ORIGINAL ARTICLE

A deep learning approach for nucleus segmentation and tumor classification from lung histopathological images

S. M. Jaisakthi[1] · Karthik Desingu[2] · P. Mirunalini[2] · S. Pavya[2] · N. Priyadharshini[2]

Abstract

Lung cancer is the leading cause of death worldwide. Early diagnosis is crucial to improve patients' chance of survival. Automated detection and analysis of cancer types can significantly improve the diagnosis process. It can aid treatment through follow-up analyses. This paper proposes a deep learning based pipeline for multi-class classification of lung tumor type (Benign (B), ADenoCarcinoma (ADC) and Squamous-Cell Carcinoma (SCC)) from histopathological images. A baseline classification method, the P_{dir} pipeline, is proposed where Whole Slide Histopathological Image (WSHI) patches are classified using the proposed Deep Convolutional Neural Network (DCNN) classifier. Since each cancer type is characterized by the difference in the structure of the nuceli, this research work proposes to improve the performance of classification by segmenting the nuclei. The P_{seg} pipeline is proposed to extract the nuclear regions from the WSHI patches using an Xception-style UNet based neural network, and this segmented region is then categorised into tumor types using the same downstream DCNN architecture. The classification system showed an accuracy of 95.40% and 99.60% using the P_{dir} and P_{seg} pipelines, respectively. The classification performed through P_{seg} pipeline exhibits significant improvement compared to the P_{dir} pipeline, supporting our hypothesis that nucleus segmentation improves classification performance. This paper posits that segmenting and retaining the nuclear regions in the input image to the tumor type classifier suppresses the importance of less relevant portions of the image during model training, pronounces nuclear region boundaries to highlight shape features, and reduces the overall computation cost of training. Ultimately, it induces a positive impact on classification performance. The enhanced performance obtained using the proposed P_{seg} pipeline supports our hypothesis.

✉ P. Mirunalini
miruna@ssn.edu.in

S. M. Jaisakthi
jaisakthi.murugaiyan@vit.ac.in

Karthik Desingu
karthik19047@cse.ssn.edu.in

S. Pavya
pavya17102@cse.ssn.edu.in

N. Priyadharshini
priyadharshini17118@cse.ssn.edu.in

[1] School of Computer Science and Engineering, Vellore Institute of Technology, Chennai Campus, Chennai, Tamilnadu 600127, India

[2] Department of Computer Science and Engineering, Sri Sivasubramaniya Nadar College of Engineering, , Chennai, Tamilnadu 603110, India

1 Introduction

Cancer causes high mortality worldwide and early diagnosis improves the chance of successful treatment. The human body consists of trillions of cells, and these cells grow and multiply to form new cells. However, in abnormal cases, cells may grow and multiply uncontrollably and result in the development of tumors. Tumors can either be cancerous (malignant) or non-cancerous (benign) and can occur in any organ of a human body, such as lung, breast, colon, and ovaries. Within the lungs, cancer cell may affect various parts such as bronchi, 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 (Nirali and Daveshwar 2019). Other less pronounced causes may

include gene changes or workplace exposure to asbestos, diesel exhaust and certain other chemicals. Early detection of tumors is curable and hence it is necessary to identify tumors early on.

Manual diagnosis of tumors by the experts is a time consuming process, and accurate categorization is challenging. This manual diagnosis can be automated specifically by learning the features from images using deep learning techniques which will reduce the efforts of human specialist without compromising accuracy. A variety of imaging modalities are available to detect lung tumors among which Computed Tomography (CT) scan, Magnetic Resonance Imaging (MRI) scan and histopathological imaging are predominantly used. Among them, the histopathological imaging procedure is considered a gold-standard in cancer diagnosis (Lad and Daveshwar 2019) as it gives a more detailed view of the disease.

In histopathological imaging technique, tissues are stained with Hematoxylin and Eosin (H&E) and the stained images are further examined under a microscope to detect malignant features in the cellular structures. It helps in distinguishing between lesions not only during the diagnostic phase, but also during follow-ups and it further helps in the prognostication and treatment decisions (Rastogi 2018). Hence, the analysis of histopathological images is invaluable in the quantitative analysis of clinical variables based on biologically relevant features.

1.1 Motivation

The tumor cells in an histopathological image exhibit high amounts of irregularity and arbitrary visual angles. Hence, categorizing histopathological images manually is subject to inter and intra observer variability (Dabeer et al. 2019), and is a time consuming process. The ability of deep learning can be leveraged to improve the diagnostic precision of computational pathology.

The stained tissues in the histology images contain nuclei and various other structures. The morphological pattern of a nucleus varies widely due to several factors, including the malignancy of the disease, nucleus type and nucleus life cycle. The nucleus can provide key information for identifying the presence or the stage of a disease, but segmenting nuclear regions from the WSHI patches is a challenging task due to extensive morphological differences.

Contrary to widespread methods to categorize tumor types that use entire WSHI patch images, passing nucleus segmented WSHI patches to the DCNN classifier will improve the classifier's precision. This is because the effort of the learning models will be reduced considerably by suppressing the weights of less relevant features stemming from other regions of the image. 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 are learned faster with more precision. 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 system established using the classification performed with the whole-slide histopathological images.

In this paper, a deep learning based method has been proposed to perform multiclass classification of lung tumors using DCNN from WSHI patches of lung tissues, as well as from the nucleus segmented WSHI patches. These are presented as baseline P_{dir} pipeline and the proposed P_{seg} pipeline, respectively. The research work then compares the two approaches to prove the positive impact of segmenting the nuclear regions on classification performance. A Python based implementation for this research work will be made available at: https://github.com/karthik-d/Lung-Tumor-Classification.

2 Related work

Nuclei segmentation is challenging for two key reasons: (i) color variation in histopathological images and (ii) difference in morphological structures. Several image processing techniques have been proposed in literature to segment nuclei from histopathological images such as the watershed method (Yang et al. 2006), multi-level thresholding followed by watershed algorithm (Abdolhoseini et al. 2019), hybrid segmentation using k-Means clustering and adaptive thresholding (Filipczuk et al. 2011). Multiscale and multimarker approach (Cui et al. 2019), graph-cuts approach (Veta et al. 2013; Lee and Kim 2016), conventional thresholding and image processing techniques (Yu et al. 2016) and morphological features (Shukla et al. 2017). All the image-processing based approaches rely on parameter based approach and are not capable of handling changes in the stains and morphological structures.

During the past decades Machine Learning (ML) based approaches has gained focus and many applied these techniques for nucleus segmentation. ML models are usually trained using handcrafted features such as color, texture, Laplacian of Gaussian response, local binary patterns of the nuclei, hough-transform, histogram of oriented gradients (HOG) and marker controlled watershed approach (Kong et al. 2011; Zhang et al. 2015; Qureshi et al. 2008; Sertel et al. 2010; Uppada et al. 2017; Hussein et al. 2022; Phoulady et al. 2016) and the learning algorithm uses either supervised or unsupervised technique.

These methods are limited by feature engineering, especially the unsupervised learning methodology produce under-segmentation results if the nuclear region has significant variation in color and texture (Kong et al. 2011). 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 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 are not perceptible. The applications of deep learning models and computer vision are growing exponentially in recent years. With the help of deep learning architectures like CNN, U-Net, ResNet, Masked RCNN, etc., better results can be achieved for complex biomedical tasks such as segmentation and classification. A multi-scale deep residual aggregation network (Vu et al. 2019) was used to segment the nuclear material and to detect nuclear locations. Feature Pyramid Network (FPN) (Cheng and Qu 2020) followed by U-Net architecture was used to separate the clustered nuclei. Nuclei contours were identified using deep learning models (Xing et al. 2016) and segmentation was performed using iterative region growing algorithm.

Spatially Constrained convolutional Neural Network was used by Sirinukunwattana et al. (2016) to detect the nucleus. Multi scale convolutional network (MSCN) and graph-partitioning-based method was used for accurate segmentation of cervical cytoplasm and nuclei (Lee and Kim 2016). Detection and segmentation of breast cancer nuclei have been performed using a Convolution Neural Network (CNN) initialized with an active contour model followed by adaptive ellipse fitting procedure to remove overlapping clumped nuclear regions (Xu et al. 2019). Colour normalization using a Convolutional Gaussian Mixture model followed by segmentation was proposed using Mask Residual Convolutional Neural Network (Mask R-CNN) (Jung et al. 2019). A comparative study of twelve nuclei segmentation methods for cytology pleura effusion images was presented and analysed (Win et al. 2018). (Kalinathan et al. 2018) Kalinathan L et al. implemented a method for segmentation of nuclei in hepatic tumors from histopathological images, where image pixels are textonized, clustered and used to train a boosting classifier that assigns a semantic class to each texton cluster. A nucleus-boundary model to predict the nuclear regions and their corresponding boundaries using a fully Convolutional Neural Network was proposed by Cui et al. (2019) after performing the segmentation steps using multilevel thresholding and series of morphological operators. Although there is a significant research work carried out using DL algorithms, there are gaps that need to be addressed.

Nucleus segmentation from histopathological images attracts research attention for the following reasons:

1. The color variation in histopathological images makes the segmentation process challenging (Jung et al. 2019; Mehmood et al. 2022).
2. Annotating all the nuclei in the histopathological images is difficult (Jung et al. 2019).
3. The varying sizes and shapes of the nuclei results in inaccurate annotations (Kalinathan et al. 2018, 2020).
4. The presence of a large number of clustered nuclei results in poor segmentation accuracy (Abdolhoseini et al. 2019).

It is evident that there is an increase in use of CNN methods in medical image analysis, due to its capability to extract increasingly complex features as the image percolates to deeper layers of the network and its ability to learn effective features to reduce the false positives. Bilinear-CNN-based model combines a bilinear convolutional module and soft attention module to perform multi tissue classification of histopathological images in lung cancer (Xu et al. 2022). A hybrid classification model comprising an inception_v3 network and a hog-and-daisy feature extraction module to classify lung cancer and normal tissue from lung pathology was proposed. The extracted features are fed into a 3-layered Fully Connected Network and the softmax layer has been used for final classification (Chen et al. 2021). Sharma and Mehra (2020) analysed both machine learning and transfer learning based approaches for breast cancer classification and concluded that transfer learning methods perform superior than machine learning approaches. Hussein et al. (2021) proposed an automated method for segmentation of ovarian and breast ultrasound images using viola-james model which improves classification accuracy. Suresh and Mohan (2020) proposed an improved system for lung cancer diagnosis using GAN generated images that uses self-learned features. Shakeel et al. (2020) used a DL network to extract number of features and the reduced feature set obtained using dimensionality reduction was boosted using an ensemble classifier to achieve high classification accuracy. Gram et al. proposed an automated classification of ADC and SCC (Graham et al. 2018), where a deep learning framework classifies the input patches as diagnostic or non-diagnostic, and then, a random forest regression model is applied on the diagnostic patches to classify each WSHI as either ADC or SCC based on statistical and morphological measurements extracted from the input WSHI. Gecer et al. used a WSHI classification pipeline for categorizing five diagnostic type of breast cancer using deep learning methods (Gecer et al. 2018). In Coudray et al. (2018), the Inception-V3 architecture was used to classify WSHI into one of three classes namely, ADC, SCC and Benign tumor cells. Medeiros et al.

(2020) extracted features from LIDC/IDRI nodule images using transfer learning models such as VGG16, VGG19, MobileNet, Xception, InceptionV3, ResNet50, Inception-ResNetV2, DenseNet169, DenseNet201, NASNetMobile and NASNetLarge. These deep features extracted were submitted to 10-fold cross-validations with naive Bayes, multi-layer perceptron, support vector machine (SVM), K-nearest neighbors KNN and random forest classifiers to analyse the results. Masud et al. extracted (Masud et al. 2021) two types of features using domain transformations from lung cancer images and acquired features trained using CNN. A simple deep learning based CNN architecture used to classify lung and colon cancer (Mangal et al. 2020). A pretrianed neural network model using AlexNet was tuned by modifying the four layers and also the authors achieved promising classification results.

From literature, it is evident that many algorithms for cancer localization and detection involves increased computational cost and time (Sirinukunwattana et al. 2016; Sun et al. 2017). Similarly, using multiple datasets with varying scan settings, results in poor accuracy (Masood et al. 2018). Likewise, authors who used less number of classes or a compact dataset, also attain low accuracies (Sun et al. 2017; Filho et al. 2018; Suresh and Mohan 2020).

2.1 Major contributions

To overcome these limitations, this paper makes the following key contributions,

- Most methods in literature focus on processing WSHI patches directly for classification. However, this work proposes to extract the nuclei from the WHSI patches for classification, and proves its positive impact on classification performance.
- Stain color normalization is performed using a deconvolution approach, which plays a crucial role in normalizing the effect of H &E staining across all images of the dataset.
- A segmentation network based on the Xception-style (Chollet 2017) UNet (Ronneberger et al. 2015) is proposed to segment the nuclear regions from Whole Slide Histopathological Image (WSHI) patches, and a DCNN based architecture is proposed to act as the downstream multiclass classifier for lung tumor subtyping.
- The dataset (Borkowski et al. 2023) used in this study is primarily used for classification. It does not contain ground truth for segmentation. Therefore, to train the proposed segmentation network, the ground truth is developed by creating binary masks that demarcate nuclear and non-nuclear regions, and validated with the help of medical experts.

- Comprehensive experiments were conducted on the dataset (Borkowski et al. 2023) comprising histopathological images. To ascertain the consistency and to avoid the generalization error of the designed architecture, we have carried out stratified k-fold cross-validation during model training.

3 Proposed method

In this paper, a DCNN based multiclass classification of lung tumor WSHI patches into one of three classes namely: ADC, SCC, and B is proposed. To improve the classification accuracy of the proposed DCNN, stain color normalization of the WSHI patch images and semantic segmentation using Xception based UNet to retain only the nuclear regions of the input was performed. Further, the difference in performance of the classifier when it is trained with segmented nuclear regions and also with entire WSHI patches is analyzed and presented. The model training process is validated using stratified k-fold cross-validation to evaluate the model by involving all images of the dataset in both training and validation sets. The workflow of the proposed system is depicted in Fig. 1 and also outlined in Algorithm 1.

Algorithm 1 Overview of the Proposed System

Input : WSHI Patch (i)
Output: Predicted tumor type label (mc)

  //**Preprocessing to obtain** i_{prep}
  $i_{temp} \leftarrow i$
  Apply random affine transformations to i_{temp}
  Resize i_{temp} to $512 \times 512 \times 3$
  Perform stain color normalization of i_{temp}
  Normalize pixel values of i_{temp} to range $[0, 1]$
  $i_{prep} \leftarrow i_{temp}$

  if *using* pipeline P_{dir} **then**
    //Entire WSHI patch classification
    $mc \leftarrow$ Classify i_{prep} using Cl_{mul}

  else if *using* pipeline P_{seg} **then**
    //Nucleus segmentation
    $i_{temp} \leftarrow i_{prep}$
    $i_{binmask} \leftarrow$ Segment i_{temp} using Seg_{net}
    $i_{seg} \leftarrow$ bitwise_and($I_{prep}, I_{binmask}$)

    //Nuclei-segmented patch classification
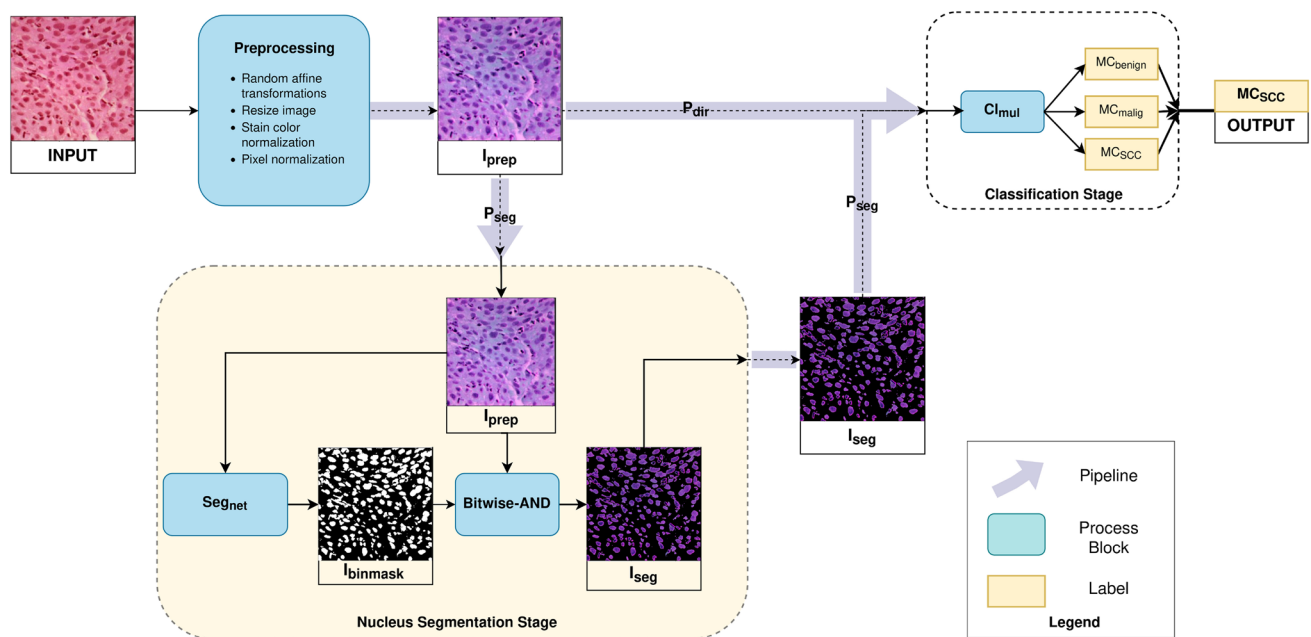 $mc \leftarrow$ Classify i_{seg} using Cl_{mul}

  end if

Fig. 1 A flowchart representing the proposed system. Pipelines P_{dir} and P_{seg}, as described in Sect. 3.1, represent the pipelines for direct and nucleus segmented classification of the WSHI patches

3.1 Preprocessing

Data preprocessing is desired to enhance the features and to reduce the noise and artifacts present in the histopathological images. In this research work, augmentation technique has been used to control overfitting, to adopt a diversity-based sampling strategy and to cater to the large input requirement of the deep learning network. Random affine transformations such as rotation, scaling and shearing are applied to augment the input image i. The input images are then scaled down to 512×512 using bi-linear interpolation.

Histopathology images often have drastically different appearances due to the use of different scanners during slide preparation, different equipment, different stain coloring, and effect from different manufacture and batches of stains. Hence, the staining must be normalized across all images of the study, before further processing. To perform stain color normalization, we have used a deconvolution based color normalization technique that uses Macenko's PCA method for stain unmixing (Macenko et al. 2009; Van Eycke et al. 2017). This is a three-step method, where Principal Component Analysis (PCA) performs stain unmixing to automatically determine the stain vectors of an H &E input image. Color deconvolution is then performed which uses the stain profile vectors of the stains to separate each stain from the image. Finally, the image is convolved with a standard stain column vector matrix. This ensures that all images across the dataset have the same standard color characteristics.

The images are then pixel normalized to envelope the pixel values between 0 and 1. This helps in improving the rate of convergence and to enhance the network's ability to standardize to diverse inputs.

The processed image i_{prep} obtained after applying the preprocessing steps is passed to two different pipelines, namely P_{dir} and P_{seg}. The direct pipeline (P_{dir}), feeds the 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} using the segmentation network Seg_{net} to obtain an image i_{seg}. This segmented image i_{seg} is then passed to the classifiers and subsequently, the performance of the two pipelines are analyzed.

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 preprocessing stage directly to the DCNN classifier Cl_{mul} which assigns a multiclass label mc.

3.2.1 Architecture of the downstream DCNN classifier — Cl_{mul}

The proposed DCNN classifier comprises six convolution blocks that use 2D depth-wise separable convolution layers (Chollet 2017). Depth-wise separable convolutions employ a
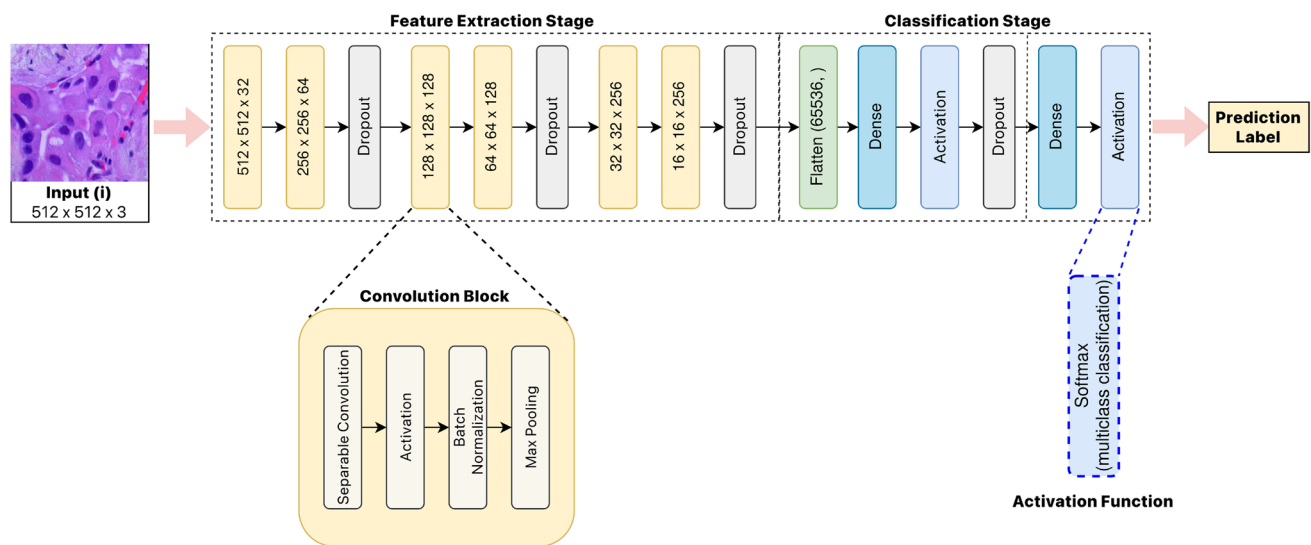
Fig. 2 The architecture of the downstream DCNN classifier as described in Sect. 3.2.1. The numbers within the boxes represent the output dimension of the respective layers. The convolution block is generically represented using an orange box and an expanded version depicts the layer structure of each convolution block

two-step convolution process where the spatial features that were learned independently from each channel are merged later using point wise convolutions. In this study, the H &E staining reveals significant color-based features which is crucial for inter-class distinguishability. Hence, employing learning parameters that operate independently on each color channel of the image is influential in leveraging these features for classification. Hence, these layers are adopted for their ability to decouple the feature learning pathways between different channels of the image (Kaiser et al. 2018).

Each block of the proposed DCNN consists of a separable convolution layer followed by 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 (overfitting).

As the images percolates through to the deeper layers of the network, the channel-size is increased progressively from 32 to 256. Consequently, the convolutions extract more subtle features that can greatly aid inter-class differentiation. To emphasize more on feature learning from smaller receptive fields than from larger areas of the WSHI patch, the receptive fields (Luo et al. 2016; Araujo et al. 2019) are varied by increasing the number of channels more aggressively in the initial convolution layers than the latter, while keeping the kernel sizes fixed at 3×3. This was done to better harvest information about atypical cell, nuclear shapes and spaces, which is contained in multiple small regions of the image that are captured with a smaller receptive field.

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 block, includes a dense layer followed by an activation layer. The representation vectors obtained through convolution are passed to an output layer with three neurons Cl_{mul} to perform multiclass classification. This DCNN architecture is represented concretely in Fig. 2.

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 have proposed an Xception-based UNet architecture called Seg_{net} to segment the nucleus.

3.3.1 Rationale of the segmentation approach

During a histopathological examination, the biopsied tissue samples are prepared on microscope slides for visual examination. 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. 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. Retaining nuclear regions by segmentation, and extracting features from these regions during classification can enhance the performance of the classifier.

3.3.2 Architecture of the nucleus segmentation network Seg_{net}

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

The UNet (Long et al. 2015) is a Convolutional Neural Network architecture, proven for segmentation of medical images. An initial experiment on our input images with a simple UNet architecture showed accuracy saturation and high training loss. Hence, we propose an Xception-style (Chollet 2017) UNet architecture, to employ residual blocks that can reduce degradation in deep networks and easily adapt to learn simplex as well as complex correlations between the layers of the network (Veit et al. 2016).

The proposed architecture consists of a contractive downsampling path followed by an expansive upsampling path. 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 are stacked, each preceded by a ReLU activation layer and followed by a batch normalization layer. This sequence is followed by a max-pooling layer and an addition-based skip connection from the input to the current block.

ReLU-only pre-activation and the full pre-activation approaches discussed in He et al. (2016) were experimented, and the former approach was more effective in reducing overfitting in our experiments. Specifically, the sequence shown in Fig. 3 was adopted in the final segmentation network, within each block of it's downsampling strata.

The latter half of the network is an expanding path that uses transposed convolutions (Dumoulin and Visin 2018) to upsample the low resolution image at the end of the contracting path progressively, back to its original resolution. In effect, the subsequent series of downsampling and upsampling operations, recovers the spatial information lost during the contraction phase and further, localizes the identified features to their corresponding positions on

Fig. 3 ReLU-only pre-activation strategy with element-wise addition based residual connection adopted within each block of the upsampling and downsampling strata of the UNet architecture
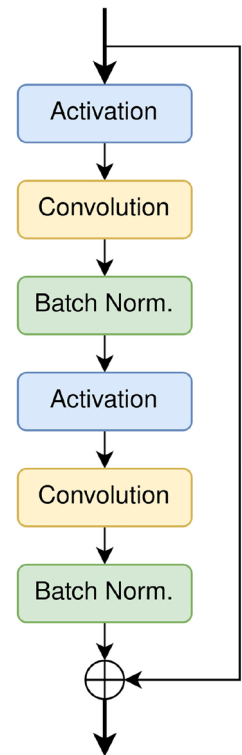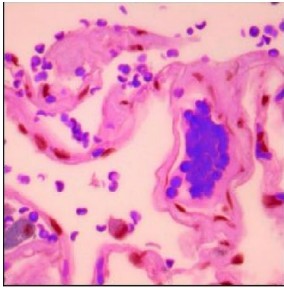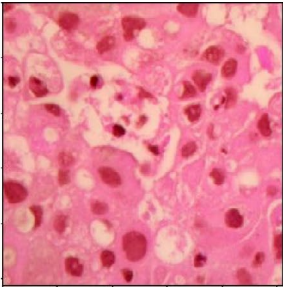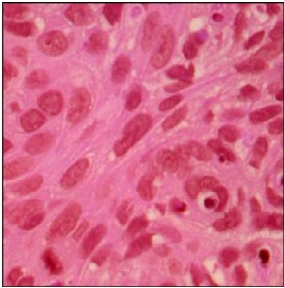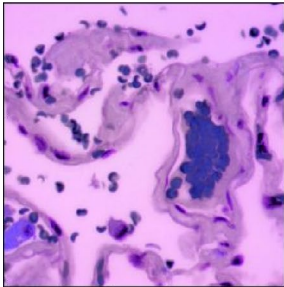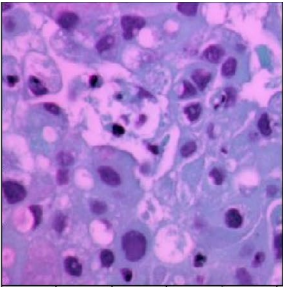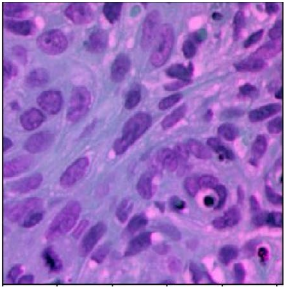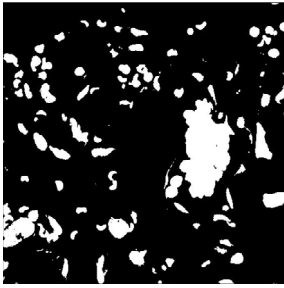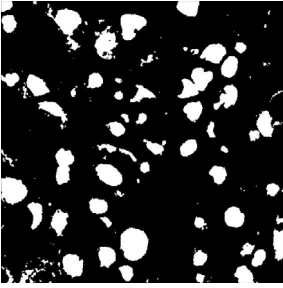
the higher resolution image produced at each step of the expansion path. To further enhance the precision of mapping between features and their respective locations on the upsampled image and avoid losing fine-grained information, skip connections are used to forward and concatenate features from corresponding stages of the contraction path with the upsampled image obtained at the same stage of expansion path. The shorter skip connections also 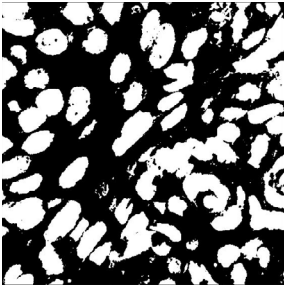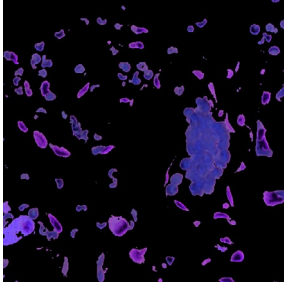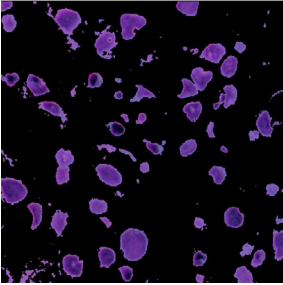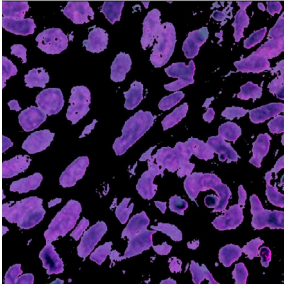help to alleviate the vanishing gradient problem (Hochreiter 1998; Drozdzal et al. 2016). Ultimately, pixel-wise classification into semantic classes is performed on these localized features once the original resolution is reached.

The transpose convolution blocks used in the expansion path have a symmetric structure with respect to the contracting path. It comprises 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, a batch normalization layer and an upsampling layer. Once again, the sequence of layers in each block follows the ReLU-only pre-activation strategy. In particular, each block of the upsampling strata follows the sequence depicted in Fig. 3.

The input WSHI patches are first resized to 512×512 with 3 channels and passed through the segmentation architecture—contracting path, followed by the expansive path. The network outputs a binary mask image, with the same dimensions as the input image, with distinct pixel

Table 1 Snapshots of WSHI patches belonging to different classes, at different processing stages of the proposed system

Image Stage	Benign	ADC	SCC
Original WSHI Patch (i)			
Preprocessed and Stain Normalized WSHI Patch (i_{prep})			
Nuclei Segmentation Mask ($i_{binmask}$)			
Nuclei Segmented WSHI Patch (i_{seg})			

values to denote semantic classes—0 s for non-nuclear and 1 s for nuclear regions.

3.3.3 Nucleus segmentation from WSHI patches

The segmentation system, namely Seg_{net}, identifies the nuclear regions of i_{prep} by producing a binary mask image called $i_{binmask}$ 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 ($i_{binmask}$) 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 different stages of WSHI patch images obtained in the classification pipeline of the proposed system for the three tumor classes is depicted in Table 1.

Fig. 4 A sample WSHI patch image from each of the three lung tumor classes in the dataset (Borkowski et al. 2023)

Benign   Adenocarcinoma   Squamous-Cell Carcinoma
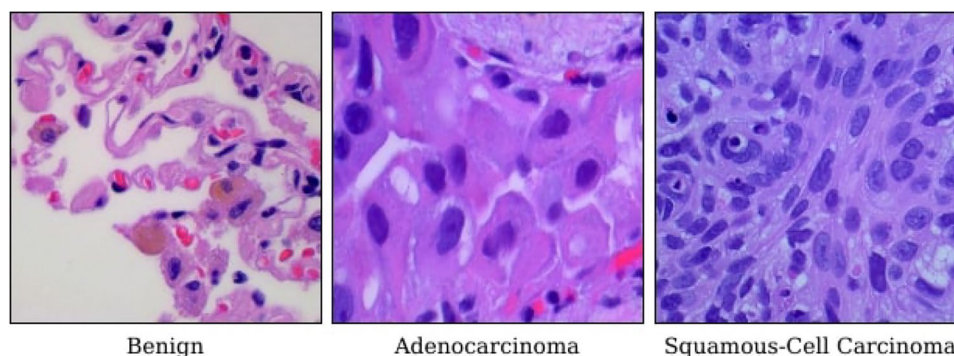
3.3.4 Classification using nucleus segmented patches

The proposed DCNN classifier Cl_{mul} whose architecture is specified in Sect. 3.2.1 is used to perform multiclass classification of the nucleus segmented WSHI patches i_{seg} obtained through pipeline P_{seg}Reaction conditions: **8** (1 mmol), **9** (1 mmol), K_2CO_3 (2 mmol), MNPFemBenzNHC@Ni complex (**7**) (100 mg) and xylene (5 mL)

4 Experiments and results

4.1 Dataset

The dataset (Borkowski et al. 2023) consists of 15000 color histopathological patch 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 B lung tumor tissues, with each image of the dataset belonging to exactly one of these classes. A few sample images from the dataset are shown in Fig. 4.

4.2 Ground truth preparation

This paper proposes the use of a DL network to classify entire or nucleus segmented WSHI patch images. The system also perform automatic semantic segmentation of WSHI patch images to retain only nuclear information using UNet architecture. This network learns to classify each pixel in the image based on the ground truth. However, the dataset (Borkowski et al. 2023) does not contain the ground truth for segmentation. Hence, the masks for all input images were created and the ground

truth was evaluated with the help of three different experts. To train the architecture, the entire dataset was split into 80% and 20% for training and testing, respectively. During the training phase, 4-fold cross validation was performed by splitting the training data into train and validation sets, comprising 75% and 25%, respectively.

To obtain ground truth data for nucleus segmentation, a set of 300 representative WSHI patch images comprising 100 images from each tumor class, was considered and validated by expert pathologists. The proposed global thresholding approach segments the nuclear information from input WSHI patches directly by leveraging the intensity variations in the visual structure of the WSHI patches due to staining. The staining attributes a darker shade to the nuclear regions, and they consequently have a higher intensity value than other components of the image.

The Otsu's method (Otsu 1979), a global thresholding procedure, is used to determine an optimal threshold $thresh_{nuc}$ to segregate the pixels in a given image into two classes with minimum intra-class variance and maximum inter-class differentiation. The input WSHI patch i is first converted to grayscale to retain only the intensity information of pixels (Liu 2020) and to discard the color details. Suzuki's contour detection algorithm (Suzuki and Be 1985), was applied to find closed-loop boundaries to identify sub-regions within the image, separating the adjacent nuclear and non-nuclear regions. 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 those of the other sub-regions are encoded as 0, thereby producing a binary segmentation mask that represents the nuclear information from i. This procedure is represented in Algorithm 2. The WSHI patch image and the corresponding binary mask serve as input and annotation respectively, to train the segmentation network.

Algorithm 2 To prepare the ground truth for nucleus segmentation

---

Input: WSHI patch i

Output: Binary mask $i_{binmask}$

$i \leftarrow$ Perform stain color normalization of i

$thresh_{nuc} \leftarrow Apply$ Otsu's method on i

$Convert$ i to grayscale

$i_{binmask} \leftarrow i$

for each $sub\text{-}region$ in i **do**

    $intensity\text{-}chk \leftarrow avg\text{-}intensity\text{-}of\text{-}sub\text{-}region$

    if $intensity\text{-}chk \geq thresh_{nuc}$ **then**

        for each $pixel$ in $sub\text{-}region$ **do**

            $pixel \leftarrow 1$

        end for

    else if $intensity\text{-}chk < thresh_{nuc}$ **then**

        for each $pixel$ in $sub\text{-}region$ **do**

            $pixel \leftarrow 0$

        end for

    end if

end for

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 (Kouwenhoven et al. 2009). GCI was estimated using Eq. 1,

$$GCI = \frac{\sum_{pairs(i>j)} X_i \cap X_j}{\sum_{pairs(i>j)} X_i \cup X_j} \tag{1}$$

where X_i represents the set of nuclear region pixels of the segmented ground truth obtained from rater i and $\sum_{pairs(i>j)} \ldots$ 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_{net} network.

4.3 Model training

This section details the methods, strategies and parameter settings when training the segmentation and classification neural networks. The architectures of these neural networks were described in Sects. 3.2.1 and 3.3.2.

4.3.1 Nucleus segmentation network Seg_{net}

The proposed Xception-style UNet architecture for Seg_{net} was trained for minimizing the dice-loss shown in Eq. 2, using the early stopping method and the NAdam optimizer. The model

converged with a loss of 0.004681. The sigmoid function was used to activate the output layer and the ReLU function was used for the hidden layers.

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

where K represents the number of semantic classes. P_k and T_k represent the set of pixels belonging to the kth class of the prediction and ground truth images respectively.

4.3.2 Multiclass classification network—Cl_{mul}

The dataset of 15000 WSHI patch images was split into training and testing sets in 80:20 ratio. The proposed DCNN architecture for classifier Cl_{mul} was trained for minimizing the Categorical Cross-Entropy (CCE) loss shown in Eq. 3 using the early stopping method and the Nesterov-accelerated Adaptive Moment Estimation (NAdam) optimizer with an initial learning rate of 0.01.

$$CCE = - \sum_{c=1}^{C} P_{true}(c).log(P_{pred}(c)) \tag{3}$$

where C is the number of output classes of the classifier and c takes on the class label values. $P_{true}(i)$ and $P_{pred}(i)$ represent the true and predicted probability of the input belonging to the ith class, respectively. The Rectified Linear Unit (ReLU) function was used to activate the hidden layers and Cl_{mul} used the softmax function for its output layers. Uniform dropout rates of 0.50 and 0.25 were used in the dense and convolution layers respectively.

4.4 Experiments

Our proposed system aims to classify the histopathological lung images into different cancer types after preprocessing the images using color normalization and augmentation. The preprocessed images were classified using deep architecture network and the performance of the classifier was compared between the WSHI-trained and nucleus segmented image trained variants. 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 classifier Cl_{mul} using entire WSHI patches as input.
2. To determine the performance of the nucleus segmentation network i_{Seg}.

3. To evaluate the performance of the DCNN based classifier Cl_{mul} using the WSHI patches with nuclear regions segmented using the segmentation network Seg_{net}.

The metrics Precision (P), Recall (R), Accuracy (Acc) and F1-Score (F1) (Baratloo et al. 2015) are used to ascertain the performance of the DCNN classifier Cl_{mul} with both, WSHI patch inputs as well as nucleus segmented inputs. These are specified in Eqs. 4, 5, 6 and 7.

$$P = \frac{\|TP\|}{\|TP + FP\|} \tag{4}$$

$$R = \frac{\|TP\|}{\|TP + FN\|} \tag{5}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{6}$$

$$Acc = \frac{\|TP + TN\|}{\|TP + TN + FP + FN\|} \tag{7}$$

In these equations, TP, FP, TN and FN represent the sets of true-positive, false-positive, true-negative and false-negative classifications respectively. $\|.\|$ 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 Eq. 8.

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

where, K represents the number of semantic classes, X_k and Y_k represent the pixels belonging to the kth class in images X and Y respectively. $\|.\|$ denotes the cardinality operation of its argument.

4.4.1 Performance of the DCNN classifier with entire WSHI patches P_{dir}

The aim of this experiment is to validate the performance of the classifier Cl_{mul} when the entire WSHI patch is fed to the network. It segregates its inputs into one of three classes: ADC, SCC and B. The confusion matrices of the classification result is shown in Table 3. To circumvent the over-fitting problem we have applied 4-fold cross validation technique, and the result is presented in Table 2. It is apparent that the accuracy obtained is consistent across runs. This shows that the proposed model is robust in generalizing to unseen samples of the dataset. Table 3, shows

Table 2 Class-wise accuracy (in %) comparison of 4-fold cross validation results for the two pipelines of lung tumor classification P_{dir} and P_{seg}

Pipeline	Class	Acc (%) for Run #			
		1	2	3	4
P_{dir}	Benign	99.20	99.40	99.40	99.30
	ADC	99.90	99.70	99.40	99.70
	SCC	86.20	88.70	87.50	86.20
P_{seg}	Benign	99.40	99.40	99.50	99.70
	ADC	99.60	99.50	99.80	99.70
	SCC	99.40	99.80	99.70	99.60

Table 3 Confusion matrix for multiclass classification performed using entire WSHI patches in the P_{dir} pipeline

Types	Predicted		
	Benign	ADC	SCC
Actual			
Benign	990	7	3
ADC	2	998	0
SCC	113	13	874

that the DCNN achieves an overall accuracy of 95.40% and a macro-averaged F1-Score of 95.37%. The proposed DCNN architecture predicts the ADC class well but there is misclassification between benign and ADC, as well as benign and SCC, potentially due to misleading features. This may have arisen due to the presence of non-nuclear information in the input images, particularly the regions that are not diagnostically contributing.

4.4.2 Performance of segmentation network Seg_{net}

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 Sect. 4.2. A corpus of 300 representative ground truth images was prepared through expert review and correction. An inter-rater reliability of 0.89 CGI was obtained across the three experts. The corpus was split into training and testing sets in 80:20 ratio and the network was trained using the training set and evaluated on the test set using the Mean-IoU metric specified in equation 8.

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 4 represents a comprehensive summary of the system's segmentation performance. The low value of overall standard deviation (0.51%) indicates consistency in the prediction, and qualifies the mean of the evaluation metric as a good representation

Table 4 Performance of the UNet-based segmentation network Seg_{net} for nucleus segmentation from WSHI patches

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

Table 5 Confusion matrix for multiclass classification performed using nuclei segmented WSHI patches in the P_{seg} pipeline

Types	Predicted		
	Benign	ADC	SCC
Actual			
Benign	995	2	3
ADC	0	997	3
SCC	1	3	996

of the model's performance across the testing set. The performance metrics of the segmentation network is presented class-wise, and it is evident that the segmentation performance is marginally lower for the SCC class with an average Mean-IoU of 93.39%. The segmentation of nuclei from images of SCC tumors is particularly demanding due to their massively clustered occurrence. Similar trends were also exhibited when the DCNN was fed with entire WSHI patch images, and misclassification was observed between the SCC and benign classes.

4.4.3 Performance of the DCNN classifier with nuclei segmented WSHI patches P_{seg}

The aim of this experiment is to validate the performance of the classifier Cl_{mul} when trained using the nucleus segmented WSHI patches. The segmentation is performed using Seg_{net}. Table 2 shows the results of 4-fold cross validation applied to the classifier when training with nucleus segmented WSHI patches. The results are consistent across different runs, proving that the model is capable of generalizing well to unseen samples of the dataset. Furthermore, hyperparameter tuning was performed based on the cross-validation results of different runs. The final confusion matrix for multiclass classification — evaluated on the test set — is presented in Table 5.

The multiclass classification achieves an accuracy of 99.60% and a macro-averaged F1-Score of 99.58%. The classification performance is superior when the nuclear regions are segmented using the segmentation network Seg_{net}. The misclassifications made by Cl_{mul} in predicting the SCC class has reduced, when compared to classification using entire WHSI patch images. This clearly proves the performance

improvement warranted by segmenting the nuclear regions of the WSHI patches before classification. However, relative to other classes, the SCC class is still characterized by 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 4. This reduced segmentation performance is likely causing the lower classification performance for this class.

The multiclass classification of the proposed system achieves high accuracy when the cancer cells are segmented and classified than when the entire WSHI patches are directly classification. When classification is performed using segmented nuclei, the classifier obtained 99.60%, whereas direct WSHI patch classification produced only 95.40%; a 4.2% difference in accuracy.

The t-test (Devore 2011) was performed to compare the accuracy of the proposed classifiers P_{dir} and Seg_{net}. Let $\hat{p}1$ and $\hat{p}2$ be the accuracy's obtained from classifiers P_{dir} and Seg_{net} respectively, and n be the number of samples. The number of samples correctly classified in P_{dir} and Seg_{net} are x1 and x2 respectively.

The formula for calculating the test statistics is given by

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{2\hat{p}(1-\hat{p})/n}} \tag{9}$$

where $\hat{p} = (x_1 + x_2)/2n$ and $\hat{p}_1 = x_1/n$, $\hat{p}_2 = x_2/n$

To prove that the global accuracy of Seg_{net} is better than that of P_{dir}, the hypothesis $H_a : p_1 < p_2$ should be true. Using the above formula we have obtained Z as -10.1709. For 5% level of significance $z_{0.5}$ is 1.645, hence the relation $-10.1709 < -1.645$ is true. From this t-test, it was concluded that Seg_{net} is more accurate than P_{dir}. Hence, classification using nucleus-segmented WSHI shows significant improvement in performance.

4.5 Overall performance and discussion

The performance of the segmentation network Seg_{net} is summarized in Table 4 and was discussed in section 4.4.3. It achieves an overall Mean-IoU 94.27%, averaged over all classes of tumors. 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 in Table 6. It achieves an overall classification accuracy of 99.60% when using the proposed P_{seg} pipeline.

There is notable improvement along all metrics, when the classification is performed using nuclei segmented WSHI patches — obtained through Seg_{net} — rather than using the entire WSHI patch input. In other words, the proposed P_{seg} pipeline, significantly outperforms the baseline P_{dir}

Table 6 Performance summary of the multiclass classification DCNN Cl_{mul}

Metric	Tumor class	Entire WSHI (P_{dir})	Nucleus segmented WSHI (P_{seg})
Precision (%)	Benign	99.00	99.50
	ADC	99.80	99.70
	SCC	87.40	99.60
	Macro average	95.40	**99.60**
Recall (%)	Benign	89.59	99.90
	ADC	98.04	99.51
	SCC	99.68	99.41
	Macro average	95.77	**99.61**
F1-Score (%)	Benign	94.06	99.69
	ADC	98.92	99.60
	SCC	93.14	99.45
	Macro Average	95.37	**99.58**
Accuracy (%)	Benign	95.83	99.80
	ADC	99.27	99.73
	SCC	95.70	99.67
	Overall	95.40	**99.60**

Results are presented as a comparison between the two pipelines — the baseline P_{dir} and the proposed P_{seg}

The numeric figures in bold indicate the summary statistics for each metric of the proposed pipeline, to highlight that is the better performing variant among the two pipelines, supporting this paper's base hypothesis

Table 7 Comparison with previous work that performs multiclass classification of lung histopathology images using the same dataset (Borkowski et al. 2023). All metrics are indicated in %, and computed only across the three lung tumor classes considered in this study

Method	Acc.	P	**R**	$F1$
Mangal et al. (2020)	97.90	–	–	–
Hatuwal and Thapa (2020)	98.05	98.09	98.06	98.07
Mehmood et al. (2022)	99.37	99.37	99.36	98.36
Masud et al. (2021)	95.32	95.39	95.49	95.44
Ours P_{dir}	95.40	95.40	95.77	95.37
Ours P_{seg}	99.60	99.60	99.61	99.58

pipeline consistently. This suggests that presenting the isolated nuclear information of the WSHI patch inputs to the classifier has a contributing impact on discerning the tumor classes. It is worth noting that when adopting the P_{seg} pipeline, the misclassification errors of Cl_{mul} between benign and malignant classes is greatly reduced, and majority of the classification error arises due to mislabelling between the ADC and SCC classes. This is of particular diagnostic importance — the system does not mislead between benign and malignant tumors.

Table 7 compares the results of our baseline and proposed pipelines for multiclass classification of lung tumors, with other existing methods that have used the same dataset (Borkowski et al. 2023) for lung tumor subtyping. Our method outperforms them along all metric measures.

Mangal et al. (2020) adopted a shallow CNN architecture to classify the lung slides into SCC, ADC and benign classes. Likewise, Hatuwal and Thapa (2020) used a simple three-hidden-layer CNN for classification. In addition, basic data augmentation techniques like flipping and zooming were applied. These implementations serve as a good comparison for the baseline P_{dir} pipeline presented in this work, which applies a DCNN classifier on entire WSHI patches without segmenting the nuclei.

Masud et al. (2021) adopt a more traditional image processing approach, wherein two independent image processing algorithms are applied to extract feature sets — 2D Fourier features and 2D Wavelet features. The feature sets are then merged and passed to a simple CNN classifier. They adopt handcrafted feature extraction to prepare the input for CNNs. In contrast, the proposed method completely decimates the need to design algorithms for extracting such handcrafted features. Using convolution based supervised learning, and nuclei segmentation, the proposed method is able to extract representative features that outperform the classification performance in Masud et al. (2021). Mehmood et al. (2022), on the other hand, use the AlexNet network (Krizhevsky et al. 2017), pre-trained on the ImageNet dataset (Deng et al. 2009). A Class-Selective Image Processing (CSIP) technique is applied only to classes that are found to underperform after a preliminary round of classification. The original images of the underperforming classes are replaced with their histogram equalized versions, and the model is retrained. This method achieves the closest (but, inferior) results to that of the proposed P_{seg} pipeline, among all the models compared. Both these methods (Masud et al. 2021; Mehmood et al. 2022) increase inference time overhead for medical diagnostic systems — the former due to image processing -based feature extraction methods, and the latter due to a more complex network architecture for classification. Although CSIP technique (Mehmood et al. 2022) can cut down training time, it does little to improve the final inference time — a more important factor in diagnostic systems.

Concretely, the proposed method adopts a simple CNN architecture for classification and completely decimates the need for handcrafted feature extraction, thereby reducing computation time overheads. In addition, the baseline classification performance is improved by segmenting and retaining only the nuclear regions of the WSHI patches, and prove our original hypothesis that nucleus segmentation improves inter-class differentiation whilst using the same classification architecture — Cl_{mul}. With this approach, the proposed

pipeline P_{seg}, consistently outperforms all related work that are based on the same dataset (Borkowski et al. 2023).

The dataset does not provide a standard division of train and test samples. Hence, different studies arbitrarily partition the dataset into train, test and validation sets. Stratified k-fold cross-validation was performed to validate model training, so as to involve different held-out folds of the training data in both training and validation sets. This corroborates the consistency of performance metrics computed for the proposed method. To normalize the effect of H &E staining across all WSHI patch images, stain color normalization was performed, that none of the other approaches have adopted.

5 Conclusion

In this paper, a deep-learning based classification pipeline was proposed for classifying lung tumors from WSHI patches into B, ADC and SCC types. The authors posited that extracting and retaining only the nuclear regions of the WSHI patches can significantly improve the DCNN classifier's ability to discern tumor classes. To this end, this work designed, compared and presented two classification pipelines — baseline P_{dir} and proposed P_{seg} — that uses the same downstream classification architecture — Cl_{mul}. The P_{dir} pipeline performs classification directly on the WSHI patches. It achieves an overall accuracy of 95.40%. On the other hand, the P_{seg} pipeline first segments the WSHI patches to retain only nuclear regions, and then feeds this image to the classifier. This approach improves the classification accuracy to 99.60%, proving our original hypothesis — the nuclei provide key information to characterize the tumor type, and segmenting them out has a positive impact on the classifier performance.

In addition to basic preprocessing techniques, stain color normalization was performed using a deconvolution method. To perform segmentation, the nucleus segmentation ground truth was prepared by applying an intensity-based image processing method to annotate nuclear regions for 300 representative images from the dataset (Borkowski et al. 2023). These images are used to train a Xception-style UNet based segmentation architecture — Seg_{net} — that handles nucleus segmentation in the P_{seg} pipeline. The training process and all presented metric evaluations were validated using stratified k-fold cross-validation to ensure consistency. It is evident from the comparisons presented in Table 7 and the t-test that our P_{seg} pipeline significantly outperforms all existing methods on the adopted dataset. The proposed methodology has outperformed state-of-the- art methods and also reduced the computational cost. As future work, our proposed methodology can also be implicated in the effective diagnosis of other diseases.

Declarations

References

Abdolhoseini M, Kluge MG, Walker FR et al (2019) Segmentation of heavily clustered nuclei from histopathological images. Sci Rep 9(1):4551. https://doi.org/10.1038/s41598-019-38813-2

Araujo A, Norris W, Sim J (2019) Computing receptive fields of convolutional neural networks. Distill. https://doi.org/10.23915/distill.00021

Baratloo A, Hosseini M, Negida A, et al (2015) Part 1: simple definition and calculation of accuracy, sensitivity and specificity. 3(2):48–49

Borkowski AA, Bui MM, Thomas LB, et al (2023) Lc25000 lung and colon histopathological image dataset. https://github.com/tampapath/lung_colon_image_set

Chen M, Huang S, Huang Z et al (2021) Detection of lung cancer from pathological images using CNN model. IEEE Int Conf Comput Sci Electr Inform Eng Intell Control Technol. https://doi.org/10.1109/CEI52496.2021.9574590

Cheng Z, Qu A (2020) A fast and accurate algorithm for nuclei instance segmentation in microscopy images. IEEE Access 8:158,679-158,689

Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 1251–1258

Coudray N, Ocampo PS, Sakellaropoulos T et al (2018) Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med 24(10):1559–1567

Cui Y, Zhang G, Liu Z et al (2019) A deep learning algorithm for one-step contour aware nuclei segmentation of histopathology images. Med Biol Eng Comput 57(9):2027–2043

Dabeer S, Khan MM, Islam S (2019) Cancer diagnosis in histopathological image: Cnn based approach. Inform Med Unlocked. https://doi.org/10.1016/j.imu.2019.100231

Deng J, Dong W, Socher R, et al (2009) Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp 248–255

Devore JL (2011) Probability and statistics for engineering and the sciences, 8th edn. Brooks/Cole

Drozdzal M, Vorontsov E, Chartrand G et al (2016) The importance of skip connections in biomedical image segmentation. Deep learning and data labeling for medical applications. Springer, pp 179–187

Dumoulin V, Visin F (2018) A guide to convolution arithmetic for deep learning. Stat 1050:11

Filho A, Silva A, Paiva A et al (2018) Classification of patterns of benignity and malignancy based on CT using topology-based phylogenetic diversity index and convolutional neural network. Pattern Recogn. https://doi.org/10.1016/j.patcog.2018.03.032

Filipczuk P, Kowal M, Obuchowicz A (2011) Automatic breast cancer diagnosis based on k-means clustering and adaptive thresholding hybrid segmentation. In: Choraś RS (ed) Image processing and communications challenges 3. Springer, Berlin, pp 295–302

Gecer B, Aksoy S, Mercan E et al (2018) Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. Pattern Recogn 84:345–356. https://doi.org/10.1016/j.patcog.2018.07.022

Graham S, Shaban M, Qaiser T et al (2018) Classification of lung cancer histology images using patch-level summary statistics. Med Imaging 1058119:327

Hatuwal B, Thapa H (2020) Lung cancer detection using convolutional neural network on histopathological images. Int J Comput Trends Technol 68:21–24. https://doi.org/10.14445/22312803/IJCTT-V68I10P104

He K, Zhang X, Ren S et al (2016) Identity mappings in deep residual networks. European conference on computer vision. Springer, Berlin, pp 630–645

Hochreiter S (1998) The vanishing gradient problem during learning recurrent neural nets and problem solutions. Internat J Uncertain Fuzziness Knowl-Based Syst 6:107–116. https://doi.org/10.1142/S0218488598000094

Hussein IJ, Burhanuddin MA, Mohammed MA et al (2021) Fully automatic segmentation of gynaecological abnormality using a new viola-jones model. Comput Mater Contin 66:3161

Hussein IJ, Burhanuddin MA, Mohammed MA et al (2022) Fully-automatic identification of gynaecological abnormality using a new adaptive frequency filter and histogram of oriented gradients (hog). Expert Syst 39(3):e12,789

Jung H, Lodhi B, Kang J (2019) An automatic nuclei segmentation method based on deep convolutional neural networks for histopathology images. BMC Biomed Eng 1(1):1–12

Kaiser L, Gomez AN, Chollet F (2018) Depthwise separable convolutions for neural machine translation. In: International Conference on Learning Representations

Kalinathan L, Kathavarayan RS, Nagendram D et al (2018) Segmentation of hepatocellular carcinoma and dysplastic liver tumors in histopathology images using area based adaptive expectation maximization. Multimed Tools Appl 77(2):1761–1782

Kalinathan L, Kathavarayan RS, Kanmani M et al (2020) Nuclei detection in hepatocellular carcinoma and dysplastic liver nodules in histopathology images using bootstrap regression. Histol Histopathol 35(10):1115–1123

Kong H, Gurcan M, Belkacem-Boussaid K (2011) Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting. IEEE Trans Med Imaging 30:1661–77. https://doi.org/10.1109/TMI.2011.2141674

Kouwenhoven E, Giezen M, Struikmans H (2009) Measuring the similarity of target volume delineations independent of the number of observers. Phys Med Biol 54(9):2863

Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. Commun ACM 60(6):84–90

Lad N, Daveshwar M (2019) Histopathological study of lung biopsy in association with immunohistochemistry. J Evol Med Dent Sci. https://doi.org/10.14260/jemds/2019/779

Lee H, Kim J (2016) 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, pp 63–69

Liu H (2020) Chapter 3 - rail transit collaborative robot systems. In: Liu H (ed) Robot systems for rail transit applications. Elsevier, pp 89–141. https://doi.org/10.1016/B978-0-12-822968-2.00003-6

Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 3431–3440

Luo W, Li Y, Urtasun R, et al (2016) Understanding the effective receptive field in deep convolutional neural networks. In: Lee D, Sugiyama M, Luxburg U et al (eds) Advances in Neural Information Processing Systems vol. 29. https://proceedings.neurips.cc/paper_files/paper/2016/file/c8067ad1937f728f51288b3eb986afaa-Paper.pdf

Macenko M, Niethammer M, Marron JS, et al (2009) A method for normalizing histology slides for quantitative analysis. In: 2009 IEEE International Symposium on Biomedical Imaging: from Nano to Macro, IEEE, pp 1107–1110

Mangal S, Chaurasia A, Khajanchi A (2020) Convolution neural networks for diagnosing colon and lung cancer histopathological images. arXiv preprint arXiv:2009.03878

Masood A, Sheng B, Li P et al (2018) Computer-assisted decision support system in pulmonary cancer detection and stage classification on CT images. J Biomed Inform. https://doi.org/10.1016/j.jbi.2018.01.005

Masud M, Sikder N, Nahid AA et al (2021) A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework. Sensors. https://doi.org/10.3390/s21030748

Medeiros R, Filho PP, Barata Rodrigues M et al (2020) Lung nodule malignancy classification in chest computed tomography images using transfer learning and convolutional neural networks. Neural Comput Appl. https://doi.org/10.1007/s00521-018-3895-1

Mehmood S, Ghazal TM, Khan MA et al (2022) Malignancy detection in lung and colon histopathology images using transfer learning with class selective image processing. IEEE Access 10:25,657-25,668. https://doi.org/10.1109/ACCESS.2022.3150924

Nirali L, Daveshwar M (2019) Histopathological study of lung biopsy in association with immunohistochemistry. J Evolution Med Dent Sci 8(48):3609

Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 9(1):62–66

Phoulady HA, Goldgof DB, Hall LO et al (2016) Nucleus segmentation in histology images with hierarchical multilevel thresholding. Med Imaging 9791:280

Qureshi H, Sertel O, Rajpoot N et al (2008) Adaptive discriminant wavelet packet transform and local binary patterns for meningioma subtype classification. In: Metaxas D, Axel L, Fichtinger G et al (eds) Medical image computing and computer-assisted intervention - MICCAI 2008. Springer, Berlin, pp 196–204

Rastogi A (2018) Changing role of histopathology in the diagnosis and management of hepatocellular carcinoma. World J Gastroenterol 24(35):4000–4013. https://doi.org/10.3748/wjg.v24.i35.4000

Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. International conference on medical image computing and computer-assisted intervention. Springer, Berlin, pp 234–241

Sertel O, Lozanski G, Shana'ah A et al (2010) Computer-aided detection of centroblasts for follicular lymphoma grading using adaptive likelihood-based cell segmentation. IEEE Trans Biomed Eng 57(10):2613–2616. https://doi.org/10.1109/TBME.2010.2055058

Shakeel PM, Burhanuddin M, Desa MI (2020) Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. Neural Comput Appl 34:9579

Sharma S, Mehra R (2020) Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images-a comparative insight. J Digit Imaging 33(3):632–654

Shukla K, Tiwari A, Sharma S et al (2017) Classification of histopathological images of breast cancerous and non cancerous cells based on morphological features. Biomed Pharmacol J 10(1):353–366

Sirinukunwattana K, Ahmed RS, Tsang YW et al (2016) Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans Med Imaging 35(5):1196–1206

Sun W, Zheng B, Qian W (2017) Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis. Comput Biol Med 89:530–539. https://doi.org/10.1016/j.compbiomed.2017.04.006

Suresh S, Mohan S (2020) Roi-based feature learning for efficient true positive prediction using convolutional neural network for lung cancer diagnosis. Neural Comput Appl 32(20):15,989-16,009

Suzuki S, Be K (1985) Topological structural analysis of digitized binary images by border following. Comput Vis Graph Image Process 30(1):32–46. https://doi.org/10.1016/0734-189X(85)90016-7

Uppada R, Rao SK, Prasad K (2017) Supervised classification of breast cancer malignancy using integrated modified marker controlled watershed approach. In: 2017 IEEE 7th International Advance Computing Conference (IACC), pp 584–589

Van Eycke YR, Allard J, Salmon I et al (2017) Image processing in digital pathology: an opportunity to solve inter-batch variability of immunohistochemical staining. Sci Rep 7(1):1–15

Veit A, Wilber M, Belongie S (2016) Residual networks behave like ensembles of relatively shallow networks. arXiv preprint arXiv:1605.06431

Veta M, van Diest PJ, Kornegoor R et al (2013) Automatic nuclei segmentation in h &e stained breast cancer histopathology images. PLoS ONE 8(7):null. https://doi.org/10.1371/journal.pone.0070221

Vu QD, Graham S, Kurc T et al (2019) Methods for segmentation and classification of digital microscopy tissue images. Front Bioeng Biotechnol 7:53. https://doi.org/10.3389/fbioe.2019.00053

Win KY, Choomchuay S, Hamamoto K et al (2018) Comparative study on automated cell nuclei segmentation methods for cytology pleural effusion images. J Healthc Eng. https://doi.org/10.1155/2018/9240389

Xing F, Xie Y, Yang L (2016) An automatic learning-based framework for robust nucleus segmentation. IEEE Trans Med Imaging 35(2):550–566. https://doi.org/10.1109/TMI.2015.2481436

Xu J, Gong L, Wang G et al (2019) Convolutional neural network initialized active contour model with adaptive ellipse fitting for nuclear segmentation on breast histopathological images. J Med Imaging (Bellingham, Wash) 6(1):017,501. https://doi.org/10.1117/1.jmi.6.1.017501

Xu R et al (2022) Histopathological tissue segmentation of lung cancer with bilinear cnn and soft attention. Biomed Res Int. https://doi.org/10.1155/2022/7966553

Yang X, Li H, Zhou X (2006) Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy. IEEE Trans Circuits Syst I Regul Pap 53(11):2405–2414. https://doi.org/10.1109/TCSI.2006.884469

Yu KH, Zhang C, Berry GJ et al (2016) Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun 7(1):12,474. https://doi.org/10.1038/ncomms12474

Zhang M, Wu T, Bennett KM (2015) Small blob identification in medical images using regional features from optimum scale. IEEE Trans Biomed Eng 62(4):1051–1062. https://doi.org/10.1109/TBME.2014.2360154