

Classification of Lung Cancer Types from Histopathological Images using Semantic Segmentation of Nuclei and Deep Learning

Mirunalini P^{1*}, Karthik Desingu^{1†}, Pavya S^{1†} and Priyadharshini N^{1†}

^{1*}Dept. of Computer Science and Engineering, Sri Sivasubramaniya Nadar College of Engineering, OMR, Chennai, 603110, Tamilnadu, India.

*Corresponding author(s). E-mail(s): miruna@ssn.edu.in;

Contributing authors: karthik19047@cse.ssn.edu.in; pavya17102@cse.ssn.edu.in ;
priyadharshini17118@cse.ssn.edu.in;

[†]These authors contributed equally to this work.

Abstract

Lung Cancer is the leading cause of cancer deaths worldwide. This form of cancer begins in the lungs and may spread to the lymph nodes or other organs in the body. There are several types of lung cancer, among which Adenocarcinoma and Squamous-Cell carcinoma are most prevalent. Diagnosis and treatment of lung cancer is very much needed to know about the cancer type. Automated detection of cancer types would significantly speed up the diagnosis and treatment. So we propose two different methods to classify the lung cancer types using Deep learning techniques. In our first method we proposed a Convolutional Neural Network (CNN) to classify the lung cancer types from the whole slide histopathological images and attained an accuracy of 99%. The cancer mainly affects the nucleus of the tissues, so in our second method we performed the classification of lung cancer types from the segmented nuclei. The nuclei were segmented from the whole slide histopathological images using threshold based and color based methods. The segmented nuclei were classified using CNN and achieved an accuracy of 99%. We have performed many experiments and evaluated the proposed methods at each level.

Keywords: Segmentation, Histopathological Images, Adenocarcinoma, Squamous Cell Carcinoma, Benign, Deep learning, Convolution Layer

1 Introduction

The lungs are a vital part of the human respiratory system. Taking in oxygen and getting rid of carbon dioxide is among the key functions of the lungs. In some abnormal cases, lung cells may grow and multiply uncontrollably and result in the development of tumors. Lung tumors can be either cancerous (malignant) or benign (non-cancerous).

Such cancerous cells are known to occur in various parts of the lungs, including the bronchi, the bronchioles and the alveoli. While the primary cause of lung cancer is smoking, lung lesions in non-smokers may also occur due to various factors such as pollution, presence of Human Immunodeficiency Virus (HIV), infections, tuberculosis, and malnutrition [31]. Other less pronounced causes

may include gene changes or workplace exposure to asbestos, diesel exhaust and certain other chemicals.

Among a variety of imaging modalities available to detect lung tumors, CT scan, MRI scan and histopathological imaging are the most predominant. However, repeated exposure to radiation and injection of potentially toxic contrast agents during CT and MRI imaging can have adverse effects on the patient's health. The histopathological imaging procedure is a relatively safer modality. The term histopathology refers to the study of tissue characteristics through the analysis of blood samples of the concerned tissue. In the histopathological imaging technique, tissues are stained with Hematoxylin and Eosin (H&E). The stained images are further examined under a microscope to detect malignant features in the cellular structures. Owing to its high accuracy in diagnosis and sub-typing of lung lesions, histopathological examination is considered as a gold-standard in cancer diagnostics. [1]. It helps in distinguishing between lesions not only during the diagnostic phase, but also during follow-ups and further helps in the prognostication and treatment decisions [35]. Hence, the analysis of histopathological images is invaluable in the quantitative analysis of clinical variables based on biologically relevant features.

The tumor cells in an histopathological image exhibit high amounts of irregularity and arbitrary visual angles. Hence, the manual analysis and detection of cancerous tumors from these images is a time-consuming process. Furthermore, it is subject to inter and intra observer variability [12]. An automated system can overcome these issues and perform faster analysis and diagnosis. Such automation can be achieved through Machine Learning (ML) systems. The accuracy of such an ML system greatly depends on quality of feature engineering. Traditional ML systems rely on hand-crafted features, where the model designer chooses the relevant features based on expert opinions and reviews. However, certain visual features may not be apparent to humans. On the other hand, a deep-learning system can overcome this problem by learning impactful features for classification automatically, irrespective of their perceptibility to the human eye. Hence, we propose to develop an automated deep-learning system to analyze the

histopathological images of the lung tumor cells and detect the presence of cancerous cells.

The histopathological images of the tumor affected tissues contain the nucleus, along with various other structures. The nuclei specifically, can exhibit a wide variety of patterns that are diagnostically significant. The appearance of the nuclei may be different due to several factors, including the nucleus type, malignancy of the disease and the nuclei life cycle. In particular, the nuclei of benign, cancerous and their sub-types lung tissues are known to possess distinguishable characteristics that may not be apparent to the human eye. Hence, analyzing the nuclei in the histopathological image may provide key information for identifying the presence and stage of the disease [22].

Instead of giving the Whole Slide Histopathological Image (WSHI) segmenting the nucleus and providing it to the Deep Convolutional Neural Networks (DCNN) classifier helps to limit the information available in the input images. This can greatly reduce the effort of the DCNN in learning to suppress the weights of less relevant features stemming from other regions of the image, thereby improve the classifier's performance. Moreover, retaining only the nuclear regions from a WSHI patch gives rise to images that mask the non-nuclear regions with sharp gradients at the nucleus boundaries. As a result of such highly pronounced boundaries, shape-based features will be learned much more robustly and rapidly. Thus, we propose to segment the nuclear regions from the histopathological images and extract features exclusively from these regions to perform classification. This approach is compared with a baseline established using the classification performed with the whole-slide histopathological images.

In this paper, we propose a deep-learning based method to perform binary (Benign and Malignant) and multiclass (Benign, Adenocarcinoma (ADC) and Squamous-Cell Carcinoma (SCC)) classification of lung tumors using DCNN from WSHI patches of tumorous lung tissues as well as from the nucleus segmented WSHI patches. We then compare the two approaches to prove the positive impact of segmenting the nuclear regions on the classification performance.

2 Related Work

The technique of using Machine Learning (ML) and Deep Learning (DL) to analyze histopathological images to draw medical inferences has proven to be very effective in recent times and is a highly active research field in the contemporary world.

In our work, we demonstrate how the field can greatly benefit from deep-learning using CNNs, which can automatically learn features from the image that may not be obvious to suspect and include during conventional image processing methods.

Yu et al. [50] have combined conventional thresholding and image processing techniques with machine learning methods, such as random forest classifiers, SVM or Naïve Bayes classifiers and achieved an Area Under the Curve (AUC) of 0.85 in distinguishing normal from tumor slides and 0.75 in distinguishing LUAD from LUSC slides. In [26] the authors performed an automatic segmentation of cervical cells in microscopic images by segmenting the cell cytoplasm through triangle thresholding and further refined by cell-wise contour refinement using graph cuts. A Computer Aided Diagnosis (CAD) based on morphological features has been implemented for detection of cancerous and non-cancerous tissues from histopathological images in [39]. A Hierarchical Conditional Random Field (HCRF) was proposed in [41] for gastric histopathological image segmentation with automatically localized abnormal regions. This method further used higher-order potentials namely, pixel-level and patch-level potentials to improve the quality of segmentation. In the research work proposed in [45], the nuclei were first located using bounding-boxes through hough-transform and marker controlled water-shield approach. About 20 features were then extracted and used to perform classification using a Support Vector Machine (SVM), a K Nearest Neighbor (KNN) and a decision tree classifiers.

The accuracy of ML based methods in classification of histopathological images highly depends on the quality of handcrafted features. Framing effective features often requires manual enumeration of intuitively meaningful characteristics of images in relation to the goal of classification and formulation of algorithms to extract them. On the contrary, DL based methods are characterized by

automatic extraction of impactful and large volume of features with the help of neural networks. These architectures can learn to extract features from images that may not even be perceptible to a human. As a result, the extent of research leveraging ML methods is limited and the general trend has shifted towards the use of DL based approaches.

A multi-scale deep residual aggregation network was proposed in [47] to segment the nuclear material from pathological images of non-small cell lung cancer tissues with the help of patch-level statistics and morphological features. A fusion module based on Feature Pyramid Network (FPN) was proposed in [8] to detect nuclear locations by combining complementary information from shallow layers with that from the deeper layers. These cropped patches were fed into a UNet architecture to separate the clustered nuclei. Two ML approaches were proposed for multiclass classification of breast cancer histopathological images in [38]. In the first approach, hand-crafted features were extracted and used to train the conventional ML classifiers. On the other hand, the second approach used transfer learning technique, wherein VGG16, VGG19 [40] and ResNet-50 were employed as feature extractors. The research work in [42] proposed a CNN architecture to extract self-learned features for diagnosis of lung cancer and feature learning has been improved using GAN generated images. A DL network was used in [37] to extract a large number of features. Spiral setting and approximation concepts were used to perform dimensionality reduction and the reduced feature set was boosted using an ensemble classifier to achieve a high classification accuracy. The automated classification of ADC and SCC was proposed in [16] using a DL framework to classify the input patches as diagnostic and non-diagnostic patches. A random forest regression model is then applied on the diagnostic patches to classify each WSHI as either ADC or SCC based on a collection of statistical and morphological measurements extracted from the input WSHI. A WSHI classification pipeline into five diagnostic categories of breast cancer using deep learning methods has been proposed in [15]. In [10], the Inception-V3 architecture [44] is applied with histopathology images obtained from The Cancer Genome Atlas (TCGA) to accurately classify the WSHI into one of three classes namely, ADC, SCC

and Benign tumor cells. An automated method for the detection and segmentation of breast cancer nuclei from high-resolution histopathological images using a Convolution Neural Network (CNN), initialized with an active contour model for subsequent nucleus segmentation and finally an adaptive ellipse fitting procedure to remove overlapping clumped nuclear regions was proposed in [49].

In this work [23], the histopathological images were first color normalized using a Convolutional Gaussian Mixture model to cluster pixels with attention to structure of the nuclei. The nuclei from these images were then segmented using the state-of-the-art Mask Residual Convolutional Neural Network (Mask R-CNN).

In [48], a comparative study of twelve nuclei segmentation methods for cytology pleural effusion images is presented. Each method involves three main steps — preprocessing, segmentation, and postprocessing. In particular, the segmentation methods are quantitatively evaluated for 35 cytology images of Pleural Effusion by computing five performance metrics. Evaluation results show that the segmentation performances of the Otsu’s thresholding, K-Means clustering, Mean-Shift, Chan–Vese, and Graph Cut methods are significantly superior with accuracy measures of 94, 94, 95, 94, and 93% respectively, characterized by high abnormal nuclei detection rates. In [2], the authors have addressed the problem in splitting heavily clustered nuclei due to its intensity variation caused by noise and uneven absorption of stains, by implementing a multilevel thresholding procedure and a subsequent watershed algorithm to separate the clustered nuclei. A novel method for unsupervised segmentation of cell nuclei in stained histology tissue is presented in [34]. The segmentation step comprised of multilevel thresholding and a series of morphological operations. A nucleus-boundary model to predict the nuclear regions and their corresponding boundaries using a fully Convolutional Neural Network was proposed in [11].

An increasing use of CNN methods is evident and this stems from its ability to extract increasingly complex features as the image percolates to deeper layers of the network and its ability to learn effective feature extraction filters during the training phase. In interest of the proposed work, the wide range of DL based methods can be segregated

into two broad approaches — ones that perform direct classification of the WSHI patches and others that segment the nuclei from these patches before classification. In this work, we experiment with both approaches and present a performance comparison of the two pipelines. Concretely, the methods proposed and work detailed in this paper have the following key highlights:

- A Segmentation Network based on the Xception-style [9] UNet [36] is proposed to segment the nuclear regions from a Whole Slide Histopathological Image (WSHI) patch. The network achieves a mean-IoU of 0.9427 on the test data.
- An algorithmic image-processing method is proposed to segment nuclear regions from WSHI patches based on intensity variations introduced by staining and using contour detection.
- Finally, a Deep Convolutional Neural Network (DCNN) is proposed to perform binary and multiclass classification of lung tumors using complete WSHI patches as well as using the nucleus segmented patches. The performance of the classifier in the two cases is compared and analyzed.

3 Overview of Proposed System

In this paper, we propose a DCNN that performs *two forms* of classification of the WSHI patch images — binary and multiclass. While binary classification identifies an input image as either benign or malignant, multiclass classification will assign one of three classes to an input image — ADC, SCC and Benign. To improve the classification accuracy of the DCNN we performed semantic segmentation to retain only the nuclear regions of the input WSHI patches. We present two different approaches to perform segmentation, one a threshold based image processing method and the other, a deep-learning based segmentation using an Xception-style UNet. Further, we study and present the difference in performance of the classifier when it is trained with segmented nuclear regions and also with entire WSHI patches. The flowchart of our proposed system is presented in Figure 1 and the overall workflow is depicted in Algorithm 1.

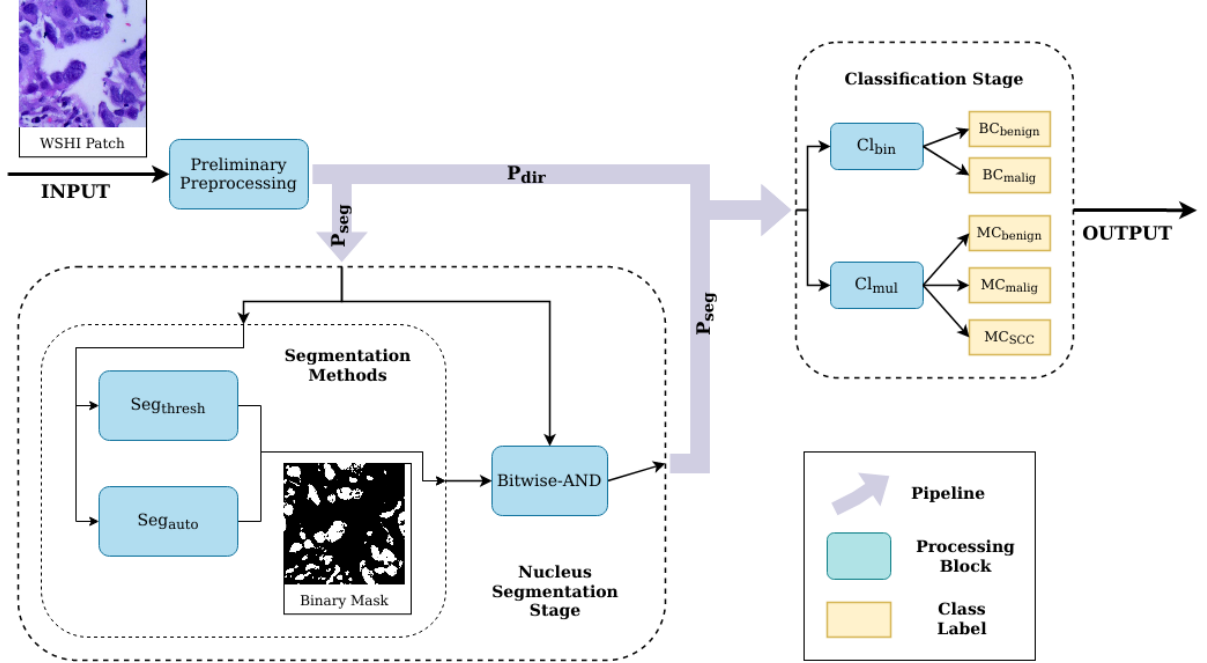


Fig. 1 A flowchart representing the proposed system. Pipelines P_{dir} and P_{seg} , as described in Section ??, represent the pipelines for direct and nucleus segmented classification of the WSHI patches.

3.1 Preliminary Preprocessing

Our proposed methodology involves training a DCNN to perform classification. To avoid over fitting, to adopt a diversity-based sampling strategy and to cater to the large input requirement of the deep learning network, data augmentation is necessary. Random affine transformations such as rotation, scaling and shearing are applied to augment the input image i .

The input images are then resampled to 512×512 , 3-channel images using bi-linear interpolation. The images are then linearly-normalized to envelope the pixel values between 0 and 1. This is known to improve the rate of convergence and enhance the network's ability to standardize to diverse inputs.

The image i_{prep} obtained after the preliminary stage of preprocessing is passed to two different pipelines — P_{dir} and P_{seg} . The direct pipeline (P_{dir}), feeds the preliminary preprocessed WSHI patch i_{prep} directly as input to the DCNN classifiers, whereas the segmentation pipeline (P_{seg}) first performs semantic segmentation of i_{prep} to obtain an image i_{seg} using either thresholding or using the U-Net based segmentation model

to contain only the nuclear regions of i_{prep} . The segmented image i_{seg} is passed to the classifiers.

3.2 Classification using Entire WSHI Patches — P_{dir} Pipeline

The first pipeline proposed for classification — P_{dir} , feeds the WSHI patch i_{prep} obtained after the preliminary preprocessing stage directly to the DCNN classifiers. The DCNN classifier Cl_{bin} assigns a binary class label bc , whereas Cl_{mul} assigns a multiclass label mc .

3.2.1 Architecture of the DCNN Classifiers

The proposed DCNN classifier comprises of six convolution blocks. Each block consists of a convolution layer followed by one each of activation, batch normalization and max-pooling layers. Dropout layers are added at the end of alternate convolution blocks to prevent excessive co-adaptation of the model, commonly regarded as overfitting. The convolution blocks extract distinctive features from the input images. The initial layers extract high-level features. As the images percolates through to the deeper layers of the network, the channel-size is increased progressively

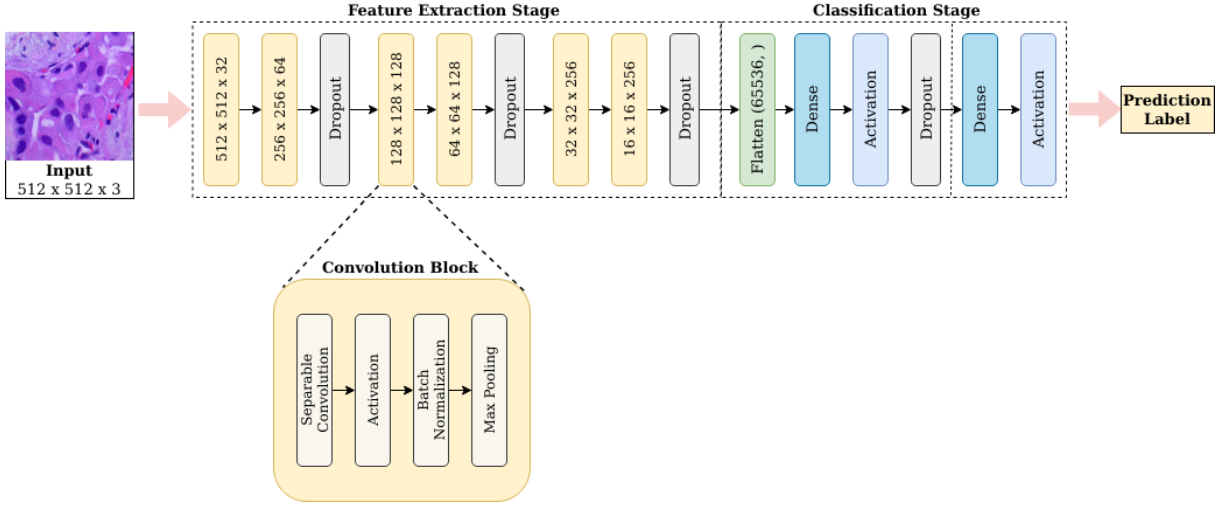


Fig. 2 The architecture of the DCNN classifiers as described in Section 3.2.1. The numbers within the boxes represent the output dimension of the respective layers. The convolution block is generically represented using an orange box and an expanded version depicts the layer structure of each convolution block.

from 32 to 256. Consequently, the convolutions extract more subtle features that can greatly aid inter-class differentiation. The feature representation vectors obtained from the convolution blocks are finally flattened using the flatten layer and fed into the classification stage. The classification stage consists of two dense blocks. The first block includes a dense layer followed by an activation layer and a dropout layer. The second block, which is the output layer, includes a dense layer followed by an activation layer.

The feature extraction layers are identical for both the classifiers — Cl_{bin} and Cl_{mul} . However, the representation vectors obtained through convolution are passed to an output layer with a single neuron in the case of Cl_{bin} to perform binary classification and to one with three neurons in the case of Cl_{mul} to perform multiclass classification.

In the design of the proposed DCNN classifier, 2D depth-wise separable convolution layers [9] were adopted for their ability to decouple the feature learning pathways between different channels of the image[24]. In contrast to regular convolutions where the spatial feature extraction and their mixing into the channels happen simultaneously, depth-wise separable convolutions employ a two-step process where the independently learned spatial features from each channel are merged later using point-wise convolutions. The WSHI patch images used for this study are 3-channel

color images. Furthermore, the H&E staining reveals significant color-based features for inter-class distinguishability. Hence, employing learning parameters that operate independently on each color channel of the image is crucial in leveraging these features for classification. In deeper layers too, learning independently from each channel and merging them in a two-step process is a more efficient use of trainable parameters than a one-step mixing of features. This is demonstrated in the Xception networks that have shown superior classification with similar number of parameters [9]. In addition, depth-wise convolutions are known to be computationally cheaper and can serve as accurate approximations of more complex regular convolution based networks [17].

The share of features learned with different convolution receptive fields was another notable design decision. The receptive fields [4, 29] were increased multiplicatively as the images progressed to deeper layers using max-pooling operations after each convolution operation, while keeping the kernel sizes fixed at 3×3 . To emphasize more on feature learning from smaller receptive fields than from larger areas of the WSHI patch, the number of channels were increased more aggressively in the initial convolution layers than the latter. This was done to better harvest information about cell atypia, nuclear shapes and spaces, which is contained in multiple small

Algorithm 1 Overview of the Proposed System**Input :** WSHI Patch (i)**Output:** Prediction labels (bc , mc)

```

//Preliminary preprocessing steps to obtain
 $i_{prep}$ 
 $i_{temp} \leftarrow i$ 
Apply random affine transformations to  $i_{temp}$ 
Resize  $i_{temp}$  to  $512 \times 512 \times 3$ 
Normalize pixel values of  $i_{temp}$  to range  $[0, 1]$ 
 $i_{prep} \leftarrow i_{temp}$ 

//Classify the entire WSHI patch
if using pipeline  $P_{dir}$  then
   $bc \leftarrow$  Classify  $i_{prep}$  using  $Cl_{bin}$ 
   $mc \leftarrow$  Classify  $i_{prep}$  using  $Cl_{mul}$ 

  //Perform nucleus segmentation first
else if using pipeline  $P_{seg}$  then
   $i_{temp} \leftarrow i_{prep}$ 
  if using threshold-based segmentation then
     $i_{binmask} \leftarrow$  Segment  $i_{temp}$  using
     $Seg_{thresh}$ 
  else if using auto-segmentation then
     $i_{binmask} \leftarrow$  Segment  $i_{temp}$  using  $Seg_{auto}$ 
  end if
   $i_{seg} \leftarrow I_{prep} \& I_{binmask}$ 

  //Classify the nucleus-segmented WSHI
  patch
   $bc \leftarrow$  Classify  $i_{seg}$  using  $Cl_{bin}$ 
   $mc \leftarrow$  Classify  $i_{seg}$  using  $Cl_{mul}$ 

end if

```

regions of the image that are captured with a smaller receptive field.

3.3 Classification using Nucleus Segmented WSHI Patches — P_{seg} Pipeline

The second proposed pipeline — P_{seg} , involves extracting nuclear regions of the input WSHI patches through semantic segmentation, before passing them to the DCNN classifier. We propose two different approaches to perform segmentation:

1. A threshold based procedural image-processing method called Seg_{thresh}

2. A deep-learning based method using a UNet network called Seg_{auto} .

3.3.1 Rationale of the Segmentation Approaches

During a histopathological examination, the biopsied tissue samples are prepared on microscope slides for visual examination, typically using methods such as chemical fixation and frozen section processing. Such slides are stained using coloring pigments to reveal and distinguish cellular components. Hematoxylin and Eosin (H&E) is the most commonly used stain combination. Hematoxylin reacts like a basic dye with specific cell components — namely the cell nucleus, ribosomes and the endoplasmic reticulum. The basic reaction stains these cell components with a purplish-blue color. On the other hand, Eosin is an acidic dye that affixes the cytoplasm, cell walls and extracellular connective matrix of the tissue with a dull reddish-pink stain. These slides can then be examined under a microscope by pathologists or micrographed to produce WSHIs.

The WSHI patch images used in this study are stained using H&E. On account of staining, the nuclear regions (as well as some other cellular components) in the histopathological images have significantly different visual properties than the remainder of the patch. The visual properties of any image pixel can be broadly categorized into two — *chroma* and *luma*. While *chroma* represents the color of a pixel, *luma* encodes the intensity level of that color in the pixel. In this method, the intensity variations are leveraged to distinguish between the nuclear and non-nuclear regions.

It is worth noting, that while Hematoxylin stains all the nuclear regions of the tissue, it stains a few other cellular components as well — namely, the ribosomes and the endoplasmic reticulum. Hence, a segmentation procedure that obtains the nuclear regions based on the intensity variations introduced by staining, will inevitably include some portions of the non-nuclear regions as well. These portions are regarded as noise. However, the intention behind the approach proposed in P_{seg} is to limit the image regions available to the classifier to contain only the most relevant information. Hence, as long as none of the nuclear regions of the image are lost to segmentation, allowing some non-nuclear regions of the image to

persist in the segmented input is not detrimental. It will not degrade the classification performance any more than passing the entire WSHI patch, as is the case in the first pipeline P_{dir} .

3.3.2 Threshold Based Segmentation (*Seg_{thresh}*)

This approach to segmenting the nuclear information from input WSHI patches directly leverages the intensity variations in the visual structure of the WSHI patches due to staining. The nature of staining leaves the pixels of the nuclear regions with a darker shade and hence a higher intensity value, relative to the remainder of the image.

First, the Otsu's method [33], a global thresholding procedure, is used to determine an optimal threshold to segregate the pixels in a given image into two classes with minimum intra-class variance and maximum inter-class differentiation. This threshold is represented as $thresh_{nuc}$.

The input WSHI patch i is first converted to grayscale to retain only the intensity information of pixels [27] and discard the color details. It is then partitioned into several sub-regions. A sub-region here, refers to the set of pixels enclosed within a bounded-region representing a subset of the image area. This is done using Suzuki's contour detection algorithm [43], which relies on the variation in intensity of pixels to achieve this. Since these variations are more pronounced at the boundary between nuclear and non-nuclear regions of the image, transitioning from a nuclear component of the image to an adjacent non-nuclear region is characterized by a sharp change in the pixel intensity values. These abrupt variations are detected by the contour detection algorithm to find closed-loop boundaries within the image that separate adjacent nuclear and non-nuclear regions. The boundaries so obtained are stored as a vector of continuous pixels forming a closed loop and can be applied on the WSHI patch to obtain sub-regions. In effect, the image is partitioned into a large number of sub-regions, some pertaining to nuclear regions and the others representing non-nuclear components of the patch.

For each sub-region in the patch image, the average intensity value is computed. The pixels that belong to sub-regions having an average intensity value higher than the threshold value $thresh_{nuc}$ are encoded with a value of 1 while

the those of the other sub-regions are encoded as 0, thereby producing a binary segmentation mask that represents the extract of nuclear information from i . This procedure is represented in Algorithm-2.

Algorithm 2 Threshold Based Segmentation

Input: Whole-Slide Histopathological Image patch i

Output: Mask to identify nuclear regions $i_{binmask}$

```

 $thresh_{nuc} \leftarrow$  Apply Otsu's method on  $i$ 
Convert  $i$  to grayscale
 $i_{binmask} \leftarrow i$ 
for each sub-region in  $i_{binmask}$  do
    if Avg-intensity sub-region  $\geq thresh_{nuc}$ 
then
        for each pixel in sub-region do
            pixel  $\leftarrow$  1
        end for
    else if Avg-intensity sub-region  $< thresh_{nuc}$ 
then
        for each pixel in sub-region do
            pixel  $\leftarrow$  0
        end for
    end if
end for

```

3.3.3 Deep-Learning Based Segmentation (*Seg_{auto}*)

In the deep learning based segmentation approach of the P_{seg} pipeline, an Xception-style UNet is trained to accept preliminary preprocessed WSHI patches i_{prep} as input and produce corresponding binary masks that represent their nuclear regions.

The UNet [28] is a Convolutional Neural Network architecture, proven for segmentation of medical images. The segmentation of nuclei information from the WSHI requires just two semantic classes — nuclear and non-nuclear regions. An initial experiment with a simple UNet architecture showed accuracy saturation and high training loss, suggestive of the degradation problem [20]. Hence, we propose an Xception-style [9] UNet architecture, to employ residual blocks that can reduce degradation in deep networks and easily adapt to learning simplex as well as complex correlations between the layers of the network [46]. Concurrent

with the standard UNet structure, the proposed architecture consists of a contractive downsampling path followed by an expansive upsampling path. The input images are first resized to 512×512 with 3 channels and passed through an entry block. This block comprises of a convolution layer with a standard kernel size of 3×3 , a batch-normalization layer and a Rectified Linear Unit (ReLU) activation layer. The block produces an output with 32 channels.

The contracting path is composed of three blocks with progressively increased filter sizes of 64, 128 and 256. The intention is to learn about the existence of distinctive finer features as the image percolates deeper into the network. Separable Convolution layers are used, which perform depth-wise spatial convolutions followed by point-wise convolutions that mix the output channels. In each block, two sets of convolution layers, each preceded by a ReLU activation layer and followed by a batch normalization layer is applied. This sequence is followed by a max-pooling layer and an addition-based skip connection from the input to the current block. The placement and ordering of the layers within each block determines the nature of activation. As discussed in [20], pre-activation and post-activation produce significantly distinct network performances in the presence of element-wise addition. Specifically, pre-activation produces better regularization by reducing overfitting due to a more direct weight propagation between subsequent blocks when representing simplex relations. Both, ReLU-only pre-activation and the full pre-activation approaches discussed in [20] were experimented with, and the former approach was found to perform better segmentation in our case. Concretely, the sequence shown in Figure ?? is adopted within each block of the downsampling strata.

The latter half of the network is an expanding path that uses transposed convolutions [14] to upsample the low resolution image at the end of the contracting path progressively, back to its original resolution. A single step of transposed convolution is achieved by performing convolutions on the image obtained by sufficiently zero-padding around each pixel of the low resolution image. In effect, the subsequent series of down-sampling and upsampling operations, recovers the spatial information lost during the contraction phase and further, localizes the identified features

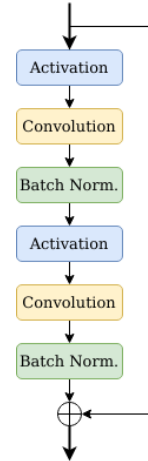


Fig. 3 ReLU-only pre-activation strategy with addition-based residual connection adopted within each block of the upsampling and downsampling strata of the UNet architecture

to their corresponding positions on the higher resolution image produced at each step of the expansion path. Ultimately, pixel-wise classification into semantic classes is performed on these localized features once the original resolution is reached. To further enhance the precision of mapping between features and their respective locations on the upsampled image and avoid losing fine-grained information, skip connections are used to forward and concatenate features from corresponding stages of the contraction path with the upsampled image obtained at the same stage of expansion path. The shorter skip connections also help to alleviate the vanishing gradient problem [13, 21].

The transpose convolution blocks used in the expansion path have a symmetric structure with respect to the contracting path. It comprised of three blocks with progressively decreased filter sizes of 256, 128 and 64. Following the symmetric structure, each block is constructed with two sets of transpose convolution layers, each followed by a ReLU activation layer and followed by a batch normalization layer. This is followed by an upsampling layer and an addition-based skip connection to forward features from the current block's input to output. Once again, the sequence of layers in each block follows the ReLU-only pre-activation strategy. Notably, the upsampling layer is different from a convolution transpose layer in that the former simply doubles the dimensions of its input while the latter performs an inverse convolution

operation and learns the kernel during the training phase. In particular, each block of the upsampling strata follows the sequence depicted in Figure ??

3.3.4 Segmentation Procedure

Both variants of the segmentation system, namely Seg_{thresh} and Seg_{auto} , identify the nuclear regions of i_{prep} by producing a binary mask image called M_{nuc} that encodes the nuclear region pixels with a class label of 1 and all the other pixels with 0. A *bitwise-AND* operation (&) is performed between the input RGB image of the WSHI patch (i_{prep}), and the corresponding mask (M_{nuc}) generated by the segmentation system. The result of this operation is an image containing only the information corresponding to the nuclear regions of the input WSHI patch as encoded in the mask and is denoted as i_{seg} . The *bitwise-AND* operation is depicted in Figure 4.

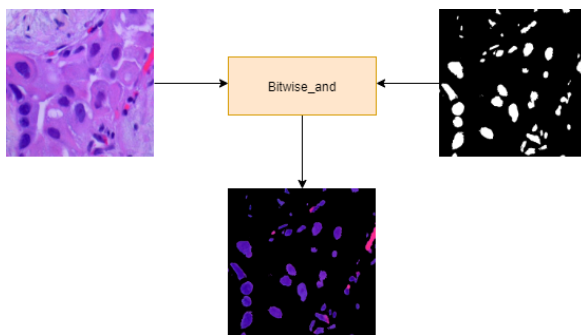


Fig. 4 Bitwise-AND operation between a binary mask and original image to obtain the corresponding segmented image.

3.3.5 Classification using Nucleus Segmented Patches

The proposed DCNN classifiers — Cl_{bin} and Cl_{mul} whose architectures are specified in Section 3.2.1 are used to perform binary and multiclass classification respectively, of the nucleus segmented WSHI patches obtained through pipeline P_{seg} .

4 Experiments and Results

4.1 Dataset

The dataset [3] consists of 15000 color histopathological images of the lung tissue. All the images are 3-channels and 768×768 pixels in size, stored in the JPEG file format. The images were generated from an original sample of Health Insurance Portability and Accountability Act (HIPAA) compliant and validated source of 750 images by augmentation. The dataset is organized into three classes with 5000 images per class. These classes represent ADC, SCC and benign lung tumor tissue, with each image of the dataset belonging to exactly one of these classes. A few sample images from the dataset are shown in Figure 5.

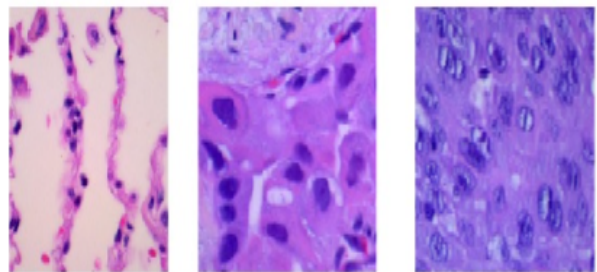


Fig. 5 WSHI Patch Images

4.2 Model Training

This paper proposes the use of deep-learning networks to classify entire or nucleus segmented WSHI patch images as well as to perform automatic semantic segmentation of WSHI patch images to retain only nuclear information. The training method and configurations are detailed in this section.

4.2.1 Binary and Multiclass Classification Networks — Cl_{bin} and Cl_{mul}

The dataset of 15000 WSHI patch images was split into training and testing set in 80:20 ratio. For the binary classifier, the images belonging to the ADC and SCC classes were together considered as malignant and were assigned a common class label.

The proposed DCNN architecture for classifiers Cl_{bin} and Cl_{mul} were trained by minimizing the equation for Categorical Cross-Entropy (CCE) loss shown in Equation 1 using the early stopping method and the Nesterov-accelerated Adaptive Moment Estimation (NAdam) optimizer with an initial learning rate of 0.01.

$$CCE = - \sum_{c=1}^C P_{true}(c) \cdot \log(P_{pred}(c)) \quad (1)$$

where C is the number of output classes of the classifier and c takes on the class label values. $P_{true}(i)$ and $P_{predicted}(i)$ represent the true and predicted probability of the input belonging to the i^{th} class, respectively.

The Rectified Linear Unit (ReLU) function was used to activate the hidden layers in both networks. Cl_{bin} used the sigmoid function for its output layer whereas, Cl_{mul} used the the softmax function. Uniform dropout rates of 0.50 and 0.25 were used in the dense and convolution layers respectively.

4.2.2 Nucleus Segmentation Network

— Seg_{auto}

To obtain ground truth data for training the segmentation neural network, 300 representative WSHI patch images with 100 images from each tumor class were segmented using the proposed threshold-based segmentation method Seg_{thresh} . The resulting nucleus segmented images were presented to three different experts to verify and make corrections. An inter-rater agreement for this corpus of 300 nucleus segmented images was computed based on regional overlap using the Generalized Conformity Index (GCI) metric [25]. GCI was estimated using Equation 2,

$$GCI = \frac{\sum_{pairs(i>j)} X_i \cap X_j}{\sum_{pairs(i>j)} X_i \cup X_j} \quad (2)$$

where X_i represents the set of nuclear region pixels of the segmented ground truth obtained from rater i and $\sum_{pairs(i>j)} \dots$ represents the summation over every unique pair of raters.

A GCI score of 0.89 was obtained across the three raters. This corpus of expert corrected segmentation labels was used as the ground truth

dataset to train the Seg_{auto} network. It was split into training and testing sets in 80:20 ratio. The proposed Xception-style UNet architecture for Seg_{auto} was trained by minimizing the equation for dice-loss shown in Equation 3 using the early stopping method and the NAdam optimizer. The sigmoid function was used to activate the output layer and the ReLU function was used for the hidden layers.

$$DiceLoss = 1 - \frac{1}{K} \sum_{k=1}^K \frac{2 \times P_k \cap T_K}{P_i + T_j} \quad (3)$$

where K represents the number of semantic classes. P_k and T_k represent the set of pixels belonging to the k^{th} class of the prediction and ground truth images respectively. $\|\cdot\|$ denotes the cardinality operation of its argument.

4.3 Experiments

Three experiments were conducted with the following goals, using the aforementioned dataset to evaluate, analyze and compare the performance of each variant of our proposed system.

1. To ascertain the performance of the DCNN based classifiers Cl_{bin} and Cl_{mul} using the entire WSHI patch image as input
2. To ascertain the performance of the nucleus segmentation network Seg_{nuc}
3. To ascertain the performance of the DCNN based classifiers Cl_{bin} and Cl_{mul} using the following inputs:
 - (a) WSHI patch image with nuclear regions segmented by the thresholding technique Seg_{thresh}
 - (b) WSHI patch image with nuclear regions segmented using the UNet Segmentation Network approach Seg_{auto}

The metrics Precision (P), Recall (R), Accuracy (Acc) and F1-Score (F1) [5] are used to ascertain the performance of the DCNN classifiers Cl_{bin} and Cl_{mul} with both, complete WSHI image inputs as well as nucleus segmented inputs. These are specified in equations 4, 5, 6 and 7.

$$P = \frac{TP}{TP+FP} \quad (4)$$

$$R = \frac{TP}{TP+FN} \quad (5)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (6)$$

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

In these equations, TP, FP, TN and FN represent the sets of true-positive, false-positive, true-negative and false-negative classifications respectively. \cdot denotes the cardinality operation of its argument.

The segmentation architecture *Seg_{nuc}* is evaluated using the Mean Intersection over Union (Mean-IoU) metric by comparing the predicted mask with the ground truth. This is specified in equation 8.

$$MeanIoU(X, Y) = \frac{1}{K} \sum_{k=1}^K \frac{X_k \cap Y_k}{X_k \cup Y_k} \quad (8)$$

where, K represents the number of semantic classes, X_k and Y_k represent the pixels belonging to the k^{th} class in images X and Y respectively.

4.4 Performance of the classifiers with entire WSHI patches

The aim of this experiment is to validate the performance of the two classifiers — Cl_{bin} and Cl_{mul} when the entire WSHI patch is fed to the network. Cl_{bin} is a DCNN based binary classifier that segregates its inputs into one of two classes — benign and malignant. Cl_{mul} is a multiclass classifier that segregates its inputs into one of three classes — ADC, SCC and benign. The confusion matrices for binary classification and multiclass classification are shown in Tables 1 and 2 respectively.

The binary classification achieves an accuracy of 95.70% and a macro-averaged F1-Score of 95%.

The multiclass classification achieves an accuracy of 94.97% and a macro-averaged F1-Score of 94.97%.

The proposed DCNN architecture predicts benign images with very good accuracy but tends to misclassify a few of the malignant images as benign. Furthermore, in the multiclass case, the

Table 1 Confusion matrix for binary classification from entire WSHI patches

	Types	Predicted	
		Benign	Malignant
Actual	Benign	1000	0
	Malignant	86	914

Table 2 Confusion matrix for multiclass classification from entire WSHI patches

	Types	Predicted		
		Benign	ADC	SCC
Actual	Benign	992	8	0
	ADC	0	1000	0
	SCC	143	0	857

ADC class is identified well but there is misclassification between benign and ADC as well as benign and SCC. This suggests that the binary classifier system is basing its decision of tumor cells being benign on features that are, in reality, not relevant to the identification of tumor type. As a result, the system has labeled some of the benign tumor cells as ADC and some of the SCC tumor cells as benign. Extraction of such misleading features may arise due to presence of non-nuclear information in the input images, particularly the regions that are not diagnostically contributing.

4.5 Performance of the *Seg_{nuc}* Nucleus Segmentation Network

The intent of this experiment is to validate the semantic segmentation performed by the UNet architecture. The UNet architecture is trained using the ground truth prepared by following the methods detailed in section-3.3. A corpus of 300 representative ground truth images was prepared through experts review and correction of the segmented images produced by the *Seg_{thresh}* technique. An inter-rater CGI of 0.89 was obtained across three experts. This corpus was split into

Table 3 Performance of the UNet Network Seg_{nuc} for Nuclear Segmentation

Class	Mean-IoU (%)			Std. Dev.
	Avg	Max	Min	
Benign	94.98	95.40	94.62	0.0022
ADC	94.84	95.16	94.60	0.0013
SCC	93.39	93.65	92.86	0.0023
Overall	94.27	95.40	92.86	0.0051

training and testing sets in 80:20 ratio. The network was trained using the training set and evaluated on the test set. The system was evaluated using the Mean-Iou specified in equation 7.

It achieved a Mean-IoU of 94.27%, averaged over all semantic classes and images in the test set. The minimum Mean-IoU was 92.86%, maximum was 95.40% and the standard deviation was 0.0051. Table 3 represents a comprehensive summary of the system's segmentation performance. The low value of overall standard deviation (0.51%) is indicative of consistency in the predictions and of the fact that the mean of the evaluation metric is a good representation of the model's performance across the testing set. The performance metrics of the segmentation network are also separately presented for each of the three classes and it is evident that the segmentation performance is marginally lower for the SCC class with an average Mean-IoU of 93.39%.

4.6 Performance of the classifiers with nuclei segmented by thresholding

The aim of this experiment is to validate the performance of the two classifiers — Cl_{bin} and Cl_{mul} when the only the nuclear regions of the WSHI patch are segmented and retained using the proposed thresholding technique and then fed to the network. Hematoxylin stains the nuclei in a dark blue color hence segmentation of nucleus based on the thresholding method helps to segment all of the nucleus from the input images. The confusion matrices for binary classification and multiclass classification for this input are shown in Tables 4 and 5 respectively.

The binary classification achieves an accuracy of 98.10% and a macro-averaged F1-Score of 98.09%.

Table 4 Confusion matrix for binary classification from nuclei segmented by thresholding

	Types	Predicted	
		Benign	Malignant
Actual	Benign	998	2
	Malignant	36	964

Table 5 Confusion matrix for multiclass classification with nuclei segmented by thresholding

	Types	Predicted		
		Benign	ADC	SCC
Actual	Benign	981	18	1
	ADC	0	984	16
	SCC	0	17	983

The multiclass classification achieves an accuracy of 98.27% and a macro-averaged F1-Score of 98.26%.

Isolating the nuclear regions from the WSHI patch images before classification has shown significant improvements in the classification performance. The malignant classification is identified with lesser errors by Cl_{bin} and the number of misclassifications of the SCC class by Cl_{mul} has substantially reduced.

However, there is increased misclassification within the malignant subtypes — ADC and SCC. This can be interpreted as a difficulty in the visual distinguishability between ADC and SCC lung tumors — a result consistent with clinical observations [6, 7]. Often, stains other than H&E may be required to reach a conclusive diagnosis in distinguishing between the two subtypes [18].

4.7 Performance of the classifiers with nuclei auto-segmented by the Seg_{nuc} network

The aim of this experiment is to validate the performance of the two classifiers — Cl_{bin} and Cl_{mul} when only the nuclear regions of the WSHI patch are segmented using deep learning technique with the help of UNet architecture. The UNet network

learns automatically all the important features necessary to segment the nucleus from the WSHI. The confusion matrix for binary classification and multiclass classification are shown in Tables 6 and 7 respectively.

Table 6 Confusion matrix for binary classification from nuclei auto-segmented using the Seg_{nuc} network

Types		Predicted	
		Benign	Malignant
Actual	Benign	1000	0
	Malignant	7	993

The binary classification achieves an accuracy of 99.65% and a macro-averaged F1-Score of 99.64%.

Table 7 Confusion matrix for multiclass classification with nuclei auto-segmented using the Seg_{nuc} network

Types		Predicted		
		Benign	ADC	SCC
Actual	Benign	997	2	1
	ADC	0	998	2
	SCC	0	6	994

The multiclass classification achieves an accuracy of 99.63% and a macro-averaged F1-Score of 99.63%.

The classification performance is superior when the nuclear regions are segmented using the segmentation network Seg_{nuc} . It is evident that the above thresholding method considers only the color features to segment the nucleus but the UNet architecture considers multiple features in addition to the color, for segmentation. Both, the misclassifications made by Cl_{bin} in predicting the malignant class and by Cl_{mul} in predicting the SCC class has further reduced. Moreover, the number of misclassifications between SCC and ADC classes has reduced in comparison to the thresholding approach.

The thresholding technique used in the previous case was purely algorithmic in that, it did

not involve any learning. The segmentation was completely based upon the intensity variations introduced by staining with H&E. However, the segmentation network trained using the ground truth prepared by the thresholding technique, later corrected and verified by experts, produces better classification results. This indicates that the deep-learning based segmentation network is able to better identify nuclear regions, by automatically learning and basing its segmentation decisions on multiple features that contribute to the visual distinguishability of nuclear regions from non-nuclear regions of the WSHI patch images.

However, relative to other classes, the SCC class is still characterized by a slightly higher misclassification errors. This observation can be linked to the slightly reduced average, maximum and minimum mean-IoUs of the segmentation network for SCC lung tissue images, as specified in Table 3. This reduced segmentation performance is likely to be contributing to the lower classification performance for this class.

4.8 Overall performance of the proposed system

The performance of the segmentation network Seg_{nuc} is summarized in Table 3 and was discussed in section 4.7. It achieves an overall Mean-IoU%, averaged over all classes of inputs.

The overall performance of the system was analyzed by comparing the accuracy along with class-wise precision, recall and F1-score of the classifiers. These results are summarized for each type of input supplied to the classifiers, in Tables 9 and 10. There is a notable improvement in all metrics, progressing from entire WSHI patch input to nuclei segmented by thresholding and finally the auto-segmented nuclei using the UNet based Seg_{nuc} . This suggests that supplying only the nuclear information segmented out of the WSHI patch inputs to the classifiers has a contributing impact on the classification performance. Furthermore, the even better classification obtained using auto-segmented nuclei from Seg_{nuc} shows that the deep-learning based segmentation framework is far more effective in identifying nuclear regions correctly, when compared to the thresholding based algorithmic approach. It is worth noting that, in both the segmentation-based approaches (sections ?? and ??) Cl_{mul} does not misclassify any of the

malignant types as benign, despite mislabelling between the ADC and SCC classes. In the case of Cl_{bin} , the number of misclassifications between benign and malignant classes has substantially reduced in the auto-segmentation case when compared to the threshold based system (reduced from 36 to 7).

We have compared the results of our proposed system for multiclass classification using auto-segmented nuclear regions as inputs with other existing methods that have used the same dataset [3] for lung cancer subtyping. Our method outperforms them along all metric measures, indicating the efficacy of segmenting the nuclear regions before classification. This comparison is presented in Table 8.

Article	Acc	P	R	F1
S. Mangal et al. [30]	97.90	-	-	-
B. K. Hatuwal et al. [19]	97.20	97.33	97.33	97.33
M. Nishio et al. [32]	99.44	99.43	99.43	99.43
Proposed Method	99.63	99.64	99.63	99.63

Table 8 Comparison with existing multiclass classification networks

References

- [1] Histopathological study of lung biopsy in association with immunohistochemistry.
- [2] Mahmoud Abdolhoseini, Murielle G Kluge, Frederick R Walker, and Sarah J Johnson. Segmentation of heavily clustered nuclei from histopathological images. *Scientific reports*, 9(1):4551, March 2019.
- [3] L. Brannon Thomas Catherine P. Wilson Lauren A. DeLand Stephen M. Mastorides Andrew A. Borkowski, Marilyn M. Bui. Lc25000 lung and colon histopathological image dataset.
- [4] André Araujo, Wade Norris, and Jack Sim. Computing receptive fields of convolutional neural networks. *Distill*, 2019. <https://distill.pub/2019/computing-receptive-fields>.
- [5] Alireza Baratloo, Mostafa Hosseini, Ahmed Negida, and Gehad El Ashal. Part 1: simple definition and calculation of accuracy, sensitivity and specificity. 2015.
- [6] Justin A Bishop, Julie Teruya-Feldstein, William H Westra, Giuseppe Pelosi, William D Travis, and Natasha Rekhtman. p40 (δ np63) is superior to p63 for the diagnosis of pulmonary squamous cell carcinoma. *Modern pathology*, 25(3):405–415, 2012.
- [7] Lukas Bubendorf, Sylvie Lantuejoul, Adrianus J de Langen, and Erik Thunnissen. Nonsmall cell lung carcinoma: Diagnostic difficulties in small biopsies and cytological specimens: Number 2 in the series “pathology for the clinician” edited by peter dorfmüller and alberto cavazza. *European Respiratory Review*, 26(144), 2017.
- [8] Zhiming Cheng and Aiping Qu. A fast and accurate algorithm for nuclei instance segmentation in microscopy images. *IEEE Access*, 8:158679–158689, 2020.
- [9] François Chollet. Xception: Deep learning with depthwise separable convolutions, 2017.
- [10] Nicolas Coudray, Paolo Santiago Ocampo, Theodore Sakellaropoulos, Navneet Narula, Matija Snuderl, David Fenyö, Andre L Moreira, Narges Razavian, and Aristotelis Tsirigos. p3 classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature medicine*, 24(10):1559–1567, 2018.
- [11] Yuxin Cui, Guiying Zhang, Zhonghao Liu, Zheng Xiong, and Jianjun Hu. A deep learning algorithm for one-step contour aware nuclei segmentation of histopathology images. *Medical & biological engineering & computing*, 57(9):2027–2043, 2019.
- [12] Sumaiya Dabeer, Maha Mohammed Khan, and Saiful Islam. Cancer diagnosis in histopathological image: Cnn based approach. *Informatics in Medicine Unlocked*, 16:100231, 2019.

Table 9 Performance summary of binary classification DCNN

Image	Acc(%)	F1 Score(%)		Precision(%)		Recall(%)	
		Benign	Malignant	Benign	Malignant	Benign	Malignant
Entire WSHI Patch	95.70	95.88	95.51	92.08	100	100	91.40
Threshold	98.10	98.13	98.07	96.52	99.79	99.80	96.40
UNet	99.65%	99.64%	99.64	99.31	100	100	99.30

Table 10 Performance summary of multiclass classification DCNN

Image	Acc(%)	F1 Score(%)			Precision(%)			Recall(%)		
		Benign	ADC	SCC	Benign	ADC	SCC	Benign	ADC	SCC
Entire WSHI Patch	94.97%	92.93	99.60	92.30	87.40	99.20	100	99.20	100	85.70
Threshold	98.27%	99.04	97.47	98.30	100	96.57	98.30	98.10	98.40	98.30
UNet	99.63%	99.85	99.50	99.55	100	99.20	99.70	99.70	99.80	99.40

- [13] Michal Drozdal, Eugene Vorontsov, Gabriel Chartrand, Samuel Kadoury, and Chris Pal. The importance of skip connections in biomedical image segmentation, 2016.
- [14] Vincent Dumoulin and Francesco Visin. A guide to convolution arithmetic for deep learning, 2018.
- [15] Baris Gecer, Selim Aksoy, Ezgi Mercan, Linda G. Shapiro, Donald L. Weaver, and Joann G. Elmore. Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. *Pattern Recognition*, 84:345–356, 2018.
- [16] Simon Graham, Muhammad Shaban, Talha Qaiser, Navid Alemi Koohbanani, Syed Ali Khurram, and Nasir Rajpoot. Classification of lung cancer histology images using patch-level summary statistics. In *Medical Imaging 2018: Digital Pathology*, volume 10581, page 1058119. International Society for Optics and Photonics, 2018.
- [17] Jianbo Guo, Yuxi Li, Weiyao Lin, Yurong Chen, and Jianguo Li. Network decoupling: From regular to depthwise separable convolutions, 2018.
- [18] Grzegorz T Gurda, Lei Zhang, Yuting Wang, Li Chen, Susan Geddes, William C Cho, Frederic Askin, Edward Gabrielson, and Qing Kay Li. Utility of five commonly used immunohistochemical markers ttf-1, napsin a, ck7, ck5/6 and p63 in primary and metastatic adenocarcinoma and squamous cell carcinoma of the lung: a retrospective study of 246 fine needle aspiration cases. *Clinical and translational medicine*, 4(1):1–13, 2015.
- [19] Bijaya Hatuwal and Himal Thapa. Lung cancer detection using convolutional neural network on histopathological images. *International Journal of Computer Trends and Technology*, 68:21–24, 10 2020.
- [20] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks, 2016.
- [21] Sepp Hochreiter. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6:107–116, 04 1998.
- [22] Hwejin Jung, Bilal Lodhi, and Jaewoo Kang. An automatic nuclei segmentation method based on deep convolutional neural networks for histopathology images. *BMC Biomedical*

Engineering, 1, 2019.

- [23] Hwejin Jung, Bilal Lodhi, and Jaewoo Kang. An automatic nuclei segmentation method based on deep convolutional neural networks for histopathology images. *BMC Biomedical Engineering*, 1(1):1–12, 2019.
- [24] Lukasz Kaiser, Aidan N. Gomez, and Francois Chollet. Depthwise separable convolutions for neural machine translation, 2017.
- [25] Erik Kouwenhoven, Marina Giezen, and Henk Struikmans. Measuring the similarity of target volume delineations independent of the number of observers. *Physics in Medicine & Biology*, 54(9):2863, 2009.
- [26] Hansang Lee and Junmo Kim. Segmentation of overlapping cervical cells in microscopic images with superpixel partitioning and cell-wise contour refinement. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 63–69, 2016.
- [27] Hui Liu. Chapter 3 - rail transit collaborative robot systems. In Hui Liu, editor, *Robot Systems for Rail Transit Applications*, pages 89–141. Elsevier, 2020.
- [28] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation, 2015.
- [29] Wenjie Luo, Yujia Li, Raquel Urtasun, and Richard Zemel. Understanding the effective receptive field in deep convolutional neural networks, 2017.
- [30] Sanidhya Mangal, Aanchal Chaurasia, and Ayush Khajanchi. Convolution neural networks for diagnosing colon and lung cancer histopathological images. *arXiv preprint arXiv:2009.03878*, 2020.
- [31] "Lad N and Daveshwar M". "histopathological study of lung biopsy in association with immunohistochemistry". *"J. Evolution Med.Dent. Sci"*, "8"("48"), "2019".
- [32] Mizuho Nishio, Mari Nishio, Naoe Jimbo, and Kazuaki Nakane. Homology-based image processing for automatic classification of histopathological images of lung tissue. *Cancers*, 13(6):1192, 2021.
- [33] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.
- [34] Hady Ahmady Phoulady, Dmitry B Goldgof, Lawrence O Hall, and Peter R Mouton. Nucleus segmentation in histology images with hierarchical multilevel thresholding. In *Medical Imaging 2016: Digital Pathology*, volume 9791, page 979111. International Society for Optics and Photonics, 2016.
- [35] Archana Rastogi. Changing role of histopathology in the diagnosis and management of hepatocellular carcinoma. *World Journal of Gastroenterology*, 24(35):4000–4013, 2018.
- [36] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.
- [37] P Mohamed Shakeel, MA Burhanuddin, and Mohammad Ishak Desa. Automatic lung cancer detection from ct image using improved deep neural network and ensemble classifier. *Neural Computing and Applications*, pages 1–14, 2020.
- [38] Shallu Sharma and Rajesh Mehra. Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images—a comparative insight. *Journal of digital imaging*, 33(3):632–654, 2020.
- [39] KK Shukla, Anoop Tiwari, Shiru Sharma, et al. Classification of histopathological images of breast cancerous and non cancerous cells based on morphological features. *Biomedical and Pharmacology Journal*, 10(1):353–366, 2017.

- [40] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2015.
- [41] Changhao Sun, Chen Li, and Xiaoyan Li. Gastric histopathology image segmentation using a hierarchical conditional random field. *CoRR*, abs/2003.01302, 2020.
- [42] Supriya Suresh and Subaji Mohan. Roi-based feature learning for efficient true positive prediction using convolutional neural network for lung cancer diagnosis. *Neural Computing and Applications*, 32(20):15989–16009, 2020.
- [43] Satoshi Suzuki and Keiichi Abe. Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, 30(1):32–46, 1985.
- [44] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.
- [45] Rajyalakshmi Uppada, S. K. Rao, and K. Prasad. Supervised classification of breast cancer malignancy using integrated modified marker controlled watershed approach. *2017 IEEE 7th International Advance Computing Conference (IACC)*, pages 584–589, 2017.
- [46] Andreas Veit, Michael Wilber, and Serge Belongie. Residual networks behave like ensembles of relatively shallow networks. *arXiv preprint arXiv:1605.06431*, 2016.
- [47] Quoc Dang Vu, Simon Graham, Tahsin Kurc, Minh Nguyen Nhat To, Muhammad Shaban, Talha Qaiser, Navid Alemi Koohbanani, Syed Ali Khurram, Jayashree Kalpathy-Cramer, Tianhao Zhao, Rajarsi Gupta, Jin Tae Kwak, Nasir Rajpoot, Joel Saltz, and Keyvan Farahani. Methods for segmentation and classification of digital microscopy tissue images. *Frontiers in Bioengineering and Biotechnology*, 7:53, 2019.
- [48] Khin Yadanar Win, Somsak Choomchuay, Kazuhiko Hamamoto, and Manasan Raveesunthornkiat. Comparative study on automated cell nuclei segmentation methods for cytology pleural effusion images. *Journal of healthcare engineering*, 2018, 2018.
- [49] Jun Xu, Lei Gong, Guanhao Wang, Cheng Lu, Hannah Gilmore, Shaoting Zhang, and Anant Madabhushi. Convolutional neural network initialized active contour model with adaptive ellipse fitting for nuclear segmentation on breast histopathological images. *Journal of medical imaging (Bellingham, Wash.)*, 6(1):017501, January 2019.
- [50] Kun-Hsing Yu, Ce Zhang, Gerald J Berry, Russ B Altman, Christopher Ré, Daniel L Rubin, and Michael Snyder. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nature Communications*, 7(1):12474, 2016.