The study proved that the deeper the network, the better the performance. At the same time, however, the resources required to do this task are highly expensive. U-ResNet is a joint framework for image registration and segmentation of 3D medical volumes (Zhang et al., 2017). Deep convolutional encoder-decoder segmentation architecture was introduced in 2016, which proved its excellence in scene understanding applications (Badrinarayanan & Cipolla, 2017). Its encoder part is similar to the VGG-16 model and the decoder part is symmetrical to the encoder side and performs the non-linear up-sampling process. The combination of DeepLabv3+ and ResNet50 has been used in the semantic segmentation of COVID-19 lung-CT images to predict and quantify the severity of affected patients (Sabeerali et al., 2022). This network segments the affected region of the lungs from the image and

the volume is calculated separately. There are different networks, especially R-CNN (Girshick et al., 2014) and their successors that enable the model to localize the object and later perform segmentation.

Segmentation does not require recognizing or identifying the object; it only divides an image into different regions based on the requirement. In the olden days, clustering-based methods were used for this purpose by focusing on contours and edges (Ruspini et al., 2019; Shukla & Naganna, 2014; Liu, 2009). They use much fewer data and thereby, less computation power is needed for the execution of their algorithms. However, the introduction of different types of neural networks outperforms them (Guo, 2018). The segmentation can be divided into semantic segmentation and instance segmentation. The former is the pixel-wise grouping of objects into corresponding classes and the latter is instance-wise grouping. Multiple instances of the same class are considered as a single object of the same colour in the former and as different in the latter method.

The research in pathology image analysis using machine learning and deep learning is growing quickly, according to the Komura and Ishikawa (2017). Their study discusses many machine-learning approaches for analysing histopathology images. They mainly concentrate on whole slide imaging (WSI) data for their study.

Fibrosis is a kind of tissue cell which are an indicator of the extent of several diseases, but there was no tool to quantify this until the FIBER-ML was developed by Facchin et al. (2022). FIBER-ML is an open-source, semi-automated, machine learning based software that can be used to quantify the fibrosis in tissue sections.

A novel deep-learning approach for predicting the outcome of colorectal patients without intermediate tissue classification was introduced by Bychkov et al. (2018). They used tumour tissue microarray samples stained with haematoxylin and eosin (H&E) from 420 patients suffering from colorectal cancer. A pipeline of pre-trained convolutional (VGG16) and recurrent (LSTM) neural networks were used in the methodology part. This state-of-the-art method predicts the patient risk score (low and high risk), which shows expert-level accuracy.

A semi-automated ML-based system has been introduced to quantify vascular density in tissue-engineered constructs (Strobel et al., 2021). This tool's fast and accurate measuring capacity makes it perfect for incorporation into tissue manufacturing workflows.

Examining scar tissue is required during or after the wound-healing process to analyse the pathological tissue conditions. An automated method using Mask RCNN has been developed (Maknuna et al., 2022) to do this task with H&E-stained scar tissue images. Another unsupervised method using K-means clustering is also used in the study to distinguish different structural features from the specimen. A machine learning approach using support vector machine followed by random forest has been used to classify prostate tissues into subtypes (Gertych et al., 2015).

1.2 Problem statement

Although the amount of tumour necrosis is a crucial factor in deciding on the further treatment plan, in most hospitals, the estimation is done manually by pathologists. Identifying different tissues with the naked eye is cumbersome in such cases. This may cause several pitfalls as mentioned below:

- inter-observer variability,
- inter-class similarity and variability,
- labour and time intensity and
- colour variations in staining.

Inter-observer variability or disagreement means the difference in results obtained by various pathologists on the same input. This may occur due to the pathologists' subjectivity, experience, workload or manual error. The morphology of tissues may not be exactly the same every time and it may

have an overlapping nature sometimes. In most cases, the necrosis and fibrosis may be scattered throughout the specimen. The different classes of tissues will have some similarities in colour, shape, etc. Thus, the classification may be difficult in such cases. We consider H&E-stained specimens in our study. Normally, haematoxylin gives a purplish-blue colour to the cell nuclei whereas eosin gives pink staining to the cytoplasm and extracellular parts (Chan et al., 2014). The rest of the tissue will appear in a combination of both these. However, based on the pigment used, the colour may vary and thus the colour itself is not enough to identify the cells.

In the research aspect, there is no dataset available for doing any kind of research in tumour necrosis quantification or segmentation. The identification of proper tissue specimens and corresponding labelling is a tedious task. This should be done under the supervision of experienced pathologists.

1.3 Proposed system

We introduce an AI-based automated system that can evaluate the response of a patient's body towards NAC using digital histopathology images. We considering osteosarcoma only, where the cell death due to NAC is treated as the key prognostic factor. This model can segment the area of necrotic tissues from a histopathology image acquired as the model input. The segmented area is then quantified with image processing methods, and this value can be used as a deciding factor for further treatment plans for the patient. The output from this tool can be used as a second opinion in the further treatment plan.

A dataset of 555 images and their corresponding masks have been created for this work with the help of a senior pathologist. All the images are not in the same texture as the staining will vary for different specimens. Some of the images from the dataset are shown in Figure 1.

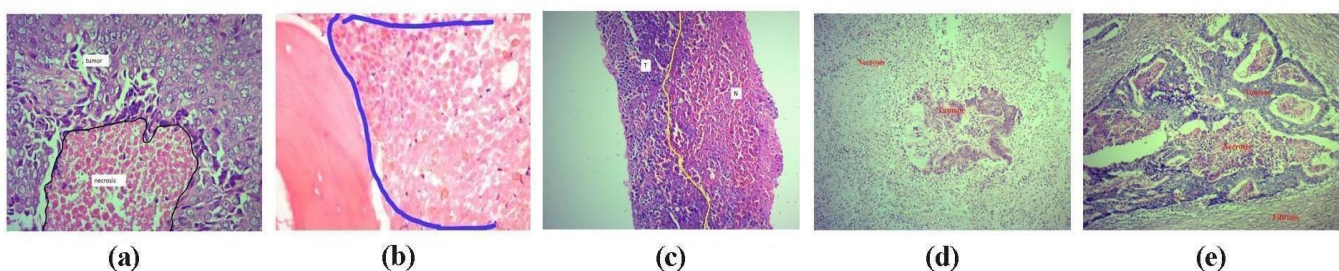


Figure 1. Different types of images collected: (a) pink colour shows the necrosis and the rest of the part is tumorous; (b) the granular part is the necrosis and the rest is bone tissues; (c) N denotes necrosis and T denotes tumour; (d) contains necrosis and tumour; (e) contains necrosis, tumour and fibrosis.

Two neural network architectures, namely U-Net++ and DeepLabv3+ have been used for training our model with the dataset created. Both these architectures have encoder and decoder parts. The encoder part can be replaced with pre-trained classification networks such as ResNet, VGG-16, AlexNet, etc., for enhanced feature extraction (Zhang et al., 2020). In this paper, we use ResNet-50 for U-Net++ and ResNet-101 for DeepLabv3+. We use GPU for training this model, as it consumes very high computation power.

The main contributions of our works are as follows:

- We develop an AI-based tool to grade osteosarcoma patients, which can be used as a second opinion for pathologists and doctors to make critical decisions on treatment plans.
- This software can be used to quantify how much cancerous tissue is dead (tissue necrosis) due to neoadjuvant chemotherapy using pathology images.
- This value of measurement is used to identify the response of the patient's body to the therapeutics.
- The area of necrotic tissues is separately highlighted from a given microscopic image.
- This tool is less error-prone as it reduces the inter-observer variability to a large extent.

- We have synthesized a dataset of 555 images that deploys the various morphologies of necrosis, tumours, and fibrosis.

This paper is organized as follows: Section 1 describes mainly the domain part, the problem of the existing system, the purpose and the contributions of the proposed system. Section 2 deals with the research methods, containing a detailed technical explanation of architectures used in the study, dataset preparation and augmentation, and describes how the models have been trained and how the parameters have been chosen to make the perfect model. The algorithm used in the study is also included in this section. Section 3 includes the analysis of the results obtained from the study along with graphs and tables, whereas Section 4 is the discussion part, where we present an evaluation of the model and the practical use of the system, its limitations, and future work-based on this study. The final section concludes the study.

2 Research Method and Preparation

2.1 Data preparation

Specimens required for the study are taken from cancer patients who have undergone NAC. They are then stained with H&E and digitized using a digital camera mounted on a Meiji microscope. The specimens include different types of tissues such as necrosis, tumour, bone, fibrosis and mucous membrane, as well as the glass background. To obtain better information and reduce the image resolution, representative sections are taken from each specimen. A low-power 10x magnification is used to focus on the maximum region of necrotic tissues within the specimen with maximum clarity with a resolution of 2592 x 1932. As these dimensions seem to be difficult to afford by the neural network, we cut these images into patches of 512x512 using the python library “patchify”. We collected 37 images and created a dataset with 555 patches from them. Manual annotations were performed on the images using a free, open-source image annotation tool called “CVAT”. Full images (37 images) were used for the annotation rather than the patches, as the cell morphology is difficult to detect in small patches. Then they were divided into patches in the same manner as the images.

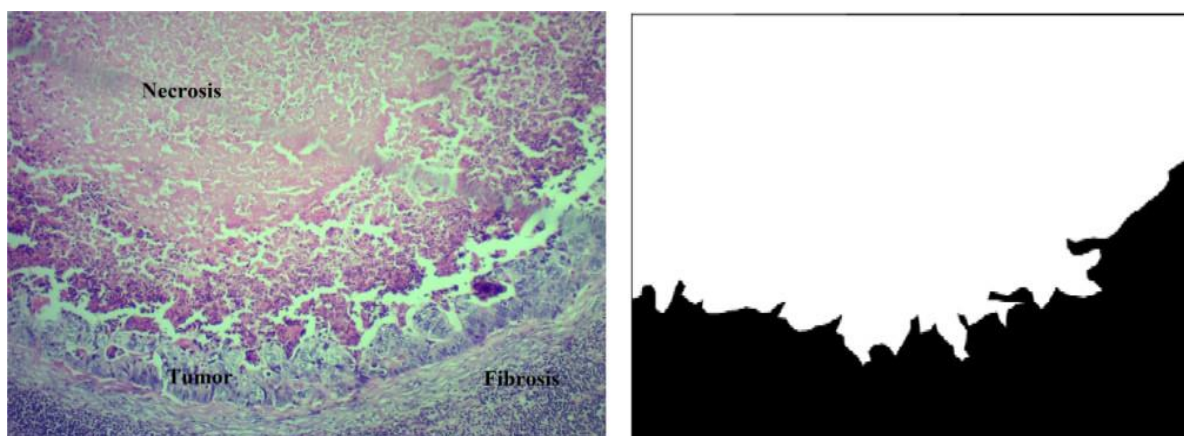


Figure 2. Image (left) and its mask obtained from CVAT (right).

2.2 Data augmentation

The microscopy images need elastic deformations as the number of images is very low compared to other data analytic projects. To do that, we used the data augmentation library “Albumentation”. The semantic segmentation is highly reliant on the pixels of the image and so the augmentation process should not change the pixel values. This library facilitates such techniques that increase the number of images without affecting their pixel values. More than seventy methods are available, but we used horizontal flip, shift scale rotate, CLAHE, random crop, pad if needed, hue saturation, random contrast, and motion blur.

Every change that is applied to an image is done on its mask too. Only the training data set requires the augmentation process.

2.3 Methods used

In this study, we employ the semantic segmentation strategy. We use two pre-trained networks here. Pre-trained networks work with the principle of *transfer learning technique*. As the name suggests, it transfers the knowledge in terms of weights from a previously learned task to solve a different but related problem. These models benefit from using very few data for training compared with other neural networks. Out of the several networks, we chose the two best algorithms that suit our dataset, which are U-Net++/ResNet50 and DeepLabv3+/ResNet101 neural networks. U-Net and DeepLab are convolutional neural networks, providing an encoder-decoder architecture which is the basement for semantic segmentation. Along with that, pre-trained classification models such as ResNet, VGG16, EfficientNet, MobileNet, etc., can be used as the backbone network for feature extraction purposes. Even though they are trained for classification tasks, they are re-purposed for the segmentation process. All these networks are trained on the ImageNet dataset (Deng et al., 2009), which contains 14,197,122 annotated images as per the findings of the WordNet hierarchy (ImageNet, 2021). In this study, we selected ResNet50 and ResNet101 for the same. The fully connected layers or dense layers of these networks are transformed into convolution layers (Chen et al., 2018a).

2.3.1 U-Net++

U-Net has been proven as one of the finest segmentation models ever developed for medical image segmentation. It won the 2015 ISBI cell tracking challenge on a dataset of light microscopy images (Ronneberger et al., 2015). Its architecture has an encoder part followed by a decoder part, where the former gradually reduces the dimension of the input image to the minimum and then the latter does the exact reverse process to get the original image (Ronneberger et al., 2015; Chandhok, 2021). The name of the architecture came from its shape, where the expanding path and the contracting path are more or less symmetric and form a “U” shape. The difference between the U-Net network and an auto-encoder network is the skip connection present in the former, which enables the flow of information from layers of the encoder part to corresponding layers of the decoder part, which causes good predictions. Without the skip connection, it is only a neural network.

Zhou et al. (2018) introduced U-Net++, which is a novel architecture for bio-medical segmentation. Its enhanced features from U-Net are:

- the dense convolution layers on the skip pathway,
- dense skip pathway, and
- deep supervision.

The convolution layers in the skip pathway reduce the semantic gap between the feature maps of encoder and decoder networks, especially in the segmentation of medical images, which provide an easier optimization problem. Here, the feature maps from the encoder part are fed into a dense convolution block rather than received directly at the decoder side as in U-Net. Lee et al. (2015) took the deep supervision approach from the idea of deeply-supervised nets or DenseNet, which allows the model to operate in normal mode and faster mode. In the normal mode, the outputs from all segmentation branches are averaged, whereas in the faster mode, the final segmentation maps are selected from any one of the segmentation branches. The architecture of this network with the above-mentioned features is shown in Figure 2 (Zhou et al., 2018).

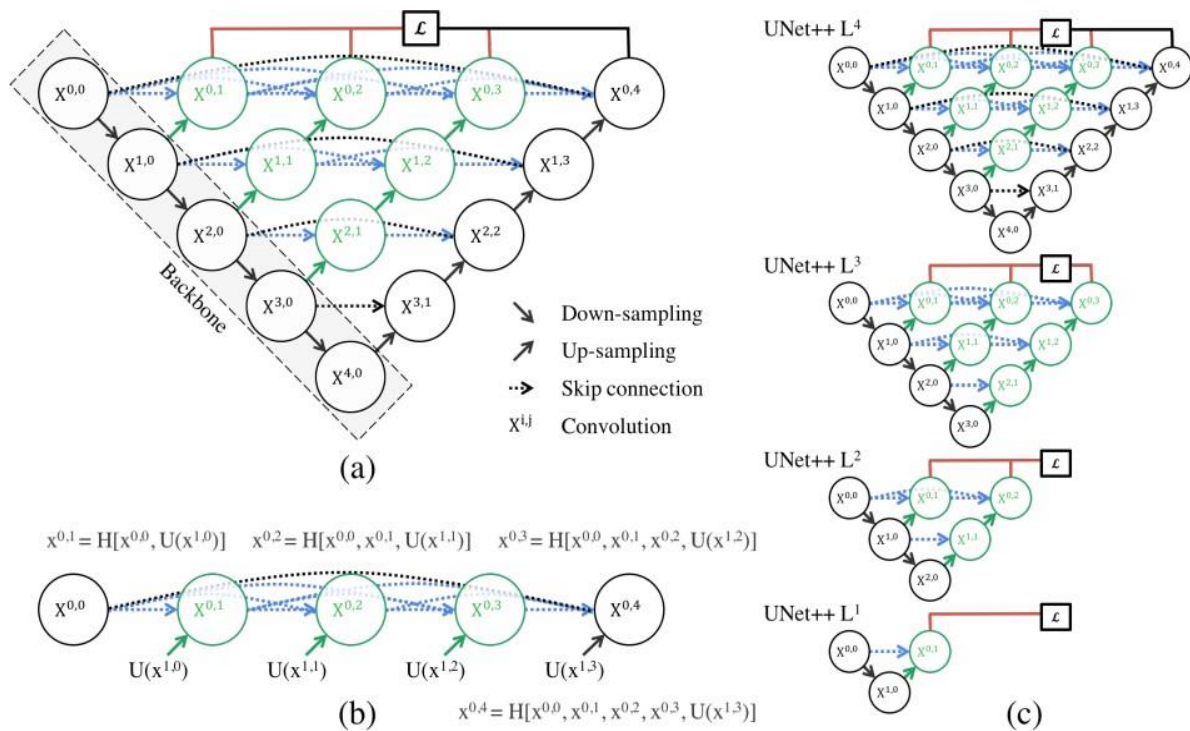


Figure 3. Architecture of U-Net++ where (a) explains the dense convolution layers in the skip connection, (b) shows the dense skip connection (c) shows the deep supervision. Black represents the normal U-Net, green and blue are for showing the dense convolution layer block and red is for the deep supervision. Source: (Zhou et al., 2018).

2.3.2 DeepLabv3+

The DeepLab series are state-of-the-art semantic segmentation models invented by Google. Within the series DeepLabv3 has proven its excellence as an image segmentation benchmark in PASCAL VOC 2012 with an accuracy of 85.7% without DenseCRF (Chen et al., 2017; 2018a; 2018b). DenseCRF is a post-processing method that is used to refine blurred boundaries in images where we can add a fully connected CRF (conditional random fields) layer and thereby increasing the efficiency of the algorithm (Yan et al., 2016). Multi-scale segmentation is made possible in this architecture by using atrous convolution in parallel. Atrous spatial pyramid pooling (ASPP), which is proposed in DeepLabv2, is also supported here. Atrous convolution, or dilated convolution, allows the networks to extract dense feature maps, thereby removing the down-sampling and up-sampling overheads. This enables the network to improve feature maps without increasing the number of parameters and eventually the computation power. Figure 3 depicts how the atrous rate increases the field of view of the image (Chen et al., 2017).

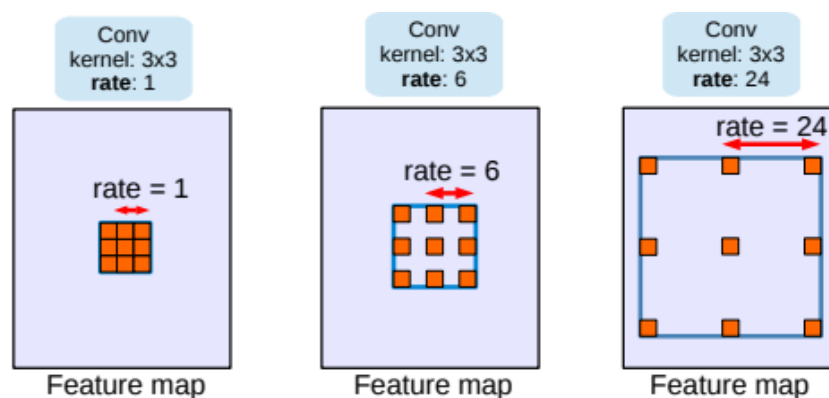


Figure 4. Atrous convolution with kernel size 3x3 and different rates. Source: (Chen et al., 2018b).

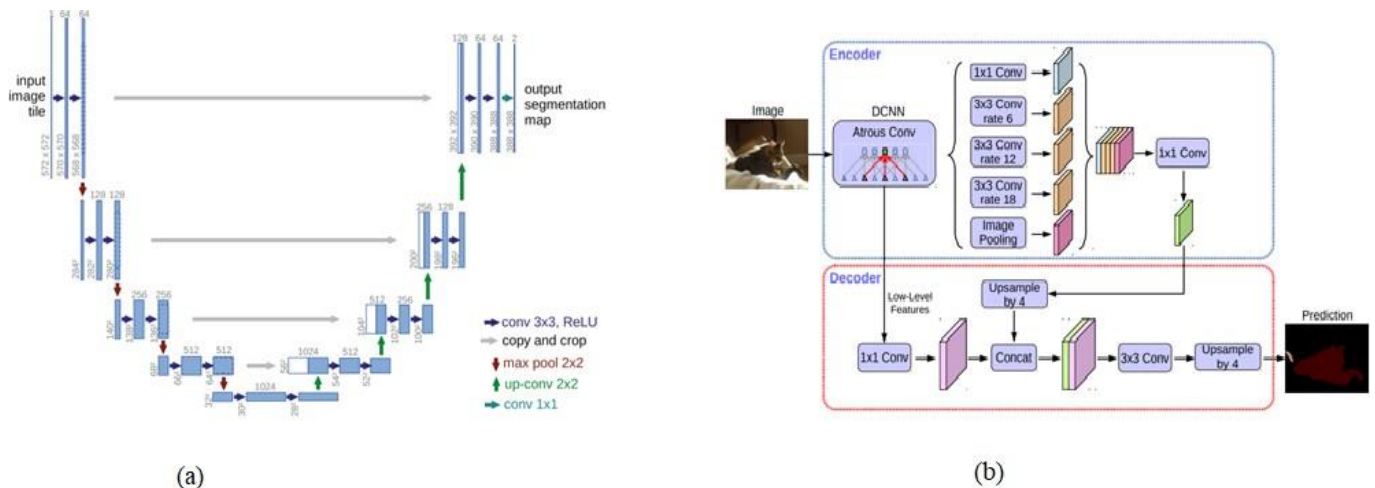


Figure 5. (a) U-Net architecture containing an encoder part on the left and a decoder part on the right with a skip connection as arrow marks in between each layer. (b) Architecture for DeepLabv3+ with Atrous convolution layers. Sources: (Zhou et al., 2018; Chen et al., 2018a).

2.3.3 ResNet50 and ResNet101

In CNN, normally there is a convention so that the model can be “the deeper, the better”. However, in some cases, even though the depth of the network can increase the model’s performance, at some point it is found that the training error increases with increasing the number of layers. This phenomenon is known as the vanishing or exploding gradient. This kind of degradation can be solved using a deep residual learning framework using the skip connection technique (He et al., 2016a). The skip connection in each residual block is explained in Figure 5.

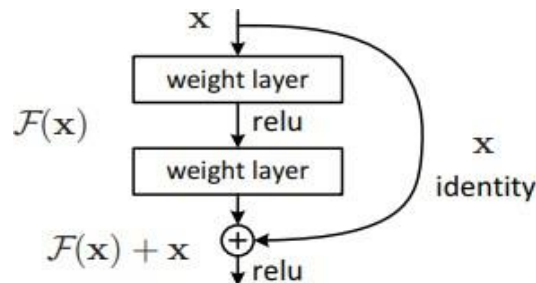


Figure 6. Short-cut connection or skip connection in a residual block, where x is the input for the layer and $F(x)+x$ is the short-cut connection, which can skip one or more layers. Source: (He et al., 2016a).

Each of these units can be denoted as follows (He et al., 2016b):

$$Y_l = h(x_l) + F(x_l, W_l) \quad (1)$$

$$x_{(l+1)} = f(Y_l)$$

where x_l and x_{l+1} are the input and output of the l^{th} unit respectively, F is a residual function, W is the set of weights and biases related to the l^{th} unit, $h(x_l) = x_l$ is an identity mapping and f is a ReLU activation function.

It allows the model to skip some layers while training and directly connects to the output. This feature was introduced by Microsoft Research in 2015 and named ResNet. Residual Net or ResNet is one of the most powerful pre-trained classification models available nowadays and was the winner of the ImageNet Large Scale Visual Recognition Challenge 2015 (ILSVRC2015). This network is trained on the ImageNet dataset and we can use its pre-trained weights in our model. So, we do not have to repeat the training section. Thus, all the inner layers can be kept frozen. Each layer of ResNet is composed of several blocks.

The loss surface will contain several irregular ups and downs if we do not use the skip connection whereas a smooth surface will be obtained by using it (Li et al., 2018).

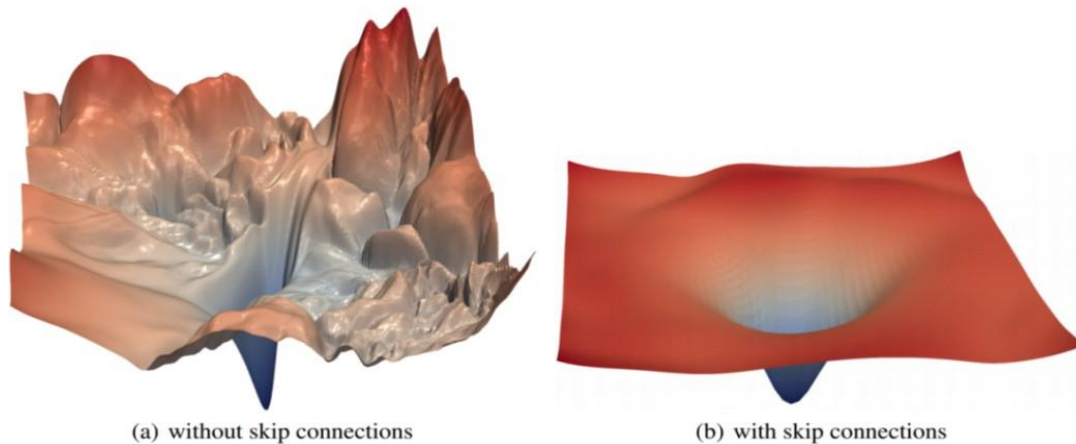


Figure 7. (a) shows the irregularities in the loss surface, and (b) is the smooth surface obtained by using the skip connection. Source: (Li et al., 2018).

In our study, we use ResNet50 along with U-Net++ and ResNet101 with DeepLabv3+, as the backbone for feature extraction. Among the several variants of ResNet, ResNet50 is one of the most commonly used ones that have 48 convolution layers, one max-pooling layer, and one average pooling layer. Meanwhile, ResNet101 is 101 layers deep as the name suggests. ResNet50 provides 25.6 million trainable parameters, whereas ResNet101 has 44.5 million (Boulch et al., 2017). ResNet can act as the down-sampling or encoder part of the U-Net. The residual functionality of ResNet makes it a good choice for feature extraction.

2.4 Solution

In this work, we trained a model that can perform binary segmentation, where the classes can be either necrosis or the background. The input images and their corresponding masks (ground truth) were used to train the model. The model training was done using the Torch library, which provides easy methods to switch between CPU and GPU. The training part was done using GPU, Quadro RTX 5000 with 16 GB RAM and the server was Intel 5120C with 128 GB RAM.

We built two segmentation models using the architecture of U-Net++/ResNet50 and DeepLabv3+/ResNet101 with our own created dataset. ResNet50 and ResNet101 were used as the encoder for U-Net++ and DeepLabv3+ respectively and the model weights obtained while training with ImageNet were transferred to our model. The backend of each model was tested with different variations of ResNet and the most appropriate one was selected. U-Net++/ResNet101 was unable to run on our system and thus we had to stay with ResNet50.

The main challenge in segmenting the necrosis was the colour variation in staining. Pre-processing methods built into U-Net++ and DeepLabv3+ were used for the data cleaning. The dimension of the image was 512x512x3 and that of the mask was 512x512x1. The entire dataset was divided into 3 sets for training, validation and testing. We used 16 workers, with a data batch size of 64.

We used dice loss and intersection over union (IoU) as the performance metrics. Dice loss was used to measure the loss and the IoU score accuracy. A combination of binary cross-entropy and dice coefficient was used as the loss function in this model, which can be denoted as follows (Van Beers et al., 2018):

$$L(Y, \hat{Y}) = -\frac{1}{N} \sum_{b=1}^N \left(\frac{1}{2} \cdot Y_b \cdot \log \hat{Y}_b + \frac{2 \cdot Y_b \cdot \hat{Y}_b}{Y_b + \hat{Y}_b} \right) \quad (2)$$

Where N is the batch size, Y_b is the ground truth, and \hat{Y}_b is the prediction value for the b^{th} image.

In semantic segmentation, accuracy cannot be considered a good performance metric, due to the class imbalance in such cases. IoU can replace it and find the degree of coincidence between the predicted boundary and the actual boundary (Van Beers et al., 2018).

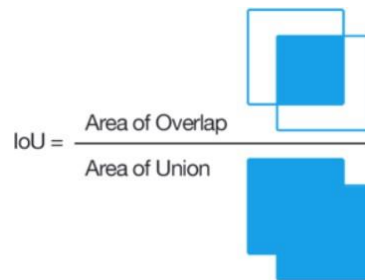


Figure 9. IoU is the area of intersection of predicted segmentation and the ground truth divided by the area of union between the same. Source: (Tiu, 2019).

Two important hyperparameters used here are the optimizer and the learning rate. Here, we use Adam as the optimizer and the learning rate is set with a value of 0.0001. Sigmoid is used as the activation function here as this is a binary model. 120 epochs are used for training the model and after 80 epochs, the learning rate is reduced to 0.00001. As the learning rate decreases, the training time increases and so does the GPU cost, but the prediction is more accurate. The best model is saved based on the highest value of the validation IoU score obtained during the training period. U-Net took almost 3 hours to run the 120 epochs whereas the DeepLab took less than 2 hours.

The following algorithms were used to build the model and quantification process.

Pseudo code 1. Model training.

```

1. Set learning rate = 0.0001
2. Set max_score = 0
3. Repeat for 120 epochs
    3.1 if validation accuracy is greater than max_score,
        then set max_score = validation accuracy
    3.2 save the model with training and validation parameters of that epoch
    3.3 if epoch = 80
        then reduce learning rate = 0.00001

```

2.4.1 Necrosis percentage calculation

The trained model can segment the necrotic tissue area from a histopathology image. The output of the model is a black-and-white image. The white portion represents the necrosis and the black denotes the rest of the image. Thus, we can find the area of necrosis by counting the non-zero pixels in the image. We use the function `countNonZero()` present in the `opencv` library. The total area of the image is 512×512 as the resolution of the image is the same. Now the percentage of the necrosis can be found using the following formula: Percentage = (number of pixels of necrosis / total number of pixels in image) * 100.

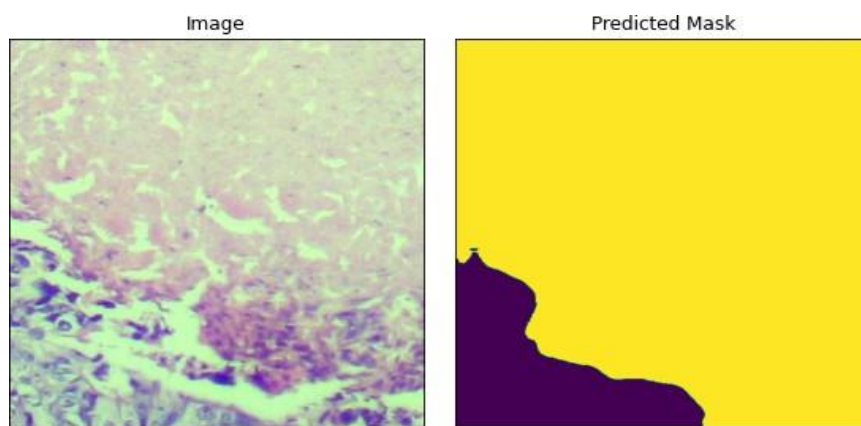


Figure 8. Original image (left) and predicted mask (right). In this case, volume = 86.9%.

Pseudo code 2. Finding percentage of necrosis from a given image.

1. Read the image of any size
2. Do the following pre-processing steps on image
 - 2.1. resize the image into 512x512
 - 2.2. convert image into tensor
 - 2.3. normalize the tensor
3. Load the saved model
4. Predicted mask = model.predict(image)
5. nec_pixels = Get the count of pixels having non-zero value, using cv2.countNonZero(pr_mask)
6. total_pixels = total number of pixels in the whole image (here 512*512)
7. Percentage = (nec_pixels / total_pixels) * 100.

3 Results

We developed a deep-learning solution for finding the quantity of tumour necrosis created due to neoadjuvant chemotherapy in cancer patients. In the current study, we built two models with pre-trained segmentation models, U-Net++ and DeepLabv3+. We used dice loss and IoU as the performance metrics. Table 1 explains the segmentation loss and accuracy during training, validation and testing.

Table 1. Dice loss and IoU obtained during segmentation.

Parameters	U-Net++	DeepLabv3+
Training loss	0.0440	0.0477
Validation loss	0.2796	0.2582
Testing loss	0.1777	0.1780
Training IoU	0.9178	0.9109
Validation IoU	0.7513	0.7457
Testing IoU	0.8264	0.8150

The values of loss and accuracy during the training and validation process were collected and plotted as graphs for both U-Net++ and DeepLabv3+, as shown in Figures 10 and 11. Loss or accuracy are the y-axis and epochs are the x-axis.

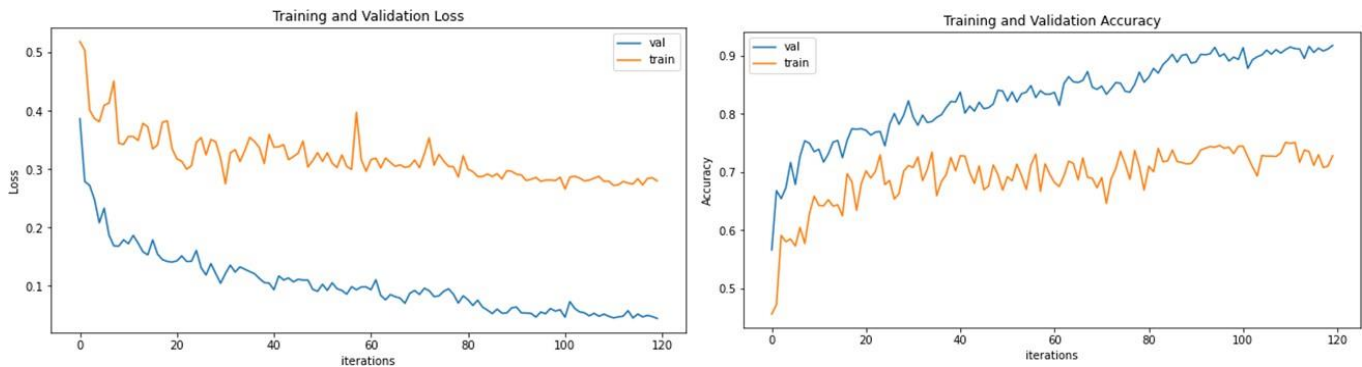


Figure 10. Training and validation loss (left) and accuracy (right) of U-Net++model.

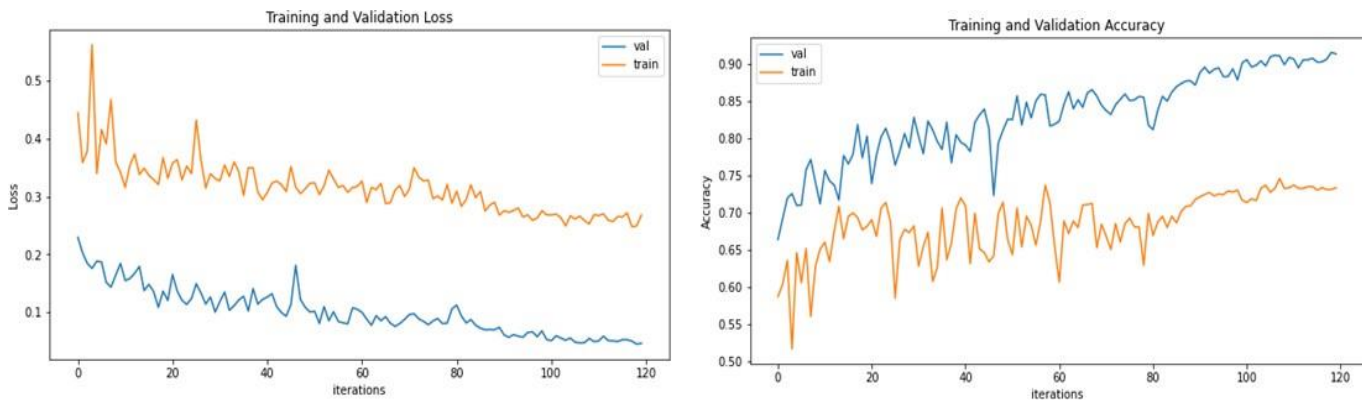


Figure 11. Training and validation loss (left) and accuracy (right) of DeepLabv3+ model.

The graphs indicate that, as the epochs increase, the loss decrease and accuracy increase in both models. The inference from the study is that both the models perform more or less the same and also have many similarities with the manual annotation. Figure 12 clearly depicts how the mask or segmented region originated from our proposed model related to the actual mask. Most of the testing images show good likeness with the original images.

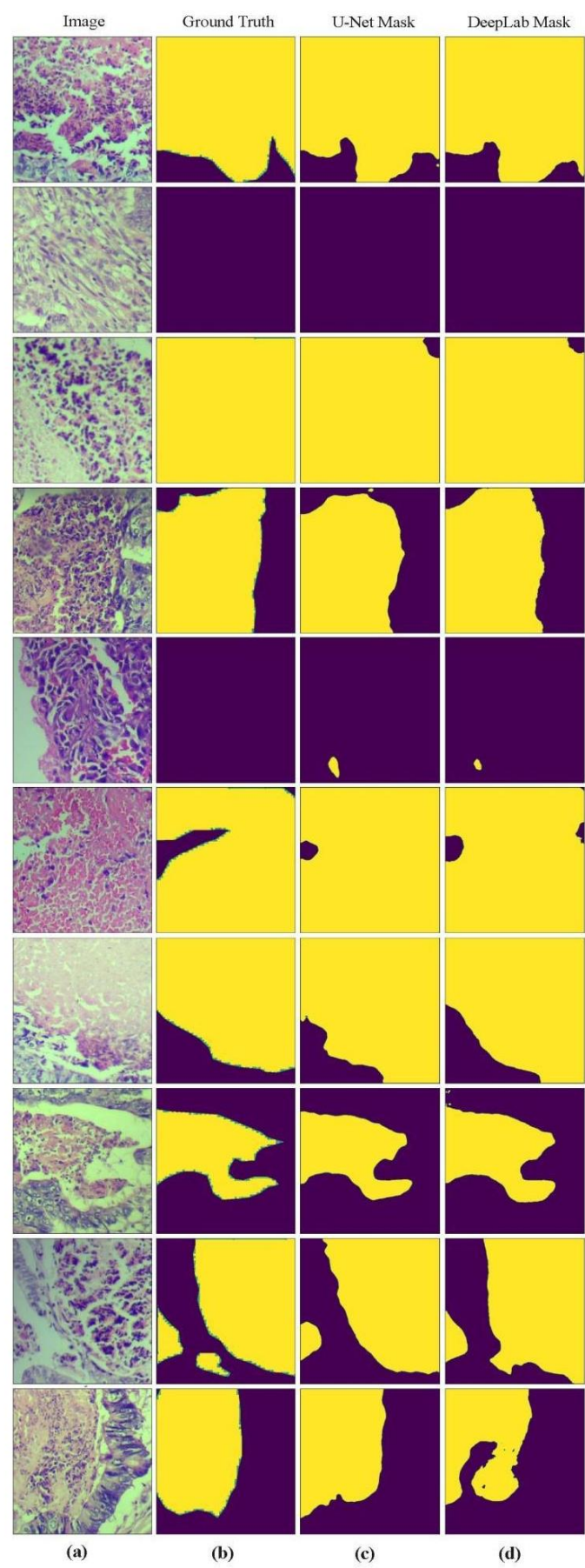


Figure 12. Segmentation results with training set: (a) original image, (b)manually annotated ground truth, (c) result from the U-Net++, (d) result from the DeepLabv3+ model.

4 Discussion

Osteosarcoma is a high-grade malignant bone tumour for which NAC is a vital component of the treatment plan. NAC is given as the first step before surgery to shrink tumour volume and thereby tumour burden. Chemotherapy brings about tumour necrosis and the percentage of necrosis is an important factor in deciding on further treatment. The quantification of necrosis is done manually by visualizing tissue sections through the microscope, “eye-balling” the percentage of necrosis per slide, and calculating the total percentage. This is a crude method that can cause significant inter-observer bias. The proposed system is an AI-powered clinical decision-making tool for cancer patients, specifically with osteosarcoma. The intended end-users of this tool are pathologists, who are now manually quantifying the volume of cancerous tissues deceased due to the chemotherapy applied to the patient's body. The pathologists can feed the histopathology images captured from the digital microscope into this tool and the output will be a segmented image and the total area of the segmented part. The evaluation metrics that we used in this work are dice loss and IoU, which are the standard metrics for segmentation work. IoU is the perfect measure to find the overlap between predicted regions and actual regions. A greater value of IoU indicates that the extent of overlap between two masks is greater and, thus, there is good accuracy in the segmentation process.

4.1 Limitation of proposed model

The dataset created is based on limited data. One of the reasons is that, before the study can proceed, the pathologist has to annotate the data, and this is a time-consuming manual process that limits the number of annotated images that can be made available. This entire model was, therefore, trained and validated on patches from images and not the entire images that were captured directly from the microscope. Thus if an entire image (high-resolution image) is used during external validation, there is a chance of wrong predictions.

The data collected for this study are mostly noise free and well-stained. Where that is not the case, noisy images can lead to wrong predictions. Another limitation of this model is that it segments only necrosis whereas the tissue might contain other elements such as fibrosis, oedema, normal tissues, tumours, etc., and those are not taken into account. Also, the GPU that we had was not sufficient to run the higher-order networks such as UNet++/ResNet101, which would have achieved higher accuracy.

Lastly, as there is no previous work in this domain, the benchmarking of the dataset is difficult. Further research that may follow this study may include more variation in data with separate annotations of different tissue elements.

4.2 Future work on proposed model

As a future extension of this work, segmentation of other kinds of tissues such as fibrosis, oedema, tumour or any other tissue component that may be of therapeutic importance can be attempted. After developing the full-fledged software, it can be implemented in different hospitals for an automated tissue quantification and classification process. The problem of segmenting high-resolution images can be solved as follows. The input image can be split into patches, the segmentation can be applied to each patch and finally, all these patches can be merged into a single original image. The same process of merging should be applied to the segmented image in the same order as the original image.

5 Conclusion

Currently, recognition of tissue necrosis and its estimation is done manually. This is a very error-prone and time-consuming process. In this study, we presented a semantic segmentation and quantification tool based on a convolutional neural network. It focuses on dead tissues resulting from neoadjuvant

chemotherapy. The entire clinical decision-making is based on this measurement. Even though image segmentation is a very common task in the deep learning field, segmentation of necrotic tissues from a histopathology image is a novel domain in it. Thus, data are very scarce in this domain. In this work, therefore, we created a dataset of 555 images and two models were trained on it: UNet++/ResNet101 and DeepLabv3+/ResNet101, UNet++ showed slightly better performance than the other. However, the DeepLab networks require less computation power compared to U-Net versions. The U-Net++ model showed 91.78% accuracy and a dice loss of 0.0440, whereas DeepLabv3+ achieved 91.09% accuracy and 0.9109 loss. Both models segmented the images in an almost similar way. The main challenge in segmenting the necrosis is the colour variation in staining. Thus, we collected specimens whose staining proportion varies greatly and trained the model with such images.

Abbreviations

NAT: Neo-adjuvant treatment

NACT: Neo-adjuvant chemotherapy

NAC: Neo-adjuvant chemotherapy

PET: Positron Emission Tomography scan

H&E: Hematoxylin & Eosin

Additional Information and Declarations

Conflict of Interests: The authors declare no conflict of interest.

Author Contributions: T.S.S.: Conceptualization, Methodology, Software, Writing – original draft. P.M.I: Supervision, Software. V.M.K.S.: Data curation. A.K.M.B.H.: Writing – review & editing.

Institutional Review Board Statement: This study was approved on 29 December 2020 by MVR Cancer Centre and Research Institute, Calicut, India Institutional Review Board, (MVRIRB2020/03/02) protocol number EC Ref. No. MVR/IEC/III/2020/02.

Informed Consent Statement: Informed consent document is not involved in this study as the patients are not directly involved in the study. The study doesn't contain any ethical concerns.





Data Availability: The data were collected by MVR Cancer Centre and Research Institute, Calicut, India, and it was annotated by Dr V. M. Kutty Sajna, a Senior Pathologist. Due to the nature of the research, due to legal issues supporting data is not publicly available.

References

- Alawad, A. A. M. (2014). Evaluation of Clinical and Pathological Response after Two Cycles of Neoadjuvant Chemotherapy on Sudanese Patients with Locally Advanced Breast Cancer. *Ethiopian Journal of Health Sciences*, 24(1), 15. <https://doi.org/10.4314/ejhs.v24i1.2>
- Badrinarayanan, V., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12), 2481–2495. <https://doi.org/10.1109/tpami.2016.2644615>
- Boulch, A. (2017). Sharesnet: reducing residual network parameter number by sharing weights. *arXiv preprint arXiv:1702.08782*. <https://doi.org/10.48550/arXiv.1702.08782>
- Burcombe, R., Makris, A., Pittam, M., Lowe, J., Emmott, J., & Wong, W. C. (2002). Evaluation of good clinical response to neoadjuvant chemotherapy in primary breast cancer using [18F]-fluorodeoxyglucose positron emission tomography. *European Journal of Cancer*, 38(3), 375–379. [https://doi.org/10.1016/s0959-8049\(01\)00379-3](https://doi.org/10.1016/s0959-8049(01)00379-3)
- Bychkov, D., Linder, N., Turkki, R., Nordling, S., Kovanen, P. E., Verrill, C., Walliander, M., Lundin, M., Haglund, C., & Lundin, J. (2018). Deep learning based tissue analysis predicts outcome in colorectal cancer. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-018-21758-3>

- Chan, J. K. (2014). The Wonderful Colors of the Hematoxylin–Eosin Stain in Diagnostic Surgical Pathology. *International Journal of Surgical Pathology*, 22(1), 12–32. <https://doi.org/10.1177/1066896913517939>
- Chandhok, S. (2021) U-Net: Training Image Segmentation Models in PyTorch. <https://pyimagesearch.com/2021/11/08/u-net-training-image-segmentation-models-in-pytorch/>
- Caparica, R., Lambertini, M., Pondé, N., Fumagalli, D., De Azambuja, E., & Piccart, M. (2019). Post-neoadjuvant treatment and the management of residual disease in breast cancer: state of the art and perspectives. *Therapeutic Advances in Medical Oncology*, 11, 175883591982771. <https://doi.org/10.1177/1758835919827714>
- Chen, L. C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*. <https://doi.org/10.48550/arXiv.1706.05587>
- Chen, L., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018a). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In *Computer Vision – ECCV 2018*, (pp. 833–851). Springer. https://doi.org/10.1007/978-3-030-01234-2_49
- Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018b). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834–848. <https://doi.org/10.1109/tpami.2017.2699184>
- Davis, A., Bell, R., & Goodwin, P. J. (1994). Prognostic factors in osteosarcoma: A critical review. *Journal of Clinical Oncology*, 12(2), 423–431. <https://doi.org/10.1200/jco.1994.12.2.423>
- Deng, J., Dong, W., Socher, R., Li, L., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition*. IEEE. <https://doi.org/10.1109/cvpr.2009.5206848>
- Elies, A., Rivière, S., Pouget, N., Becette, V., Dubot, C., Donnadiou, A., Rouzier, R., & Bonneau, C. (2018). The role of neoadjuvant chemotherapy in ovarian cancer. *Expert Review of Anticancer Therapy*, 18(6), 555–566. <https://doi.org/10.1080/14737140.2018.1458614>
- ESMO. (2014). Bone sarcomas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology*, 25, iii113–iii123. <https://doi.org/10.1093/annonc/mdu256>
- Estienne, T., Vakalopoulou, M., Christodoulidis, S., Battistella, E., Lerousseau, M., Carré, A., Klausner, G., Sun, R., Robert, C., Mougiakakou, S., Paragios, N., & Deutsch, E. (2019). U-ReSNet: Ultimate Coupling of Registration and Segmentation with Deep Nets. In *International conference on medical image computing and computer-assisted intervention, Lecture Notes in Computer Science*, (pp. 310–319). Springer. https://doi.org/10.1007/978-3-030-32248-9_35
- Facchin, C., Certain, A., Yoganathan, T., Delacroix, C., Garcia, A., Gaillard, F., Lenoir, O., Tharaux, P., Tavitian, B., & Balvay, D. (2022). FIBER-ML, an Open-Source Supervised Machine Learning Tool for Quantification of Fibrosis in Tissue Sections. *American Journal of Pathology*, 192(5), 783–793. <https://doi.org/10.1016/j.ajpath.2022.01.013>
- Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M. S., Parkin, D. M., Forman, D., & Bray, F. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, 136(5), E359–E386. <https://doi.org/10.1002/ijc.29210>
- Gertych, A., Ing, N., Ma, Z., Fuchs, T. J., Salman, S., Mohanty, S. K., Bhele, S., Velásquez-Vacca, A., Amin, M. B., & Knudsen, K. E. (2015). Machine learning approaches to analyze histological images of tissues from radical prostatectomies. *Computerized Medical Imaging and Graphics*, 46, 197–208. <https://doi.org/10.1016/j.compmedimag.2015.08.002>
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. <https://doi.org/10.1109/cvpr.2014.81>
- Guo, Y., Liu, Y., Georgiou, T., & Lew, M. S. (2018). A review of semantic segmentation using deep neural networks. *International Journal of Multimedia Information Retrieval*, 7(2), 87–93. <https://doi.org/10.1007/s13735-017-0141-z>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016a). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778). IEEE. <https://doi.org/10.1109/CVPR.2016.90>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016b). Identity mappings in deep residual networks. In *European conference on computer vision* (pp. 630–645). Springer, Cham. https://doi.org/10.1007/978-3-319-46493-0_38
- Ikeda, T., Jinno, H., Matsui, A., Masamura, S., & Kitajima, M. (2002). The role of neoadjuvant chemotherapy for breast cancer treatment. *Breast Cancer*, 9(1), 8–14. <https://doi.org/10.1007/bf02967540>
- Kann, B. H., Thompson, R., Thomas, C. R., Dicker, A., & Aneja, S. (2019). Artificial Intelligence in Oncology: Current Applications and Future Directions. *Oncology*, 30(2), 46–53. <https://www.cancernetwork.com/view/artificial-intelligence-oncology-current-applications-and-future-directions>
- Komura, D., & Ishikawa, S. (2017). Machine Learning Methods for Histopathological Image Analysis. *Computational and Structural Biotechnology Journal*, 16, 34–42. <https://doi.org/10.1016/j.csbj.2018.01.001>
- Lee, C. Y., Xie, S., Gallagher, P., Zhang, Z., & Tu, Z. (2015). Deeply-supervised nets. In *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, (pp. 562–570). PMLR. <https://proceedings.mlr.press/v38/lee15a.html>
- Li, H., Xu, Z., Taylor, G., Studer, C., & Goldstein, T. (2018). Visualizing the loss landscape of neural nets. In *Advances in neural information processing systems*, vol. 31, NeurIPS Proceedings. <https://proceedings.neurips.cc/paper/2018/file/a41b3bb3e6b050b6c9067c67f663b915-Paper.pdf>

- Liu, D., & Yu, J. Z. (2009). Otsu Method and K-means. International Conference Hybrid Intelligent Systems. In *2009 Ninth International Conference on Hybrid Intelligent Systems*. IEEE. <https://doi.org/10.1109/his.2009.74>
- Luchini, C., Pea, A., & Scarpa, A. (2022). Artificial intelligence in oncology: current applications and future perspectives. *British Journal of Cancer*, 126(1), 4–9. <https://doi.org/10.1038/s41416-021-01633-1>
- Maknuna, L., Kim, H. S., Lee, Y., Choi, Y., Kim, H. J., Yi, M., & Kang, H. (2022). Automated Structural Analysis and Quantitative Characterization of Scar Tissue Using Machine Learning. *Diagnostics*, 12(2), 534. <https://doi.org/10.3390/diagnostics12020534>
- Picci, P., Sangiorgi, L., Rougraff, B. T., Neff, J. R., Casadei, R., & Campanacci, M. (1994). Relationship of chemotherapy-induced necrosis and surgical margins to local recurrence in osteosarcoma. *Journal of Clinical Oncology*, 12(12), 2699–2705. <https://doi.org/10.1200/jco.1994.12.12.2699>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28
- Ruspini, E. H., Bezdek, J. C., & Keller, J. M. (2019). Fuzzy Clustering: A Historical Perspective. *IEEE Computational Intelligence Magazine*, 14(1), 45–55. <https://doi.org/10.1109/mci.2018.2881643>
- Sabeerali, K. P., Saleena, T. S., Muhamed Ilyas, P., & Mohan, N. (2022). AI-Powered Semantic Segmentation and Fluid Volume Calculation of Lung CT Images in COVID-19 Patients. In *Emergent Converging Technologies and Biomedical Systems* (pp. 93–101). Springer. https://doi.org/10.1007/978-981-16-8774-7_9
- Shukla, S., & Naganna, S. (2014). A review on K-means data clustering approach. *International Journal of Information and Computation Technology*, 4(17), 1847–1860.
- Strobel, H. A., Schultz, A., Moss, S. J., Eli, R., & Hoving, J. B. (2021). Quantifying Vascular Density in Tissue Engineered Constructs Using Machine Learning. *Frontiers in Physiology*, 12. <https://doi.org/10.3389/fphys.2021.650714>
- Tiu, E. (2019). Metrics to Evaluate your Semantic Segmentation Model. <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>
- van Beers, F. (2018). *Using intersection over union loss to improve binary image segmentation*. University of Groningen.
- Wang, Y., Wang, G., Wei, L., Huang, L., Wang, J., Wang, S., Li, X., Shen, D., Bao, D., & Gao, J. (2011). Neoadjuvant chemotherapy for locally advanced cervical cancer reduces surgical risks and lymph-vascular space involvement. *Chinese Journal of Cancer*, 30(9), 645–654. <https://doi.org/10.5732/cjc.011.10050>
- Yan, Z., Zhang, H., Jia, Y., Breuel, T., & Yu, Y. (2016). Combining the best of convolutional layers and recurrent layers: A hybrid network for semantic segmentation. *arXiv preprint arXiv:1603.04871*. <https://doi.org/10.48550/arXiv.1603.04871>
- Zhang, Q., Cui, Z., Niu, X., Geng, S., & Qiao, Y. (2017). Image Segmentation with Pyramid Dilated Convolution Based on ResNet and U-Net. In *ICONIP 2017: Neural Information Processing*, (pp. 364–372). Springer. https://doi.org/10.1007/978-3-319-70096-0_38
- Zhang, R., Du, L., Xiao, Q., & Liu, J. (2020). Comparison of Backbones for Semantic Segmentation Network. *Journal of Physics: Conference Series*, 1544(1), 012196. <https://doi.org/10.1088/1742-6596/1544/1/012196>
- Zhou, Z., Siddiquee, M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In *DLMIA 2018, ML-CDS 2018: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, (pp. 3–11). Springer. https://doi.org/10.1007/978-3-030-00889-5_1

Editorial record: The article has been peer-reviewed. First submission received on 28 September 2022. Revisions received on 27 November 2022 and 22 January 2023. Accepted for publication on 30 January 2023. The editors coordinating the peer-review of this manuscript were Mazin Abed Mohammed , Seifedine Kadry , and Oana Geman . The editor in charge of approving this manuscript for publication was Zdenek Smutny .

Special Issue: Deep Learning Blockchain-enabled Technology for Improved Healthcare Industrial Systems.

Acta Informatica Pragensia is published by Prague University of Economics and Business, Czech Republic.

ISSN: 1805-4951